ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

ADIUVANTIBUS

G. ALEXITS, E. EGERVÁRY, L. FEJÉR, CH. JORDAN,
L. KALMÁR, L. RÉDEI, A. RÉNYI, F. RIESZ,
B. SZ.-NAGY, P. TURÁN, O. VARGA

REDIGIT

G. HAJÓS

TOMUS VI FASCICULI 1—2

MAGYAR TUDOMÁNYOS AKADÉMIA
BUDAPEST, 1955

ACTA MATH. HUNG.

ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

A MAGYAR TUDOMÁNYOS AKADÉMIA III. OSZTÁLYÁNAK MATEMATIKAI KÖZLEMÉNYEI

SZERKESZTŐSÉG ÉS KIADÓHIVATAL: BUDAPEST, V., ALKOTMÁNY U. 21

Az Acta Mathematica orosz, francia, angol és német nyelven közöl értekezéseket a matematika köréből.

Az Acta Mathematica változó terjedelmű füzetekben jelenik meg, több füzet alkot egy kötetet.

A közlésre szánt kéziratok géppel írva, a következő címre küldendők:

Acta Mathematica, Budapest 62, Postafiók 440.

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Az Acta Mathematica előfizetési ára kötetenként belföldre 80 forint, külföldre 110 forint. Megrendelhető a belföld számára az „Akadémiai Kiadó"-nál (Budapest, V., Alkotmány utca 21. Bankszámla 04-878-111-46), a külföld számára pedig a „Kultúra" Könyv- és Hírlap Külkereskedelmi Vállalatnál (Budapest, VI., Sztálin út 21. Bankszámla 43-790-057-181) vagy külföldi képviseleteinél és bizományosainál.

„*Acta Mathematica*" публикует трактаты из области математических наук на русском, французском, английском и немецком языках.

„*Acta Mathematica*" выходит отдельными выпусками разного объема. Несколько выпусков составляют один том.

Предназначенные для публикации рукописи (в напечатанном на машинке виде) следует направлять по адресу:

Acta Mathematica, Budapest 62, Postafiók 440.

По этому же адресу направлять всякую корреспонденцию для редакции и администрации.

Подписная цена „*Acta Mathematica*" — 110 форинтов за том. Заказы принимает предприятие по внешней торговле книг и газет „*Kultúra*" (*Budapest, VI., Sztálin út 21.* Текущий счет № 43-790-057-181), или его заграничные представительства и уполномоченные.

ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

ADIUVANTIBUS

G. ALEXITS, E. EGERVÁRY, L. FEJÉR, CH. JORDAN,
L. KALMÁR, L. RÉDEI, A. RÉNYI, F. RIESZ,
B. SZ.-NAGY, P. TURÁN, O. VARGA

REDIGIT

G. HAJÓS

TOMUS VI

MAGYAR TUDOMÁNYOS AKADÉMIA
BUDAPEST, 1955

ACTA MATH. HUNG.

Technikai szerkesztő: Molnár Ferenc

A kiadásért felelős: Az Akadémiai Kiadó igazgatója

Műszaki felelős: Farkas Sándor

Szegedi Nyomda Vállalat 55-3605

Felelős vezető: Vincze György

INDEX

TOMUS VI

Aczél, J., Lösung der Vektor-Funktionalgleichung der homogenen und inhomogenen n-dimensionalen einparametrigen „Translation", der erzeugenden Funktion von Kettenreaktionen und des stationären und nicht-stationären Bewegungsintegrals.
Я. Ацел, Решение векторного функционального уравнения однородных и неоднородных n-мерных однопараметрических „трансляций", в связи с производящей фунцией цепных реакций и интегралом стационарных и нестационарных движений 131

Alexits, G., Über die Konvergenz der Orthogonalpolynom-entwicklungen.
Г. Алексич, О сходимости разложения в ряд по ортогональным многочленам 1

Alexits, G., Sur la caractérisation de certaines classes de fonctions au sens de la théorie constructive des fonctions.
Г. Алексич, Охарактеризование некоторых классов функций с помощью конструктивной теории функций 41

Carlitz, L., A special symmetric equation in a finite field.
Л. Карлиц, Одно особенное симметрическое уравнение в конечном поле 445

Czipszer, J. and Gehér, L., Extension of functions satisfying a Lipschitz condition.
Я. Ципсер и Л. Гехер, О распространении функций, удовлетворяющих условиям Липшица 213

Császár, Á., Sur la structure des espaces de probabilité conditionnelle.
А. Часар, О структуре пространств условной вероятности 337

Császár, Á., Sur les fonctions localement monotones au sens généralisé.
А. Часар, О функциях, локально монотонных в общем смысле 451

Erdős, P. and Turán, P., On the role of the Lebesgue functions in the theory of the Lagrange interpolation.
П. Ердёш и П. Туран, Роль функций Лебега в теории интерполяции Лагранжа 47

Fejes Tóth, L., Extremum properties of the regular polytopes.
Л. Фееш Тот, Об экстремальных свойствах регулярных политопов . . 143

Fejér, L., Verschiedene Bemerkungen elementarer Natur über die Grundpolynome, die bei den parabolischen Interpolationen auftreten.
Л. Фейер, Несколько элементарных замечаний о фундаментальных многочленах параболического интерполирования 227

Fisz, M., Refinement of a probability limit theorem and its application to Bessel functions.
М. Фиш, Уточнение одной вероятностной предельной теоремы и ее применение к бесселевым функциям 199

Fleischer, I., A note on subdirect products.
И. Флейшер, Замечание о подпрямых произведениях 463

FREUD, G., Über das gliedweise Differenzieren einer orthogonalen Polynomreihe.
Г. Фрайд, О почленном дифференцировании рядов по ортогональным многочленам 221
FREUD, G., Über differenzierte Folgen der Lagrangeschen Interpolation.
Г. Фрайд, О дифференцировании интерполяционных последовательностей Лагранжа 467
GEHÉR, L. and CZIPSZER, J., Extension of functions satisfying a Lipschitz condition.
Л. Гехер и Я. Ципсер, О распространении функций, удовлетворяющих условиям Липшица 213
GEHÉR, L., On approximately differentiable functions of two variables.
Л. Гехер, Об аппроксимативной дифференцируемости функций двух переменных 439
GILBERT, W. M., Projections of probability distributions.
В. М. Джилберт, О проекциях распределений вероятностей 195
HÁJEK, J. and RÉNYI, A., Generalization of an inequality of Kolmogorov.
Я. Хаек и А. Реньи, Обобщение одного неравенства А. Н. Колмогорова 281
ILIEFF, L., Über trigonometrische Integrale, welche ganze Funktionen mit nur reellen Nullstellen darstellen.
Л. Илиев, О тригонометрических интегралах, представляющих целые функции, имеющие только вещественных корней 191
MIKOLÁS, M., Über gewisse Eigenschaften orthogonaler Systeme der Klasse L^2 und die Eigenfunktionen Sturm—Liouvillescher Differentialgleichungen.
М. Миколаш, О некоторых свойствах ортогональных систем и о собственных значениях дифференциальных уравнений типа Штурма—Лиувилля 147
MIKOLÁS, M., Zur Theorie der Gammafunktion, der Riemannschen Zetafunktion und verwandter Funktionen. I.
М. Миколаш, К теории гамма-функции, ζ-функции Римана и других, родственных им функций. I 381
RÉDEI, L., Zetafunktionen in der Algebra.
Л. Редеи, ζ-функции в алгебре 5
RÉDEI, L., Neuer Beweis des Hajósschen Satzes über die endlichen Abelschen Gruppen.
Л. Редеи, Новое доказательство теоремы Хайоша о конечных абелевых группах 27
RÉDEI, L., Die gruppentheoretischen Zetafunktionen und der Satz von Hajós.
Л. Редеи, ζ-функции в теории групп и теорема Хайоша 271
RÉNYI, A. and HÁJEK, J., Generalization of an inequality of Kolmogorov.
А. Реньи и Я. Хаек, Обобщение одного неравенства А. Н. Колмогорова 281
RÉNYI, A., On a new axiomatic theory of probability.
А. Реньи, Новое аксиоматическое построение теории вероятностей . . 285
STEINFELD, O., Bemerkung zu einer Arbeit von T. Szele.
О. Штеинфельд, Замечание к одной статье Т. Селе 479
SURÁNYI, J. and TURÁN, P., Notes on interpolation. I. (On some interpolatorical properties of the ultraspherical polynomials.)
Я. Шураньи и П. Туран, Замечание об интерполяции. I. (О некоторых интерполяционных свойствах ультрасферических многочленов) 67
SZÁSZ, F., On groups every cyclic subgroup of which is a power of the group.
Ф. Сас, О группах, каждая циклическая подгруппа которых является некоторой степенью группы 475
SZÜSZ, P., Bemerkungen zur Approximation einer reellen Zahl durch Brüche.
П. Сюс, Замечания к аппроксимации вещественных чисел дробями . . . 203
TAKÁCS, L., On processes of happenings generated by means of a Poisson process.
Л. Такач, О процессах происшествия, порожденных процессом Пуассона 81

Takács, L., Investigation of waiting time problems by reduction to Markov processes.
Л. Такач, Исследование проблем времени ожидания с помощью марковских процессов 101
Takács, L., On stochastic processes connected with certain physical recording apparatuses.
Л. Такач, О стохастических процессах, связанных с некоторыми физическими регистрирующими приборами 363
T. Sós, Vera and Turán, P., On some new theorems in the theory of diophantine approximations.
Вера Т. Шош и П. Туран, О некоторых новых теоремах теории диофантовых приближений 241
Turán, P., and Erdős, P., On the role of the Lebesgue functions in the theory of the Lagrange interpolation.
П. Туран и П. Ердёш, Роль функций Лебега в теории интерполяции Лагранжа 47
Turán, P. and Surányi, J., Notes on interpolation. I. (On some interpolatorical properties of the ultraspherical polynomials.
П. Туран и Я. Шуранъи, Замечание об интерполяции. I. (О некоторых интерполяционных свойствах ультрасферических многочленов) 67
Turán, P. and T. Sós, Vera, On some new theorems in the theory of diophantine approximations.
П. Туран и Вера Т. Шош, О некоторых новых теоремах теории диофантовых приближений 241
Turán, P., On the instability of systems of differential equations.
П. Туран, О неустойчивости систем дифференциальных уравнений . . . 257

ÜBER DIE KONVERGENZ DER ORTHOGONALPOLYNOMENTWICKLUNGEN

Von

G. ALEXITS (Budapest), Mitglied der Akademie

1. Vor kurzem haben wir gezeigt,[1] daß die Fourierreihe einer Funktion $f \in L^2$ im Intervall $(a, b) \subset (0, 2\pi)$ fast überall konvergiert, wenn der quadratische Stetigkeitsmodul

$$\omega_2(\delta, f; a, b) = \sup_{|h| \leqq \delta} \left\{ \int_a^b [f(x+h) - f(x)]^2 dx \right\}^{\frac{1}{2}}$$

von der Größenordnung $O(1/\sqrt{\lambda(1/\delta)})$ ist, wo $\lambda(x)$ eine beliebige monoton zunehmende Funktion mit

$$\int_1^\infty \frac{dx}{x\lambda(x)} < \infty \tag{1}$$

bedeutet. Ungefähr zur selben Zeit erschien eine Arbeit von S. B. STEČKIN,[2] in welcher derselbe Satz in etwas verschiedener Fassung bewiesen wurde. In der vorliegenden Note setzen wir uns den Beweis einer ähnlichen Konvergenzbedingung bezüglich der Orthogonalpolynomentwicklungen stetiger Funktionen zum Ziel.

2. Bezeichne $w(x) \geqq 0$ eine in $(-1, 1)$ definierte, L-integrierbare Funktion. Sie bestimmt bekanntlich eindeutig ein System $\{p_n(x)\}$ von normierten Orthogonalpolynomen, wo $p_n(x)$ ein Polynom n-ten Grades bedeutet. Die Orthogonalreihe

$$\sum_{n=0}^{\infty} c_n p_n(x) \sim f(x) \tag{2}$$

[1] G. ALEXITS, Über den Einfluß der Struktur einer Funktion auf die Konvergenz fast überall ihrer Fourierreihe, *Acta Math. Acad. Sci. Hung.*, **4** (1953), S. 95—101.

[2] С. Б. Стечкин, О теореме Колмогорова—Селиверстова, Известия Акад. Наук СССР, **17** (1953), S. 499—512.

heißt die Entwicklung von $f(x)$, wenn ihre Koeffizienten durch

$$c_n = \int_{-1}^{1} w(x) f(x) p_n(x) dx$$

bestimmt sind. Wir werden den folgenden Satz beweisen:

Genügt die Gewichtsfunktion $w(x)$ *in* $(-1, 1)$ *fast überall der Bedingung*

$$0 < w(x) = O\left(\frac{1}{\sqrt{1-x^2}}\right)$$

und sind die durch $w(x)$ *bestimmten Orthonormalpolynome* $p_0(x), p_1(x), \ldots$ *in einem ganz im Inneren von* $(-1, 1)$ *liegenden Intervall* (a, b) *in ihrer Gesamtheit beschränkt, erfüllt ferner der in* (a, b) *gebildete Stetigkeitsmodul*

$$\omega(\delta, f; a, b) = \sup_{\substack{|h| \leqq \delta \\ a \leqq x \leqq b}} |f(x+h) - f(x)|$$

der L_w^2*-integrierbaren Funktion* $f(x)$ *die Bedingung*

$$\omega(\delta, f; a, b) = O\left(\frac{1}{\sqrt{\lambda(1/\delta)}}\right),$$

wo $\lambda(x)$ *eine monoton zunehmende, der Bedingung* (1) *genügende Funktion bedeutet, so konvergiert die Entwicklung* (2) *in* (a, b) *fast überall.*

Bezeichnet insbesondere $\{p_n(x)\}$ eines der klassischen Orthogonalpolynomsysteme, so folgt aus unserem Satz als Korollar die Behauptung, daß die klassischen Orthogonalpolynomentwicklungen einer Funktion, die einer Dini-Lipschitzschen Bedingung α-ter Ordnung mit $\alpha > \frac{1}{2}$ genügt, fast überall konvergieren. Dann läßt sich nämlich (a, b) im Inneren von $(-1, 1)$ beliebig annehmen und $\lambda(x) = (\log x)^{2\alpha}$ mit $\alpha > \frac{1}{2}$ wählen.

3. Zum Beweise bilden wir $(-1, 1)$ durch die Transformation $\vartheta = \arccos x$ auf das Intervall $(0, \pi)$ ab, und betrachten statt $f(x)$ die gerade Funktion $f^*(\vartheta) = f(\cos \vartheta)$. Ist $|\chi| \leqq \delta$, so ist auch $|\cos(\vartheta + \chi) - \cos \vartheta| \leqq |\chi| \leqq \delta$ und daher gilt

$$\omega(\delta, f^*; \arccos b, \arccos a) \leqq \omega(\delta, f; a, b) = O\left(\frac{1}{\sqrt{\lambda(1/\delta)}}\right).$$

Da $(\arccos b, \arccos a)$ ganz im Inneren von $(0, \pi)$ liegt, läßt sich eine Funktion $g^*(\vartheta)$ folgenderweise definieren:

$$g^*(\vartheta) = \begin{cases} f^*(\vartheta), \text{ falls } \vartheta \in (\arccos b, \arccos a), \\ f^*(\arccos b), \text{ falls } \vartheta \in (0, \arccos b), \\ f^*(\arccos a), \text{ falls } \vartheta \in (\arccos a, \pi). \end{cases}$$

Offensichtlich ist dann

$$\omega(\delta, g^*; 0, \pi) = \omega(\delta, f^*; \arccos b, \arccos a) = O\left(\frac{1}{\sqrt{\lambda(1/\delta)}}\right). \tag{3}$$

Dann gilt aber — wie wir es in der eingangs erwähnten Note gezeigt haben[3] — für die Fourierkoeffizienten $\alpha_0, \alpha_1, \ldots$ der Funktion $g^*(\vartheta)$ die Kolmogoroff-Seliverstoff-Plessnersche Beziehung

$$\sum_{n=1}^{\infty} \alpha_n^2 (\log n)^2 < \infty.$$

An anderer Stelle haben wir aber bewiesen,[4] daß diese Bedingung, unter den getroffenen Annahmen über die Gewichtsfunktion $w(x)$, die Konvergenz der Entwicklung

$$\sum_{n=0}^{\infty} \gamma_n p_n(x) \sim g(x)$$

in $(-1, 1)$ fast überall nach sich zieht, wo $g(x)$ die aus $g^*(\vartheta)$ durch die Transformation $x = \cos\vartheta$ hervorgehende Funktion bezeichnet. Damit wäre unser Satz schon bewiesen, wenn wir wüßten, daß die Reihen $\Sigma c_n p_n(x)$ und $\Sigma \gamma_n p_n(x)$ in jedem Intervall $(a+\varepsilon, b-\varepsilon)$ mit beliebig kleinem $\varepsilon > 0$ äquikonvergent sind. Dies ergibt sich aber leicht. Die Partialsummen der beiden Entwicklungen stehen nämlich nach der Christoffel—Darbouxschen Formel in folgendem Zusammenhang:

$$s_n(f, x) - s_n(g, x) = O(1) \int_{-1}^{1} [f(t) - g(t)] w(t) \frac{p_n(t) p_{n+1}(x) - p_{n+1}(t) p_n(x)}{t - x} dt.$$

Da aber für $\vartheta \in (\arccos b, \arccos a)$ die Beziehung $f^*(\vartheta) = g^*(\vartheta)$ besteht, ist $f(t) - g(t) = 0$, wenn $t \in (a, b)$, also

$$s_n(f, x) - s_n(g, x) = O(1)\left(\int_{-1}^{a} + \int_{b}^{1}\right) \frac{f(t) - g(t)}{t - x} w(t)[p_n(t) p_{n+1}(x) - p_{n+1}(t) p_n(x)] dt.$$

Hier besteht die rechte Seite aus Entwicklungskoeffizienten der Funktion

$$F(t) = \begin{cases} 0 \text{ für } t \in (a, b), \\ \dfrac{f(t) - g(t)}{t - x} \text{ für } t \notin (a, b), \end{cases}$$

multipliziert mit den Polynomen $p_{n+1}(x)$ bzw. $p_n(x)$, deren absolute Werte aber, wegen $x \in (a+\varepsilon, b-\varepsilon)$, laut Annahme unter einer festen Zahl bleiben.

[3] A. a. O. [1], S. 98—99.

[4] G. Alexits, Sur la convergence et la sommabilité presque partout des séries de polynomes orthogonaux, *Acta Sci. Math.*, **12** (1950), S. 223—225.

Daraus folgt, da $F(t)$ eine L_w^2-integrierbare Funktion ist, deren Entwicklungskoeffizienten mithin gegen Null konvergieren:

$$s_n(f, x) - s_n(g, x) = o(1).$$

Hiemit ist die Äquikonvergenz der Reihen $s_n(f, x)$ und $s_n(g, x)$ im Intervall $(a+\varepsilon, b-\varepsilon)$ und folglich auch unser Satz bewiesen.

(Eingegangen am 12. November 1954.)

О СХОДИМОСТИ РАЗЛОЖЕНИЯ В РЯД ПО ОРТОГОНАЛЬНЫМ МНОГОЧЛЕНАМ

Г. Алексич (Будапешт)

(Резюме)

Пусть $w(x)$ функция определенная в интервале $(-1, 1)$ и удовлетворяющая условию

$$0 < w(x) = O\left(\frac{1}{\sqrt{1-x^2}}\right).$$

Обозначим через $\{p_n(x)\}$ полную ортогональную нормированную систему полиномов определенную весовой функцией $w(x)$. Обозначим далее через $\omega(\delta, f; a, b)$ модуль непрерывности функции $f(x)$ в интервале $(a, b) \subset (-1, 1)$.

В настоящей работе доказывается аналог для ортогональных полиномов теоремы о рядах Фурье, доказанной независимо друг от друга Стечкиным и автором.

Если $\{p_n(x)\}$ является равномерно ограниченной в интервале (a, b) лежащего внутри $(-1, 1)$, и модуль непрерывности функции $f(x)$ удовлетворяет условию

$$\omega(\delta, f; a, b) = O\left(\frac{1}{\sqrt{\lambda(1/\delta)}}\right),$$

где $\lambda(x)$ функция, монотонно возрастающая для которой

$$\int_1^\infty \frac{dx}{x\lambda(x)} < \infty,$$

то разложение

$$f(x) \sim \sum_{n=0}^{\infty} c_n p_n(x)$$

сходится почти всюду в (a, b).

ZETAFUNKTIONEN IN DER ALGEBRA

Von

LADISLAUS RÉDEI (Szeged), Mitglied der Akademie

Dem Andenken von BERNHARD RIEMANN *gewidmet anläßlich der Hundertjahrfeier des Einreichens seiner Habilitationsschrift an die Göttinger Universität*

§ 1. Einleitung

Auch schon bisher sind viele wichtige Verallgemeinerungen und Analoga der Riemannschen Zetafunktionen bekannt geworden. In dieser Arbeit werden wir im Zusammenhang mit beliebigen (algebraischen) Strukturen (d. h. Gruppen, Ringen usw.) eine neue Art „Zetafunktionen" einführen. Näher werden wir hier nur den Fall von endlichen Abelschen Gruppen untersuchen, der zu Resultaten führen wird, die uns von großem Interesse und fortsetzungsfähig zu sein scheinen. Vielversprechend ist auch der Fall der unendlichen Abelschen Gruppen (insbesondere der Torsionsgruppen), den wir aber bisher noch kaum untersucht haben. Die Nützlichkeit unserer Zetafunktionen für sonstige Strukturen scheint uns dagegen fraglich zu sein. Selbst schon die endlichen Abelschen Gruppen werden zu ziemlich zusammengesetzten Untersuchungen Anlaß geben. Am merkwürdigsten wird dabei vielleicht, daß alle von uns bezüglich der Zetafunktionen für die endlichen Abelschen Gruppen entdeckten Grundtatsachen vom Fundamentalsatz dieser Gruppen unabhängig sind. Diese Unabhängigkeit erschwert einerseits die Untersuchungen, gibt ihnen aber andererseits auch den Reiz.

Für eine endliche Menge M bezeichnen wir mit (M) die Anzahl der Elemente von M. Ist M die leere Menge, die wir mit $\mathbf{0}$ bezeichnen, so soll natürlich $(M)=0$ sein. Für eine unendliche Menge M setzen wir $(M)=\infty$. Ist M insbesondere eine Struktur, so nennen wir (M) die Ordnung von M.

Stets soll n eine nichtnegative ganze Zahl und $\mathfrak{N}$ die Menge der Zahlen $1,\ldots,n$ bezeichnen. (Für $n=0$ ist also $\mathfrak{N}=\mathbf{0}$.)

Für eine komplexe Variabel z bezeichnen wir mit Rz den reellen Teil von z.

Sind $A_1, \ldots, A_n$ irgendwelche (nicht notwendig verschiedene) Unterstrukturen[1] einer beliebigen algebraischen Struktur S, so setzen wir für $Rz > 0$

$$\varrho(z) = \varrho(z; A_1, \ldots, A_n) = \sum_{\mathfrak{M} \subseteq \mathfrak{N}} (-1)^{(\mathfrak{M})} (A_{\mathfrak{M}})^{-z}, \tag{1}$$

wobei wir mit $A_{\mathfrak{M}}$ die durch die Elemente der A_i ($i \in \mathfrak{M}$) erzeugte Unterstruktur von S bezeichnen, d. h. $A_{\mathfrak{M}}$ ist die engste Unterstruktur von S, die die A_i ($i \in \mathfrak{M}$) enthält. (Wenn $\mathfrak{M} = \mathbf{0}$ ist, so kann $A_{\mathfrak{M}}$ sinnlos werden, aber dann setzen wir $(A_{\mathfrak{M}}) = 1$.) Man sieht, daß (1) aus 2^n Gliedern besteht. Gleich bemerken wir, daß die $\mathfrak{M}$ mit $(\mathfrak{M}) = \infty$ wegen $Rz > 0$ keinen Beitrag zu (1) liefern, weshalb man in (1) nur die $\mathfrak{M}$ mit endlichem $A_{\mathfrak{M}}$ zu berücksichtigen hat.

Wir lassen auch den Fall zu, daß S mit Operatoren versehen ist; dann sollen als „Unterstrukturen" nur die „zulässigen Unterstrukturen" gelten.

Handelt es sich um (abzählbar) unendlich viele Unterstrukturen $A_1, A_2, \ldots$ von S, so definieren wir für ein z mit $Rz \geqq 1$

$$\varrho(z) = \varrho(z; A_1, A_2, \ldots) = \lim_{n \to \infty} \varrho(z; A_1, \ldots, A_n) \tag{2}$$

stets, wenn die rechte Seite existiert und von der Reihenfolge der A_i unabhängig ist. Ist dabei $\varrho(z)$ in einem Teilgebiet der Halbebene $Rz \geqq 1$ analytisch und regulär, so definieren wir $\varrho(z)$ (von diesem Gebiet ausgehend) durch analytische Fortsetzung.

Als ein sehr einfaches Beispiel betrachten wir den Fall, wo S eine Gruppe und zwar das direkte Produkt von $A_1, A_2, \ldots$ bedeutet und $(A_i) = p_i$ die i-te Primzahl der natürlichen Zahlenfolge ist. Nach (1) gilt dann offenbar

$$\varrho(z; A_1, \ldots, A_n) = \prod_{i=1}^{n} (1 - p_i^{-z}).$$

Hieraus folgt nach (2)

$$\varrho(z) = \varrho(z; A_1, A_2, \ldots) = \prod_{i=1}^{\infty} (1 - p_i^{-z}).$$

Jetzt ist also $\varrho(z)^{-1}$ eben die Zetafunktion von RIEMANN.

Deshalb fühlen wir uns veranlaßt ganz allgemein die Bezeichnung

$$\zeta(z) = \zeta(z; A_1, A_2, \ldots) = \varrho(z)^{-1} = \varrho(z; A_1, A_2, \ldots)^{-1} \tag{3}$$

einzuführen und diese $\zeta(z)$ *Zetafunktionen* (in der Algebra) zu nennen. Je nachdem die vorgelegte Struktur eine Gruppe, ein Ring, ... ist, sprechen wir entsprechend auch über *gruppentheoretische, ringtheoretische, ... Zetafunktionen*. (In der Definition (3) lassen wir auch den Fall zu, wo an Stelle von $A_1, A_2, \ldots$ eine endliche Folge $A_1, \ldots, A_n$ tritt.)

[1] „Unterstrukturen" einer Gruppe, eines Ringes usw. sollen bzw. die Untergruppen Unterringe usw. der gegebenen Struktur bedeuten.

Ein Fall für die ringtheoretischen Zetafunktionen entsteht z. B. folgendermaßen. Es bezeichne S einen Ring, den wir als den Modul S^+ seiner Elemente auffassen, versehen mit den Elementen von S als links- und rechtsseitigen Operatoren, wobei als „Operation" die links- bzw. rechtsseitige Multiplikation in S gelten soll. Das kommt darauf hinaus, daß in (3) für die A_i eben nur die Ideale von S zuzulassen sind und $A_{\mathfrak{M}}$ (in (1)) die Summe der Ideale A_i ($i \in \mathfrak{M}$) bedeutet. Beachtet man nur die linksseitigen Operationen, so hat man für die A_i die Linksideale von S zuzulassen. Im „gewöhnlichen" ringtheoretischen Fall ist dagegen ein Ring S ohne Operatoren angegeben; für die A_i sind dann die Unterringe von S zuzulassen, ferner bezeichnet $A_{\mathfrak{M}}$ jetzt den durch die Elemente der A_i ($i \in \mathfrak{M}$) erzeugten Unterring von S.

Unsere Definitionen lassen sich für den gruppen- und ringtheoretischen Fall auch „dualisieren". Das soll einfach darin bestehen, daß wir in (1) $A_{\mathfrak{M}}$ als den Durchschnitt der A_i ($i \in \mathfrak{M}$) deuten und $(A_{\mathfrak{M}})$ durch den Index $(S : A_{\mathfrak{M}})$ von $A_{\mathfrak{M}}$ in S ersetzen. Die Erweiterung dieses „dualen" Falles für unendlich viele $A_i, A_2, \ldots$ soll ähnlich wie bei (2) durch Grenzübergang und analytische Fortsetzung geschehen. Die so entstehenden $\zeta(z)$ nennen wir ebenfalls Zetafunktionen, genauer *duale Zetafunktionen* (in der Algebra).

Leicht sieht man, daß die Dedekindschen Zetafunktionen der algebraischen Zahlentheorie als sehr einfacher Spezialfall hierher gehören. Man bezeichne nämlich mit S den Ring der ganzen Zahlen eines absolut algebraischen Zahlkörpers vom endlichen Grade und nehme für $A_1, A_2, \ldots$ die sämtlichen Primideale von S. Für die entsprechende (duale) Zetafunktion gilt dann offenbar

$$\zeta(z; A_1, \ldots, A_n) = \prod_{i=1}^{n} (1 - N(A_i)^{-z})^{-1},$$

wobei $N(A_i)$ den Index $(S : A_i)$, d. h. die Norm von A_i bezeichnet. Hiernach ist jetzt $\zeta(z; A_1, A_2, \ldots)$ in der Tat mit der Dedekindschen Zetafunktion identisch.

In dieser Arbeit wollen wir, wie gesagt, näher nur die gruppentheoretischen Zetafunktionen für die endlichen Abelschen Gruppen (kurz den endlich-Abelschen Fall) untersuchen. In diesem Fall sind die zwei Arten Zetafunktionen nichtverschieden, denn die anfangs definierten Zetafunktionen angewendet auf die Charaktergruppe der vorgelegten Gruppe stimmen mit den dualen Zetafunktionen für die letztere Gruppe überein, wie man das leicht sieht. Im unendlich Abelschen Falle ist es natürlich anders, aber wir bemerken, daß die Riemannsche Zetafunktion auch als Spezialfall unter den dualen Zetafunktionen für die unendlichen Abelschen Gruppen enthalten ist. Sind nämlich $A_1, A_2, \ldots$ die sämtlichen Untergruppen von Primzahlindex in der unendlichen zyklischen Gruppe, so stimmt die duale Zetafunktion $\zeta(z; A_1, A_2, \ldots)$ mit der eben betrachteten Dedekindschen Zetafunktion für den Fall des rationalen Zahlkörpers, d. h. mit der Riemannschen Zetafunktion überein.

Kehren wir uns nunmehr dem endlich-Abelschen Fall zu. Mit G bezeichnen wir eine endliche Abelsche Gruppe. Die $A_1, \ldots, A_n$ in (1) sollen beliebige Untergruppen von G bedeuten. Natürlich ist (1) so zu verstehen, daß im Fall $n=0$

$$\varrho(z; A_1, \ldots, A_n) = 1 \qquad (n=0) \tag{4}$$

gilt. Man beachte, daß alle $A_1, \ldots, A_n$ Untergruppen von $A_1 \ldots A_n$ sind, weshalb man sich stets auf den Fall $G = A_1 \ldots A_n$ beschränken dürfte, was wir doch der größeren Allgemeinheit halber nicht tun wollen.

Die Definition (1) ist jetzt für alle komplexen z sinnvoll, und zwar ist $\varrho(z)$ eine ganze transzendente Funktion. Schon aus diesem Grunde empfiehlt sich, durchweg $\varrho(z)$ statt $\zeta(z) = \varrho(z)^{-1}$ zu untersuchen. Hierdurch werden sich auch unsere Formeln einfacher gestalten.

Man bemerke

$$\varrho(0; A_1, \ldots, A_n) = 0 \qquad (n \geqq 1). \tag{5}$$

Wir schreiben (1) auch in einer mehr expliziten Form hin:

$$\varrho(z) = \varrho(z; A_1, \ldots, A_n) = \sum_{k=0}^{n} \sum_{1 \leqq i_1 \leqq \ldots \leqq i_k \leqq n} (-1)^k (A_{i_1} \ldots A_{i_k})^{-z}, \tag{6}$$

wobei wir berücksichtigt haben, daß das Produkt $A_{i_1} \ldots A_{i_k}$ eben die durch die Elemente der Faktoren erzeugte Untergruppe von G ist. (Auch für die nichtabelschen G gilt (6) mit der Änderung, daß man statt $A_{i_1} \ldots A_{i_k}$ mit gewohnter Bezeichnung $\{A_{i_1}, \ldots, A_{i_k}\}$ schreibt.)

Zum Beispiel lautet (6) für $n = 1, 2, 3$ so:

$$\varrho(z; A) = 1 - \frac{1}{(A)^z}, \tag{7}$$

$$\varrho(z; A, B) = 1 - \frac{1}{(A)^z} - \frac{1}{(B)^z} + \frac{1}{(AB)^z}, \tag{8}$$

$$\varrho(z; A, B, C) = 1 - \frac{1}{(A)^z} - \frac{1}{(B)^z} - \frac{1}{(C)^z} + \frac{1}{(AB)^z} + \frac{1}{(AC)^z} + \frac{1}{(BC)^z} - \frac{1}{(ABC)^z}. \tag{9}$$

An diese Beispiele wollen wir einige Bemerkungen knüpfen. Den sehr einfachen Fall (7) außer Acht gelassen, bemerken wir, daß auch Fall (8) kein nennenswertes Problem bietet. Bekanntlich gilt nämlich

$$(AB) = \frac{(A)(B)}{(A \cap B)}. \tag{10}$$

Hieraus ersieht man leicht, daß die Tripel $(A), (B), (AB)$ mit den sämtlichen Tripeln ad, bd, abd zusammenfallen, wo die Parameter a, b, d voneinander unabhängig alle natürlichen Zahlen durchlaufen. Hiernach stimmen die $\varrho(z; A, B)$ in ihrer Gesamtheit mit den Funktionen

$$1 - (ad)^{-z} - (bd)^{-z} + (abd)^{-z} \qquad (a, b, d = 1, 2, \ldots)$$

überein.

Dagegen ist es um (9) nicht mehr so einfach. Vor allem gilt nämlich für (ABC) eine ähnliche Formel wie (10) nicht. So z. B. ist im allgemeinen

$$(ABC) + \frac{(A)(B)(C)(A \cap B \cap C)}{(A \cap B)(A \cap C)(B \cap C)},$$

wovon man sich leicht überzeugen kann. Überhaupt scheint uns, daß die in (9) vorkommenden sieben Gruppenordnungen $(A), (B), (C), (AB), (AC), (BC), (ABC)$ ein nicht leicht übersehbares System bilden, weshalb von der Funktion (9) kein triviales Verhalten zu erwarten ist. Noch mehr bezieht sich das auf (1) im allgemeinen.

Fassen wir jetzt nur die reellen $z(\geqq 1)$ ins Auge. Von (8) liest man sofort ab, daß

$$0 \leqq \varrho(z; A, B) < 1 \qquad (z \geqq 1)$$

gilt. Ferner sieht man, daß „=" nur dann gilt, wenn mindestens das eine von (A) und (B) gleich 1 (und z beliebig) ist. Mit einiger Mühe bekommt man, daß auch

$$0 \leqq \varrho(z; A, B, C) < 1 \qquad (z \geqq 1)$$

zutrifft, daß aber „=" außer den trivialen Fällen „(A) oder (B) oder (C) gleich 1 und z beliebig" auch noch im Fall gilt, wo $z = 1$ ist und A, B, C die drei echten Untergruppen der nichtzyklischen Gruppe vierter Ordnung sind. In diesem Zusammenhang bemerken wir gleich schon hier, daß im allgemeinen die Untersuchung der Nullstellen von $\varrho(z; A_1, \ldots, A_n)$ für $z = 1, 2, \ldots$ sich für ein sehr interessantes Problem erweisen wird (s. unten).

Man würde allgemein zunächst nur fragen, ob es absolute Konstanten $c, C(>0)$ gibt, so daß stets

$$-c < \varrho(z; A_1, \ldots, A_n) < C \qquad (z \geqq 1) \tag{11}$$

gilt. Die Antwort wird verneinend, aber insbesondere für die ganzen z gewinnen wir die überraschenden Ungleichungen

$$0 \leqq \varrho(z; A_1, \ldots, A_n) < 1 \qquad (n \geqq 1;\ z = 1, 2, \ldots), \tag{12}$$

die wir als ein sehr bemerkenswertestes Resultat ansehen und den *Trägheitssatz* der Theorie der endlichen Abelschen Gruppen nennen. Um diesen zu würdigen bemerken wir, daß es sich in ihm um einen in dreierlei Hinsicht scharfen Satz handelt. *Erstens* gilt nämlich (12) für die endlichen nichtabelschen Gruppen nicht mehr, sogar ist für diese die linke Seite von (12) mit jedem c falsch, weshalb (12) eine echt Abelsche Tatsache ist. *Zweitens* ist (12) auch in dem Sinne scharf, daß in ihm sehr oft, und zwar in keineswegs trivialen Fällen „=" gilt. *Drittens* ist (12) auch in der Beziehung scharf, daß es für die reellen aber nichtganzen $z(>1)$ nicht mehr allgemein gilt, was wir für sehr auffallend finden; es wird sich sogar zeigen, daß dann beide Seiten von (11) falsch sind, wie man auch c, C wählt.

Eine interessante Folgerung aus obigem Trägheitssatz wird sein, daß zwischen unseren Zetafunktionen und dem Satz von HAJÓS [2][2] über die endlichen Abelschen Gruppen ein enger Zusammenhang besteht. Und zwar läßt sich dieser tiefliegende Satz in einer äquivalenten Form als die notwendige und hinreichende Bedingung aussprechen, damit gewisse Zetafunktionen $\zeta(z)$ an der Stelle $z=1$ einen Pol, d. h. die entsprechenden reziproken Funktionen $\varrho(z)=\zeta(z)^{-1}$ eine Nullstelle haben.

Das gesagte beleuchtet die Bedeutung unserer Zetafunktionen genügend, weshalb wir meinen, daß durch sie in der Algebra im allgemeinen, insbesondere aber jedenfalls in der Theorie der endlichen Abelschen Gruppen ein wichtiges Forschungsgebiet eröffnet wird.

Im § 2 stellen wir unsere Resultate bezüglich der gruppentheoretischen Zetafunktionen im endlich-Abelschen Fall zusammen (auf den Zusammenhang mit dem Satz von HAJÓS gehen wir aber erst in einer anderen Arbeit [3] ein). Im § 3 erfolgen die Beweise. Im § 4 machen wir einige Schlußbemerkungen auch über den allgemeinen Fall.

§ 2. Der endlich-Abelsche Fall

Mit ausnahme vom § 4 soll eine endliche Abelsche Gruppe G mit dem Einselement ε vorgelegt sein, ferner sollen $A_1, \ldots, A_n$ durchweg Untergruppen von G bezeichnen, die im allgemeinen nicht verschieden zu sein brauchen. Alle unseren nachfolgenden Sätze formulieren wir bequemlichkeitshalber bezüglich $\varrho(z)=\zeta^{-1}(z)$ statt $\zeta(z)$. Die Sätze 1 bis 7 werden fast trivial sein, dagegen drücken die Sätze 8 bis 12 tiefliegende Sachverhalte aus.

SATZ 1. (Vertauschungsregel.) *Es ist* $\varrho(z; A_1, \ldots, A_n)$ *von der Reihenfolge der* A_n *unabhängig.*

SATZ 2. (Kürzungsregel.) *Enthält* A_1 *ein* A_i $(i=2, \ldots, n)$, *so gilt*

$$\varrho(z; A_1, \ldots, A_n) = \varrho(z; A_2, \ldots, A_n) \qquad (n \geqq 2). \tag{13}$$

KOROLLAR. *Behält man aus den Untergruppen* $A_1, \ldots, A_n$ *zunächst nur die verschiedenen und dann aus diesen nur die minimalen (d. h. die keine der übrigen enthalten), so gilt für die übriggebliebenen Untergruppen* $B_1, \ldots, B_r$:

$$\varrho(z; A_1, \ldots, A_n) = \varrho(z; B_1, \ldots, B_r).$$

Insbesondere wenn ein A_i *gleich* ε *ist, so gilt*

$$\varrho(z; A_1, \ldots, A_n) = 0. \tag{14}$$

(Es ist klar, daß umgekehrt $\varrho(z)=\varrho(z; A_1, \ldots, A_n)$ $(n \geqq 1)$ nichtkonstant ist, wenn alle A_i von ε verschieden sind, da dann $\varrho(+\infty)=1$ ist, dagegen nach (5) $\varrho(0)=0$ gilt.)

[2] Mit [ ] verweisen wir auf das Literaturverzeichnis am Ende unserer Arbeit.

SATZ 3. (Erste additive Eigenschaft.) *Für jede Untergruppe B von A_1 gilt*

$$\varrho(z; A_1, \ldots, A_n) = \varrho(z; B, A_2, \ldots, A_n) + (B)^{-z}\varrho(z; A_1/B, A_2B/B, \ldots, A_nB/B) \qquad (n \geqq 1; B \subseteq A_1), \tag{15}$$

wobei „/" die Faktorgruppe bezeichnet,

KOROLLAR. *Sind die A_i von ε verschieden, so gilt*

$$\varrho(z; A_1, \ldots, A_n) = \sum_{B_1, \ldots, B_n} (B)^{-z}\varrho(z; B_1'B/B, \ldots, B_n'B/B) \qquad (n \geqq 1), \tag{15'}$$

wobei B_i die von A_i verschiedenen Glieder einer beliebigen Kompositionsreihe von A_i durchläuft, B_i' das B_i enthaltende benachbarte Glied dieser Kompositionsreihe bezeichnet ($i = 1, \ldots, n$) und B für $B_1 \ldots B_n$ steht. (Es genügt diejenigen Glieder der rechten Seite von (15') beizubehalten, in denen B kein B_i' enthält, da die übrigen Glieder wegen (14) verschwinden. In den verbleibenden Gliedern sind alle Gruppen $B_i'B/B$ offenbar von Primzahlordnung.)

SATZ 4. (Zweite additive Eigenschaft.) *Es gilt*

$$\varrho(z; A_2, \ldots, A_n) = \varrho(z; A_1, \ldots, A_n) + (A_1)^{-z}\varrho(z; A_1A_2/A_1, \ldots, A_1A_n/A_1) \qquad (n \geqq 1). \tag{16}$$

SATZ 5. (Dritte additive Eigenschaft.) *Es gilt*

$$\sum_{\mathfrak{M} \subseteq \mathfrak{N}} (A_{\mathfrak{M}})^{-z}\varrho(z; A_x A_{\mathfrak{M}}/A_{\mathfrak{M}} A_y A_{\mathfrak{M}}/A_{\mathfrak{M}}, \ldots) = 1, \tag{17}$$

wobei $x, y, \ldots$ jedesmal die sämtlichen Elemente der Menge $\mathfrak{N} - \mathfrak{M}$ bezeichnen. (Man bemerke, daß die linke Seite von (17) insbesondere die den Fällen $\mathfrak{M} = \mathbf{0}$, $\mathfrak{N}$ entsprechenden Glieder $\varrho(z; A_1, \ldots, A_n)(A_1 \ldots A_n)^{-z}$ enthält.)

Zur Beleuchtung von (17) schreiben wir den Fall $n = 3$ explizit hin:

$$\varrho(z; A, B, C) + (A)^{-z}\varrho(z; AB/A, AC/A) + (B)^{-z}\varrho(z; AB/B, BC/B) + (C)^{-z}\varrho(z; AC/C, BC/C) + (AB)^{-z}\varrho(z; ABC/AB) + (AC)^{-z}\varrho(z; ABC/AC) + (BC)^{-z}\varrho(z; ABC/BC) + (ABC)^{-z} = 1.$$

SATZ 6. (Multiplikative Eigenschaft.) *Haben die Produkte $A_1 \ldots A_i$, $A_{i+1} \ldots A_n$ nur das Element ε gemein, so gilt*

$$\varrho(z; A_1, \ldots, A_n) = \varrho(z; A_1, \ldots, A_i)\,\varrho(z; A_{i+1}, \ldots, A_n). \tag{18}$$

SATZ 7. (Additiv-multiplikative Eigenschaft.) *Es gilt*

$$\varrho(z; A_1, \ldots, A_n) = \varrho(z; A_2, \ldots, A_n) + \varrho(z; A_1, A_3, \ldots, A_n) - \varrho(z; A_1A_2, A_3, \ldots, A_n) \qquad (n \geqq 2). \tag{19}$$

Um den Satz 8 formulieren zu können, definieren wir eine Funktion $\psi(z, G)$ folgenderweise. Ist $G = P$ eine p-Gruppe und t die Anzahl ihrer Invarianten, so sei

$$\psi(z, P) = \prod_{l=0}^{t-1} (1 - p^{l-z}) = (1 - p^{-z})(1 - p^{1-z}) \ldots (1 - p^{t-1-z}) \qquad (t \geqq 1). \tag{20}$$

Ist ferner G das direkte Produkt von $P_1, \dots, P_s$, wobei P_k eine p_k-Gruppe ist und $p_1, \dots, p_s$ verschiedene Primzahlen sind, so sei

$$\psi(z, G) = \prod_{k=1}^{s} \psi(z, P_k). \tag{21}$$

(Im Fall $G = \varepsilon$ ist natürlich $\psi(z, G) = 1$ zu verstehen. Im Fall $G \neq \varepsilon$ ist $\psi(z, G)$ nichtkonstant.)

SATZ 8. (Hauptsatz der Zetafunktionen für endliche Abelsche Gruppen.) *Sind $A_1, \dots, A_n$ Untergruppen der endlichen Abelschen Gruppe G, so gilt*

$$\varrho(z; A_1, \dots, A_n) = \sum_{A_1, \dots, A_n \subseteq H \subseteq G} (H)^{-z} \psi(z, G/H) \qquad (n \geqq 1), \tag{22}$$

wobei man über alle Untergruppen H von G zu summieren hat, die kein A_i enthalten (somit von G verschieden sind).

BEMERKUNG 1. Aus diesem Satz folgt, daß die rechte Seite von (22) für alle G gleich ist, die das Produkt $A_1 \dots A_n$ enthalten, was man leicht auch direkt beweisen könnte. Hat man es mit nur einem $\varrho(z; A_1, \dots, A_n)$ zu tun, so wendet man (22) am bequemsten mit $G = A_1 \dots A_n$ an. Auch die allgemeinere Annahme $G \supseteq A_1 \dots A_n$ kann in (22) von Vorteil sein, wenn man nämlich gleichzeitig mehrere $\varrho(z; A_1, \dots, A_n)$ betrachtet, in denen die A_i irgendwelche Untergruppen einer gegebenen endlichen Abelschen Gruppe G sind.

BEMERKUNG 2. Ist ein A_i gleich ε, so ist die Summe in (22) leer, weshalb dann (22) eben in (14) übergeht. Sind dagegen alle A_i von ε verschieden, so ist in (22) der Fall $H = \varepsilon$ gewiß zuzulassen. Hiernach folgt aus (22):

$$\varrho(z; A_1, \dots, A_n) = \psi(G) + \sum_{A_1, \dots, A_n \subseteq H \subset G, \neq \varepsilon} (H)^{-z} \psi(z, G/H) \tag{23}$$

$$(n \geqq 1; A_1, \dots, A_n \neq \varepsilon).$$

BEMERKUNG 3. Sind insbesondere alle A_i von p-ter Ordnung (p Primzahl), so folgt aus (20), (23) (mit Anwendung auf $G = A_1 \dots A_n$) offenbar

$$\varrho(z; A_1, \dots, A_n) = \prod_{l=0}^{t-1} (1 - p^{l-z}) + \sum_{k=1}^{t-1} \nu_k p^{-kz} \prod_{l=0}^{t-k-1} (1 - p^{l-z}) \tag{24}$$

$$(n \geqq 1; (A_1) = \dots = (A_n) = p; p^t = (A_1 \dots A_n)),$$

wobei ν_k die Anzahl derjenigen Untergruppen p^k-ter Ordnung von $A_1 \dots A_n$ bezeichnet, die kein A_i enthalten. Das erste Produkt in (24) verschwindet dann und nur dann, wenn z eine der ganze Zahlen $0, \dots, t-1$ ist; in diesen Fällen verschwindet auch das zweite Produkt stets, wenn außerdem $t-k-1 \geqq z$, d. h. $k \leqq t-z-1$ gilt, weshalb aus (24) folgt

$$\varrho(z; A_1, \dots, A_n) = \sum_{k=t-z}^{t-1} \nu_k p^{-kz} \prod_{l=0}^{t-k-1} (1 - p^{l-z}) \tag{25}$$

$$(n \geqq 1; (A_1), \dots, (A_n) | p; p^t = (A_1 \dots A_n); z = 0, \dots, t-1),$$

wobei auch der Fall zugelassen werden durfte, daß einige der A_i gleich ε sind, da dann $\nu_1 = \cdots = \nu_{t-1} = 0$ ist, somit beide Seiten von (25) verschwinden. Z. B. für $z = 1, 2$ gelten also

$$\left.\begin{aligned} &(26) \qquad \varrho(1; A_1, \ldots, A_n) = \frac{(p-1)\nu_{t-1}}{p^t} && (t \geqq 2) \\ &(27) \qquad \varrho(2; A_1, \ldots, A_n) = \frac{(p^2-1)\nu_{t-1} + (p^2-1)(p^2-p)\nu_{t-2}}{p^{2t}} && (t \geqq 3) \end{aligned}\right\}$$

$$((A_1), \ldots, (A_n) | p; p^t = (A_1 \ldots A_n)).$$

BEMERKUNG 4. Wir nennen Satz 8 deshalb einen Hauptsatz, weil er für die Berechnung von $\varrho(z)$ eine der Definition (1) gegenüber wesentlich neue Möglichkeit bietet und einen Ausgangspunkt für weitere Schlüße bildet, worüber in den folgenden Sätzen 9 bis 12 die Rede sein wird.

SATZ 9. (Trägheitssatz.) *Es gilt*

$$(28) \qquad 0 \leqq \varrho(z; A_1, \ldots, A_n) < 1 \qquad (n \geqq 1; z = 1, 2, \ldots).$$

ZUSATZ. *Die rechte Seite von* (28) *gilt sogar mit*

$$\text{„} \leqq 1 - (\min((A_1), \ldots, (A_n)))^{-z} \text{“ } \textit{statt} \text{ „} < 1 \text{“}.$$

BEMERKUNG 5. Wir nennen Satz 9 einen Trägheitssatz, weil nach ihm die $\varrho(z)$ in (28) nur sehr kleiner Schwankungen fähig sind. In der Einleitung haben wir schon gesagt, daß Satz 9 in dreierlei Hinsicht „scharf“ ist. Das wollen wir hier mit Beispielen bestätigen. *Erstens* bedeute jetzt G die Diedergruppe von der Ordnung $2p$ (p ungerade Primzahl). Diese hat p Untergruppen 2-ter Ordnung, die wir mit $A_1, \ldots, A_p$ bezeichnen. Wir wollen $\varrho(z; A_1, \ldots, A_p)$ berechnen. Dies besteht nach (1) aus gleichvielen Gliedern vom Vorzeichen „+“ bzw. „−“. Eins davon ist 1, weitere p Glieder sind gleich -2^{-z}, alle übrigen sind gleich $\pm(2p)^{-z}$, weshalb offenbar

$$\varrho(z; A_1, \ldots, A_p) = 1 - \frac{p}{2^z} + \frac{p-1}{(2p)^z}$$

gilt. Man sieht, daß die rechte Seite für kein $z = 1, 2, \ldots$ (sogar für kein $z \geqq 1$) von unten beschränkt ist, weshalb (28) für die endlichen *nichtabelschen* Gruppen falsch ist. *Zweitens* bezeichne G eine elementare Abelsche Gruppe von der Ordnung p^t (p Primzahl, $t \geqq 2$). Sind dann $A_1, \ldots, A_n$ die sämtlichen Untergruppen p-ter Ordnung von G, so sind in (25) alle ν_k gleich 0, weshalb jetzt in (28) für $z = 1, \ldots, t-1$ das Zeichen „=“ gilt. Man könnte mit Leichtigkeit ähnliche aber weniger einfache Beispiele in großer Zahl konstruieren (s. auch Bemerkung 6 und Satz 10). *Drittens* betrachten wir wieder das vorige $\varrho(z; A_1, \ldots, A_n)$, jetzt aber für die reellen $z (\geqq 1)$. Nach (24) gilt

$$\varrho(z; A_1, \ldots, A_n) = \prod_{l=0}^{t-1} (1 - p^{l-z}).$$

Man sieht, daß dies für die reellen z im Innern der Intervalle $(1,2),(2,3),\ldots,(t-2,t-1)$ vom abwechselnden Vorzeichen ist, daß also die linke Seite von (28) für die nichtganzen $z(>1)$ im allgemeinen falsch ist; da t beliebig groß sein kann, so sieht man sogar, daß — wie auch c, C gewählt werden — keine Seite von (11) allgemein gilt.

BEMERKUNG 6. Wir wollen genauer zusehen, was für Möglichkeiten vorliegen, damit eine positive ganze Zahl z eine Nullstelle von einem $\varrho(z)$, d. h. ein Pol der zugehörigen Zetafunktion $\zeta(z)=\varrho(z)^{-1}$ ist. (Diese Frage wird nicht nur an sich, sondern — wie in der Einleitung erwähnt wurde — auch im Zusammenhang mit dem Satz von HAJÓS interessant.) Man liest von (20) ab, daß $\psi(z,P)$ für $z=0,1,\ldots$ nichtnegativ, und zwar für $z=0,\ldots,t-1$ gleich 0, für $z=t,t+1,\ldots$ positiv ist. Ähnliches gilt nach (21) auch über $\psi(z,G)$, wenn wir t folgenderweise definieren:

$$t=t(G)=\max(t_1,\ldots,t_s), \tag{29}$$

wobei t_i die Anzahl der Invarianten des Faktors P_i in der Darstellung von G als ein direktes Produkt $P_1\ldots P_s$ von p_i-Gruppen P_i bedeutet ($p_1,\ldots,p_s$ verschiedene Primzahlen). Man bemerke auch, daß stets $t(G/H)\leqq t(G)$ gilt. Nun gilt nach dem Gesagten, daß der Summand in (22) für $z=0,1,\ldots$ nichtnegativ, und zwar für $z=0,\ldots,t(G/H)-1$ gleich 0, für $z=t(G/H)$, $t(G/H)+1,\ldots$ positiv ist. So kommen wir zum folgenden:

SATZ 10. *Um die positiven ganzzahligen Nullstellen von $\varrho(z;A_1,\ldots,A_n)$ zu bestimmen, wobei $A_1,\ldots,A_n$ ($n\geqq 1$) von 1 verschiedene Untergruppen einer endlichen Abelschen Gruppe G sind, verfahre man so: Man nehme alle Untergruppen H von G, die kein A_i enthalten, bilde für sie*

$$t_0=\min_H t(G/H), \tag{30}$$

wobei $t(G/H)$ nach (29) zu verstehen ist. Dann verschwindet $\varrho(z;A_1,\ldots,A_n)$ für $z=1,\ldots,t_0-1$ und ist für $z=t_0,t_0+1,\ldots$ positiv.

KOROLLAR. *Die sämtlichen positiven ganzen Nullstellen eines nichtkonstanten $\varrho(z;A_1,\ldots,A_n)$ bilden eine lückenlose Folge $1,\ldots,t_0-1$, wobei für t_0 (30) gilt.* (Im Fall $t_0=1$, und nur in diesem, gibt es keine solchen Nullstellen.)

BEMERKUNG 7. Wir formulieren eine leichte Folgerung vom Satz 10 besonders: *Für beliebige Untergruppen $A_1,\ldots,A_n$ von G gilt $\varrho(1;A_1,\ldots,A_n)=0$ dann und nur dann, wenn jede Untergruppe H von G mit zyklischer Faktorgruppe G/H mindestens ein A_i enthält.* (Hierbei darf unter den A_i auch ε vorkommen.)

SATZ 11. (Monotone Eigenschaft.) *Stets gilt*

$$\varrho(z;A_1,\ldots,A_n)\leqq\varrho(z;A_1,\ldots,A_{n-1})\qquad(n\geqq 1;\ z=1,2,\ldots). \tag{31}$$

Die vorigen bilden die Grundlagen einer elementaren Theorie der betrachteten Zetafunktionen. Wir haben nicht untersucht, wie weit für sie durch Einsetzung von funktionentheoretischen Hilfsmitteln die Begründung einer höheren Theorie möglich ist. Leicht scheint das wegen der sich in der freien Wahl des Untergruppensystems $A_1, \ldots, A_n$ offenbarenden, ungeheuer großen Allgemeinheit unserer Zetafunktionen nicht zu sein. Hier wollen wir nur eine ganz leichte funktionentheoretische Anwendung machen, die so entstehen wird, daß man den Differentialquotienten $\varrho'(0)$ sowohl nach (1), als auch nach (22) berechnet und die zwei Resultate einander gleichsetzt. So werden wir zum folgenden Satz kommen:

SATZ 12. *Man definiere zu* (1) *ähnlich das Produkt*

$$\tau(A_1, \ldots, A_n) = \prod_{\mathfrak{M} \subseteq \mathfrak{N}} (A_{\mathfrak{M}})^{(-1)^{(\mathfrak{M})}} \qquad (n \geqq 1), \tag{32}$$

wobei $A_1, \ldots, A_n$ Untergruppen einer endlichen Abelschen Gruppe G sind. Bezeichnen $p_1, \ldots, p_s$ die verschiedenen Primteiler in der Ordnung von G, so gilt

$$\tau(A_1, \ldots, A_n) = \prod_{k=1}^{s} p_k^{-\sum_{H_k}' \prod_{l=1}^{t(G/H_k)-1} (1-p_k^l)}, \tag{33}$$

wobei H_k diejenigen Untergruppen von G durchläuft, die kein A_i enthalten und für die die Faktorgruppe G/H_k eine p_k-Gruppe ist, ferner $t(G/H_k)$ die Anzahl der Invarianten von G H_k bezeichnet.

Man bemerke, daß (33) beliebig groß sein und auch beliebig nahe an 0 liegen kann. Ist nämlich G eine elementare Abelsche p-Gruppe von der Ordnung p^t und sind $A_1, \ldots, A_n$ die sämtlichen Untergruppen p-ter Ordnung von G, so geht (33) in

$$\tau(A_1, \ldots, A_n) = p^{-(1-p)(1-p^2)\ldots(1-p^{t-1})}$$

über, woraus die Behauptung folgt.

§ 3. Beweis

Der Satz 1 ist eine Folgerung aus der Definition (1).

Für das nächtsfolgende bezeichnen wir mit $\mathfrak{N}_1$ die Menge der Zahlen $2, \ldots, n$ $(n \geqq 1)$; insbesondere für $n = 1$ ist also $\mathfrak{N}_1 = \mathfrak{0}$. Dann gilt nach (1) offenbar

$$\varrho(z; A_1, \ldots, A_n) = \sum_{\mathfrak{M} \subseteq \mathfrak{N}_1} (-1)^{(\mathfrak{M})} (A_{\mathfrak{M}})^{-z} - \sum_{\mathfrak{M} \subseteq \mathfrak{N}_1} (-1)^{(\mathfrak{M})} (A_1 A_{\mathfrak{M}})^{-z} \quad (n \geqq 1). \tag{34}$$

Um nunmehr Satz 2 zu beweisen dürfen wir wegen Satz 1 annehmen, daß z. B. $A_1 \supseteq A_2$ gilt. Wir fassen die sämtlichen $\mathfrak{M} (\subseteq \mathfrak{N}_1)$ in Paare $\mathfrak{M}_1, \mathfrak{M}_2$,

so daß jedesmal $\mathfrak{M}_1, \mathfrak{M}_2$ voneinander nur im Element 2 unterscheiden. Dann gilt $(\mathfrak{M}_1)-(\mathfrak{M}_2) = \pm 1$, $A_1 A_{\mathfrak{M}_1} = A_1 A_{\mathfrak{M}_2}$, folglich verschwindet jetzt die zweite Summe auf der rechten Seite von (34). Hiernach geht (34) in (13) über, womit Satz 2 bewiesen ist.

Die erste Hälfte des Korollars folgt unmittelbar aus Satz 2. Da für $A = \varepsilon$ offenbar $\varrho(z; A) = 0$ gilt, so ist auch (14) richtig.

Zum Beweis von Satz 3 wenden wir (34) mit B statt A_1 an. Nach Subtraktion beider Gleichungen ergibt sich

$$\varrho(z; A_1, \ldots, A_n) - \varrho(z; B, A_2, \ldots, A_n) = \sum_{\mathfrak{M} \subseteq \mathfrak{N}_1} (-1)^{(\mathfrak{M})} ((B A_{\mathfrak{M}})^{-z} - (A_1 A_{\mathfrak{M}})^{-z}).$$

Wenn nun $B \subseteq A_1$ gilt, so können wir die rechte Seite in

$$(B)^{-z} \sum_{\mathfrak{M} \subseteq \mathfrak{N}_1} (-1)^{(\mathfrak{M})} ((B A_{\mathfrak{M}}/B)^{-z} - (A_1 A_{\mathfrak{M}}/B)^{-z})$$

umformen. Dies ist wegen $A_1 = B A_1$ offenbar gleich

$$(B)^{-z} \sum_{\mathfrak{M} \subseteq \mathfrak{N}} (-1)^{(\mathfrak{M})} (B A_{\mathfrak{M}}/B)^{-z},$$

d. h. nach (1) gleich

$$(B)^{-z} \varrho(z; B A_1/B, \ldots, B A_n/B).$$

Nach Berücksichtigung von $B A_1 = A_1$ folgt hieraus Satz 3.

Das Korollar ließe sich aus Satz 3 durch Induktion gewinnen, bequemer beweisen wir es aber direkt. Die rechte Seite von (15′) läßt sich nach (1) so schreiben:

$$\sum_{B_1, \ldots, B_n} \sum_{\mathfrak{M} \subseteq \mathfrak{N}} (B)^{-z} (-1)^{(\mathfrak{M})} (B'_{\mathfrak{M}} B/B)^{-z},$$

wobei $B'_{\mathfrak{M}}$ das Produkt der B'_i $(i \in \mathfrak{M})$ bezeichnet. Hierfür schreibt sich:

$$\sum_{B_1, \ldots, B_n} \sum_{\mathfrak{M} \subseteq \mathfrak{N}} (-1)^{(\mathfrak{M})} (B'_{\mathfrak{M}} B)^{-z}.$$

Es ist

$$B'_{\mathfrak{M}} B = B'_{\mathfrak{M}} B_1, \ldots, B_n = B'_{\mathfrak{M}} B_{\mathfrak{N}-\mathfrak{M}},$$

wobei $B_{\mathfrak{K}}$ für $\mathfrak{K} \subseteq \mathfrak{M}$ das Produkt der B_i $(i \in \mathfrak{K})$ bezeichnet. Hiernach geht die vorige Summe in

$$\sum_{B_1, \ldots, B_n} \sum_{\mathfrak{M} \subseteq \mathfrak{N}} (-1)^{(\mathfrak{M})} (B'_{\mathfrak{M}} B_{\mathfrak{N}-\mathfrak{M}})^{-z}$$

über. Jedes $B'_{\mathfrak{M}} B_{\mathfrak{N}-\mathfrak{M}}$ entsteht so aus $B_1, \ldots, B_n$, daß man die B_i $(i \in \mathfrak{M})$ durch B'_i ersetzt. Hieraus ist leicht zu sehen, daß die vorige Summe nach Streichung entgegengesetzt gleicher Glieder in

$$\sum_{\mathfrak{M} \subseteq \mathfrak{N}} (-1)^{(\mathfrak{M})} (A_{\mathfrak{M}})^{-z}$$

übergeht. Das besagt nach (1) die Richtigkeit von (15′), womit wir das Korollar vom Satz 3 bewiesen haben.

Satz 4 ist im Fall $n=1$ trivial. Der Fall $n \geqq 2$ entsteht aus dem Fall $n=2$ vom Satz 3 so, daß man diesen für A_1, A_2, A_1 statt A_1, B_1 anwendet und Satz 2 berücksichtigt.

Um Satz 5 zu beweisen drücken wir die linke Seite von (17) nach (1) so aus:

$$\sum_{\mathfrak{M} \subseteq \mathfrak{N}} (A_{\mathfrak{M}})^{-z} \sum_{\mathfrak{M}_1 \subseteq \mathfrak{N}-\mathfrak{M}} (-1)^{(\mathfrak{M}_1)} (A_{\mathfrak{M}_1} A_{\mathfrak{M}} / A_{\mathfrak{M}})^{-z}.$$

Hierfür schreibt sich

$$\sum_{\mathfrak{M} \subseteq \mathfrak{N}} (A_{\mathfrak{M}})^{-z} \sum_{\mathfrak{M}_1 \subseteq \mathfrak{N}-\mathfrak{M}} (-1)^{(\mathfrak{M}_1)} (A_{\mathfrak{M}_1} A_{\mathfrak{M}})^{-z}.$$

Nach Umordnung der Glieder entsteht hieraus (wegen $A_{\mathfrak{M}_1} A_{\mathfrak{M}} = A_{\mathfrak{M}_1 \cup \mathfrak{M}}$)

$$\sum_{\mathfrak{A} \subseteq \mathfrak{N}} \sum_{\mathfrak{B} \subseteq \mathfrak{A}} (-1)^{(\mathfrak{B})} (A_{\mathfrak{A}})^{-z}.$$

Die innere Summe ist 1, wenn $\mathfrak{A}=0$ ist, und verschwindet für $\mathfrak{A} \neq \mathbf{0}$. Damit haben wir Satz 5 bewiesen.

Zum Beweis von Satz 6 zerlegen wir $\mathfrak{N}$ in zwei komplementäre Untermengen $\mathfrak{N}', \mathfrak{N}''$, so daß diese aus $1, \ldots, i$ bzw. $i+1, \ldots, n$ bestehen. Nach (1) gilt dann

$$\varrho(z; A_1, \ldots, A_n) = \sum_{\mathfrak{M}_1 \subseteq \mathfrak{N}'} \sum_{\mathfrak{M}_2 \subseteq \mathfrak{N}''} (-1)^{(\mathfrak{M}_1)+(\mathfrak{M}_2)} (A_{\mathfrak{M}_1 \cup \mathfrak{M}_2})^{-z}. \tag{35}$$

Es ist $A_{\mathfrak{M}_1 \cup \mathfrak{M}_2} = A_{\mathfrak{M}_1} A_{\mathfrak{M}_2}$. Andererseits gilt $A_{\mathfrak{M}_1} \subseteq A_1 \ldots A_i$, $A_{\mathfrak{M}_2} \subseteq A_1 \ldots A_n$, somit folgt aus der Annahme: $(A_{\mathfrak{M}_1 \cup \mathfrak{M}_2}) = (A_{\mathfrak{M}_1})(A_{\mathfrak{M}_2})$. Hiernach geht (35) in (18) über, womit Satz 6 bewiesen ist.

Um Satz 7 zu beweisen bezeichnen wir mit $\mathfrak{N}_2$ die Menge der Zahlen $3, \ldots, n$ ($n \geqq 2$). Nach (1) gilt dann

$$\varrho(z; A_1, \ldots, A_n) = \sum_{\mathfrak{M} \subseteq \mathfrak{N}} (-1)^{(\mathfrak{M})} ((A_{\mathfrak{M}})^{-z} - (A_1 A_{\mathfrak{M}})^{-z} - (A_2 A_{\mathfrak{M}})^{-z} + (A_1 A_2 A_{\mathfrak{M}})^{-z}).$$

Ähnlich gelten

$$\varrho(z; A_1, A_2, A_3, \ldots, A_n) = \sum_{\mathfrak{M} \subseteq \mathfrak{N}_2} (-1)^{(\mathfrak{M})} ((A_{\mathfrak{M}})^{-z} - (A_1 A_2 A_{\mathfrak{M}})^{-z}),$$

$$\varrho(z; A_i, A_3, \ldots, A_n) = \sum_{\mathfrak{M} \subseteq \mathfrak{N}_2} (-1)^{(\mathfrak{M})} ((A_{\mathfrak{M}})^{-z} - (A_i A_{\mathfrak{M}})^{-z}) \qquad (i = 1, 2).$$

Aus diesen Gleichungen folgt Satz 7.

Wir wollen Satz 8 beweisen. Da die rechte Seite von (22) und wegen des Korollars vom Satz 2 auch die linke Seite von der Multiplizität der A_i unabhängig ist, so dürfen wir annehmen, daß alle A_i verschieden sind.

Für jede Untergruppe K von G bezeichnen wir mit $N_k(K)$ die Anzahl derjenigen Mengen $\mathfrak{M} (\subseteq \mathfrak{N})$, für die

$$(\mathfrak{M}) = k, \; A_{\mathfrak{M}} = K \qquad (K \subseteq G; \; k = 0, \ldots, n) \tag{36}$$

ist. Nach (1) gilt dann

$$\varrho(z) = \varrho(z; A_1, \ldots, A_n) = \sum_{k=0}^{n} \sum_{K \subseteq G} (-1)^k N_k(K) (K)^{-z}. \tag{37}$$

Mit $\mathfrak{R}$ bezeichnen wir die Menge derjenigen A_i, die in K enthalten sind. Nimmt man aus $\mathfrak{R}$ auf jede Art k Untergruppen heraus und bildet ihr Produkt, so entsteht jede Untergruppe H von K genau $N_k(H)$-mal. Hiernach gilt

$$\binom{(\mathfrak{R})}{k} = \sum_{H \subseteq K} N_k(H) \qquad (K \subseteq G;\ k = 0, \ldots, n), \tag{38}$$

wobei links der Binomialkoeffizient steht und die Summe über alle Untergruppen H von K zu erstrecken ist.

Wir betrachten mit DELSARTE [1] irgendzwei Funktionen $f(G), F(G)$, die für alle endlichen Abelschen Gruppen G erklärt sind, so daß beide Funktionswerte in einen Modul gehören und nur vom Typus von G abhängen (d. h. $f(G) = f(G'), F(G) = F(G')$ gilt, wenn G, G' isomorph sird). Gilt für alle G

$$F(G) = \sum_{H \subseteq G} f(G),$$

wobei man über alle Untergruppen H von G zu summieren hat, so gilt nach dem wichtigen Satz von DELSARTE [1] folgende Verallgemeinerung der Möbiusschen Umkehrungsformel:

$$f(G) = \sum_{H \subseteq G} \mu(G/H)\, F(H),$$

wobei die Funktion μ von MÖBIUS—DELSARTE folgende Bedeutung hat: Ist Q eine elementare endliche Abelsche p-Gruppe, d. h. das direkte Produkt von t Gruppen p-ter Ordnung, so ist

$$\mu(Q) = (-1)^t p^{\binom{t}{2}} \qquad (t = 0, 1, \ldots).$$

Ist ferner G das direkte Produkt von $Q_1, \ldots, Q_s$, wobei Q_i eine elementare endliche Abelsche p_i-Gruppe ist und $p_1, \ldots, p_s$ verschiedene Primzahlen sind, so ist

$$\mu(G) = \mu(Q_1) \ldots \mu(Q_s).$$

Ist endlich G eine sonstige endliche Abelsche Gruppe, d. h. eine solche, deren mindestens eine Invariante durch ein Primzahlquadrat teilbar ist, so ist $\mu(G) = 0$.

Hiernach folgt aus (38)

$$N_k(K) = \sum_{H \subseteq K} \mu(K/H) \binom{(\mathfrak{H})}{k} \qquad (K \subseteq G; k = 0, \ldots, n),$$

wobei $\mathfrak{H}$ die Menge der in H enthaltenen A_i bezeichnet. Nach Einsetzung in (37) entsteht

$$\varrho(z) = \sum_{k=0}^{n} \sum_{K \subseteq G} \sum_{H \subseteq K} (-1)^k \mu(K/H) \binom{(\mathfrak{H})}{k} (K)^{-z}. \tag{39}$$

Es ist

$$\sum_{k=0}^{n}\binom{(\mathfrak{H})}{k}$$

gleich 1, wenn $(\mathfrak{H})=0$ ist, d. h. wenn H kein A_i enthält. Im anderen Fall dagegen verschwindet diese Summe. Wird dies in (39) berücksichtigt, so entsteht nach leichter Umordnung der Glieder

$$\varrho(z)=\sum_{H}{}' \sum_{H\subseteq K\subseteq G} \mu(K/H)\,(K)^{-z},$$

wo man über die kein A_i enthaltenden Untergruppen H von G und alle Gruppen K zwischen H und G zu summieren hat.

Wegen $(K)=(H)(K/H)$ läßt sich hierfür

$$\varrho(z)=\sum_{H}{}' (H)^{-z} f(z; G/H) \tag{40}$$

schreiben mit

$$f(z, G)=\sum_{L\subseteq G} \mu(L)\,(L)^{-z}, \tag{41}$$

wobei man über alle Untergruppen L von G zu summieren hat. Wir haben nur noch zu zeigen, daß $f(z, G)$ der bei (20), (21) definierten Funktion $\psi(z, G)$ gleich ist, denn dann stimmt (40) mit (22) überein.

Haben $P_1, \ldots, P_s$ die Bedeutung wie in (21), so folgt aus (41) und der Definition von $\mu(G)$ leicht

$$f(z, G)=\prod_{i=1}^{s} f(z, P_i).$$

Dies mit (21) verglichen sehen wir, daß es genügt $f(z, P)=\psi(z, P)$ zu zeigen, wobei P eine endliche Abelsche p-Gruppe ist. Wir bezeichnen mit Q die größte, in P enthaltene elementare Gruppe. Aus (41) und der Definition von μ folgt

$$f(z, P)=\sum_{L\subseteq Q} \mu(L)\,(L)^{-z}=f(z, Q). \tag{42}$$

Wir setzen

$$(Q)=p^t \qquad (t\geqq 1).$$

Dann ist t zugleich die Anzahl der Invarianten von P. Wir setzen zur Abkürzung

$$\begin{bmatrix} k \\ l \end{bmatrix}=\frac{(p^k-1)(p^{k-1}-1)\cdots(p^{k-l+1}-1)}{(p-1)(p^2-1)\cdots(p^l-1)} \qquad (0\leqq l\leqq k)$$

und erweitern diese Definition durch die Vereinbarung $\begin{bmatrix} k \\ l \end{bmatrix}=0$ für $l<0$ und für $l>k$. Bekanntlich ist $\begin{bmatrix} t \\ l \end{bmatrix}$ die Anzahl der Untergruppen p^l-ter Ordnung von Q $(l=0, \ldots, t)$. Für jede solche Gruppe L gilt $\mu(L)=(-1)^l\, p^{\binom{l}{2}}$, somit

folgt aus (42)

$$f(z, P) = \sum_{l=0}^{t} (-1)^l p^{\binom{l}{2}} \begin{bmatrix} t \\ l \end{bmatrix} p^{-lz}.$$

Wir bezeichnen die rechte Seite kurz mit $g(z, t)$. Es genügt zu zeigen, daß $g(z, t)$ für $t = 0, 1, \ldots$ mit der rechten Seite von (20) übereinstimmt. Für $t = 0$ ist das klar, da $g(z, 0) = 1$ ist. Da ferner der Summand für $l < 0$ und für $l > t$ verschwindet, so gilt

$$g(z, t) = \sum_{l} (-1)^l p^{\binom{l}{2}} \begin{bmatrix} t \\ l \end{bmatrix} p^{-lz},$$

wobei man über alle $l = 0, \pm 1, \pm 2, \ldots$ summieren darf. Wegen

$$\begin{bmatrix} t+1 \\ l \end{bmatrix} = p^l \begin{bmatrix} t \\ l \end{bmatrix} + \begin{bmatrix} t \\ l-1 \end{bmatrix}$$

bekommt man

$$g(z+1, t+1) = \sum_{l} (-1)^l p^{\binom{l}{2}} \begin{bmatrix} t \\ l \end{bmatrix} p^{-lz} + \sum_{l} (-1)^l p^{\binom{l}{2}} \begin{bmatrix} t \\ l-1 \end{bmatrix} p^{-lz-l} \quad (t \geqq 0).$$

Wird im zweiten Summand $l+1$ statt l geschrieben, so erkennt man, daß

$$g(z+1, t+1) = (1 - p^{-z-1}) g(z, t) \qquad (t \geqq 0)$$

ist. Dies bestätigt (20) mit vollständiger Induktion nach t, womit Satz 8 bewiesen ist.

Bezüglich des Satzes 9 beweisen wir zuerst die linke Seite von (28). Diese ist eine direkte Folgerung aus Satz 8, da $\psi(z, G)$ nach (20), (21) für $z = 1, 2, \ldots$ (sogar für $z = 0, 1, \ldots$) nichtnegativ ist. Aus der linken Seite von (28) folgt wegen Satz 5 auch die rechte Seite. Somit haben wir Satz 9 bewiesen.

Zum Beweis des Zusatzes nehmen wir Satz 11 vorweg. Aus diesem folgt

$$\varrho(z; A_1, \ldots, A_n) \leqq \varrho(z; A_i) \qquad (z = 1, 2, \ldots;\ i = 1, \ldots, n).$$

Die rechte Seite ist nach (1) gleich $1 - (A_i)^{-z}$. Somit haben wir den Zusatz von Satz 9 bewiesen.

Die Richtigkeit von Satz 10 (und die des Korollars) folgt aus Satz 8, wie wir es in der Bemerkung 6 schon auseinandergesetzt haben.

Satz 11 folgt sofort aus Satz 4 und der ersten Hälfte von Satz 9.

Um Satz 12 zu beweisen bemerken wir zunächst, daß nach (20)

$$\psi(0, P) = 0, \quad \psi'(0, P) = \log p \prod_{l=1}^{t-1} (1 - p^l) \qquad (t \geqq 1) \tag{43}$$

gilt. Aus (43_1) folgt nach (21), daß $\psi'(0, G)$ verschwindet, wenn (G) durch mindestens zwei verschiedene Primzahlen teilbar (d. h. $s \geqq 2$) ist. Auch gilt $\psi(0, G) = 0$, wenn $(G) \neq 1$ ist. Differenziert man also (1) und (22) an der

Stelle $z=0$, so folgt aus (43_2) nach Gleichsetzung der rechten Seiten

$$-\sum_{\mathfrak{M}\subseteq\mathfrak{N}}(-1)^{(\mathfrak{M})}\log(A_{\mathfrak{M}})=\sum_{H}{}'\log p\prod_{l-1}^{t-1}(1-p^l),$$

wobei man über die Untergruppen $H(\neq G)$ von G zu summieren hat, die kein A_i enthalten und für die die Ordnung (G/H) eine Primzahlpotenz p^t $(t\geqq 1)$ ist (so daß also p und t sich von Glied zu Glied ändern). Hieraus folgt leicht die Gleichheit der rechten Seiten von (32), (33), womit wir Satz 12 bewiesen haben.

§ 4. Schlußbemerkungen

Wir wollen noch einige kurze Bemerkungen über unsere Zetafunktionen machen. Zunächst betrachten wir wieder den endlich-Abelschen Fall. Wie wir schon in der Einleitung bemerkt haben, halten wir den auf dem Hauptsatz (Satz 8) beruhenden Trägheitssatz (Satz 9) für das wichtigste Ergebnis der von uns in den §§ 2, 3 entwickelten „Theorie der Zetafunktionen für endliche Abelsche Gruppen". Über diese Theorie haben wir schon bemerkt, daß sie vom Fundamentalsatz der endlichen Abelschen Gruppen unabhängig ist, was wir jetzt bezüglich der Sätze 8—12 klarmachen wollen (für die Sätze 1—7 ist das ohnehin schon klar). Wohl bedeutet t in (20) die Anzahl der Invarianten der p-Gruppe P und diese Invarianten selbst bilden einen vom Fundamentalsatz abhängigen Begriff, aber t läßt sich auch unabhängig vom Fundamentalsatz durch $p^t=(Q)$ definieren, wobei Q die Gruppe derjenigen Elemente von P bedeutet, die von der Ordnung $\leqq p$ sind. Da ferner die (eindeutige) Zerlegung von G in seine Primärkomponenten $P_1,\ldots,P_s$ nichts mit dem Fundamentalsatz zu tun hat, so läßt sich auch die Funktion $\psi(z,G)$ in (21) vom Fundamentalsatz unabhängig definieren. Da endlich der Beweis des Satzes von DELSARTE [1] über die verallgemeinerte Umkehrungsformel von Möbius keinen unentbehrlichen Gebrauch vom Fundamentalsatz macht, so gilt ähnliches über den Beweis von unserem Satz 8. Es sind also auch die Sätze 9 bis 12, als Folgerungen aus Satz 8, vom Fundamentalsatz ebenfalls unabhängig. Diese Bemerkung ist auch deshalb interessant, weil eine ähnliche Unabhängigkeit auch der Satz von HAJÓS aufzeigt (vgl. [3]).

Es ist bemerkenswert, daß man zum Beweis vom Satz 9 nicht des vollen Inhalts vom Satz 8 bedarf. Und zwar genügt es hierzu Satz 8 nur für elementare p-Gruppen zu beweisen, woraus für denselben Spezialfall auch die linke Seite von (28) folgt. Hieraus entsteht dann der allgemeine Fall auf Grund der elementaren Sätze 3, 6 durch eine leichte Induktion.

Die Sätze 1 bis 7 lassen sich mit einigen Änderungen auf den Fall von beliebigen endlichen Gruppen und Ringen verallgemeinern — was wir dem Leser überlassen können — mit den Sätzen 8 bis 12 geht das aber

nicht mehr. (Vgl. den Anfang der Bemerkung 5 im § 2.) Überhaupt scheinen uns die Zetafunktionen für ganz beliebige Strukturen von sehr kompliziertem Verhalten zu sein.

Dagegen bieten die Zetafunktionen für die Abelschen Torsionsgruppen gewiß einen sehr interessanten Fall. (Eine Gruppe nennen wir eine Torsionsgruppe, wenn ihre Elemente von endlicher Ordnung sind.) Bezeichnet G eine solche (unendliche) Gruppe, so ist es eine Grundaufgabe zu untersuchen, für welche $z (Rz \geqq 1)$ ein

$$\varrho(z) = \varrho(z; A_1, A_2, \ldots) \tag{44}$$

existiert, wobei $A_1, A_2, \ldots$ ein abzählbar unendliches System von (endlichen) Untergruppen von G ist.

Diesbezüglich zeigen wir hier nur, daß (44) jedenfalls für alle $z = 1, 2, \ldots$ existiert. Für jedes solche z ist nämlich die Folge

$$\varrho(z; A_1, \ldots, A_n) \qquad (n = 1, 2, \ldots)$$

nach den Sätzen 9, 11 beschränkt und monoton, weshalb sie konvergent ist. Wir haben noch zu zeigen, daß ihr Grenzwert von der Reihenfolge der $A_1, A_2, \ldots$ unabhängig ist. Zu diesem Zweck bezeichne $B_1, B_2, \ldots$ eine andere Anordnung von $A_1, A_2, \ldots$. Zu jedem n gibt es dann ein m, so daß alle $A_1, \ldots, A_n$ unter den $B_1, \ldots, B_m$ vorkommen. Nach Satz 11 gilt dann

$$\varrho(z; A_1, \ldots, A_n) \geqq \varrho(z; B_1, \ldots, B_m),$$

woraus

$$\lim_{n \to \infty} \varrho(z; A_1, \ldots, A_n) \geqq \lim_{n \to \infty} \varrho(z; B_1, \ldots, B_n)$$

folgt. Aus Symmetriegründen gilt dies auch mit „$\leqq$“, also auch mit „$=$“. Das beendet den Beweis, womit wir die Existenz von (44) für $z = 1, 2, \ldots$ ausgewiesen haben. Auch folgt aus diesem Beweis, daß Satz 9 nebst Zusatz für (44) seine Gültigkeit behält, d. h.

$$0 \leqq \varrho(z; A_1, A_2, \ldots) \leqq 1 - (\min((A_1), (A_2), \ldots)^{-z} (< 1) \qquad (z = 1, 2, \ldots)$$

ist.

Im Fall unendlicher Abelscher Gruppen haben wir (44) für die (reellen oder komplexen) Stellen $z \neq 1, 2, \ldots$ bisher gar nicht allgemein untersucht, weshalb wir über das etwaige analytische Verhalten der entsprechenden Zetafunktionen $\zeta(z) = \varrho(z; A_1, A_2, \ldots)^{-1}$ noch nicht orientiert sind. Die Untersuchung scheint für die (unendlichen) Abelschen Torsionsgruppen mit lauter endlichen p-Untergruppen aussichtsvoll zu sein. Allerdings gibt es unendliche Abelsche Gruppen, für die die Zetafunktionen sich trivial verhalten. So ist es z. B. mit der Prüferschen Gruppe $C(p^\infty)$ (p-Primzahl) definiert durch die Erzeugenden $\alpha_1, \alpha_2, \ldots$ und die zwischen ihnen bestehenden Gleichungen

$$\alpha_1^p = 1, \quad \alpha_{i+1}^p = \alpha_i, \quad \alpha_i \alpha_k = \alpha_k \alpha_i \qquad (i, k = 1, 2, \ldots).$$

Da jetzt die sämtlichen echten Untergruppen die Kette

$$\{\alpha_1\} \subset \{\alpha_2\} \subset \cdots$$

bilden, so folgt aus Satz 2 leicht, daß in diesem Fall stets

$$\varrho(z; A_1, A_2, \ldots) = \varrho(z; A_i) = 1-(A_i)^{-z}$$

gilt, wobei A_i das minimale unter den $A_1, A_2, \ldots$ bezeichnet. Dieses äußerst einfache Benehmen der Zetafunktionen für die Gruppen $C(p^\infty)$ steht im Einklang damit, daß diese Gruppen sonst auch in vieler Hinsicht eine Sonderstellung in der Gruppentheorie einnehmen. Soviel steht fest, daß die Zetafunktionen sehr empfindlich gegen die Kompliziertheit der zugrunde gelegten algebraischen Struktur reagiert.

(Eingegangen am 22. Dezember 1954.)

Literaturverzeichnis

[1] S. Delsarte, Fonctions de Möbius sur les groupes Abelian finis, *Annals of Math.*, **49** (1948), S. 600—609.

[2] G. Hajós, Über einfache und mehrfache Bedeckung des n-dimensionalen Raumes mit einem Würfelgitter, *Math. Zeitschr.*, **47** (1942), S. 427—467.

[3] L. Rédei, Die gruppentheoretischen Zetafunktionen und der Satz von Hajós, *Acta Math. Acad. Sci. Hung.*, **6** (1955) (im Erscheinen).

ζ-ФУНКЦИИ В АЛГЕБРЕ

Л. Редеи (Сегед)

(Резюме)

Пусть M — произвольное множество; обозначим через (M) $(= 0, 1, 2, \ldots, \infty)$ число элементов от M. Итак, если M является (алгебраической) структурой (группой, кольцом, полугруппой и т. п.), то (M) означает порядок от M.

Пусть n — неотрицательное целое число. Обозначим через $\mathfrak{N}$ множество чисел $1, \ldots, n$. (Если $n = 0$, то $\mathfrak{N} = \mathbf{0}$ является пустым множеством.)

Пусть дана теперь структура S, и посмотрим систему $A_1, \ldots, A_n$ ее подструктур. (Если S группа или кольцо и. т. п., то каждое из множеств A_i является также группой, или кольцом и. т. п.) При любом $\mathfrak{M}(\subseteq \mathfrak{N})$ обозначим через $A_\mathfrak{M}$ подструктуру из S порожденную элементами подструктур A_i $(i \in \mathfrak{M})$. Рассмотрим систему

(1) $$(A_\mathfrak{M}) \qquad (\mathbf{0} \subset \mathfrak{M} \subseteq \mathfrak{N}).$$

(Она не легко обозрима даже в том частном случае, когда S конечная абелева группа и $n \geqq 3$.)

Чтобы цолучить сведения о системе (1), определим с помощью комплексной переменной z для $Rz > 0$ функцию

(2) $$\varrho(z) = \varrho(z; A_1, \ldots, A_n) = \sum_{\mathfrak{M} \subseteq \mathfrak{N}} (-1)^{(\mathfrak{M})} (A_\mathfrak{M})^{-z},$$

где Rz означает вещественную часть z и в случае $\mathfrak{M}=\mathbf{0}$ следует считать $(A_{\mathfrak{M}})=1$. (Конечно, можно пренебрегать членами соответствующими случаями $(A_{\mathfrak{M}})=\infty$, и для положительных c следует считать $c^{-z}=e^{-z\log c}$, где $\log c$ означает вещественное значение логарифма.) Будем называть функции

$$\zeta(z)=\zeta(z;A_1,\ldots,A_n)=\varrho(z)^{-1}=\varrho(z;A_1,\ldots,A_n)^{-1} \tag{3}$$

ζ-функциями (в алгебре). Распространим наши определения и на случай бесконечных (счетных) систем подструктур $A_1, A_2, \ldots$ так, чтобы было

$$\varrho(z;A_1,A_2,\ldots)=\lim_{n\to\infty}\varrho(z;A_1,\ldots,A_n), \tag{4}$$

если только правая сторона существует и ее значение независимо от порядка членов в последовательности $A_1, A_2, \ldots$. С помощью аналитического продолжения можно распространять определение в некоторых случаях на всю плоскость. Можно охватывать и структуры с операторами, ограничиваясь подструктурами, допустимыми относительно операторов. Если S группа или кольцо, то можно определение „дуализировать", взяв вместо $A_{\mathfrak{M}}$ пересечение подструктур A_i $(i\subseteq\mathfrak{M})$, и вместо порядкового числа $(A_{\mathfrak{M}})$ индекс $(S:A_{\mathfrak{M}})$ („двойственная ζ-функция"). (Если однако S конечная абелева группа, то двойственные ζ-функции в совокупности совпадают с „обыкновенными" ζ-функциями.)

Если, в частности, A_i $(i=1,2,\ldots)$ есть группа порядка p_i, где p_i означает i-ое простое число, а S является прямым произведением групп $A_1, A_2, \ldots$, то очевидно

$$\zeta(z;A_1,A_2,\ldots)=\varrho(z;A_1,A_2,\ldots)^{-1}\prod_{i=1}^{\infty}(1-p_i^{-z})^{-1},$$

а это есть как раз ζ-функция Римана. Если S бесконечная циклическая группа, а A_i $(i=1,2,\ldots)$ ее подгруппы индексов p_i, то соответствующая двойственная ζ-функция есть опять ζ-функция Римана. Если же P кольцо целых чисел алгебраического числового тела конечной степени, а $A_1, A_2, \ldots$ все простые идеалы P (иными словами: все максимальные допустимые подмодули операторного модуля P^+, где операция означает умножение на элементы P), то соответствующая двойственная ζ-фуикция есть ζ-функция Дедекинда, принадлежащей P.

Автором был до сих пор более подробно рассмотрен только случай, когда $S=G$ есть конечная абелева группа. Кроме нескольких почти тривиальных соотношений (теоремы 1—7) имеет место „основная теорема ζ-функций конечных абелевых групп":

Имеет место соотношение

$$\varrho(z;A_1,\ldots,A_n)=\sum_H(H)^{-z}\psi(G/H), \tag{5}$$

где H пробегает подгруппы G не содержащие ни одну из подгрупп $A_1,\ldots,A_n$, G/H означает факторгруппу а функция ψ определена следующим образом: Если P конечная абелева p-группа с конечным числом t инвариантов, то

$$\psi(P)=(1-p^{-z})(1-p^{1-z})\cdots(1-p^{t-s-z}),$$

если же $P_1,\ldots,P_k$ все различные силовские подгруппы конечной абелевой группы, то

$$\psi(P_1\cdots P_k)=\psi(P_1)\cdots\psi(P_k).$$

Наиболее важным следствием является „теорема инерции конечных абелевых групп", соответственно которой всегда имеет место

$$0\leqq\varrho(z;A_1,\ldots,A_n)<1 \qquad (z=1,2,\ldots). \tag{6}$$

Эта теорема является оптимальной с трех точек зрения. Во первых, если $z(\geqq 1)$ вещественное нецелое число, то ни одно из неравенств

$$c<\varrho(z;A_1,\ldots,A_n)<C$$

не выполняется всегда, как бы ни задать вещественные числа c, C. Во вторых для любого целого положительного числа z на левой стороне (6) имеет место равенство во бесконечном числе (нетривиальных) случаев. Во третьях, если $A_1, \ldots, A_n$ подгруппы конечной некоммутативной группы, то левая сторона (6) вообще несправедлива, даже

$$c < \varrho(z; A_1, \ldots, A_n) \qquad (z = 1, 2, \ldots)$$

не выполняется ни при каком c.

Дальнейшим следствием из (5) является (в случае конечных абелевых групп) „свойство монотонности“

$$\varrho(z; A_1, \ldots, A_n) \leqq \varrho(z; A_1, \ldots, A_{n-1}) \qquad (z = 1, 2, \ldots).$$

Можно ожидать интересные результаты и от исследования (бесконечных) периодических абелевых групп. Для них $\varrho(z; A_1, A_2, \ldots)$ существует при $z = 1, 2, \ldots$ и в аналогии с (6) имеет место

$$0 \leqq \varrho(z; A_1, A_2, \ldots) < 1 \qquad (z = 1, 2, \ldots).$$

NEUER BEWEIS DES HAJÓSSCHEN SATZES ÜBER DIE ENDLICHEN ABELSCHEN GRUPPEN

Von

LADISLAUS RÉDEI (Szeged), Mitglied der Akademie

Herrn Prof. Dr. LADISLAUS KALMÁR *am 50. Geburtstage mit Freundschaft zugeeignet*

Durchgängige Bezeichnungen:

G eine endliche Abelsche Gruppe mit mindestens zwei Elementen.

ε Einselement von G.

$\alpha, \beta, \ldots$ (mit oder ohne Indizes) Elemente von G.

$a, b, \ldots$ (mit oder ohne Indizes) ganze Zahlen.

p, q, r (mit oder ohne Indizes) positive Primzahlen.

$\{\ldots\}$ die durch die eingeklammerten Elemente erzeugte Untergruppe von G.

$o(\ldots)$ die Ordnung eines Gruppenelementes.

$O(\ldots)$ die Ordnung einer Gruppe.

$O\{\ldots\} = O(\{\ldots\})$.

A (mit oder ohne Indizes) ein Komplex (d. h. eine Untermenge) von G.
Unter dem Produkt $A_1 \ldots A_k$ der Komplexe $A_1, \ldots, A_k$ verstehen wir wie üblich den Komplex aller verschiedenen $\alpha_1 \ldots \alpha_k$ $(\alpha_i \in A_i;\ i = 1, \ldots, k)$. Erscheint in dieser Form jedes Element von $A_1 \ldots A_k$ nur einmal, so nennen wir dieses Komplexprodukt *direkt*. Bei uns werden ausschließlich direkte Komplexprodukte vorkommen, auch wenn wir es nicht ausdrücklich betonen.

$(\alpha)_e$ der Komplex mit den Elementen $\varepsilon, \alpha, \ldots, \alpha^{e-1}$ $(o(\alpha) \geqq e \geqq 2)$. Wir nennen $(\alpha)_e$ ein *Simplex* (von G). Dies ist dann und nur dann eine (zyklische) Gruppe, wenn $o(\alpha) = e$ ist. Insbesondere nennen wir $(\alpha)_p$ ein *Primsimplex*. Gilt $G = (\alpha_1)_{e_1} \cdots (\alpha_n)_{e_n}$ und ist die rechte Seite ein direktes Produkt, so nennen wir sie eine *Hajóssche Zerlegung* (von G). Diese nennen wir im Fall $e_i = p_i$ $(i = 1, \ldots, n)$ *prim*.

§ 1. Einleitung

Der berühmte, vor 14 Jahren entdeckte Satz von HAJÓS [2][1] (s. auch HAJÓS [1]) lautet so:

SATZ 1. *In jeder Hajósschen Zerlegung von G ist mindestens ein Faktor eine Gruppe.*

Diesem Satz verleihen neben seiner Wichtigkeit auch die unerwarteten Schwierigkeiten des Beweises einen besonderen Reiz. Wohl ist nämlich der Satz dem Wortlaut nach eine der einfachsten in der Gruppentheorie, auch bieten sich zum Beweis viele trivial erscheinende Wege an, zur peinlichen Überraschung zeigen sich aber die meisten von ihnen sehr bald als Irrwege, die nur da sind, um einem zu erschweren, einen richtigen Weg zum Beweis zu finden.

HAJÓS überwand die Schwierigkeiten mit einer genialen Idee, die nach Heranziehen des Ringes $\mathfrak{Z}(G)$ von G über dem Ring $\mathfrak{Z}$ der ganzen Zahlen hauptsächlich in einer feinen Nullteileruntersuchung bestand. Sein glänzender Beweis ist bis heute der einzige geblieben, doch wurde dieser durch RÉDEI [5] und noch weiter durch SZELE [9] bedeutend vereinfacht. (S. auch das Buch von RÉDEI [7], wo dieser vereinfachte Beweis möglichst kurz ausgearbeitet ist.) Wir bemerken, daß wegen der Abstraktheit des von tiefen Gedanken erfüllten Beweises von HAJÓS sein Satz auch noch heute sehr geheimnisvoll ist.

Ich habe früher (vgl. RÉDEI [6]) den Satz rein gruppentheoretisch, prinzipiell sehr einfach aber in der Ausführung ziemlich kompliziert auf den Spezialfall der p-Gruppen reduziert, konnte jedoch damals auf diesem Wege nicht zu einem vollständigen Beweis kommen. Es schien deshalb, daß die eigentlichen Schwierigkeiten des Satzes in diesem Spezialfall konzentriert sind, welche Meinung auch im Beweis von Hajós ihre Unterstützung findet, da dieser sich für p-Gruppen nicht kürzer gestaltet.

Nach vielem Mißerfolg, aber von den Schönheiten des Satzes immer neu angeregt fand ich endlich für den Fall von p-Gruppen einen überraschend einfachen Beweis. Vollständigkeitshalber führe ich den Beweis auch für nicht-p-Gruppen aus; dieser Teil des Beweises ist im wesentlichen eine Reproduktion meiner vorigen Arbeit [6], doch mit einigen Vereinfachungen.

Am Ende meines Beweises für p-Gruppen benütze ich einen kurzen ringtheoretischen Schluß im auch von HAJÓS verwendeten Gruppenring $\mathfrak{Z}(G)$, sonst ist mein Beweis gruppentheoretisch. Als ich das Herrn SZELE erzählte, hat er bemerkt, daß sich dieser ringtheoretische Schluß wahrscheinlich durch einen gruppentheoretischen ersetzen läßt. Das ist wenn auch etwas kompliziert in der Tat möglich, das ich im § 5 zeigen werde. Es ist die prinzipielle Wichtigkeit zu betonen, daß somit für den Satz von HAJÓS durch die

[1] [ ] bezieht sich auf das Literaturverzeichnis am Ende unserer Arbeit.

wertvolle Mitwirkung von Herrn SZELE auch ein rein gruppentheoretischer Beweis entstanden ist.

Der Satz von HAJÓS brachte zurzeit auch die Überraschung mit, daß ihn HAJÓS unabhängig vom Fundamentalsatze der endlichen Abelschen Gruppen bewiesen hat und dieser nicht geeignet ist seinen Beweis zu verkürzen. Dasselbe bezieht sich auch auf meinen Beweis, weshalb beide Sätze mit großer Wahrscheinlichkeit voneinander völlig unabhängig sind. Das ist sehr merkwürdig, da es natürlich möglich ist, mit Hilfe des Fundamentalsatzes den Satz von HAJÓS in Angriff zu nehmen; desto enttäuschender ist seine Machtlosigkeit zum Beweis.

Bezüglich des Verhältnisses beider Sätze zueinander soll auch bemerkt werden, daß sich der Satz von HAJÓS als eine Umkehrung des Fundamentalsatzes ansehen läßt, da man diesen folgenderweise aussprechen kann: Jede endliche Abelsche Gruppe hat eine Hajóssche Zerlegung, in der alle Faktoren Gruppen sind.

Nach diesen Bemerkungen liegt vor uns eine Synthese von zwei gleich wichtigen, einander ergänzenden und (mit großer Wahrscheinlichkeit) voneinander unabhängigen Sätzen, nämlich des Fundamentalsatzes und des Satzes von HAJÓS.

Es ist auch eine merkwürdige Gegenüberstellung beider Sätze, daß der Fundamentalsatz sich von p-Gruppen auf nicht-p-Gruppen trivial verallgemeinern läßt, dagegen in meinem Beweis eben der Fall von p-Gruppen sehr einfach und die Verallgemeinerung auf nicht-p-Gruppen kompliziert ist. Vielleicht läßt auch diese Unverhältnismäßigkeit hoffen, daß der zweite Teil meines Beweises vereinfachbar ist.

Ein wesentliches Moment ist es, daß wir den Hajósschen Satz in seiner von uns bald nach seinem Entstehen entdeckten „scharfen" Form beweisen werden. (Vgl. RÉDEI [3]. Es ist ein eigenartiger Zug vom ganzen Problemkreis des Minkowski—Hajósschen Satzes, daß auch die beiden, dem Satz von HAJÓS äquivalenten Minkowskischen Vermutungen je eine „scharfe" Formulierung zulassen. Vgl. RÉDEI [4], wo diese drei Sätze in „schwacher" und „scharfer" Form angeführt sind.)

Um die scharfe Form des Satzes von HAJÓS aussprechen zu können, betrachten wir *Systeme*

$$\begin{pmatrix} a_1 & \dots & a_n \\ e_1 & \dots & e_n \end{pmatrix} \qquad (e_1, \dots, e_n \geqq 2;\ n \geqq 1). \tag{1}$$

Wir nennen (1) ein *geordnetes Normalsystem* (für $\{a_1, \dots, a_n\}$), wenn

$$O\{a_1, \dots, a_k\} = e_1 \dots e_k \qquad (k = 1, \dots, n)$$

ist. Mit anderen Worten bedeutet das, daß zur Untergruppenkette

$$\varepsilon \subset \{a_1\} \subset \{a_1, a_2\} \subset \dots \subset \{a_1, \dots, a_n\}$$

die Indizesfolge $e_1, \dots, e_n$ gehört. (Für Primzahlen $e_1, \dots, e_n$ handelt es sich

um Kompositionsreihen.) Unter einem *Normalsystem* schlechthin verstehen wir ein System (1), das durch eine Vertauschung der „Spalten“ α_i, e_i $(i=1,\ldots,n)$ in ein geordnetes Normalsystem überführbar ist. (Statt „Normalsystem“ sagen wir auch „normales System“.) Nunmehr verstehen wir unter dem *scharfen Hajósschen Satz* den folgenden:

SATZ 2. *Dann und nur dann ist*

$$G=(\alpha_1)_{e_1}\cdots(\alpha_n)_{e_n} \tag{2}$$

eine Hajóssche Zerlegung, wenn (1) *ein Normalsystem für* G *ist.*

Dieser Satz ist in der Tat eine Verschärfung vom Satz 1, denn in einem Normalsystem (1) muß nach seiner Definition für mindestens ein i die Gleichung $o(\alpha_i)=O\{\alpha_i\}=e_i$ erfüllt, d. h. der Faktor $(\alpha_i)_{e_i}$ eine Gruppe sein.

Gleich hier beweisen wir den Teil „dann“ vom Satz 2. Im Fall eines Normalsystems für G sind die Produkte

$$\alpha_1^{k_1}\cdots\alpha_n^{k_n} \qquad (k_i=0,\ldots,e_i-1\,;\ i=1,\ldots,n)$$

offenbar die sämtlichen verschiedenen Elemente von G, was eben besagt, daß (2) eine Hajóssche Zerlegung von G ist.

Da hiernach der Teil „dann“ vom Satz 2 eine Trivialität ist, somit das Gewicht von ihm im Teil „nur dann“ liegt, den wir allein noch zu beweisen haben, so werden wir im folgenden „Satz 2“ oft mit seinem Teil „nur dann“ identifizieren.

Übrigens werden wir sehen, daß Satz 2 eine fast triviale Folgerung vom Satz 1 ist (vgl. Hilfssatz 1 nebst Korollar im § 2).

Natürlich ist Satz 2 insofern Satz 1 überlegen, als er die Hajósschen Zerlegungen vollständig charakterisiert, während Satz 1 nur eine notwendige Bedingung für sie ausspricht. Diese Überlegenheit macht Satz 2 für einen induktiven Beweis besonders geeignet.

Wir bemerken noch, daß die leichte aber wichtige Reduktion von HAJÓS auf den Fall von Primsimplexen auch in unserem Beweis unentbehrlich wird. Sonst haben beide Beweise keine Berührungspunkte.[2]

§ 2. Reduktion des Beweises

HILFSSATZ 1. *Gilt Satz 1 für eine gegebene Hajóssche Zerlegung* (2) *von* G *und gilt Satz 2 für die Gruppen von kleinerer Ordnung als* G, *so gilt Satz 2 für die gegebene Zerlegung* (2).

KOROLLAR. *Hieraus folgt die Äquivalenz der Sätze 1, 2.*

[2] Bezüglich anderer neuerer Forschungen im Problemkreis des Satzes von HAJÓS s. man RÉDEI [8].

Zum Beweis legen wir uns eine Hajóssche Zerlegung (2) von G vor, in der z. B. $(\alpha_1)_{e_1}$ eine Gruppe ist. Dann ist diese gleich $\{\alpha_1\}$, ferner gilt $O\{\alpha_1\} = o(\alpha_1) = e_1$. Aus (2) folgt $G = \{\alpha_1, \ldots, \alpha_n\}$. Nun ist im Fall $n = 1$ die Behauptung klar. Im Fall $n \geqq 2$ gewinnen wir aus dem Gesagten, daß

$$(\alpha_2\{\alpha_1\})_{e_2} \cdots (\alpha_n\{\alpha_1\})_{e_n}$$

eine Hajóssche Zerlegung der Faktorgruppe $G/\{\alpha_1\}$ ist. Wegen der Voraussetzung ist also

$$\begin{pmatrix} \alpha_2\{\alpha_1\} & \cdots & \alpha_n\{\alpha_1\} \\ e_2 & \cdots & e_n \end{pmatrix}$$

ein Normalsystem (für $G/\{\alpha_1\}$). Hieraus und aus $O\{\alpha_1\} = e_1$ folgt, daß (1) ein Normalsystem für G ist, was wir zu beweisen hatten.

HILFSSATZ 2. (HAJÓS) *Ist Satz 1 für die primen Hajósschen Zerlegungen von G richtig, so ist er für alle Hajósschen Zerlegungen von G richtig.*

Denn legen wir uns eine Hajóssche Zerlegung (2) vor. Diese läßt sich mit wiederholter Anvendung der Formel

$$(\alpha)_{ef} = (\alpha)_e(\alpha^e)_f \qquad (o(\alpha) \geqq ef;\ e, f \geqq 2)$$

in eine prime Hajóssche Zerlegung verwandeln. Zum Beweis vom Hilfssatz 2 genügt es also folgendes zu zeigen. Gilt eine direkte Produktzerlegung

$$(\varrho)_e = AB \qquad (A \neq \varepsilon),$$

wo A eine Gruppe ist, so ist auch $(\varrho)_e$ eine Gruppe. Um dies zu beweisen nehmen wir ein Element $\alpha(\neq \varepsilon)$ von A. Wegen $\alpha A = A$ besteht

$$\alpha(\varrho)_e = (\varrho)_e.$$

Hieraus folgt $\alpha = \varrho^k$ $(0 < k < e)$, $\varrho^e = \alpha\varrho^{e-k} = \varrho^l$ $(0 \leqq l < e)$, $\varrho^{e-l} = \varepsilon$, also wegen $o(\varrho) \geqq e$ auch $o(\varrho) = e$, womit der Beweis beendet ist.

Aus den Hilfssätzen 1, 2 kommen wir zur folgenden:

REDUKTION DES BEWEISES. *Gilt Satz 1 für die primen Hajósschen Zerlegungen von G und gilt Satz 2 für die Gruppen von kleinerer Ordnung als G, so gilt letzterer auch für G.*

Aus Hilfssatz 2 folgt nämlich, daß Satz 1 für G gilt. Weiter folgt hieraus nach Hilfssatz 1 in der Tat die Richtigkeit von Satz 2 für G.

§ 3. Beweis für p-Gruppen

LEMMA. *Ist für $\alpha_1, \ldots, \alpha_n$ $(n \geqq 1)$*

$$O\{\alpha_1, \ldots, \alpha_n\} = p^n \tag{3}$$

und gelten alle Ungleichungen

$$O\{\alpha_{i_1}, \ldots, \alpha_{i_k}\} \geqq p^k \qquad (1 \leqq i_1 < \cdots < i_k \leqq n;\ 1 \leqq k \leqq n-1), \tag{4}$$

so gibt es Potenzen $s_1, \ldots, s_n$ *von* p, *so daß*

$$\begin{pmatrix} \alpha_1^{s_1} & \cdots & \alpha_n^{s_n} \\ p & \cdots & p \end{pmatrix} \tag{5}$$

ein Normalsystem (für $\{\alpha_1, \ldots, \alpha_n\}$*) ist.*

Im Fall $n = 1$ ist die Behauptung trivial. Im Fall $n \geqq 2$ machen wir die Induktionsannahme, daß die Behauptung für die Systeme $\alpha_1, \ldots, \alpha_n$ mit kleinerem $o(\alpha_1) \cdots o(\alpha_n)$ richtig ist. *Erstens* betrachten wir den Fall, daß in (4) mindestens einmal „=" gilt. Dann dürfen wir mit passender Reihenfolge der $\alpha_1, \ldots, \alpha_n$ annehmen, daß für ein m $(1 \leqq m \leqq n-1)$

$$O\{\alpha_1, \ldots, \alpha_m\} = p^m \tag{6}$$

ist. Da ferner (4) für m statt n noch mehr gilt, so gibt es nach der Induktionsannahme Potenzen $s_1, \ldots, s_m$ von p, so daß

$$\begin{pmatrix} \alpha_1^{s_1} & \cdots & \alpha_m^{s_m} \\ p & \cdots & p \end{pmatrix} \tag{7}$$

ein Normalsystem ist. Hieraus und aus (6) folgt auch

$$\{\alpha_1, \ldots, \alpha_m\} = \{\alpha_1^{s_1}, \ldots, \alpha_m^{s_m}\}. \tag{8}$$

Wird andererseits

$$\hat{\alpha} = \alpha\{\alpha_1, \ldots, \alpha_m\} \tag{9}$$

gesetzt, aufgefaßt als Element der Faktorgruppe $G/\{\alpha_1, \ldots, \alpha_m\}$, so folgen aus (3), (4), (6) offenbar

$$O\{\hat{\alpha}_{m+1}, \ldots, \hat{\alpha}_n\} = p^{n-m},$$
$$O\{\hat{\alpha}_{i_1}, \ldots, \hat{\alpha}_{i_k}\} = p^k \qquad (m+1 \leqq i_1 < \cdots < i_k \leqq n;\ 1 \leqq k \leqq n-m-1).$$

Da ferner stets $o(\hat{\alpha}) \mid o(\alpha)$ ist, so folgt hieraus wieder nach der Induktionsannahme die Existenz von Potenzen $s_{m+1}, \ldots, s_n$ von p, so daß

$$\begin{pmatrix} \hat{\alpha}_{m+1}^{s_{m+1}} & \cdots & \hat{\alpha}_n^{s_n} \\ p & \cdots & p \end{pmatrix} \tag{10}$$

ein Normalsystem ist. Wegen (8), (9) setzen sich die Normalsysteme (7), (10) zum Normalsystem (5) zusammen, womit die Behauptung für diesen Fall bewiesen ist. *Zweitens* haben wir nur noch den Fall zu betrachten, daß (statt (4)) sogar

$$O\{\alpha_{i_1}, \ldots, \alpha_{i_k}\} \geqq p^{k+1} \qquad (1 \leqq i_1 < \cdots < i_k \leqq n;\ 1 \leqq k \leqq n-1) \tag{11}$$

gilt. Insbesondere ist dann $O\{\alpha_2, \ldots, \alpha_n\} \geqq p^n$, also wegen (3)

$$O\{\alpha_1^p, \alpha_2, \ldots, \alpha_n\} = p^n.$$

Hiernach und nach (11) behalten (3), (4) ihre Gültigkeit bei Ersetzung von α_1 durch α_1^p. Wegen (4) gilt dabei $o(\alpha_1) \geqq p$, also $o(\alpha_1^p) < o(\alpha_1)$, somit folgt aus der Induktionsannahme mit Anwendung auf das System $\alpha_1^p, \alpha_2, \ldots, \alpha_n$ die Richtigkeit der Behauptung auch für diesen Fall (jetzt sogar mit $s_1 \geqq p$). Das beweist unser Lemma.

Nunmehr können wir Satz 2 für p-Gruppen beweisen. Nach der Reduktion im § 2 genügt es zu zeigen, daß in jeder Hajósschen Zerlegung

$$G = (\alpha_1)_p \dots (\alpha_n)_p \tag{12}$$

mindestens ein Faktor eine Gruppe ist.

Wegen der Annahme gilt (3), ferner ist jedes Teilprodukt von (12) ebenfalls direkt, weshalb auch (4) erfüllt ist. Nach dem Lemma gibt es also Potenzen $s_1, \dots, s_n$ von p, so daß (5) ein Normalsystem ist. Gilt dabei $s_1 = \dots = s_n = 1$, so sind wir mit dem Beweis fertig. Im anderen Falle dürfen wir annehmen, daß mit einem m

$$p | s_1, \dots, s_m;\ s_{m+1} = \dots = s_n = 1 \qquad (1 \leqq m \leqq n) \tag{13}$$

gilt. Da (5) ein Normalsystem ist, so folgt unter Berücksichtigung von (12) aus dem Teil „dann" vom Satz 2, daß auch

$$G = (\alpha_1^{s_1})_p \dots (\alpha_m^{s_m})_p (\alpha_{m+1})_p \dots (\alpha_n)_p \tag{14}$$

eine Hajóssche Zerlegung einer Gruppe ist. Wir werden in einen Widerspruch geraten, womit der Beweis beendet wird.

Einen Augenblick nehmen wir den im § 1 erwähnten Gruppenring $\mathfrak{J}(G)$ zu Hilfe. Gleichzeitig ändern wir unsere Bezeichnungen so ab, daß wir mit $(\alpha)_e$ das Element $\varepsilon + \alpha + \dots + \alpha^{e-1}$ von $\mathfrak{J}(G)$ bezeichnen, nach dem die rechte Seite von (12) ebenfalls ein Element von $\mathfrak{J}(G)$, und zwar die Summe der Elemente der Gruppe G bedeutet. Da diese α_1 enthält, so ist das $(\varepsilon - \alpha_1)$-fache der rechten Seite von (12) gleich 0. Wegen (13) und der Formel

$$(\varepsilon - \alpha)(\alpha)_e (\alpha^e)_f = \varepsilon - \alpha^{ef}$$

gilt also noch mehr

$$(\varepsilon - \alpha_1^{s_1}) \dots (\varepsilon - \alpha_m^{s_m}) (\alpha_{m+1})_p \dots (\alpha_n)_p = 0.$$

Nach Ausmultiplizieren tritt links das Glied ε auf, welches also gegen ein anderes Glied aufgehen muß. Das steht aber in einem Widerspruch damit, daß das Simplexprodukt (14) direkt ist, womit wir Satz 2 für p-Gruppen bewiesen haben.

§ 4. Beweis für nicht p-Gruppen

HILFSSATZ 3. *Es gilt*

$$(\alpha_1)_{p_1} \dots (\alpha_n)_{p_n} = (\alpha_1^{t_1})_{p_1} \dots (\alpha_n^{t_n})_{p_n} \qquad (p_i \nmid t_i;\ i = 1, \dots, n), \tag{15}$$

vorausgesetzt, daß die linke Seite eine Hajóssche Zerlegung einer Gruppe ist, ferner gilt dann

$$p_i | o(\alpha_i) \qquad (i = 1, \dots, n).$$

Es genügt nämlich folgendes zu zeigen: Es gilt

$$(\alpha)_p A = (\alpha^t)_p A \qquad (p \nmid t), \tag{16}$$

wenn die linke Seite die Zerlegung einer Gruppe in ein direktes Produkt

und $\varepsilon \in A$ ist, ferner gilt dann $p|o(\alpha)$. Um dies zu beweisen berücksichtigen wir, daß $\alpha \in (\alpha)_p A$, also $\alpha(\alpha)_p A = (\alpha)_p A$ ist. Wegen $\alpha(\alpha)_p = \alpha, \ldots, \alpha^p$ folgt hieraus $\alpha^p A = A$, also allgemeiner

$$\alpha^{p^i} A = A \qquad (i = 0, \pm 1, \ldots).$$

Hiernach muß $p|o(\alpha)$ sein, denn im anderen Fall würde $\alpha A = A$ folgen, was aber wegen der Annahme unmöglich ist. Da ferner $\alpha^k A$ nur von der Restklasse k (mod p) abhängt, so ist auch (16) richtig. Das beendet den Beweis vom Hilfssatz 3.

HILFSSATZ 4. *Im Falle eines geordneten Normalsystems*

$$\begin{pmatrix} \alpha_1 & \ldots & \alpha_m \\ p & \ldots & p \end{pmatrix} \qquad (p^2 | o(\alpha_2), \ldots, o(\alpha_m); m \geqq 1)$$

(für $\{\alpha_1, \ldots, \alpha_n\}$*) läßt sich jedes Element* $\alpha (\neq \varepsilon)$ *von* $\{\alpha_1, \ldots, \alpha_m\}$ *in der Form*

$$\alpha = \alpha_1^{k_1} \ldots \alpha_m^{k_m} \qquad (p \nmid k_1; p \nmid k_i \text{ oder } k_i = 0; i = 2, \ldots, m)$$

schreiben.

Wir machen die Induktionsannahme, daß die Behauptung für die kleineren Werte von m richtig ist. Aus der Voraussetzung folgt das Bestehen einer Gleichung

$$\alpha = \alpha_s^{l_s} \ldots \alpha_m^{l_m} \qquad (1 \leqq l_s \leqq p-1; 0 \leqq l_i \leqq p-1; i = s+1, \ldots, m).$$

Im Fall $s = 1$ sind wir mit dem Beweis fertig. Im Fall $s \geqq 2$ berücksichtigen wir, daß α_s^p wegen der Voraussetzung ein von ε verschiedenes Element von $\{\alpha_1, \ldots, \alpha_{s-1}\}$ ist. Da ferner auch

$$\begin{pmatrix} \alpha_1 & \ldots & \alpha_{s-1} \\ p & \ldots & p \end{pmatrix}$$

ein geordnetes Normalsystem ist, so folgt aus der Induktionsannahme das Bestehen einer Gleichung

$$\alpha_s^p = \alpha_1^{k_1} \ldots \alpha_{s-1}^{k_{s-1}} \qquad (p \nmid k_1; p \nmid k_i \text{ oder } k_i = 0; i = 2, \ldots, s-1).$$

Aus beiden Gleichungen folgt

$$\alpha = \alpha_1^{k_1} \ldots \alpha_{s-1}^{k_{s-1}} \alpha_s^{-p+l_s} \alpha_{s+1}^{l_{s+1}} \ldots \alpha_m^{l_m},$$

womit Hilfssatz 4 bewiesen ist.

HILFSSATZ 5. *Ist*

$$G = (\alpha_1)_{p_1} \ldots (\alpha_n)_{p_n} \tag{17}$$

eine Hajóssche Zerlegung und sind

$$\begin{pmatrix} \alpha_i & \alpha_{i_1} \ldots \alpha_{i_g} \\ p_i & p_i \ldots p_i \end{pmatrix}, \begin{pmatrix} \alpha_j & \alpha_{j_1} \ldots \alpha_{j_k} \\ p_j & p_j \ldots p_j \end{pmatrix} \qquad (i \neq j; 1 \leqq i, i_1, \ldots, j, j_1, \ldots, \leqq n) \tag{18}$$

zwei geordnete Normalsysteme, wobei die in (18) *vorkommenden* $\alpha_{i_x}, \alpha_{j_y}$ *von Primzahlpotenzordnung sind, so sind* $\alpha_i, \alpha_{i_1}, \ldots, \alpha_j, \alpha_{j_1}, \ldots$ *verschieden.*

Jedenfalls sind nämlich $\alpha_i, \alpha_{i_1}, \ldots$, desgleichen auch $\alpha_j, \alpha_{j_1}, \ldots$ untereinander verschieden. Ferner ist wegen der Voraussetzung kein Glied dieser Folgen dem ersten Glied der anderen Folge gleich. Wenn also die Behauptung falsch ist, so dürfen wir (nötigenfalls nach Streichung einiger hinterer Spalten der Systeme (18)) annehmen, daß die gesagten zwei Folgen mindestens zweigliedrig und ihre letzten Glieder $\alpha_{i_g} = \alpha_{j_h}$ gleich sind, sie aber sonst keine weiteren gleichen Glieder enthalten. Wegen $\alpha_{i_g} = \alpha_{j_h}$ müssen p_i, p_j gleich sein. Wir setzen

$$p = p_i = p_j, \quad \alpha = \alpha_{i_g}^{p} = \alpha_{j_h}^{p} \ (\neq \varepsilon).$$

Da die Systeme (18) normal sind, so gilt

$$\alpha \in \{\alpha_i, \alpha_{i_1}, \ldots, \alpha_{i_{g-1}}\}, \ \{\alpha_j, \alpha_{j_1}, \ldots, \alpha_{j_{h-1}}\}.$$

Es ist auch klar, daß die $o(\alpha_i), o(\alpha_{i_x}), o(\alpha_j), o(\alpha_{j_y})$ Potenzen von p sind. Hiernach können wir Hilfssatz 4 auf das Element α und beide geordneten Normalsysteme anwenden, die aus (18) nach Streichung der letzten Spalte entstehen. So kommen wir zu einer Gleichung von der Form

$$(\alpha =) \alpha_i^a \ldots \alpha_{i_{g-1}}^b = \alpha_j^c \ldots \alpha_{j_{h-1}}^d \ (\neq \varepsilon), \tag{19}$$

wobei jeder Exponent entweder durch p nicht teilbar oder gleich 0 ist. Wählt man also die t_i im Hilfssatz 3 passend, so treten beide Seiten von (19) beim Ausmultiplizieren der rechten Seite von (15) auf. Da aber diese (mit (17) zusammen) ein direktes Produkt ist und $\alpha_i, \ldots, \alpha_{i_{g-1}}, \alpha_j, \ldots, \alpha_{j_{h-1}}$ verschieden sind, so ist (19) unmöglich. Dieser Widerspruch beweist Hilfssatz 5.

Nunmehr können wir Satz 2 auch im Fall beweisen, daß G keine p-Gruppe ist. Zu diesem Zweck machen wir die Induktionsannahme, daß Satz 2 für die Gruppen von kleinerer Ordnung als G richtig ist, ferner nehmen wir an, daß Satz 1 für eine prime Hajóssche Zerlegung (17) von G falsch ist. Nach der Reduktion am Ende des § 2 genügt es hieraus einen Widerspruch abzuleiten.

Aus Hilfssatz 3 folgt sofort, daß (mit veränderten Bezeichnungen) auch eine (prime) Hajóssche Zerlegung

$$G = (\alpha_1 \beta_1)_{p_1} \ldots (\alpha_m \beta_m)_{p_m} (\gamma_1)_{r_1} \ldots (\gamma_n)_{r_n} \qquad (m, n \geqq 0; m + n > 0), \tag{20}$$

gilt, wobei

$$o(\alpha_i) = p_i \neq o(\beta_i) = q_i \qquad (i = 1, \ldots, m), \tag{21}$$

$$o(\gamma_i) \text{ eine durch } r_i^2 \text{ teilbare Potenz von } r_i \text{ ist } (i = 1, \ldots, n). \tag{22}$$

Mit (20) zusammen sind nach Hilfssatz 3 auch

$$G = (\alpha_1^{a_1} \beta_1^{b_1})_{p_1} \ldots (\alpha_m^{a_m} \beta_m^{b_m})_{p_m} (\gamma_1^{c_1})_{r_1} \ldots (\gamma_n^{c_n})_{r_n} \tag{23}$$

$$(p_i \nmid a_i, r_k \nmid c_k; i = 1, \ldots, m; k = 1, \ldots, n)$$

lauter Hajóssche Zerlegungen.

Wenn in (20) $m=0$ wäre, so liefert ein geeignetes Teilprodukt von (20) eine Hajóssche Zerlegung einer Sylowgruppe von G. Das ist aber wegen (22) unmöglich, da doch Satz 2 und noch mehr Satz 1 für p-Gruppen richtig ist. Folglich muß $m \geqq 1$ gelten.

Wir zeigen, daß es zu jedem $i(=1,\ldots,m)$ mindestens ein geordnetes Normalsystem

$$\begin{pmatrix} \alpha_i & \gamma_s & \gamma_t & \ldots & \gamma_w \\ p_i & p_i & p_i & \ldots & p_i \end{pmatrix} \qquad (1 \leqq s, t, \ldots, x \leqq n) \tag{24}$$

und ein von i verschiedenes $i'(=1,\ldots,m)$ mit

$$\beta_{i'} \in \{\alpha_i, \gamma_s, \gamma_t, \ldots, \gamma_w\} \qquad (i' \neq i) \tag{25}$$

gibt.

Als einfacher Spezialfall von (23) gilt nämlich (20) auch nach Ersetzung des i-ten Faktors durch $(\alpha_i)_{p_i}$. Da dieser nach (21) eine Gruppe ist, so ist für die entstandene Hajóssche Zerlegung von G Satz 1, also wegen der Induktionsannahme und Hilfssatz 1 auch Satz 2 richtig. Hiernach ist das dieser Zerlegung entsprechende System (1) normal. Letzteres geht nach passender Spaltenvertauschung in ein geordnetes Normalsystem über, das offenbar von der Form

$$\begin{pmatrix} \alpha_i & \gamma_s & \gamma_t & \ldots & \gamma_w \alpha_{i'} & \beta_{i'} & \ldots \\ p_i & p_i & p_i & \ldots & p_i & p_{i'} & \ldots \end{pmatrix} \qquad (i' \neq i) \tag{26}$$

sein muß (die $\gamma_s, \gamma_t, \ldots, \gamma_w$ können auch fehlen), wobei sich mit Notwendigkeit auch $m \geqq 2$ ergab. Hieraus folgt, daß auch (24) ein geordnetes Normalsystem und

$$(\alpha_{i'}\beta_{i'})^{p_{i'}} \in \{\alpha_i, \gamma_s, \gamma_t, \ldots, \gamma_x\}$$

ist. Hier fällt $\alpha_{i'}$ wegen (21) heraus. Ferner ist nach (24) der Exponent $p_{i'}$ prim zu $o(\beta_{i'})$, weshalb in der Tat auch (25) richtig ist.

Nun denken wir uns zu jedem $i(=1,\ldots,m)$ ein geordnetes Normalsystem (24) und ein $i'(=1,\ldots,m)$ mit der Eigenschaft (25) festgewählt. Insbesondere ist dann

$$i \to i' \qquad (i=1,\ldots,m)$$

eine Abbildung der Menge der Zahlen $1,\ldots,m$ in sich ohne Fixelemente. Folglich gibt es eine Untermenge dieser Menge mit mindestens zwei Elementen, die dabei permutiert wird. Diese Permutation läßt sich nach passender Umordnung der Faktoren von (20) durch

$$i \to i' \qquad (i=1,\ldots,m_0;\ 2 \leqq m_0 \leqq m) \tag{27}$$

angeben.

Für jedes $i(=1,\ldots,m_0)$ bezeichnen wir mit M_i die Menge der Elemente in der ersten Zeile von (24). Da nach (23) insbesondere

$$G = (\alpha_1)_{p_1} \ldots (\alpha_m)_{p_m} (\gamma_1)_{r_1} \ldots (\gamma_n)_{r_n}$$

gilt, so folgt aus Hilfssatz 5, daß die $M_i (i = 1, \ldots, m_0)$ paarweise elementfremd sind.

Wird andererseits Hilfssatz 4 auf das geordnete Normalsystem (24) und das Element $\beta_{i'}$ in (25) angewendet, so entsteht eine Gleichung von der Form

$$\beta_{i'} = \alpha_i^{a_i} \mu_1^{c_1} \ldots \mu_k^{c_k} \qquad (i = 1, \ldots, m_0),$$

wobei $\mu_1, \ldots, \mu_k (\neq \alpha_i)$ verschiedene Elemente von M_i sind und jeder Exponent prim zur Ordnung der Basis ist. Nach Multiplizieren dieser Gleichungen entsteht wegen (27) und des vorher Gesagten eine Gleichung

$$(\alpha_1^{a_1} \beta_1^{-1}) \ldots (\alpha_{m_0}^{a_{m_0}} \beta_{m_0}^{-1}) \gamma_1^{d_1} \ldots \gamma_n^{d_n} = \varepsilon,$$

wobei jedes d_k gleich 0 oder zu $o(\gamma_k)$ prim ist. Diese Gleichung steht aber im Widerspruch damit, daß (23) ein direktes Produkt ist. Das beendet den Beweis des Satzes von HAJÓS in seiner im Satz 2 gefaßten scharfen Form.

§ 5. Schlußbemerkungen

Wir wollen kurz skizzieren, wie sich der einzige ringtheoretische Schluß im obigen Beweise des Satzes von HAJÓS der Bemerkung von Herrn SZELE folgend gruppentheoretisch ausführen läßt, wodurch der ganze Beweis einen rein gruppentheoretischen Charakter erhält. Hierzu haben wir folgende Behauptung auf gruppentheoretischem Wege zu beweisen: Die Annahme, daß (12), (14) Hajóssche Zerlegungen von G sind und (13) gilt, ist widersprüchig.

Nach (12) gilt nämlich allgemeiner

$$G = \prod_{i=1}^{m} \alpha_i^{pa_i + x_i} (\alpha_i)_p \prod_{k=m+1}^{n} (\alpha_k)_p \tag{28}$$

$$\left(a_i = 0, \ldots, \frac{s_i}{p} - 1;\ x_i = 0, 1;\ i = 1, \ldots, m \right).$$

(Das ist sogar für beliebige ganze Zahlen a_i, x_i richtig.) Das Produkt auf der rechten Seite bezeichnen wir mit $P(x) = P(x_1, \ldots, x_m)$ und nennen es gerade oder ungerade, je nachdem $x_1 + \cdots + x_m$ gerade oder ungerade ist. (In der Wahrheit hängt $P(x)$ auch von $a_1, \ldots, a_m$ ab.) Es ist klar, daß (nach Ausmultiplizieren) die Elemente von $P(x)$ von der Form sind:

$$\alpha_1^{y_1} \ldots \alpha_m^{y_m} \alpha_{m+1}^{z_{m+1}} \ldots \alpha_n^{z_n} \qquad \begin{pmatrix} 0 \leqq y_i \leqq s_i; & i = 1, \ldots, m \\ 0 \leqq z_k \leqq p-1; & k = m+1, \ldots, n \end{pmatrix}. \tag{29}$$

Ist jedes y_i gleich 0 oder s_i, so gehört (29) offenbar nur einem $P(x)$ zu, und zwar ist

$$\alpha_1^{s_1 x_1} \ldots \alpha_m^{s_m x_m} \alpha_{m+1}^{z_{m+1}} \ldots \alpha_n^{z_n} \tag{30}$$

in einem $P(x)$ enthalten. Wegen (14) sind diese Elemente (30) auch verschieden. Wir teilen sie in zwei Komplexe A_0, A_1 ein, je nachdem $x_1 + \cdots + x_m$

gerade oder ungerade ist. Wegen des Gesagten ist jedenfalls $A_0 \neq A_1$. Der gewünschte Widerspruch wird so entstehen, daß wir $A_0 = A_1$ zeigen. Wir betrachten hierzu auch die übrigen Elemente (29), in denen nämlich mindestens ein y_i mit $0 < y_i < s_i$ vorkommt. Man sieht leicht, daß jedes solche Element gleichvielen geraden und ungeraden $P(x)$ zugehört. Da die Anzahl der geraden und ungeraden $P(x)$ gleich ist und nach (28) alle $P(x)$ gleich G sind, so folgt $A_0 = A_1$ in der Tat, womit der Beweis beendet ist. (Freilich war obiger ringtheoretischer Beweis viel eleganter.)

Unser Lemma im § 3 verdient auch einiges selbständiges Interesse. Seine unmittelbare Verallgemeinerung würde so lauten: Wenn

$$O\{\alpha_1, \ldots, \alpha_n\} = p_1 \ldots p_n,$$
$$p_{i_1} \ldots p_{i_k} \mid O\{\alpha_{i_1}, \ldots, \alpha_{i_k}\} \qquad (1 \leqq i_1 < \cdots < i_k \leqq n;\ 1 \leqq k \leqq n-1)$$

gelten, so gibt es ganze Zahlen $s_1, \ldots, s_n$, für die

$$\begin{pmatrix} \alpha_1^{s_1} & \ldots & \alpha^{s_n} \\ p_1 & \ldots & p_n \end{pmatrix}$$

ein Normalsystem ist. Wir wissen nicht, ob diese Verallgemeinerung richtig ist. Wenn das auch der Fall ist, so könnte man dadurch unseren Beweis (für nicht-p-Gruppen) doch nicht verkürzen. Insofern ist unser Lemma „scharf", als seine weitere Verallgemeinerung auf den Fall beliebiger ganzen Zahlen $e_1, \ldots, e_n (\geqq 2)$ statt $p_1, \ldots, p_n$ sogar schon dann falsch ist, wenn dabei $e_1, \ldots, e_n$ Potenzen von p sind, was wir mit folgendem Beispiel zeigen: Es bestehe $o(\varrho) = p^3$, $o(\sigma) = p$, $\sigma \notin \{\varrho\}$. Für $\alpha_1 = \varrho$, $\alpha_2 = \varrho\sigma$ gelten dann

$$O\{\alpha_1, \alpha_2\} = p^2 \cdot p^2,\ p^2 \mid O\{\alpha_1\}, O\{\alpha_2\},$$

trotzdem ist

$$\begin{pmatrix} \alpha_1^{s_1} & \alpha_2^{s_2} \\ p^2 & p^2 \end{pmatrix}$$

für keine ganzen Zahlen s_1, s_2 ein Normalsystem. Hieraus folgt, daß in unserem Beweis die im Hilfssatz 2 gefaßte Reduktion auf den Fall von Primsimplexen unterbehrlich ist.

(Eingegangen am 29. Januar 1955.)

Literaturverzeichnis

[1] G. Hajós, Többméretű terek egyszeres befedése kockaraccsal, *Matematikai és fizikai lapok*, **48** (1941), S. 37—64 (ungarisch mit deutschem Auszug).

[2] G. Hajós, Über einfache und mehrfache Bedeckung des n-dimensionalen Raumes mit einem Würfelgitter, *Math. Zeitschrift*, **47** (1942), S. 427—467.

[3] L. Rédei, Jelentés az 1942. évi Kőnig Gyula jutalomról: Hajós György munkáinak ismertetése, *Matematikai és fizikai lapok*, **49** (1942), S. 1—16. (Ein ungarisch verfaßter Bericht über die Arbeiten von G. Hajós.)
[4] L. Rédei, Vereinfachter Beweis des Satzes von Minkowski—Hajós, *Acta Sci. Math.*, **13** (1949), S. 21—35.
[5] L. Rédei, Kurzer Beweis des gruppentheoretischen Satzes von Hajós, *Comm. Math. Helv.*, **23** (1949), S. 272—282.
[6] L. Rédei, Die Reduktion des gruppentheoretischen Satzes von Hajós auf den Fall von p-Gruppen, *Monatshefte für Math.*, **53** (1949), S. 221—226.
[7] L. Rédei, *Algebra I* (Budapest, 1954) (ungarisch).
[8] L. Rédei, Die gruppentheoretischen Zetafunktionen und der Satz von Hajós, *Acta Math. Acad. Sci. Hung.*, **6** (1955) (im Erscheinen).
[9] T. Szele, Neuer Vereinfachter Beweis des gruppentheoretischen Satzes von Hajós, *Publicationes Math. (Debrecen)*, **1** (1949), S. 56—62.

НОВОЕ ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ ХАЙОША О КОНЕЧНЫХ АБЕЛЕВЫХ ГРУППАХ

Л. Редеи (Сегед)

(Резюме)

Пусть $G(\neq\varepsilon)$ — конечная абелева группа, ε же единица G. Обозначим порядок G через $O(G)$ а порядок элемента $\alpha(\in G)$ через $o(\alpha)$. Если $o(\alpha)\geqq e\geqq 2$, то обозначим множество элементов $\varepsilon, \alpha, \ldots, \alpha^{e-1}$ через $(\alpha)_e$. Оно образует группу очевидно тогда и только тогда, когда $o(\alpha)=e$. Знаменимая теорема Хайоша, доказывающая гипотезу Минковского „о предельном случае" системы линейных неравенств, утверждает, что

$$G=(\alpha_1)_{e_1}\cdots(\alpha_k)_{e_k},\quad O(G)=e_1\ldots e_k \tag{1}$$

возможно лишь если хотя бы одно из множеств $(\alpha_i)_{e_i}$ образует группу. Для этой теоремы имелось до сих пор только оригинальное доказательство Хайоша (см. Хайош [2]), упрощенное Редеи [5] и Селе [9]. Излагаемое ниже простое доказательство использует прежде всего простой факт, также принадлежащий Хайошу, что теорему достаточно доказать для частного случая

$$G=(\alpha_1)_{p_1}\cdots(\alpha_1)_{p_n},\quad O(G)=p_1\cdots p_n, \tag{2}$$

где $p_1, \ldots, p_n$ простые числа.

Дальнейшим шагом является следующая нетрудная

Лемма. Если $\alpha_1, \ldots, \alpha_n\in G$, p простое число и

$$O(\{\alpha_1, \ldots, \alpha_n\})=p, \tag{3}$$

$$O(\{\alpha_{i_1}, \ldots, \alpha_{i_k}\})\geqq p^k \qquad (1\leqq i_1<\cdots<i_k\leqq n;\ 1\leqq k\leqq n-1), \tag{4}$$

то, удобно располагая элементы $\alpha_1, \ldots, \alpha_n$, можно указать такие степени $s_1, \ldots, s_n(\geqq 1)$ числа p, что

$$O(\{\alpha_1^{s_1}, \ldots, \alpha_k^{s_k}\})=p^k \qquad (1\leqq k\leqq n). \tag{5}$$

($\{\alpha, \beta, \ldots\}$ означает группу, порожденную элементами $\alpha, \beta, \ldots$.)

Если, в частности, G является p-группой, то в соответствии с вышесказанным можно ограничиться исследованием случая

$$G=(\alpha_1)_p\cdots(\alpha_n)_p,\quad O(G)=p^n. \tag{6}$$

Очевидно, что из (6) следуют соотношения (3) и (4). Так как далее (6) справедливо при любом расположении факторов, то можно предположить и (5) выполненным. Если теперь $s_1=1$, то из частного случая (5) соответствующего $k=1$ уже вытекает $o(\alpha_1)=p$ так, что $(\alpha_1)_p$ образует группу. В оставшемся случае можно считать (при удобно выбранном m), что

(7) $$p|s_1, \ldots, s_m;\ s_{m+1}=\cdots=s_n=1 \qquad (1\leqq m\leqq n).$$

С другой стороны из (5) сразу следует

(8) $$G=(\alpha_1^{s_1})_p \ldots (\alpha_n^{s_n})_p.$$

Однако мы из соотношений (6)—(8) выведем противоречие.

Для этой цели используем групповое кольцо $\mathfrak{J}(G)$ группы G, где $\mathfrak{J}$ означает кольцо целых чисел. Обозначим через $[\alpha]_e$ сумму $\varepsilon+\alpha+\cdots+\alpha^{e-1}$ ($\in\mathfrak{J}(G)$). Так как $\alpha_1 G=G$ и $(\varepsilon-\alpha_1)[\alpha_1]_p=\varepsilon-\alpha_1^p$, то из (6) сразу следует

$$(\varepsilon-\alpha_1^p)[\alpha_2]_p\cdots[\alpha_n]_p=0,$$

ввиду (7) тем более

$$(\varepsilon-\alpha_1^{s_1})\cdots(\varepsilon-\alpha_m^{s_m})[\alpha_{m+1}]_p\cdots[\alpha_n]_p=0.$$

На левой стороне равенства появится (после умножения) член ε, которой и равен, с точностью до знака, дальнейшему члену. Ввиду (7) это действительно противоречит (8).

Случай групп не являющихся p-группами можно трактовать опираясь на этот случай (см. Редеи [6]). К сожалению эта часть доказательства, только принципиально проста, подробное доказательство довольно сложно.

SUR LA CARACTÉRISATION DE CERTAINES CLASSES DE FONCTIONS AU SENS DE LA THÉORIE CONSTRUCTIVE DES FONCTIONS

Par

G. ALEXITS (Budapest), membre de l'Académie

Un problème très attirant de la théorie constructive des fonctions consiste dans la caractérisation de la structure des fonctions par l'ordre de grandeur qu'on obtient en les approchant avec les moyennes (C, α) de leurs séries de Fourier. Il y a quelques années nous avons réussi à caractériser de cette manière[1] les classes[2] Lip $(1, p)$ en étudiant l'approximation dans l'espace L^p. Or l'approximation dans l'espace L^p ne donne aucun renseignement sur la vitesse de la convergence atteinte aux différents points de l'intervalle $(0, 2\pi)$. Toutefois, une caractérisation de ces classes par l'ordre de grandeur de l'approximation dans le sens habituel aurait un intérêt, parce qu'elle éclaircirait en détail les propriétés structurelles des fonctions appartenantes à ces classes. De plus, la méthode qui pourrait servir à une telle caractérisation permetterait, en même temps, d'étudier aussi l'ordre de grandeur de l'approximation des fonctions qui ne sont dérivables qu'en certains points. Une étude méthodique de l'approximation aux points individuels pourrait conduire, peut-être, à des résultats plus profonds que ceux qu'on obtient à l'aide des recherches concernant la vitesse de l'approximation en moyenne dans l'intervalle entier $(0, 2\pi)$.

Dans les suivants, nous allons résoudre le problème d'une telle caractérisation des classes Lip $(1, p)$ avec $p > 1$ et essayons de faire les premiers pas vers une étude systématique de l'ordre de grandeur obtenu point par

[1] G. ALEXITS, Sur l'ordre de grandeur de l'approximation d'une fonction périodique par les sommes de Fejér, *Acta Math. Acad. Sci. Hung.*, 3 (1952), p. 29—40 (en français) et p. 41—42 (en russe).

[2] Une fonction $f(x)$ appartient à la classe Lip (α, p), si

$$\sup_{|h| \leqq \delta} \left\{ \int_0^{2\pi} |f(x+h) - f(x)|^p dx \right\}^{\frac{1}{p}} = O(\delta^\alpha)$$

où $0 < \alpha \leqq 1$ et $p \geqq 1$.

point en approximant des fonctions qui ne sont dérivables qu'en certains points. La méthode que nous allons employer consiste dans l'application d'un lemme que nous avons démontré ailleurs[3] et qui nous a servi à nos recherches plus ou moins semblables.

1. Caractérisation des classes Lip $(1, p)$

Désignons par

$$\sigma_n^\alpha(x) = \frac{a_0}{2} + \sum_{k=1}^{n} \frac{\binom{n-k+\alpha}{n-k}}{\binom{n+\alpha}{n}} (a_k \cos kx + b_k \sin kx)$$

la n-ième moyenne (C, α) de la série de Fourier d'une fonction intégrable $f(x)$. La fonction conjuguée sera désignée par $\tilde{f}(x)$ et

$$\tilde{\sigma}_n^\alpha(x) = \sum_{k=1}^{n} \frac{\binom{n-k+\alpha}{n-k}}{\binom{n+\alpha}{n}} (b_k \cos kx - a_k \sin kx)$$

est la n-iéme moyenne (C, α) de la série conjuguée. Le lemme dont nous avons fait mention et que nous avons démontré ailleurs est le suivant:

Si, pour un $\alpha > 0$, on a en un point x

$$|\tilde{f}(x) - \tilde{\sigma}_n^\alpha(x)| \leqq \frac{K}{n},$$

alors il existe une constante C_α ne dépendant que de α telle que $|\sigma_n^{\alpha'}(x)| \leqq \leqq C_\alpha K$. Inversement, si les dérivées $\sigma_n^{\alpha'}(x)$ des moyennes $\sigma_n^\alpha(x)$ satisfont au point x à la condition $|\sigma_n^{\alpha'}(x)| \leqq K$, alors

$$|\tilde{f}(x) - \tilde{\sigma}_n^\alpha(x)| \leqq \frac{C_\alpha K}{n}.$$

A l'aide de ce lemme, nous pouvons obtenir une caractérisation complète des classes Lip $(1, p)$ de la manière suivante:

THÉORÈME 1. *Pour que la fonction $f(x)$ appartienne à la classe* Lip $(1, p)$ *où $p \geqq 1$, il faut qu'il existe, pour tout $\alpha > 0$, une fonction $F_\alpha(x) \in L^p$ telle que*

$$|f(x) - \sigma_n^\alpha(x)| \leqq \frac{F_\alpha(x)}{n}, \quad |\tilde{f}(x) - \tilde{\sigma}_n^\alpha(x)| \leqq \frac{F_\alpha(x)}{n},$$

et il suffit qu'une au moins de ces inégalités soit satisfaite pour un $\alpha > 0$ arbitrairement fixé.

[3] L. c. [1] et G. ALEXITS, A Fourier-sor Cesàro-közepeivel való approximáció nagyságrendjéről, *Mat. Fiz. Lapok*, **48** (1941), p. 410—421 (hongrois) et p. 421—422 (français).

Soit, en effet, $f \in \mathrm{Lip}(1, p)$. Il existe alors une fonction absolument continue $g(x)$ qui ne diffère de $f(x)$ qu'aux points d'un ensemble de mesure nulle et dont la dérivée $g'(x)$ est L^p-intégrable.[4] Alors $\sigma_n^{\alpha'}(x)$ est la n-ième moyenne (C, α) de la série de Fourier de $g'(x)$; il existe[5] donc pour tout $\alpha > 0$ une fonction L^p-intégrable $\Phi_\alpha(x)$ telle que

$$|\sigma_n^{\alpha'}(x)| \leqq \Phi_\alpha(x) \qquad (n = 1, 2, \ldots).$$

En appliquant notre lemme, il s'ensuit que

$$|\tilde{g}(x) - \tilde{\sigma}_n^{\alpha}(x)| \leqq \frac{C_\alpha \Phi_\alpha(x)}{n}.$$

Mais $\tilde{g}(x)$ ne diffère de $\tilde{f}(x)$ qu'en un ensemble de mesure nulle; en changeant donc les valeurs de $\Phi_\alpha(x)$ aux points de cet ensemble, on obtient aussi

$$|\tilde{f}(x) - \tilde{\sigma}_n^{\alpha}(x)| \leqq \frac{C_\alpha \Phi_\alpha(x)}{n}.$$

C'est la deuxième des inégalités que nous voulons démontrer. Quant à la première, on n'a qu'à observer que $f(x)$ et $\tilde{f}(x)$ appartiennent en même temps à la classe $\mathrm{Lip}(1, p)$. En effet, puisque $g' \in L^p$, il s'ensuit[6] pour sa conjuguée: $\tilde{g}' \in L^p$. Or cette relation entraîne[4] que l'intégrale de $\tilde{g}'(x)$, c'est à dire que $\tilde{g}(x)$ appartient à $\mathrm{Lip}(1, p)$. Mais $\tilde{g}(x)$ ne diffère qu'en un ensemble de mesure nulle de $\tilde{f}(x)$, par conséquent $\tilde{f} \in \mathrm{Lip}(1, p)$. On en obtient pour la conjuguée $f(x)$ de la fonction $\tilde{f}(x)$ comme auparavant

$$|f(x) - \sigma_n^{\alpha}(x)| \leqq \frac{C_\alpha \Psi_\alpha(x)}{n}$$

où $\Psi_\alpha(x)$ est L^p-intégrable. En posant $F_\alpha(x) = \max\{C_\alpha \Phi_\alpha(x), C_\alpha \Psi_\alpha(x)\}$, la nécessité de notre condition est démontrée. — Quant à sa suffisance, soit une de nos inégalités, par exemple la première, satisfaite. On en tire par l'application de notre lemme

$$|\tilde{\sigma}_n^{\alpha'}(x)| \leqq C_\alpha F_\alpha(x).$$

Les sommes $\tilde{\sigma}_n^{\alpha'}(x)$ sont donc, d'après des théorèmes connus,[7] les moyennes (C, α) de la série de Fourier d'une fonction L^p-intégrable et de même les sommes $\sigma_n^{\alpha'}(x)$ conjuguées aux $\tilde{\sigma}_n^{\alpha'}(x)$. Les moyennes $\sigma_n^{\alpha'}(x)$ convergent donc presque partout vers une fonction L^p-intégrable $g'(x)$ dont l'intégrale $g(x)$, ayant presque partout une dérivée L^p-intégrable, appartient à la classe

[4] G. H. Hardy—J. E. Littlewood, Some properties of fractional integrals. I, *Math. Zeitschrift*, **27** (1928), p. 565—606.

[5] G. H. Hardy—J. E. Littlewood, A maximal theorem with function-theoretic applications, *Acta Math.*, **54** (1930), p. 81—116.

[6] M. Riesz, Sur les fonctions conjuguées, *Math. Zeitschrift*, **27** (1928), p. 218—249.

[7] V. p. ex. A. Zygmund, *Trigonometrical series* (Warszawa—Lwów, 1935), p. 87.

Lip $(1, p)$. Mais $f(x)$ et $g(x)$ ont toutes les deux $\sigma_n^\alpha(x)$ pour n-ième moyenne (C, α) de leur série de Fourier, elles ne diffèrent donc qu'en un ensemble de mesure nulle, par conséquent $f \in \text{Lip}(1, p)$ et notre théorème est entièrement démontré.

Ce théorème n'est plus valable pour $p = 1$, parce qu'alors aucune de nos inégalités n'est nécessaire et ce n'est que la deuxième inégalité qui assure l'appartenance de $f(x)$ à la classe Lip $(1, 1)$.

2. L'ordre de grandeur de l'approximation des fonctions conjuguées à des fonctions dérivables en certains points

Les précédents contiennent, au fond, la caractérisation des fonctions dérivables presque partout dont les dérivées sont soumises à des conditions d'intégrabilité. En renonçant à ces conditions, nous ne pouvons plus caractériser les classes plus larges, mais la dérivabilité seule, même si elle n'est satisfaite qu'en un seul point, entraîne déjà une convergence rapide des moyennes $\sigma_n^\alpha(x)$ à ce point.

THÉORÈME 2. *Si la fonction $f(x)$ est dérivable au point x_0, alors on a pour tout $\alpha > 1$*

$$\tilde{f}(x_0) - \tilde{\sigma}_n^\alpha(x_0) = O\left(\frac{1}{n}\right).$$

En effet, puisque la dérivée $f'(x_0)$ existe, les dérivées $\sigma_n^{\alpha\prime}(x_0)$ convergent[8] vers $f'(x_0)$, par conséquent

$$\sigma_n^{\alpha\prime}(x_0) = O(1)$$

et la proposition résulte immédiatement de notre lemme.

On voit que $f(x)$ étant dérivable aux points d'un ensemble, le degré d'approximation obtenu à l'aide des moyennes $\tilde{\sigma}_n^\alpha(x)$ peut être évalué par $F(x)/n$ où $F(x)$ est une fonction finie sur cet ensemble, mais nous ne savons rien sur la sturcture de $F(x)$. On obtient un peu plus en supposant $f \in \text{Lip}$ $(1, p)$ dans un intervalle $(a, b) \subset (0, 2\pi)$.

THÉORÈME 3. *Si $f \in \text{Lip}(1, p)$ dans un intervalle (a, b) pour un $p > 1$, alors il existe, pour tout $\varepsilon > 0$, une fonction $F_\varepsilon(x) \in L^p$, telle que*

$$|\tilde{f}(x) - \tilde{\sigma}_n^1(x)| \leqq \frac{F_\varepsilon(x)}{n}$$

aux points x de l'intervalle $(a + \varepsilon, b - \varepsilon)$.

Comme $f \in \text{Lip}(1, p)$ dans l'intervalle (a, b), il existe[4] une fonction $g(x)$ équivalente à $f(x)$ en (a, b) et telle que sa dérivée $g'(x)$ soit $\in L^p(a, b)$.

[8] V. p. ex. l. c. [7], p. 55.

Posons $\overline{g}(x)=g(x)$ pour $x\in(a,b)$ et $\overline{g}(x)=0$ ailleurs. Alors $\overline{g}'(x)$ existe presque partout et $\overline{g}'\in L^p$. En désignant donc par $\sigma_n(\overline{g},x)$ la n-ième moyenne de FEJÉR de la série de Fourier de $\overline{g}(x)$, on obtient comme auparavant

$$|\sigma_n'(\overline{g},x)|\leqq G(x)$$

où $G\in L^p$. Mais, en désignant par $K_n(t,x)$ le noyau de FEJÉR, on obtient aussi

$$|\sigma_n^{1\prime}(x)-\sigma_n'(\overline{g},x)|\leqq\int_0^{2\pi}|f(t)-\overline{g}(t)||K_n'(t,x)|dt=$$

$$=\left(\int_0^a+\int_b^{2\pi}\right)|f(t)-\overline{g}(t)|O\left(\frac{1}{(t-x)^3}\right)dt.$$

On a donc pour tout $\varepsilon>0$ et pour tous les points $x\in(a+\varepsilon,b-\varepsilon)$

$$|\sigma_n^{1\prime}(x)-\sigma_n'(\overline{g},x)|=O\left(\frac{1}{\varepsilon^3}\right)\int_0^{2\pi}|f(t)-\overline{g}(t)|dt=K_\varepsilon.$$

Par conséquent

$$|\sigma_n^{1\prime}(x)|\leqq K_\varepsilon+|\sigma_n'(\overline{g},x)|\leqq K_\varepsilon+G(x)=F_\varepsilon(x)$$

où la fonction $F_\varepsilon(x)$ est évidemment L^p-intégrable. Il s'ensuit par l'application de notre lemme:

$$|\tilde{f}(x)-\tilde{\sigma}_n^1(x)|\leqq\frac{F_\varepsilon(x)}{n},$$ c. q. f. d.

(Reçu le 7 février 1955.)

ОХАРАКТЕРИЗОВАНИЕ НЕКОТОРЫХ КЛАССОВ ФУНКЦИЙ С ПОМОЩЬЮ КОНСТРУКТИВНОЙ ТЕОРИИ ФУНКЦИЙ

Г. Алексич (Будапешт)

(Резюме)

Функции принадлежащие классу Lip $(1,p)$ были охарактеризованы авторам быстротой сходимости в пространстве L^p средних Фейера их сопряженных рядов Фурье. В настоящей работе автор занимается охарактеризованием класса Lip $(1,p)$ с помощью порядка величины почечной аппроксимации. Обозначим для этой цели через $\sigma_n^\alpha(x)$ средние (C,α) порядка $\alpha>0$ ряда Фурье интегрируемой функции $f(x)$, а через $\tilde{f}(x)$ и $\tilde{\sigma}_n^\alpha(x)$ сопряженные фуикций $f(x)$ и $\sigma_n^\alpha(x)$.

I. Необходимым для того, чтобы $f(x)$ принадлежала классу Lip $(1,p)$ $(p>1)$, является существование для всех $\alpha>0$ такой функции $F_\alpha(x)$ интегрируемой L^p, для которой

$$|f(x)-\sigma_n^\alpha(x)|\leqq\frac{F_\alpha(x)}{n},\quad|\tilde{f}(x)-\tilde{\sigma}_n^\alpha(x)|\leqq\frac{F_\alpha(x)}{n};$$

выполнение хотя бы одного из этих неравенств является достаточным для того, чтобы было $f \in \mathrm{Lip}\,(1, p)$.

Соотношение $f \in \mathrm{Lip}\,(1, p)$ по существу означает, что $f(x)$ почти всюду имеет производную, интегрируемую L^p. Если отбросим требование интегрируемости, то можем оценить лишь быстроту сходимости аппроксимации.

II. Если $f(x)$ имеет производную в точке x_0, то для всех $\alpha > 1$

$$\tilde{f}(x_0) - \tilde{\sigma}_n^{\alpha}(x_0) = O\left(\frac{1}{n}\right).$$

Если $f \in \mathrm{Lip}\,(1, p)$ в некотором подинтервале интервала $(0, 2\pi)$, то имеет место следующее более сильное утверждение:

III. Если $f \in \mathrm{Lip}\,(1, p)$ в интервале $(a, b) \subset (0, 2\pi)$ $(p > 1)$, то при любом $\varepsilon > 0$ можно задать функцию $F_\varepsilon(x)$ интегрируемую L^p, для которой

$$|\tilde{f}(x) - \tilde{\sigma}_n^1(x)| \leqq \frac{F_\varepsilon(x)}{n}$$

во всех точках интервала $(a + \varepsilon, b - \varepsilon)$.

ON THE ROLE OF THE LEBESGUE FUNCTIONS IN THE THEORY OF THE LAGRANGE INTERPOLATION [1]

By

P. ERDŐS (Jerusalem) and P. TURÁN (Budapest), member of the Academy

To Prof. dr. L. FEJÉR *on the occasion of his 75th birthday*

1. Let there be given a triangular matrix

$$A \equiv \begin{matrix} x_{11} & & & \\ x_{12} & x_{22} & & \\ \vdots & & \ddots & \\ x_{1n} & x_{2n} & \dots & x_{nn} \\ \vdots & & & \ddots \end{matrix} \tag{1.1}$$

where for $n = 1, 2, \dots$ we have

$$1 \geqq x_{1n} > x_{2n} > \cdots > x_{nn} \geqq -1. \tag{1.2}$$

Then, as it is well known, for given values $y_{\nu n}$ there is exactly one polynomial $g(x)$ of degree $\leqq n-1$ such that

$$g(x_{\nu n}) = y_{\nu n} \qquad (\nu = 1, 2, \dots, n).$$

If the values $y_{\nu n}$ are the values $f(x_{\nu n})$ of a function $f(x)$ defined in $[-1, +1]$, then we call the corresponding $g(x)$ polynomial "the nth interpolatory polynomial of $f(x)$ belonging to A" and denote it by $L_n(f, A)$ or — if misunderstanding cannot arise — by $L_n(f)$. The abscissae $x_{\nu n}$ are called the nth fundamental points of the matrix A and are sometimes denoted also by x_ν. It is well known that $L_n(f, A)$ can be written in the form

$$L_n(f, A) = \sum_{\nu=1}^{n} f(x_{\nu n})\, l_{\nu n}(x, A), \tag{1.3}$$

where the polynomials $l_{\nu n}(x, A)$, the so-called fundamental-functions belonging

[1] A part of the results (assertions a) and b) of this paper) was the subject of a lecture made by one of us at a colloquium for the constructive function-theory in Eger (Hungary), 29 Nov. 1953; they were found twenty years ago. The new results showing they are best possibles were a subject of another lecture in Pécs (Hungary), 18 Sept. 1954.

to A, depend only upon A and have the representation

$$l_{\nu n}(x, A) = \frac{\omega_n(x, A)}{\omega_n'(x_{\nu n}, A)\ (x - x_{\nu n})} \tag{1.4}$$

where

$$\omega_n(x, A) = \prod_{\nu=1}^{n} (x - x_{\nu n}). \tag{1.5}$$

As easy to verify, we have

$$\sum_{\nu=1}^{n} l_{\nu n}(x, A) \equiv 1 \tag{1.6}$$

and if $h(x)$ denotes an arbitrary polynomial of degree $\leqq n-1$, then

$$L_n(h, A) \equiv h(x). \tag{1.7}$$

From (1.3) it follows that for an $f(x)$ bounded in $[-1, +1]$ we have for $-1 \leqq x \leqq +1$

$$|L_n(f, A)| \leqq \left(\sum_{\nu=1}^{n} |l_{\nu n}(x, A)| \right) \sup_{-1 \leqq x \leqq +1} |f(x)|. \tag{1.8}$$

2. One would be inclined to think that if A is such that for an arbitrarily small $\varepsilon > 0$

$$x_{\nu n} - x_{\nu+1, n} \leqq \varepsilon, \quad x_{0n} = -x_{n+1, n} = 1 \qquad (\nu = 0, 1, \ldots, n;\ n > n_0(\varepsilon)), \tag{2.1}$$

then the sequence $L_n(f, A)$ converges uniformly in $[-1, +1]$ to $f(x)$ whenever[2] $f(x) \in C$. It was a great surprise at the end of the last century when RUNGE and BOREL discovered that the sequence $L_n(f, B)$, belonging to the "most classical" matrix B defined by

$$x_{\nu n} = 1 - \frac{2\nu}{n+1} \qquad (\nu = 1, 2, \ldots, n;\ n = 1, 2, \ldots),$$

can diverge in a whole subinterval of $[-1, +1]$ for such a simple function as $f(x) = \dfrac{1}{1+x^2}$. This would leave open the possibility, the situation can be saved by choosing another "better" A matrix. But G. FABER[3] discovered in 1910 the shaking fact that *no* matrix A can assure the uniform convergence of the polynomials $L_n(f, A)$ for every $f \in C$. His proof showed that it is essential for this phenomenon a property of the quantity

$$M_n(A) \equiv \max_{-1 \leqq x \leqq +1} \lambda_n(x, A) = \max_{-1 \leqq x \leqq +1} \sum_{\nu=1}^{n} |l_{\nu n}(x, A)|, \tag{2.2}$$

namely that for *all* A-matrices we have

$$\overline{\lim_{n \to \infty}}\ M_n(A) = +\infty. \tag{2.3}$$

[2] C denotes, as usual, the class of functions continuous for $-1 \leqq x \leqq +1$.

[3] G. FABER, Über die interpolatorische Darstellung stetiger Funktionen, *Jahresber. der Deutsch. Math. Ver.*, **23** (1914), pp. 190–210.

We call the quantity $\lambda_n(x, A)$ the nth Lebesgue function, the quantity $M_n(A)$ the nth Lebesgue constant belonging to the matrix A. Moreover, S. BERNSTEIN[4] proved the still more surprising fact that to every matrix A there belongs an x_0 in $-1 \leqq x_0 \leqq +1$ and an $f_0 \in C$ with

$$\overline{\lim_{n\to\infty}} |L_n(f_0, A)|_{x=x_0} = +\infty \tag{2.4}$$

what roots in the relation

$$\overline{\lim_{n\to\infty}} \lambda_n(x_0, A) = +\infty. \tag{2.5}$$

3. These facts show that for the divergence theory of the Lagrange interpolation the functions $\lambda_n(x, A)$ are essential. But it was observed by FEJÉR[5] that these Lebesgue functions $\lambda_n(x, A)$ play also a role in the convergence theory. His simple reasoning reproduced for the sake of completeness in **9** gives the following theorem:

If the quantities $M_n(A)$ of (2.2) satisfy the inequality

$$M_n(A) < c_1 n^{\beta} \tag{3.1}$$

with a fixed $0 < \beta < 1$ and numerical[6] c_1, then the polynomials $L_n(f, A)$ converge for $-1 \leqq x \leqq +1$ uniformly to $f(x)$ if[7] $f \in \mathrm{Lip}\,\gamma$, $\gamma > \beta$.

4. These results are responsible for the rather general opinion that the convergence-divergence theory of the Lagrange interpolation is by and large identical with the study of the Lebesgue constants $M_n(A)$. We have set ourselves the task to investigate to which extent this is true. We have found that going a little beyond the mere continuity this fails to be true; the result can quite vaguely expressed saying that *there is a "rough" and a "fine" convergence-divergence theory for the Lagrange interpolation*. To be more exact, let us consider, if

$$0 < \beta < 1, \tag{4.1}$$

the class $A(\beta)$ of all A-matrices for which with arbitrarily small positive ε we have

$$\overline{\lim_{n\to\infty}} M_n(A) n^{-\beta-\varepsilon} < c_2(\varepsilon), \tag{4.2}$$

$$\overline{\lim_{n\to\infty}} M_n(A) n^{-\beta+\varepsilon} > c_3(\varepsilon), \tag{4.3}$$

[4] S. BERNSTEIN, Sur la limitation des valeurs etc., *Bull. Acad. Sc. De l'URSS*, **8** (1931), pp. 1025—1050.

[5] L. FEJÉR, Lagrangesche Interpolation und die zugehörigen konjugierten Punkte, *Math. Ann.*, **106** (1932), pp. 1—55.

[6] Later on $c_2, c_3, \ldots$ denote generally also numerical constants. If some c_ν depends upon some parameters, the dependence will be explicitely stated.

[7] As usual, the class $\mathrm{Lip}\,\gamma$ denotes the totality of those functions which satisfy uniformly a Lipschitz-condition with the exponent γ in $-1 \leqq x \leqq +1$.

i. e. for our matrices the Lebesgue constants $M_n(A)$ increase roughly speaking like n^β. We call for a fixed β with (4. 1) the Lip γ class of functions

"a good class of functions $f(x)$ for $A(\beta)$",	if for all $A \in A(\beta)$ and $f \in$ Lip γ $L_n(f, A)$ tend uniformly in $[-1, +1]$ to $f(x)$ for $n \to \infty$,

respectively

"a bad class of functions $f(x)$ for $A(\beta)$",	if for all $A \in A(\beta)$ there is an $f_1(x) \in$ Lip γ such that $L_n(f_1, A)$ is unbounded in $[-1, +1]$ for $n \to \infty$.

If a Lip γ class is a good or a bad one, then its convergence-divergence behaviour is by and large determined by the numbers $M_n(A)$; one would be inclined to think that all Lip γ classes with $\gamma < \beta$ are bad and all with $\gamma > \beta$ are good classes and thus the finer structure of the matrix A cannot essentially influence the convergence-divergence behaviour for the respective Lip γ class. A closer investigation showed, howewer, that this is not the case, there are values γ depending only upon β for which the Lip γ classes are neither good nor bad ones. This means that the convergence-divergence behaviour is certainly not determined for the respective Lip γ class alone by the Lebesgue constants $M_n(A)$; hence the convergence-divergence behaviour for the respective Lip γ class depends upon the finer structure of A and thus the determination of the convergence-divergence behaviour belongs to a "finer" theory. Thus even the existence of this "finer" theory is somewhat surprising; moreover we shall see that *for fixed* $0 < \beta < 1$ *all* Lip γ *classes with*

$$\gamma < \frac{\beta}{\beta + 2} \tag{4. 4}$$

are bad classes, all Lip γ *classes with*

$$\gamma > \beta \tag{4. 5}$$

are good ones and the Lip γ *classes with*

$$\frac{\beta}{\beta + 2} < \gamma < \beta \tag{4. 6}$$

form the exact domain of the finer convergence-divergence theory.

Similar questions arise in connection with orthogonal expansions, singular integrals, mechanical quadrature and generally with linear operations. Also other scales than the Lip-classes can be used. Further, the convergence behaviour can instead of uniform convergence refer e. g. to pointwise convergence. Finally, perhaps the matrix-class $A(\beta)$ can be defined more suitably than in (4. 2)—(4. 3).

5. Our result is given under (4. 4), (4. 5) and (4. 6). In order to prove it we brake the assertion into four parts. β is fixed with

$$0 < \beta < 1. \tag{5. 1}$$

a) If $\gamma < \frac{\beta}{\beta+2}$ and $A \in A(\beta)$, then there is an $f_2 \in \operatorname{Lip}\gamma$ such that the sequence $L_n(f_2, A)$ is unbounded for $-1 \leqq x \leqq 1$, i. e. the class $\operatorname{Lip}\gamma$ is bad in this case.

b) If $\gamma > \beta$ and $A \in A(\beta)$, then the sequence $L_n(f, A)$ converges uniformly in $[-1, +1]$ to $f(x)$ whenever $f \in \operatorname{Lip}\gamma$, i. e. the class $\operatorname{Lip}\gamma$ is good in this case.

c) If $\gamma > \frac{\beta}{\beta+2}$, then there is a matrix $A_0 \in A(\beta)$ such that the sequence $L_n(f, A_0)$ converges uniformly in $[-1, +1]$ to $f(x)$ whenever $f \in \operatorname{Lip}\gamma$, i. e. the class $\operatorname{Lip}\gamma$ is certainly not a bad class.

d) If $\gamma < \beta$, then there is a matrix $A_1 \in A(\beta)$ and $f_3 \in \operatorname{Lip}\gamma$ such that the sequence $L_n(f_3, A_1)$ is unbounded for $-1 \leqq x \leqq +1$, i. e. the class $\operatorname{Lip}\gamma$ is certainly not a good one.[8]

6. To prove the assertion a) we need the simple

LEMMA I. *If $A \in A(\beta)$, then we have for $\nu = 1, 2, \ldots, (n-1)$*

$$x_{\nu n} - x_{\nu+1,\,n} > c_4(\varepsilon) n^{-\beta-2-\varepsilon}. \tag{6.1}$$

For the proof we consider the quantity

$$\frac{1}{x_{\nu n} - x_{\nu+1,\,n}}.$$

Owing to the definition of the fundamental functions and the mean-value theorem

$$\frac{1}{x_{\nu n} - x_{\nu+1,\,n}} = \frac{l_{\nu n}(x_{\nu n}, A) - l_{\nu n}(x_{\nu+1,\,n}, A)}{x_{\nu n} - x_{\nu+1,\,n}} = l'_{\nu n}(\zeta, A), \tag{6.2}$$

where

$$x_{\nu+1,\,n} \leqq \zeta \leqq x_{\nu n}.$$

But using MARKOV's well-known theorem we have

$$|l'_{\nu n}(\zeta)| \leqq (n-1)^2 \max_{-1 \leqq x \leqq +1} |l_{\nu n}(x, A)|,$$

i. e. a fortiori

$$|l'_{\nu n}(\zeta, A)| < n^2 M_n(A). \tag{6.3}$$

From (4.2) we have $M_n(A) \leqq c_5(\varepsilon) n^{\beta+\varepsilon}$, i. e. (6.2) and (6.3) give

$$\frac{1}{x_{\nu n} - x_{\nu+1,\,n}} < c_5(\varepsilon) n^{\beta+2+\varepsilon}$$

which proves (6.1)

[8] By modifications of the construction we could prove that in cases c) and d) also a matrix with $c_1 n^\beta \leqq \max_{-1 \leqq x \leqq +1} \sum_{\nu=1}^{n} |l_{\nu n}(x)| \leqq c_2 n^\beta$ could have been constructed with the other required properties. Also the investigation of the limiting cases is of interest.

7. Now we turn to the proof of the assertion a). Let

$$\sum_{\nu=1}^{n} |l_{\nu n}(x, A)|$$

maximal in $[-1, +1]$ for $x = \zeta_n$. With the convention

$$x_{0n} = -x_{n+1,n} = 1, \quad l_{0n}(x, A) \equiv l_{1n}(x, A), \quad l_{n+1,n}(x, A) \equiv l_{nn}(x, A)$$

we consider the function $\varphi_n(x)$ defined by the broken line with the vertices at the points

$$P_\nu \equiv (x_{\nu n}, \operatorname{sign} l_{\nu n}(\zeta_n, A)) \qquad (\nu = 0, 1, \ldots, n, n+1).$$

Then we have obviously

$$L_n(\varphi_n, A)_{x=\zeta_n} = \sum_{\nu=1}^{n} |l_{\nu n}(\zeta_n, A)| = M_n(A). \tag{7. 1}$$

According to the Lemma I all slopes of $\varphi_n(x)$ are absolutely

$$< c_5(\varepsilon) n^{\beta+2+\varepsilon}, \tag{7. 2}$$

and for $-1 \leqq x \leqq +1$

$$|\varphi_n(x)| \leqq 1. \tag{7. 3}$$

Now we can construct $f_2(x)$ in a way which is a suitably modified form of the resonance principle of LEBESGUE. Since $\gamma < \frac{\beta}{\beta+2}$, we may choose ε so small that

$$0 < \varepsilon < \frac{\beta}{10} \tag{7. 4}$$

and

$$\frac{\beta - 2\varepsilon}{\beta + 2 + 2\varepsilon} > \gamma; \tag{7. 5}$$

we fix ε. According to (4. 3) there is an infinite sequence

$$2^2 \leqq n_1 < n_2 < \cdots$$

such that for $\nu = 1, 2, \ldots$

$$M_{n_\nu}(A) = \lambda_{n_\nu}(\zeta_{n_\nu}, A) > \frac{c_3(\varepsilon)}{2} n_\nu^{\beta-\varepsilon}. \tag{7. 6}$$

Now we select a suitable sub-sequence of the n_ν's which we shall denote by r_ν's. Let

$$r_1 = n_1$$

and we suppose

$$r_1, r_2, \ldots, r_{\nu-1}$$

are already defined. We distingish two cases.

CASE I. Denoting

$$F_{\nu-1}(x) = \sum_{j=1}^{\nu-1} \frac{1}{r_j^{\beta-2\varepsilon}} \varphi_{r_j}(x)$$

the sequence

$$L_\mu(F_{\nu-1}, A) \qquad (\mu = 1, 2, \ldots)$$

is unbounded. In this case for

$$-1 \leqq x < x+h \leqq 1$$

we have

$$|F_{\nu-1}(x+h) - F_{\nu-1}(x)| \leqq \sum_{j=1}^{\nu-1} \frac{1}{r_j^{\beta-2\varepsilon}} |\varphi_{r_j}(x+h) - \varphi_{r_j}(x)|,$$

i. e. using (7. 2)

$$|F_{\nu-1}(x+h) - F_{\nu-1}(x)| \leqq c_5(\varepsilon) h \sum_{j=1}^{\nu-1} r_j^{2+3\varepsilon} < c_6(\varepsilon) h,$$

i. e. $F_{\nu-1}(x)$ belongs even to Lip 1.

CASE II. The sequence $L_\mu(F_{\nu-1}, A)$ is bounded, i. e. for $[-1, +1]$ and $\mu = 1, 2, \ldots$

(7. 7) $$|L_\mu(F_{\nu-1}, A)| \leqq C_{\nu-1}.$$

Then let r_ν be the smallest integer satisfying the conditions

(7. 8) $$r_\nu > r_{\nu-1}^2 \qquad \text{(i. e. also } > 2r_{\nu-1}),$$

(7. 9) $$r_\nu > C_{\nu-1}^{\frac{2}{\varepsilon}}.$$

We may suppose that we have for all ν's the Case II and we assert that in this case we may choose as $f_2(x)$ of assertion a)

(7. 10) $$f_2(x) = \sum_{j=1}^{\infty} \frac{1}{r_j^{\beta-2\varepsilon}} \varphi_{r_j}(x).$$

In order to show that $f_2(x) \in \operatorname{Lip} \gamma$ we write for an x and h satisfying

$$-1 \leqq x < x+h \leqq +1, \quad 0 < h < \frac{1}{100}$$

$$f_2(x+h) - f_2(x) = \sum_{j=1}^{p} + \sum_{j=p+1}^{\infty} \frac{1}{r_j^{\beta-2\varepsilon}} (\varphi_{r_j}(x+h) - \varphi_{r_j}(x))$$

where the index p is defined uniquely by

(7. 11) $$r_p \leqq \left(\frac{1}{h}\right)^{\frac{1}{\beta+2+2\varepsilon}} < r_{p+1}.$$

Using for $\nu \leqq p$ (7. 2) and for $\nu \geqq p+1$ (7. 3), we get

$$|f_2(x+h) - f_2(x)| \leqq c_5(\varepsilon) h \sum_{j=1}^{p} r_j^{2+3\varepsilon} + 2 \sum_{j=p+1}^{\infty} r_j^{-\beta+2\varepsilon} <$$

$$< 2c_5(\varepsilon) \left\{ h p r_p^{2+3\varepsilon} + \frac{1}{r_{p+1}^{\beta-2\varepsilon}} \sum_{j=p+1}^{\infty} \left(\frac{r_{p+1}}{r_j}\right)^{\beta-2\varepsilon} \right\},$$

i. e. using (7. 8)

(7. 12) $$|f_2(x+h) - f_2(x)| \leqq c_6(\varepsilon, \beta) \left\{ h p r_p^{2+3\varepsilon} + \frac{1}{r_{p+1}^{\beta-2\varepsilon}} \right\}.$$

(7. 8) gives $r_\nu > 2^{\nu-1} r_1 \geqq 2^\nu$, i. e. $p < 2 \log r_p$; hence from this and (7. 11)

$$h p r_p^{2+3\varepsilon} < c_7(\varepsilon) h r_p^{2+4\varepsilon} < c_7(\varepsilon) h^{\frac{\beta-2\varepsilon}{\beta+2+2\varepsilon}}, \quad \frac{1}{r_{p+1}^{\beta-2\varepsilon}} < h^{\frac{\beta-2\varepsilon}{\beta+2+2\varepsilon}},$$

i. e. (7. 12) and (7. 5) give at once

$$|f_2(x+h) - f_2(x)| \leqq c_8(\varepsilon, \beta) h^{\frac{\beta-2\varepsilon}{\beta+2+2\varepsilon}} < c_8(\varepsilon, \beta) h^\gamma.$$

Thus our $f_2(x)$ belongs to the class Lip γ as stated.

8. We have also to show that the polynomials $L_n(f_2, A)$ are unbounded in $[-1, +1]$. Let $s \geqq 2$ and write

$$f_2(x) = F_{s-1}(x) + \frac{1}{r_s^{\beta-2\varepsilon}} \varphi_{r_s}(x) + \Phi_s(x) \tag{8. 1}$$

with

$$\Phi_s(x) = \sum_{j=s+1}^{\infty} \frac{1}{r_j^{\beta-2\varepsilon}} \varphi_{r_j}(x).$$

(7. 7) gives for $-1 \leqq x \leqq +1$

$$|L_{r_s}(F_{s-1}, A)| \leqq C_{s-1}. \tag{8. 2}$$

From (7. 1) and (4. 3) we have

$$r_s^{-\beta+2\varepsilon} |L_{r_s}(\varphi_{r_s}, A)|_{x=\zeta_{r_s}} > c_9(\varepsilon) r_s^\varepsilon. \tag{8. 3}$$

Finally, (1. 8) gives for $-1 \leqq x \leqq +1$

$$|L_{r_s}(\Phi_s, A)| \leqq M_{r_s}(A) \max_{-1 \leqq x \leqq +1} |\Phi_s(x)|;$$

hence from (4. 2) and (7. 3), by (7. 8),

$$|L_{r_s}(\Phi_s, A)| \leqq c_{10}(\varepsilon) r_s^{\beta+\varepsilon} \sum_{j=s+1}^{\infty} \frac{1}{r_j^{\beta-2\varepsilon}} < c_{11}(\varepsilon, \beta) r_s^{-\beta+2\varepsilon}.$$

From this, (8. 2) and (8. 3) we obtained

$$|L_{r_s}(f_2, A)|_{x=\zeta_{r_s}} > c_9(\varepsilon) r_s^\varepsilon - C_{s-1} - c_{11}(\varepsilon, \beta) r_{s+1}^{-\beta+2\varepsilon}$$

what proves the unboundedness, using also (7. 9).

9. Next we turn to the proof of the assertion b) in **5.** This is based on an idea of S. BERNSTEIN adapted to interpolation by L. FEJÉR;[5] as told, we only reproduce it for the sake of completeness. Let $P_{n-1}(x)$ be the best-approximating polynomial of $(n-1)^{\text{th}}$ degree of $f(x)$ in $[-1, +1]$ in CHEBYSEV's sense. If $f \in$ Lip γ, then according to S. BERNSTEIN we have in $[-1, +1]$

$$|f(x) - P_{n-1}(x)| \leqq c_{12} n^{-\gamma}. \tag{9. 1}$$

Owing to (1. 7) we have

$$L_n(P_{n-1}, A) \equiv P_{n-1}(x),$$

i. e.

$$L_n(f, A) - f(x) = L_n(f - P_{n-1}, A) + (P_{n-1}(x) - f(x))$$

and thus

$$|L_n(f, A)-f| \leqq c_{12} n^{-\gamma} + |L_n(f-P_{n-1}, A)|. \tag{9.2}$$

But using (1.3) and (9.1) we obtain

$$|L_n(f-P_{n-1}, A)| \leqq c_{12} n^{-\gamma} \sum_{\nu=1}^{n} |l_{\nu n}(x, A)|. \tag{9.3}$$

Choosing in (4.2) ε so small that $\beta+\varepsilon<\gamma-\varepsilon$ and fixing it, we obtain

$$\sum_{\nu=1}^{n} |l_{\nu n}(x, A)| \leqq c_{13}(\varepsilon) n^{\beta+\varepsilon} < c_{13}(\varepsilon) n^{\gamma-\varepsilon}$$

and from (9.3)

$$|L_n(f-P_{n-1}, A)| \leqq c_{14}(\varepsilon) n^{-\varepsilon}.$$

This and (9.2) prove already the assertion b).

10. Next we turn to the proof of the assertion c) in **5.** We shall show that the matrix whose n^{th} row consists of

$$x_{1n} = \cos\left(\frac{3\pi}{2n} - \frac{1}{n^{1+\beta}}\right), \quad x_{\nu n} = \cos\frac{2\nu-1}{2n}\pi \qquad (\nu = 2, 3, \ldots, n) \tag{10.1}$$

belongs to $A(\beta)$ and fulfils the requirements for A_0 of the assertion c). In what follows, we shall speak about one line of the matrix so that instead of $x_{\nu n}$ and $l_{\nu n}(x, A_0)$ we may write x_ν and $l_\nu(x, A_0)$. We have obviously

$$x_1-x_2 = 2\sin\frac{1}{2n^{1+\beta}} \sin\left(\frac{3\pi}{2n} - \frac{1}{2n^{1+\beta}}\right) = \frac{3\pi}{2} n^{-\beta-2} + O(n^{-2-2\beta}), \tag{10.2}$$

i. e. for $n > n_0$

$$x_1-x_2 < 2\pi n^{-\beta-2}. \tag{10.3}$$

Hence

$$\frac{n^{\beta+2}}{2\pi} < \frac{1}{x_1-x_2} = \frac{l_1(x_1, A_0)-l_1(x_2, A_0)}{x_1-x_2} = \left|\frac{l_1(x_1, A_0)-l_1(x_2, A_0)}{x_1-x_2}\right| \leqq$$

$$\leqq \max_{-1\leqq x\leqq+1} \left|\frac{d}{dx} l_1(x, A_0)\right| < n^2 \max_{-1\leqq x\leqq 1} |l_1(x, A_0)| \leqq n^2 M_n(A_0),$$

using again MARKOV's theorem. Hence

$$M_n(A_0) \geqq \frac{1}{2\pi} n^{\beta}. \tag{10.4}$$

If we show that

$$M_n(A_0) \leqq c_{15} n^{\beta}, \tag{10.5}$$

then A_0 belongs indeed to $A(\beta)$.

11. For the proof of (10. 5) we need some lemmas. We shall denote by T the matrix the nth row of which consists of the numbers

$$x_\nu^* = \cos\frac{2\nu-1}{2n}\pi \qquad (\nu = 1, 2, \ldots, n) \tag{11. 1}$$

and by $l_\nu(x, T)$ the fundamental functions belonging to T. Then we need the

LEMMA II. *We have for* $-1 \leqq x \leqq +1$

$$\sum_{\nu=1}^{n} |l_\nu(x, T)| \leqq c_{16} \log n.$$

This lemma is well known.[9] We need further the

LEMMA III. *If* $n \geqq 4$, *we have for* $-1 \leqq x \leqq +1$

$$\sum_{\nu=3}^{n} |l_\nu(x, A_0)| \leqq c_{17} \log n.$$

PROOF. For $\nu \geqq 2$ we have

$$\frac{l_\nu(x, A_0)}{l_\nu(x, T)} = \frac{x-x_1}{x_\nu-x_1} \cdot \frac{x_\nu-x_1^*}{x-x_1^*} = \left(1-\frac{x_1^*-x_1}{x_1^*-x}\right)\left(1+\frac{x_1^*-x_1}{x_1-x_\nu}\right). \tag{11. 2}$$

Since for $\nu \geqq 3$, $n \geqq 3$

$$0 < 1+\frac{x_1^*-x_1}{x_1-x_\nu} < 1+\frac{\cos\frac{\pi}{2n}-\cos\frac{3\pi}{2n}}{\cos\frac{3\pi}{2n}-\cos\frac{5\pi}{2n}} = 1+\frac{1}{2\cos\frac{\pi}{n}} \leqq 1+\frac{1}{\sqrt{2}},$$

we have

$$\left|\frac{l_\nu(x, A_0)}{l_\nu(x, T)}\right| \leqq \left(1+\frac{1}{\sqrt{2}}\right)\left|1-\frac{x_1^*-x_1}{x_1^*-x}\right|. \tag{11. 3}$$

If $-1 \leqq x \leqq x_1$, then owing to

$$0 < 1-\frac{x_1^*-x_1}{x_1^*-x} < 1$$

we have from (11. 3) for $-1 \leqq x \leqq x_1$

$$|l_\nu(x, A_0)| \leqq 2|l_\nu(x, T)|,$$

i. e. summing for $\nu = 3, 4, \ldots, n$ and using Lemma II we obtain

$$\sum_{\nu=3}^{n} |l_\nu(x, A_0)| \leqq 2c_{17} \log n, \tag{11. 4}$$

i. e. Lemma III is proved for $-1 \leqq x \leqq x_1$. To prove it also for $x_1 \leqq x \leqq 1$ we write (11. 3) in the form

$$|l_\nu(x, A_0)| < 2\left|\frac{x-x_1}{x-x_1^*}\right| |l_\nu(x, T)| = 2|x-x_1|\left|\frac{l_\nu(x, T)}{x-x_1^*}\right|.$$

[9] See S. N. BERNSTEIN, Quelques remarques sur l'interpolation, *Math. Ann.*, **79** (1918), pp. 1—12.

Let us observe that for $\nu \geqq 3$ and $x_1 \leqq x \leqq x_1^*$ we have

$$\operatorname{sign} l_\nu(x, T) = (-1)^\nu \tag{11.5}$$

and for $x_1^* \leqq x \leqq 1$

$$\operatorname{sign} l_\nu(x, T) = (-1)^{\nu+1}. \tag{11.6}$$

Hence for $x_1 \leqq x \leqq x_1^*$, $\nu \geqq 3$

$$|l_\nu(x, A_0)| < 2|x-x_1| \frac{(-1)^\nu l_\nu(x, T)}{|x-x_1^*|},$$

i. e.

$$\sum_{\nu=3}^{n} |l_\nu(x, A_0)| < 2|x-x_1| \left| \frac{\sum_{\nu=3}^{n} (-1)^\nu l_\nu(x, T)}{x-x_1^*} \right|. \tag{11.7}$$

A similar reasoning shows, (11.7) holds also for $x_1^* \leqq x \leqq 1$, i. e. (11.7) holds for $x_1 \leqq x \leqq 1$. Applying the mean-value theorem to the expression

$$\frac{\sum_{\nu=3}^{n} (-1)^\nu l_\nu(x, T)}{x-x_1^*}$$

and also

$$|x-x_1| < 1-\cos\frac{3\pi}{2n} < \frac{10^2}{n^2},$$

we obtain from (11.7) for $x_1 \leqq x \leqq 1$

$$\sum_{\nu=3}^{n} |l_\nu(x, A_0)| < 2\frac{10^2}{n^2} \max_{-1\leqq x\leqq+1} \left| \frac{d}{dx} \sum_{\nu=3}^{n} (-1)^\nu l_\nu(x, T) \right|.$$

Using again MARKOV's theorem the right side is

$$< \frac{200}{n^2}(n-1)^2 \max_{-1\leqq x\leqq+1} \left| \sum_{\nu=3}^{n} (-1)^\nu l_\nu(x, T) \right| < 200 \max_{-1\leqq x\leqq 1} \sum_{\nu=1}^{n} |l_\nu(x, T)| \leqq$$

$$\leqq 200\, c_{16} \log n$$

by Lemma II. By this and (11.4) Lemma III is proved.

12. It follows from Lemma III that (10.5) will be proved if we succeed in showing that for $-1 \leqq x \leqq +1$ we have

$$|l_1(x, A_0)| \leqq c_{18} n^\beta, \quad |l_2(x, A_0)| \leqq c_{18} n^\beta. \tag{12.1}$$

This will follow as a byproduct from the following Lemma which we shall need in **13.**

LEMMA IV. *For* $-1 \leqq x \leqq +1$ *we have*

$$|l_1(x, A_0) + l_2(x, A_0)| \leqq c_{19}.$$

PROOF. From (1.4) we have

$$\frac{l_1(x, A_0)}{l_1(x, T)} = \prod_{j=2}^{n} \frac{x_1^*-x_j}{x_1-x_j} = \frac{T_n'(x_1^*)}{l_n'(x_2)} \left\{ (x_1-x_1^*) \frac{T_n'(x_2)}{T_n(x_1)} \right\}. \tag{12.2}$$

The factor in the bracket is

$$\left\{\cos\frac{\pi}{2n}-\cos\left(\frac{3\pi}{2n}-\frac{1}{n^{1+\beta}}\right)\right\}\frac{n}{\sin\frac{3\pi}{2n}}\cdot\frac{1}{\cos\left(\frac{3\pi}{2}-n^{-\beta}\right)}=-2\frac{\pi}{2n}\left\{1+O\left(\frac{1}{n^{\beta}}\right)\right\}\cdot \tag{12.3}$$

$$\cdot\frac{\pi}{n}\left\{1+O\left(\frac{1}{n^{\beta}}\right)\right\}\left\{\frac{2}{3\pi}n^{2}\right\}\left\{1+O\left(\frac{1}{n^{2}}\right)\right\}\left\{n^{\beta}\right\}\left\{1+O\left(\frac{1}{n^{2\beta}}\right)\right\}=-\frac{2\pi}{3}n^{\beta}+O(1).^{10}$$

Taking in account that

$$\frac{x_2-x_1^*}{x_2-x_1}=\frac{\cos\frac{3\pi}{2n}-\cos\frac{\pi}{2n}}{\cos\frac{3\pi}{2n}-\cos\left(\frac{3\pi}{2n}-n^{-1-\beta}\right)}=\frac{\pi^2n^{-2}\left\{1+O\left(\frac{1}{n^2}\right)\right\}}{\frac{3\pi}{2}n^{-2-\beta}\{1+O(n^{-\beta})\}}=$$

$$=\frac{2\pi}{3}n^{\beta}\{1+O(n^{-\beta})\}=\frac{2\pi}{3}n^{\beta}+O(1),$$

this and (12.3) means that

$$(x_1-x_1^*)\frac{T_n'(x_2)}{T_n(x_1)}=-\frac{x_2-x_1^*}{x_2-x_1}+O(1)$$

and from (12.2), using the explicit form of $l_1(x, T)$,

$$l_1(x, A_0)=\frac{T_n(x)}{x-x_1^*}\cdot\frac{1}{T_n'(x_2)}\left\{-\frac{x_2-x_1^*}{x_2-x_1}+O(1)\right\}= \tag{12.4}$$

$$=-\frac{x_2-x_1^*}{x_2-x_1}\cdot\frac{T_n(x)}{(x-x_1^*)T_n'(x_2)}+O(1)|l_1(x, T)|\left|\frac{T_n'(x_1^*)}{T_n'(x_2)}\right|.$$

As well known,[5] we have for $-1\leqq x\leqq+1$

$$|l_\nu(x, T)|\leqq\sqrt{2}\qquad(\nu=1, \ldots, n), \tag{12.5}$$

i. e. the second term on the right of (12.4) is

$$O(1)\left|\frac{T_n'(x_1^*)}{T_n'(x_2)}\right|=O(1)\frac{\sin\frac{3\pi}{2n}}{\sin\frac{\pi}{2n}}=O(1);$$

thus from (12.4)

$$l_1(x, A_0)=-\frac{x_2-x_1^*}{x_2-x_1}\cdot\frac{T_n(x)}{x-x_1^*}\cdot\frac{1}{T_n'(x_2)}+O(1). \tag{12.6}$$

Further we have

$$\frac{l_2(x, A_0)}{l_2(x, T)}=\frac{x-x_1}{x-x_1^*}\cdot\frac{x_2-x_1^*}{x_2-x_1},$$

[10] The O-sign refers to $n\to\infty$ but always uniformly for $-1\leqq x\leqq+1$.

i. e. using the explicit form of $l_2(x, T)$

$$l_2(x, A_0) = \frac{x_2 - x_1^*}{x_2 - x_1} \cdot \frac{x - x_1}{x - x_1^*} \cdot \frac{T_n(x)}{x - x_2} \cdot \frac{1}{T_n'(x_2)}.$$

From this and (12. 6)

$$|l_1(x, A_0) + l_2(x, A_0)| = O(1) + \left|\frac{x_2 - x_1^*}{x_2 - x_1}\right| \left|\frac{T_n(x)}{T_n'(x_2)}\right| \left|-\frac{1}{x - x_1^*} + \frac{x - x_1}{x - x_1^*} \cdot \frac{1}{x - x_2}\right| =$$

$$= O(1) + \left|\frac{x_2 - x_1^*}{x_2 - x_1}\right| \left|\frac{T_n(x)}{T_n'(x_2)\,(x - x_1^*)}\right| \left|\frac{x_1 - x_2}{x - x_2}\right| = \tag{12. 7}$$

$$= O(1) + |x_2 - x_1^*| \frac{1}{|T_n'(x_2)|} \left|\frac{T_n(x)}{(x - x_1^*)(x - x_2)}\right| = O(1) + O\left(\frac{1}{n^4}\right) \left|\frac{T_n(x)}{(x - x_1^*)(x - x_2)}\right|.$$

Since we have for $-1 \leqq x \leqq \cos \frac{2\pi}{n}$

$$\left|\frac{T_n(x)}{(x - x_1^*)(x - x_2)}\right| \leqq \frac{1}{\left(\cos \frac{\pi}{2n} - \cos \frac{2\pi}{n}\right)\left(\cos \frac{3\pi}{2n} - \cos \frac{2\pi}{n}\right)} = O(n^4)$$

for $\cos \frac{2\pi}{n} \leqq x \leqq \cos \frac{\pi}{n}$,

$$\left|\frac{T_n(x)}{(x - x_1^*)(x - x_2)}\right| \leqq \frac{1}{\cos \frac{\pi}{2n} - \cos \frac{\pi}{n}} \cdot \max_{-1 \leqq x \leqq +1} \left|\frac{T_n(x)}{x - x_2}\right| = O(n^4)$$

and similarly for $\cos \frac{\pi}{n} \leqq x \leqq +1$, Lemma IV follows from (12. 7) indeed.

In this proof is at the same time a proof for (12. 1) contained. We may write namely (12. 6) in the form

$$l_1(x, A_0) = -\frac{x_2 - x_1^*}{x_2 - x_1} \frac{T_n'(x_1^*)}{T_n'(x_2)} l_1(x, T) + O(1) = \frac{x_2 - x_1^*}{x_2 - x_1} \cdot \frac{\sin \frac{3\pi}{2n}}{\sin \frac{\pi}{2n}} l_1(x, T) + O(1).$$

Taking in account (12. 5), further

$$\frac{\sin \frac{3\pi}{2n}}{\sin \frac{\pi}{2n}} = O(1)$$

and

$$\frac{x_2 - x_1^*}{x_2 - x_1} = O(n^\beta),$$

the relation $l_1(x, A_0) = O(n^\beta)$ follows indeed. For $l_2(x, A_0)$ the assertion follows using Lemma IV. Thus our matrix belongs to $A(\beta)$ indeed.

13. In order to finish the proof of the assertion c) in **5** we have to show that the polynomials $L_n(f, A_0)$ converge uniformly in $[-1, +1]$ to $f(x)$ whenever $f \in \operatorname{Lip}\gamma$, $\gamma > \frac{\beta}{\beta+2}$. Let η be so small that

$$\gamma > \frac{\beta}{\beta+2} + \eta \tag{13. 1}$$

and fixed. Let further

$$\left[\frac{n}{\log n}\right] = m \tag{13. 2}$$

and $P_m(x)$ should stand for the best-approximating polynomial of $f(x)$ in $[-1, +1]$ in CHEBYSEV's sense. Then we have again

$$L_n(P_m, A_0) \equiv P_m(x)$$

and proceeding exactly as in **9** we obtain, using the abbreviation

$$f(x_\nu) - P_m(x_\nu) = y_\nu, \tag{13. 3}$$

that

$$L_n(f, A_0) - f = (P_m - f) + \sum_1^n y_\nu l_\nu(x, A_0). \tag{13. 4}$$

Using again S. BERNSTEIN's theorem mentioned in **9** we have

$$|f - P_m| \leqq c_{20} m^{-\gamma} < c_{21} n^{-\gamma} \log^\gamma n, \tag{13. 5}$$

i. e. from (13. 4) we get

$$|L_n(f, A_0) - f| \leqq c_{21} n^{-\gamma} \log^\gamma n + |y_1 l_1(x, A_0) + y_2 l_2(x, A_0)| + \sum_{\nu=3}^n |y_\nu| |l_\nu(x, A_0)|. \tag{13. 6}$$

Taking in account (13. 5) and Lemma III the last sum is $O(m^{-\gamma} \log n) = O(n^{-\gamma} \log^{\gamma+1} n)$; hence from (13. 6) we have

$$\begin{aligned} &|L_n(f, A_0) - f| \leqq \\ &\leqq c_{21} n^{-\gamma} \log^{\gamma+1} n + |y_1| |l_1(x, A_0) + l_2(x, A_0)| + |y_1 - y_2| |l_2(x, A_0)|. \end{aligned} \tag{13. 7}$$

Taking in account Lemma IV and (12. 1), (13. 7) gives

$$|L_n(f, A_0) - f| \leqq c_{22} \{ n^{-\gamma} \log^{\gamma+1} n + n^\beta |y_1 - y_2| \}. \tag{13. 8}$$

For the estimation of $|y_1 - y_2|$ we write

$$|y_1 - y_2| \leqq |f(x_1) - f(x_2)| + |P_m(x_1) - P_m(x_2)|. \tag{13. 9}$$

Since $f \in \operatorname{Lip}\gamma$ and

$$|x_1 - x_2| = \cos\left(\frac{3\pi}{2n} - \frac{1}{n^{1+\beta}}\right) - \cos\frac{3\pi}{2n} < \frac{100}{n^{2+\beta}}, \tag{13. 10}$$

we have, using (13. 1),

$$|f(x_1) - f(x_2)| \leqq c_{23}\left(\frac{100}{n^{2+\beta}}\right)^\gamma \leqq c_{24} n^{-(2+\beta)\left(\frac{\beta}{\beta+2}+\eta\right)} = c_{24} n^{-\beta} n^{-(\beta+2)\eta}. \tag{13. 11}$$

Further, using again MARKOV's theorem and (13. 10),

$$|P_m(x_1)-P_m(x_2)| \leqq \frac{100}{n^{2+\beta}} \max_{-1\leqq x\leqq+1} |P'_m(x)| \leqq \frac{100m^2}{n^{2+\beta}} \max_{-1\leqq x\leqq+1} |P_m(x)| <$$

$$< \frac{100}{n^\beta \log^2 n}(1+\max_{-1\leqq x\leqq+1}|f(x)|).$$

From this, (13. 8) and (13. 11) it follows that the polynomials $L_n(f, A_0)$ converge in $[-1, +1]$ uniformly to $f(x)$, i. e. the proof of the assertion c) is finished.

14. Finally we shall prove the assertion d) in **5.** Since the proof follows the line of that of a) we shall not go into all details. It is obviously sufficient to define A_1 only for even $n=2k$. With our β we define

$$\vartheta = \frac{\beta^2}{2}\left(\frac{\log k}{k}\right)^2 \tag{14. 1}$$

and the $2k^{th}$ line of our matrix A_1 should be given by the zeros $\zeta_1, \ldots, \zeta_{2k}$ of

$$\omega_{2k}(x, A_1) \equiv \omega_{2k}(x) \equiv T_k(1+\vartheta-(2+\vartheta)x^2) = 0 \quad (T_k(\cos\vartheta) = \cos k\vartheta), \tag{14. 2}$$

i. e. with $x_\nu = \cos\dfrac{2\nu-1}{2k}\pi$

$$\zeta_\nu = \sqrt{\frac{1+\vartheta-x_\nu}{2+\vartheta}} \qquad (\nu = 1, 2, \ldots, k),$$

$$\zeta_{-\nu} = -\sqrt{\frac{1+\vartheta-x_\nu}{2+\vartheta}} \qquad (\nu = 1, 2, \ldots, k). \tag{14. 3}$$

We have obviously

$$\left(\frac{\sqrt{\vartheta}}{2} <\right) \sqrt{\frac{\vartheta}{2+\vartheta}} < \zeta_1 < \zeta_2 < \cdots < \zeta_k < 1. \tag{14. 4}$$

Since for $1 \leqq |\nu| \leqq k$

$$|\omega'_{2k}(\zeta_\nu)| = \left|\frac{dT_k(u)}{du}\right|_{u=1+\vartheta-(2+\vartheta)\zeta_\nu^2} \cdot 2(2+\vartheta)|\zeta_\nu| = \frac{k}{\sqrt{1-x_{|\nu|}^2}} 2(2+\vartheta)|\zeta_\nu|,$$

we get — denoting the fundamental functions by $l_{\pm\nu}(x, A_1)$ $(\nu = 1, 2, \ldots, k)$ —

$$\lambda_{2k}(x, A_1) = \sum_{1\leqq|\nu|\leqq k} |l_\nu(x, A_1)| = \frac{1}{2(2+\vartheta)k} \sum_{1\leqq|\nu|\leqq k} \frac{\sqrt{1-x_{|\nu|}^2}}{|\zeta_\nu|} \frac{|\omega_{2k}(x)|}{|x-\zeta_\nu|}. \tag{14. 5}$$

15. First we are going to show (by rough estimations)

$$\sum_{|\nu|=1}^{k} |l_\nu(x, A_1)| < c_{25}(\beta)\, k^\beta \log^2 k. \tag{15. 1}$$

We may obviously suppose $0 \leqq x \leqq 1$. Since for all of our ν's we have

$$\frac{\sqrt{1-x_\nu}}{|\zeta_\nu|} = \sqrt{2+\vartheta}\sqrt{\frac{1-x_\nu}{1+\vartheta-x_\nu}} < 2$$

and

$$\sqrt{1+x_\nu} < 10\frac{k-\nu+1}{k},$$

we have

$$M_{2k}(A_1) < \frac{4}{k^2}\sum_{1\leqq|\nu|\leqq k}\frac{k-|\nu|+1}{|x-\zeta_\nu|}|\omega_{2k}(x)|. \tag{15.2}$$

What can be said on $(\zeta_{\nu+1}-\zeta_\nu)$? For

$$k \geqq |\nu| \geqq \frac{k}{2}$$

we have from (14.3)

$$|\zeta_{\nu+1}-\zeta_\nu| = \frac{1}{\sqrt{2+\vartheta}}\cdot\frac{x_\nu - x_{\nu+1}}{\sqrt{1+\vartheta-x_\nu}+\sqrt{1+\vartheta-x_{\nu+1}}} > \frac{k-|\nu|+1}{20k^2}. \tag{15.3}$$

For

$$\frac{k}{2} > |\nu| \geqq 1$$

we have from (15.3)

$$|\zeta_{\nu+1}-\zeta_\nu| > c_{26}(\beta)\frac{\frac{\nu}{k^2}}{\sqrt{\frac{\nu^2}{k^2}+\frac{\log^2 k}{k^2}}} > \frac{c_{27}(\beta)}{k\log k}. \tag{15.4}$$

Let first be

$$0 \leqq x \leqq \zeta_1.$$

Then we have from (15.2)

$$\lambda_{2k}(x, A_1) < \frac{8}{k^2}|\omega_{2k}(x)|\sum_{\nu=1}^{k}\frac{k-\nu+1}{\zeta_\nu - x}. \tag{15.5}$$

Owing to (15.4) we have

$$\sum_{\nu=1}^{k}\frac{k-\nu+1}{\zeta_\nu-x} \leqq \frac{k}{\zeta_1-x} + k\sum_{2\leqq\nu<\frac{k}{2}}\frac{1}{\zeta_\nu-x} + k\sum_{\frac{k}{2}\leqq\nu\leqq k}\frac{1}{\zeta_\nu-x} <$$

$$< \frac{k}{\zeta_1-x} + c_{28}k^2 + k\sum_{2\leqq\nu<\frac{k}{2}}\frac{1}{(\nu-1)c_{27}(\beta)\frac{1}{k\log k}} < \frac{k}{\zeta_1-x} + c_{29}(\beta)k^2\log^2 k,$$

i. e. from (15.5) for such values x

$$\lambda_{2k}(x, A_1) < c_{30}(\beta)\left\{\frac{1}{k}\left|\frac{\omega_{2k}(x)}{x-\zeta_1}\right| + |\omega_{2k}(x)|\log^2 k\right\}. \tag{15.6}$$

But for $0 \leqq x \leqq \xi_1$ owing to (14.1) we have for $k > c_{31}$

$$|\omega_{2k}(x)| \leqq T_k(1+\vartheta) < \frac{1}{2}\{1+(1+\vartheta+\sqrt{(1+\vartheta)^2-1})^k\} <$$

$$< (1+\vartheta+\sqrt{2\vartheta+\vartheta^2})^k < e^{k\left(\vartheta+\sqrt{2\vartheta}\sqrt{1+\frac{\vartheta}{2}}\right)} < 2e^{k\sqrt{2\vartheta}} = 2k^\beta. \tag{15.7}$$

Further, taking in account that for $0 \leqq x \leqq \zeta_1$

$$\left|\frac{d\omega_{2k}(x)}{dx}\right| = 2(2+\vartheta)|x|\left|\frac{dT_k(u)}{du}\right|_{u=1+\vartheta-(2+\vartheta)x^2} =$$

$$= k(2+\vartheta)|x|\left\{(u+\sqrt{u^2-1})^{k-1}\left(1+\frac{u}{\sqrt{u^2-1}}\right)+\right.$$

$$\left.+(u-\sqrt{u^2-1})^{k-1}\left(1-\frac{u}{\sqrt{u^2-1}}\right)\right\}_{u=1+\vartheta-(2+\vartheta)x^2} =$$

$$= k(2+\vartheta)|x|\left\{\frac{(u+\sqrt{u^2-1})^k-(u-\sqrt{u^2-1})^k}{\sqrt{u^2-1}}\right\}_{u=1+\vartheta-(2+\vartheta)x^2} \leqq$$

$$\leqq k(2+\vartheta)\zeta_1\left\{\quad\right\}_{u=1+\vartheta} < \frac{k(2+\vartheta)}{\sqrt{2\vartheta}}\sqrt{\frac{2\sin^2\frac{\pi}{4k}+\vartheta}{2}}\cdot 2k^\beta < 10k^{\beta+1},$$

we have for $0 \leqq x \leqq \zeta_1$

$$\frac{1}{k}\left|\frac{\omega_{2k}(x)}{x-\zeta_1}\right| \leqq \frac{1}{k}\max_{0\leqq x\leqq\zeta_1}|\omega'_{2k}(x)| < 10k^\beta. \tag{15.8}$$

(15.6), (15.7) and (15.8) give at once (15.1) for $0 \leqq x \leqq \xi_1$. For $\zeta_1 \leqq x \leqq 1$ the proof of (15.1) runs on similar lines and also a much better estimation of $\lambda_{2k}(x, A_1)$ could have been given; we omit the details. Thus (15.1) is proved.

16. Next we show that

$$\lambda_{2k}(0, A_1) > c_{32}(\beta)\, k^\beta. \tag{16.1}$$

Since for $k > c_{33}(\beta)$

$$|\omega_{2k}(0)| = T_k(1+\vartheta) > \frac{1}{2}\,(1+\vartheta+\sqrt{2\vartheta+\vartheta^2})^k > \frac{1}{4}\,k^\beta,$$

we obtain from (14.5)

$$\lambda_{2k}(0, A_1) > \frac{k^{\beta-1}}{4(2+\vartheta)}\sum_{\nu=1}^{k}\frac{\sqrt{1-x_\nu^2}}{\zeta_\nu^2} > \frac{k^{\beta-1}}{4}\sum_{\frac{k}{10}\leqq\nu\leqq\frac{9}{10}k}\frac{\sqrt{1-x_\nu^2}}{1+\vartheta-x_\nu} > c_{33}(\beta)k^\beta.$$

This and (15.1) show that our matrix A_1 belongs indeed to $A(\beta)$.

17. What we actually proved, is a little more and this is what we shall need. We showed

$$\sum_{1\leqq|\nu|\leqq\frac{1}{2}k}|l_\nu(0, A_1)| > c_{34}(\beta)k^\beta,$$

i. e.

$$\sum_{1\leqq|\nu|\leqq\frac{1}{2}k}|l_\nu(\Theta_{2k}, A_1)| \equiv \max_{-\frac{1}{4}\leqq x\leqq\frac{1}{4}}\sum_{1\leqq|\nu|\leqq\frac{1}{2}k}|l_\nu(x, A_1)| \geqq c_{35}(\beta)k^\beta. \tag{17.1}$$

Having this, the construction of $f_3(x)$ runs as follows. Introducing the notation

$$y_{\nu,2k} = \begin{cases} \operatorname{sign} l_{\nu,2k}(\Theta_{2k}, A_1) & \text{for } 1 \leqq |\nu| \leqq \frac{k}{2}, \\ 0 & \text{for } \frac{k}{2} < |\nu| \leqq k, \end{cases} \tag{17.2}$$

we consider the function $\psi_{2k}(x)$ defined by the broken line with the vertices at the points

$$P_\nu^* = (\zeta_{\nu,2k},\ y_{\nu,2k}) \qquad (1 \leqq |\nu| \leqq k);$$

we use also here the convention of **7** to complete the graph of $\psi_{2k}(x)$ in the whole interval $[-1, +1]$ by affixing two horizontals at the ends. We have obviously

$$L_{2k}(\psi_{2k}, A_1)_{x=\Theta_{2k}} = \sum_{1 \leqq |\nu| \leqq \frac{k}{2}} |l_{\nu,2k}\Theta_{2k}, A_1)| > c_{35}(\beta) k^\beta. \tag{17.3}$$

According to (15.4) the slopes of $\psi_{2k}(x)$ are absolutely

$$\leqq \frac{1}{c_{27}(\beta)} k \log k,$$

for $-\frac{51}{100} \leqq x \leqq \frac{51}{100}$ and $k \geqq c_{36}(\beta)$ we have

$$|\psi_{2k}(x)| \leqq 1, \tag{17.4}$$

for $\frac{51}{100} \leqq |x| \leqq 1$ (x real)

$$\psi_{2k}(x) = 0$$

and for $-1 \leqq x < x+h \leqq +1$

$$|\psi_{2k}(x+h) - \psi_{2k}(x)| \leqq c_{37}(\beta) h k \log k. \tag{17.5}$$

Now our function $f_3(x)$ will be of the form

$$f_3(x) = \sum_{\nu=1}^{\infty} \frac{1}{r_\nu^{\beta-\varepsilon}} \psi_{2r_\nu}(x)$$

with sufficiently quickly increasing r_ν-indices, as in **7.** The proof of unboundedness of $L_n(f_3, A_1)$ needs only slight changes compared to that of **8,** so we can omit the details. As to the Lipschitz exponent of $f_3(x)$ we have (p to be disposed later)

$$|f_3(x+h) - f_3(x)| \leqq \sum_{\nu \leqq p} \frac{1}{r_\nu^{\beta-\varepsilon}} |\psi_{2r_\nu}(x+h) - \psi_{2r_\nu}(x)| +$$

$$+ \sum_{\nu \geqq p+1} \frac{1}{r_\nu^{\beta-\varepsilon}} (|\psi_{2r_\nu}(x+h)| + |\psi_{2r_\nu}(x)|)$$

or owing to (17.5) and (17.4) "essentially" (owing to the quick increase of the r_ν's)

$$|f_3(x+h)-f_3(x)| \leqq c_{38}(\beta)\left\{ h r_p^{1-\beta+2\varepsilon} + \frac{1}{r_{p+1}^{\beta-\varepsilon}} \right\}.$$

Choosing p so that

$$r_p \leqq \frac{1}{h} < r_{p+1},$$

everything follows since ε is arbitrary small. Hence assertion d) is also proved.

(Received 22 January 1955)

РОЛЬ ФУНКЦИЙ ЛЕБЕГА В ТЕОРИИ ИНТЕРПОЛЯЦИИ ЛАГРАНЖА

П. Ердёш (Иерусалим) и П. Туран (Будапешт)

(Резюме)

Главным результатом настоящей работы является доказательство того, что в теории сходимости и расходимости интерполяции Лагранжа следует различать между „грубой" и „тонкой" теориями. „Грубой" является часть теории, зависящая только от быстроты возрастания „лебеговых констант"

$$M_n(A) = \max_{-1 \leqq x \leqq +1} \sum_{\nu=1}^{n} |l_\nu(x)|,$$

где A означает основную матрицу интерполяции, „тонкие" же другие результаты. Главная задача состояла в отделении этих двух теорий. Из различных возможных точек зрения мы здесь остановимся на той, при которой матрицы A для произвольно малого положительного числа ε, и для фиксированного $0 < \beta < 1$ подчинены условиям

$$\overline{\lim_{n\to\infty}} \frac{M_n(A)}{n^{\beta+\varepsilon}} < c_1(\varepsilon) \; (< \infty), \tag{1}$$

$$\underline{\lim_{n\to\infty}} \frac{M_n(A)}{n^{\beta-\varepsilon}} > c_2(\varepsilon) \; (> 0), \tag{2}$$

а следует определять те классы функций $\mathrm{Lip}\,\alpha$ в $[-1, +1]$, для которых при любой матрице удовлетворяющей условиям (1) и (2) интерполяция пригодна, и те классы, для которых интерполяция при любых таких матрицах непригодна. Оказывалось, что если исходить из классы матриц удовлетворяющих (1) и (2), то „тонкая" теория есть теория сходимости-расходимости тех класс $\mathrm{Lip}\,\alpha$, для которых

$$\frac{\beta}{\beta+2} < \alpha < \beta.$$

Более точно, имеет место следующее:

а) Если для A справедливы (1) и (2) и имеет место $\alpha < \frac{\beta}{\beta+2}$, то существует функция $f_1 \in \mathrm{Lip}\,\alpha$, для которой взятые по A интерполяционные полиномы Лагранжа $L_n(f_1, A)$ неограничены на $[-1, +1]$. (Интерполяция непригодна.)

b) Если A удовлетворяет условиям (1) и (2) и имеет место $\alpha > \beta$, то для любой $f \in \mathrm{Lip}\,\alpha$ интерполяционные полиномы Лагранжа $L_n(f, A)$ соответствующие A стремятся к f равномерно на $[-1, +1]$. (Интерполяция пригодна.)

c) Если $\alpha > \frac{\beta}{\beta+2}$, то уже существует матрица A_0, удовлетворяющая (1) и (2), для которой интерполяционные полиномы $L_n(f, A_0)$ равномерно сходятся к f, если только $f \in \mathrm{Lip}\,\alpha$.

d) Если $\alpha < \beta$, то уже существуют матрица A_1 удовлетворяющая (1) и (2), и функция $f_2 \in \mathrm{Lip}\,\alpha$ такие, что интерполяционные полиномы Лагранжа $L_n(f_2, A_1)$ неограничены в интервале $[-1, +1]$.

Аналогичные вопросы возникают также для разложений в ортогональный ряд, для механических квадратур и для других линейных операций.

NOTES ON INTERPOLATION. I.
(ON SOME INTERPOLATORICAL PROPERTIES OF THE ULTRASPHERICAL POLYNOMIALS)

By

J. SURÁNYI (Budapest) and P. TURÁN (Budapest), member of the Academy

To G. SZEGŐ *on his 60th birthday*

1. Let be given n points $x_1, \ldots, x_n$ with

$$(1.1) \qquad -1 \leqq x_n < x_{n-1} < \cdots < x_1 \leqq 1.$$

It is a classical fact that there is exactly one polynomial $f(x)$ of degree $\leqq n-1$ such that $f(x)$ assumes at the place x_ν a prescribed value y_ν. The explicite form of this polynomial $L_n(f)$ was given by NEWTON and LAGRANGE. The investigation of the convergence-properties of these polynomials, which began at the end of the last century by RUNGE and BOREL, revealed a curious and unexpected situation with the results of G. FABER and S. BERNSTEIN. We consider the system A of points

$$(1.2) \qquad -1 \leqq x_{nn} < x_{n-1,\,n} < \cdots < x_{2n} < x_{1n} \leqq 1 \qquad (n = 1, 2, \ldots)$$

and denote by C the class of functions $f(x)$ continuous for $-1 \leqq x \leqq 1$. Then the above named authors proved that in whatever way we choose the system A, we can find an $f_0(x) \in C$ such that the polynomials $L_n(f_0)$ belonging to A are unbounded in $[-1, +1]$.[1]

2. As FEJÉR[2] showed, the situation changes when we turn from the interpolatory polynomials $L_n(f)$ to another sort of interpolatory polynomials. Already HERMITE[3] showed, that prescribing beside the points (1.1) the non-negative integers $m_1, \ldots, m_n$ and the real or complex numbers $y_{\mu\nu}$, there is a uniquely determined polynomial $f(x)$ of degree

$$(2.1) \qquad \leqq m_1 + \cdots + m_n - 1$$

[1] For this theory see e. g. И. П. Натансон, Конструктивная теория функций (Москва—Ленинград, 1949).

[2] L. FEJÉR, Über Interpolation, *Gött. Nachr.*, (1916), pp. 66—91.

[3] CH. HERMITE, Sur la formule d'interpolation de Lagrange, *Journ. für Math.*, **84** (1878), pp. 70—79.

such that

$$f^{(\nu)}(x_\mu) = y_{\mu\nu} \qquad (\mu = 1, 2, \ldots, n\ ;\ \nu = 0, 1, 2, \ldots, m_\mu - 1). \tag{2.2}$$

In the case

$$m_1 = m_2 = \cdots = m_n = 2 \tag{2.3}$$

FEJÉR gave a particularly simple form of these polynomials which enabled him in certain cases to perform the convergence-proofs in a very elegant manner. The new feature of the results is that for certain A-systems, choosing

$$y_{\mu 1} = 0 \qquad (\mu = 1, 2, \ldots, n),$$

the $H_n(f)$ Hermite—Fejér interpolatory polynomials of degree $\leqq 2n-1$ converge uniformly to $f(x)$ if only $f \in C$. In the more general cases of Hermite interpolation results are rather scarce, partly owing to the lack of suitably simple form of the interpolatory polynomials.

3. In all previously mentioned problems we prescribed at each point x_μ the value of the function and of some *consecutive* derivatives. In the most general case we consider a polynomial $f_m(x)$ of degree $\leqq (m_1 + \cdots + m_n - 1) = m$ such that for each x_μ we prescribe the value of some m_μ derivatives with given indices *(not necessarily consecutive ones!)*. This general form of interpolation was treated as far as we know only in a paper of GEORGE BIRKHOFF.[4] His point of view — as the title of the paper shows — was so general that one cannot expect better formulae than those of HERMITE. One cannot even see from this paper the new feature of this general interpolation, viz. *for each* n by a suitable choice of the x_ν's and the indices of the prescribed derivatives the problem can be unsolvable or can have an infinity of solutions. On this account, for the special case $n = 2$ a direct treatment was given by PÓLYA.[5] Hence the questions concerning this most general form of interpolation are roughly speaking the following ones:

1) Does there exist such an $f_m(x)$ at all? (Problem of existence.)

2) If there is such an $f_m(x)$, is it uniquely determined? (Problem of uniqueness.)

3) If $f_m(x)$ exists and is unique, how can it be represented in the most suitable form? (Problem of representation.)

4) Fixing for each n the indices of the prescribed derivatives we choose the points (1.2), i.e. the system A so that $f_m(x)$ should be uniquely determined, if the values of the prescribed derivatives are all 0 (if possible). Form-

[4] G. D. BIRKHOFF, General mean value and remainder theorems with applications to mechanical differentiation and quadrature, *Trans. Amer. Math. Soc.*, **7** (1906), pp. 107—136.

[5] G. PÓLYA, Bemerkung zur Interpolation und zur Näherungstheorie der Balkenbiegung, *Zeitschr. für angew. Math. und Mech.*, **11** (1931), pp. 445—449.

ing for an $f(x) \in C$ and each n these interpolatory polynomials, does the sequence converge uniformly in $[-1,+1]$ to $f(x)$ or not? (Problem of convergence.)

4. In this sequence of papers we shall try to get orientation concerning this general sort of interpolation by considering the simplest case only, when $m=2n-1$ and we are seeking a polynomial $f_{2n-1}(x)$ of degree $\leqq 2n-1$ where the values of the function and the *second* derivatives are prescribed at the points (1. 1), i. e.

$$\left.\begin{aligned} f_{2n-1}(x_\nu) &= y_\nu = \text{given}, \\ f''_{2n-1}(x_\nu) &= y^*_\nu = \text{given} \end{aligned}\right\} \qquad (\nu = 1, 2, \ldots, n). \tag{4. 1}$$

The first question is how to choose the points (1. 1) in order to get the simplest proofs. It seems to us now that the best choice from this point of view are the zeros of the polynomial

$$\Pi_n(x) = (1-x^2)P'_{n-1}(x) \tag{4. 2}$$

where $P_{n-1}(x)$ stands for the $(n-1)^{\text{th}}$ Legendre polynomial with the normalisation

$$P_{n-1}(1) = 1. \tag{4. 3}$$

Denoting them by $\xi_{\mu\nu}$ we have[6]

$$-1 = \xi_{n,n} < \xi_{n-1,n} < \cdots < \xi_{2n} < \xi_{1n} = 1. \tag{4. 4}$$

Forming these points for $n=2, 3, \ldots$ and denoting this by B-system, in the second paper we shall consider the convergence behaviour of the corresponding interpolatory polynomials (4. 1) when $f(x) \in C$

$$y_\nu = f(\xi_{\nu n}), \quad y^*_\nu = 0 \qquad (\nu = 1, 2, \ldots, n;\ n = 2, 3, \ldots). \tag{4. 5}$$

Here, in this first paper we shall consider the questions of existence and uniqueness only, but not exclusively for the B-matrix; already this will show how much simpler the proofs are in the case of the B-matrix.

5. Let

$$n = 2k+1 \tag{5. 1}$$

and

$$1 \geqq x_1 > x_2 > \cdots > x_k > x_{k+1} = 0 > x_{k+2} > \cdots > x_{2k+1} \geqq -1 \tag{5. 2}$$

with

$$x_j = -x_{2k+2-j} \qquad (j = k+2, \ldots, 2k+1). \tag{5. 3}$$

We shall show the

THEOREM I. *If $n=2k+1$ and the points $x_1, \ldots, x_n$ satisfy* (5. 2) *and* (5. 3), *there is in general no polynomial $f(x)$ of degree $\leqq 2n-1$ such that for*

[6] In this paper n is fixed; thus the zeros of $\Pi_n(x)$ will be denoted simply by ξ_ν in the sequel.

given y_ν and y_ν^*

(5. 4) $$f(x_\nu)=y_\nu,\quad f''(x_\nu)=y_\nu^* \qquad (\nu=1,2,\ldots,n).$$

If there exists such a polynomial, then there is an infinity of them.

Hence in the case of an odd number of distinct symmetrical points x_ν both the problem of existence and uniqueness have a negative solution. This means that in this case there is a universal linear relation between the values $\pi_{2n-1}(x_\nu)$ and $\pi''_{2n-1}(x_\nu)$ $(\nu=1,\ldots,n)$ of an arbitrary polynomial $\pi_{2n-1}(x)$ of degree $\leqq 2n-1$, the coefficients being *independent* of $\pi_{2n-1}(x)$. The proof will show that the same state of affairs holds if the second derivative is replaced by any derivative of fixed even order for $\nu=1,2,\ldots,n$.

6. The simple proof of Theorem I shall be given in **10.** The possibility of an infinity of solutions raises the question on the dimensionality of the number of them. Taking as points x_ν those defined in (4. 4) we shall show directly in **11** *that in the case of infinitely many solutions in Theorem I for $n\geqq 3$ the general form of the solutions is*

(6. 1) $$f(x)=f_0(x)+cf_1(x),$$

where $f_0(x)$ and $f_1(x)$ are fixed polynomials of degree $\leqq 2n-1$ and c an arbitrary complex number.

7. In the case $n=2k$ the situation changes. In **12** we shall show directly[7] the

THEOREM II. *If $n=2k$, then to prescribed values y_ν and y_ν^* there is a uniquely determined polynomial $f(x)$ of degree $\leqq 2n-1$ such that*

(7. 1) $$f(\xi_\nu)=y_\nu,\quad f''(\xi_\nu)=y_\nu^* \qquad (\nu=1,2,\ldots,n),$$

if ξ_ν stands, as in (4. 4), *for the zeros of $H_n(x)$.*

This means, of course, that in the case

(7. 2) $$y_\nu=y_\nu^*=0 \qquad (\nu=1,2,\ldots,n;\ n \text{ even})$$

the only solution of (7. 1) is $f(x)\equiv 0$. This result is curious for a certain reason. It is namely well known that $H_n(x)$ satisfies the differential-equation (for each integer $n\geqq 2$)

(7. 3) $$(1-x^2)H_n''(x)+n(n-1)H_n(x)=0.$$

Thus for $n\geqq 4$

$$H_n''(\xi_\nu)=0 \qquad (\nu=2,3,\ldots,(n-2),(n-1)).$$

Hence there is a non-trivial polynomial, even of nth degree, satisfying "almost" the requirements (7. 2), i. e. except the two conditions

(7. 5) $$f''(\xi_1)=f''(\xi_n)=0.$$

[7] A verification of the further results of this paper in the frame of BIRKHOFF's approach seems to us to be a very difficult task.

Superficially thinking one would incline to conclude that there is a non-trivial polynomial of degree $(n+2)$ satisfying *all* conditions (7. 2). The Theorem II and the result in **11** show that this is not true; it follows that for an even n no polynomial of degree $\leqq 2n-1$ exists with the property (7. 2) except $f \equiv 0$.

8. In **13** we shall investigate the existence- and uniqueness questions of the interpolation problem (4. 1) when the points x_ν are the zeros $\eta_{\nu n}^{(\lambda)}$ of the nth ultraspherical polynomial $P_n^{(\lambda)}(x)$, to the knowledge of which G. SZEGŐ contributed very much. Since λ and n are fixed we shall denote the zeros in question simply by η_{ν} $(\nu = 1, 2, \ldots, n)$ and instead $P_n^{(\lambda)}(x)$ we shall write $Q_n(x)$ only. We consider only the case $\lambda \geqq -\frac{1}{2}$; since in the case $\lambda = -\frac{1}{2}$ $Q_n(x)$ is identical with $H_n(x)$, we may suppose

$$\lambda > -\frac{1}{2}. \tag{8. 1}$$

The polynomial $Q_n(x)$ satisfies for all non-negative integer n's the equation

$$(1-x^2)Q_n''(x) - (2\lambda+1)xQ_n'(x) + n(n+2\lambda)Q_n(x) = 0; \tag{8. 2}$$

it is uniquely determined in the case $\lambda \neq 0$ by the normalisation

$$Q_n(1) = \binom{n+2\lambda-1}{n}, \tag{8. 3}$$

and for $\lambda = 0$ by

$$Q_n(1) = 1. \tag{8. 4}$$

It is well known that all the zeros are real, simple and lying in $-1 < x < 1$. Since for odd n these zeros fulfil (5. 2) and (5. 3), we have only to confine ourselves in this § to the case

$$n = 2k \geqq 4. \tag{8. 5}$$

We shall see that in this case the interpolation problem (4. 1) is generally uniquely determined. A particular behaviour is shown only in the case

$$\lambda + \frac{3}{2} = \text{odd integer} \qquad n \geqq \lambda + \frac{5}{2}, \tag{8. 6}$$

i. e. when the points x_ν are the zeros of some derivative of odd order of the Legendre polynomials. In this case neither existence nor uniqueness of the polynomials with (4. 1) holds good. More particular is the behaviour of these polynomials, if in (4. 1) we have

$$y_\nu = y_\nu^* = 0 \qquad (\nu = 1, 2, \ldots, n). \tag{8. 7}$$

In this case generally $f(x) \equiv 0$ is the only polynomial of degree $\leqq 2n-1$; but in the case (8. 6) there is a uniquely determined non-trivial polynomial

$f_0(x)$ of degree $\left(n+\lambda+\frac{1}{2}\right)$ such that for all polynomials $f_1(x)=cf_0(x)$ and only for these

$$f_1(\eta_\nu)=f_1''(\eta_\nu)=0 \qquad (\nu=1,2,\ldots,n) \tag{8.8}$$

holds with any numerical c. Thus we shall show the

THEOREM III. *The ultraspherical polynomials* $P_n^{(\lambda)}(x)$, *with* $\lambda+\frac{1}{2}=$ *even integer* >0 *and* $n\geqq\lambda+\frac{5}{2}$, *with the zeros* $\eta_1,\ldots,\eta_n$, *have the property that there exist non-trivial polynomials* $f(x)$ *of degree* $\left(n+\lambda+\frac{1}{2}\right)$ *such that* $\eta_1,\ldots,\eta_n$ *are zeros and at the same time places of inflexion of the polynomial* $f(x)$.

9. Before turning to the proofs we mention a corollary of Theorem I (similar conclusions could have been drawn from the other results of this paper, too). Let $n=2k+1$ and the numbers x_ν satisfy the condition (5.2)—(5.3); to a prescribed integrable $F(x)$ we consider the solution of the equation

$$f^{(2n)}(x)=F(x) \tag{9.1}$$

with the conditions

$$f(x_\nu)=y_\nu, \quad f''(x_\nu)=y_\nu^* \qquad (\nu=1,2,\ldots,n) \tag{9.2}$$

where y_ν and y_ν^* are prescribed numbers. If $g_0(x)$ is any solution of (9.1) (without requiring (9.2)), then $g(x)=f(x)-g_0(x)$ satisfies $g^{(2n)}(x)=0$, i. e.

$$f(x)=g_0(x)+\pi_{2n-1}(x)$$

where $\pi_{2n-1}(x)$ is a polynomial of degree $\leqq 2n-1$ such that

$$\pi_{2n-1}(x_\nu)=y_\nu-g_0(x_\nu)=\text{prescribed},$$
$$\pi_{2n-1}''(x_\nu)=y_\nu^*-g_0''(x_\nu)=\text{prescribed}.$$

But owing to Theorem I this is generally unsolvable. Hence we obtained the

COROLLARY.[8] *If* n *is odd and the points* x_ν *are real and symmetrical to* $x=0$, *then the equation* (9.1) *is generally unsolvable under the restrictions* (9.2).

10. Now we turn to the proof of Theorem I. We decompose the polynomial

$$f(x)=\sum_{\nu=0}^{2n-1} c_\nu x^\nu$$

[8] A corollary of similar character is mentioned in BIRKHOFF's quoted paper, p. 114.

into an even and an odd part:

$$s(x)=\sum_{\nu=0}^{n-1} c_{2\nu}x^{2\nu}, \qquad t(x)=\sum_{\nu=0}^{n-1} c_{2\nu+1}x^{2\nu+1}.$$

The first part of the conditions (5. 4) will then give

$$\begin{aligned} f(x_\nu)&=s(x_\nu)+t(x_\nu)=y_\nu,\\ f(-x_\nu)&=s(x_\nu)-t(x_\nu)=y_{n-\nu+1} \end{aligned} \qquad \left(\nu=1,2,\ldots,\left[\frac{n}{2}\right]\right). \tag{10. 1}$$

Since $s''(x)$ and $t''(x)$ are also even and odd polynomial, respectively, from the second part of the conditions follows

$$\begin{aligned} f''(x_\nu)&=s''(x_\nu)+t''(x_\nu)=y_\nu^*,\\ f''(-x_\nu)&=s''(x_\nu)-t''(x_\nu)=y_{n-\nu+1}^* \end{aligned} \quad \left(\nu=1,2,\ldots,\left[\frac{n}{2}\right]\right). \tag{10. 2}$$

From (10. 1), (10. 2) we get for $t(x)$

$$t(x_\nu)=\frac{y_\nu-y_{n-\nu+1}}{2}=u_\nu,\ t''(x_\nu)=\frac{y_\nu^*-y_{n-\nu+1}^*}{2}=u_\nu^* \left(\nu=1,2,\ldots,\left[\frac{n}{2}\right]\right), \tag{10. 3}$$

and of course

$$t(0)=t''(0)=0. \tag{10. 4}$$

Since the last conditions are automatically fulfilled, in the case of an odd n we have $n-1$ linear equations for the n coefficients of $t(x)$. Hence apart from special values x_ν the system (10. 3) has an infinity of solutions.

For $s(x)$ the equations (10. 1) and (10. 2) give

$$\begin{aligned} s(x_\nu)&=\frac{y_\nu+y_{n-\nu+1}}{2}=v_\nu,\\ s''(x_\nu)&=\frac{y_\nu^*+y_{n-\nu+1}^*}{2}=v_\nu^* \end{aligned} \qquad \left(\nu=1,2,\ldots,\left[\frac{n}{2}\right]\right). \tag{10. 5}$$

Since in the case of an odd n, 0 occurs among the x_ν's, owing to (10. 4), we have the additional conditions

$$s(0)=y_{\left[\frac{n}{2}\right]+1}, \qquad s''(0)=y_{\left[\frac{n}{2}\right]+1}^*.$$

The last two equations determine c_0 and c_2. For the remaining $n-2$ coefficients the system (10. 5) furnishes $n-1$ equations. This system becomes contradictory for appropriate values of the v_ν's and v_ν^*'s.

Since the u_ν's and v_ν's resp. u_ν^*'s and v_ν's also determine the values y_ν and y_ν^*, these results prove Theorem I.

11. As told in **6** we shall consider with an odd n in the case of $x_\nu=\xi_\nu$ (see (4. 4)) the general structure of the solutions. The assertion (6. 1) will be

obviously proved if we show that in our case *all* polynomials $f(x)$ of degree $\leqq 2n-1$ satisfying

$$f(\xi_\nu)=f''(\xi_\nu)=0 \qquad (\nu=1,2,\ldots,n) \tag{11.1}$$

are

$$c\Pi_n(x)(P_{n-1}(x)-3) \tag{11.2}$$

with arbitrary numerical c. If $f(x)$ satisfies (11.1), the first part of the conditions implies

$$f(x)=\Pi_n(x)q_{n-1}(x)$$

where $q_{n-1}(x)$ is a polynomial of degree $\leqq n-1$. Then the second part of the conditions means

$$\Pi_n''(\xi_\nu)q_{n-1}(\xi_\nu)+2\Pi_n'(\xi_\nu)q_{n-1}'(\xi_\nu)=0 \qquad (\nu=1,2,\ldots,(n-1),n). \tag{11.3}$$

But the equation (7.3) gives for $\nu=2,3,\ldots,(n-2),(n-1)$

$$\Pi_n''(\xi_\nu)=0,$$

i. e. owing to the simplicity of the ξ_ν's we have

$$q_{n-1}'(\xi_\nu)=0 \qquad (\nu=2,3,\ldots,(n-2),(n-1)).$$

But this means that $q_{n-1}'(x)$ has all its zeros common with

$$\frac{\Pi_n(x)}{1-x^2}=P_{n-1}'(x),$$

i. e.

$$q_{n-1}'(x)=cP_{n-1}'(x)$$

with a numerical c. Hence if $c\neq 0$

$$q_{n-1}(x)=c(P_{n-1}(x)+c_1)$$

with a numerical c_1, i. e.

$$f(x)=c\Pi_n(x)(P_{n-1}(x)+c_1). \tag{11.4}$$

To determine c and c_1, we use the requirements $f''(\pm 1)=0$. As well known and easy to verify

$$P_{n-1}'(1)=\frac{n(n-1)}{2}=(-1)^nP_{n-1}'(-1) \tag{11.5}$$

and thus, using the known identity

$$\Pi_n(x)=-n(n-1)\int_{-1}^{x}P_{n-1}(t)dt, \tag{11.6}$$

we have

$$\Pi_n'(1)=-n(n-1)=(-1)^{n-1}\Pi_n'(-1), \tag{11.7}$$

$$\Pi_n''(1)=-\frac{n^2(n-1)^2}{2}=(-1)^n\Pi_n''(-1). \tag{11.8}$$

Hence from $f''(1)=0$ we have

$$0=c\{\Pi_n''(1)(1+c_1)+2\Pi_n'(1)P_{n-1}'(1)\}$$

and using (11. 5), (11. 7) and (11. 8)

$$c_1=-3,$$

$$f(x)=c\Pi_n(x)(P_{n-1}(x)-3). \tag{11. 9}$$

Since n is now odd, it is easy to verify with (11. 5), (11. 7) and (11. 8) that the polynomial (11. 9) satisfies also the condition $f''(-1)=0$. If $c=0$, then $q_{n-1}(x)=\text{const.}$, i. e. $f(x)$ would be $c_2\Pi_n(x)$ with a numerical c_2; but then $f''(1)\neq 0$.[9] Hence *only* the polynomials (11. 9) fulfil our requirement and thus (6. 1) is proved.

12. The proof of Theorem II runs parallel to the reasoning of **11.** This proves also in this case, n being even, that if

$$f(\xi_\nu)=f''(\xi_\nu)=0, \tag{12. 1}$$

then $f(x)$ has necessarily the form

$$f(x)=c\Pi_n(x)(P_{n-1}(x)-3) \tag{12. 2}$$

with numerical c, without using the requirement $f''(-1)=0$. But since n is now even, this condition together with (11. 5), (11. 7) and (11. 8) assures that $c=0$, i. e.

$$f(x)\equiv 0. \tag{12. 3}$$

Now this means that writing out (12. 1) as a linear system, the determinant is $\neq 0$. Considering the general problem

$$f(\xi_\nu)=y_\nu,\quad f''(\xi_\nu)=y_\nu' \qquad (\nu=1,\ldots,n) \tag{12. 4}$$

this shows that the corresponding linear system is always uniquely soluble, i. e. Theorem II is proved.

13. Finally we turn, as told, to the case when the x_ν's are given by the numbers η_ν defined in **8.** Let first be

$$-\frac{1}{2}<\lambda\neq 0, \qquad n=2k. \tag{13. 1}$$

As we have seen previously, everything depends on determination of all polynomials $g(x)$ of degree $\leqq 2n-1$ for which

$$g(\eta_\nu)=g''(\eta_\nu)=0 \qquad (\nu=1,2,\ldots,n). \tag{13. 2}$$

From the first part of the conditions we get

$$g(x)=Q_n(x)r_{n-1}(x)$$

[9] This last sentence holds obviously also in the case of an even n.

with a suitable polynomial $r_{n-1}(x)$ of degree $\leqq n-1$. The second part gives

$$Q_n''(\eta_\nu) r_{n-1}(\eta_\nu) + 2 Q_n'(\eta_\nu) r_{n-1}'(\eta_\nu) = 0 \qquad (\nu = 1, 2, \ldots, n). \tag{13. 3}$$

But from (8. 2) we have

$$(1-\eta_\nu^2) Q_n''(\eta_\nu) - (2\lambda+1)\eta_\nu Q_n'(\eta_\nu) = 0;$$

i. e. from (13. 3), using also the fact that the zeros of $Q_n(x)$ are simple, i. e. $Q_n'(\eta_\nu) \neq 0$, we get

$$\left(\lambda + \frac{1}{2}\right)\eta_\nu r_{n-1}(\eta_\nu) + (1-\eta_\nu^2) r_{n-1}'(\eta_\nu) = 0. \tag{13. 4}$$

Since the polynomial

$$(1-x^2) r_{n-1}'(x) + \left(\lambda + \frac{1}{2}\right) x r_{n-1}(x)$$

is of degree $\leqq n$ and by (13. 4) all its zeros are the same as those of $Q_n(x)$, we obtain

$$(1-x^2) r_{n-1}'(x) + \left(\lambda + \frac{1}{2}\right) x r_{n-1}(x) = c\, Q_n(x) \tag{13. 5}$$

with numerical c.

14. We have to investigate whether or not the equation (13. 5) has a polynomial solution of degree $\leqq n-1$ (n even). It would be easy to give the solution in the form of an integral, but it would not be easy to decide on this form whether or not the solution is a polynomial. Instead of it, we shall use the idea to try to solve the equation (13. 5) by

$$r_{n-1}(x) = \sum_{\nu=0}^{n-1} c_\nu Q_\nu(x). \tag{14. 1}$$

We shall use the identities[10] for $\nu \geqq 1$

$$(1-x^2) Q_\nu'(x) =$$

$$= \frac{1}{2(\nu+\lambda)} \{(\nu+2\lambda-1)(\nu+2\lambda) Q_{\nu-1}(x) - \nu(\nu+1) Q_{\nu+1}(x)\}, \tag{14. 2}$$

$$x Q_\nu(x) = \frac{1}{2(\nu+\lambda)} \{(\nu+1) Q_{\nu+1}(x) + (\nu+2\lambda-1) Q_{\nu-1}(x)\}. \tag{14. 3}$$

Substituting (14. 1) into (13. 5) we obtain

$$c\, Q_n(x) = \sum_{\nu=1}^{n-1} c_\nu (1-x^2) Q_\nu'(x) + \sum_{\nu=0}^{n-1} c_\nu \left(\lambda + \frac{1}{2}\right) x Q_\nu(x)$$

[10] See G. SZEGŐ, *Orthogonal polynomials* (New York, 1939), pp. 83 and 82.

and using (14. 2), (14. 3)

$$c\,Q_n(x)=\sum_{\nu=1}^{n-1}\frac{c_\nu}{2(\nu+\lambda)}\{(\nu+2\lambda-1)(\nu+2\lambda)\,Q_{\nu-1}(x)-\nu(\nu+1)\,Q_{\nu+1}(x)\}+$$

$$+c_0\frac{\lambda+\frac{1}{2}}{2\lambda}\,Q_1(x)+\sum_{\nu=1}^{n-1}c_\nu\frac{\lambda+\frac{1}{2}}{2(\nu+\lambda)}\{(\nu+1)\,Q_{\nu+1}(x)+(\nu+2\lambda-1)\,Q_{\nu-1}(x)\}=$$

$$(14.\,4)\qquad =c_0\frac{2\lambda+1}{4\lambda}\,Q_1(x)+\sum_{\nu=1}^{n-1}\frac{c_\nu}{2(\nu+\lambda)}(\nu+1)\left(\lambda+\frac{1}{2}-\nu\right)Q_{\nu+1}(x)+$$

$$+\sum_{\nu=1}^{n-1}\frac{c_\nu}{2(\nu+\lambda)}(\nu+2\lambda-1)\left(\nu+3\lambda+\frac{1}{2}\right)Q_{\nu-1}(x)=c_0\frac{2\lambda+1}{4\lambda}\,Q_1(x)+$$

$$+\sum_{\nu=2}^{n}\frac{c_{\nu-1}}{2(\nu+\lambda-1)}\,\nu\left(\lambda+\frac{3}{2}-\nu\right)Q_\nu(x)+$$

$$+\sum_{\nu=0}^{n-2}\frac{c_{\nu+1}}{2(\nu+\lambda+1)}(\nu+2\lambda)\left(\nu+3\lambda+\frac{3}{2}\right)Q_\nu(x).$$

We have to compare the coefficients of $Q_\nu(x)$ in (14. 4). Comparing the coefficients of $Q_{n-1}(x)$ we find for $n\geqq 4$

$$0=\frac{c_{n-2}}{2(n+\lambda-2)}(n-1)\left(\lambda+\frac{5}{2}-n\right),$$

i. e. if $n\geqq 4$ and $n\neq\lambda+\frac{5}{2}$

$$(14.\,5)\qquad c_{n-2}=0.$$

Comparing the coefficients of $Q_0(x)$ in (14. 4) we obtain

$$0=\frac{c_1}{\lambda+1}\lambda\left(3\lambda+\frac{3}{2}\right),$$

i. e. since now $\lambda\neq 0$

$$(14.\,6)\qquad c_1=0.$$

If $n\geqq 4$ and $2\leqq\nu\leqq n-2$, the comparison of the coefficients of $Q_\nu(x)$ in (14. 4) gives

$$(14.\,7)\qquad 0=\frac{c_{\nu-1}}{\nu+\lambda-1}\,\nu\left(\lambda+\frac{3}{2}-\nu\right)+\frac{c_{\nu+1}}{\nu+\lambda+1}(\nu+2\lambda)\left(\nu+3\lambda+\frac{3}{2}\right)$$

and considering the coefficient of $Q_1(x)$ in (14. 4) we see that (14. 7) holds also for $\nu=1$; thus (14. 7) holds for

$$(14.\,8)\qquad n\geqq 4,\quad 1\leqq\nu\leqq n-2.$$

Evidently $c_{\nu+1}$ can be expressed always by $c_{\nu-1}$ from (14. 7); thus starting from (14. 6) we obtain

$$(14.\,9)\qquad c_1=c_3=c_5=\cdots=c_{n-1}=0.$$

Investigating $c_0, c_2, c_4, \ldots, c_{n-2}$ we consider first the case when λ is not only restricted by (13. 1) but also by

$$\lambda + \frac{3}{2} \neq \text{odd integer.} \tag{14. 10}$$

Then $c_{\nu-1}$ can be expressed for odd ν's by $c_{\nu+1}$ from (14. 7) and hence starting from (14. 5) we obtain

$$c_{n-2} = c_{n-4} = \cdots = c_2 = c_0 = 0,$$

i. e. $r_{n-1}(x) \equiv 0$. Since (14. 10) implies that $n \neq \lambda + \frac{5}{2}$, and the case $\lambda = 0$ can be settled with reference to the continuity in λ, we obtained that *if* $n \geqq 4$ *even and*

$$\lambda + \frac{3}{2} \neq \textit{odd integer}, \quad \lambda > -\frac{1}{2}, \tag{14. 11}$$

then we can always find a uniquely determined $g(x)$ *polynomial of degree* $\leqq 2n-1$ *so that*

$$g(\eta_{\nu n}^{(\lambda)}) = \textit{prescribed}, \quad g''(\eta_{\nu n}^{(\lambda)}) = \textit{prescribed} \qquad (\nu = 1, 2, \ldots, n). \tag{14. 12}$$

15. If

$$\lambda > -\frac{1}{2}, \quad \lambda + \frac{3}{2} = \text{odd integer } (\geqq 3), \quad n = \text{even} \geqq \lambda + \frac{9}{2} \ (\geqq 6), \tag{15. 1}$$

the situation changes, since $c_{\nu-1}$ can be expressed by $c_{\nu+1}$ from (14. 7) only for $\nu > \lambda + \frac{3}{2}$, i. e. only

$$c_{n-2} = c_{n-4} = \cdots = c_{\lambda+\frac{5}{2}} = 0 \tag{15. 2}$$

remains true and $c_{\lambda+\frac{1}{2}}$ is indetermined. Giving any value to $c_{\lambda+\frac{1}{2}}$ (14. 7) is again applicable with

$$\nu = 1, 2, \ldots, \lambda - \frac{1}{2},$$

i. e. the values of

$$c_0, c_2, c_4, \ldots, c_{\lambda-\frac{3}{2}}$$

are with $c_{\lambda+\frac{1}{2}} \neq 0$ uniquely determined and $\neq 0$. Hence in this case all $g(x)$ polynomials of degree $\leqq 2n-1$ satisfying (14. 12) are of the form

$$g(x) = ck(x) \tag{15. 3}$$

where $k(x)$ is a uniquely determined polynomial of the exact degree $\left(n+\lambda+\frac{1}{2}\right)$. If

$$\lambda > -\frac{1}{2}, \quad \lambda+\frac{3}{2} = \text{odd integer}, \quad n = \lambda+\frac{5}{2} \ (\geqq 4), \tag{15.4}$$

then (14.5) ceases to be true, but

$$c_1 = c_3 = \cdots = c_{\lambda+\frac{3}{2}} = 0$$

holds. c_{n-2} is now arbitrary, fixing its value, (14.7) is applicable again with $\nu = 1, 2, \ldots, \left(\lambda-\frac{1}{2}\right)$, i. e. again $c_0, c_2, \ldots, c_{n-4} = c_{\lambda-\frac{3}{2}}$ are uniquely determined. Hence (15.3) holds and the degree of $k(x)$ is $\lambda+\frac{1}{2}+n$ also in the case (15.4). If the restrictions are

$$\lambda > -\frac{1}{2}, \quad \lambda+\frac{3}{2} = \text{odd integer} \ (\geqq 3), \quad 4 \leqq n = \text{even} \leqq \lambda+\frac{1}{2},$$

then repeating the same argument, it follows again $g(x) \equiv 0$. Thus we obtained that *if*

$$\lambda+\frac{3}{2} = \textit{odd integer} \ (\geqq 3), \tag{15.5}$$

then in the case $4 \leqq n = \textit{even} \leqq \lambda+\frac{1}{2}$ *the only solution of the problem* (14.12) *is* $g(x) \equiv 0$, *but for* $n \geqq \lambda+\frac{5}{2}$ *there is an infinity of solutions with the general form*

$$g(x) = c\,k(x),$$

where c *is an arbitrary constant and* $k(x)$ *a uniquely determined polynomial of exact degree* $\left(n+\lambda+\frac{1}{2}\right)$.

(Received 22 January 1955)

ЗАМЕЧАНИЕ ОБ ИНТЕРПОЛЯЦИИ. I.
(О НЕКОТОРЫХ ИНТЕРПОЛЯЦИОННЫХ СВОЙСТВАХ УЛЬТРАСФЕРИЧЕСКИХ МНОГОЧЛЕНОВ)

Я. Шураньи и П. Туран (Будапешт)

(Резюме)

Авторы занимаются определением многочленов степени, не превышающей натурального числа $2n-1$, для которых значение функции и второй производной заданы на различных местах. Оказывается, что если n — число нечетное и основные точки расположены симметрично относительно O, то задача никогда не является единственно определенной. При четном n рассматривается случай, когда основные точки являются корнями n-ого ультрасферического многочлена, соответствующего некоторому знамению параметра λ; под ультрасферическим многочленом степени n мы понимаем решение в виде многочлена степени n уравнения (8. 2). В этих случаях задача вообще единственно определена. Неопределенность возникает лишь если $\lambda + \frac{3}{2}$ есть нечетное целое число ($\geqq 3$). Особенно удобны для целей этих исследований основные точки, являющиеся корнями многочлена

$$\Pi_n(x) = \int_{-1}^{x} P_{n-1}(t)\, dt,$$

где $P_{n-1}(t)$ — многочлен Лежандра степени $n-1$. Мы вернемся к исследованию сходимости интерполяционного процесса вышеописанного типа, если основные точки корни многочлена $\Pi_n(x)$, в другом работе.

ON PROCESSES OF HAPPENINGS GENERATED BY MEANS OF A POISSON PROCESS

By
LAJOS TAKÁCS (Budapest)
(*Presented by* A. RÉNYI)

§ 1. Introduction

In the second part of our paper [8] we have discussed the following problem:

Let us consider m mutually independent stochastic processes, each defined as follows: let us consider a Poisson process in the time interval $0 \leqq u < \infty$, in which the density of the occurrence of events is p; some of these events produce a happening, the duration of which is a random variable. The first event taking place in the time interval $(0 \leqq u < \infty)$ produces a happening having the duration χ_1, the first of the events taking place after the end of first happening produces a happening of duration χ_2, etc. We suppose that the variables χ_k $(k=1, 2, 3, \ldots)$ are independent and equidistributed, with distribution function $P(\chi_k \leqq x) = H(x)$, further that χ_k can assume positive values only, that is $H(0) = 0$ and that the expectation $M(\chi_k) = \alpha$ is finite. The secondary stochastic process obtained in this manner from the Poisson process is called *the process of happenings*. Let us mention that such processes occur in practice, for instance, in connection with the counting of particles by means of a Geiger—Müller counter.

In [8] it has been shown that the probability of a happening starting between the moments u and $u+\Delta u$ is[1] $f(u)\Delta u + o(\Delta u)$ where $f(u)$ satisfies the following integral equation:

$$f(u) = p - p\int_0^u f(u-x)[1-H(x)]dx \tag{1}$$

which has a uniquely defined continuous solution $f(u)$. If the Laplace—Stieltjes

[1] By $o(\Delta u)$ we denote a function for which $\lim_{\Delta u \to 0} \frac{o(\Delta u)}{\Delta u} = 0$.

transform of $H(x)$ is

$$\psi(s) = \int_0^\infty e^{-sx} dH(x) \qquad (\Re(s) \geqq 0), \tag{2}$$

the Laplace transform of $f(u)$ will be

$$\int_0^\infty e^{-su} f(u) du = \frac{p}{s + p - p\psi(s)} \tag{3}$$

which converges for $\Re(s) > 0$; relying upon this $f(u)$ can be uniquely determined.

The probability of a happening going on at the moment u is

$$F(u) = 1 - \frac{1}{p} f(u). \tag{4}$$

It has been shown that the following limit exists:

$$f^* = \lim_{u \to \infty} f(u) = \frac{p}{1 + p\alpha} \tag{5}$$

and consequently also

$$F^* = \lim_{u \to \infty} F(u) = \frac{p\alpha}{1 + p\alpha}. \tag{6}$$

Let us now consider simultaneously m processes of happenings of the type defined above. We shall say that at the moment u the system is in the state E_k $(k = 0, 1, \ldots, m)$ if at this moment in $m - k$ processes no happening is taking place, and in k processes a happening is going on. In [8] the system we just referred to was examined first in the time interval $(0, t)$ and following that in any time interval having the length t, in case the process is going on since an infinite time. We call this latter case the stationary case. The expected numbers $m_{mk}(t)$ and $m^*_{mk}(t)$ of the transitions $E_{k-1} \to E_k$ during the time interval $(0, t)$ or, in the stationary case, during a time interval of the length t, respectively, have been determined. We say that at the moment u an (at least k-fold) coincidence takes place if the system is in one of the states E_j $(j = k, k+1, \ldots, m)$. We have determined $\tau_{mk}(t)$ and $\tau^*_{mk}(t)$, the mean total length of duration of the coincidences falling into the time intervall $(0, t)$ in the non-stationary case and in the stationary case, respectively.

We remark that the results concerning the stationary processes are to be understood that we determine the probabilities and expectations for the time interval $(u, u + t)$ and succeedingly we pass to the limit $u \to \infty$.

In the present paper we choose the special case $k = 1$ as subject of a more detailed study. We shall determine in the non-stationary case the distribution of the number of transitions $E_0 \to E_1$ in the time interval $(0, t)$; let us denote the probability of the event that this number is $\leqq n$ by $W_1(t, n)$ $(n = 0, 1, 2, \ldots)$. Further, we determine the distribution of the total length of

time in the time interval $(0, t)$, while the system is in the states E_j $(j = 1, 2, \ldots, m)$; let us denote this distribution function by $\Omega_1(t, z)$. These functions will be determined also for the stationary case and will be denoted by $W_1^*(t, n)$ and $\Omega_1^*(t, z)$, respectively.

Let us denote in succession the distances between a transition $E_o \to E_1$ and the immediately following transition $E_1 \to E_o$ by the random variables $\xi_1, \xi_2, \ldots$. It is clear that $\xi_1, \xi_2, \ldots$ are independent random variables with the same distribution function $P(\xi_k \leqq x) = D(x)$ and expectation

$$\vartheta = \int_0^\infty [1 - D(x)]\,dx. \tag{7}$$

Let us denote in the case $k = 1$ the above-mentioned mean values by $m_1(t)$, $m_1^*(t)$, $\tau_1(t)$ and $\tau_1^*(t)$, respectively. Relying on the results of [8] we have the following equations:

$$m_1(t) = \sum_{n=0}^{\infty} [1 - W_1(t, n)] = mp \int_0^t [1 - F(u)]^m\,du, \tag{8}$$

$$m_1^*(t) = \sum_{n=0}^{\infty} [1 - W_1^*(t, n)] = mp\left(\frac{1}{1+p\alpha}\right)^m t \tag{9}$$

and

$$\tau_1(t) = \int_0^t [1 - \Omega_1(t, z)]\,dz = \int_0^t [1 - (1 - F(u))^m]\,du, \tag{10}$$

$$\tau_1^*(t) = \int_0^t [1 - \Omega_1^*(t, z)]\,dz = \left[1 - \left(\frac{1}{1+p\alpha}\right)^m\right] t. \tag{11}$$

By comparison of (8) and (10), and of (9) and (11) it results

$$m_1(t) + mp\tau_1(t) = mpt \tag{12}$$

and

$$m_1^*(t) + mp\tau_1^*(t) = mpt. \tag{13}$$

The expectation ϑ can be determined using the renewal theory in the following manner:

$$\vartheta = \lim_{t \to \infty} \frac{\tau_1^*(t)}{m_1^*(t)} = \frac{(1+p\alpha)^m - 1}{mp}. \tag{14}$$

In the first part of our paper we shall discuss the determination of the above-mentioned probabilities. The method employed is, as far as we know, new and at the same time very simple, and seems to be applied successfully in the solution of many problems of similar type. In the second part of our paper we shall discuss an apparently different problem by applying the same method. As we shall see, however, this problem is closely connected with the previous one, notably, it constitutes a limiting case of the previous problem.

§ 2. Preliminaries

In this section we prove some auxiliary theorems which will be needed in the following discussion.

Let us consider a stochastic process in the time interval $0 \leqq u < \infty$ describing the state of a system, with the only possible states A and B. Let the system be at time $u = 0$ in state A and we suppose that the system passes in a random manner from state A to B, and conversely. Let the durations of the intervals in which the system is staying in the states A or B be independent random variables. Let the successive durations, while the system is staying in state A, be denoted by the random variables $\eta_0, \eta_1, \eta_2, \ldots$. Let us suppose that the η_k's have the same distribution function

$$C(x) = \begin{cases} 1 - e^{-px} & \text{if } x \geqq 0, \\ 0 & \text{if } x < 0. \end{cases} \tag{15}$$

Further, we suppose that the successive durations of the system staying in state B, which we denote by $\xi_1, \xi_2, \xi_3, \ldots$, have the same distribution function $D(x)$ for which $D(0) = 0$. Further, we suppose that the variables $\eta_0, \eta_1, \eta_2, \ldots; \xi_1, \xi_2, \ldots$ are mutually independent.

This definition determines the structure of the process uniquely. The regeneration points of the process are formed by points in which the system is in state A. Namely in such time points the probability that the following transition $A \to B$ occurs within the next time interval of length x is $C(x)$, independently of the past of the system.

It is clear that the length of time between consecutive transitions $A \to B$, which we denote by $\zeta_1, \zeta_2, \zeta_3, \ldots$, are also equally distributed independent random variables with the same distribution function

$$G(x) = \int_0^x [1 - e^{-p(x-y)}] dD(y), \tag{16}$$

since $\zeta_n = \eta_n + \xi_n$.

Let us denote by $W(t, n)$ the probability that the number of transition $A \to B$ occurring in the time interval $(0 \leqq u < t)$ is $\leqq n$, and by $W^*(t, n)$ in the stationary case. We shall determine these distribution functions as well as their mean values and moments.

We shall determine in the non-stationary case the distribution function $\Omega(t, z)$ of the total length of the duration within the time interval $(0, t)$, while the system is in state B. We shall determine the corresponding distribution function also in the stationary case and denote it by $\Omega^*(t, z)$. Further, we shall also determine the averages of these probability distributions.

Non-stationary case

The determination of $W(t, n)$. Let us denote the instants of the transitions $A \to B$ by $u_1, u_2, \ldots, u_n, \ldots$. Then the probability of a number $\leqq n$ of transition $A \to B$ occurring in the time interval $(0 \leqq u \leqq t)$ is equal to the probability of the $n+1$-st transition occurring after the time t, that is to say

$$W(t, n) = P(t < u_{n+1}) = 1 - P(u_{n+1} \leqq t) = 1 - C(t) * G_n(t), \tag{17}$$

where $G_n(t)$ denotes the n-fold convolution of the distribution function $G(t)$ which can be determined starting from $G_1(t) = G(t)$ by the following recursive formula:

$$G_n(t) = \int_0^t G_{n-1}(t-x)\, d\, G(x) \qquad (n = 2, 3, \ldots). \tag{18}$$

Namely, $u_{n+1} = \eta_0 + \zeta_1 + \zeta_2 + \cdots + \zeta_n$ is the sum of $n+1$ independent random variables. We have $P(\eta_0 \leqq x) = C(x)$ and the distribution function

$$P(\zeta_1 + \zeta_2 + \cdots + \zeta_n \leqq x) = G_n(x)$$

satisfies to the recursive formula (18).

If the Laplace—Stieltjes transform of $G(x)$ is denoted by $\varphi(s)$, i. e.

$$\varphi(s) = \int_0^\infty e^{-sx} d\, G(x), \tag{19}$$

if follows from (18) that the Laplace—Stieltjes transform of $\zeta_1 + \zeta_2 + \cdots + \zeta_n$ is

$$\int_0^\infty e^{-sx}\, d\, G_n(x) = [\varphi(s)]^n . \tag{20}$$

Taking into consideration that

$$\int_0^\infty e^{-sx}\, d\, C(x) = \frac{p}{p+s}, \tag{21}$$

it follows from (17) that the Laplace transform of $W(t, n)$ will be

$$\int_0^\infty e^{-st}\, W(t, n)\, dt = \frac{1 - [\varphi(s)]^{n+1}}{s} \tag{22}$$

whence $W(t, n)$ can be determined uniquely.

Determination of the mean value. If the expected number of transitions $A \to B$ occurring in the time interval $(0, t)$ is finite, what clearly holds in our case, we have

$$m(t) = \sum_{n=0}^{\infty} [1 - W(t, n)] = \sum_{n=0}^{\infty} [C(t) * G_n(t)]. \tag{23}$$

The Laplace—Stieltjes transform of (23) is

$$\int_0^\infty e^{-st}\, d\, m(t) = \frac{p}{p+s}\,\frac{1}{1-\varphi(s)}\,. \tag{24}$$

It is well known from the renewal theory (W. FELLER [3]) that the non-decreasing average function $m(t)$ can be determined uniquely from (24).

Determination of the moments. If the r-th moment of $W(t, n)$ is finite, it can be determined as follows:

$$m_r(t) = \sum_{n=0}^{\infty} [(n+1)^r - n^r]\,[1 - W(t, n)]. \tag{25}$$

The Laplace—Stieltjes transform of this moment is

$$\int_0^\infty e^{-st}\, d\, m_r(t) = \frac{p}{p+s} \sum_{n=0}^{\infty} [(n+1)^r - n^r]\,[\varphi(s)]^n = \tag{26}$$

$$= \frac{p}{p+s} \sum_{j=1}^{r} \left\{ {r \atop j} \right\} \frac{j!\,[\varphi(s)]^{j-1}}{[1-\varphi(s)]^j} \qquad (\Re(s) > 0)$$

wherein $\left\{ {r \atop j} \right\}$ denotes the Stirling numbers of the second kind. Here we have used the well-known relation

$$n^r = \sum_{j=1}^{r} \left\{ {r \atop j} \right\} j! \binom{n}{j}. \tag{27}$$

By (26) $m_r(t)$ can be determined uniquely.

Stationary case

The determination of $W^*(t, n)$. If the examined process has already been going on for an infinitely long time and we raise the question: what is the probability that in a time interval of length t a number $\leqq n$ of transition $A \to B$ are taking place, it is necessary to know what is the situation at the starting point of the interval. The probability of the next transition occurring within a period of time $\leqq x$, counted from a given starting moment, is

$$G^*(x) = \frac{1}{\mu} \int_0^x [1 - G(y)]\, dy \tag{28}$$

wherein μ denotes the average of $G(x)$, that is

$$\mu = \int_0^\infty [1 - G(x)]\, dx. \tag{29}$$

This can be proved in the following manner. Let us denote by ζ_u the distance between the time point u and the next transition $A \to B$. For the process ζ_u a theorem due to J. L. DOOB [2] (Theorem 12) is applicable which states that under the present conditions there exists the limiting distribution $\lim_{u\to\infty} P(\zeta_u \leqq x) = G^*(x)$ where $G^*(x)$ is given by (28).

If the distribution function $G^*(x)$ is known, $W^*(t, n)$ can be determined in the same way as $W(t, n)$, only we must replace η_0 by such a random variable which has the distribution function $G^*(x)$. We can write

$$W^*(t, n) = 1 - G^*(t) * G_n(t). \tag{30}$$

As in our case

$$\int_0^\infty e^{-sx}\, d\, G^*(x) = \frac{1-\varphi(s)}{\mu s}\,, \tag{31}$$

we have by (20)

$$\int_0^\infty e^{-st}\, W^*(t, n)\, dt = \frac{1}{s} - \frac{[1-\varphi(s)]\,[\varphi(s)]^n}{\mu s^2} \qquad (\Re(s) > 0) \tag{32}$$

whence $W^*(t, n)$ can be determined uniquely.

Determination of the mean value. In the stationary case the expected number of events occurring during the period of time t is

$$m^*(t) = \sum_{n=0}^{\infty} [1 - W^*(t, n)] = \frac{t}{\mu}\,, \tag{33}$$

what results by simple calculation, as it might have been expected in advance.

Determination of the moments. In case the moments of higher order are finite, they can be determined in a manner similar to what has been outlined in what precedes. The r-th moment is

$$m_r^*(t) = \sum_{n=0}^{\infty} [(n+1)^r - n^r][1 - W^*(t, n)], \tag{34}$$

its Laplace—Stieltjes transform is similarly to (28),

$$\int_0^\infty e^{-st}\, d\, m_r^*(t) = \frac{1}{\mu s} \sum_{j=1}^{r} \mathfrak{S}_r^j \frac{j!\,[\varphi(s)]^{j-1}}{[1-\varphi(s)]^{j-1}} \qquad (\Re(s) > 0) \tag{35}$$

and hence $m_r^*(t)$ can be determined.

Duration of staying in the state B

As we have seen the distribution function $G(x)$, of the length of time between two consecutive transitions $A \to B$, can be determined from (16). On the other hand, if we know $G(x)$, $D(x)$ can be determined from (16). As the Laplace—Stieltjes transform of $G(x)$ is $\varphi(s)$ and the Laplace—Stieltjes transform of $C(x)$ is $p/(p+s)$, it follows from (16)

$$\int_0^\infty e^{-sx}\, d\,D(x) = \frac{\varphi(s)\,(p+s)}{p} \qquad (\Re(s) \geqq 0) \tag{36}$$

whence the distribution function $D(x)$ can be determined uniquely. Let us denote the n-fold convolution of $D(x)$ by $D_n(x)$. This can be determined very simply by aid of (36) since according to that formula the Laplace—Stieltjes transform of $D_n(x)$ is

$$\int_0^\infty e^{-sx}\, d\,D_n(x) = \left[\frac{\varphi(s)\,(p+s)}{p}\right]^n \qquad (\Re(s) \geqq 0) \tag{37}$$

whence $D_n(x)$ can be determined. We put $D_0(x) = 0$ if $x < 0$, and $D_0(x) = 1$ if $x \geqq 0$.

In order to determine the probability distribution functions $\Omega(t, z)$ and $\Omega^*(t, z)$ we refer to our paper [8]: in the non-stationary case we get

$$\Omega(t, z) = \begin{cases} \sum_{n=0}^{\infty} e^{-p(t-z)}\, \dfrac{[p(t-z)]^n}{n!}\, D_n(z) & \text{if } 0 \leqq z \leqq t, \\ 1 & \text{if } z \geqq t \end{cases} \tag{38}$$

and in the stationary case

$$\Omega^*(t, z) = \tag{39}$$

$$= \begin{cases} \dfrac{1}{1+p\vartheta} \sum_{n=0}^{\infty} e^{-p(t-z)}\, \dfrac{[p(t-z)]^n}{n!} \left\{ D_n(z) + p \int_0^z [1 - D(x)]\, D_n(z-x)\, dx \right\} & \text{if } 0 \leqq z < t, \\ 1 & \text{if } z \geqq t, \end{cases}$$

where ϑ denotes the mean value of the duration of a single staying of state B:

$$\vartheta = \int_0^\infty [1 - D(x)]\, dx. \tag{40}$$

We remark that $\Omega^*(t, z)$ has a jump

$$\frac{p}{1+p\vartheta} \int_t^\infty [1 - D(x)]\, dx \tag{41}$$

at $z=t$. The probabilities $\Omega(t,z)$ and $\Omega^*(t,z)$ can also be obtained in a simple manner by aid of Laplace transformation, notably, we would confine ourselves to the case in which $t-z=a$ (a constant) and the Laplace transform with respect to z of $\Omega(a+z,z)$ and $\Omega^*(a+z,z)$, are formed with the parameter a. In this case the following equation holds:

$$(42)\qquad \int_0^\infty e^{-sz}\Omega(a+z,z)\,dz=\frac{1}{s}\sum_{n=0}^{\infty}e^{-pa}\frac{(pa)^n}{n!}\left[\frac{\varphi(s)(p+s)}{p}\right]^n=\frac{1}{s}e^{-a[p-(p+s)\varphi(s)]},$$

and similarly

$$(43)\qquad \int_0^\infty e^{-sz}\Omega^*(a+z,z)\,dz=\frac{(p+s)(1-\varphi(s))}{(1+p\vartheta)s^2}e^{-a[p-(p+s)\varphi(s)]}.$$

By inversion of the Laplace transform we obtain $\Omega(a+z,z)$ and $\Omega^*(a+z,z)$, respectively. From these the desired probability distribution functions $\Omega(t,z)$ and $\Omega^*(t,z)$ may be obtained by the substitution $a=t-z$.

The average duration of the system staying in state B is in the non-stationary case

$$(44)\qquad \tau(t)=\int_0^t[1-\Omega(t,z)]\,dz$$

and its Laplace—Stieltjes transform is

$$(45)\qquad \int_0^\infty e^{-st}\,d\tau(t)=\frac{1}{s}-\frac{1}{p+s}\sum_{n=0}^{\infty}[\varphi(s)]^n=\frac{1}{s}-\frac{1}{(p+s)[1-\varphi(s)]}\quad(\Re(s)>0).$$

In the stationary case we have

$$(46)\qquad \tau^*(t)=\int_0^t[1-\Omega^*(t,z)]\,dz=\frac{pt}{1+p\vartheta},$$

as results by simple calculation.

We shall not deal with the question of determining the moments of higher order; in this respect we refer to [8].

It appears from our above discussion that if the distribution function $G(x)$ is known, it will be possible, by considering its Laplace—Stieltjes transform $\varphi(s)$, to determine all the probabilities in question. In what follows we shall discuss cases, in which the distribution function $G(x)$ is not known, but, indirectly, we are able to determine the average function $m(t)$. If $m(t)$ is known, $\varphi(s)$ can be determined according (24) by

$$(47)\qquad \varphi(s)=\int_0^\infty e^{-sx}\,dG(x)=1-\frac{p}{p+s}\frac{1}{\int_0^\infty e^{-st}\,dm(t)},$$

the mean value of $G(x)$ is evidently

$$\mu = \int_0^\infty x\,dG(x) = \lim_{t\to\infty} \frac{t}{m(t)}. \tag{48}$$

If we denote by $P(t)$ the probability that the system is at time t in state A and $\lim_{t\to\infty} P(t) = P$ exists, then relying our results [8] we have $\mu = 1/pP$.

The main idea of our paper is the following: there exist cases in which the determination of $G(x)$ either can not be carried out or raises great difficulties, whilst at the same time $m(t)$ can be calculated in a simple manner, succeedingly $\varphi(s)$, the Laplace—Stieltjes transform of $G(x)$, is obtained very simply and thus also $G(x)$ itself can be determined.

It must be mentioned that in the next two sections we shall use the same notations as in the general case treated above.

§ 3. The solution of the problem

As easily seen, in the system consisting of m mutually independent processes of happenings the moments in which the system is in state E_0 form the regeneration points of the process. The durations of the time intervals in which the system is in state E_0 are independent random variables with the same distribution function $C(x) = 1 - e^{-mpx}$. The lengths of successive time intervals in which the system is in state *non-*E_0 (that is to say in one of the states $E_1, E_2, \ldots, E_m$) are also independent random variables with the same unknown distribution function $D(x)$.

This is the same process as discussed in the last chapter. To the state E_0 corresponds, in the preceding chapter the state A, to the state *non-*E_0 the state B and to the transition $E_0 \to E_1$ the transition $A \to B$. Let us denote the common distribution functions of the distances between consecutive transitions $E_0 \to E_1$ by $G(x)$.

We consider the Laplace—Stieltjes transform of $G(x)$

$$\varphi(s) = \int_0^\infty e^{-sx}\,dG(x) \qquad (\Re(s) \geqq 0). \tag{49}$$

If this transform is known, it will be possible, using the results of the previous chapter, to determine the probability in question.

Thus $W_1(t, n)$, the probability of $\leqq n$ transition $E_0 \to E_1$ taking place in the interval $(0, t)$, can be determined uniquely by the following expression:

$$\int_0^\infty e^{-st}\,W_1(t, n)\,dt = \frac{1}{s} - \frac{mp}{mp+s}\,\frac{[\varphi(s)]^n}{s} \tag{50}$$

which converges for $\Re(s) > 0$.

In the stationary case $W_1^*(t, n)$ can be determined from the transform

$$\int_0^\infty e^{-st} W_1^*(t, n)dt = \frac{1}{s} - \frac{[1-\varphi(s)][\varphi(s)]^n}{\mu s^2} \qquad (\Re(s) > 0). \tag{51}$$

Here μ denotes the average of $G(x)$.

In our case, in accordance with (36), the following equation holds:

$$\int_0^\infty e^{-sx} dD(x) = \frac{mp+s}{mp}\varphi(s). \tag{52}$$

Let the mean value of $D(x)$ be $\vartheta = \mu - 1/mp$ and the n-fold convolution of $D(x)$ be $D_n(x)$; in this case, in accordance with (38), the following equation holds:

$$\Omega_1(t, z) = \sum_{n=0}^\infty e^{-mp(t-z)} \frac{[mp(t-z)]^n}{n!} D_n(z) \qquad (0 \leqq z \leqq t) \tag{53}$$

and by (39)

$$\Omega_1^*(t, z) = \frac{1}{1+mp\vartheta} \sum_{n=0}^\infty e^{-mp(t-z)} \frac{[mp(t-z)]^n}{n!} \cdot$$
$$\cdot \left\{ D_n(z) + mp \int_0^z [1-D(x)] D_n(z-x)\, dx \right\} \qquad (0 \leqq z < t). \tag{54}$$

As can be seen, it is possible to answer all the questions if $\varphi(s)$ is known. This will be obtained by aid of (47) $m_1(t)$. According to (8), taking (4) into account,

$$m_1(t) = \frac{mp}{p^m} \int_0^t [f(u)]^m du \tag{55}$$

and the Laplace—Stieltjes transform of this function is

$$\int_0^\infty e^{-st} dm_1(t) = \frac{mp}{p^m} \int_0^\infty e^{-st} [f(t)]^m dt \qquad (\Re(s) > 0). \tag{56}$$

Hence, according to (47), we obtain

$$\varphi(s) = 1 - \frac{p^m}{mp+s} \left\{ \int_0^\infty e^{-st} [f(t)]^m dt \right\}^{-1}. \tag{57}$$

Thus, $\varphi(s)$ having been determined, the probabilities in question can also be obtained in a simple manner. We would remark that with a given

$H(x)$ the solution $f(u)$ is furnished by (1). The mean value of $G(x)$ is

$$\mu = \int_0^\infty x\,dG(x) = \frac{(1+p\alpha)^m}{mp} \tag{58}$$

where $\alpha = M\{\chi_k\}$, i. e. the expectation of χ_k.

We remark that, generally, our method can not be employed for studying the transitions $E_k \to E_{k+1}$ or $E_{k+1} \to E_k$ $(k \neq 0)$ because the periods elapsing between such transitions depend on the length of time during which the happenings going on at the starting moment endure. The only exception is the case in which the variables χ_k are exponentially distributed which may be treated easily by the same method. For the case $k = 0$, however, the method described is suitable for the general solution of our problem.

§ 4. Processes of happenings generated by a Poisson process

For the application of the method mentioned above we shall give another example which can be discussed by aid of this method in a very simple manner, whereas otherwise its discussion would be very complicated, even in special cases. As will be seen, this example stands in close connection with the preceding one.

Let us consider a Poisson process of density λ in the time interval $0 \leqq u < \infty$. We suppose that each event produces a happening lasting for a period of time χ_k $(k = 1, 2, \ldots)$, where the χ_k's are independent random variables having the same distribution function $P(\chi_k \leqq x) = H(x)$. Let us suppose that $\chi_k > 0$, that is $H(0) = 0$. This process has been examined by A. RÉNYI [7], who proved that the number of happenings going on at a given moment has a Poisson distribution.

We shall say that the process is at the moment u in the state E_k if the number of happenings just going on is k. Let $W_1(t, n)$ denote the probability of the event that the number of transitions $E_0 \to E_1$ in the time interval $(0, t)$ is $\leqq n$ and let us designe the same probability in the stationary case in an interval of the length t by $W_1^*(t, n)$. The distribution function of the total length of time, while the process is in the state *non-*E_0 (that is one of the states E_j, $j = 1, 2, \ldots$) in the time interval $(0, t)$, will be denoted by $\Omega_1(t, z)$ and, correspondingly, in the stationary case by $\Omega_1^*(t, z)$.

As easily seen, those instants in which the system is in state E_0 form the Markov points of the process. The lengths of successive time intervals in which the system is in state E_0 are independent random variables with the common distribution function $C(x) = 1 - e^{-\lambda x}$. The durations in which the system is in state *non-*E_0 are also independent random variables with the common unknown distribution function $D(x)$. This process agrees with the

process discussed in the second chapter if we let correspond A to E_0, B to *non*-E_0 and the transitions $A \to B$ to the transitions $E_0 \to E_1$. Accordingly, the distances between consecutive transitions $E_0 \to E_1$ are independent random variables having the common distribution function $G(x) = C(x) * D(x)$. Let us denote the mean value of $G(x)$ by μ. Let us consider the Laplace—Stieltjes transform of $G(x)$

$$\varphi(s) = \int_0^\infty e^{-sx} dG(x). \tag{59}$$

Thus we obtain by means of (16)

$$\int_0^\infty e^{-sx} dD(x) = \frac{\lambda + s}{\lambda} \varphi(s). \tag{60}$$

Let us denote the n-fold convolution of $D(x)$ by $D_n(x)$ and the average of $D(x)$ by $\vartheta = \mu - \frac{1}{\lambda}$.

Relying on the previous chapters,

$$\int_0^\infty e^{-st} W_1(t, n) dt = \frac{1}{s} - \frac{\lambda}{\lambda + s} \frac{[\varphi(s)]^n}{s}, \tag{61}$$

$$\int_0^\infty e^{-st} W_1^*(t, n) dt = \frac{1}{s} - \frac{[1 - \varphi(s)][\varphi(s)]^n}{\mu s^2} \tag{62}$$

and

$$\Omega_1(t, z) = \sum_{n=0}^\infty e^{-\lambda(t-z)} \frac{[\lambda(t-z)]^n}{n!} D_n(z) \qquad (0 \leqq z \leqq t), \tag{63}$$

$$\Omega_1^*(t, z) = \begin{cases} \dfrac{1}{1 + \lambda\vartheta} \displaystyle\sum_{n=0}^\infty e^{-\lambda(t-z)} \frac{[\lambda(t-z)]^n}{n!} \cdot \\ \qquad \cdot \left\{ D_n(z) + \lambda \int_0^z [1 - D(x)] D_n(z-x) dx \right\} & \text{if } 0 \leqq z < t, \\ 1 & \text{if } z \geqq t. \end{cases} \tag{64}$$

It appears that the determination of the probabilities in question has been reduced to the determination of $\varphi(s)$ and this, in turn, is indirectly obtained by determining the average function.

The above problem has been discussed by C. LEVERT and W. L. SCHEEN [6], L. KOSTEN [5] and W. FELLER [4] in connection with Geiger—Müller counters, whilst in connection with the distribution of random intervals along a straight line it has been discussed by C. DOMB [1]. These authors discussed only the case in which $\chi_k = \alpha$ (constant). LEVERT and SCHEEN substituted

the time interval of length t by the periphery of a circle and calculated by means of complicated multi-dimensional integrals, for the problem thus modified, the corresponding probability $W_1(t, n)$ (this, of course, differs from ours); further, they determined the factorial moments of this probability. KOSTEN criticises the discussion of the problem by LEVERT and SCHEEN; for the stationary case he indicates the exact solution $W_1^*(t, n)$ and calculates its factorial moment. His method is based on the solution by means of the Heaviside operator calculus, of equilibrium equations composed of a system of partial differential equations and integral equations. W. FELLER reduces the problem to the theory of summation of random variables and so he obtains simpler results than the previous authors. DOMB determines the distribution function $\Omega_1^*(t, z)$, or at least its Laplace transform, by aid of recursive formulae which he solves by aid of Laplace transformation.

We shall show how simply the above results may be obtained with our method.

All that we require is the determination of $\varphi(s)$. This can be determined, according to (47), by aid of the average function. If we denote by $m_1(t)$ the expexted number of transitions $E_0 \to E_1$ occurring in time interval $(0, t)$, then $m_1(t+\Delta t) - m_1(t)$ is equal to the expected number of transitions $E_0 \to E_1$ occurring in the time interval $(t, t+\Delta t)$. Let us denote by $P_0(t)$ the probability that the system is in state E_0 at the moment t. We have

$$P_0(t) = \begin{cases} e^{-\lambda t} & \text{if} \quad 0 \leqq t \leqq \alpha, \\ e^{-\lambda \alpha} & \text{if} \quad t \geqq \alpha. \end{cases} \tag{65}$$

Namely, the system is at time t in state E_0, if in the interval $(0, t)$ or $(t-\alpha, t)$, respectively, there does not occur any event.

We obtain

$$m_1(t+\Delta t) - m_1(t) = P_0(t)\lambda \Delta t + o(\Delta t). \tag{66}$$

If the system is at time t in state E_0, then in the time interval $(t, t+\Delta t)$ there is one event occurring with the probability $\lambda \Delta t + o(\Delta t)$ and more than one with the probability $o(\Delta t)$. In other cases the expected number of the transitions $E_0 \to E_1$ in time interval $(t, t+\Delta t)$ is $o(\Delta t)$.

We have from (66)

$$m_1'(t) = \begin{cases} e^{-\lambda t}\lambda & \text{if} \quad 0 \leqq t \leqq \alpha, \\ e^{-\lambda \alpha}\lambda & \text{if} \quad t \geqq \alpha. \end{cases} \tag{67}$$

Hence

$$\int_0^\infty e^{-st} dm_1(t) = \frac{\lambda}{\lambda + s} \frac{s + \lambda e^{-(\lambda+s)\alpha}}{s} \tag{68}$$

and with the aid of (47) we obtain

$$\varphi(s) = \frac{\lambda e^{-\alpha(\lambda+s)}}{s + \lambda e^{-\alpha(\lambda+s)}}. \tag{69}$$

Knowing this the probabilities in question can be determined by means of the formulae (61)—(64).

After these preliminary remarks we are passing to the general case, where $H(x)$, the distribution function of χ_k, is arbitrary. Considering the happenings beginning in the interval $(0, t)$, it appears that, as it has been shown by A. RÉNYI [7], the probability of the system being at the moment t in the state E_0 is

$$P_0(t) = e^{-\lambda \int_0^t [1 - H(x)]\, dx}. \tag{70}$$

This formula can be simply deduced as follows:

$$P_0(t) = e^{-\lambda t} + \lambda \int_0^t e^{-\lambda(t-x)} P_0(x) H(x) dx, \tag{71}$$

because the system may be in state E_0 at the time t in one of both following ways: Either in the time interval $(0, t)$ does not occur any event, the probability whereof is $e^{-\lambda t}$; or in the time interval $(0, t)$ the first transition $E_0 \to E_1$ occurs at the time $t - x$ (the density function for which is $e^{-\lambda(t-x)}\lambda$) and the produced happening terminated not later than at the time t, the probability whereof is $H(x)$, and all the happenings beginning in the time interval $(t - x, t)$ finish before the moment t, the probability whereof is $P_0(x)$. Putting $g(t) = e^{\lambda t} P_0(t)$, we have

$$g(t) = 1 + \lambda \int_0^t g(x) H(x) dx \tag{72}$$

where $g(0) = 1$. By differentiation we obtain

$$g'(t) = \lambda g(t) H(t). \tag{73}$$

Hence

$$g(t) = e^{\lambda \int_0^t H(x)\, dx} \tag{74}$$

and finally

$$P_0(t) = e^{-\lambda \int_0^t [1 - H(x)]\, dx}. \tag{75}$$

Now let us determine the expected number of transitions $E_0 \to E_1$ occurring in the time interval $(t, t + \Delta t)$. This yields the following expression:

$$m_1(t + \Delta t) - m_1(t) = P_0(t) \lambda \Delta t + o(\Delta t). \tag{76}$$

If the system is in state E_0 at time t, then in the time interval $(t, t + \Delta t)$ there occurs one event with probability $\lambda \Delta t + o(\Delta t)$ and more than one event with probability $o(\Delta t)$. In other cases the expected number of the transitions $E_0 \to E_1$ in time interval $(t, t + \Delta t)$ is $o(\Delta t)$.

From (76) we obtain

$$m_1'(t) = \lambda P_0(t) \tag{77}$$

and thus

$$\int_0^\infty e^{-st}\, dm_1(t) = \lambda \int_0^\infty e^{-st} P_0(t)\, dt = \lambda \int_0^\infty e^{-st-\lambda \int_0^t [1-H(x)]\, dx}\, dt, \tag{78}$$

whence, by (47), we obtain

$$\varphi(s) = 1 - \frac{1}{\lambda + s} \left\{ \int_0^\infty e^{-st-\int_0^t [1-H(x)]\, dx}\, dt \right\}^{-1}. \tag{79}$$

Knowing this, the probabilities in question can be calculated by aid of the formulae (61)—(64).

Now we have

$$\mu = \int_0^\infty x\, dG(x) = \frac{e^{\alpha\lambda}}{\lambda} \tag{80}$$

where $\alpha = M\{\chi_k\}$.

The process discussed in this last chapter stands in close connection with the system consisting of m processes of happenings discussed earlier. Notably, this process constitutes the limiting case of the earlier discussed one, reached when $m \to \infty$, $p \to 0$ and $mp = \lambda$ (constant). To prove this, it is sufficient to show that, while passing to this limit the formula (57) goes over into the formula (79).

Using (1) it can be shown that the following equation holds:

$$f(t) = p - p^2 \int_0^t [1 - H(x)]\, dx + o(p^2) \tag{81}$$

and thus we obtain

$$\left[\frac{f(t)}{p}\right]^m = \left[1 - p \int_0^t [1 - H(x)]\, dx + o(p)\right]^m \to e^{-\lambda \int_0^t [1-H(x)]\, dx} \quad \text{as } m \to \infty. \tag{82}$$

By employing this latter formula (79) will follow from (57). The fact that the second process is the limiting case of the preceding one is intuitively also evident.

EXAMPLE. Let $\chi_k = \alpha$, i. e. a constant, that is to say, put $H(x) = 0$ if $x < \alpha$ and $H(x) = 1$ if $x \geqq \alpha$. In this case it follows according to (79)

$$\varphi(s) = \frac{\lambda e^{-\alpha(\lambda+s)}}{s + \lambda e^{-\alpha(\lambda+s)}} \tag{83}$$

in conformity with (69). Hence

$$G(x)=\sum_{j=1}^{\left[\frac{x}{\alpha}\right]}(-1)^{j-1}\frac{e^{-j\alpha\lambda}\lambda^j}{j!}(x-j\alpha)^j \tag{84}$$

and, according to (60), the Laplace—Stieltjes transform of $D(x)$ will be

$$\int_0^\infty e^{-sx}dD(x)=\frac{\lambda+s}{\lambda}\frac{\lambda e^{-\alpha(\lambda+s)}}{s+\lambda e^{-\alpha(\lambda+s)}} \tag{85}$$

whence

$$D(x)=\sum_{j=1}^{\left[\frac{x}{\alpha}\right]}(-1)^{j-1}e^{-j\alpha\lambda}\left[\frac{\lambda^j(x-j\alpha)^j}{j!}+\frac{\lambda^{j-1}(x-j\alpha)^{j-1}}{(j-1)!}\right] \tag{86}$$

and similarly

$$D_n(x)=\sum_{j=n}^{\left[\frac{x}{\alpha}\right]}(-1)^{j-n}\binom{j-1}{n-1}e^{-j\alpha\lambda}\left[\sum_{k=0}^{n}\binom{n}{k}\frac{\lambda^{j-k}(x-j\alpha)^{j-k}}{(j-k)!}\right]. \tag{87}$$

According to (61) we obtain

$$\int_0^\infty e^{-st}W_1(t,n)dt=\frac{1}{s}-\frac{\lambda}{\lambda+s}\frac{[\varphi(s)]^n}{s}= \tag{88}$$

$$=\frac{1}{s}-\frac{\lambda}{\lambda+s}\sum_{j=n}^{\infty}(-1)^{j-n}\binom{j-1}{n-1}\frac{\lambda^j e^{-j\alpha(\lambda+s)}}{s^{j+1}}$$

whence

$$W_1(t,n)=1+(-1)^n\sum_{j=n}^{\left[\frac{t}{\alpha}\right]}\binom{j-1}{n-1}e^{-j\alpha\lambda}\left[e^{-\lambda(t-j\alpha)}-\sum_{k=0}^{j}(-1)^k\frac{\lambda^k(t-j\alpha)^k}{k!}\right] \tag{89}$$

and according to (62)

$$\int_0^\infty e^{-st}W_1^*(t,n)dt=\frac{1}{s}-\frac{1-\varphi(s)}{\mu s}\frac{[\varphi(s)]^n}{s}= \tag{90}$$

$$=\frac{1}{s}-\frac{1}{\mu}\sum_{j=n}^{\infty}(-1)^{j-n}\binom{j}{n}\frac{\lambda^j e^{-\alpha j(\lambda+s)}}{s^{j+2}}$$

wherein μ denotes the average of $G(x)$, the value whereof is, by simple calculation,

$$\mu=\frac{e^{\alpha\lambda}}{\lambda}. \tag{91}$$

Hence

$$W_1^*(t, n) = 1 - \sum_{j=n}^{\left[\frac{t}{\alpha}\right]} (-1)^{j-n} \binom{j}{n} \frac{e^{-\alpha\lambda(j+1)} \lambda^{j+1} (t - j\alpha)^{j+1}}{(j+1)!} \tag{92}$$

in conformity with the result of KOSTEN [5].

The distribution functions $\Omega_1(t, z)$ and $\Omega_1^*(t, z)$ may be written in accordance with (63) and (64) by aid of the expression (87) of $D_n(x)$. Now $\vartheta = (e^{\alpha\lambda} - 1)/\lambda$ is the average of $D(x)$. It is only an appearence that the expressions $\Omega_1(t, z)$ and $\Omega_1^*(t. z)$ contain an infinite number of terms, because $D_n(x) = 0$ if $n < x/\alpha$.

INSTITUTE FOR APPLIED MATHEMATICS
OF THE HUNGARIAN ACADEMY OF SCIENCES

(Received 17 August 1952)

Bibliography

[1] C. DOMB, The problem of random intervals on a line, *Proc. Cambridge Phil. Soc.*, **43** (1947), pp. 329—341.

[2] J. L. DOOB, Renewal theory from the point of view of the theory of probability, *Transactions of the American Mathematical Society*, **63** (1948), pp. 422—438.

[3] W, FELLER, On the integral equation of renewal theory, *Annals of Mathematical Statistics*, **12** (1941), pp. 243—267.

[4] W. FELLER, *On probability problems in the theory of counters*, Courant Anniversary Volume (New-York, 1948), pp. 105—115.

[5] L. KOSTEN, On the frequency distribution of the number of discharges counted by a Geiger—Müller counter in a constant interval, *Physica*, **10** (1943), pp. 749—756.

[6] C. LEVERT and W. L. SCHEEN, Probability fluctuation of discharges in a Geiger—Müller counter produced by cosmic radiation, *Physica*, **10** (1943), pp. 225—238.

[7] A. RÉNYI, On some problems concerning Poisson processes, *Publicationes Mathematicae (Debrecen)*, **2** (1951), pp. 66—79.

[8] L. TAKÁCS, Occurrence and coincidence phenomena in case of happenings with arbitrary distribution law of duration, *Acta Math. Acad. Sci. Hung.*, **2** (1951), pp. 275—298.

О ПРОЦЕССАХ ПОРОИСШЕСТВИЯ, ПОРОЖДЕННЫХ ПРОЦЕССОМ ПУАССОНА

Л. Такач (Будапешт)

(Резюме)

Рассмотрим m независимых друг от друга процессов происшествия в моментах $0 \leqq u < \infty$. Каждый из этих процессов определяется следующим образом: Все те события (наступающие в интервале времени $0 \leqq u < \infty$) процесса Пуассона с плотностью событий p, во время наступления которых нет в ходе происшествие, пускают в ход происшествия, сроков $\chi_1, \chi_2, \chi_3, \ldots$ соответственно. Предположим, что сроки χ_n— независимые случайные величины с общей функцией распределения $H(x)$.

Говорим, что система состоящая из m процессов происшествия находится в момент u в состоянии E_j ($j = 0, 1, 2, \ldots, m$), если число текущих происшествий равно j.

В работе исследуется распределение числа переходов $E_0 \to E_1$ наступающих в интервале времени t, и распределение времени проведенного вне состоянии E_0. Решим поставленные проблемы впервые для интервала времени $(0, t)$, а потом для $(u, u+t)$, при предельном переходе $u \to \infty$.

В предыдущей работе мы определили средние этих распределений, а теперь, с помощью нового метода, выведем из функции средних функции распределений. Определение этих функций распределения приводится к определению распределения интервалов времени между последовательными переходами $E_0 \to E_1$. Если ета функция распределения обозначается через $G(x)$, а ее преобразованная Лапласа—Стильтьеса через $\varphi(s)$, то справедливо соотношение

$$\varphi(s) = 1 - \frac{p^m}{mp+s}\left[\int_0^\infty e^{-st}(f(t))^m\,dt\right]^{-1},$$

где $f(t)$ есть однозначно определенное решение следующего интегрального уравнения типа Волтерра:

$$f(t) = p - p\int_0^t f(t-x)\,[1-H(x)]\,dx.$$

Во второй части работы исследуется процесс Пуассона плотности вероятностей λ, рассмотренная в моментах времени $0 \leqq u < \infty$. Предполагается, что все события пускают в ход происшествия, сроков $\chi_1, \chi_2, \chi_3, \ldots$ соответственно, где χ_n последовательность независимых, одинаково распределенных случайных величин; обозначим через $H(x)$ их общую функцию распределения. Говорим, что система находится в состоянии E_j ($j = 0, 1, 2, \ldots$), если число текущих происшествий равно j. Этот процесс получается из вышеописанного, если провести предельный переход $m \to \infty$ и $p \to 0$ при $mp = \lambda$. Исследуем также в этом случае распределение переходов $E_0 \to E_1$ и распределение времени проведенного вне состоянии E_0. Определение этих распределений приводится к определению распределения интервалов времени между последовательными переходами $E_0 \to E_1$. Обозначив эту функцию распределения через $G(x)$, получим для ее преобразованной Лапласа—Стильтьеса

$$\varphi(s) = 1 - \frac{1}{\lambda+s}\left[\int_0^\infty e^{-st-\lambda\int_0^t [1-H(x)]\,dx}\,dt\right]^{-1}.$$

Вышеуказанные процессы играют важную роль при решении проблем, связанных с теорией счетчиков частиц.

INVESTIGATION OF WAITING TIME PROBLEMS BY REDUCTION TO MARKOV PROCESSES

By
LAJOS TAKÁCS (Budapest)
(*Presented by* A. RÉNYI)

Introduction

By *waiting time* problems we mean questions connected with the following problem:

Let us consider a service (a counter with a single server) at which customers are arriving in the moments $0 \leqq t_1 < t_2 < \cdots < t_n < \cdots$. The sequence $\{t_n\}$ of the moments of arrivals is determined by some probability law. The arriving customers will be attended by the server in the order of their arrivals as follows: if a customer arrives at the service in a moment when nobody is waiting or attended, then he will be attended immediately. If the server is busy when a customer arrives, then this person starts to be served at the moment when the preceding finished. We suppose that the durations of the service time of the customers are random variables. We denote them by $\chi_1, \chi_2, \ldots, \chi_n, \ldots$.

If the probability laws governing the sequences of random variables, $\{t_n\}$ and $\{\chi_n\}$, furthermore, the state of the system in the moment $t=0$ are known (the service is free or in what time will it be free), then the mentioned stochastic process is uniquely defined.

We may investigate the following questions connected with the mentioned stochastic process: What is the waiting time of the n-th arriving customer? What is the waiting time of a customer eventually arriving at the moment t? To what limit does the distribution function of the waiting time converge, when $n \to \infty$ and $t \to \infty$, respectively? What is the number of waiting customers in a given moment? What kind of law is valid for the service periodes and breaks?

We shall deal with these questions in the case when $t_1, t_2, \ldots, t_n, \ldots$ are the moments of events occurring in a Poisson process. We generally suppose the Poisson process homogeneous in time, i. e. the density of events

$\lambda(t) \equiv \lambda$ (constant), but we give the results also for the inhomogeneous case, when we can do it. Concerning the sequence $\{\chi_n\}$ of service times we shall suppose that they are equidistributed mutually independent positive random variables.

There is also another interesting case. Suppose the differences $t_n - t_{n-1}$ $(n = 1, 2, \ldots;\ t_0 = 0)$ (the time intervals between two consecutive arrivals) are mutually independent positive random variables with the same (arbitrary) distribution law. We shall not consider this case in details. Originally we have also dealt with this question in our work, however, the paper of D. V. LINDLEY which appeared in the meantime made our investigations superfluous.

In the case when the sequence $\{t_n\}$ consists of the moments of the events of a Poisson process with density λ, A. JA. KHINTCHINE has determined the characteristic function of the waiting time, for the stationary state. This result of KHINTCHINE may be obtained simply by our methods. Further we shall discuss the problem also for non-stationary state, which seemed hitherto to be a very complicated case. Namely, the process is in general a non-Markov process and the "markovization" seems to lead to the solution of infinite system of integro-differential equations. However, we have stated that the process is also of Markov type, if we characterize the state by a real number: the waiting time. In this case the problem can be described by a single integro-differential equation.

We have to mention here the case of a many servers system too. In this case the waiting time problem has been dealt with only under simplifying assumptions, essentially reducing the problem to the case of a single server. Here A. K. ERLANG, F. POLLACZEK and others obtained several results.

§ 1. Formulation of the problem

Let us consider a stochastic process of the Poisson type in the moments $t \geqq 0$. Let $\lambda(t)$ be the density of the occurrence of the events, where $\lambda(t)$ is a real-valued non-negative, continuous and bounded function of the time parameter t. In the case of a Poisson process we suppose that *the numbers of the events occurring in non-overlapping time intervals are mutually independent random variables, moreover, that the probability that in the time interval* $(t, t+\Delta t)$ *an event occurs, is* $\lambda(t)\Delta t + o(\Delta t)$ *and that of occurrence of more than one event is* $o(\Delta t)$.

Let us suppose that in the moments when an event occurs in the Poisson process, a customer arrives at a service. If there is no person waiting, then he will be served immediately. If there are waiting customers, then he must wait until the persons arrived earlier will have been attended and he begins being served in the moment when the preceding person finished. Let the durations $\chi_1, \chi_2, \ldots, \chi_n, \ldots$ of the service times be mutally independent

random variables with the same distribution function $P(\chi_n \leqq x) = H(x)$. We suppose χ_n to be positive, i. e. $H(0) = 0$.

This process occurs in several cases in practice, e. g. in the case of callings arriving at a telephone centre, landing of aeroplanes, waiting of persons before a cash-box and other waiting processes.

The above terminology has been chosen for the sake of clarity. Evidently, the problem may be also formulated abstractly.

Let the duration of waiting of a person arriving eventually in the moment t be denote by the random variable $\eta(t)$. The process described by the probability function $\eta(t)$ will be called *homogeneous in time*, if $\lambda(t) \equiv \lambda$ (a constant not depending on t). If $\lambda(t)$ depends effectively on the time parameter t, then we call the process *inhomogeneous in time.*

Let $P(\eta(t) \leqq x) = F(t, x)$, i. e. $F(t, x)$ is the probability that the waiting time of a person arriving in the moment t will not exceed x. Further let $P(\eta(t_n - 0) \leqq x) = F_n(x)$, i. e. $F_n(x)$ is the probability that the waiting time of the n-th arriving person is at most x. Here t_n $(n = 1, 2, \ldots)$ denotes the moment of arriving of the n-th person.

First we shall deal with the determination of $F(t, x)$ and $F_n(x)$. Then we investigate the condition under which the limiting distributions $\lim_{t \to \infty} F(t, x)$ and $\lim_{n \to \infty} F_n(x)$ exist and show how they can be obtained. It will be shown that, under general conditions, these limiting distributions exist and they agree with each other. This common limiting distribution function gives the stationary solution, i. e. the probability that — when the process is going on since an infinitely long time — the waiting time in any moment or the waiting time of any person does not exceed x.

Having discussed these problems, we shall deal with the determination of the duration of service periodes and of the number of waiting persons.

The theory of waitings was discussed by several authors. We mention the book by Th. C. FRY [8] in which, using the results of A. K. ERLANG, and supposing a stationary state he has dealt with the laws for waitings. Namely, he has determined the average waiting time in case of service times with constant and exponential distribution, and the probability that the number of the waiting persons in a given moment equals j $(j = 0, 1, 2, \ldots)$. (In the case of exponentially distributed service times FRY [8] considers also the more general problem of a many servers system. Here he gives also the distribution function of the waiting time.) In his paper [14] F. POLLACZEK discusses the case when there are $\nu \geqq 1$ servers and the servers attend only every ν-th arriving person. He deduces the distribution function of waiting times when there is a given number of persons arriving in a given time interval at the service and the service times are distributed arbitrarily. In this paper POLLACZEK gives several asymptotic formulae for the distribution func-

tion and their expected values. From among the other works of POLLACZEK we mention only [15] in which he considers some waiting problems arising in connection with the landing of aeroplanes. In his work [11] A. N. KOLMOGOROV discusses the case when the moments of the arriving and leaving persons form a Poisson process, i. e. the duration of service times is exponentially distributed. Far-reaching results are due to A. J. KHINTCHINE [9], dealing with the stated problem in the stationary case with arbitrary service times: namely the determination of the characteristic function of waiting time. Further we mention the work of D. V. LINDLEY [13] containing KHINTCHINE's result on the stationary state as a special case. Furthermore, the work of D. G. KENDALL [10] is to be mentioned in which a very detailed bibliography of the theory of waiting time problems may be found.

In the sequel we shall discuss a problem which is more general than the preceding ones. Namely, we shall treat the non-stationary process, when it is homogeneous in time, and prove results for the inhomogeneous process too. We shall give conditions for the existence of a stationary state and the equilibrium distributions.

We think we have solved the general problem more simply and shortly than many authors the special cases of the problem. This simple solution may be explained as follows: Up to now the general opinion was (see e. g. W. FELLER [5], p. 246) that the waiting time problem leads to a Markov process only in the special case when the service time is distributed exponentially. In this case let the state E_j denote that the number of persons waiting and being served is equal to j. Then the changes of states actually form a Markov process. If, however, the service time has not an exponential distribution, then this process ceases to be of Markov type. It becomes of Markov type by extending the notion of the "state" by including in the definition of the state a continuous variable, namely the time passed since the begin of service of the person just being served. In this case, however, quoting FELLER, "The new equations can usually be written down but they are so complicated that it is doubtful whether anything is gained". However, we have noticed that the process may be considered as of Markov type even if we characterize the state of the system by a single variable, the waiting time of the person arriving eventuelly in the moment considered. This description may be carried out equally in the cases of service times with exponential and arbitrary distribution, respectively. It is a great advantage that in this procedure the states E_j do not play any role. Consequently, we have to deal only with a single integro-differential equation instead of a system with infinitely many equations.

§ 2. The distribution function of the waiting time

Let $\eta_t(t)$ be the waiting time of a person arriving eventually at the service in the moment t. If in the moment $t=0$ the server is free, then $\eta_t(0)=0$. If in the moment $t=0$ the server is not free, then we denote by $\eta_t(0)=\eta_{t_0}$ the random variable giving the moment when the server ceases to be busy. If in a given moment $\eta_t(t)=0$, then it remains zero until a person arrives at the service.

If $\{t_n\}$ and $\{\chi_n\}$ denote the sequence of moments of arrivals and the sequence of the corresponding service times, respectively, then the value of $\eta_t(t)$ will have the jump χ_n $(n=1,2,\ldots)$ in the moments t_n $(n=1,2,\ldots)$. In the meantime the value of $\eta_t(t)$ decreases linearly, but, if it assumes the value 0, then it remains zero until the moment of arrival of the next person.

Thus $\eta_t(t)$ is defined as follows: $\eta_t(0)=\eta_{t_0}$ and if $t_n<t<t_{n+1}(n=0,1,2,\ldots;$ $t_0=0)$, then

$$\eta_t(t)=\begin{cases}\eta_t(t_n)-(t-t_n) & \text{if} \quad \eta_t(t_n)>t-t_n,\\ 0 & \text{if} \quad \eta_t(t_n)\leqq t-t_n\end{cases} \tag{1}$$

and if $t=t_n$, then

$$\eta_t(t_n)=\eta_t(t_n-0)+\chi_n. \tag{2}$$

If we know the probability laws governing the sequences $\{t_n\}$ and $\{\chi_n\}$ of random variables, then by the initial condition $\eta_t(0)=\eta_{t_0}$ and by equations (1) and (2) $\eta_t(t)$ may be determined in any moment t.

We emphasize that the probability function $\eta_t(t)$ is defined for all $t\geqq 0$ and $\eta_t(t_n-0)$ yields the waiting time of the n-th arriving person. The course of $\eta_t(t)$ is illustrated by Fig. 1.

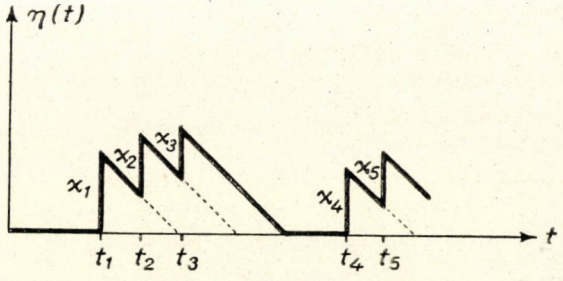


Fig. 1

If t_n agree with the moments of the events in a Poisson process and χ_n are mutually independent random variables, then the probability function $\eta_t(t)$ describes a Markov process. In fact, if we know the value of $\eta_t(t)$ in a given moment, this determines the future stochastic behaviour of $\eta_t(t)$ completely.

If the moments t_n form no Poisson process, but the time differences t_n-t_{n-1} $(n=1,2,\ldots)$ are mutually independent random variables of the same

distribution, then the process $\eta(t)$ will not be of Markov type, but the moments t_n will be the Markov points (or, in an other nomenclature, "the regeneration points") of the process. Thus, also in this case we can reduce the process to the study of Markov processes.

In the sequel we shall suppose that t_n denote the moments of the events in a Poisson process of density $\lambda(t)$ and χ_n are equidistributed mutually independent random variables with the common distribution function $H(x)$. In general, we shall suppose that $\eta(0)=\eta_0$ is arbitrary with the distribution function $P(\eta_0 \leqq x)=F_0(x)$ or, in particular, $\eta_0 \equiv 0$ with $F_0(x)=0$ for $x<0$ and $F_0(x)=1$ for $x \geqq 0$. Now, for the distribution function $P(\eta(t) \leqq x)=F(t,x)$ we prove the following

THEOREM 1. *The distribution function $F(t,x)$ of the random variable $\eta(t)$ satisfies the integro-differential equation*

$$\frac{\partial F(t,x)}{\partial t}=\frac{\partial F(t,x)}{\partial x}-\lambda(t)F(t,x)+\lambda(t)\int_0^x H(x-y)\,d_y F(t,y); \tag{3}$$

all the derivatives occurring in (3) *are either right-hand derivatives* $(x \geqq 0)$ *or left-hand derivatives* $(x>0)$. *The distribution function $F(t,x)$ has a jump of value $F(t,0)$ at $x=0$ and for $x>0$, is continuous for all t. Under the initial condition $F(0,x)=1$ $(x \geqq 0)$* (3) *has a unique solution $F(t,x)$.*

PROOF. For $0<\Delta t$ we have the equation

$$F(t+\Delta t,x)=(1-\lambda(t)\Delta t)\,F(t,x+\Delta t)+\lambda(t)\Delta t\int_0^x H(x-y)\,d_y F(t,y)+o(\Delta t). \tag{4}$$

Indeed, the event $\eta(t+\Delta t) \leqq x$ may happen in several mutually excluding ways, namely:

1. In the interval $(t, t+\Delta t)$ no event occurs (the probability is $1-\lambda(t)\Delta t+o(\Delta t)$), then we have $\eta(t) \leqq x+\Delta t$ with the probability $F(t,x+\Delta t)$.

2. In the interval $(t, t+\Delta t)$ one event occurs; its probability is $\lambda(t)\Delta t+o(\Delta t)$ and if $\eta(t)=y$ $(0 \leqq y<x)$, then we must have $\chi \leqq x-y$ whose probability is $H(x-y)$ and the distribution function of y is $F(t,y)$.

3. In the interval $(t,t+\Delta t)$ more than one event occurs the probability whereof is $o(\Delta t)$.

Later we shall show that for $x>0$ the distribution function $F(t,x)$ is a continuous function of x. At $x=0$, $F(t,x)$ has in general a jump of value $F(t,0)$. So we have for $\Delta t>0$

$$F(t,x+\Delta t)=F(t,x)+\frac{\partial F(t,x)}{\partial x}\Delta t+o(\Delta t) \tag{5}$$

where $\partial F/\partial x$ denotes a right-hand derivative which exists for $x \geqq 0$. Putting

(5) in (4), the limit-transition $\Delta t \to 0$ yields for $x \geqq 0$

$$\frac{\partial F(t,x)}{\partial t} = \frac{\partial (F,t,x)}{\partial x} - \lambda(t)F(t,x) + \lambda(t)\int_0^x H(x-y)\,d_y F(t,y)$$

which agrees with (3). This equation shows that $F(t,x)$ is differentiable with respect to t from the right and the derivative agrees with the right-hand side of the equation. Similarly, it may be shown that for $x > 0$ the left-hand derivative of $F(t,x)$ with respect to t also exists and it agrees with the right-hand side of (3), the only difference is that here $\partial F/\partial x$ denotes the left-hand derivative with respect to x. For such values x for which the right and left-hand derivatives of $F(t,x)$ with respect to x coincide (i. e. for almost every x), $F(t,x)$ is differentiable with respect to t and its derivative agrees with the right-hand side of (3).

The process described by the probability function $\eta(t)$ is a so-called Markov process of mixed type, since it consists of a mixture of discontinuous and continuous state-changings. Processes of this type have been dealt with by W. FELLER [4], he has namely shown that, under certain conditions, a unique solution $F(t,x)$ exists which is a distribution function in x. FELLER also gives the solution in the form of uniformly convergent infinite series. Unfortunately, his general discussion can not be immediately applied to this special case.

Instead of the solution $F(t,x)$ we pass now to the study of its Laplace—Stieltjes transform. In this way we shall not only prove our Theorem but also obtain the solution itself. Let us introduce the following functions converging for $\Re(s) \geqq 0$:

$$\Phi(t,s) = \int_0^\infty e^{-sx}\,dF(t,x) \tag{6}$$

and

$$\psi(s) = \int_0^\infty e^{-sx}\,dH(x). \tag{7}$$

Let us form the Laplace—Stieltjes transform of equations (3) on both sides. Then we obtain for $\Re(s) \geqq 0$

$$\frac{\partial \Phi(t,s)}{\partial t} = \Phi(t,s)[s - \lambda(t) + \lambda(t)\psi(s)] - sF(t,0). \tag{8}$$

It is easy to prove that the derivative $\partial\Phi/\partial t$ exists for all values of t and $F(t,0) = \lim_{x \to +0} F(t,x)$, the jump of $F(t,x)$ at $x = 0$. The coefficients of this differential equation are continuous functions of t and $\Phi(0,s) = 1$. Under this initial conditions, (8) has, for $\Re(s) \geqq 0$ as well known, a unique solution

$$\Phi(t,s) = e^{st - [1-\psi(s)]\Lambda(t)}\left[1 - s\int_0^t e^{-su + [1-\psi(s)]\Lambda(u)} F(u,0)\,du\right] \tag{9}$$

where

$$\Lambda(t) = \int_0^t \lambda(u)\,du. \tag{10}$$

Knowing $\Phi(t, s)$, the distribution function $F(t, x)$ may be determined uniquely, as it is known by the theory of the Laplace transform. So formula (9) determines $F(t, x)$ uniquely, provided that $F(u, 0)$ may be given uniquely for the values $0 \leqq u \leqq t$. The fact that $F(t, 0)$ can be determined uniquely for every t may be elucidated by regarding the process from the point of view of the server. The time of the server consists of alternating breaks and service periodes and $F(t, 0)$ is the probability that there is a break in the moment t. Knowing the distribution function of the duration of service periodes, $F(t, 0)$ may be determined uniquely. In case of a process inhomogeneous in time, this calculation is very complicated but it may be performed easily for homogeneous processes, which also will be done in the next chapter.

It remains to prove that the distribution function $F(t, x)$ is continuous with respect to x for every value t. $F(t, x)$ may be presented also in the following form:

$$F(t, x) = e^{\Lambda(t)}\left[1 + \int_0^t F(t-u, 0)\,\lambda(t-u)\,e^{\Lambda(t-u)}\,H(u+x)\,du + \right.$$
$$\left. + \int_0^t \left(\int_0^{x+u} F(t-u, x+u-y)\,\lambda(t-u)\,e^{\Lambda(t-u)}\,dH(y)\right)du\right]. \tag{11}$$

In fact, $\eta(t) \leqq x$ if 1) in the interval $(0, t)$ no event occurs, 2) in the moment $t-u$ (where $0 \leqq u \leqq t$) a service begins without waiting with a duration $\leqq u+x$ and in the interval $(t-u, t)$ no event occurs, 3) in the moment $t-u$ (where $0 \leqq u \leqq t$) a person arrives at the service and his waiting time is smaller than $x+u-y$, where y denotes the service time and in the interval $(t-u, t)$ nobody arrives at the service.

From the representation (11), which is at the same time a recurrent formula for determining $F(t, x)$, we can easily show the continuity of $F(t, x)$ for $x > 0$.

By our definition $F(t, x)$ is a non-decreasing function of x. $F(t, \infty) = 1$ follows from the fact that for the Laplace—Stieltjes transform of $F(t, x)$ we have $\Phi(t, 0) = 1$.

Owing to what has been said above we can also state the following

THEOREM 2. *The Laplace—Stieltjes transform*

$$\Phi(t, s) = \int_0^\infty e^{-sx}\,dF(t, x) \tag{12}$$

of the solution $F(t, x)$ *of the integro-differential equation* (1) *in Theorem* 1

may be presented in the following form:

$$\Phi(t,s)=e^{st-[1-\psi(s)]A(t)}\left[1-s\int_0^t e^{-su+[1-\psi(s)]A(u)}F(u,0)\,du\right] \tag{13}$$

where $F(u,0)$ denotes the probability that in the moment u the service is free.

We should have dealt similarly with the case when η_0 is not identically zero but it is an arbitrary random variable. If we suppose that the distribution function $F_0(x)$ of η_0 is continuous for $x>0$, then the above discussion can be applied step by step with the sole modification that in (13) the initial condition is $\Phi(0,s)=\Phi_0(s)$ instead of $\Phi(0,s)=1$, where $\Phi_0(s)$ denotes the Laplace—Stieltjes transform of $F_0(x)$ and $F(t,0)$ is also to be substituted by an expression corresponding to the new case.

Next we pass to the determination of the limiting distribution $\lim_{t\to\infty} F(t,x)=F^*(x)$. Here we shall prove the following theorems:

THEOREM 3. *Let the expected value of the random variable χ_n be*

$$\alpha=\int_0^\infty x\,dH(x)=\int_0^\infty [1-H(x)]\,dx \tag{14}$$

and let

$$\lim_{t\to\infty}\lambda(t)=\lambda \tag{15}$$

be a positive constant. Now if $\lambda\alpha<1$, then the limiting distribution function $\lim_{t\to\infty} F(t,x)=F^*(x)$ exists, is independent of the initial distribution $F_0(x)$ and is uniquely determined by the equations

$$F^*(0)=1-\lambda\alpha \tag{16}$$

and

$$\frac{dF^*(x)}{dx}=\lambda\left[F^*(x)-\int_0^x H(x-y)\,dF^*(y)\right] \tag{17}$$

where $\frac{dF^*(x)}{dx}$ denotes the right-hand derivative.

If $\lambda\alpha\geqq 1$, then the limitinq distributon $F^*(x)$ does not exist, howewer, $\lim_{t\to\infty} F(t,x)=0$ for every x.

THEOREM 4. *If $\lim_{t\to\infty}\lambda(t)=\lambda$ and $\lambda\alpha<1$, then the limiting distribution $\lim_{t\to\infty} F(t,x)=F^*(x)$ exists, is independent of the initial distribution $F_0(x)$ and is uniquely determined by the Laplace—Stieltjes transform*

$$\Phi^*(s)=\int_0^\infty e^{-sx}\,dF^*(x),\quad (\Re(s)\geqq 0), \tag{18}$$

satisfying

$$\Phi^*(s) = \frac{1-\lambda\alpha}{1-\lambda\dfrac{1-\psi(s)}{s}}. \tag{19}$$

PROOF. First we shall establish that, independently of the initial distribution $F_0(x)$, the limit $\lim_{t\to\infty} F(t,0) = 1-\lambda\alpha$ exists if $\lambda\alpha < 1$ and $\lim_{t\to\infty} F(t,0) = 0$ if $\lambda\alpha \geqq 1$. If $\eta_0 \equiv 0$ and $\lambda(t) \equiv \lambda$ (constant), then we shall show in the next chapter that the limit $\lim_{t\to\infty} F(t,0) = F^*(0) = 1-\lambda\alpha$ is valid if $\lambda\alpha < 1$. If η_0 is arbitrary, then we consider the moment when the initial reservation of the service ceases. Then, no matter how many persons are waiting there, the probability of the state E_0 will be the same when $t \to \infty$, independently of the initial state, i. e. also for an arbitrary η_0 the limit $\lim_{t\to\infty} F(t,0) = F^*(0) = 1-\lambda\alpha$ is valid. Namely, in the next chapter we shall show that the state E_0 is ergodic, hence, our statement will be a consequence of the theory of Markov chains. If $\lim_{t\to\infty} \lambda(t) = \lambda$, then, from some sufficiently large value of t on, the relation $\lambda_1 \leqq \lambda(t) \leqq \lambda_2$ holds for arbitrarily given values of λ_1 and λ_2 around λ, and independently of the initial distribution, $F_{\lambda_2}(t,0) \leqq F(t,0) \leqq F_{\lambda_1}(t,0)$ is valid for sufficiently large values of t, where $F_\lambda(t,0)$ denotes the probability of the state $\eta(t) = 0$ in case of a process of density λ. This follows from the ergodicitiy of the state E_0, using the same argument as above. Consequently, if $\lambda_2\alpha < 1$, then $1-\lambda_2\alpha \leqq \lim_{t\to\infty} F(t,0) \leqq 1-\lambda_1\alpha$. Since λ_1 and λ_2 may be chosen arbitrarily near to λ, therefore by $\lambda\alpha < 1$ we have $\lim_{t\to\infty} F(t,0) = F^*(0) = 1-\lambda\alpha$. If $\lambda\alpha \geqq 1$, then, as it will be seen, in case of $\lambda(t) \equiv \lambda$ (constant) we have $\lim_{t\to\infty} F(t,0) = 0$ which is true in general too.

In the sequel, first we shall suppose $\eta_0 \equiv 0$, i. e. $F_0(x) \equiv 1$ if $x \geqq 0$.

Now let $s = -i\omega$ in formula (13) (where ω is real). So we obtain the characteristic function of the distribution function $F(t,x)$ and denote it by $\Phi_t(\omega)$. Then

$$\Phi_t(\omega) = e^{-i\omega t-[1-\psi(-i\omega)]A(t)}\left[1 + i\omega\int_0^t e^{i\omega u+[1-\psi(-i\omega)]A(u)}F(u,0)\,du\right]. \tag{20}$$

For every t, $\Phi_t(\omega)$ is the characteristic function of a non-negative random variable. So we may apply a theorem due to A. ZYGMUND [18] stating that if, in an arbitrary interval containing the point $\omega = 0$, we have $\lim_{t\to\infty} \Phi_t(\omega) = \Phi(\omega)$ and $\Phi(\omega)$ is continuous at $\omega = 0$, then the limiting distribution function $\lim_{t\to\infty} F(t,x) = F^*(x)$ exists and its characteristic function is equal to $\Phi(\omega)$.

Now it will be shown that for the values ω satisfying $\Re[\psi(-i\omega)] < 1$ we have

$$\lim_{t\to\infty} \Phi_t(\omega) = -\frac{F^*(0)}{1+\lambda\dfrac{1-\psi(-i\omega)}{i\omega}}. \tag{21}$$

Let ω be a fixed value for which $\Re[\psi(-i\omega)] < 1$ and, for the sake of brevity, let $\varphi(t) = i\omega t + [1-\psi(-i\omega)]\,\Lambda(t) = i\omega t + \gamma\Lambda(t)$ where $\Re(\gamma) > 0$. Then

$$\Phi_t(\omega) = e^{-\varphi(t)}\left[1 + i\omega\int_0^t e^{\varphi(u)}F(u,0)\,du\right].$$

Now $\lim_{t\to\infty} e^{-\varphi(t)} = 0$ and we have to prove

$$\lim_{t\to\infty} e^{-\varphi(t)}\int_0^t e^{\varphi(u)}F(u,0)\,du = \frac{F^*(0)}{\varphi^*} \tag{22}$$

where $\varphi^* = \lim_{t\to\infty}\varphi'(t) = i\omega + [1-\psi(-i\omega)]\lambda$. For a fixed value of T we have

$$\lim_{t\to\infty} e^{-\varphi(t)}\int_0^T e^{\varphi(t)}F(u,0)\,du = 0, \tag{23}$$

so we have to verify only that, for a sufficiently large t,

$$\left| e^{-\varphi(t)}\int_T^t e^{\varphi(u)}F(u,0)\,du - \frac{F^*(0)}{\varphi^*}\right| \tag{24}$$

may be made arbitrarily small.

First of all we prove that, for a sufficiently large T,

$$|e^{-\varphi(t)}|\int_T^t |e^{\varphi(u)}|\,du < \frac{2}{\lambda\delta} \tag{25}$$

where $\delta = \Re(\gamma)$, i. e. (25) is a bounded function. In fact, we have $\Re(\varphi(t)) = \delta\Lambda(t)$ and hence

$$|e^{-\varphi(t)}|\int_T^t |e^{\varphi(u)}|\,du = e^{-\delta\Lambda(t)}\int_T^t e^{\delta\Lambda(u)}du \leqq \frac{2e^{-\delta\Lambda(t)}}{\lambda}\int_T^t e^{\delta\Lambda(u)}\lambda(u)du \leqq \frac{2}{\lambda\delta} \tag{26}$$

where T has been chosen such that for $t \geqq T$ we have $\lambda(t) \geqq \lambda/2$.

Now the inequality

$$\left| e^{-\varphi(t)}\int_T^t e^{\varphi(u)}F(u,0)du - \frac{F^*(0)}{\varphi^*}\right| \leqq \left| e^{-\varphi(t)}\int_T^t e^{\varphi(u)}[F(u,0)-F^*(0)]\,du\right| + \tag{27}$$

$$+F^*(0)\left| e^{-\varphi(t)}\int_T^t e^{\varphi(u)}du - \frac{1}{\varphi^*}\right|$$

is valid. The first term of the right-hand side is $\leqq \frac{2}{\lambda\delta} \max_{(T,t)} |F(u,0)-F^*(0)|$ which may be made arbitrarily small for sufficiently large values T. Making use of the equations

$$e^{-\varphi(t)} \int_T^t e^{\varphi(u)} \varphi'(u)\, du = 1 - e^{\varphi(T)-\varphi(t)} \tag{28}$$

and $\varphi'(u) = \varphi^* + \gamma[\lambda(u)-\lambda]$, the second term may be written as

$$e^{-\varphi(t)} \int_T^t e^{\varphi(u)}\, du - \frac{1}{\varphi^*} = -\frac{e^{\varphi(T)-\varphi(t)}}{\varphi^*} - \frac{\gamma e^{-\varphi(t)}}{\varphi^*} \int_T^t e^{\varphi(u)}[\lambda(u)-\lambda]\, du. \tag{29}$$

For sufficiently large t and T, both the first and the second term on the right-hand side may be made arbitrarily small. For $\lambda(u)-\lambda$ may be made arbitrarily small for $u \geqq T$ and (25) is valid. Thus (21) has been established.

It remains to show that there is a neighbourhood of $\omega = 0$ such that $\Re\left(\int_0^\infty e^{i\omega x}\, dH(x) \right) < 1$. If $H(x)$ is a degenerated distribution, i. e. if χ_n is equal to a constant α different from zero, then this is obvious. If, on the other hand, $H(x)$ is no degenerated distribution, then, from some result discussed in the book of A. N. KOLMOGOROV and B. V. GNEDENKO [12, p. 60], it follows that there is a number $\omega_0 \neq 0$ such that in the interval $(-\omega_0, \omega_0)$ we have $|\psi(-i\omega)| < 1$, i. e. $\Re\psi(-i\omega) < 1$ also holds. Furthermore, we have to show that (21) is continuous in the point $\omega = 0$. Since the characteristic function $\psi(-i\omega)$ is uniformly continuous for all ω and the limit

$$\lim_{\omega \to 0} \frac{1-\psi(-i\omega)}{i\omega} = -\alpha \tag{30}$$

exists, (21) is also continuous at $\omega = 0$. Thus we have shown that the limiting distribution $F^*(x)$ exists. But if the limiting distribution $F^*(x)$ exists, then, from (3) it follows that $F^*(x)$ satisfies the integro-differential equation (17).

Evidently, the Laplace—Stieltjes transform of $F^*(x)$ is

$$\Phi^*(s) = \int_0^\infty e^{-sx}\, dF^*(x) = \frac{F^*(0)}{1-\lambda\dfrac{1-\psi(s)}{s}} = \frac{1-\lambda\alpha}{1-\lambda\dfrac{1-\psi(s)}{s}}. \tag{31}$$

Here $\Phi^*(0) = 1$, since

$$\lim_{s \to 0} \frac{1-\psi(s)}{s} = \int_0^\infty x\, dH(x) = \alpha. \tag{32}$$

If $\lambda\alpha \geqq 1$, then $\lim_{t \to \infty} F(t,0) = 0$ and it is easy to prove that $\lim_{t \to \infty} \Phi_t(\omega) = 0$ whence $\lim_{t \to \infty} F(t,x) = 0$ for all x.

If the distribution function $F_0(x)$ of η_0 is arbitrary, then the limiting distribution $\lim_{t\to\infty} F(t,x)$ also exists and agrees with $F^*(x)$. Indeed, the state E_0 being ergodic, the system reaches the state E_0 with in a finite time with probability one, and, as proved above, starting from the state E_0, $F(t,x)$ tends to the limiting distribution $F^*(x)$ when $t\to\infty$.

Formula (31) is equivalent to the solution given for the stationary case by A. JA. KHINTCHINE [9].

REMARK. If, following KHINTCHINE, we suppose that, in case $\lambda\alpha<1$, the stationary state with a distribution function $F^*(x)$ exists, then we may also obtain equation (31) starting from equation (3). For, from (3) the validity of (17) follows immediately. Using it, for the Laplace—Stieltjes transform $\Phi^*(s)$ of $F^*(x)$ we obtain

$$\Phi^*(s)=\frac{F^*(0)}{1-\lambda\dfrac{1-\psi(s)}{s}}.$$

Since $F^*(x)$ is a distribution function, we must have $\lim_{s\to 0}\Phi^*(s)=1$ and, on the other hand, $\lim_{s\to 0}[1-\psi(s)]/s=\alpha$ holds, so we get

$$F^*(0)=1-\lambda\alpha.$$

Now let us suppose that the moments t_n agree with the moments of the events in a homogeneous Poisson process of density λ. Let $F_n(x)$ denote the distribution function of the waiting time of the n-th person, i. e. let $P(\eta_1(t_n-0)\leqq x)=F_n(x)$. If the original occupation of the service is η_0 with the distribution function $P(\eta_0\leqq x)=F_0(x)$, then the terms of the sequence $\{F_n(x)\}$ may be obtained successively by the following recurrence formula:

$$F_n(x)=\int_0^\infty K(x,y)\,dF_{n-1}(y) \qquad (n=1,2,\ldots) \tag{33}$$

where

$$K(x,y)=H(x-y)+\int_{x-y}^{\infty} e^{-\lambda(u+y-x)}\,dH(u). \tag{34}$$

This follows simply from the fact that the sequence $\{\eta_1(t_n-0)\}$ of random variables forms a Markov chain with the transition probabilities

$$P(\eta_1(t_n-0)\leqq x\,|\,\eta_1(t_{n-1}-0)=y)=K(x,y).$$

We investigate now the behaviour of $F_n(x)$ when $n\to\infty$. We shall prove the following

THEOREM 5. *If for the expected value $M(\chi_n)=\alpha$ we have $\lambda\alpha<1$, then, independently of the initial distribution function $F_0(x)$, the distribution functions $F_n(x)$ converge to $F^*(x)$ determined in Theorems* 3 *and* 4 *when $n\to\infty$. If $\lambda\alpha\geqq 1$, then $\lim_{n\to\infty} F_n(x)=0$ for all x.*

PROOF. First let us consider the case when $\eta_0 \equiv 0$. Then it is clear that $F_{n+1}(x) \leqq F_n(x)$. In fact, $F_n(x)$ is the probability that the waiting time of the n-th person does not exceed x, if the service of the first person has begun without waiting, and $F_{n+1}(x)$ may be regarded as the probability that the waiting time of the n-th person does not exceed x in a waiting where the service of the first person begins eventually after some waiting. Since the sequence $\{F_n(x)\}$ decreases monotonically and is bounded from below, there exists a limit $\lim_{x \to \infty} F_n(x)$ for all x.

Now, let $G(t, x)$ denote the distribution function of the waiting time of the person arriving last in the time interval $(0, t)$. Further, let $K_t(x, y)$ be the conditional probability that a person, eventually arriving at the service in the moment t, has a waiting time $\leqq x$, provided that the person arriving last in the time interval $(0, t)$ had a waiting time y. Then we have

$$F(t, x) = \int_0^t K_t(x, y) d_y G(t, y). \tag{35}$$

It can be easily shown that

$$\lim_{t \to \infty} K_t(x, y) = K(x, y) \tag{36}$$

where $K(x, y)$ is defined by (33) and

$$\lim_{t \to \infty} G(t, y) = \lim_{n \to \infty} F_{n+1}(y). \tag{37}$$

Indeed, if we denote by t_n the moment of the arrival of the person arriving last in the time interval $(0, t)$, then in case $t \to \infty$ the random variable $t - t_n$ is distributed exponentially with a density function $\lambda e^{-\lambda x}$ $(x \geqq 0)$ and this agrees with the distribution of the differences between consecutive arrivals. Further, if $t \to \infty$, then $n \to \infty$ with probability one. Hence by (35) we obtain

$$\lim_{t \to \infty} F(t, x) = \lim_{n \to \infty} \int_0^\infty K(x, y) d F_{n+1}(y) = \lim_{n \to \infty} F_{n+2}(x) \tag{38}$$

where the last equation follows from (33). We have shown before that the limit $\lim_{n \to \infty} F_n(x)$ exists and so, by (38), it agrees with $\lim_{t \to \infty} F(t, x)$, that is, $\lim_{n \to \infty} F_n(x) = F^*(x)$ if $\lambda a < 1$ and $\lim_{n \to \infty} F_n(x) = 0$ if $\lambda a \geqq 1$.

If η_0 is arbitrary, then a limiting distribution also exists and agrees with $F^*(x)$. For, as it will be proved in the next chapter, the state E_0 is ergodic. Thus, starting from an arbitrary initial state, we shall have, with probability one, a moment in which the system has the state E_0. After this moment the system behaves as if it started originally with the condition $\eta_0 \equiv 0$.

EXAMPLES. 1. Let us have $\chi_n \equiv \alpha$ (constant); then $\psi(s) = e^{-s\alpha}$ and by (19) in case $\lambda\alpha < 1$ we have

$$\Phi^*(s) = \frac{1-\lambda\alpha}{1-\lambda\dfrac{1-e^{-\alpha s}}{s}} = (1-\lambda\alpha)s\sum_{j=0}^{\infty}(-1)^j\frac{\lambda^j e^{-s\alpha j}}{(s-\lambda)^{j+1}} \tag{39}$$

whence

$$F^*(x) = (1-\lambda\alpha)\sum_{j=0}^{\left[\frac{x}{\alpha}\right]}(-1)^j e^{\lambda(x-\alpha j)}\frac{\lambda^j(x-j\alpha)^j}{j!}. \tag{40}$$

2. Let us suppose that χ_n has an exponential distribution, i. e. $H(x) = 1-e^{-x/\alpha}$ (if $x \geqq 0$); then $\psi(s) = 1/(1+\alpha s)$ and hence

$$\Phi^*(s) = \frac{(1-\lambda\alpha)(1+\alpha s)}{(1-\lambda\alpha)+\alpha s} \tag{41}$$

and

$$F^*(x) = 1-\lambda\alpha e^{-\frac{1}{\alpha}\frac{\lambda\alpha}{\alpha}x}. \tag{42}$$

§ 3. Investigations concerning homogeneous processes

In this chapter we always suppose that the moments t_n are the moments of events in a homogeneous Poisson process of density λ. For the sake of simplicity let $\eta_0 \equiv 0$, i. e. let the starting occupation of the service be zero. It is clear that the time of the server is composed of pauses and service periodes. Let $\vartheta_1, \vartheta_2, \ldots$ and $\xi_1, \xi_2, \ldots$ be the durations of consecutive pauses and service periodes, respectively. They are all mutually independent random variables. The ϑ_n are distributed exponentially with the distribution function $P(\vartheta_n \leqq x) = 1-e^{-\lambda x}$ for $x \geqq 0$. The ξ_n have a common distribution law $P(\xi_n \leqq x) = G(x)$ which will be determined below. We shall also determine $F(t, 0)$, i. e. the probability that in the moment t the server is free. We shall prove that $\lim_{t\to\infty} F(t, 0) = 1-\lambda\alpha$ when $\lambda\alpha < 1$ and $\lim_{t\to\infty} F(t, 0) = 0$ when $\lambda\alpha \geqq 1$. Furthermore, in the stationary case ($t \to \infty$, $\lambda\alpha < 1$) we shall determine the probability that the system will stay in the state E_j ($j = 0, 1, 2, \ldots$), i. e. that the number of persons waiting and being served be equal to j.

1. *Investigation of the service period. Determination of* $G(x)$. Suppose that a person without waiting arrives at the service. In the moment of the arrival he will be attended immediately. Let us suppose that the duration of this service time is y. The probability that in this time interval n persons arrive at the counter is

$$e^{-\lambda y}\frac{(\lambda y)^n}{n!}, \tag{43}$$

i. e. it has a Poisson distribution. If $n=0$, then the service period consists of a single service and the distribution function of this is $H(x)$. If $n \geqq 1$, then the server after having attended the first person, starts to attend one of the waiting persons. The other waiting persons may be imagined standing aside. Having finished the service of this person and the persons arriving eventually in the meanwhile (the distribution function of this duration is $G(x)$) the server begins to serve one of the persons standing aside. This person and the other persons arriving in the meanwhile having been served again the service begins to attend one of the persons standig aside again. This will be continued until there are persons standing aside. From the point of wiew of the service it is perfectly indifferent whether the persons will be attend in order of their arrivals or not. This fact touches the persons only, namely, the distribution function of the waiting time is changed by this fact but $F(t, 0)$ and the expected waiting time remain unaltered. Let $G_n(x)$ denote the distribution function of the sums of n mutually independent random variables whose distribution agrees with that of ξ, i. e. the n-fold convolution of $G(x)$. $G_n(x)$ may be calculated by the following recurrence formula:

$$G_n(x) = \int_0^x G_{n-1}(x-y)\, dG(y), \qquad (n = 2, 3, \ldots) \tag{44}$$

where $G_1(x) = G(x)$. By what has been told we may write

$$G(x) = \int_0^x \sum_{n=0}^{\infty} e^{-\lambda y} \frac{(\lambda y)^n}{n!} G_n(x-y)\, dH(y). \tag{45}$$

In fact, the length of the service periode does not exceed x if the service of the first arriving person lasts a time y $(0 < y \leqq x)$; the distribution function of it is $H(y)$ and during his service time there are n $(n = 0, 1, 2, \ldots)$ arriving persons (the probability of it is given by (43)) and the service duration of these and those having arrived in the meanwhile does not exceed $x-y$, the probability whereof is $G_n(x-y)$ as seen. We note that $G_0(x) = 1$ if $x \geqq 0$ and $G_0(x) = 0$ if $x < 0$.

Let the Laplace—Stieltjes transform of $G(x)$ be

$$\Gamma(s) = \int_0^{\infty} e^{-sx}\, dG(x) \tag{46}$$

which converges for $\Re(s) \geqq 0$.

Passing from equation (45) *to the Laplace transform, we obtain*

$$\Gamma(s) = \sum_{n=0}^{\infty} \frac{(\Gamma(s))^n}{n!} \int_0^{\infty} e^{-(s+\lambda)x} (\lambda x)^n\, dH(x) = \sum_{n=0}^{\infty} (-1)^n \frac{\lambda^n (\Gamma(s))^n \psi^{(n)}(s+\lambda)}{n!} = \tag{47}$$

$$= \psi[s+\lambda-\lambda\Gamma(s)],$$

i. e. for $\Re(s) \geqq 0$, $\Gamma(s)$ *the Laplace—Stieltjes transform of* $G(x)$ *satisfies the functional equation*

$$\Gamma(s) = \psi[s + \lambda - \lambda\Gamma(s)]. \tag{48}$$

A sketchy proof of this equation has been already given by D. G. KENDALL [10].

Now we prove the following theorem:

THEOREM 6. *The Laplace transform of the distribution function* $G(x)$ *of the service period* $\Gamma(s)$ *is the uniquely defined analytic solution of the functional equation*

$$\Gamma(s) = \psi[s + \lambda - \lambda\Gamma(s)], \tag{49}$$

valid for $\Re(s) \geqq 0$, *where* $\Gamma(s)$ *is subject to the condition* $\Gamma(\infty) = 0$. $G(x)$ *can be determined uniquely by* $\Gamma(s)$.

Denote by p^* *the smallest positive number for which*

$$\psi(\lambda(1 - p^*)) = p^*, \tag{50}$$

then $\lim_{x \to \infty} G(x) = p^*$. *If* $\lambda\alpha \leqq 1$, *then* $p^* = 1$ *and* $G(x)$ *is a proper distribution function, while if* $\lambda\alpha > 1$, *then* $p^* < 1$ *and* $G(x)$ *is an improper distribution function, namely, in such cases the service period can be infinite with probability* $(1 - p^*)$.

PROOF. It is clear that $G(x)$ exists for each finite x and, as we have shown, $\Gamma(s)$ satisfies the functional equation (49). Further it is clear that $\Gamma(s)$ is an analytic function for $\Re(s) > 0$, and $\Gamma(\infty) = 0$. Knowing $\Gamma(s)$, $G(x)$ can be determined uniquely. Therefore, it is sufficient to show that $\Gamma(s)$ is uniquely determined by $\Gamma(\infty) = 0$ and by its analytic character. It is sufficient to know $\Gamma(s)$ only for non-negative real values s, because it can be uniquely extended for $\Re(s) \geqq 0$. The function $\psi(s)$ decreases monotonically for $0 \leqq s < \infty$; $\psi(0) = 1$ and $\psi(\infty) = 0$. $\psi(s)$ is differentiable for $0 \leqq s < \infty$ and $\psi'(s)$ increases monotonically; $\psi'(0) = -\alpha$ and $\psi'(\infty) = 0$. Let $\Gamma(s) = x$ in (49), then $s = \Gamma^{-1}(x)$. $\psi(s)$ is a monotonous function, therefore the inverse function $\psi^{-1}(x)$ is uniquely determined for $0 \leqq x \leqq 1$. Hence from (49) we get

$$\Gamma^{-1}(x) = \psi^{-1}(x) - \lambda(1 - x),$$

i. e. $\Gamma^{-1}(x)$ is uniquely determined for $0 \leqq x \leqq 1$. Now we investigate whether $\Gamma(s)$ can also be determined uniquely.

$\Gamma^{-1}(0) = \infty$ and $\Gamma^{-1}(1) = 0$ and for $0 < x < 1$ the derivative

$$\frac{d\Gamma^{-1}(x)}{dx} = \lambda + \frac{1}{\psi'(\psi^{-1}(x))}$$

exists and is a monotonically increasing function. In $x = 1$ this derivative has the value $\lambda - \frac{1}{\alpha}$. If $\lambda\alpha \leqq 1$, then this derivative is not positive and

$\Gamma^{-1}(x)=0$ has no root in the interval $0 \leqq x<1$, but $\Gamma^{-1}(1)=0$. If $\lambda\alpha>1$, then this derivative is negative and $\Gamma^{-1}(x)=0$ has one and clearly only one root in the interval $0 \leqq x<1$. Let $x=p^*$ denote the smallest positive root of $\Gamma^{-1}(x)=0$. In the interval $(0 \leqq x \leqq p^*)$ the function $\Gamma^{-1}(x)$ is a monotonically decreasing continuous function; therefore its inverse function $\Gamma(s)$ exists, it is the Laplace transform of the sought $G(x)$ for $0 \leqq s<\infty$. This function $\Gamma(s)$ is uniquely determined by the prescribed conditions. The functional equation has in general more solutions but only one of them satisfies the mentioned conditions. Now $\Gamma(0)=p^*$, that is $\lim_{x\to\infty} G(x)=p^*$. Further the roots of the equation $\Gamma^{-1}(x)=0$ agree with those of the equation $\psi(\lambda(1-x))=x$. This completes the proof of the theorem.

EXAMPLE. If $H(x)=1-e^{-x/\alpha}$ $(x \geqq 0)$, then $\psi(s)=1/(1+\alpha s)$ and hence

$$\Gamma(s)=\frac{1+\alpha(s+\lambda)-\sqrt{[1+\alpha(s+\lambda)]^2-4\lambda\alpha}}{2\lambda\alpha}. \tag{51}$$

Now (49) has two solutions corresponding to the positive and negative sign of the root, respectively. We must take the solution with negative square root because of the condition $\Gamma(\infty)=0$. Inverting (51), we obtain

$$G'(x)=e^{-\frac{(1+\lambda\alpha)x}{\alpha}} I_1\left(\frac{2\sqrt{\lambda\alpha}x}{\alpha}\right)\frac{1}{\sqrt{\lambda\alpha}x} \tag{52}$$

where $I_1(x)=J_1(ix)/i$ and $J_1(x)$ denotes the Bessel function of genus 1. Here $G(\infty)=1$ if $\lambda\alpha \leqq 1$, while $G(\infty)=1/\lambda\alpha$ if $\lambda\alpha>1$.

(49) can be used excellently for the determination of the moments of $G(x)$.

Now we shall determine the average service periode

$$\mu=\int_0^\infty x\,dG(x) \tag{53}$$

directly from (45). From (45) we obtain

$$\mu=\alpha+\sum_{n=1}^{\infty} n\mu\int_0^\infty e^{-\lambda y}\frac{(\lambda y)^n}{n!}\,dH(y). \tag{54}$$

This can be seen directly too: μ is composed of the average service time α of the first person, and if n persons arrive during the service time of the first — the probability whereof is

$$\int_0^\infty e^{-\lambda y}\frac{(\lambda y)^n}{n!}\,dH(y)$$

— then, by the discussion given above, the average service time of these and the others coming in the meanwhile is equal to $n\mu$.

Performing the summation in (54), we have

$$\mu = \alpha + \lambda\alpha\mu. \tag{55}$$

If $\lambda\alpha < 1$, *then*

$$\mu = \frac{\alpha}{1-\lambda\alpha}, \tag{56}$$

however, if $\lambda\alpha \geqq 1$, *then* $\mu = \infty$.

We have mentioned that in case $\lambda\alpha > 1$ the probability that the service period is finite equals $p^* = \Gamma(0)$. Then the question arises: what is the conditional expected value of the service period under the assumption that the service time is finite? This conditional expected value is

$$\mu^* = -\frac{\Gamma'(0)}{\Gamma(0)}. \tag{57}$$

Here $\Gamma(0) = p^*$ and by (49) $\Gamma'(0) = \psi'(\lambda(1-p^*))[1-\lambda\Gamma'(0)]$, from which $\Gamma'(0)$ can be determined. Finally, we obtain

$$\mu^* = \frac{-\psi'(\lambda(1-p^*))}{p^*[1+\lambda\psi'(\lambda(1-p^*))]}. \tag{58}$$

In case of an exponentially distributed service time $p^* = 1/\lambda\alpha$ if $\lambda\alpha > 1$ and

$$\mu^* = \frac{\alpha}{\lambda\alpha - 1}. \tag{59}$$

In case of a constant service time ($\chi_n = \alpha$), p^* is the smallest positive root of the equation $\lambda\alpha p^* e^{-\lambda\alpha p^*} = \lambda\alpha e^{-\lambda\alpha}$ and

$$\mu^* = \frac{\alpha}{1-\lambda\alpha p^*}. \tag{60}$$

Concerning the service period an other problem arises: How many persons will be served in a service period? Let f_j denote the probability that a service periode consists of j ($j = 1, 2, \ldots$) services. The probability that during the service of a single person n persons arrive at the service counter, is given by

$$p_n = \int_0^\infty e^{-\lambda x}\frac{(\lambda x)^n}{n!}\,dH(x) \qquad (n = 0, 1, 2, \ldots), \tag{61}$$

as mentioned above. Now it is clear that

$$f_j = \sum_{\substack{n_1+n_2+\ldots+n_j = j-1 \\ n_1+n_2+\ldots+n_k \geqq k \ (k=1,2,\ldots,j-1)}} p_{n_1}p_{n_2}\ldots p_{n_j}. \tag{62}$$

However, it seems more convenient to determine the generating function

$$F(\omega) = \sum_{j=1}^{\infty} f_j \omega^j. \tag{63}$$

We have the recurrence formulae

$$(64)\qquad \begin{cases} f_1 = p_0, \\ f_j = \sum_{n=1}^{j-1} p_n \sum_{j_1+j_2+\ldots+j_n=j-1} f_{j_1} f_{j_2} \cdots f_{j_n}. \end{cases}$$

The reasoning is perfectly analogous to that used to establish formula (45) for $G(x)$. Passing to the generating function from (64) and using

$$(65)\qquad \psi(\lambda(1-\omega)) = \sum_{n=0}^{\infty} p_n \omega^n,$$

we obtain the following functional equation for $F(\omega)$:

$$(66)\qquad F(\omega) = \omega\psi[\lambda - \lambda F(\omega)].$$

THEOREM 7. *The generating function $F(\omega)$ of the probabilities f_j is the uniquely determined, for $|\omega| \leqq 1$ analytic solution of the functional equation*

$$(67)\qquad F(\omega) = \omega\psi[\lambda - \lambda F(\omega)]$$

subject to the condition $F(0) = 0$. Let $x = p^*$ be the smallest positive real root of the equation $\psi(\lambda(1-x)) = x$; then $\lim_{\omega \to 1} F(\omega) = p^*$.

PROOF. It is clear that the probabilities f_j exist. Hence the generating function $F(\omega)$ exists also for $|\omega| \leqq 1$, $F(0) = 0$, $F(\omega)$ is an analytic function for $|\omega| \leqq 1$ and $F(\omega)$ satisfies the functional equation (67). Knowing $F(\omega)$, the probabilities $\{f_j\}$ can be determined uniquely. Therefore, it is sufficient to show that the conditions mentioned determine $F(\omega)$ uniquely. It is sufficient to determine $F(\omega)$ for real values with $0 \leqq \omega \leqq 1$, for then $F(\omega)$ can be determined for all ω with $|\omega| \leqq 1$. Let $x = F(\omega)$ in (67), then $\omega = F^{-1}(x)$ and by (67) we have

$$F^{-1}(x) = \frac{x}{\psi(\lambda(1-x))}.$$

Now $F^{-1}(0) = 0$, $F^{-1}(1) = 1$ and $\dfrac{dF^{-1}(x)}{dx}$ is monotonically decreasing for $0 \leqq x \leqq 1$. The value of the derivative in $x = 1$ is $1 - \lambda\alpha$. If this derivative is non-negative, that is $\lambda\alpha \leqq 1$, then $F^{-1}(x)$ is monotonically increasing for $0 \leqq x \leqq 1$ and $F^{-1}(1) = 1$. If this derivative is positive, that is $\lambda\alpha > 1$, then $F^{-1}(x) = 1$ has one and only one root $x = p^*$ in the interval $0 < x < 1$; consequently, $F^{-1}(x)$ is monotonically increasing for $0 \leqq x \leqq p^*$ and $F^{-1}(p^*) = 1$. Hence, if $x = p^*$ denotes the smallest positive root of $F^{-1}(x) = 1$, then $F^{-1}(x)$ can be uniquely inverted for $0 \leqq x \leqq p^*$. $F(0) = 0$, $F(\omega)$ is a continuous function of ω and $\lim_{\omega \to 1} F(\omega) = p^*$. By the above procedure $F(\omega)$ is determined uniquely. p^* denotes the probability that the service period consists of a finite number of services.

EXAMPLE. Let $\chi_n = \alpha$ (constant). Then $\psi(s) = e^{-s\alpha}$ and it is clear that

$$G(x) = \sum_{j=1}^{\left[\frac{x}{\alpha}\right]} f_j. \tag{68}$$

So the determination of $G(x)$ is reduced to that of the probabilities f_j. Now $F(\omega)$ satisfies the following functional equation:

$$F(\omega) = \omega e^{-\lambda\alpha[1-F(\omega)]}. \tag{69}$$

If we use the notations $y = \lambda\alpha F(\omega)$ and $x = \omega e^{-\lambda\alpha}\lambda\alpha$, then we obtain the relation

$$ye^{-y} = x. \tag{70}$$

Hence we have to determine the functions $y = y(x)$, i. e. the inverse function of (70). As it is well known, the branch passing through the origin of $y = y(x)$ may be represented as the power series

$$y(x) = \sum_{j=1}^{\infty} \frac{j^{j-1}}{j!} x^j \tag{71}$$

for $|x| \leqq 1/e$, i. e. for the values $|\omega| \leqq 1$ we obtain

$$F(\omega) = \sum_{j=1}^{\infty} \frac{e^{-\lambda\alpha j}(\lambda\alpha j)^{j-1}}{j!} \omega^j. \tag{72}$$

We remark that from (69) we obtain the differential equation

$$\omega F'(\omega) - F(\omega) = \lambda\alpha\omega F(\omega) F'(\omega)$$

whose solution corresponding to the initial condition $F(0) = 0$ is given also by (72).

So, by (72), we obtain

$$G(x) = \sum_{j=1}^{\left[\frac{x}{\alpha}\right]} \frac{e^{-\lambda\alpha j}(\lambda\alpha j)^{j-1}}{j!}. \tag{73}$$

By (62) we may put

$$f_j = e^{-\lambda\alpha j}(\lambda\alpha)^{j-1} \sum_{\substack{n_1+n_2+\ldots+n_j = j-1 \\ n_1+n_2+\ldots+n_k \geqq k \ (k=1,2,\ldots,j-1)}} \frac{1}{n_1!\, n_2! \ldots n_j!} = \frac{e^{-\lambda\alpha j}(\lambda\alpha j)^{j-1}}{j!}, \tag{74}$$

hence, (73) may be obtained in this way too.

If χ_n is exponentially distributed, then

$$F(\omega) = \frac{(1+\lambda\alpha) - \sqrt{(1+\lambda\alpha)^2 - 4\lambda\alpha\omega}}{2\lambda\alpha}$$

from which

$$f_j = \frac{1}{2}\binom{2j}{j} \frac{(\lambda\alpha)^{j-1}}{(1+\lambda\alpha)^{2j-1}}.$$

2. *Determination of* $F(t, 0)$. In our paper [16] we have dealt with stochastic processes in which pauses and happenings with distribution function $1-e^{-\lambda x}$ and $G(x)$, respectively, are alternating. We have shown that if the process starts in the moment $t=0$ with a pause, then the probability that in the interval $(t, t+\Delta t)$ a happening (here: service period) begins is $f(t)\Delta t+o(\Delta t)$ where $f(t)$ is continuous and is defined uniquely by the integral equation of Volterra type

$$f(t)=\lambda-\lambda\int_0^t f(t-x)\,[1-G(x)]\,dx. \tag{75}$$

Now we have $f(t)\Delta t+o(\Delta t)=F(t,0)\lambda\Delta t+o(\Delta t)$. In fact, in the interval $(t, t+\Delta t)$ a service period begins when in the moment t there is no waiter, the probability whereof is $F(t, 0)$, and in the interval $(t, t+\Delta t)$ somebody arrives at the service, the probability whereof is $\lambda\Delta t+o(\Delta t)$. Hence

$$F(t, 0)=\frac{f(t)}{\lambda}. \tag{76}$$

By (75), for $\Re(s)>0$ we have

$$\int_0^\infty e^{-st}F(t, 0)\,dt=\frac{1}{s+\lambda-\Gamma(s)}. \tag{77}$$

Hence $F(t, 0)$ may be determined uniquely ($F(0, 0)=1$).

In our paper [16] we have proved that the limit

$$\lim_{t\to\infty} f(t)=\frac{\lambda}{1-\lambda\mu} \tag{78}$$

exists, if μ is finite; μ is the mean value of $G(x)$. Hence by (76) and (55) we have

$$\lim_{t\to\infty} F(t, 0)=F^*(0)=\frac{1}{1+\lambda\mu}=1-\lambda\alpha \quad \text{for} \quad \lambda\alpha<1. \tag{79}$$

So we have proved the existence of the limit $F^*(0)$ in the case homogeneous in time. It was used in the preceding chapter.

3. *Determination of the probability of the state* E_j. Let us consider the homogeneous process in the stationary case. Let P_j ($j=0, 1, 2, \ldots$) denote the probability that, in a moment chosen arbitrarily, the system has the state E_j ($j=0, 1, 2, \ldots$), i. e. for $t\to\infty$, the number of persons waiting and being served is j. P_0 is the probability that the system is in the state E_0, i. e. there is no person waiting or being served. As we have seen, in the case $\lambda\alpha<1$ we have

$$P_0=1-\lambda\alpha. \tag{80}$$

If $\lambda\alpha\geqq 1$, then, as we have proved, $\mu=\infty$ and by the theory of recurrent events we have $P_0=0$. (See W. FELLER [6], [7].)

In his work [9] A. JA. KHINTCHINE has shown that in the stationary state (i. e. for $t \to \infty$) the probability that the system has the state E_j at the end of a service time, agrees with the probability that in an arbitrary moment the system has the state E_j. Thus it suffices to establish these probabilities for the endpoints of the service times only. The random variables, $\zeta_1, \zeta_2, \zeta_3, \ldots$ describe the states at the endpoints of the consecutive service times, where $\zeta_n = j$ if at the end of the n-th service the system has the state E_j. The process described by the random variables $\{\zeta_n\}$ forms a Markov chain with the transition probabilities

$$P(\zeta_n = k+j-1 \,|\, \zeta_{n-1} = k) = p_j = \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^j}{j!}\, dH(x) \tag{81}$$

for $j = 0, 1, 2, \ldots$ and $k \geqq 1$, and

$$P(\zeta_n = j \,|\, \zeta_{n-1} = 0) = p_j \tag{82}$$

for $k = 0$.

It may be easily proved that in our case we can apply a well-known theorem (Cf. W. FELLER [7], p. 325): If the states of an irreducible chain are ergodic, then, independently of the initial value ζ_1, the limiting probabilities

$$\lim_{n \to \infty} P(\zeta_n = j) = P_j \qquad (j = 0, 1, 2, \ldots) \tag{83}$$

exist and the values P_j may be determined uniquely from the system of equations

$$P_j = P_{j+1} p_0 + P_j p_1 + \cdots + P_1 p_j + P_0 p_j, \qquad (j = 0, 1, 2, \ldots). \tag{84}$$

In our case the Markov chain is irreducible, for each state may be reached from any other state in a sufficient number of steps. As known, all the states of an irreducible chain belong to the same class. In case $\lambda\alpha < 1$, as we have seen, the state E_0 is ergodic, hence the same holds for every state. If $\lambda\alpha \geqq 1$, then the state E_0 is a recurrent-null state, hence so are all other states, i. e. $P_j = 0$ for every j.

Let

$$\Pi(\omega) = \sum_{j=0}^{\infty} P_j \omega^j \tag{85}$$

and

$$\pi(\omega) = \sum_{j=0}^{\infty} p_j \omega^j = \int_0^\infty e^{-\lambda(1-\omega)x}\, dH(x) = \psi(\lambda(1-\omega)); \tag{86}$$

then by making use of these generating functions, from (79) we obtain

$$\omega \Pi(\omega) = \pi(\omega)[\Pi(\omega) + P_0 \omega - P_0]. \tag{87}$$

Hence

$$\Pi(\omega) = P_0 \frac{\pi(\omega)}{1 - \dfrac{1-\pi(\omega)}{1-\omega}}. \tag{88}$$

If $\omega \to 0$, then

$$\frac{1-\pi(\omega)}{1-\omega} \to \lambda\alpha, \tag{89}$$

and hence $\Pi(0) = P_0/(1-\lambda\alpha)$. Considering that $\Pi(0) = 1$ or $P_0 = 1-\lambda\alpha$ if $\lambda\alpha < 1$, we obtain

$$\Pi(\omega) = \frac{(1-\lambda\alpha)\psi(\lambda(1-\omega))}{1 - \dfrac{1-\psi(\lambda(1-\omega))}{1-\omega}}. \tag{90}$$

EXAMPLE. 1) Let χ_n be an exponentially distributed random variable, i. e. $H(x) = 1-e^{-x/\alpha}$. Then $\psi(s) = 1/(1+\alpha s)$ and we have

$$\Pi(\omega) = \frac{1-\lambda\alpha}{1-\lambda\alpha\omega}. \tag{91}$$

Hence

$$P_j = (1-\lambda\alpha)(\lambda\alpha)^j. \tag{92}$$

2) $\chi_n \equiv \alpha$ (constant). Then $\psi(s) = e^{-s\alpha}$ and

$$\Pi(\omega) = \frac{(1-\lambda\alpha)e^{-\lambda\alpha(1-\omega)}}{1 - \dfrac{1-e^{-\lambda\alpha(1-\omega)}}{1-\omega}}. \tag{93}$$

Hence

$$P_j = (1-\lambda\alpha)\sum_{k=1}^{j,\,j-1} (-1)^{j-k} e^{k\lambda\alpha}\left(\frac{(k\lambda\alpha)^{j-k}}{(j-k)!} + \frac{(k\lambda\alpha)^{j-k-1}}{(j-k-1)!}\right). \tag{94}$$

§ 4. Waiting time problems in case of a non-Poisson process

Now, concerning the moments t_n of the arrivals of the persons at the service, let us suppose that the consecutive time differences $\tau_n = t_n - t_{n-1}$ $(n = 1, 2, \ldots;\ t_0 = 0)$ are equidistributed mutually independent positive random variables with the common distribution function $P(\tau_n \leqq x) = Q(x)$. In the case of a Poisson process homogeneous in time we had $Q(x) = 1-e^{-\lambda x}$ $(x \geqq 0)$. Now we shall deal with the more general case when $Q(x)$ is arbitrary. In the sequel we also suppose that the service times $\chi_1, \chi_2, \chi_3, \ldots$ are equidistributed mutually independent random variables with the common distribution function $H(x)$.

In Chapter 2 we have already mentioned that the random function $\eta(t)$ so defined is in general not a Markov process but the moments t_n are Markov points of the process. Let $\eta(t_n - 0) = \eta_n$, i. e. the waiting time of the

n-th arriving person, and $\eta_1(0) = \eta_0$. According to what has been said, the sequence $\{\eta_n\}$ of random variables is a Markov chain. Knowing η_0, the random variables η_n may be determined step by step from the following equation:

$$\eta_{n+1} = \begin{cases} \eta_n + \chi_n - \tau_{n+1} & \text{if} \quad \tau_{n+1} - \chi_n < \eta_n, \\ 0 & \text{if} \quad \tau_{n+1} - \chi_n \geqq \eta_n. \end{cases} \tag{95}$$

Let $P(\eta_n \leqq x) = F_n(x)$ denote the distribution function of the random variable η_n. Starting from $F_0(x)$, the sequence $\{F_n(x)\}$ may be determined step by step by means of the recurrence formula

$$F_{n+1}(x) = \int_0^\infty K(x, y)\, dF_n(y) \qquad (n = 0, 1, 2, \ldots) \tag{96}$$

where

$$K(x, y) = \int_0^\infty [1 - Q(u + y - x)]\, dH(u). \tag{97}$$

In fact, the transition probabilities of the Markov process described by the sequence $\{\eta_n\}$ of variables are $P(\eta_{n+1} \leqq x \mid \eta_n = y) = K(x, y)$, so that (96) holds. The computations may be perfected conveniently by the following two-step recurrence formulae:

$$K_n(x) = \int_0^\infty H(x - y)\, dF_n(y) \tag{98}$$

and

$$F_{n+1}(x) = \int_0^\infty [1 - Q(y - x)]\, dK_n(y). \tag{99}$$

A further question is: When does the limiting distribution function $\lim_{n \to \infty} F_n(x) = F(x)$ exist and how can $F(x)$ be determined? Here we quote the work by D. V. LINDLEY [14] in which the following theorem is proved:

For the existence of the limiting distribution function $\lim_{n \to \infty} F_n(x) = F(x)$ *it is necessary and sufficient that either* $M(\chi_n) \leqq M(\tau_n)$ *or* $\tau_n - \chi_n \equiv 0$ *holds. If* $M(\chi_n) \geqq M(\tau_n)$ *and* $\tau_n - \chi_n \not\equiv 0$, *then* $\lim_{n \to \infty} F_n(x) = 0$ *for every* x. *The limiting distribution* $F(x)$, *if it exists, is independent of the initial distribution* $F_0(x)$ *and is the unique solution of the integral equation*

$$F(x) = \int_0^\infty K(x, y)\, dF(y). \tag{100}$$

In his investigations LINDLEY starts from equation (95) and proves that it characterizes a random walk problem which can be discussed by making use of the laws of large numbers and recurrent events.

The theorem stated above may be proved by general theorems concerning Markov chains or it may be reduced to the proof of the ergodicity of the state E_0 with a finite recurrence time if $\alpha < \Theta$ where $\alpha = M(\chi_n)$ and $\Theta = M(\tau_n)$.

It is also interesting to investigate the following problem. If the process $\eta(t)$ is defined for all t as it was described in Chapter 2, then determine conditions for the existence of the limiting distribution function $\lim_{t\to\infty} F(t,x) = F^*(x)$ and a method to obtain it. Here we state the following

THEOREM 8. *If* $0 < \alpha < \Theta < \infty$ *and* $Q(x)$ *does not agree with a distribution function which has only jumps at some integral multiplies of a given number, then there exists the limiting distribution function* $\lim_{t\to\infty} F(t,x) = F^*(x)$, *independently of the initial distribution* $F_0(x)$ *and may be given by*

$$F^*(x) = \frac{1}{\Theta}\int_0^x [1 - Q(y)]K(y+x)\,dy, \tag{101}$$

where

$$K(x) = \int_0^\infty F(x-y)\,dH(y) \tag{102}$$

and $F(x)$ *is the unique solution of the integral equation* (100).

For the proof let us introduce the new variable $\zeta(t)$ which denotes the distance between the moment t and the immediately preceding moment t_n $(n = 0, 1, 2, \ldots)$. The process $\zeta(t)$ so defined essentially agrees with a special case of the renewal process investigated by J. L. DOOB [3]. Under our conditions we may apply Theorem 12 of DOOB. According to this theorem, the limiting distribution function

$$Q^*(x) = \lim_{t\to\infty} P(\zeta(t) \leqq x) = \frac{1}{\Theta}\int_0^x [1 - Q(u)]\,du \tag{103}$$

exists.

If $t \to \infty$ and t_n denotes the moment of arrival of the last person arriving in the interval $(0, t)$, then with probability one we have $n \to \infty$ and $\lim_{n\to\infty} P(\eta(t_n) \leqq x) = K(x)$, where $K(x)$ is given by (102). So it is clear that

$$F^*(x) = \int_0^x K(x+y)\,dQ^*(y) \tag{104}$$

which is the same as (101).

The uniqueness as well as the independence of the initial distribution $F_0(x)$ follow from the theorem stated for $F(x)$.

We remark that if $F^*(x)$ will be defined as the distribution function of $\eta(t)$ in the moment $t=t^*$ chosen "at random" in a process of an infinitely long time, then it suffices to suppose regarding $Q(x)$ only that $0<\Theta<\infty$, where Θ is its mean value.

EXAMPLE. Apart from the waiting times here we may also quote as example the filling and evacuating of cisterns. This example is taken from the book of G. P. BOEV [1] (p. 311, example 170). Let $\eta(t)$ denote the provision of the cistern in the moment t. Cistern are filled with periodes of a year. If we suppose that the quantity of leaving water is the same in the unit time until the evacuating follows, then we have a phenomenon which may be described by the above model. In the stationary case the distribution function of the provisions in the cistern will be given by $F^*(x)$ in (101).

If we suppose that the water leaving in the unit time is proportional to the provisions, then the model described in our paper [17] may be applied.

INSTITUTE FOR APPLIED MATHEMATICS
OF THE HUNGARIAN ACADEMY OF SCIENCES

(Received 1 December 1952)

Bibliography

[1] Г. П. Боев, Теория Вероятностей (Москва, 1950).
[2] G. DOETSCH, *Handbuch der Laplace Transformation*, Bd. 1 (Basel, 1949).
[3] J. L. DOOB, Renewal theory from the point of view of the theory of probability, *Transactions of the American Mathematical Society*, **63** (1948), pp. 422—438.
[4] W. FELLER, Zur Theorie der stochastischen Prozesse, *Mathematische Annalen*, **113** (1936), pp. 13—60.
[5] W. FELLER, On theory of stochastic processes with particular reference to applications, *Proceedings of the First Berkeley Symposium*, (1949), pp. 403—432.
[6] W. FELLER, Fluctuation theory of recurrent events, *Transactions of the American Mathematical Society*, **67** (1949), pp. 98—119.
[7] W. FELLER, *An introduction to probability theory and its applications* (New-York, 1950).
[8] TH. C. FRY, *Probability and its engineering uses* (New-York, 1928), pp. 372—387.
[9] А. Я. Хинчин, Математическая теория стационарной очереди, Математический Сборник, **39** (1932), pp. 73–84.
[10] D. G. KENDALL, Some problems in the theory of queues, *Journal of the Royal Statistical Society*, Ser. B, **13** (1951), pp. 151—185.
[11] A. N. KOLMOGOROV, Sur le problème d'attente, *Recueil Mathématique*, Математический Сборник, **38** (1931), pp. 101—106.
[12] A. N. KOLMOGOROV—B. V. GNEDENKO, *Független valószínűségi változók összegeinek határeloszlásai* (Budapest, 1951).
[13] D. V. LINDLEY, The theory of queues with a single counter, *Proceedings of the Cambridge Philosophical Society*, **48** (1952), pp. 277—289.

[14] F. POLLACZEK, Über eine Aufgabe der Wahrscheinlichkeitsrechnung, *Mathematische Zeitschrift*, **32** (1950), pp. 64—100; II, pp. 729—850.

[15] F. POLLACZEK, Répartition des délais d'attente des avions arrivant à un aéroport qui possède s pistes d'atterissages, *Comptes Rendus Acad. Sci. Paris*, **232** (1951), p. 1901.

[16] L. TAKÁCS, Occurrence and coincidence phenomena in case of happenings with arbitrary distribution law of duration, *Acta Math. Acad. Sci. Hung.*, **2** (1951), pp. 275—298.

[17] L. TAKÁCS, On stochastic processes connected with some physical recording apparatuses, *Acta Math. Acad. Sci. Hung.*, **6** (1955) (to be appear).

[18] A. ZYGMUND, A remark on characteristic functions, *Proceedings of the Second Berkeley Symposium*, (1951), pp. 369 372.

ИССЛЕДОВАНИЕ ПРОБЛЕМ ВРЕМЕНИ ОЖИДАНИЯ С ПОМОЩЬЮ МАРКОВСКИХ ПРОЦЕССОВ

Л. Такач (Будапешт)

(Резюме)

Лица прибывающие в моментах времени $t_1, t_2, \ldots, t_n, \ldots$ соответственно, обслуживаются в порядке очереди единственным приказчиком. Пусть времена обслуживания $\chi_1, \chi_2, \ldots, \chi_n, \ldots$ следующих друг за другом лиц являются независимыми положительными случайными величинами с одном и том же функцией распределения $H(x)$. Пусть $\Re(s) \geqq 0$

$$\psi(s) = \int_0^\infty e^{-sx}\, dH(x)$$

и $M(\chi_n) = \alpha$.

Обозначим через $\eta(t)$ время ожидания лица при условии что он прибывает в момент, и через η_n время ожидания лица прибывающего n-ым. Пусть далее $P(\eta(t) \leqq x) = F(t, x)$ и $P(\eta_n \leqq x) = F_n(x)$.

1. Пусть последовательность $\{t_n\}$ моментов прибытия образует процесс Пуассона с плотностью событий $\lambda(t)$. Тогда $\eta(t)$ является марковским процессом и $F(t, x)$ удовлетворяет следующему интегро-дифференциальному уравнению:

$$\frac{\partial F(t, x)}{\partial t} = \frac{\partial F(t, x)}{\partial x} - \lambda(t)\left[F(t, x) - \int_0^x H(x-y)\, d_y F(t, y)\right].$$

С помощью этого уравнения можно однозначно определять $F(t, x)$, зная $F(t, 0)$. (Нетрудно определять $F(t, 0)$, если $\lambda(t) \equiv \lambda$ (постоянна). Пусть $\lim_{t\to\infty} \lambda(t) = \lambda$; тогда $\lim_{t\to\infty} F(t, 0) = 1 - \lambda\alpha$ если $\lambda\alpha < 1$, и $\lim_{t\to\infty} F(t, 0) = 0$ если $\lambda\alpha \geqq 1$. В случае $\lambda\alpha < 1$ существует предельное распределение $\lim_{t\to\infty} F(t, x) = F^*(x)$, и для преобразованная Лапласа—Стильтьеса от $F^*(x)$ имеем

$$\int_0^\infty e^{-sx}\, dF^*(x) = \frac{1-\alpha}{1 - \lambda\dfrac{1-\psi(s)}{s}},$$

в соответствии решением заданным раньше Хинчиным для стационарного случая.

2. Пусть последовательность $\{t_n\}$ моментов прибытия образует процесс Пуассона с (постоянной) плотностью событий $\lambda(t) \equiv \lambda$. Тогда последовательность переменных $\{\eta_n\}$ образует однородную цепь Маркова, а функции распределения $F_n(x)$ определяются с помощью следующей рекуррентной формулы:

$$F_n(x) = \int_0^\infty K(x, y)\, dF_{n-1}(y) \qquad (n = 2, 3, \ldots),$$

где

$$K(x, y) = H(x-y) + \int_{x-y}^\infty e^{-\lambda(u+y-x)}\, dH(u).$$

Если $n \to \infty$, то $F_n(x) \to F^*(x)$.

Пусть $G(x)$ — функция распределения периода обслуживания, и пусть

$$\Gamma(s) = \int_0^\infty e^{-sx}\, dG(x).$$

Тогда $\Gamma(s)$ является решением функционального уравнения

$$\Gamma(s) = \psi(s + \lambda - \lambda\Gamma(s)),$$

удовлетворяющим условию $\Gamma(\infty) = 0$ и аналитическим для $\Re(s) \geqq 0$. Пусть p^* — найменьшее положительное число, для которого $p^* = \psi(\lambda(1-p^*))$, тогда $\lim_{x\to\infty} G(x) = p^*$. Если $\lambda\alpha \leqq 1$, то $p^* = 1$, а если $\lambda\alpha > 1$, то $0 < p^* < 1$.

Пусть f_j вероятность того, что период обслуживания состоит из j отдельных актов обслуживания, и пусть $F(\omega) = \sum_{j=1}^\infty f_j\, \omega^j$. Тогда $F(\omega)$ является решением функционального уравнения

$$F(\omega) = \omega\, \psi(\lambda - \lambda F(\omega)),$$

удовлетворяющим условию $F(0) = 0$ и аналитическим для $|\omega| \leqq 1$. Далее, имеет место соотношение $\lim_{\omega\to 1} F(\omega) = p^*$, где значение p^* то же как и прежде.

Пусть P_j — вероятность того, что при $t \to \infty$ число ожидающих равняется j. Пусть $\Pi(\omega) = \sum_{j=0}^\infty P_j\, \omega^j$; тогда для $\lambda\alpha < 1$ имеет место

$$\Pi(\omega) = \frac{(1-\lambda\alpha)\,\psi(\lambda(1-\omega))}{1 - \dfrac{1-\psi(\lambda(1-\omega))}{1-\omega}}.$$

3. На конце статьи делаются некоторые замечания относительно случая, когда разности $t_n - t_{n-1}$ одинаково распределенные независимые случайные величины.

LÖSUNG DER VEKTOR-FUNKTIONALGLEICHUNG DER HOMOGENEN UND INHOMOGENEN n-DIMENSIONALEN EINPARAMETRIGEN „TRANSLATION", DER ERZEUGENDEN FUNKTION VON KETTENREAKTIONEN UND DES STATIONÄREN UND NICHT-STATIONÄREN BEWEGUNGSINTEGRALS

Von
J. ACZÉL (Debrecen)
(*Vorgelegt von* A. RÉNYI)

§ 1

Die folgende Vektor-Funktionalgleichung der n-dimensionalen einparametrigen (homogenen) Translation (s. z. B. [1][1]), d. h. der n-dimensionalen Transformationen mit einem additiven Parameter (auf die — vgl. § 4 — unter gewissen Voraussetzungen durch Einführung eines additiven Parameters jede einparametrige Transformation zurückgeführt werden kann), ist zugleich die Funktionalgleichung der erzeugenden Funktion von Kettenreaktionen im homogenen Fall [2] und auch die Integrale der Vektor-Differentialgleichung (d. h. des Differentialgleichungssystems) einer stationären Bewegung erfüllen diese Gleichung (s. z. B. [3]):

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u] = \mathbf{f}(\mathbf{x}, t+u). \tag{1}$$

(Übrigens folgen auch die Iterierten $\mathbf{f}_v(\mathbf{x}) = \mathbf{f}(\mathbf{x}, v)$ einer Vektorfunktion dieser Gleichung, vgl. § 4.)

Die bezügliche Gleichung für die erzeugende Funktion der inhomogenen Kettenreaktionen [2] lautet

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, s, t), t, u] = \mathbf{f}(\mathbf{x}, s, u). \tag{2}$$

Man sieht auch leicht, daß die Integrale des Differentialgleichungssystems (der Vektor-Differentialgleichung) einer nicht-stationären Bewegung

$$\frac{d}{dt}\mathbf{f}(\mathbf{x}, s, t) = \mathbf{v}(\mathbf{f}, t)$$

[1] Eckige Klammern verweisen auf das Literaturverzeichnis am Ende der Arbeit.

der Gleichung (2) genüge leisten. Es scheint auch zweckmäßig zu sein, als Verallgemeinerung der gewöhnlichen Translationen die der Gleichung (2) genügenden Zuordnungen „inhomogene Translationen" zu nennen.

In dem § 3 werden wir Folgerungen bezüglich der oben vorgeführten Probleme (z. B. die Zurückführbarkeit der einparametrigen Transformationen auf koordinaten-parallele Verschiebungen, usw.) aus der in dem § 2 vollbrachten Lösung der Funktionalgleichungen (1), (2) ziehen.

Im eindimensionalen Fall wurde die Gleichung (1) und die ähnliche Gleichung für nicht-additiven Parameter schon in den Arbeiten [4], [5] gelöst.

Die Voraussetzungen über die Natur der Funktionen $\mathbf{f}$ werden wir so wählen, daß die Beweise möglichst einfach und klar werden, deshalb werden wir nicht immer die schwächsten Bedingungen suchen. Es ist aber bemerkenswert, daß wir die Gültigkeit unserer Funktionalgleichungen nur auf einer $(n-1)$-dimensionalen Hyperebene voraussetzen und hieraus ihre Gültigkeit schon im ganzen Raum folgt.

§ 2

Bezüglich der Gleichung (1) beweisen wir den

SATZ 1. *Die Vektorfunktionen*

$$\mathbf{f}(\mathbf{x}, t) = \mathbf{g}[\mathbf{h}(\mathbf{x}) + t], \tag{3}$$

wo $\mathbf{h}(\mathbf{x})$ *eine beliebige ein-eindeutige Vektorfunktion mit der Umkehrfunktion* $\mathbf{g}$ *ist, also*

$$\mathbf{h}[\mathbf{g}(\mathbf{y})] = \mathbf{y}, \tag{4}$$

und die Addition $\mathbf{y} + t$ *zwischen Vektoren und Skalaren als Addition des Skalars zur ersten Komponenten des Vektors:*

$$\mathbf{y} + t = \{y_1, y_2, \ldots, y_n\} + t = \{y_1 + t, y_2, \ldots, y_n\}$$

definiert wird, erfüllen alle die Gleichung

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u] = \mathbf{f}(\mathbf{x}, t + u). \tag{1}$$

Wird umgekehrt (1) *auf der Koordinatenhyperebene*[2] $\mathbf{x} = \mathbf{x}_a = \{a, x_2, \ldots, x_n\}$ *(a konstant) vorausgesetzt, und gibt es ferner für jedes* $\mathbf{x}$ *ein*

$$\mathbf{x}_t = \{t, x_2, \ldots, x_n\} = \mathbf{h}(\mathbf{x}) \tag{5}$$

derart, daß

$$\mathbf{g}(\mathbf{x}_t) = \mathbf{f}(\mathbf{x}_a, t) = \mathbf{x} \tag{6}$$

ist, so folgt

$$\mathbf{f}(\mathbf{x}, t) = \mathbf{g}[\mathbf{h}(\mathbf{x}) + t] \tag{3}$$

(und die Gültigkeit der Gleichung (1)*) für jedes* $\mathbf{x}$ (und t, u).

[2] Natürlich kann statt x_1 auch x_2 usw. konstant gehalten werden, wenn dies die Erfüllung der übrigen Voraussetzungen besser sichert.

Die oben stehende Definition der Addition von Vektoren und Skalaren mag etwas seltsam scheinen, sie leistet aber hier guten Dienst (vgl. auch § 3). Auch die übliche Addition von komplexen Zahlen bzw. Quaternionen mit Skalaren (reellen Zahlen) entspricht dieser Vorschrift.

BEWEIS. Daß die Funktion (3) die Gleichung (1) erfüllt, sieht man unmittelbar:

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u] = \mathbf{g}[\mathbf{h}(\mathbf{f}(\mathbf{x}, t)) + u] = \mathbf{g}(\mathbf{h}[\mathbf{g}(\mathbf{h}(\mathbf{x}) + t)] + u) = \\ = \mathbf{g}[\mathbf{h}(\mathbf{x}) + t + u] = \mathbf{f}(\mathbf{x}, t + u),$$

wo auch auf (4) Rücksicht genommen wurde.

Gilt umgekehrt (1) für $\mathbf{x} = \mathbf{x}_a$ (und alle t, u):

$$\mathbf{f}[\mathbf{f}(\mathbf{x}_a, t), u] = \mathbf{f}(\mathbf{x}_a, t + u),$$

dann ist wegen (5)—(6) und der Definition von $\mathbf{x}_t + u = \mathbf{x}_{t+u}$:

$$\mathbf{f}[\mathbf{g}(\mathbf{x}_t), \mathbf{u}] = \mathbf{g}(\mathbf{x}_t + u),$$

also

$$\mathbf{f}(\mathbf{x}, u) = \mathbf{g}[\mathbf{h}(\mathbf{x}) + u]$$

für jedes $\mathbf{x}$ gültig, w. z. b. w.

Der entsprechende Satz für die Gleichung (2) lautet:

SATZ 2. *Die Funktionen*

(7) $$\mathbf{f}(\mathbf{x}, t, u) = \mathbf{g}[\mathbf{h}(\mathbf{x}, t), u]$$

(wo $\mathbf{y} = \mathbf{h}(\mathbf{x}, t)$ *eine beliebige Vektorfunktion mit dem Parameter* t *ist, die sich bezüglich* x *umkehren läßt:* $\mathbf{x} = \mathbf{g}(\mathbf{y}, t)$, *also*

(8) $$\mathbf{h}[\mathbf{g}(\mathbf{y}, t), t] = \mathbf{y}$$

gilt) erfüllen immer die Gleichung

(2) $$\mathbf{f}[\mathbf{f}(\mathbf{x}, s, t), t, u] = \mathbf{f}(\mathbf{x}, s, u).$$

Wird umgekehrt (2) *auf der Koordinatenhyperebene* $\mathbf{x} = \mathbf{x}_a$ *erfüllt und gibt es für jedes* $\mathbf{x}$ *und* t *ein*

(9) $$\mathbf{x}_s = \{s, x_2, \ldots, x_n\} = \mathbf{h}(\mathbf{x}, t)$$

derart, daß

(10) $$\mathbf{g}(\mathbf{x}_s, t) = \mathbf{f}(\mathbf{x}_a, s, t) = \mathbf{x}$$

ist, so gilt

(7) $$\mathbf{f}(\mathbf{x}, t, u) = \mathbf{g}[\mathbf{h}(\mathbf{x}, t), u]$$

(und deshalb auch die Gleichung (2)*) für jedes* $\mathbf{x}$ (und t, u, s).

Es ist bemerkenswert, daß hier gar keine neue Operation zwischen Vektoren und Skalaren definiert werden mußte.

BEWEIS. Die Funktionen der Gestalt (7) erfüllen offenbar die Gleichung (2):

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, s, t), t, u] = \mathbf{g}(\mathbf{h}[\mathbf{f}(\mathbf{x}, s, t), t], u) = \mathbf{g}[\mathbf{h}(\mathbf{g}[\mathbf{h}(\mathbf{x}, s), t], t), u] = \\ = \mathbf{g}[\mathbf{h}(\mathbf{x}, s), u] = \mathbf{f}(\mathbf{x}, s, u)$$

mit Rücksicht auf (8).

Ist umgekehrt (2) für $\mathbf{x} = \mathbf{x}_a$ (und alle s, t, u) erfüllt:

$$\mathbf{f}[\mathbf{f}(\mathbf{x}_a, s, t), t, u] = \mathbf{f}(\mathbf{x}_a, s, u),$$

so gilt wegen (9)—(10)

$$\mathbf{f}[\mathbf{g}(\mathbf{x}_s, t), t, u] = \mathbf{g}(\mathbf{x}_s, u),$$

also

$$\mathbf{f}(\mathbf{x}, t, u) = \mathbf{g}[\mathbf{h}(\mathbf{x}, t), u]$$

für jedes $\mathbf{x}$, w. z. b. w.

Die Frage nach der Eindeutigkeit der Darstellungen (7) und (3) läßt sich folgenderweise erledigen:

SATZ 3. *Ist*

$$\mathbf{g}[\mathbf{h}(\mathbf{x}, t), u] = \bar{\mathbf{g}}[\bar{\mathbf{h}}(\mathbf{x}, t), u], \tag{11}$$

wo $\mathbf{h}$ *bzw.* $\bar{\mathbf{h}}$ *die einzigen Umkehrfunktionen von* $\mathbf{g}$ *bzw.* $\bar{\mathbf{g}}$ *sind, so gilt*

$$\bar{\mathbf{g}}(\mathbf{y}, u) = \mathbf{g}[\mathbf{k}(\mathbf{y}), u] \tag{12}$$

$$(\bar{\mathbf{h}}(\mathbf{x}, u) = \mathbf{l}[\mathbf{h}(\mathbf{x}, u)], \quad wo \quad \mathbf{l} = \mathbf{k}^{-1}, \quad d.h. \quad \mathbf{k}[\mathbf{l}(\mathbf{x})] = \mathbf{x})$$

und diese erfüllen auch (11). *D. h. in* (7) *kann* $\mathbf{g}$ *durch die in* (12) *gegebenen* $\bar{\mathbf{g}}$ *und nur durch diese ersetzt werden, also ist die Funktion* $\mathbf{g}$ *in* (7) *nur bis auf eine n-dimensionale Vektorfunktion* $\mathbf{k}$ *bestimmt.*

BEWEIS. Aus (11) folgt, wenn wir $\bar{\mathbf{h}}(\mathbf{x}, t) = \mathbf{y}$, $\mathbf{x} = \bar{\mathbf{g}}(\mathbf{y}, t)$ schreiben mit $t = t_0$:

$$\bar{\mathbf{g}}(\mathbf{y}, u) = \mathbf{g}(\mathbf{h}[\bar{\mathbf{g}}(\mathbf{y}, t_0), t_0], u) = \mathbf{g}[\mathbf{k}(\mathbf{y}), u].$$

Hieraus folgt auch

$$\bar{\mathbf{h}}(\mathbf{x}, u) = \mathbf{l}[\mathbf{h}(\mathbf{x}, u)],$$

und diese Funktionen erfüllen auch wirklich (11):

$$\bar{\mathbf{g}}[\bar{\mathbf{h}}(\mathbf{x}, t), u] = \mathbf{g}(\mathbf{k}[\bar{\mathbf{h}}(\mathbf{x}, t)], u) = \mathbf{g}[\mathbf{k}(\mathbf{l}[\mathbf{h}(\mathbf{x}, t)]), u] = \mathbf{g}[\mathbf{h}(\mathbf{x}, t), u],$$

w. z. b. w.

Im homogenen Fall gilt ähnlich der

SATZ 4. *Aus*

$$\mathbf{g}[\mathbf{h}(\mathbf{x}) + t] = \bar{\mathbf{g}}[\bar{\mathbf{h}}(\mathbf{x}) + t] \tag{13}$$

folgt

$$\bar{\mathbf{g}}(\mathbf{y}) = \mathbf{g}[\varkappa(\mathbf{y}^{(n-1)}) + y_1], \tag{14}$$

(wo die Bezeichnungen $\mathbf{y} = \{y_1, y_2, \ldots, y_n\}$, $\mathbf{y}^{(n-1)} = \{y_2, \ldots, y_n\}$ *verwendet wurden), und diese Funktion erfüllt mit ihrer Umkehrfunktion* $\bar{\mathbf{h}}(\mathbf{x})$ *die Gleichung*

(13) *tatsächlich. D. h. in* (3) *kann* $\mathbf{g}$ *durch die in* (14) *gegebenen* $\bar{\mathbf{g}}$ *und nur durch diese ersetzt werden, also ist die Funktion* $\mathbf{g}$ *in* (3) *nur bis auf eine n-dimensionale Vektorfunktion eines* $(n-1)$*-dimensionalen Vektors (d. h. n Funktionen von* $(n-1)$ *Veränderlichen) bestimmt.*

Der BEWEIS dieses Satzes ist merkwürdigerweise mehr verwickelt als der des Satzes 3. (14) folgt aus (13) folgenderweise: Wir bezeichnen in (13) $\bar{\mathbf{h}}(\mathbf{x})+t=\mathbf{y}$, $\mathbf{x}=\bar{\mathbf{g}}(\mathbf{y}-t)$, dann gilt $\bar{\mathbf{g}}(\mathbf{y})=\mathbf{g}(\mathbf{h}[\bar{\mathbf{g}}(\mathbf{y}-t)]+t)$, wo t noch beliebig gewählt werden kann. Wählen wir $t=y_1$, so erhalten wir tatsächlich

$$\bar{\mathbf{g}}(\mathbf{y})=\mathbf{g}(\mathbf{h}[\bar{\mathbf{g}}(\mathbf{y}^{(n-1)})]+y_1)=\mathbf{g}[\varkappa(\mathbf{y}^{(n-1)})+y_1].$$

Aber schon die Konstruktion der Umkehrfunktion $\bar{\mathbf{h}}(\mathbf{x})$ von $\bar{\mathbf{g}}(\mathbf{y})$ ist nicht mehr trivial:

Aus

$$\mathbf{x}=\bar{\mathbf{g}}(\mathbf{y})=\mathbf{g}[\varkappa(\mathbf{y}^{(n-1)})+y_1]=\mathbf{g}\left[\begin{pmatrix}\varkappa_1(\mathbf{y}^{(n-1)})+y_1\\ \varkappa^{(n-1)}(\mathbf{y}^{(n-1)})\end{pmatrix}\right],\quad \mathbf{y}=\bar{\mathbf{h}}(\mathbf{x})$$

folgt

$$\varkappa(\mathbf{y}^{(n-1)})+y_1=\mathbf{h}(\mathbf{x}),$$

d. h.

$$\varkappa_1(\mathbf{y}^{(n-1)})+y_1=h_1(\mathbf{x}),$$
$$\varkappa^{(n-1)}(\mathbf{y}^{(n-1)})=\mathbf{h}^{(n-1)}(\mathbf{x})$$

und falls wir die Umkehrfunktion von $\mathbf{x}^{(n-1)}=\varkappa^{(n-1)}(\mathbf{y}^{(n-1)})$ mit $\mathbf{y}^{(n-1)}=\lambda^{(n-1)}(\mathbf{x}^{(n-1)})$ bezeichnen, so wird

$$\begin{cases}\bar{\mathbf{h}}^{(n-1)}(\mathbf{x})=\mathbf{y}^{(n-1)}=\lambda^{(n-1)}[\mathbf{h}^{(n-1)}(\mathbf{x})],\\ \bar{h}_1(\mathbf{x})=y_1=h_1(\mathbf{x})-\varkappa_1(\mathbf{y}^{(n-1)})=h_1(\mathbf{x})-\varkappa_1(\lambda^{(n-1)}[\mathbf{h}^{(n-1)}(\mathbf{x})]\end{cases}$$

die gesuchte Umkehrfunktion von (14) sein. Nun erfüllt sie aber zusammen mit (14) die Gleichung (13) tatsächlich

$$\bar{\mathbf{g}}[\bar{\mathbf{h}}(\mathbf{x})+t]=\bar{\mathbf{g}}\left[\begin{pmatrix}\bar{\mathbf{h}}_1(\mathbf{x})+t\\ \bar{\mathbf{h}}^{(n-1)}(\mathbf{x})\end{pmatrix}\right]=\bar{\mathbf{g}}\left[\begin{pmatrix}h_1(\mathbf{x})-\varkappa_1(\mathbf{y}^{(n-1)})+t\\ \lambda^{(n-1)}[\mathbf{h}^{(n-1)}(\mathbf{x})]\end{pmatrix}\right]=$$
$$=\mathbf{g}\left[\begin{pmatrix}\varkappa_1(\lambda^{(n-1)}[\mathbf{h}^{(n-1)}(\mathbf{x})])+h_1(\mathbf{x})-\varkappa_1(\lambda^{(n-1)}[\mathbf{h}^{(n-1)}(\mathbf{x})])+t\\ \varkappa^{(n-1)}(\lambda^{(n-1)}[\mathbf{h}^{(n-1)}(\mathbf{x})])\end{pmatrix}\right]=$$
$$=\mathbf{g}\left[\begin{pmatrix}h_1(\mathbf{x})+t\\ \mathbf{h}^{(n-1)}(\mathbf{x})\end{pmatrix}\right]=\mathbf{g}[\mathbf{h}(\mathbf{x})+t],$$

(da $\varkappa^{(n-1)}[\lambda^{(n-1)}(\mathbf{x}^{(n-1)})]=\mathbf{x}^{(n-1)}$ ist), w. z. b. w.

§ 3

Die Lösungen (3) bzw. (7) der Gleichungen (1) bzw. (2) bedeuten, daß *mittels der Koordinatentransformation*

$$\bar{\mathbf{x}}=\mathbf{h}(\mathbf{x}),\quad \bar{\mathbf{y}}=\mathbf{h}(\mathbf{y}),$$

bzw. der Zuordnung des beweglichen Punktes

$$\bar{\mathbf{x}}(t)=\mathbf{h}(\mathbf{x},t)$$

zum Punkte $\mathbf{x}$ *(und ähnlich* $\overline{\mathbf{y}}(t) = \mathbf{h}(\mathbf{y}, t)$ *zu* $\mathbf{y}$*) die Transformationen mit einem additiven Parameter*

$$\mathbf{y} = \mathbf{f}(\mathbf{x}, t),$$

bzw. die „inhomogenen Translationen"

$$\mathbf{y} = \mathbf{f}(\mathbf{x}, t, u)$$

sich in die zur x_1 *Achse parallelen Verschiebung*

$$\overline{\mathbf{y}} = \overline{\mathbf{x}} + t,$$

bzw. in die Transformation

$$\overline{\mathbf{y}}(u) = \overline{\mathbf{x}}(t) \tag{15}$$

überführen lassen, falls die Transformationen $\mathbf{f}(\mathbf{x}, t)$ *bzw.* $\mathbf{f}(\mathbf{x}, t, u)$ *auf der Koordinatenhyperebene* $x_1 = a$ *eine Schar bilden, d. h. den Gleichungen* (1) *bzw.* (2) *genügen und dort auch die Auflösbarkeitsbedingungen* (5)—(6) *bzw.* (9)—(10) *erfüllt sind (und dann bleiben die Schareigenschaften* (1) *bzw.* (2) *im ganzen Raum erhalten).*

(15) läßt sich so deuten, daß der zu $\mathbf{x}$ geordnete bewegliche Punkt im Zeitpunkt u eben dort ist, wo der zu $\mathbf{y} = \mathbf{f}(\mathbf{x}, t, u)$ geordnete im Zeitpunkt t. Der homogene Fall schließt sich hierin durch bewegliche Punkte mit axenparallelen Bahnen ein.

Bei unserer Deutung gibt der Satz 4 aus einer geeigneten Koordinatentransformation alle solche an und Ähnliches gilt auch vom Satz 3.

Für die wahrscheinlichkeitstheoretische Anwendung unserer Gleichungen ist die Deutung unserer Resultate weniger an der Hand liegend. Jedenfalls kann bemerkt werden, daß die erzeugenden Funktionen der Wahrscheinlichkeiten bei Kettenreaktionen dem Werte-n-tuppel $(1, 1, \ldots, 1)$ immer den Funktionenwert $(1, 1, \ldots, 1)$ zuordnen, deshalb wird bei den oben beschriebenen Zuordnungen dem Einheitspunkt immer ein bezüglich der neuen Transformation fixer Punkt entsprechen. Ähnliches gilt oft auch bezüglich des Anfangspunktes $(0, 0, \ldots, 0)$.

Wir wenden uns endlich den stationären und nicht-stationären Bewegungen zu. Wie bekannt ([3]), erfüllen die Ortsvektoren eines beweglichen Punktes $\mathbf{f}(\mathbf{x}, t)$ mit dem Anfangspunkt $\mathbf{x}$, die der Differentialgleichung der stationären Bewegung

$$\frac{d\mathbf{f}(\mathbf{x}, t)}{dt} = \mathbf{v}[\mathbf{f}(\mathbf{x}, t)] \tag{16}$$

genügen, die Gleichung (1). Ähnlich erfüllen die Ortsvektoren des beweglichen Punktes $\mathbf{f}(\mathbf{x}, s, t)$, der sich zur Ausgangszeit s am Ort $\mathbf{x}$ befindet, falls die Bewegung nichtstationär ist (also der Gleichung

$$\frac{d\mathbf{f}(\mathbf{x}, s, t)}{dt} = \mathbf{v}[\mathbf{f}(\mathbf{x}, s, t), t] \tag{17}$$

genügt), die Gleichung (2). Den Zusammenhang dieser Funktionen $\mathbf{v}$ mit unseren Vektorfunktionen geben die folgenden zwei Sätze an.

SATZ 5. *Die Funktionen* $\mathbf{v}(\mathbf{x})$ *in der Differentialgleichung der stationären Bewegung*

$$\frac{d\mathbf{f}(\mathbf{x}, t)}{dt} = \mathbf{v}[\mathbf{f}(\mathbf{x}, t)], \tag{16}$$

(deren Lösungen $\mathbf{f}(\mathbf{x}, t)$ *der Gleichung* (1) *genüge leisten), lassen sich (mittels* $\mathbf{f}$ *bzw.) mittels der im Ausdruck* (3) *von* $\mathbf{f}$ *figurierenden Funktion* $\mathbf{h}$ *folgenderweise ausdrücken:*

$$\mathbf{v}(\mathbf{x}) = \left.\frac{d\mathbf{f}(\mathbf{x}, t)}{dt}\right|_{t=0} = \mathbf{H}'(\mathbf{x})^{-1}\mathbf{e},$$

wo $\mathbf{e}$ *der Vektor* $e_1 = 1, e_2 = \cdots = e_n = 0$ *ist und* $\mathbf{H}'(\mathbf{x})$ *die als regulär vorausgesetzte derivierte Matrix der Vektorfunktion* $\mathbf{h}(\mathbf{x})$ *und* $\mathbf{H}'(\mathbf{x})^{-1}$ *ihre Inverse ist.*

BEWEIS. Aus (3) folgt, wie wir schon sahen,

$$\mathbf{h}[\mathbf{f}(\mathbf{x}, t)] = \mathbf{h}(\mathbf{x}) + t.$$

Derivieren wir diese Gleichung nach t, so können wir

$$\mathbf{H}'[\mathbf{f}(\mathbf{x}, t)]\frac{d\mathbf{f}(\mathbf{x}, t)}{dt} = \mathbf{e} \tag{18}$$

schreiben. Aus (3) (oder auch aus (1)) folgt aber

$$\mathbf{f}(\mathbf{x}, 0) = \mathbf{x}$$

und aus (16) erhalten wir für $t = 0$

$$\mathbf{v}(\mathbf{x}) = \left.\frac{d\mathbf{f}(\mathbf{x}, t)}{dt}\right|_{t=0},$$

folglich wird aus (18) für $t = 0$

$$\mathbf{v}(\mathbf{x}) = \mathbf{H}'(\mathbf{x})^{-1}\mathbf{e},$$

w. z. b. w.

SATZ 6. *Die Funktionen* $\mathbf{v}(\mathbf{x}, t)$ *von* (17) *stehen mit den Funktionen* $\mathbf{f}(\mathbf{x}, s, t)$ *und den Funktionen* $\mathbf{h}(\mathbf{x}, t)$ *von* (7) *im folgenden Zusammenhang*

$$\mathbf{v}(\mathbf{x}, t) = \left.\frac{d\mathbf{f}(\mathbf{x}, s, t)}{dt}\right|_{s=t} = -\mathbf{H}'(\mathbf{x}, t)^{-1}\frac{d\mathbf{h}(\mathbf{x}, t)}{dt},$$

wo $\mathbf{H}'$ *wieder die derivierte Matrix von* $\mathbf{h}(\mathbf{x}, t)$ *bezüglich* $\mathbf{x}$ *bezeichnet.*

BEWEIS. Aus (7) (oder (2)) folgt

$$\mathbf{f}(\mathbf{x}, t, t) = \mathbf{x}, \tag{19}$$

und aus

$$\mathbf{v}[\mathbf{f}(\mathbf{x}, s, t), t] = \frac{d\mathbf{f}(\mathbf{x}, s, t)}{dt} \tag{17}$$

für $s = t$

$$\mathbf{v}(\mathbf{x}, t) = \left.\frac{d\mathbf{f}(\mathbf{x}, s, t)}{dt}\right|_{s=t}. \tag{20}$$

Andererseits derivieren wir wieder die mit (7) äquivalente Gleichung

$$\mathbf{h}[\mathbf{f}(\mathbf{x}, s, t), t] = \mathbf{h}(\mathbf{x}, s)$$

nach t

$$\mathbf{H}'[\mathbf{f}(\mathbf{x}, s, t), t] \frac{d\mathbf{f}(\mathbf{x}, s, t)}{dt} + \frac{d\mathbf{h}(\mathbf{y}, t)}{dt}\Big|_{\mathbf{y}=\mathbf{f}(\mathbf{x}, s, t)} = 0$$

und substituieren $s = t$, dann wird mit Rücksicht auf (19) und (20)

$$\mathbf{v}(\mathbf{x}, t) = -\mathbf{H}'(\mathbf{x}, t)^{-1} \frac{d\mathbf{h}(\mathbf{x}, t)}{dt}.$$

Dies und (20) ergeben die Behauptung des Satzes 6.

Die Sätze 6 bzw. 5 lassen sich für die Komponenten folgenderweise aussprechen:

$$v_k = -\frac{\Delta_k}{\Delta}, \quad wo \quad \Delta = \frac{\partial(h_1, \ldots, h_n)}{\partial(x_1, \ldots, x_n)} \quad und \quad \Delta_k = \frac{\partial(h_1, \ldots, h_{k-1}, h_k, h_{k+1}, \ldots, h_n)}{\partial(x_1, \ldots, x_{k-1}, t, x_{k+1}, \ldots, x_n)}$$

bzw.

$$\Delta_k = \frac{\partial(h_1 - t, h_2, \ldots, h_{k-1}, h_k, h_{k+1}, \ldots, h_n)}{\partial(x_1, x_2, \ldots, x_{k-1}, t, x_{k+1}, \ldots, x_n)}$$

ist.

§ 4

Die bisherigen Interpretationen unserer Gleichungen liefen parallel im homogenen und inhomogenen Fall. Wir schließen mit zwei Bemerkungen die sich nur auf die homogene Gleichung (1) beziehen.

Erstens sei bemerkt, daß auch *die iterierten* $\mathbf{f}_t(\mathbf{x}) = \mathbf{f}(\mathbf{x}, t)$ *einer Vektorfunktion* $\mathbf{f}(\mathbf{x})$, d. h. eines Funktionen-n-tuppels $f_1(x_1, \ldots, x_n), f_2(x_1, \ldots, x_n), \ldots, f_n(x_1, x_2, \ldots . x_n)$, *der Gleichung* (1) *genügen,* da man von den Iterierten die Erfüllung der Relation

$$\mathbf{f}_u[\mathbf{f}_t(\mathbf{x})] = \mathbf{f}_{t+u}(\mathbf{x})$$

zu fordern pflegt.

Zweitens sei darauf hingewiesen wie sich die Transformationen mit einem nicht-additiven Parameter auf den Fall mit additivem Parameter, d. h. auf die „homogenen Translationen" der Gleichung (1) zurückführen lassen.

SATZ 7. *Falls die n-dimensionalen, einparametrigen Transformationen* $f(x, t)$ *mit einem (nicht-additiven) Parameter eine Schar in dem Sinne*

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u] = \mathbf{f}(\mathbf{x}, t \circ u) \tag{21}$$

bilden, [wo $t \circ u$ *eine stetige rechts kürzbare Operation (skalare Funktion) von* t *und* u *ist], und* $\mathbf{f}(\mathbf{x}, t)$ *effektiv ist, d. h. aus* $\mathbf{f}(\mathbf{x}, t) \equiv \mathbf{f}(\mathbf{x}, u)$ $t = u$ *folgt, so kann ein additiver Parameter eingeführt werden, d. h. es existiert eine (skalare) stetige und streng monotone Funktion* $\varphi(t)$ *derart, daß die neue Transformation*

$$\bar{\mathbf{f}}(\mathbf{x}, t) = \mathbf{f}[\mathbf{x}, \varphi(t)] \tag{22}$$

schon die Gleichung

$$\bar{\mathbf{f}}[\bar{\mathbf{f}}(\mathbf{x}, t), u] = \bar{\mathbf{f}}(\mathbf{x}, t+u)$$

der homogenen Translation erfüllt.

BEWEIS. Aus (21) folgt

$$\mathbf{f}[\mathbf{x}, (s \circ t) \circ u] = \mathbf{f}[\mathbf{f}(\mathbf{x}, s \circ t), u] = \mathbf{f}(\mathbf{f}[\mathbf{f}(\mathbf{x}, s), t], u) = \\ = \mathbf{f}[\mathbf{f}(\mathbf{x}, s), t \circ u] = \mathbf{f}[\mathbf{x}, s \circ (t \circ u)],$$

was wegen der Effektivität von $\mathbf{f}$ die *Assoziativität*

$$(s \circ t) \circ u = s \circ (t \circ u)$$

von $t \circ u$ bedeutet. Es ist ebenfalls aus der Effektivität zu sehen, daß aus $t \circ u_1 = t \circ u_2$ wegen

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u_1] = \mathbf{f}(\mathbf{x}, t \circ u_1) = \mathbf{f}(\mathbf{x}, t \circ u_2) = \mathbf{f}[\mathbf{f}(\mathbf{x}, t), u_2]$$

$u_1 = u_2$ folgt, also $t \circ u$ auch von *links kürzbar* ist. Da nun, wie wir gleich beweisen werden, jede stetige assoziative und beiderseitig kürzbare Operation (Funktion) $t \circ u$ in der Gestalt

$$t \circ u = \varphi[\varphi^{-1}(t) + \varphi^{-1}(u)] \tag{23}$$

darstellbar ist, folgt aus (21)

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u] = \mathbf{f}(\mathbf{x}, \varphi[\varphi^{-1}(t) + \varphi^{-1}(u)]),$$

also mit der Bezeichnung (22) und $\varphi^{-1}(t) = v$, $\varphi^{-1}(u) = w$

$$\bar{\mathbf{f}}[\bar{\mathbf{f}}(\mathbf{x}, v), w] = \bar{\mathbf{f}}(\mathbf{x}, v+w),$$

w. z. b. w.

Die noch übrig gebliebene Verifikation von (23) folgt aus dem

SATZ 8. *Ist die Operation $t \circ u$ zwischen zwei reellen Zahlen t, u eines Intervalles stetig, assoziativ und beiderseitig kürzbar, so gibt es eine stetige, streng monotone Funktion $\varphi(t)$ derart, daß*

$$t \circ u = \varphi[\varphi^{-1}(t) + \varphi^{-1}(u)], \tag{23}$$

(φ^{-1} ist die inverse Funktion von φ).

Der BEWEIS benutzt den Satz (s. [6]), daß für jede stetige *streng monoton wachsende* assoziative Operation (23) gilt. Wir müssen also noch beweisen, daß *aus unseren Voraussetzungen das streng monotone Wachsen von $t \circ u$ in beiden Veränderlichen folgt.* Die strenge Monotonität selbst folgt natürlich aus der Stetigkeit und beiderseitigen Kürzbarkeit von $t \circ u$, wir müssen aber die Möglichkeit des streng monotonen Abnehmens ausschließen. Gäbe es z. B. ein u_0 derart, daß für jedes Paar $t_1 < t_2$

$$t_1 \circ u_0 - t_2 \circ u_0 > 0 \tag{24}$$

wäre, so hat jedes u dieselbe Eigenschaft, denn falls es noch ein u gebe für das

$$t_1 \circ u - t_2 \circ u < 0,$$

so gäbe es wegen der Stetigkeit auch ein u_1 zwischen u_0 und u mit

$$t_1 \circ u_1 - t_2 \circ u_1 = 0,$$

was der rechtsseitigen Kürzbarkeit widersprechen würde. Gibt es also ein u_0, derart, daß für jedes Paar $t_1 < t_2$ (24) gilt, so folgt für jedes u aus $t_1 < t_2$

$$t_1 \circ u > t_2 \circ u$$

und

$$t_2 \circ (u \circ v) < t_1 \circ (u \circ v) = (t_1 \circ u) \circ v < (t_2 \circ u) \circ v = t_2 \circ (u \circ v),$$

was unmöglich ist. Ähnlich kann $t \circ u$ auch in der zweiten Veränderlichen nicht abnehmen. Damit ist der Beweis vollendet.

(Eingegangen am 4. Mai 1953.)

Bemerkung bei der Korrektur. (Am 9. Juni 1955.) Ein dem Satz 2 analoger Satz für $n = 1$ ist in der eleganten Arbeit von S. GOLAB, Sur les objets géométriques non différentiels, *Bulletin de l'Acad. Polon. des Sci. et des Lettres Classe des Sci. Math. et Nat.* A, (1949), S. 67—72 zu finden. Dies zeigt zugleich, daß (2) auch in der Theorie der geometrischen Objekten eine Anwendung hat. Auch (1) kann in dieser Theorie angewendet werden. Auf diese Frage wollen wir noch zurückkehren.

Literaturverzeichnis

[1] G. VRANCEANU, *Leçons de géométrie différentielle I* (Bucarest, 1947), S. 65—77.

[2] Б. А. Севастьянов, Теория ветвящихся случайных процессов, Успехи Мат. Наук, **6** (1951), в. 6 (46), S. 47—99, insb. 54.

[3] В. В. Степанов, Курс дифференциальных уравнений (Москва—Ленинград, 1950), S. 267—270.

[4] J. ACZÉL, L. KALMÁR, J. G. MIKUSINSKI, Sur l'équation de translation, *Studia Math.*, **12** (1951), S. 112—116.

[5] J. ACZÉL, Über einparametrige Transformationen, *Publ. Math. (Debrecen)*, **1** (1950), S. 243—247.

[6] J. ACZÉL, Sur les opérations définies pour nombres réels, *Bulletin de la Soc. Math. de France*, **76** (1948), S. 59—64.

РЕШЕНИЕ ВЕКТОРНОГО ФУНКЦИОНАЛЬНОГО УРАВНЕНИЯ ОДНОРОДНЫХ И НЕОДНОРОДНЫХ n-МЕРНЫХ ОДНОПАРАМЕТРИЧЕСКИХ „ТРАНСЛАЦИЙ“, В СВЯЗИ С ПРОИЗВОДЯЩЕЙ ФУНКЦИЕЙ ЦЕПНЫХ РЕАКЦИЙ И ИНТЕГРАЛОМ СТАЦИОНАРНЫХ И НЕСТАЦИОНАРНЫХ ДВИЖЕНИЙ

Я. А ц е л (Дебрецен)

(Р е з ю м е)

Целью настоящей работы является доказать, что функции

$$\mathbf{f}(\mathbf{x}, t) = \mathbf{g}[\mathbf{h}(\mathbf{x}) + t] \qquad (\mathbf{y} + t = \{y_1, y_2, \ldots, y_n\} + t = \{y_1 + t, y_2, \ldots, y_n\})$$

и

$$\mathbf{f}(\mathbf{x}, t, u) = \mathbf{g}[\mathbf{h}(\mathbf{x}, t), u]$$

($\mathbf{h}$ произвольный, $\mathbf{h}[\mathbf{g}(\mathbf{y})] = \mathbf{h}[\mathbf{g}(\mathbf{y}, t), t] = \mathbf{y}$) являются самыми общими решениями функциональных уравнений

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u] = \mathbf{f}(\mathbf{x}, t + u)$$

и

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, s, t), t, u] = \mathbf{f}(\mathbf{x}, s, u)$$

соответственно, если требовать выполненность условий разрешимости (5)—(6) и (9)—(10) соответственно. Эти уравнения имеют применения, указанные в заглавии работы. Излагаются также некоторые дальнейшие обобщения, следствия и применения.

EXTREMUM PROPERTIES OF THE REGULAR POLYTOPES

By
L. FEJES TÓTH (Budapest)
(*Presented by* G. HAJÓS)

It is more than a century ago that L. SCHLÄFLI[1] discovered the higher dimensional analogues of the five Platonic solids. He found that besides the regular simplex, cross polytope and measure polytope, i. e. the analogues of the regular tetrahedron $\{3, 3\}$, octahedron $\{3, 4\}$ and the cube $\{4, 3\}$, there are only three further convex regular polytopes, all in four dimensions. These are the self dual $\{3, 4, 3\}$ made up from 24 $\{3, 4\}$, 3 around each edge, $\{3, 3, 5\}$ bounded by 600 $\{3, 3\}$, 5 around each edge, and its dual, $\{5, 3, 3\}$ having 120 $\{5, 3\}$-cells (dodecahedra), 3 around each edge.

The great variety of the well-known extremum properties of the regular polygons, and especially of the equilateral triangle, suggests analogous properties of the regular solids and polytopes. In fact, many extremum properties of the regular polygons can easily be transferred to the regular simplex and a part of them to the cross polytope or measure polytope.[2] Also $\{5, 3\}$ and $\{3, 5\}$ (icosahedron) can be characterised by different extremum postulates.[3] But — though the theory of the regular polytopes has a vast literature[4] — little, if anything, is known in this direction about the "non-trivial" regular polytopes $\{3, 4, 3\}$, $\{3, 3, 5\}$ and $\{5, 3, 3\}$ mentioned above.[5]

In the present paper we shall give a simple proof[6] of the fact that among those convex four-dimensional polytopes of given insphere which are topologically isomorphic with $\{3, 3, 3\}$ (regular simplex), or $\{3, 3, 4\}$ (cross

[1] Theorie der vielfachen Kontinuität, *Denkschriften der schweizerischen naturforschenden Gesellschaft*, **38** (1901), pp. 1—237; — *Ges. Abh.* I (Basel, 1950), pp. 167—387.

[2] See e. g. the remarks on the end of the paper.

[3] L. FEJES TÓTH, *Lagerungen in der Ebene, auf der Kugel und im Raum* (Berlin—Göttingen—Heidelberg, 1953).

[4] H. S. M. COXETER, *Regular polytopes* (London, 1948).

[5] L. FEJES TÓTH, On close-packings of spheres in spaces of constant curvature, *Publicationes Math.* (*Debrecen*), **3** (1953), pp. 158—167.

[6] Compare our proof with M. GOLDBERG, The isoperimetric problem for polyhedra, *Tôhoku Math. J.*, **42** (1935), pp. 226—236, and H. HADWIGER, Zur isoperimetrischen Ungleichung für k-dimensionale konvexe Polyeder, *Nagoya Math. J.*, **5** (1953), pp. 39—44.

polytope), or $\{3, 3, 5\}$ (600-cell), or $\{3, 4, 3\}$ (24-cell), the corresponding regular polytope has the minimal volume.

It was J. STEINER who first compared polyhedra of the same topologic type, probably in order to characterize all the five regular solids by a unique extremum condition.[7] We can guess that the consideration of isomorphic polytopes has the same effect also in higher dimensions. For example, there is no doubt that the above proposition holds for $\{4, 3, 3\}$ (measure polytope) and $\{5, 3, 3\}$ (120-cell) too, but the proof of this conjecture seems to be not so simple. On the other hand, for the remaining polytopes our proof yields a little more:

Among all convex four-dimensional polytopes of given insphere bounded by 5, 16 *or* 600 *tetrahedra or* 24 *polyhedra isomorphic with the regular octahedron the corresponding regular polytope has the minimal volume.*

Let P be in 4-space a convex polytope bounded by n tetrahedra $t_1, \ldots, t_n$ and containing the unit sphere S. Let τ_i be the spherical tetrahedron arising from t_i by central projection upon S. The sum of the volumes[8] τ_i equals the surface (3-dimensional content) of S:

$$\sum_{i=1}^{n} \tau_i = 2\pi^2.$$

We assert that

$$t_i \geqq v(\tau_i)$$

where $v(\tau)$ denotes the volume of a regular tetrahedron whose hyperplane (3-space) touches S at the centre of the tetrahedron and the central projection of which has a volume equal to τ. This is equivalent to the fact that among the tetrahedra t, of given volume, whose hyperplanes do not intersect S, the regular tetrahedron touching S at its centre has maximal projection.

It is easy to show that if t is not entirely contained in a sufficiently large sphere concentric with S the projection τ of t becomes arbitrarily small. Therefore we can restrict ourself to tetrahedra lying in a fixed sphere and (with respect to the continuity of the functional τ) the existence of an extremal tetrahedron follows from the theorem of WEIERSTRASS.

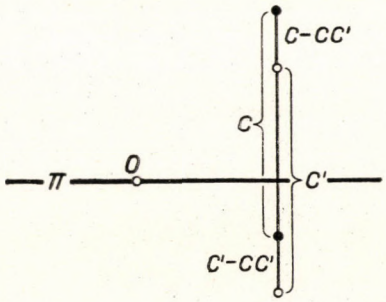
Obviously, we can suppose that the hyperplane of t touches S at a point O since otherwise we could increase the volume of its projection by a translation. Let Π be a plane through O, orthogonal to an edge of t. Let us replace all chords c of t orthogonal to Π by the chords c' of the same length and lying on the same line symmetrically with respect to Π. This process, called symmetrisation of STEINER,[9] carries t into a new tetrahedron t' of the

[7] Such a characterisation is impossible by comparing polyhedra of given number of faces or vertices.

[8] We denote a body and its volume by the same symbol.

[9] The symmetrisation of a tetrahedron and octahedron was already used by STEINER in order to establish the isoperimetric property of $\{3, 3\}$ and $\{3, 4\}$.

same volume $t' = t$ symmetric with respect to Π. The fact that any inner point of the segment $c'—cc'$ lies nearer to O than any point of $c—cc'$ involves a volume preserving representation of t upon t' which lets the points of tt' invariant and carries the other ones into points lying nearer to O. But, since among two volume elements of t of equal volume that one lying nearer to O has a greater projection upon S, the projection of t' is greater than or equal to the projection of t. Equality holds only if t is symmetric with respect to Π.



It follows that the best tetrahedron must be symmetric with respect to any plane through O orthogonal to an edge, i. e. it is regular and has O for its centre. This completes the proof of our assertion.

The above remark on the volume elements of t makes it evident that $v(\tau)$ is a convex function of τ $(0 \leqq \tau < \pi^2)$. Hence, by Jensen's inequality

$$4P \geqq \sum_{i=1}^{n} t_i \geqq \sum_{i=1}^{n} v(\tau_i) \geqq n \cdot v\left(\frac{1}{n} \sum_{i=1}^{n} \tau_i\right) = n \cdot v\left(\frac{2\pi^2}{n}\right).$$

The case of equality is obvious and implies just the desired extremum property of the regular 5-, 16- and 600-cell.

The proof of our theorem concerning $\{3, 4, 3\}$ runs analogously by replacing the word "tetrahedron" by "octahedron" or, more precisely, by "polyhedron isomorphic with the regular octahedron". But in this case we symmetrise first on a plane orthogonal to a diagonal of the octahedron. Under such a symmetrisation the topologic type of the octahedron remains unaltered. Three successive symmetrisations on such planes passing through O yield an octahedron symmetric with respect to O and having mutually orthogonal diagonals. Consequently, we can restrict ourself to the simple case of octahedra of this property. Symmetrising such an octahedron first on a plane through O orthogonal to an edge, then on a plane through O orthogonal to a diagonal radiating from one endpoint of this edge, we obtain an octahedron of the same property. It follows that the best octahedron must be symmetric also with respect to any plane through O orthogonal to an edge, i. e. it must be regular and to have O for its centre.

An analogous proof concerning $\{4, 3, 3\}$ and $\{5, 3, 3\}$ miscarries on the fact that in case of a polyhedron isomorphic with $\{4, 3\}$ or $\{5, 3\}$ the symmetrisation on any plane generally alters the type of the polyhedron.

We close our discussions with some remarks.

The process used in the proof of the inequality $t \geqq v(\tau)$ shows at the same time that among the tetrahedra t of given volume the regular tetrahedron $\bar{t}$ concentric with a sphere S contains the greatest part of the volume of S:

$$tS \leqq \bar{t}S.$$

This is a generalisation of the well-known facts that among the tetrahedra inscribed in or circumscribed about to a sphere the regular tetrahedron has the maximal and minimal volume, respectively.

Another consequence of the symmetrisation is the fact that among the tetrahedra the regular one has the least quotient R/r where R denotes the circumradius and r the inradius.[10]

Analogous propositions hold for polyhedra isomorphic with $\{3, 4\}$. Further, the theorem concerning R/r can be transferred by aid of polar reciprocation from $\{3, 4\}$ to $\{4, 3\}$.[11] All these proofs can be extended without difficulty to higher dimensions. So we obtain, for instance: If R and r are the circumradius and inradius of an n-dimensional simplex, then

$$R \geqq nr.$$

In case of an n-dimensional polytope isomorphic with the cross polytope or measure polytope we have

$$R^2 \geqq nr^2.$$

(Received 19 February 1954)

ОБ ЭКСТРЕМАЛЬНЫХ СВОЙСТВАХ РЕГУЛЯРНЫХ ПОЛИТОПОВ

Л. Фееш Тот (Будапешт)

(Резюме)

Автором доказывается, что из выпуклых политопов четырех измерений, топологически изоморфных регулярному политопу символа $\{3, 3, 3\}$ или $\{3, 3, 4\}$ или $\{3, 3, 5\}$, и содержащих данную сферу, сам регулярный политоп имеет минимальный объем.

[10] A very simple direct proof of this was given by I. ÁDÁM. (See the book quoted in footnote [3], p. 28.)

[11] Cf. the book cited in [3], p. 131.

ÜBER GEWISSE EIGENSCHAFTEN ORTHOGONALER SYSTEME DER KLASSE L^2 UND DIE EIGENFUNKTIONEN STURM—LIOUVILLESCHER DIFFERENTIALGLEICHUNGEN

Von
MIKLÓS MIKOLÁS (Budapest)
(*Vorgelegt von* B. Sz.-Nagy)

§ 1. Einleitung

1.1. Vielfach benutzt man diejenige Eigenschaft des *trigonometrischen Orthogonalsystems,* daß man von ihm durch Differentiation und Integration wieder Orthogonalsysteme erhält, welche überdies mit demselben Intervalle (der Länge 2π) verknüpft sind. Ähnlich ist der Sachverhalt, wie man sofort einsieht, z. B. bei den Tschebyscheffschen Polynomen erster und zweiter Art, allgemeiner bei den Jacobischen, gleichwie bei den Laguerreschen und Hermiteschen Polynomen; d. h. in diesen Fällen ist die abgeleitete Folge, bzw. ein Integralsystem der mit der Gewichtsfunktion multiplizierten Folge wieder inbezug auf dasselbe Intervall nebst einer passenden Belegung orthogonal. Die nach den „klassischen" Systemen gebildeten Orthogonalentwicklungen gehen also bei Differentiation und Integration in *Orthogonalreihen* desselben Typus über; m. a. W.: sie behalten ihren Charakter bei den genannten Operationen. Diese Tatsache ist für die Anwendungen ebenso vorteilhaft, wie die entsprechende im Falle der *Potenzreihen.*

1.2. Auf Grund der bisherigen Überlegungen hat die Frage ein gewisses Interesse, inwiefern alles dies sich verallgemeinern läßt. — Dementsprechend wollen wir uns in der vorliegenden Arbeit mit den folgenden Problemen beschäftigen:

I. Es sei $\{\varphi_\lambda(x)\}$ ein beliebiges Orthogonalsystem der Klasse $L^2_w(a, b)$, verknüpft mit der nicht negativen Belegungsfunktion $w(x)$.[1] Es seien gewisse

[1] Es kommen nur solche Belegungsfunktionen (Gewichtsfunktionen) in Betracht, welche nur über einer Nullmenge des Grundintervalles $[a, b]$ verschwinden und darin L-integrabel sind. — $L^2_w(a, b)$ bezeichnet die Gesamtheit der in (a, b) meßbaren Funktionen $f(x)$, für welche $\int_a^b w(x)f(x)^2\,dx$ existiert. Ferner benutzt man $(w(x) \equiv 1)$ die übliche Bezeichnung $L^p(a, b)$ für die Klasse der in (a, b) meßbaren Funktionen $f(x)$ mit dabei L-integrablen $|f(x)|^p$. — Hier und durchwegs in dieser Arbeit kommt nur der *Lebesguesche Integralbegriff* zur Anwendung.

Konstanten $c_1, c_2 \ldots$ und eine Funktion $Q(x) \geqq 0$ $(a \leqq x \leqq b)$ angegeben. Gesucht wird eine notwendige und hinreichende Bedingung dafür, daß die Folge $\{\Phi_\lambda(x) + c_\lambda\}$ mit $\Phi_\lambda(x) = \int_a^x w(t)\varphi_\lambda(t)\,dt$ $(a \leqq x \leqq b;\ \lambda = 1, 2, \ldots)$, bzw. das erweiterte System 1, $\{\Phi_\lambda(x) + c_\lambda\}$ inbezug auf (a, b) nebst der Belegungsfunktion $Q(x)$ orthogonal ist.[2]

II. Es sei $\{\varphi_\lambda(x)\}$ ein Orthogonalsystem von in (a, b) differenzierbaren Funktionen für (a, b), verknüpft mit der ebenda nicht negativen Belegungsfunktion $w(x)$. Es sei noch eine Funktion $g(x) \geqq 0$ $(a \leqq x \leqq b)$ angegeben. Gesucht wird eine notwendige und hinreichende Bedingung dafür, daß die Ableitungsfolge $\{\varphi_\lambda'(x)\}$ ein Orthogonalsystem zur Belegung $g(x)$ für das Intervall (a, b) bildet.

Eng schließen sich die weiteren Aufgaben heran:

III. Es ist eine möglicherweise weite Klasse solcher Orthogonalsysteme anzugeben und näher zu untersuchen, welche gleichzeitig die in I und II genannten Eigenschaften besitzen.

IV. Wie lassen sich die „klassischen" Orthogonalsysteme, d. h. die Jacobischen, Laguerreschen, Hermiteschen Polynome und die Sturm—Liouvilleschen Eigenfunktionen auf Grund der in dieser Richtung gefundenen Resultate charakterisieren?

V. Sei eine (allgemeine) Fourierreihe betrachtet, von welcher man durch Differentiation (Integration) wieder eine Orthogonalreihe erhält; was kann über die Konvergenz bzw. Summabilität der ursprünglichen und abgeleiteten (integrierten) Reihe ausgesprochen werden?

1.3. Diese Probleme scheinen in der gegebenen Fassung neu zu sein. Was IV betrifft, so sind in den letzten Jahrzehnten mehrere Untersuchungen publiziert worden, welche von den üblichen Definitionen (Rodriguessche Formeln, Orthogonalisierung, Erzeugungsfunktionen, Eigenwertaufgaben) abweichende, eigenartige Kennzeichnung einer oder anderer Klasse von *Orthogonalpolynomen* enthalten.[3] J. EGERVÁRY gab z. B. in einer schönen Arbeit gewisse charakteristische geometrische Eigenschaften der Legendreschen und Tschebyscheffschen Polynome an,[4] während W. HAHN und H. L. KRALL den Satz bewiesen haben: *unter allen orthogonalen Polynomsystemen* besitzen die klassischen im wesentlichen[5] allein die Eigenschaft, daß die aus ihnen durch

[2] Wie wir sehen werden, ist der Fall $c_1 = c_2 = \cdots = 0$ bei den klassischen Orthogonalpolynomen ausgezeichnet.

[3] Vgl. die Bibliographie [28].

[4] Vgl. [5]. — Der Satz rührt von L. FEJÉR her. (Vgl. EGERVÁRY, 1. c., S. 18.)

[5] D. h. von trivialen linearen Transformationen abgesehen. — Vgl. [14], [15], [21], [22], [23], ferner [30], [28], [2].

Differentiation (oder wiederholtes Differenzieren) entstehende Polynomfolge gleichfalls nebst einer geeigneten Belegung inbezug auf dasselbe Intervall orthogonal ist. Nach der Entstehung dieser Arbeit ist eine Notiz von J. ACZÉL erschienen, welche sich auf eine vereinfachte Behandlung der klassischen Orthogonalpolynome mittels einer Rodriguesschen Formel bezieht.[6] — In diesem Zusammenhange hat man noch einen ganz speziellen Satz von B. GAGAEFF zu erwähnen,[7] wonach es außer $1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots$ wesentlich kein anderes Orthogonalsystem gibt, welches sich bei Derivation bis auf Zahlenfaktoren reproduziert.[8]

V umfaßt offenbar die klassische Frage über die gliedweise Differenzierbarkeit und Integrierbarkeit einer gewöhnlichen Fourierreihe; die zweite ist bekanntlich stets bestattet, bezüglich der ersten aber folgt aus dem Fejérschen Grundsatze kurz und elegant, daß die stetige Differenzierbarkeit der Summenfunktion die C_1-Summierbarkeit (zur Derivierten) der abgeleiteten Reihe nach sich zieht.[9] Dieses schöne Fejérsche Ergebnis hat etwa zehn Jahre später zu mehrerlei weiteren Untersuchungen über trigonometrische Fourierreihen Anlaß gegeben.

1.4. Die erwähnten Probleme werden wir vor allem im Falle vollständiger Systeme behandeln.

Bezüglich I und II ergibt sich in § 2 die notwendige und hinreichende Bedingung, daß $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ einer *Integralgleichung* der Form

$$\varphi_\lambda(x) = \varrho_\lambda \int_a^x Q(t)[\Phi_\lambda(t) + c_\lambda]\,dt + K_\lambda \tag{1.1}$$

mit bestimmten konstanten ϱ_λ, c_λ, K_λ *fast überall in* (a, b) genügt (vgl. Satz 1),[10] bzw. (was im Falle stetiger $\varphi_\lambda(x)$ auf dasselbe hinausgeht), daß in (a, b) eine Relation vom Typus

$$[g(x)\varphi_\lambda'(x)]' = \varrho_\lambda w(x)\varphi_\lambda(x) \tag{1.2}$$

nebst Randbedingungen erfüllt ist. (Sätze 2 und 3.)

[6] Vgl. [1].

[7] Vgl. [10], [12].

[8] Beim Lektorieren machte mich Herr Professor B. SZŐKEFALVI-NAGY auf eine mir vorher unzugängliche, neuestens publizierte Arbeit von D. C. LEWIS ([25]) aufmerksam; hierin beweist der Verfasser einige, meinen Sätzen 3 und 4 teilweise entsprechende Behauptungen, seine Resultate stehen aber nur mit meinem Problem II in Berührung. Ich will Prof. SZ.-NAGY für die wertvollen Ratschläge auch an dieser Stelle meinen besonderen Dank ausdrücken. — Wie ich bei der Korrektur bemerkt habe, findet sich ein dem ersten Teile meines Satzes 3 ähnliches, etwas spezielleres Ergebnis schon in einer anderen Arbeit von GAGAEFF ([11]), welche aber von LEWIS nicht zitiert wird.

[9] Vgl. [6], [7].

[10] Wir geben gleich in 2.1 eine Verschärfung des Satzes 1, welche sich auf *Lebesgue—Stieltjes-Integrale* bezieht.

M. a. W.: es soll jede Funktion $y=\varphi_\lambda(x)$ $(\lambda=1, 2, \ldots)$ eine Lösung der *Sturm—Liouvilleschen Differentialgleichung*

$$[g(x)y']'=\varrho_\lambda w(x)y \tag{1.3}$$

sein.

In § 3 zeigt man u. a. ohne Mühe (vgl. III), daß die, passende Randbedingungen befriedigenden Lösungen einer Gleichung von der Form (1. 3) ein die Eigenschaften I und II *gleichzeitig* besitzendes Orthogonalsystem bilden (Satz 4). Die Gesamtheit der betreffenden Folgen $\{\varphi_\lambda(x)\}$ — *„gewöhnlicher SL-Systeme"* — ist somit als eine natürliche Erweiterungsklasse zu betrachten. Es wird auch die naheliegende Frage untersucht, wann die ersten bzw. höheren Ableitungen eines gewöhnlichen *SL*-Systems $\{\varphi_\lambda(x)\}$ gleichfalls zu dieser Klasse gehören (Satz 5); die Bemerkung, daß die Derivierte einer Lösung von (1. 3) wieder einer Differentialgleichung vom Sturm—Liouvilleschen Typus genügt, wobei die Koeffizienten einfach mit $g(x)$, $w(x)$ zusammenhängen (Lemma 1), spielt da und später eine wesentliche Rolle.

Die beiden letzten Abschnitte (§§ 4 und 5) enthalten *Anwendungen* bezüglich IV und V.

Sätze 6, 7, 8 liefern je eine solche *Charakterisierung* (in der Gesamtheit aller Orthogonalsysteme der Klasse L^2) für das trigonometrische System bzw. für die klassischen Orthogonalpolynome, ferner für beliebige Eigenfunktionen von (1. 3) erster oder zweiter Art, welche nur *Eigenschaften allgemeiner Natur*, wie Beschränktheit, Stetigkeit bzw. Differenzierbarkeit, Orthogonalität, Vollständigkeit benutzt.[11]

Satz 9 ergibt endlich eine *Verallgemeinerung des Fejérschen Satzes* über die Differentiation gewöhnlicher Fourierschen Sinusreihen, deren Beweis natürlich weitere Hilfsmittel aus der Theorie der Sturm—Liouvilleschen Randwertprobleme (z. B. asymptotische Formeln, das Äquikonvergenztheorem von HAAR) und aus der allgemeinen Theorie der reellen Funktionen erfordert.

§ 2. Untersuchung der Probleme I, II

2.1. In den folgenden bezeichnen, wie üblich, $[a, b]$ ein abgeschlossenes, (a, b) ein offenes Intervall, welche endlich oder unendlich sein mögen. Es kommen durchwegs nur reelle Zahlen und Funktionen in Betracht.

Zunächst gilt bezüglich der Frage I der folgende

SATZ 1. *Es sei* $w(x)\in L(a, b)$ *eine in* (a, b) *fast überall positive Funktion,* $\varphi_\lambda(x)\in L^2_w(a, b)$ $(\lambda=1, 2, \ldots)$ *eine Funktionenfolge, welche kein konstan-*

[11] Insbesondere wird die Gültigkeit einer Differentialgleichung, oder einer anderen speziellen Relation weder bei den Elementen des Systems, noch bei den Gewichtsfunktionen verlangt.

tes Element enthält.[12] *Wir nehmen an, daß entweder* 1) $\{\varphi_\lambda(x)\}$ *selbst oder* 2) *die Folge* $\gamma_0, \varphi_1(x), \varphi_2(x), \ldots$, *mit einem konstanten* $\gamma_0 \neq 0$, *ein w-vollständiges und w-orthonormales System*[13] *inbezug auf* (a, b) *bildet.*

I) *Sei* $Q(x)$ *eine in* (a, b) *fast überall positive Funktion,* $Q(x) \in L(a, b)$, *und seien* $c_1, c_2, \ldots$ *gewisse Konstanten; man betrachte die Folge*

$$1,\ \Phi_1(x)+c_1,\ \Phi_2(x)+c_2, \ldots \tag{2.1}$$

mit

$$\Phi_\lambda(x) = \int_a^x w(t)\varphi_\lambda(t)\,dt \qquad (\lambda = 1, 2, \ldots). \tag{2.2}$$

Das Funktionensystem (2.1) *ist dann und nur dann Q-orthogonal auf* (a, b), *wenn fast überall in* (a, b)

$$\varphi_\lambda(x) = \varrho_\lambda \int_a^x Q(t)\,[\Phi_\lambda(t)+c_\lambda]\,dt + K_\lambda \qquad (\lambda = 1, 2, \ldots) \tag{2.3}$$

gilt, wobei

$$\varrho_\lambda = -\left\{\int_a^b Q(x)\,[\Phi_\lambda(x)+c_\lambda]^2\,dx\right\}^{-1} \quad \textit{und}$$

$$c_\lambda = -\frac{\displaystyle\int_a^b Q(x)\,\Phi_\lambda(x)\,dx}{\displaystyle\int_a^b Q(x)\,dx} \qquad (\lambda = 1, 2, \ldots) \tag{2.4}$$

sind, während K_λ *die Null oder* $-\gamma_0^2\varrho_\lambda \int_a^b w(t) \int_a^t Q(u)\,[\Phi_\lambda(u)+c_\lambda]\,du\,dt$ *bedeutet, je nachdem der Fall* 1) *bzw.* 2) *vorliegt.*

II) *Wird* $c_\lambda = 0$ *und* $Q(x)\,\Phi_\lambda(x) \in L(a, b)$ $(\lambda = 1, 2, \ldots)$ *statt* $Q(x) \in L(a, b)$ *vorausgesetzt, so bildet* $\{\Phi_\lambda(x)\}$ *im Falle* 2) *dann und nur dann ein Q-orthogonales System auf* (a, b), *wenn* (2.3) *darin bis auf eine Nullmenge mit* $K_\lambda = -\gamma_0^2\varrho_\lambda \int_a^b w(t) \int_a^t Q(u)\,\Phi_\lambda(u)\,du\,dt$ *und nebst der Bedeutung* (2.4) *von* ϱ_λ *gilt.*[14]

[12] Da und später heißen zwei Funktionen nur dann verschieden, wenn ihre Erklärung über einer Menge von positivem Maße abweicht.

[13] D. h. vollständig, usw., in $L_w^2(a, b)$.

[14] Beispiele: 1*—2* alle normierten Sturm—Liouvilleschen Eigenfunktionen erster bzw. zweiter Art, welche zu $[Q(x)^{-1}y']' = \varrho w(x)y$, $Q(x) > 0$, $w(x) > 0$ $(a \leqq x \leqq b)$ angehören, 3* das trigonometrische Orthonormalsystem (auf einem Intervalle der Länge 2π), 4* die normierten Jacobischen Polynome. Für 1* ist I), Fall 1), für 2*, 3* aber I), Fall 2), während für 4* $(Q(x) = (1-x)^{-\alpha-1}(1+x)^{-\beta-1}$, $\alpha > -1$, $\beta > -1)$ nur II), Fall 2) anwendbar. (Vgl. [20], S. 265 ff., [29].)

BEWEIS. Es seien alle Voraussetzungen für $w(x)$ und $\{\varphi_\lambda(x)\}$ erfüllt; dann gilt in beiden Fällen 1) und 2)

$$\int_a^b w(x)\varphi_m(x)\varphi_n(x)dx = \begin{cases} 1, & \text{falls } m=n, \\ 0 & \text{für } m \neq n \end{cases} \qquad (m, n = 1, 2, \ldots) \tag{2.5}$$

und im Falle 2) überdies

$$\int_a^b w(x)\varphi_\lambda(x)dx = 0 \qquad (\lambda = 1, 2, \ldots), \tag{2.6}$$

$$\gamma_0^2 \int_a^b w(x)dx = 1, \quad \gamma_0^2 = \left[\int_a^b w(x)\,dx\right]^{-1}. \tag{2.7}$$

I) 1° *(Notwendigkeit.)* Nehmen wir an, daß das System (2. 1) unter der Bezeichnung (2. 2) bei gewissen Werten von c_λ inbezug auf (a, b) nebst einer angegebenen Belegungsfunktion $Q(x) \in L(a, b)$ orthogonal ist, d. h.

$$\int_a^b Q(x)[\Phi_\lambda(x) + c_\lambda]dx = 0 \qquad (\lambda = 1, 2, \ldots) \tag{2.8}$$

und

$$\int_a^b Q(x)[\Phi_m(x) + c_m][\Phi_n(x) + c_n]dx = 0 \qquad (m \neq n;\ m, n = 1, 2, \ldots). \tag{2.9}$$

Dann können wir aus (2. 8) sofort auf den Wert (2. 4) von c_λ $(\lambda = 1, 2, \ldots)$ schließen; c_λ soll also weiterhin zur kurzen Bezeichnung des dortigen Ausdrucks dienen.

Die linke Seite von (2. 9) integrieren wir nach Teilen und zwar in der für Lebesguesche Integrale passenden Weise.[15] Man erhält

$$\int_a^b Q(x)[\Phi_m(x) + c_m][\Phi_n(x) + c_n]dx = \int_a^b Q(t)[\Phi_m(t) + c_m]dt \cdot [\Phi_n(b) + c_n] -$$
$$- \int_a^b w(x)\varphi_n(x) \int_a^x Q(t)[\Phi_m(t) + c_m]dt\,dx \qquad (m, n = 1, 2, \ldots), \tag{2.10}$$

wobei das Glied in der Mitte vermöge (2. 8) verschwindet. Durch Kombination von (2. 10) und (2. 9) ergibt sich also

$$\int_a^b w(x)\varphi_m(x) \int_a^x Q(t)[\Phi_m(t) + c_m]dt\,dx =$$
$$= -\int_a^b Q(x)[\Phi_m(x) + c_m]^2 dx \qquad (m = 1, 2, \ldots) \tag{2.11}$$

[15] Vgl. [26], S. 54.

für $m = n$ und

$$\int_a^b w(x)\varphi_n(x)\int_a^x Q(t)[\Phi_m(t)+c_m]\,dt\,dx = 0 \qquad (m \neq n;\ m, n = 1, 2, \ldots) \tag{2.12}$$

sonst.

Aus (2. 5) und (2. 12) folgt nun

$$\int_a^b w(x)\varphi_n(x)\left\{\varphi_m(x) - \varrho_m \int_a^x Q(t)[\Phi_m(t)+c_m]\,dt\right\}dx = 0 \tag{2.13}$$

$$(m \neq n;\ m, n = 1, 2, \ldots)$$

bei *jeder* Wahl der Zahlenfaktoren ϱ_m; indem man zweckmäßig auch für $m = n$

$$\int_a^b w(x)\varphi_m(x)\left\{\varphi_m(x) - \varrho_m \int_a^x Q(t)[\Phi_m(t)+c_m]\,dt\right\}dx = 0 \tag{2.14}$$

setzt, so führt (2. 14) mit Rücksicht auf (2. 5) und (2. 11) zu

$$\varrho_m = -\frac{\int_a^b w(x)\varphi_m(x)^2\,dx}{\int_a^b Q(x)[\Phi_m(x)+c_m]^2 dx} = -\left\{\int_a^b Q(x)[\Phi_m(x)+c_m]^2 dx\right\}^{-1} \qquad (m = 1, 2, \ldots), \tag{2.15}$$

entsprechend (2. 4).

Mit dieser (für die Folgenden festgehaltenen) Bedeutung von ϱ_m schreiben wir zur Abkürzung

$$F_m(x) = \varphi_m(x) - \varrho_m \int_a^x Q(t)[\Phi_m(t)+c_m]\,dt \qquad (m = 1, 2, \ldots);$$

dann geht (2. 13) in

$$\int_a^b w(x)\varphi_n(x)F_m(x)\,dx = 0 \tag{2.16}$$

über, was aber jetzt für *beliebige positiv-ganze* m, n (auch bei $m = n$) gilt.

Wenn man in (2. 16) m fix denkt, während n alle natürlichen Zahlen durchlaufen läßt, so sieht man, daß alle verallgemeinerten Fourier-Koeffizienten von $F_m(x)$ inbezug auf $\{\varphi_\lambda(x)\}$ und die Belegungsfunktion $w(x)$ verschwinden. Daraus folgt im Falle 1), daß

$$F_m(x) \equiv 0 \qquad m = 1, 2, \ldots),$$

d. h.

$$\varphi_m(x) = \varrho_m \int_a^x Q(t)[\Phi_m(t)+c_m]\,dt \qquad (m = 1, 2, \ldots) \tag{2.17}$$

f. ü. in (a, b).

Im Falle 2) folgt aber (z. B. durch die Parsevalsche Formel bzw. durch die Konvergenz „im Mittel“ der Fourierentwicklung nach $\{\varphi_\lambda(x)\}$ von $F_m(x)$[16]), daß

$$F_m(x) \equiv \text{const} = K_m,$$

d. h.

$$\varphi_m(x) = \varrho_m \int_a^x Q(t)\,[\Phi_m(t) + c_m]\,dt + K_m \qquad (m = 1, 2, \ldots) \tag{2. 18}$$

f. ü. in (a, b).

(2. 17) und (2. 18) ergeben die erste Hälfte der Behauptung I).

2° *(Hinreichen.)* Umgekehrt, es sei die Gültigkeit von (2. 3) f. ü. in (a, b) und mit der im Satze gegebenen Bedeutung der Konstanten $c_\lambda, \varrho_\lambda, K_\lambda$ angenommen.

Dann folgt (2. 8) aus der Wahl von c_λ $(\lambda = 1, 2, \ldots)$, so daß die oben ausgeführte partielle Integration wieder die Formel

$$\int_a^b Q(x)\,[\Phi_m(x) + c_m]\,[\Phi_n(x) + c_n]\,dx =$$

$$= -\int_a^b w(x)\varphi_n(x) \int_a^x Q(t)\,[\Phi_m(t) + c_m]\,dt\,dx \qquad (m, n = 1, 2, \ldots)$$

ergibt. Nun ist aber vermöge (2. 3) und (2. 5) in beiden Fällen

$$\varrho_m \int_a^b w(x)\,\varphi_n(x) \int_a^x Q(t)\,[\Phi_m(t) + c_m]\,dt\,dx = \int_a^b w(x)\varphi_n(x)\,[\varphi_m(x) - K_m]\,dx =$$

$$= -K_m \int_a^b w(x)\varphi_n(x)\,dx = 0 \qquad (m \neq n;\ m, n = 1, 2, \ldots), \tag{2. 19}$$

weil im vorletzten Gliede bei 1) der erste, bei 2) nach (2. 6) der zweite Faktor verschwindet.

Beachtet man noch, daß offenbar $\varrho_m < 0$ $(m = 1, 2, \ldots)$ ist, so kann man aus den beiden letzten Relationen auf (2. 9), d. h. auf die behauptete Orthogonalitätseigenschaft von (2. 1) schließen.

II) Wird $c_\lambda = 0$ $(\lambda = 1, 2, \ldots)$ gewählt, und statt $Q(x)$ nur $Q(x)|\Phi_\lambda(x)|$ $(\lambda = 1, 2, \ldots)$ in $[a, b]$ L-integrabel angenommen, ferner der Fall 2) betrachtet, so fällt (2. 8) weg, das mittlere Glied in (2. 10) verschwindet aber wegen (2. 6) wieder; es folgt also, daß alle weiteren Formeln und Feststellungen mit $c_\lambda = 0$ stehen bleiben.

Damit ist Satz 1 bewiesen.

[16] Vgl. z. B. [16], S. 16.

Wir bemerken, daß alle Integrale im Satze 1 sich bequem als diejenigen von Lebesgue—Stieltjes schreiben lassen, was offenbar gewisse Verallgemeinerungen ermöglicht.

Bezeichnen $w^*(x)$ und $Q^*(x)$ unbestimmte Integrale von $w(x)$ bzw. $Q(x)$ in (a, b), so ist

$$\Phi_\lambda(x) = \int_a^x \varphi_\lambda(t)\, dw^*(t) \qquad (\lambda = 1, 2, \ldots), \tag{2.20}$$

und der Teil I) unseres Satzes lautet: *falls* $\int_a^b dQ^*(x)$ *vorhanden und positiv ist, bildet* $1, \{\Phi_\lambda(x) + c_\lambda\}$ *dann und nur dann ein Orthogonalsystem inbezug auf die Verteilung* $dQ^*(x)$ *und* (a, b), *wenn über einer Menge* $E \in (a, b)$, *welche fast alle Punkte* x *von* (a, b) *enthält,*

$$\varphi_\lambda(x) = \varrho_\lambda \int_{x_0}^x [\Phi_\lambda(t) + c_\lambda]\, dt + q_\lambda(x_0) \qquad (\lambda = 1, 2, \ldots) \tag{2.21}$$

gilt, wobei

$$\varrho_\lambda = -\left\{ \int_a^b [\Phi_\lambda(x) + c_\lambda]^2\, dQ^*(x) \right\}^{-1},$$
$$c_\lambda = \left[\int_a^b dQ^*(x) \right]^{-1} \int_a^b \Phi_\lambda(x)\, dQ^*(x) \qquad (\lambda = 1, 2, \ldots) \tag{2.22}$$

und im Falle 1) $x_0 = a$, $\varphi_\lambda(a) = 0$ $(\lambda = 1, 2, \ldots)$, *während im Falle* 2) x_0 *beliebig in* E *gewählt werden soll.*

Teil II) erlaubt die folgende Verschärfung: *ist* $c_\lambda = 0$ $(\lambda = 1, 2, \ldots)$ *und strebt das Produkt* $\Phi_n(x) \int_{x_0}^x \Phi_m(t)\, dQ^*(t)$ *für jedes feste* m, n, $x_0 \in (a, b)$ *für* $x \to a + 0$, $x \to b - 0$ *gegen Null, so bildet* $\{\Phi_\lambda(x)\}$ *im Falle* 2) *genau dann ein Orthogonalsystem, wenn* (2.21) *für jedes* $x \in E$, $x_0 \in E$ *mit* $c_\lambda = 0$, $\varrho_\lambda = -\left[\int_a^b \Phi_\lambda(x)^2\, dQ^*(x) \right]^{-1}$ $(\lambda = 1, 2, \ldots)$ *erfüllt ist.*[17]

2.2. Wenn die Funktionen $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$, $w(x)$, $Q(x)$ in (a, b) *stetig* und die beiden letzten noch positiv vorausgesetzt werden,[18] so ist dies auf die angewandte Schlußweise nur in dem Sinne von Einfluß, daß in den

[17] Die Bedingungen dieser allgemeineren Fassung II) passen nicht nur für die Jacobischen, sondern auch für die (mit unendlichem (a, b) verknüpften) Laguerreschen bzw. Hermiteschen Polynome.

[18] Da die Stetigkeit in den Endpunkten nicht verlangt wird, sollen die Integrabilitätsforderungen ja für sich beibehalten werden.

obigen Gleichungen die Ausnahmemengen vom Maße Null wegfallen; dann ist (2. 3) wegen der Differenzierbarkeit der unbestimmten Integrale offenbar äquivalent mit einer Differentialgleichung mit Randwertvorschriften.

Man erhält so den

SATZ 2. *Es sei* $w(x)\in L(a,b)$ *eine in* (a,b) *positive und stetige Funktion,* $\varphi_\lambda(x)$ $(\lambda=1,2,\ldots)$ *eine Folge von nicht konstanten, in* $[a,b]$ *stetigen Funktionen. Wir nehmen an, daß entweder* 1) $\{\varphi_\lambda(x)\}$ *selbst oder* 2) *die Folge* $\gamma_0, \varphi_1(x), \varphi_2(x), \ldots$ *mit einem konstanten* $\gamma_0 \neq 0$, *ein w-vollständiges und w-orthonormales System inbezug auf* (a,b) *bildet.*

I) *Es bezeichne* $Q(x)\in L(a,b)$ *noch eine in* (a,b) *positive und stetige Funktion;* $c_1, c_2, \ldots$ *seien gewisse Konstanten. Man betrachte die Folge:* $1, \Phi_1(x)+c_1, \Phi_2(x)+c_2, \ldots$ *mit* (2. 2).

Das letztgenannte System ist dann und nur dann Q-orthogonal für (a,b), *wenn* 1° $\varphi_\lambda(x)$ *und sogar* $Q(x)^{-1}\varphi_\lambda(x)$ $(\lambda=1,2,\ldots)$ *darin stetig diffenrenzierbar,* 2° *die Relationen*

$$\left[\frac{1}{Q(x)}\varphi_\lambda'(x)\right]' = \varrho_\lambda w(x)\varphi_\lambda(x) \qquad (a<x<b;\ \lambda=1,2,\ldots), \tag{2.23}$$

und 3° *die Randbedingungen*

$$\left.\begin{aligned} &\varphi_\lambda(a+0)=K_\lambda \qquad (2.24)\\ &\lim_{x\to a+0}\frac{\varphi_\lambda'(x)}{Q(x)}=\varrho_\lambda c_\lambda \qquad (2.25)\end{aligned}\right\} \qquad (\lambda=1,2,\ldots)$$

mit der im Satze 1 *enthaltenen Bedeutung von* $K_\lambda, \varrho_\lambda, c_\lambda$ *erfüllt sind.*[19]

II) *Ist* $c_\lambda=0$ $(\lambda=1,2,\ldots)$ *und wird die Integrabilität von* $Q(x)|\Phi_\lambda(x)|$ $(\lambda=1,2,\ldots)$ *in* $[a,b]$, *statt derjenigen von* $Q(x)$ *angenommen, so bildet nebst* 2) $\{\Phi_\lambda(x)\}$ *genau dann ein Q-orthogonales System für* (a,b), *falls* (2. 23) *und* $\lim_{x\to a+0}[Q(x)^{-1}\varphi_\lambda'(x)]=0$ $(\lambda=1,2,\ldots)$ *gelten.*[20]

2. 3. Was die Frage II der Einleitung betrifft, so zeigen wir den

SATZ 3. *Es sei* $w(x)\in L(a,b)$ *eine in* (a,b) *positive und stetige Funktion,* $\varphi_\lambda(x)\in L_w^2(a,b)$ $(\lambda=1,2,\ldots)$ *eine Folge ebenda zweimal stetig differenzierbarer Funktionen. Wir nehmen an, daß* $\{\varphi_\lambda(x)\}$ *ein w-vollständiges, w-orthogonales System inbezug auf* (a,b) *bildet.* $g(x)$ *bezeichne eine in* (a,b) *positive, stetig differenzierbare Funktion mit* $w(x)^{-1}[g(x)\varphi_\lambda'(x)]'\in L_w^2(a,b)$ $(\lambda=1,2,\ldots)$ *und*

[19] Es sei hervorgehoben, daß diese Bedingungen 1°—3° offenbar noch $\varphi_\lambda(b-0)=K_\lambda$ und im Falle 2) auch $\lim_{x\to b-0}[Q(x)^{-1}\varphi_\lambda'(x)]=\varrho_\lambda c_\lambda$ $(\lambda=1,2,\ldots)$ implizieren.

[20] Für Beispiele s. die Fußnote [14].

es seien die Relationen

$$\lim_{x\to a+0}[g(x)\varphi_m(x)\varphi_n'(x)]=\lim_{x\to b-0}[g(x)\varphi_m(x)\varphi_n'(x)] \qquad (m,n=1,2,\ldots) \tag{2.26}$$

(mit endlichen Grenzwerten) befriedigt.[21]

Damit nun $\{\varphi_\lambda'(x)\}$ ein g-orthogonales System inbezug auf (a,b) bildet, ist notwendig und hinreichend, daß die Relationen

$$[g(x)\varphi_\lambda'(x)]'=\varrho_\lambda w(x)\varphi_\lambda(x) \qquad (a<x<b;\ \lambda=1,2,\ldots) \tag{2.27}$$

mit

$$\varrho_\lambda=-\frac{\int_a^b g(x)\varphi_\lambda'(x)^2dx}{\int_a^b w(x)\varphi_\lambda(x)^2dx} \qquad (\lambda=1,2,\ldots) \tag{2.28}$$

bestehen.[22]

BEWEIS. Es seien unsere Forderungen erfüllt; dann gilt

$$\int_a^b w(x)\varphi_m(x)\varphi_n(x)\,dx=0 \qquad (m\neq n;\ m,n=1,2,\ldots). \tag{2.29}$$

1° *(Notwendigkeit.)* Es werde angenommen, daß $\{\varphi_\lambda'(x)\}$ inbezug auf (a,b) g-orthogonal ist, d. h.

$$\int_a^b g(x)\varphi_m'(x)\varphi_n'(x)\,dx=0 \qquad (m\neq n;\ m,n=1,2,\ldots). \tag{2.30}$$

Dann liefert eine Teilintegration des linksstehenden Ausdrucks wegen (2. 26) die Formel

$$\int_a^b g(x)\varphi_m'(x)\varphi_n'(x)\,dx=-\int_a^b[g(x)\varphi_m'(x)]'\varphi_n(x)\,dx \qquad (m,n=1,2,\ldots), \tag{2.31}$$

[21] Die erste Annahme: $w(x)^{-1}[g(x)\varphi_\lambda'(x)]'\in L_w^2(a,b)$ gilt z. B. von selbst, wenn die übrigen verlangten Eigenschaften von $w(x), g(x), \{\varphi_\lambda(x)\}$ im ganzen *(abgeschlossenen)* $[a,b]$ bestehen. — Was (2. 26) betrifft, so kann sie in mannigfaltiger Weise erfüllt sein, z. B. 1° wenn jedes $\varphi_\lambda(x)$ nebst der Stetigkeit von $g(x)\varphi_\lambda'(x)$ in $[a,b]$ für $x\to a+0$, $x\to b-0$ gegen 0 strebt, oder 2° durch die Beschränktheit von $\varphi_\lambda(x)$ und

$$\lim_{x\to a+0}[g(x)\varphi_\lambda'(x)]=\lim_{x\to b-0}[g(x)\varphi_\lambda'(x)]=0 \qquad (\lambda=1,2,\ldots),$$

ferner 3° nebst im ganzen $[a,b]$ stetigen $g(x)$, $\varphi_\lambda(x)$, $\varphi_\lambda'(x)$ und $g(a)+0$ durch $\varphi_\lambda(a)=M\varphi_\lambda(b)$, $\varphi_\lambda'(a)=M_1\varphi_\lambda'(b)$, $MM_1=g(a)^{-1}g(b)$ (M und M_1 Konstanten), usw. — Verlangt man statt der Differenzierbarkeitsbedingungen die Totalstetigkeit von $\varphi_\lambda(x)$ und $g(x)\varphi_\lambda'(x)$ in (a,b), so kann eine entsprechende Behauptung mit *fast überall* gültiger (2. 27) ausgesprochen werden.

[22] Satz 3 ist bei allen betrachteten klassischen Orthogonalsystemen (einschließlich der Laguerreschen und Hermiteschen Polynome) anwendbar.

woraus man insbesondere für $m = n$

$$\int_a^b [g(x)\varphi_m'(x)]'\varphi_m(x)\,dx = -\int_a^b g(x)\varphi_m'(x)^2\,dx, \tag{2.32}$$

sonst aber durch (2.30)

$$\int_a^b [g(x)\varphi_m'(x)]'\varphi_n(x)\,dx = 0 \qquad (m \neq n;\ m, n = 1, 2, \ldots) \tag{2.33}$$

erhält.

Auf Grund von (2.29) und (2.33) gilt nun

$$\int_a^b \{[g(x)\varphi_m'(x)]' - \varrho_m w(x)\varphi_m(x)\}\varphi_n(x)\,dx = 0 \quad (m \neq n;\ m, n = 1, 2, \ldots) \tag{2.34}$$

mit jedem reellen ϱ_m $(m = 1, 2, \ldots)$; verlangt man diese Gleichung auch für $m = n$, so ergibt sich infolge (2.32) eben die Wertbestimmung (2.28) von ϱ_λ.

Es sei diese Bedeutung von ϱ_m in Evidenz gehalten und der Kürze halber

$$G_m(x) = [g(x)\varphi_m'(x)]' - \varrho_m w(x)\varphi_m(x) \qquad (m = 1, 2, \ldots) \tag{2.35}$$

gesetzt; dann hat man also

$$\int_a^b \frac{G_m(x)}{w(x)}\varphi_n(x)w(x)\,dx = 0 \qquad (m, n = 1, 2, \ldots). \tag{2.36}$$

Wie man sieht, handelt es sich um die Fourier-Koeffizienten bezüglich $\{\varphi_\lambda(x)\}$ und der Belegung $w(x)$ der Funktion $G_m(x)/w(x)$, welche für jedes feste m wegen $\varphi_m(x) \in L_w^2(a, b)$ und $w(x)^{-1}[g(x)\varphi_m'(x)]' \in L_w^2(a, b)$ gewiß der Klasse $L_w^2(a, b)$ angehört. Da, der Voraussetzung nach, das genannte System w-vollständig ist, hat (2.36) die Relation

$$G_m(x) \equiv 0 \qquad (a < x < b;\ m = 1, 2, \ldots),$$

also, in der Tat,

$$[g(x)\varphi_m'(x)] = \varrho_m w(x)\varphi_m(x) \qquad (a < x < b;\ m = 1, 2, \ldots) \tag{2.37}$$

zur Folge.

2° *(Hinreichen.)* Gehe man, umgekehrt, von der Hypothese aus, daß (2.27) mit (2.28) besteht.

Dann kommen wir, durch Anwendung von (2.27) und (2.33), welche letzte wegen (2.26) ihre Gültigkeit behält, zur Relation

$$\int_a^b g(x)\varphi_m'(x)\varphi_n'(x)\,dx = -\varrho_m \int_a^b w(x)\varphi_m(x)\varphi_n(x)\,dx;$$

daraus geht aber nach (2.29) sofort die fragliche Orthogonalitätseigenschaft von $\{\varphi_\lambda'(x)\}$ hervor.

§ 3. Gewöhnliche SL-Systeme

3.1. Im Anschluß zu den Sätzen 2 und 3 ist es zweckmäßig eine Benennung einzuführen.

DEFINITION. *Es seien $g(x)$ und $w(x)$ in $a<x<b$ positive, stetig differenzierbare Funktionen, $\varphi_\lambda(x)\in L^2_w(a,b)$ $(\lambda=1,2,\ldots)$ sei eine Folge mit nichtkonstanten, linear unabhängigen, ebenda mit stetigen zweiten Ableitungen versehenen Elementen;*[23] *ferner seien $\varrho_\lambda\neq 0$ $(\lambda=1,2,\ldots)$ verschiedene reelle Zahlen.*

Man sagt, die Folge $\{\varphi_\lambda(x)\}$ sei vom gewöhnlichen Sturm—Liouvilleschen (kurz SL-) Typus (oder ein gewöhnliches SL-System) für (a,b) mit den Eigenwerten $\{\varrho_\lambda\}$ zu den Gewichtsfunktionen $(g(x),w(x))$, wenn 1) *die Relationen*

$$[g(x)\varphi_\lambda'(x)]'=\varrho_\lambda w(x)\varphi_\lambda(x) \qquad (a<x<b;\ \lambda=1,2,\ldots) \tag{3.1}$$

bestehen, 2) *die Funktionen $R_\lambda(x)=g(x)\varphi_\lambda'(x)$ $(\lambda=1,2,\ldots)$ in den Punkten a und b von rechts bzw. links stetig,*[24] 3) *die Grenzwerte von $\varphi_m(x)R_n(x)$ $(m,n=1,2,\ldots)$ für $x\to a+0$, $x\to b-0$ vorhanden und die Randbedingungen*

$$\lim_{x\to a+0}[\varphi_m(x)R_n(x)]=\lim_{x\to b-0}[\varphi_m(x)R_n(x)] \qquad (m,n=1,2,\ldots) \tag{3.2}$$

erfüllt sind.

Es ist nun mühelos einzusehen, daß die eben definierten Funktionenfolgen eine *solche Klasse orthogonaler Systeme bilden, welche fast alle klassischen Systeme umfaßt* und andererseits, daß man von ihnen *durch Differentiation und Integration wieder Orthogonalsysteme* erhält. Es gilt nämlich der folgende (zugleich die Übersicht der bisherigen Ergebnisse erleichternde)

SATZ 4. *Es sei $\varphi_\lambda(x)$ $(\lambda=1,2,\ldots)$ eine Folge vom gewöhnlichen SL-Typus für (a,b) mit den Eigenwerten $\{\varrho_\lambda\}$ zu den Gewichtsfunktionen $(g(x),w(x))$.*

Dann gelten die Behauptungen:

I) *$\{\varphi_\lambda(x)\}$, $\{\varphi_\lambda'(x)\}$ und zugleich $\{\Phi_\lambda(x)+c_\lambda\}$ mit*

$$\Phi_\lambda(x)=\int_a^x w(t)\varphi_\lambda(t)\,dt, \quad c_\lambda=\frac{1}{\varrho_\lambda}R_\lambda(a) \qquad (\lambda=1,2,\ldots)^{25} \tag{3.3}$$

[23] $g(x)$ und $w(x)$ mögen also in den Endpunkten $x=a$, $x=b$ zufällig verschwinden; sie können, ebenso wie $\varphi_\lambda(x)$ $(\lambda=1,2,\ldots)$, dabei auch unendlich werden.

[24] Wird die (zu starke) Forderung gestellt, daß die Stetigkeit von $g(x)$ und $\varphi_\lambda'(x)$ $(\lambda=1,2,\ldots)$ im ganzen $[a,b]$ gilt, so ist 2) natürlich überflüssig.

[25] Gehört $w(t)\varphi_\lambda(t)$ nicht zur Klasse $L(a,b)$, so soll $\int_a^x w(t)\varphi_\lambda(t)\,dt=\lim_{A\to a+0}\int_A^x w(t)\varphi_\lambda(t)\,dt$ $(a<A\leqq x<b)$ verstanden werden.

bilden je ein Orthogonalsystem für (a, b) und zwar, der Reihe nach, zu den Belegungsfunktionen $w(x)$, $g(x)$ und $Q(x) = g(x)^{-1}$.

Wenn $Q(x) \in L(a, b)$ und $\varphi_\lambda(x)$ im ganzen $[a, b]$ stetig mit $\varphi_\lambda(a) = \varphi_\lambda(b)$ $(\lambda = 1, 2, \ldots)$ ist, so kann $\{\Phi_\lambda(x) + c_\lambda\}$, wenn aber nebst $w(x) \in L(a, b)$ $R_\lambda(a) = R_\lambda(b)$ $(\lambda = 1, 2, \ldots)$ besteht, so mag $\{\varphi_\lambda(x)\}$ durch 1 erweitert werden.

II) *Es ist*

$$(3.4) \qquad \varrho_\lambda = -\left[\int_a^b g(x)\varphi_\lambda'(x)^2\,dx\right]\left[\int_a^b w(x)\varphi_\lambda(x)^2\,dx\right]^{-1} \qquad (\lambda = 1, 2, \ldots)^{26}$$

oder

$$(3.5) \qquad \varrho_\lambda = -\left[\int_a^b w(x)\varphi_\lambda(x)^2\,dx\right]\left[\int_a^b Q(x)(\Phi_\lambda(x) + c_\lambda)^2\,dx\right]^{-1} \qquad (\lambda = 1, 2, \ldots).$$

Man hat ferner, falls jedes $\varphi_\lambda(x)$ in $[a, b]$ stetig und $\int_a^b Q(x)\,dx$ vorhanden ist,

$$(3.6) \qquad c_\lambda = \left[\int_a^b Q(x)\,dx\right]^{-1}\left[\frac{\varphi_\lambda(b) - \varphi_\lambda(a)}{\varrho_\lambda} - \int_a^b Q(x)\Phi_\lambda(x)\,dx\right] \qquad (\lambda = 1, 2, \ldots).$$

III) *Sind die Funktionen $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ auch in $[a, b]$ stetig, so sind die Folgen $\{\Phi_\lambda(x) + c_\lambda\}$ und $\{R_\lambda(x)\}$ gleichfalls vom gewöhnlichen SL-Typus für (a, b) mit den Eigenwerten $\{\varrho_\lambda\}$, und zwar zu den Gewichtsfunktionen $(w(x)^{-1}, g(x)^{-1} = Q(x))$.*

BEWEIS. Wir nehmen an, daß unsere Bedingungen erfüllt sind; es sei $a < A < B < b$.

Wenn man in (3.1) $\lambda = m$, $\lambda = n$ setzt und die entstehenden Gleichungen mit $\varphi_n(x)$ bzw. $\varphi_m(x)$ multipliziert, so folgt durch Substraktion

$$(3.7) \qquad \{g(x)[\varphi_m'(x)\varphi_n(x) - \varphi_m(x)\varphi_n'(x)]\}' = (\varrho_m - \varrho_n)w(x)\varphi_m(x)\varphi_n(x);$$

integriert man noch von A bis B, so ergibt sich die Relation

$$(3.8) \qquad (\varrho_m - \varrho_n)\int_A^B w(x)\varphi_m(x)\varphi_n(x)\,dx = \varphi_n(x)R_m(x) - \varphi_m(x)R_n(x)\Big|_{x=A}^{x=B}$$
$$(m, n = 1, 2, \ldots).$$

Die rechte Seite strebt laut (3.2) für $A \to a + 0$, $B \to b - 0$ gegen Null; somit hat auch das linksstehende Integral einen Limes und wegen $\varrho_m \neq \varrho_n$

[26] Wie aus (3.4) und (3.5) hervorgeht, müssen Zähler und Nenner positiv, d. h. $\varrho_\lambda < 0$ $(\lambda = 1, 2, \ldots)$ sein. — Bei konkreten Folgen $\{\varphi_\lambda(x)\}$ können somit die Eigenwerte sofort berechnet und z. B. ihr asymptotisches Verhalten untersucht werden. Umgekehrt: sucht man $y = \varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ als Eigenfunktionen des durch $[g(x)y']' = \text{const.}\, y$ $(a < x < b)$ und Randbedingungen vom (3.2)-Typ angegebenen Randwertproblems, so ist das einschlägige *Variationsprinzip* eben auf (3.4) oder (3.5) gegründet. (Vgl. z. B. [24]; [19], S. 214 ff.)

$(m \neq n)$ kann man auf

$$\int_a^b w(x)\varphi_m(x)\varphi_n(x)\,dx = 0 \qquad (m \neq n;\ m, n = 1, 2, \ldots) \tag{3.9}$$

schließen.

Andererseits gibt eine partielle Integration und die Benutzung von (3.1)

$$\int_A^B g(x)\varphi_m'(x)\varphi_n'(x)\,dx = R_m(x)\varphi_n(x)\Big|_{x=A}^{x=B} - \varrho_m \int_A^B w(x)\varphi_m(x)\varphi_n(x)\,dx \tag{3.10}$$
$$(m, n = 1, 2, \ldots),$$

woher durch (3.2) und (3.9) bei $A \to a+0$, $B \to b-0$

$$\int_a^b g(x)\varphi_m'(x)\varphi_n'(x)\,dx = 0 \qquad (m \neq n;\ m, n = 1, 2, \ldots) \tag{3.11}$$

folgt.

Endlich besteht, wie (3.1) und (3.3) zeigt,

$$\varrho_\lambda[\Phi_\lambda(x) + c_\lambda] = \lim_{A \to a+0}\left[\varrho_\lambda \int_A^x w(t)\varphi_\lambda(t)\,dt + R_\lambda(A)\right] = R_\lambda(x) \tag{3.12}$$
$$(a < x < b;\ \lambda = 1, 2, \ldots),$$

so daß man durch (3.11) und (3.12) zu

$$\int_a^b Q(x)[\Phi_m(x) + c_m][\Phi_n(x) + c_n]\,dx = \frac{1}{\varrho_m \varrho_n}\int_a^b g(x)\varphi_m'(x)\varphi_n'(x)\,dx = 0 \tag{3.13}$$
$$(m \neq n;\ m, n = 1, 2, \ldots)$$

gelangt. — (3.9), (3.11), (3.13) enthalten eben die fraglichen Orthogonalitätseigenschaften von $\{\varphi_\lambda(x)\}$, $\{\varphi_\lambda'(x)\}$ bzw. $\{\Phi_\lambda(x) + c_\lambda\}$.

(3.12) impliziert auch die Gleichungen

$$\left.\begin{aligned} &\varrho_\lambda \int_A^B w(t)\varphi_\lambda(t)\,dt = R_\lambda(B) - R_\lambda(A) && (3.14)\\ &\varrho_\lambda \int_A^B Q(t)[\Phi_\lambda(t) + c_\lambda]\,dt = \varphi_\lambda(B) - \varphi_\lambda(A) && (3.15)\end{aligned}\right\} \quad (\lambda = 1, 2, \ldots).$$

Wenn nun $R_\lambda(a) = R_\lambda(b)$ oder $\varphi_\lambda(a) = \varphi_\lambda(b)$ $(\lambda = 1, 2, \ldots)$ gilt, so strebt die erste bzw. zweite Differenz rechter Hand für $A \to a+0$, $B \to b-0$ gegen Null, d. h. das entsprechende Integral über (a, b) muß verschwinden; dies bedeutet aber nichts anderes, als daß das Element 1 nebst $w(x) \in L(a, b)$ bzw. $Q(x) \in L(a, b)$ dem Orthogonalsysteme $\{\varphi_\lambda(x)\}$ bzw. $\{\Phi_\lambda(x) + c_\lambda\}$ angehört.

II) Aus (3. 10) entspringt bei $m = n = \lambda$ nach einem Grenzübergange durch (3. 2)

$$\varrho_\lambda \int_a^b w(x)\varphi_\lambda(x)^2\,dx = -\int_a^b g(x)\varphi_\lambda'(x)^2\,dx \qquad (\lambda = 1, 2, \ldots), \tag{3. 16}$$

was sofort (3. 4) liefert.[27]

Um (3. 5) zu bekommen, konstatieren wir vor allem auf Grund von (3. 12) die Gleichheit

$$[\Phi_\lambda(x) + c_\lambda] \cdot \left\{\varrho_\lambda \int_A^x Q(t)[\Phi_\lambda(t) + c_\lambda]\,dt + \varphi_\lambda(A)\right\}\Big|_{x=A}^{x=B} = \frac{1}{\varrho_\lambda}\,\varphi_\lambda(x)R_\lambda(x)\Big|_{x=A}^{x=B} \tag{3. 17}$$

$$(\lambda = 1, 2, \ldots).$$

Somit gelangt man aber durch Anwendung einer Teilintegration und des Grenzübergangs $A \to a + 0$, $B \to b - 0$ zur Formel

$$\int_a^b w(x)\varphi_\lambda(x)^2\,dx = -\varrho_\lambda \int_a^b Q(x)[\Phi_\lambda(x) + c_\lambda]^2\,dx \qquad (\lambda = 1, 2, \ldots), \tag{3. 18}$$

woher der fragliche zweite Ausdruck für ϱ_λ ablesbar ist.

Schließlich findet man, von der schon benutzten Relation (vgl. (3. 15))

$$\varrho_\lambda \int_a^b Q(t)[\Phi_\lambda(t) + c_\lambda]\,dt = \varphi_\lambda(b) - \varphi_\lambda(a) \tag{3. 19}$$

ausgehend, unmittelbar die Wertbestimmung (3. 6) für c_λ $(\lambda = 1, 2, \ldots)$, wofern nur die dort angegebenen Voraussetzungen gemacht werden.

III) Seien $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ in $[a, b]$ stetig angenommen; die Integrale

$$\int_a^b Q(x)R_\lambda(x)^2\,dx = \varrho_\lambda^2 \int_a^b Q(x)[\Phi_\lambda(x) + c_\lambda]^2\,dx$$

sind gleich $\int_a^b g(x)\varphi_\lambda'(x)^2\,dx$ $(\lambda = 1, 2, \ldots)$, existieren also laut (3. 16) nach Hypothese.

In diesem Falle ergeben (3. 1) und (3. 2) durch eine einfache Rechnung, daß auch die Folgen $\{R_\lambda(x)\}$ und $\{\Phi_\lambda(x) + c_\lambda\}$ die verlangten Eigenschaften 1)—3) gewöhnlicher SL-Systeme besitzen, w. z. b. w.

3. 2. Der Teil III) des eben bewiesenen Satzes legt die Frage nahe: *wann hat ein gewöhnliches SL-System eine Ableitungsfolge bzw. höhere Deriviertenfolge derselben Art?*[28]

[27] Das linksstehende Integral existiert nach Hypothese!

[28] Wie man weiß, liefern $\{\sin \lambda x\}$, $\{\cos \lambda x\}$, ferner die Jacobischen, Hermiteschen, Laguerreschen Polynome je Beispiele solcher SL-Systeme.

In dieser Richtung stützt man sich vor allem auf die folgende Bemerkung, welche sich durch einfache Umformungen verifizieren läßt:

LEMMA 1. *Genügt eine in (a, b) dreimal stetig differenzierbare Funktion $y = \varphi_\lambda(x)$ nebst ebenda einmal bzw. zweimal stetig differenzierbarer, positiver $w(x)$ und $g(x)$ der Gleichung* (3. 1), *so ist $y = \varphi'_\lambda(x)$ eine Lösung der Sturm—Liouvilleschen Differentialgleichung*

$$\left[\frac{g(x)^2}{w(x)}y'\right]' = \left\{\varrho_\lambda - \left[\frac{g'(x)}{w(x)}\right]'\right\} g(x)\,y \qquad (a < x < b). \tag{3. 20}$$

(3. 20) setzt einerseits in Evidenz die g-Orthogonalität von $\{\varphi'_\lambda(x)\}$, falls $\{\varphi_\lambda(x)\}$ eine gewöhnliche SL-Folge zu $(g(x), w(x))$ bildet (vgl. Satz 4, I)), andererseits aber zeigt sie, daß eine besondere Wahl der sukzessiven Gewichtsfunktionen, nämlich der Fall, wo diese (wie $w(x)$, $g(x)$, $g(x)^2 w(x)^{-1}$) *eine geometrische Progression* bilden, bei SL-Systemen ausgezeichnet ist.

In der Tat, gilt

SATZ 5. *Es sei $\{\varphi_\lambda(x)\}$ vom gewöhnlichen SL-Typus für (a, b) mit den Eigenwerten $\{\varrho_\lambda\}$ zu den Gewichtsfunktionen $(g(x), w(x))$; $g_k(x)$ $(k = 1, 2, \ldots)$ seien in (a, b) positive, zweimal stetig differenzierbare Funktionen derart, daß $w(x)$, $g_1(x)$, $g_2(x)$, ... eine geometrische Progression bildet. Wir nehmen an, daß die zur Anwendung kommenden Ableitungen von $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ in (a, b) vorhanden und stetig sind, und ferner, daß die Folgen $\{\varphi_\lambda^{(k)}(x)\}$, für jedes in Betracht kommende k, keine Konstante enthalten.*

I) *Dann hat man dafür, damit auch $\{\varphi'_\lambda(x)\}$ eine gewöhnliche SL-Folge für (a, b) zu $(g_2(x), g_1(x))$ bildet, folgende notwendige und hinreichende Bedingungen:*

1) *Es sei $g'_1(x)w(x)^{-1} = Px + Q$ $(a < x < b)$ mit konstanten P und Q;* 2) *die Funktionen $g_2(x)\varphi''_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ seien für $x \to a+0$, $x \to b-0$ stetig;* 3) *das Produkt $(Px+Q)g_1(x)\varphi'_\lambda(x)^2$ habe denselben Limes η_λ $(\lambda = 1, 2, \ldots)$ für $x \to a+0$, $x \to b-0$; bei jedem λ, wofür $\eta_\lambda \neq 0$ ist, besitze $\lim_{x \to a+0} \operatorname{sg} \varphi'_\lambda(x) \cdot \lim_{x \to b-0} \operatorname{sg} \varphi'_\lambda(x)$ denselben Wert.*[29]

Unter diesen Voraussetzungen ist die zu $\{\varphi'_\lambda(x)\}$ gehörige Eigenwertfolge $\{\varrho_\lambda - P\}$.

II) *Man wähle $\nu \geqq 2$. Die Systeme $\{\varphi'_\lambda(x)\}, \{\varphi''_\lambda(x)\}, \ldots, \{\varphi_\lambda^{(\nu)}(x)\}$ sind dann und nur dann gleichzeitig vom gewöhnlichen SL-Typus und zwar, der Reihe nach, zu den Gewichtsfunktionen $(g_2(x), g_1(x))$, $(g_3(x), g_2(x)), \ldots, (g_{\nu+1}(x), g_\nu(x))$, wenn die folgenden Bedingungen erfüllt sind:*

[29] Dies bedeutet offenbar, daß jedes $\varphi'_\lambda(x)$ (für die betreffenden λ) in der Nähe von a und b das gleiche Vorzeichen hat.

1) $g_1'(x)w(x)^{-1} = Px + Q$, $g_1(x)w(x)^{-1} = px^2 + qx + r$ $(a < x < b)$ *mit konstanten* P, Q, p, q, r; 2) $g_1(x)\left(\frac{g_1(x)}{w(x)}\right)^k \varphi_\lambda^{(k+1)}(x)$ *ist in den Punkten* a *und* b *von rechts bzw. links stetig* $(k = 1, 2, \ldots, \nu;\ \lambda = 1, 2, \ldots)$; 3) *der Ausdruck*

$$g_1(x)\left(\frac{g_1(x)}{w(x)}\right)^{k-1}\left[\frac{g_1'(x)}{w(x)} + (k-1)\left(\frac{g_1(x)}{w(x)}\right)'\right]\varphi_\lambda^{(k)}(x)^2$$

strebt gegen denselben (endlichen) Grenzwert $\eta_{\lambda k}$ $(k = 1, 2, \ldots, \nu;\ \lambda = 1, 2, \ldots)$ *für* $x \to a + 0$, $x \to b - 0$; 4) *bei jedem fixen* k *und sämtlichen* λ *mit* $\eta_{\lambda k} \neq 0$ *ist das Produkt* $\lim_{x \to a+0} \operatorname{sg} \varphi_\lambda^{(k)}(x) \cdot \lim_{x \to b-0} \operatorname{sg} \varphi_\lambda^{(k)}(x)$ *konstant* $(k = 1, 2, \ldots, \nu)$.

Unter diesen Annahmen ist die mit dem Systeme $\{\varphi_\lambda^{(k)}(x)\}$ *verknüpfte Eigenwertfolge* $\{\varrho_\lambda - kP - k(k-1)p\}$ $(k = 1, 2, \ldots, \nu)$.

KOROLLARIEN. (*) Die Anwendung des Satzes 4 ergibt u. a.: neben $\{\varphi_\lambda(x)\}$ und $\{\varphi_\lambda'(x)\}$ ist auch $\{q_\lambda''(x)\}$ orthogonal für (a, b) (zur Belegung $g_2(x)$), falls die Bedingungen II) 1)—4) erfüllt sind, ferner bilden ja auch die höheren Ableitungen bis zu $\varphi_\lambda^{(\nu)}(x)$ $(\lambda = 1, 2, \ldots)$ Orthogonalsysteme für (a, b) (zu den Belegungen $g_k(x)$; $k = 2, 3, \ldots, \nu$), wenn die Forderungen III) 1)—4)) befriedigt sind.

(**) Folgende Bedingungen sind offenbar *hinreichend*, weil sie die Erfüllung der entsprechenden oben erwähnten Bedingungen nach sich ziehen: 1) $g_1(x)$ sei eine Konstante und 2) $\varphi_\lambda''(x)w(x)^{-1}$ $(\lambda = 1, 2, \ldots)$ für $x \to a + 0, x \to b - 0$ stetig; oder samt II) 1) und 2) noch drittens: $\lim [g_1(x)\varphi_\lambda'(x)^2]$ und $\lim [xg_1(x)\varphi_\lambda'(x)^2]$ für $x \to a + 0$, $x \to b - 0$ verschwinde (beide statt II) 1)—4); neben III) 1) und 2) noch drittens: $x^m g_1(x)\varphi_\lambda^{(k)}(x)^2$ habe den Grenzwert Null für $x \to a + 0$, $x \to b - 0$ bei allen Werten $m = 0, 1, \ldots, 2\nu - 1$; $k = 1, 2, \ldots, \nu$ (statt III) 1)—4)).

(***) Man erkennt leicht, daß die Bedingung III) 1) folgendes impliziert: $g_1(x)$ soll eine Konstante oder ein Ausdruck der Form

$$\exp\left(\int \frac{Px + Q}{px^2 + qx + r}\, dx\right)$$

sein, d. h. sich jedenfalls mit Hilfe von rationalen Funktionen, Logarithmus bzw. Arcustangens und Exponentialfunktion aufbauen lassen, während andererteils die Darstellung $w(x) = (px^2 + qx + r)^{-1} g_1(x)$ gelten muß.

BEWEIS. Es seien die über die SL-Folge[30] $\{\varphi_\lambda(x)\}$, ferner über $\{\varrho_\lambda\}$, $g_1(x)$, $w(x)$ gemachten Forderungen in Evidenz gehalten.

I) 1° *(Hinreichen.)* Wir nehmen an, daß auch die Bedingungen I) 1)—4) befriedigt sind und haben zu zeigen, daß $\{\varphi_\lambda'(x)\}$ die Eigenschaften einer

[30] Wir benutzen diese Bezeichnung, der Kürze halber, statt „Folge vom gewöhnlichen SL-Typus".

SL-Folge für (a, b) zu $(g_2(x), g_1(x))$ besitzt. (Vgl. die oben angegebene Definition.)

Vor allem bestätigt man, daß (3. 16) nach Hypothese besteht, womit die Existenz der Integrale $\int_a^b g(x)\varphi'_\lambda(x)^2 dx$ $(\lambda = 1, 2, \ldots)$ gesichert ist.

Wegen I) 1) ergibt nun (3. 20)

$$[g_2(x)\varphi''_\lambda(x)]' = (\varrho_\lambda - P)g_1(x)\varphi'_\lambda(x), \tag{3.21}$$

entsprechend (3. 1);[31] andererteils hat man die Identität für jedes m, n

$$g_2(x)\varphi'_m(x)\varphi''_n(x) = \frac{g_1(x)}{w(x)}\varphi'_m(x)\{[g_1(x)\varphi'_n(x)]' - g'_1(x)\varphi'_n(x)\}, \tag{3.22}$$

welche durch (3. 1) und die Bedingung I) 1) in

$$g_2(x)\varphi'_m(x)\varphi''_n(x) = \varrho_n g_1(x)\varphi_n(x)\varphi'_m(x) - g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x) \tag{3.23}$$

transformiert werden kann. Das erste Glied rechter Hand hat nach (3. 2) übereinstimmende (endliche) Grenzwerte für $x \to a+0$ und $x \to b-0$; dasselbe gilt bezüglich des zweiten Gliedes. Denn zunächst ist auf Grund von I) 3) (angewendet mit $\lambda = m, \lambda = n$)

$$\begin{aligned} &\lim_{x\to a+0}[g_1(x)^2(Px+Q)^2\varphi'_m(x)^2\varphi'_n(x)^2] = \\ &= \lim_{x\to b-0}[g_1(x)^2(Px+Q)^2\varphi'_m(x)^2\varphi'_n(x)^2] = \eta_m\eta_n, \end{aligned} \tag{3.24}$$

also jedenfalls

$$\begin{aligned} &\lim_{x\to a+0}|g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)| = \\ &= \lim_{x\to b-0}|g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)| = \sqrt{\eta_m\eta_n}. \end{aligned} \tag{3.25}$$

Ist $\eta_m\eta_n = 0$, so kommt man unmittelbar zu

$$\lim_{x\to a+0}[g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)] = \lim_{x\to b-0}[g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)]. \tag{3.26}$$

Widrigenfalls (d. h. bei $\eta_m\eta_n > 0$) kann weder $\operatorname{sg}\varphi'_m(x)$, noch $\operatorname{sg}\varphi'_n(x)$ in den Endpunkten a und b den Grenzwert Null haben, so daß aus I) 4)

$$\frac{\lim_{x\to a+0}\operatorname{sg}\varphi'_m(x)}{\lim_{x\to b-0}\operatorname{sg}\varphi'_m(x)} = \frac{\lim_{x\to a+0}\operatorname{sg}\varphi'_n(x)}{\lim_{x\to b-0}\operatorname{sg}\varphi'_n(x)} \tag{3.27}$$

folgt und sogar, da wegen I) 3)

$$\lim_{x\to a+0}\operatorname{sg}(Px+Q) = \lim_{x\to b-0}\operatorname{sg}(Px+Q) \tag{3.28}$$

[31] Dies zeigt zugleich, die zu $\{\varphi'_\lambda(x)\}$ gehörige Eigenwertfolge sei $\{\varrho_\lambda - P\}$, wie behauptet.

besteht, auch

$$\lim_{x\to a+0} \operatorname{sg}\left(g_1(x)(Px+Q)\varphi_m'(x)\varphi_n'(x)\right) = \lim_{x\to b-0} \operatorname{sg}\left(g_1(x)(Px+Q)\varphi_m'(x)\varphi_n'(x)\right). \tag{3.29}$$

Aus (3.29) und (3.25) folgt wieder (3.26).

Somit gelangt man endlich durch (3.23) zu

$$\lim_{x\to a+0}\left[g_2(x)\varphi_m'(x)\varphi_n''(x)\right] = \lim_{x\to b-0}\left[g_2(x)\varphi_m'(x)\varphi_n''(x)\right] \quad (a<x<b;\ m,n=1,2,\ldots). \tag{3.30}$$

Unter Hinzunahme der Bedingung I) 2) sieht man also, daß sämtliche fragliche Eigenschaften von $\{\varphi_\lambda'(x)\}$ bestehen.

2° *(Notwendigkeit.)* Sei $\{\varphi_\lambda'(x)\}$ als ein System vom gewöhnlichen SL-Typus für (a, b) zu $(g_2(x), g_1(x))$ angenommen.

Dann muß für jedes λ eine Relation der Form

$$[g_2(x)\varphi_\lambda''(x)]' = C_\lambda g_1(x)\varphi_\lambda'(x) \qquad (a<x<b;\ C_\lambda = \text{const.}), \tag{3.31}$$

ferner (3.30), sowie die einseitige Stetigkeit der Funktionen $g_2(x)\varphi_\lambda''(x)$ $(\lambda = 1, 2, \ldots)$ in a und b gelten.

Das letzte ist eben die Bedingung I) 2); der Vergleich von (3.20) und (3.31) ergibt die Forderung: $\left[\frac{g_1'(x)}{w(x)}\right]' = \text{const.}$, d. h. I) 1). Um noch I) 3) und 4) zu erhalten, fassen wir die aus (3.22) und (3.1) fließende Formel:

$$g_2(x)\varphi_m'(x)\varphi_n''(x) = \varrho_n g_1(x)\varphi_n(x)\varphi_m'(x) - g_1(x)\frac{g_1'(x)}{w(x)}\varphi_m'(x)\varphi_n'(x) \quad (a<x<b;\ m,n=1,2,\ldots) \tag{3.32}$$

ins Auge. Daraus folgt für $m = n = \lambda$ $(\lambda = 1, 2, \ldots)$

$$g_1(x)\frac{g_1'(x)}{w(x)}\varphi_\lambda'(x)^2 = \varrho_\lambda g_1(x)\varphi_\lambda(x)\varphi_\lambda'(x) - g_2(x)\varphi_\lambda'(x)\varphi_\lambda''(x), \tag{3.33}$$

wobei (nach Hypothese) sowohl das zweite, als das dritte Glied für $x\to a+0$, $x\to b-0$ übereinstimmende (endliche) Grenzwerte hat. Somit kann man unter Beachtung von I) 1) auf I) 3) schließen.

Gleichfalls impliziert aber (3.32) die Gleichungen

$$\lim_{x\to a+0}\left[g_1(x)(Px+Q)\varphi_m'(x)\varphi_n'(x)\right] = \lim_{x\to b-0}\left[g_1(x)(Px+Q)\varphi_m'(x)\varphi_n'(x)\right] = \xi_{mn} \quad (m,n=1,2,\ldots). \tag{3.34}$$

Sind nun m und n zwei beliebige natürliche Zahlen mit $\eta_m \neq 0$, $\eta_n \neq 0$, so hat man durch (3.24) $\xi_{mn} \neq 0$ und wegen (3.34)

$$\lim_{x\to a+0} \operatorname{sg}\left(g_1(x)(Px+Q)\varphi_m'(x)\varphi_n'(x)\right) = \lim_{x\to b-0} \operatorname{sg}\left(g_1(x)(Px+Q)\varphi_m'(x)\varphi_n'(x)\right) \neq 0; \tag{3.35}$$

da dies für $m = n$ zeigt, daß

$$\lim_{x \to a+0} \operatorname{sg}(g_1(x)(Px+Q)) = \lim_{x \to b-0} \operatorname{sg}(g_1(x)(Px+Q))$$

sein soll, erhalten wir

$$\frac{\lim\limits_{x \to a+0} \operatorname{sg} \varphi_m'(x)}{\lim\limits_{x \to b-0} \operatorname{sg} \varphi_m'(x)} = \frac{\lim\limits_{x \to b-0} \operatorname{sg} \varphi_n'(x)}{\lim\limits_{x \to a+0} \operatorname{sg} \varphi_n'(x)}$$

oder

$$\text{(3.36)} \quad \lim_{x \to a+0} \operatorname{sg} \varphi_m'(x) \cdot \lim_{x \to b-0} \operatorname{sg} \varphi_m'(x) = \lim_{x \to a+0} \operatorname{sg} \varphi_n'(x) \cdot \lim_{x \to b-0} \operatorname{sg} \varphi_n'(x).$$

(3.36) besagt dasselbe, als I) 4).

II) Die Untersuchung des zweiten Teils begründen wir zweckmäßig auf den folgenden kleinen *Hilfssatz:*

Es sei k positiv-ganz und $\{\varphi_\lambda^{(k-1)}(x)\}$ eine Folge vom gewöhnlichen SL-Typus für (a, b) zu $(g_k(x), g_{k-1}(x))$. In diesem Falle bildet $\{\varphi_\lambda^{(k)}(x)\}$ dann und nur dann eine SL-Folge für (a, b) zu $(g_{k+1}(x), g_k(x))$, wenn $g_1'(x)w(x)^{-1} + (k-1)[g_1(x)w(x)^{-1}]'$ ein Polynom höchstens ersten Grades ist und die Bedingungen II) 2)—4) *für das betreffende* k bestehen.

In der Tat, wendet man den soeben bewiesenen Teil I) auf das System $\{\varphi_\lambda^{(k-1)}(x)\}$ an, wobei $g_k(x)$, $g_{k-1}(x)$ statt der dortigen $g_1(x)$ bzw. $w(x)$ geschrieben werden sollen, so ergibt sich unmittelbar die Notwendigkeit und Hinlänglichkeit der genannten vier Bedingungen.

Nun ist der Fall $\nu = 2$ der Behauptung II) leicht zu erledigen.

1° Sei angenommen, daß II) 1)—4) für $k = 1$, $k = 2$ erfüllt sind. Aus II) 1) und II) 2)—4) mit $k = 1$ können wir gemäß des Teils I) schließen, $\{\varphi_\lambda'(x)\}$ sei eine SL-Folge für (a, b) zu $(g_2(x), g_1(x))$; dies und die Bedingungen II) 1), II) 2)—4) mit $k = 2$ involvieren aber nach dem Hilfssatze, daß auch $\{\varphi_\lambda''(x)\}$ ein gewöhnliches SL-System für (a, b) zu $(g_3(x), g_2(x))$ bildet.

2° Setzen wir voraus, $\{\varphi_\lambda'(x)\}$ und $\{\varphi_\lambda''(x)\}$ seien vom gewöhnlichen SL-Typus für (a, b) zu $(g_2(x), g_1(x))$ bzw. $(g_3(x), g_2(x))$. Dann ergibt der Teil I) die Gültigkeit der Bedingungen I) 1)—4), welche durch den Hilfssatz zu II) 2)—4) mit $k = 1$, $k = 2$, ferner zu

$$\text{(3.37)} \quad \frac{g_1'(x)}{w(x)} = \text{linear}, \quad \frac{g_1'(x)}{w(x)} + \left[\frac{g_1'(x)}{w(x)}\right]' = \text{linear}$$

erweitert werden können. Die beiden letzten liefern II) 1).

Ist $\nu > 2$, so führt eine Induktion nebst Benutzung des Hilfssatzes einfach zum Ziele.

Endlich ist es nicht schwer, unsere Behauptung über die Eigenwerte zu verifizieren.

Es sei $\{\varphi_\lambda^{(k-1)}(x)\}$ eine SL-Folge für (a, b) zu $(g_k(x), g_{k-1}(x))$ und $\{\varphi_\lambda^{(k)}(x)\}$ auch eine solche zu $(g_{k+1}(x), g_k(x))$ $(k=1, 2, \ldots, \nu)$; dann sieht man durch Anwendung des Teils II) sofort ein, daß die zu $\{\varphi_\lambda^{(k)}(x)\}$ gehörige Eigenwertfolge nur $\{\varrho_\lambda - kP - k(k-1)p\}$ sein kann. $y=\varphi_\lambda^{(k)}(x)$ genügt also der Differentialgleichung

$$(3.38)\quad [g_{k+1}(x)\varphi_\lambda^{(k+1)}(x)]' = [\varrho_\lambda - kP - k(k-1)p]g_k(x)\varphi_\lambda^{(k)}(x) \quad (k=1, 2, \ldots, \nu).$$

§ 4. Anwendungen: Charakterisierung der „klassischen" Orthogonalsysteme

4. 1. Unsere bisherigen Sätze zeigen, daß ein Orthogonalsystem $\{\varphi_\lambda(x)\}$ stetiger Funktionen, welches die betreffenden Eigenschaften besitzt, nur aus solchen Elementen bestehen kann, die einer Sturm—Liouvilleschen Differentialgleichung genügen; andererteils, wie erwähnt, haben jedenfalls die klassischen orthogonalen Systeme einer oder andere jener (mit Orthogonalität, Differentiation und Integration verknüpften) Eigenschaften.

Dies ermöglicht uns offenbar, eine Klasse von Sturm—Liouvilleschen Eigenfunktionen, insbesondere das trigonometrische System und die klassischen orthogonalen Polynome in einfacher Weise zu charakterisieren.

SATZ 6. *Es seien $g(x), w(x)$ in einem endlichen Intervalle (a, b) stetig differenzierbare und positive*[32] *Funktionen mit $g(x)^{-1} \in L(a, b)$, $w(x) \in L(a, b)$, für welche eines der beiden Randwertprobleme: „gesucht die in $[a, b]$ beschränkten Integrale $y=y(x)$ von*

$$(4.1)\qquad [g(x)y']' = \varrho w(x)y \qquad (a<x<b;\ \varrho \text{ ein reelles Parameter})$$

mit hierin stetiger $g(x)y'$ und mit (I) $y(a+0)=y(b-0)=0$ *oder* (II) $\lim_{x\to a+0}[g(x)y'(x)] = \lim_{x\to b-0}[g(x)y'(x)] = 0$" *nicht-triviale Lösungen besitzt. Wir bezeichnen die von Null verschiedenen (einfachen) Eigenwerte des betreffenden Problems mit $(0>)\varrho_1>\varrho_2>\cdots$ und die entsprechenden w-normierten Eigenfunktionen, der Reihe nach, mit $\varphi_1(x), \varphi_2(x), \ldots$; das letzte System sei bei* (I) *selbst, bei* (II) *samt* $\gamma_0 = \left[\int_a^b w(x)dx\right]^{-\frac{1}{2}} (>0)$ *für (a, b) w-abgeschlossen. Man schreibt noch* $\Phi_\lambda(x) = \int_a^x w(t)\varphi_\lambda(t)dt$ $(\lambda=1, 2, \ldots)$.

Dann läßt sich $\{\varphi_\lambda(x)\}$ bis auf Vorzeichen und Reihenfolge folgendermaßen charakterisieren: im Falle der Randbedingung (I) *durch*

1)* *$\{\varphi_\lambda(x)\}$ bildet auf (a, b) ein w-vollständiges, w-orthonormales System von nicht-konstanten, stetigen Funktionen;*

[32] Das Verschwinden der Gewichtsfunktionen $g(x), w(x)$ in a und b ist also erlaubt.

2)* *das System* 1, $\{\Phi_\lambda(x)+c_\lambda\}$ *ist* $g(x)^{-1}$*-orthogonal inbezug auf* (a, b), *wobei* c_λ $(\lambda=1, 2, \ldots)$ *geeignet gewählte Konstanten bedeuten;*

im Falle von (II) *aber durch*

1)** $\varphi_\lambda(x)$ $(\lambda=1, 2, \ldots)$ *sind in* (a, b) *stetig;* $\gamma_0, \{\varphi_\lambda(x)\}$ *bildet hier ein w-vollständiges, w-orthonormales System;*

2)** $\{\Phi_\lambda(x)\}$ *ist* $g(x)^{-1}$*-orthogonal inbezug auf* (a, b).

BEWEIS. 1° Wir nehmen an, daß alle Bedingungen des Satzes erfüllt sind.

Dann ist $\{\varphi_\lambda(x)\}$ *ein gewöhnliches SL-System,* verknüpft mit der Eigenwertfolge $\{\varrho_\lambda\}$ und den Belegungsfunktionen $(g(x), w(x))$, so daß Satz 4 angewendet werden kann. Unter Beachtung der Annahmen bezüglich der Normiertheit und Abgeschlossenheit von $\{\varphi_\lambda(x)\}$ bzw. $\gamma_0, \{\varphi_\lambda(x)\}$, erhält man genau die Behauptungen 1)*, 2)* (mit $c_\lambda=\frac{1}{\varrho_\lambda}\lim_{x\to a+0}[g(x)\varphi_\lambda'(x)]$) oder 1**, 2)**, je nach dem Falle (I) bzw. (II).

2° Es sei $\{\overline{\varphi}_\lambda(x)\}$ ein Funktionensystem mit den Eigenschaften 1)* und 2)*. Nun wendet man Satz 2, Teil I) mit $Q(x)=g(x)^{-1}$ an und erhält, daß 1. $\overline{\varphi}_\lambda(x)$ und zugleich $g(x)\overline{\varphi}_\lambda'(x)$ eine stetige Ableitung für $a<x<b$, $\lambda=1, 2, \ldots$ hat; 2. die Relationen

$$[g(x)\overline{\varphi}_\lambda'(x)]'=\bar{\varrho}_\lambda w(x)\overline{\varphi}_\lambda(x) \qquad (a<x<b;\ \lambda=1, 2, \ldots) \tag{4.2}$$

und 3. die Randbedingungen

$$\overline{\varphi}_\lambda(a+0)=\overline{\varphi}_\lambda(b-0)=0,\quad \lim_{x\to a+0}[g(x)\overline{\varphi}_\lambda'(x)]=\bar{\varrho}_\lambda\bar{c}_\lambda \tag{4.3}$$

mit

$$\bar{\varrho}_\lambda=-\left\{\int_a^b\frac{1}{g(x)}\left[\bar{\Phi}_\lambda(x)+\bar{c}_\lambda\right]^2dx\right\}^{-1},\quad \bar{c}_\lambda=-\left[\int_a^b\frac{dx}{g(x)}\right]^{-1}\int_a^b\frac{1}{g(x)}\bar{\Phi}_\lambda(x)dx \tag{4.4}$$

$$(\lambda=1, 2, \ldots)$$

befriedigt sind.

Daher schließt man aber wegen der w-Vollständigkeit von $\{\overline{\varphi}_\lambda(x)\}$ unmittelbar, daß $\{\bar{\varrho}_\lambda\}$ mit dem Spektrum $\{\varrho_\lambda\}$ zusammenfällt, und weiter, daß $\{\overline{\varphi}_\lambda(x)\}$, abgesehen von Vorzeichen und Reihenfolge, mit dem Systeme $\{\varphi_\lambda(x)\}$ identisch sein muß.

3° Falls $\{\overline{\varphi}_\lambda(x)\}$ die Eigenschaften 1)**, 2)** besitzt, so folgt aus Satz 2, Teil II) ähnlicherweise, daß $\{|\overline{\varphi}_\lambda(x)|\}$ nach einer geeigneten Umordnung in $\{|\varphi_\lambda(x)|\}$ übergeht, w. z. b. w.

4. 2. Satz 6 liefert z. B. eine Charakterisierung der trigonometrischen Systeme $\left\{\sqrt{\frac{2}{\pi}}\sin\lambda x\right\}$ und $\frac{1}{\sqrt{\pi}}$, $\left\{\sqrt{\frac{2}{\pi}}\cos\lambda x\right\}$ $(a=0, b=\pi, g(x)=w(x)\equiv 1)$, ferner des normierten Systems der Jacobischen Polynome $(a=-1, b=1, g(x)=$

$= (1-x)^{\alpha+1}(1+x)^{\beta+1}$, $w(x) = (1-x)^{\alpha}(1+x)^{\beta}$, $-1 < \alpha < 0$, $-1 < \beta < 0$). Das erste ist ein Beispiel für den Fall (4. 1) (I), das zweite und dritte entsprechen aber dem Falle (4. 1) (II).

Insbesondere ergibt sich: *für die Tschebyscheffschen Polynome erster Art* $T_\lambda(x) = \cos(\lambda \arccos x)$ $(\lambda = 0, 1, 2, \ldots)$ *sind, abgesehen von konstanten Faktoren, folgende Eigenschaften charakteristisch:* 1) $T_\lambda(x)$ $(\lambda = 0, 1, 2, \ldots)$ *sind in* $(-1, 1)$ *stetig und bilden hier ein* $(1-x^2)^{-\frac{1}{2}}$*-vollständiges,* $(1-x^2)^{-\frac{1}{2}}$*-orthogonales System;* 2) $T_\lambda^*(x) = \int_{-1}^{x} (1-u^2)^{-\frac{1}{2}} T_\lambda(u)\,du$ $(\lambda = 1, 2, \ldots)$ *ist* $(1-x^2)^{\frac{1}{2}}$*-orthogonal inbezug auf* $(-1, 1)$.

Es sei noch, allgemeiner, *der Fall in* $[a, b]$ *stetiger und positiver* $g(x)$ *und* $w(x)$ hervorgehoben; dann bietet Satz 6, wie man sofort einsieht, *eine Charakterisierung jedes Systems von Sturm—Liouvilleschen Eigenfunktionen verknüpft mit*

$$[g(x)y']' = \varrho w(x) y \qquad (y = y(x);\ a \leqq x \leqq b)$$

und den Randbedingungen $y(a) = y(b) = 0$ *oder* $y'(a) = y'(b) = 0$.

4. 3. Die Funktionenfolge

(4. 5) $$1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots$$

hat die Eigenschaft, daß nicht nur sie selbst, sondern auch ihre erste und höhere Deriviertensysteme inbezug auf ein beliebiges Intervall von Länge 2π (zur Belegung 1) *orthogonal* sind; ein jedes dieser Systeme wird *vollständig*, wenn man 1 als Element hinzufügt. Wir wollen zeigen, daß diese Festsetzungen schon für (4.5) charakteristisch sind, ja sogar.

SATZ 7. *Es seien* $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ *in einem endlichen Intervalle* $[a, b]$ *zweimal stetig differenzierbare Funktionen* $\not\equiv 0$, *welche hier ein vollständiges Orthogonalsystem bilden und den Anfangsbedingungen* $\varphi_\lambda(a)\varphi_\lambda'(a) = 0$ $(\lambda = 1, 2, \ldots)$ *genügen; es sei auch* $1, \{\varphi_\lambda'(x)\}$, *gleichwie* $1, \{\varphi_\lambda''(x)\}$ *orthogonal inbezug auf* $[a, b]$.[33]

Dann stimmt $\{\varphi_\lambda(x)\}$ *bis auf Reihenfolge und triviale lineare Transformationen mit* (4. 5) *überein.*

BEWEIS. Unsere Annahmen implizieren

(4. 6) $$\begin{cases} \int_a^b \varphi_\lambda'(x)\,dx = 0, & \varphi_\lambda(a) = \varphi_\lambda(b) \\ \int_a^b \varphi_\lambda''(x)\,dx = 0, & \varphi_\lambda'(a) = \varphi_\lambda'(b) \end{cases} \qquad (\lambda = 1, 2, \ldots)$$

[33] Statt der Orthogonalität von $1, \{\varphi_\lambda'(x)\}$ und $1, \{\varphi_\lambda''(x)\}$ mag nur diejenige von $\{\varphi_\lambda'(x)\}$ nebst den Randbedingungen $\varphi_\lambda(a) = \varphi_\lambda(b)$, $\varphi_\lambda'(a) = \varphi_\lambda'(b)$ $(\lambda = 1, 2, \ldots)$ verlangt werden. — In diesem Satze ist die Orthogonalität offenbar in engerem Sinne, d. h. mit der Belegung 1, zu verstehen.

und erlauben die Anwendung des Satzes 3 mit $g(x)=w(x)\equiv 1$. Man sieht, daß die Bedingungen (2.26) wegen (4.6) erfüllt sind; somit ergibt die Orthogonalität von $\{\varphi_\lambda(x)\}$, $\{\varphi'_\lambda(x)\}$ inbezug auf $[a,b]$ und die Vollständigkeit des ersten Systems

$$\varphi''_\lambda(x)=\varrho_\lambda\varphi_\lambda(x) \qquad (a<x<b;\ \lambda=1,2,\ldots) \tag{4.7}$$

mit

$$\varrho_\lambda=-\left[\int_a^b \varphi'_\lambda(x)^2\,dx\right]\left[\int_a^b \varphi_\lambda(x)^2\,dx\right]^{-1} \qquad (\lambda=1,2,\ldots). \tag{4.8}$$

Durch die Transformation

$$t=\frac{2\pi}{b-a}(x-a), \quad x=a+\frac{b-a}{2\pi}t \tag{4.9}$$

ergibt es sich also: $y=\varphi_\lambda\left(a+\frac{b-a}{2\pi}t\right)$ $(\lambda=1,2,\ldots)$ genügen der Differentialgleichung

$$\frac{d^2y}{dt^2}=\varrho_\lambda\left(\frac{b-a}{2\pi}\right)^2 y \qquad (0<t<2\pi) \tag{4.10}$$

nebst den Bedingungen

$$y(0)=y(2\pi), \quad y'(0)=y'(2\pi). \tag{4.11}$$

Das Randwertproblem $\ddot{y}=\varrho y$ mit (4.11) hat aber bekanntlich die Eigenwerte $\varrho=-n^2$ $(n=0,1,2,\ldots)$, so daß

$$\varrho_\lambda\cdot\left(\frac{b-a}{2\pi}\right)^2=-n^2, \quad \varrho_\lambda=-\left(\frac{2\pi}{b-a}\right)^2 n^2 \tag{4.12}$$

für jede $\lambda=1,2,\ldots$ mit einem geeigneten ganzen n gelten muß; jedem Eigenwerte gehören zwei linear unabhängige, bis auf konstante Faktoren bestimmte Eigenfunktionen: $\cos nt$, $\sin nt$ an.[34]

Wegen der Vollständigkeit von $\{\varphi_\lambda(x)\}$ inbezug auf $[a,b]$ und auf Grund der Bedingungen $\varphi_\lambda(a)\varphi'_\lambda(a)=0$ $(\lambda=1,2,\ldots)$ muß die Folge $\left\{\varphi_\lambda\left(a+\frac{b-a}{2\pi}t\right)\right\}$, von konstanten Faktoren abgesehen, mitdem Systeme 1, $\cos t$, $\sin t$, $\cos 2t$, $\sin 2t,\ldots$ in $0\leqq t\leqq 2\pi$ übereinstimmen; d. h. ist

$$\varphi_1(x)=C_1,\ \varphi_2(x)=C_2\cos\left(\frac{2\pi}{b-a}(x-a)\right),\ \varphi_3(x)=C_3\sin\left(\frac{2\pi}{b-a}(x-a)\right),$$

$$\varphi_4(x)=C_4\cos\left(\frac{4\pi}{b-a}(x-a)\right),\ \varphi_5(x)=C_5\sin\left(\frac{4\pi}{b-a}(x-a)\right),\ldots$$

für $a\leqq x\leqq b$, falls die Indices und die Konstanten C passend gewählt sind.

Hiermit ist Satz 7 bewiesen.

[34] Vgl. z. B. [9].

4. 4. Betreffs *der klassischen orthogonalen Polynomsysteme* seien die folgenden Tatsachen hervorgehoben:[35] 1) Die Jacobischen Polynome $J_\lambda^{(\alpha,\beta)}(x) = \frac{(-1^\lambda)}{(2\lambda)!!}(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d^\lambda}{dx^\lambda}(1-x)^{\alpha+\lambda}(1+x)^{\beta+\lambda}$ $(\alpha > -1, \beta > -1;$ $\lambda = 0, 1, 2, \ldots)$, die (verallgemeinerten) Laguerreschen Polynome $\mathfrak{L}_\lambda^{(\alpha)}(x) = \frac{1}{\lambda!}x^{-\alpha}e^x\frac{d^\lambda}{dx^\lambda}(x^{\alpha+\lambda}e^{-x})$ $(\alpha > -1; \lambda = 0, 1, 2, \ldots)$ und die Hermiteschen Polynome $H_\lambda(x) = (-1)^\lambda e^{x^2}\frac{d^\lambda}{dx^\lambda}e^{-x^2}$ $(\lambda = 0, 1, 2, \ldots)$ sind, der Reihe nach, mit den folgenden Grundintervallen und Belegungsfunktionen verknüpft: $(-1, 1)$, $w(x) = (1-x)^\alpha(1+x)^\beta$; $(0, \infty)$, $w(x) = x^\alpha e^{-x}$; bzw. $(-\infty, \infty)$, $w(x) = e^{-x^2}$. Alle drei Systeme sind inbezug auf das zugehörige Intervall und $w(x)$ vollständig. 2) Es gelten die Relationen:

$$\frac{d}{dx}J_\lambda^{(\alpha,\beta)}(x) = \frac{1}{2}(\lambda+\alpha+\beta+1)J_{\lambda-1}^{(\alpha+1,\beta+1)}(x), \quad \frac{d}{dx}\mathfrak{L}_\lambda^{(\alpha)}(x) = -\mathfrak{L}_{\lambda-1}^{(\alpha+1)}(x),$$

$$\frac{d}{dx}H_\lambda(x) = 2\lambda H_{\lambda-1}(x) \qquad (\lambda = 1, 2, \ldots).$$

Wie daraus erhellt, sind auch die ersten, zweiten usw. Ableitungen eines genannten Systems *vollständig* und *orthogonal* inbezug auf dasselbe Intervall, und zwar, der Reihe nach, zu solchen Belegungen, die (samt $w(x)$) eine geometrische Progression bilden und in den Endpunkten des Grundgebietes verschwinden, sonst aber positiv bleiben. Unser nächstes Ziel ist zu zeigen, daß diese Eigenschaften, erweitert mit gewissen natürlichen Bedingungen über Beschränktheit, schon *charakteristisch* sind; ja sogar, es genügt im wesentlichen *die ersten drei (und nicht wenigere) Ableitungssysteme* zu betrachten. Schärfer und genauer:

SATZ 8. *Es seien $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ in (a, b) viermal stetig differenzierbare Funktionen, welche darauf ein w-vollständiges, w-orthogonales System bilden. Wir nehmen an,* 1) *$\{\varphi_\lambda'(x)\}$, $\{\varphi_\lambda''(x)\}$, $\{\varphi_\lambda'''(x)\}$ seien auch vollständig und orthogonal inbezug auf (a, b) und gewisse Belegungsfunktionen $g_1(x)$, $g_2(x)$ bzw. $g_3(x)$;* 2) *$w(x)$, $g_1(x)$, $g_2(x)$, $g_3(x)$ seien in (a, b) positiv, zweimal stetig differenzierbar und bilden eine geometrische Progression;* 3) *Sei $w(x) \in L(a, b)$, $g_1(a+0) = g_1(b-0) = 0$, während $[g_1(x)w(x)^{-1}]'$, $g_1'(x)w(x)^{-1}$, gleichwie jedes $\varphi_\lambda(x)$ und ihre genannten Ableitungen in einem beliebigen endlichen Intervalle $[x_1, x_2] \in [a, b]$ beschränkt bleiben sollen,*[36] *schließlich, wenn $[a, b]$ ins Unendliche reicht,* 4) *sei $w(x) = o(|x|^{-M})$ für jedes $M > 0$, während $[g_1(x)w(x)^{-1}]', g_1'(x)w(x)^{-1}$, ferner alle $\varphi_\lambda(x)$ und ihre genannten Ableitungen für $|x| \to \infty$ höchstens wie eine Potenz von $|x|$ unendlich werden sollen.*

[35] Vgl. z. B. [29].

[36] Ist $[a, b]$ endlich, so mag der Ausdruck: „in einem beliebigen endlichen Intervalle $[x_1, x_2] \in [a, b]$" einfach mit „in $[a, b]$" ersetzt werden.

Dann stimmt $\{\varphi_\lambda(x)\}$ bis auf Reihenfolge und triviale lineare Transformationen mit einem der klassischen Polynomsysteme überein, und zwar mit demjenigen von Jacobi, Laguerre oder Hermite, je nachdem $[a, b]$ endlich ist, bzw. auf einer Seite oder beiden Seiten ins Unendliche reicht. — Zugleich müssen $w(x)$ und $g_1(x)$, abgesehen von geeigneten linearen Transformationen, mit den entsprechenden Gewichtsfunktionen identisch sein.

BEWEIS. Es sei angenommen, daß alle genannten Bedingungen befriedigt sind.

1° In einem beliebigen endlichen und abgeschlossenen Intervalle, welches $[a, b]$ angehört, bleibt der Quotient $g_1(x)w(x)^{-1}$ wegen

$$(4.13)\qquad \frac{g_1(x)}{w(x)} = \frac{g_1(x_0)}{w(x_0)} + \int_{x_0}^{x} \left[\frac{g_1(u)}{w(u)}\right]' du \leqq \frac{g_1(x_0)}{w(x_0)} + (B-A) \max_{A \leqq u \leqq B} \left|\left[\frac{g_1(u)}{w(u)}\right]'\right|$$
$$(a < A \leqq x_0 < x \leqq B < b)$$

beschränkt (vgl. 3)).

Im Falle eines unendlichen (a, b) hat man (vgl. 4))

$$(4.14)\qquad \left[\frac{g_1(x)}{w(x)}\right]' = O(|x|^\mu)$$

mit einem geeigneten $\mu \geqq 0$; dies ergibt die Abschätzung

$$\frac{g_1(x)}{w(x)} \leqq \frac{g_1(x_0)}{w(x_0)} + K\int_{x_0}^{x} |u|^\mu du + K_1(x-x_0)$$

mit $K > 0$, $K_1 > 0$ und für jedes $x \in (a, b)$, $x_0 \in (a, b)$, so daß man

$$(4.15)\qquad \frac{g_1(x)}{w(x)} = O(|x|^{\mu+1})$$

erhält.

Die Bedingung 2) liefert

$$(4.16)\qquad g_2(x) = \frac{g_1(x)^2}{w(x)}, \quad g_3(x) = \frac{g_1(x)^3}{w(x)^2};$$

daher folgt wegen (4. 13) und 3)

$$(4.17)\qquad \lim g_2(x) = 0, \quad \lim g_3(x) = 0$$

in jedem Endpunkte von (a, b), welcher im Endlichen liegt.

Ist (a, b) unendlich, so implizieren die Bedingungen 3), 4) und (4. 15), (4. 16)

$$(4.18)\qquad \left.\begin{matrix} g_1(x) \\ g_2(x) \\ g_3(x) \end{matrix}\right\} = o(|x|^{-N}) \qquad (N > 0 \text{ beliebig}).$$

Wir erwähnen ferner die aus (4. 16) fließenden Formeln

$$\left.\begin{aligned} &\frac{g_2(x)}{g_1(x)}=\frac{g_1(x)}{w(x)}, \quad \frac{g_2'(x)}{g_1(x)}=\frac{g_1'(x)}{w(x)}+\left[\frac{g_1(x)}{w(x)}\right]' && (4.19)\\ &\frac{g_3(x)}{g_2(x)}=\frac{g_1(x)}{w(x)}, \quad \frac{g_3'(x)}{g_2(x)}=\frac{g_1'(x)}{w(x)}+2\left[\frac{g_1(x)}{w(x)}\right]' && (4.20)\end{aligned}\right\} \quad (a<x<b),$$

welche samt den vorigen zeigen, daß die Funktionen $g_2(x)g_1(x)^{-1}$, $g_2'(x)g_1(x)^{-1}$ und $g_3(x)g_2(x)^{-1}$, $g_3'(x)g_2(x)^{-1}$ in jedem endlichen abgeschlossenen Intervalle in $[a, b]$ beschränkt sind, während im Falle eines unendlichen (a, b) für $|x|\to\infty$ höchstens wie eine Potenz von $|x|$ unendlich werden können.

2° Wegen unserer Voraussetzungen ist Satz 3 anwendbar. Z. B. gilt $w(x)^{-1}[g_1(x)\varphi_\lambda'(x)]'\in L_w^2(a, b)$, da

$$\frac{1}{\sqrt{w(x)}}[g_1(x)\varphi_\lambda'(x)]'=\sqrt{w(x)}\left[\frac{g_1'(x)}{w(x)}\varphi_\lambda'(x)+\frac{g_1(x)}{w(x)}\varphi_\lambda''(x)\right] \quad (a<x<b),$$

und der Ausdruck rechter Hand, laut der Bedingungen 1)—4) und 1°, zur Klasse $L^2(a, b)$ gehört; es folgt ähnlicherweise aus 3), 4) und (4. 18)

$$\lim_{x\to a+0}[g_1(x)\varphi_m(x)\varphi_n'(x)]=\lim_{x\to b-0}[g_1(x)\varphi_m(x)\varphi_n'(x)]=0 \qquad (m, n=1, 2, \ldots).$$

Also ergeben die w-Orthogonalität, w-Vollständigkeit von $\{\varphi_\lambda(x)\}$ und die g_1-Orthogonalität von $\{\varphi_\lambda'(x)\}$ die Differentialgleichungen

$$[g_1(x)\varphi_\lambda'(x)]'=\varrho_\lambda w(x)\varphi_\lambda(x) \qquad (a<x<b;\ \lambda=1, 2, \ldots) \tag{4. 21}$$

mit

$$\varrho_\lambda=-\left[\int_a^b g_1(x)\varphi_\lambda'(x)^2\,dx\right]\left[\int_a^b w(x)\varphi_\lambda(x)^2\,dx\right]^{-1} \qquad (\lambda=1, 2, \ldots). \tag{4. 22}$$

Nach einer geeigneten Abänderung der Bezeichnungen und durch Benützung von (4. 19), (4. 20) schließt man aus der g_1-Orthogonalität, g_1-Vollständigkeit von $\{\varphi_\lambda'(x)\}$ und der g_2-Orthogonalität von $\{\varphi_\lambda''(x)\}$

$$[g_2(x)\varphi_\lambda''(x)]'=\varrho_\lambda^* g_1(x)\varphi_\lambda'(x) \qquad (a<x<b;\ \lambda=1, 2, \ldots) \tag{4. 23}$$

mit

$$\varrho_\lambda^*=-\left[\int_a^b g_2(x)\varphi_\lambda''(x)^2\,dx\right]\left[\int_a^b g_1(x)\varphi_\lambda'(x)^2\,dx\right]^{-1} \qquad (\lambda=1, 2, \ldots), \tag{4. 24}$$

während die g_2-Orthogonalität, g_2-Vollständigkeit von $\{\varphi_\lambda''(x)\}$ und die g_3-Orthogonalität von $\{\varphi_\lambda'''(x)\}$ zu den Relationen

$$[g_3(x)\varphi_\lambda'''(x)]'=\varrho_\lambda^{**} g_2(x)\varphi_\lambda''(x) \qquad (a<x<b;\ \lambda=1, 2, \ldots) \tag{4. 25}$$

mit

$$\varrho_\lambda^{**}=-\left[\int_a^b g_3(x)\varphi_\lambda'''(x)^2\,dx\right]\left[\int_a^b g_2(x)\varphi_\lambda''(x)^2\,dx\right]^{-1} \qquad (\lambda=1, 2, \ldots) \tag{4. 26}$$

führen.

3° Auf Grund von (4. 21), (4. 23), (4. 25) und der oben benutzten Randbedingungen ist leicht zu erkennen, daß $\{\varphi_\lambda(x)\}$, $\{\varphi_\lambda'(x)\}$, $\{\varphi_\lambda''(x)\}$, nach

zufälliger Weglassung gewisser Elemente, *gewöhnliche SL-Systeme bilden* welche der Reihe nach mit den Eigenwerten $\{\varrho_\lambda\}$, $\{\varrho_\lambda^*\}$, $\{\varrho_\lambda^{**}\}$ und den Belegungsfunktionen $(g_1(x), w(x))$, $(g_2(x), g_1(x))$ bzw. $(g_3(x), g_2(x))$ verknüpft sind; dann hat man aber, laut Satz 5, Teil II), u. a. die Relationen

$$\left.\begin{aligned} \frac{g_1(x)}{w(x)} &= px^2+qx+r \\ \frac{g_1'(x)}{w(x)} &= Px+Q \end{aligned}\right\} \qquad (a<x<b) \tag{4.27, 4.28}$$

mit konstanten p, q, r, P und Q.

Die Koeffizienten p, q, r können, wegen der Positivität von $g_1(x)$ in (a, b), offenbar nicht gleichzeitig verschwinden, d. h.

$$p^2+q^2+r^2>0; \tag{4.29}$$

ferner muß

$$P \neq 0 \tag{4.30}$$

sein, weil man sonst durch die Randbedingungen 3)

$$\lim_{x\to b-0} g_1(x) = Q\int_a^b w(t)\,dt = 0,$$

also (vgl. (4.28)) $g_1(x)\equiv 0$ folgern könnte, in Gegensatz zu 2).

Aus (4.27) und (4.28) resultiert

$$\frac{g_1'(x)}{g_1(x)} = \frac{Px+Q}{px^2+qx+r}, \tag{4.31}$$

d. h. $y=g_1(x)$ genügt notwendigerweise der sogenannten *Pearsonschen Differentialgleichung*[37]

$$y' = \frac{Px+Q}{px^2+qx+r}\,y. \tag{4.32}$$

Setzen wir die Ausdrücke von $g_1(x)$ und $g_1'(x)$ aus (4.27) und (4.28) in (4.21), so folgt

$$(px^2+qx+r)\varphi_\lambda''(x)+(Px+Q)\varphi_\lambda'(x)-\varrho_\lambda\varphi_\lambda(x)=0 \qquad (a<x<b;\ \lambda=1,2,\ldots), \tag{4.33}$$

d. h. $y=\varphi_\lambda(x)$ soll eine Lösung *der Differentialgleichung vom Fuchsschen Typus*[38]

$$(px^2+qx+r)y''+(Px+Q)y-\varrho_\lambda y=0 \qquad (a<x<b) \tag{4.34}$$

sein.

4° Nun gebrauchen wir folgendes

LEMMA 2.[39] *Es sei $y=y(x)$ eine für $a<x<b$ stetig differenzierbare und nicht verschwindende Funktion, welche hier* (4.32) *nebst den Randbedin-*

[37] Im Falle der sogenannten *Pearsonschen Wahrscheinlichkeitsverteilungen* genügt die Verteilungsfunktion bekanntlich einer Differentialgleichung von der Form (4.32). Vgl. [3].

[38] Vgl. z. B. [19], S. 131.

[39] Vgl. z. B. [18], Ch. VI, VII.

gungen $y(a+0)=y(b-0)=0$ *genügt; im Falle eines unendlichen* (a,b) *bestehe überdies* $y(x)=o(|x|^{-N})$ *für jedes* $N>0$. *Dann hat man (mit geeignetem konstantem* $C\neq 0$)

1) $p\neq 0$, $px^2+qx+r\equiv p(x-a)(x-b)$ *und für* $a<x<b$

$$y(x)=C(b-x)^{\varkappa_1}(x-a)^{\varkappa_2} \tag{4.35}$$

mit

$$\varkappa_{1,2}=\frac{P}{2p}\pm\frac{1}{p(b-a)}\left(Q-\frac{Pq}{2p}\right)>0, \tag{4.36}$$

falls $[a,b]$ *endlich ist;*

2) $p=0$, $Pq<0$, $a=-\dfrac{r}{q}$ *und für* $a<x<\infty$

$$y(x)=C(x-a)^{\tau}e^{\frac{P}{q}x} \tag{4.37}$$

mit

$$\tau=\frac{1}{q}\left(Q-\frac{Pr}{q}\right)>0, \tag{4.38}$$

wenn a *endlich ist und* $b=\infty$;

3) $p=0$, $Pq>0$, $b=-\dfrac{r}{q}$ *und für* $-\infty<x<b$

$$y(x)=C(b-x)^{\tau}e^{\frac{P}{q}x} \tag{4.39}$$

mit (4.38), *wenn* $a=-\infty$ *und* b *endlich ist;*

4) $p=q=0$, $Pr<0$ *und für* $-\infty<x<\infty$

$$y(x)=C\exp\left(\frac{P}{2r}x^2+\frac{Q}{r}x\right), \tag{4.40}$$

falls $a=-\infty$, $b=\infty$.

Denn (4.32) läßt sich sofort durch logarithmische Integration lösen; mit Rücksicht auf (4.29) und (4.30) hat man die folgenden Fälle zu unterscheiden (wobei x ein Stetigkeitsintervall des rationalen Koeffizienten rechter Hand in (4.32) durchläuft, die Konstante C aber beliebig gewählt werden mag):

1) ist $p\neq 0$, $q^2>4pr$, so erhält man

$$y=C|x-\xi_1|^{\varkappa_1}|x-\xi_2|^{\varkappa_2} \tag{4.41}$$

mit

$$\xi_{1,2}=\frac{1}{2p}(-q\pm\sqrt{q^2-4pr}),\quad \varkappa_{1,2}=\frac{P}{2p}\pm\left(Q-\frac{Pq}{2p}\right)\frac{1}{p(\xi_1-\xi_2)};$$

2) ist $p \neq 0$, $q^2 = 4pr$, so ergibt sich

$$y = C\left|x + \frac{q}{2p}\right|^{\frac{P}{q}} \exp\left[\left(\frac{Pq}{p} - 2Q\right)\frac{1}{2px+q}\right]; \tag{4.42}$$

3) ist $p \neq 0$, $q^2 < 4pr$, so resultiert

$$y = C|px^2+qx+r|^{\frac{P}{2p}} \exp\left[\left(2Q - \frac{Pq}{p}\right)\frac{1}{\sqrt{4pr-q^2}} \operatorname{arc\,tg}\frac{2px+q}{\sqrt{4pr-q^2}}\right]; \tag{4.43}$$

4) im Falle $p = 0$, $q \neq 0$ folgt

$$y = Ce^{\frac{P}{q}x}\left|x + \frac{r}{q}\right|^{\tau} \tag{4.44}$$

mit $\tau = \frac{1}{q}\left(Q - \frac{Pr}{q}\right)$;

5) im Falle $p = q = 0$ ist

$$y = C\exp\left(\frac{P}{2r}x^2 + \frac{Q}{r}x\right). \tag{4.45}$$

Von (4.42)—(4.45) liest man leicht ab, daß die Annahmen unseres Lemmas und die Fälle 2), 3) unverträglich sind, und weiter, daß diese die Gültigkeit einer von den fraglichen Behauptungen 1)—4) nach sich ziehen, je nachdem $[a, b]$ endlich, einerseits bzw. beiderseits unendlich ist.[40]

5° Nun kehren wir auf (4.31) und (4.33) zurück.

Ist $[a, b]$ endlich, so hat man nach (4.31) und dem eben bewiesenen Lemma die Behauptungen 1) mit $y(x) = g_1(x)$ und $C > 0$; wir setzen

$$\varkappa_1 = \alpha + 1, \quad \varkappa_2 = \beta + 1, \tag{4.46}$$

so daß (vgl. (4.36)) $\alpha > -1$, $\beta > -1$. Also ist

$$g_1(x) = C(b-x)^{\alpha+1}(x-a)^{\beta+1} \qquad (a < x < b;\ \alpha > -1,\ \beta > -1,\ C > 0) \tag{4.47}$$

und wegen (4.27)

$$w(x) = -\frac{C}{p}(b-x)^{\alpha}(x-a)^{\beta} \tag{4.48}$$

$$(a < x < b;\ \alpha > -1,\ \beta > -1,\ C > 0,\ p < 0).$$

Andererseits, falls man den Teil 1) des Lemmas und die lineare Transformation

$$x = a + \frac{b-a}{2}(t+1), \quad t = -1 + \frac{2}{b-a}(x-a)$$

[40] Die allgemeinen Lösungen (4.41)—(4.45) liefern unmittelbar eine partikuläre Integralkurve von (4.32), welche durch einen gegebenen nicht-singulären Punkt der xy-Ebene geht.

auf (4.33) anwendet, so folgt (vgl. (4.46))

$$(4.49)\qquad \begin{aligned}&(1-t^2)\frac{d^2}{dt^2}\varphi_\lambda\left(a+\frac{b-a}{2}(t+1)\right)-\\&-[(\alpha+\beta+2)t+(\alpha-\beta)]\frac{d}{dt}\varphi_\lambda\left(a+\frac{b-a}{2}(t+1)\right)+\\&+\frac{\varrho_\lambda}{p}\varphi_\lambda\left(a+\frac{b-a}{2}(t+1)\right)=0 \qquad (-1<t<1;\ \alpha>-1,\ \beta>-1);\end{aligned}$$

d. h. $y=\varphi_\lambda\left(a+\frac{b-a}{2}(t+1)\right)$ genügt der Differentialgleichung

$$(4.50)\qquad (1-t^2)\frac{d^2y}{dt^2}-[(\alpha+\beta+2)t+(\alpha-\beta)]\frac{dy}{dt}+Ky=0$$

oder (in Sturm—Liouvillescher Form geschrieben)

$$(4.51)\qquad \frac{d}{dt}\left[(1-t)^{\alpha+1}(1+t)^{\beta+1}\frac{dy}{dt}\right]=-K(1-t)^\alpha(1+t)^\beta$$

mit $K=\frac{\varrho_\lambda}{p}$ für $-1<t<1$ und $\lambda=1,2,\ldots$.

Aber das Randwertproblem: „gesucht die für $-1<t<1$ beschränkten Lösungen von (4.50) oder (4.51)“ hat bekanntlich die (einfachen) Eigenwerte $K=-(\lambda-1)(\lambda+\alpha+\beta)$ $(\lambda=1,2,\ldots)$ und die Eigenfunktionen const$\cdot J_{\lambda-1}^{(\alpha,\beta)}(t)$ $(\lambda=1,2,\ldots;$[41] da unsere Voraussetzungen 3) die Beschränktheit alle $\varphi_\lambda\left(a+\frac{b-a}{2}(t+1)\right)$ in $[-1,1]$ implizieren, ergibt sich bei einer geeigneten Wahl der Indices

$$(4.52)\qquad \varrho_\lambda=pK=p(\lambda-1)(\lambda+\alpha+\beta) \qquad (\lambda=1,2,\ldots)$$

und (wegen der angenommenen Vollständigkeit)

$$(4.53)\qquad \varphi_\lambda\left(a+\frac{b-a}{2}(t+1)\right)=\text{const}\cdot J_{\lambda-1}^{(\alpha,\beta)}(t) \qquad (-1<t<1;\ \lambda=1,2,\ldots),$$

d. h.

$$(4.54)\qquad \varphi_\lambda(x)=\text{const}\cdot J_{\lambda-1}^{(\alpha,\beta)}\left(\frac{2}{p-a}(x-a)-1\right) \qquad (a<x<b;\ \lambda=1,2,\ldots)$$

mit $\alpha>-1$, $\beta>-1$.

6° Ist a endlich, $b=\infty$, so ergibt die Anwendung des Teils 2) unseres Lemmas mit $y(x)=g_1(x)$, daß $p=0$, $Pq<0$, $a=-\frac{r}{q}$ und für $a<x<\infty$

$$(4.55)\qquad g_1(x)=C(x-a)^{\alpha+1}e^{\frac{P}{q}x}$$

[41] Vgl. z. B. [19], S. 465 oder [4], S. 258—260; unsere Bezeichnungen sind diejenigen von [29]. — Die betreffende Behauptung kann am kürzesten aus der Vollständigkeit des Jacobischen Polynomsystems inbezug auf $[-1,1]$ und die entsprechenden Gewichtsfunktionen gefolgert werden. ([29], S. 39.)

mit $C>0$, $\alpha+1=\tau=\frac{1}{q}\left(Q-\frac{Pr}{q}\right)$ und (vgl. (4.38))

$$\alpha>-1. \tag{4.56}$$

Kombiniert man (4.55) mit (4.27), so folgt noch

$$w(x)=\frac{C}{q}(x-a)^{\alpha}e^{\frac{P}{q}x} \tag{4.57}$$

mit $q>0$. (4.33) kann nun in der Form

$$q(x-a)\varphi_\lambda''(x)+(Px+Q)\varphi_\lambda'(x)-\varrho_\lambda\varphi_\lambda(x)=0 \qquad (a<x<\infty) \tag{4.58}$$

geschrieben werden, welche sich durch

$$x=a-\frac{q}{P}t, \qquad t=\frac{P}{q}(a-x) \tag{4.59}$$

in die Relation

$$\begin{aligned} t\frac{d^2}{dt^2}\varphi_\lambda\left(a-\frac{q}{P}t\right)+[(\alpha+1)-t]\frac{d}{dt}\varphi_\lambda\left(a-\frac{q}{P}t\right)+ \\ +\frac{\varrho_\lambda}{P}\varphi_\lambda\left(a-\frac{q}{P}t\right)=0 \qquad (0<t<\infty;\ \alpha>-1) \end{aligned} \tag{4.60}$$

transformiert; d. h. $y=\varphi_\lambda\left(a-\frac{q}{P}t\right)$ genügt der Differentialgleichung

$$t\frac{d^2y}{dt^2}+[(\alpha+1)-t]\frac{dy}{dt}+Ky=0 \tag{4.61}$$

oder

$$\frac{d}{dt}\left(t^{\alpha+1}e^{-t}\frac{dy}{dt}\right)=-Kt^{\alpha}e^{-t}y \tag{4.62}$$

mit $K=\frac{\varrho_\lambda}{P}$ für $0<t<\infty$ und $\lambda=1,2,\ldots$.

Das durch (4.61) bzw. (4.62) und die Bedingungen: „$y=y(t)$ sei beschränkt für $t\to+0$ und $O(t^{\nu})$ mit geeignetem $\nu\geqq 0$ für $t\to+\infty$“ gegebene Randwertproblem hat aber die (einfachen) Eigenwerte $K=\lambda-1$ $(\lambda=1,2,\ldots)$ und die Eigenfunktionen $\text{const}\cdot\mathfrak{L}_{\lambda-1}^{(\alpha)}(t)$ $(\lambda=1,2,\ldots)$;[42] da die Voraussetzungen 4) $\varphi_\lambda\left(a-\frac{q}{P}t\right)=O(t^{\nu})$ $(0\leqq t<\infty)$ implizieren, gelangt man nach einer geeigneten Umordnung zu

$$\varrho_\lambda=PK=P(\lambda-1) \qquad (\lambda=1,2,\ldots), \tag{4.63}$$

[42] Vgl. z. B. [19], S. 429 und [4], S. 261. — Das ist eine einfache Konsequenz der Vollständigkeit von $\{\mathfrak{L}_\lambda^{(\alpha)}(t)\}$ inbezug auf $(0,\infty)$ und $w(t)=t^{\alpha}e^{-t}$ $(\alpha>-1)$ ([29], S. 104).

und (wegen der Vollständigkeit)

$$\varphi_\lambda\left(a-\frac{q}{P}t\right)=\text{const}\cdot\mathfrak{L}_{\lambda-1}^{(\alpha)}(t) \qquad (0<t<\infty;\ \lambda=1,2,\ldots), \tag{4.64}$$

d. h.

$$\varphi_\lambda(x)=\text{const}\cdot\mathfrak{L}_{\lambda-1}^{(\alpha)}\left(\frac{P}{q}(a-x)\right) \qquad (a<x<\infty;\ \lambda=1,2,\ldots). \tag{4.65}$$

7° Falls $a=-\infty$ und b endlich ist, bedarf man nur einer einfachen Veränderung in den Bezeichnungen von 6°. Wir erhalten

$$\left.\begin{aligned} g_1(x)&=C(b-x)^{\alpha+1}e^{\frac{P}{q}x} \qquad (4.66)\\ w(x)&=-\frac{C}{q}(b-x)^{\alpha}e^{\frac{P}{q}x} \qquad (4.67)\end{aligned}\right\} \quad (C>0,\ q<0,\ P<0,\ \alpha>-1),$$

$$\varphi_\lambda(x)=\text{const}\cdot\mathfrak{L}_{\lambda-1}^{(\alpha)}\left(\frac{P}{q}(b-x)\right) \qquad (-\infty<x<b;\ \alpha>-1,\ \lambda=1,2,\ldots). \tag{4.68}$$

8° Schließlich, im Falle $a=-\infty$, $b=\infty$ kann Teil 4) des Lemmas benutzt werden. Bezüglich (4.31) bekommt man $p=q=0$, $Pr<0$ und für $-\infty<x<\infty$

$$g_1(x)=C\exp\left(\frac{P}{2r}x^2+\frac{Q}{r}x\right); \tag{4.69}$$

daher und aus (4.27)

$$w(x)=\frac{C}{r}\exp\left(\frac{P}{2r}x^2+\frac{Q}{r}x\right) \tag{4.70}$$

mit $r>0$.

Andererseits reduziert (4.27) sich nun auf

$$r\varphi_\lambda''(x)+(Px+Q)\varphi_\lambda'(x)-\varrho_\lambda\varphi_\lambda(x)=0 \qquad (-\infty<x<\infty), \tag{4.71}$$

was durch

$$x=t\sqrt{-\frac{2r}{P}}-\frac{Q}{P}, \quad t=x\sqrt{-\frac{P}{2r}}+\frac{Q}{\sqrt{-2Pr}} \tag{4.72}$$

in die Relation

$$\begin{aligned}\frac{d^2}{dt^2}\varphi_\lambda\left(t\sqrt{-\frac{2r}{P}}-\frac{Q}{P}\right)-2t\frac{d}{dt}\varphi_\lambda\left(t\sqrt{-\frac{2r}{P}}-\frac{Q}{P}\right)+\\ +\frac{2\varrho_\lambda}{P}\varphi_\lambda\left(t\sqrt{-\frac{2r}{P}}-\frac{Q}{P}\right)=0 \qquad (-\infty<t<\infty)\end{aligned} \tag{4.73}$$

übergeht.

Also sind $y=\varphi_\lambda\left(t\sqrt{-\frac{2r}{P}-\frac{Q}{P}}\right)$ $(\lambda=1,2,\ldots)$ Lösungen von

$$\frac{d^2y}{dt^2}-2t\frac{dy}{dt}+Ky=0, \tag{4.74}$$

$$\frac{d}{dt}\left(e^{-t^2}\frac{dy}{dt}\right)=-Ke^{-t^2}y \tag{4.75}$$

mit $K=\frac{2\varrho_\lambda}{P}$ für $-\infty<t<\infty$.

Das Problem: „gesucht sind Funktionen $y=y(t)$, welche (4.74) bzw. (4.75) und den Randbedingungen $y(t)=O(|t|^\mu)$ (mit einem geeigneten $\mu\geqq 0$) für $t\to\pm\infty$ genügen" — hat, wie bekannt, die (einfachen) Eigenwerte $K=2(\lambda-1)$ $(\lambda=1,2,\ldots)$ und die Eigenfunktionen $\mathrm{const}\cdot H_{\lambda-1}(t)$ $(\lambda=1,2,\ldots)$;[43] wegen unserer Voraussetzungen 4) folgt also aus (4.75)

$$\varrho_\lambda=\frac{P}{2}K=P(\lambda-1) \qquad (\lambda=1,2,\ldots), \tag{4.76}$$

ferner

$$\varphi_\lambda\left(t\sqrt{-\frac{2r}{P}-\frac{Q}{P}}\right)=\mathrm{const}\cdot H_\lambda(t) \qquad (-\infty<t<\infty;\ \lambda=1,2,\ldots), \tag{4.77}$$

d. h.

$$\varphi_\lambda(x)=\mathrm{const}\cdot H_\lambda\left(x\sqrt{-\frac{P}{2r}}+\frac{Q}{\sqrt{-2Pr}}\right) \qquad (-\infty<x<\infty;\ \lambda=1,2,\ldots). \tag{4.78}$$

Somit ist der Beweis vollendet.

§ 5. Verallgemeinerung einiger Sätze über die Differentiation gewöhnlicher Fourierreihen

5. 1. Wie bekannt, darf man eine trigonometrische Fourierreihe stets gliedweise integrieren, dagegen nur ausnahmsweise formal differenzieren: die Größenordnung der Fourierkoeffizienten wird im ersten Falle verbessert, im zweiten aber verschlimmert. Im Anschluß einer Fragestellung von ABEL hat jedoch L. FEJÉR bewiesen, daß die gliedweise differenzierte Fourierreihe einer Funktion $f(x)$ $(0\leqq x\leqq 2\pi)$ im Inneren des Grundintervalles C_1-summierbar ist und $f'(x)$ zur C_1-Summe hat, falls diese Ableitung z. B. in $(0,2\pi)$ existiert und stetig ist.[44]

[43] Vgl. z. B. [19], S. 415 und [4], S. 261; für die Vollständigkeit von $\{H_\lambda(t)\}$ s. auch [29], S. 104.

[44] [6], [7]. Die stetige Differenzierbarkeit von $f(x)$ impliziert natürlich die Konvergenz der Fourierreihe selbst. — Spätere Arbeiten von YOUNG, DE LA VALLÉE POUSSIN, GRONWALL, ZYGMUND ([32]) haben sich mit gewissen Verschärfungen und mit dem Falle höherer Ableitungen beschäftigt.

Die Ergebnisse der vorigen Abschnitte ergeben nun einen allgemeinen Satz über Konvergenz und Differentiation solcher Orthogonalreihen, welche eine orthogonale Ableitung haben.

SATZ 9. *Es seien $w(x)$ und $g(x)$ in einem endlichen $[a, b]$ positive, zweimal stetig differenzierbare Gewichtsfunktionen. $\{\varphi_\lambda(x)\}$ sei ein w-vollständiges, w-orthonormales System ebenda mit stetigen dritten Ableitungen versehener Funktionen mit $\varphi_\lambda(a) = \varphi_\lambda(b) = 0$ $(\lambda = 1, 2, \ldots)$ und $\{\varphi'_\lambda(x)\}$ sei g-orthogonal in (a, b); bezüglich der Reihenfolge nehmen wir an, daß die g-Norm von $\varphi'_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ zunimmt. — Es bezeichnen γ_λ $(\lambda = 1, 2, \ldots)$ die Fourierkoeffizienten einer Funktion $f(x) \in \mathrm{Lip}\, 1$ $(a \leqq x \leqq b)$[45] inbezug auf $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$.*

Dann ist die Fourierreihe $\sum_{\lambda=1}^{\infty} \gamma_\lambda \varphi'_\lambda(x)$ in (a, b) konvergent und sie stellt dort in jedem inneren Teilintervalle gleichmäßig $f(x)$ dar, während $\sum_{\lambda=1}^{\infty} \gamma_\lambda \varphi'_\lambda(x)$ für fast jedes $x \in (a, b)$ C_1-summierbar ist und $f'(x)$ zur C_1-Summe hat. Letzteres gilt gewiß in solchen Punkten von (a, b), wo $f'(x)$ existiert und stetig ist, und die Summierbarkeit ist in jedem abgeschlossenen inneren Stetigkeitsintervall der Ableitung $f'(x)$ eine gleichmäßige.

Der BEWEIS stützt sich u. a. auf wohlbekannte Tatsachen über *Sturm—Liouvillesche Orthogonalsysteme*, insbesondere auf die asymptotische Formel von HOBSON und den *Haarschen Äquikonvergenzsatz;* wir brauchen noch eine Eigenschaft der auftretenden Kernfunktionen, nämlich die Tatsache, daß sie in den Endpunkten zum Werte Null C_1-limitierbar sind.

1° Die vorausgesetzten Orthogonalitätseigenschaften von $\{\varphi_\lambda(x)\}$, $\{\varphi'_\lambda(x)\}$ haben nach dem Satze 3 zur Folge, daß die Relationen

$$[g(x)\varphi'_\lambda(x)]' = \varrho_\lambda w(x)\varphi_\lambda(x) \qquad (a \leqq x \leqq b;\ \lambda = 1, 2, \ldots) \tag{5.1}$$

mit

$$\left.\begin{aligned} \varrho_\lambda &= -\int_a^b g(x)\varphi'_\lambda(x)^2 dx \qquad (5.2)\\ \varphi_\lambda(a) &= \varphi_\lambda(b) = 0 \qquad (5.3) \end{aligned}\right\} \quad (\lambda = 1, 2, \ldots)$$

gelten; wegen der w-Vollständigkeit von $\{\varphi_\lambda(x)\}$ sind also $y = \mathrm{const}\cdot\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ *alle* Eigenfunktionen des Sturm—Liouvilleschen Randwertproblems

$$[g(x)y']' = \varrho w(x)y \qquad (a \leqq x \leqq b);\quad y(a) = y(b) = 0 \tag{5.4}$$

und es ist $(0 >)\ \varrho_1 > \varrho_2 > \varrho_3 > \cdots$ das zugehörige Spektrum.

2° Nun liefert Lemma 1

$$[g(x)^2 w(x)^{-1}\varphi''_\lambda(x)]' = \left\{\varrho_\lambda - \left[\frac{g'(x)}{w(x)}\right]'\right\} g(x)\varphi'_\lambda(x) \tag{5.5}$$

$$(a \leqq x \leqq b;\ \lambda = 1, 2, \ldots)$$

[45] Diese Annahme mag durch die Totalstetigkeit von $f(x)$ ersetzt werden.

mit (vgl. (5. 1), (5. 3))

$$\left.\begin{aligned}(5.6)\qquad & [g(x)\varphi_\lambda'(x)]'_{x=a} = g'(a)\varphi_\lambda'(a) + g(a)\varphi_\lambda''(a) = 0\\ (5.7)\qquad & [g(x)\varphi_\lambda'(x)]'_{x=b} = g'(b)\varphi_\lambda'(b) + g(b)\varphi_\lambda''(b) = 0\end{aligned}\right\} \qquad (\lambda = 1, 2, \ldots);$$

man sieht zugleich, daß bei $\varrho \neq 0$ für jede Lösung $y = y(x)$ der Differentialgleichung

$$(5.8)\qquad [g(x)^2 w(x)^{-1} y']' = \left\{\varrho - \left[\frac{g'(x)}{w(x)}\right]'\right\} g(x) y \qquad (a \leqq x \leqq b),$$

welche den Randbedingungen

$$(5.9)\qquad [g(x)y]'_{x=a} = [g(x)y]'_{x=b} = 0$$

genügt, die primitive Funktion $Y = \int_a^x y(t)dt$ (5. 4) befriedigt. Im Falle $\varrho = 0$ resultiert aber leicht $[g(x)y(x)]' \equiv 0$ $(a \leqq x \leqq b)$.

Daraus erhellt durch 1°, daß die Folge $g(x)^{-1}, \varphi_1'(x), \varphi_2'(x), \ldots$ als ein komplettes (g-vollständiges) Eigenfunktionensystem der *SL*-Aufgabe (5. 8), (5. 9) aufgefaßt werden kann.

3° Wir wenden auf (5. 5), (5. 6), (5. 7) die sogenannte *Liouvillesche Normaltransformation* an.[46] Man setze

$$(5.10)\qquad u = \int_a^x \left(\frac{w(t)}{g(t)}\right)^{\frac{1}{2}} dt \quad (a \leqq x \leqq b), \qquad H = \int_a^b \left(\frac{w(t)}{g(t)}\right)^{\frac{1}{2}} dt$$

und für die inverse Funktion

$$(5.11)\qquad x = V(u) \qquad (0 \leqq u \leqq H).$$

Schreibt man noch

$$\left.\begin{aligned}(5.12)\qquad & \varphi_\lambda'(x) = \sqrt{-\varrho_\lambda}\, g(x)^{-\frac{3}{4}} w(x)^{\frac{1}{4}} \varphi_\lambda^*(u)\\ (5.13)\qquad & \varphi_\lambda^*(u) = (-\varrho_\lambda)^{-\frac{1}{2}} \left[\varphi_\lambda'(x) g(x)^{\frac{3}{4}} w(x)^{-\frac{1}{4}}\right]_{x=V(u)}\end{aligned}\right\} \qquad (\lambda = 1, 2, \ldots),$$

ferner

$$(5.14)\qquad \varphi_0^*(u) = C_0 [g(x)w(x)]^{-\frac{1}{4}}_{x=V(u)}, \quad C_0 = \left(\int_a^b g(x)^{-1} dx\right)^{-\frac{1}{2}},$$

dann sind $y = \varphi_\lambda^*(u)$ $(\lambda = 0, 1, 2, \ldots)$ laut 2°, von konstanten Faktoren abgesehen, sämtliche Eigenfunktionen einer Differentialgleichung der Form

$$(5.15)\qquad \frac{d^2y}{du^2} + (\Xi(u) - \varrho)y = 0 \qquad (0 \leqq u \leqq H)$$

[46] Vgl. z. B. [19], S. 261—262.

mit

$$\Xi(u)=\left[\left(\frac{g'(x)}{w(x)}\right)'-\frac{1}{2}\left(\frac{g(x)}{w(x)}\right)'\frac{\Psi'(x)}{\Psi(x)}-\frac{g(x)}{w(x)}\frac{\Psi''(x)}{\Psi(x)}\right]_{x=V(u)}, \tag{5.16}$$

$$\Psi(x)=g(x)^{\frac{3}{4}}\,w(x)^{-\frac{1}{4}},$$

welche den (5.9) entsprechenden SL-Randbedingungen dritter Art genüge leisten. (Die Eigenwerte sind offenbar wieder $\varrho=0,\varrho_1,\varrho_2,\ldots.$) Da wir ohne Mühe

$$\int_0^H \varphi_\lambda^*(u)^2\,du=1 \qquad (\lambda=0,1,2,\ldots) \tag{5.17}$$

finden, so bilden $\varphi_\lambda^*(u)$ $(\lambda=0,1,2,\ldots)$ ein vollständiges Orthonormalsystem in $(0,H)$.

4° Nach HOBSON gilt nun die Formel[47]

$$\varphi_\lambda^*(u)=\sqrt{\frac{2}{H}}\cos\frac{\pi\lambda}{H}u+\frac{\omega^*(u)}{\lambda}\sin\frac{\pi\lambda}{H}u+O\left(\frac{1}{\lambda^2}\right) \qquad (\lambda=1,2,\ldots), \tag{5.18}$$

wobei $\omega^*(u)$ eine von λ unabhängige, in $0\leqq u\leqq H$ stetige Funktion bedeutet und das letzte Glied ebenda gleichmäßig beschränkt ist.[48]

Daraus folgt vermöge des genannten Äquikonvergenztheorems von HAAR,[4] daß für eine *beliebige* $F(t)\in L(0,H)$ die Limesrelation

$$\lim_{n\to\infty}\left\{\sum_{\lambda=0}^{n}\varphi_\lambda^*(u)\int_0^H F(t)\varphi_\lambda^*(t)\,dt-\right.$$

$$\left.-\left(\frac{1}{H}\int_0^H F(t)\,dt+\frac{2}{H}\sum_{\lambda=1}^{n}\cos\frac{\pi\lambda}{H}u\int_0^H F(t)\cos\frac{\pi\lambda}{H}t\,dt\right)\right\}= \tag{5.19}$$

$$=\lim_{n\to\infty}\int_0^H F(t)\left\{\sum_{\lambda=0}^{n}\varphi_\lambda^*(t)\varphi_\lambda^*(u)-\frac{1}{H}\left(1+2\sum_{\lambda=1}^{n}\cos\frac{\pi\lambda}{H}t\cos\frac{\pi\lambda}{H}u\right)\right\}dt=0$$

besteht, und sogar gleichmäßig für $0\leqq u\leqq H$.

5° Es sei die Reihe

$$\sum_{\lambda=1}^{\infty}\left\{\varphi_\lambda^*(t)\varphi_\lambda^*(u)-\frac{2}{H}\cos\frac{\pi\lambda}{H}t\cos\frac{\pi\lambda}{H}u-\right.$$

$$\left.-\sqrt{\frac{2}{H}}\frac{1}{\lambda}\left[\omega^*(t)\sin\frac{\pi\lambda}{H}t\cos\frac{\pi\lambda}{H}u-\omega^*(u)\cos\frac{\pi\lambda}{H}t\sin\frac{\pi\lambda}{H}u\right]\right\} \tag{5.20}$$

$$(0\leqq t\leqq H;\ 0\leqq u\leqq H)$$

[47] Der genaue explizite Ausdruck von $\omega^*(u)$ ist leicht angebbar, aber für unsere Zwecke unnötig.

[48] Vgl. [17]; [19], S. 262.

[49] [13] M. I., S. 355.

betrachtet. Da ihr allgemeines Glied $O\left(\frac{1}{\lambda^2}\right)$ ist (vgl. (5. 18)), konvergiert sie nach WEIERSTRASS für die betreffenden t und u gleichmäßig und stellt eine stetige Funktion $M(t, u)$ dar. Somit gilt durch (5. 18)

$$(5.21)\qquad \sum_{\lambda=1}^{\infty}\left(\varphi_\lambda^*(0)\varphi_\lambda^*(t)-\frac{2}{H}\cos\frac{\pi\lambda}{H}t\right)=M(t,0)+\frac{\pi}{\sqrt{2H}}\,\omega^*(t)\left(1-\frac{t}{H}\right)$$
$$(0<t\leqq H),$$

$$(5.22)\qquad \sum_{\lambda=0}^{\infty}\left(\varphi_\lambda(H)\varphi_\lambda^*(t)-\frac{2}{H}(-1)^\lambda\cos\frac{\pi\lambda}{H}t\right)=M(t,H)+\frac{\pi}{H\sqrt{2H}}\,t\omega^*(t)$$
$$(0\leqq t<H).$$

Die Reihensummen linker Hand sind also in $(0, H)$ vorhanden, sogar stetig und in $[0, H]$ beschränkt; die Partialsummen beider Reihen bleiben in $[0, H]$, wie man auf Grund von $\left|\sum_{\lambda=1}^{n}\frac{\sin\lambda\xi}{\lambda}\right|\leqq\int_0^\pi\frac{\sin t}{t}\,dt<2\ (-\infty<\xi<\infty,\ n=1,2,\ldots)$ leicht nachweist,[50] absolut unter einer von n unabhängigen oberen Schranke.

Alles dies ermöglicht die Anwendung des Lebesgueschen Satzes über gliedweise Integration[51] im Falle von (5. 19); man erhält

$$\int_0^H F(t)\left\{\varphi_\lambda^*(0)\varphi_\lambda^*(t)-\frac{1}{H}+\sum_{\lambda=1}^{\infty}\left(\varphi_\lambda^*(0)\varphi_\lambda^*(t)-\frac{2}{H}\cos\frac{\pi\lambda}{H}t\right)\right\}dt=0,$$

$$\int_0^H F(t)\left\{\varphi_\lambda^*(H)\varphi_\lambda^*(t)-\frac{1}{H}+\sum_{\lambda=1}^{\infty}\left(\varphi_\lambda^*(H)\varphi_\lambda^*(t)-\frac{2}{H}(-1)^\lambda\cos\frac{\pi\lambda}{H}t\right)\right\}dt=0$$

und daraus, daß die Faktoren in Klammern neben $F(t)$ für $0<t<H$ identisch verschwinden müssen. Geht man endlich zu den arithmetischen Mitteln der Partialsummen über, so folgen wegen

$$(5.23)\qquad \lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\left(1+2\sum_{\lambda=1}^{k}\cos\lambda\xi\right)=\lim_{n\to\infty}\frac{1}{n+1}\left(\frac{\sin\frac{1}{2}(n+1)\xi}{\sin\frac{1}{2}\xi}\right)^2=0$$
$$(0<\xi<2\pi)$$

die Relationen

$$(5.24)\qquad \lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\sum_{\lambda=0}^{k}\varphi_\lambda^*(0)\varphi_\lambda^*(t)=\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\sum_{\lambda=0}^{k}\varphi_\lambda^*(H)\varphi_\lambda^*(t)=0$$
$$(0<t<H).$$

Sie gelten übrigens offenbar gleichmäßig in jedem Intervalle $[\varepsilon, H-\varepsilon]$ mit $\varepsilon>0$.

[50] Die oft benutzte gleichmäßige Beschränktheit der betreffenden Sinuspolynome kam bekanntlich zuerst bei FEJÉR zu ausgebreiteter Anwendung. (Vgl. [8]; ROGOSINSKI [27].)

[51] Vgl. z. B. [26], S. 37 ff.

6° Da, der Voraussetzung nach, $f(x) \in \mathrm{Lip}\,1$ $(a \leqq x \leqq b)$, ist diese Funktion in $[a, b]$ stetig und *beschränkter Schwankung* und sogar, laut des wohlbekannten Lebesgueschen Satzes über monotone Funktionen,[52] ebenda *fast überall differenzierbar.*

Eine Teilintegration ergibt also (vgl. (5. 1), (5. 12), (5. 13))

$$\gamma_\lambda \sqrt{-\varrho_\lambda} = \sqrt{-\varrho_\lambda} \int_a^b f(t) w(t) \varphi_\lambda(t)\, dt = (-\varrho_\lambda)^{\frac{1}{2}} \Big\{ f(t) g(t) \varrho_\lambda'(t) \Big|_{t=a}^{t=b} +$$

$$+ \int_a^b f'(t) g(t) \varphi_\lambda'(t)\, dt \Big\} = f(a) (g(a) w(a))^{\frac{1}{4}} \varphi_\lambda^*(0) - \tag{5.25}$$

$$- f(b) (g(b) w(b))^{\frac{1}{4}} \varphi_\lambda^*(H) + \int_0^H \Big[f'(x) g(x)^{\frac{3}{4}} w(x)^{-\frac{1}{4}} \Big]_{x = V(t)} \varphi_\lambda^*(t)\, dt$$

$$(\lambda = 1, 2, \ldots),$$

wobei die Erklärung der Integrande passend ausgedehnt zu denken ist.[53] Es sei zunächst bemerkt, daß aus (5. 14), (5. 10)

$$f(a) (g(a) w(a))^{\frac{1}{4}} \varphi_0^*(0) - f(b) (g(b) w(b))^{\frac{1}{4}} \varphi_0^*(H) + \tag{5.26}$$

$$+ \int_0^H \Big[f'(x) g(x)^{\frac{3}{4}} w(x)^{-\frac{1}{4}} \Big]_{x = V(t)} \varphi_0^*(t)\, dt = 0$$

folgt.

Nun hat man formal

$$\sum_{\lambda=1}^{\infty} \gamma_\lambda \varphi_\lambda'(x) = g(x)^{-\frac{3}{4}} w(x)^{\frac{1}{4}} \Big\{ f(a) (g(a) w(a))^{\frac{1}{4}} \sum_{\lambda=0}^{\infty} \varphi_\lambda^*(0) \varphi_\lambda^*(u) -$$

$$- f(b) (g(b) w(b))^{\frac{1}{4}} \sum_{\lambda=0}^{\infty} \varphi_\lambda^*(H) \varphi_\lambda^*(u) \Big\} + \tag{5.27}$$

$$+ g(x)^{-\frac{3}{4}} w(x)^{\frac{1}{4}} \sum_{\lambda=0}^{\infty} \varphi_\lambda^*(u) \int_0^H \Big[f'(x) g(x)^{\frac{3}{4}} w(x)^{-\frac{1}{4}} \Big]_{x = V(t)} \varphi_\lambda^*(t)\, dt.$$

Das erste Glied rechts ist wegen (5. 24) C_1-summierbar zur C_1-Summe 0 für jedes $x \in (a, b)$, d. h. $u \in (0, H)$. Die Reihe im zweiten Gliede ist die Fouriersche Entwicklung der Funktion $[f'(x) g(x)^{\frac{3}{4}} w(x)^{-\frac{1}{4}}]_{x = V(u)}$ inbezug auf das Orthonormalsystem $\{\varphi_\lambda^*(u)\}$; da die trigonometrische Fourierreihe dieser Funktion nach der *Lebesgueschen Verallgemeinerung des Fejérschen Satzes*[54]

[52] Vgl. z. B. [26], S. 5—11.

[53] Etwa in der Weise, daß man den Punkten derjenigen Nullmenge, wo $f'(x)$ nicht existiert, den Funktionswert 0 zuordnet.

[54] Vgl. z. B. [16], S. 62.

fast überall, insbesondere an allen inneren Stetigkeitsstellen von $f'(x)$, zum Funktionswerte C_1-summierbar ist, ergibt das Haarsche Theorem (5. 19) dasselbe auf die fragliche Reihe.[55] Ähnlich schließt man auf gleichmäßige Summierbarkeit in (inneren) Stetigkeitsstrecken mit Hilfe von 5° bzw. des *Fejérschen Approximationssatzes.*

Somit sind die Behauptungen über $\Sigma\gamma_\lambda\varphi'_\lambda(x)$ von (5. 27) ablesbar.

7° Um die Reihe $\Sigma\gamma_\lambda\varphi'_\lambda(x)$ zu untersuchen, bemerken wir vor allem, daß die Benutzung von (5. 1), (5. 3) nebst der Transformation (5. 10), (5. 11) für $\overline{\varphi}_\lambda(u) = [(g(x)w(x))^{\frac{1}{4}}\varphi_\lambda(x)]_{x=V(u)}$ die asymptotische Formel

$$(5.28)\qquad \overline{\varphi}_\lambda(u) = \sqrt{\frac{2}{H}}\sin\frac{\pi\lambda}{H}u + \frac{\overline{\omega}(u)}{\lambda}\cos\frac{\pi\lambda}{H}u + O\left(\frac{1}{\lambda^2}\right) \qquad (\lambda = 1, 2, \ldots)$$

liefert, wobei $\overline{\omega}(u)$ in $[0, H]$ stetig, von λ frei und das Zeichen O in u gleichmäßig zu verstehen ist.[56]

Daraus schließt man einfach, daß unsere Reihe, welche einer Funktion $f(x) \in \mathrm{Lip}\,1$ angehört, in (a, b) konvergent ist und zwar in inneren Teilintervallen gleichmäßig konvergiert, ferner gleichmäßig C_1-summierbar ist. Was ihre Summe betrifft, so beachte man die nach (5. 19) für *jedes* in $[0, H]$ stetige $F(u)$ gültige Relation

$$(5.29)\qquad \lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\sum_{\lambda=0}^{k}\varphi^*_\lambda(u)\int_0^H F(t)\varphi^*_\lambda(t)dt = F(u) \qquad (0 \leqq u \leqq H).$$

Multipliziert man mit $(g(x)w(x))^{\frac{1}{4}}$, differenziert[57] nach x und schreibt $F(u) = \left[(g(x)w(x))^{-\frac{1}{4}}\int_a^x f(t)w(t)dt\right]_{x=V(u)}$, so folgt (vgl. (5. 1), (5. 13), (5. 14))

$$(5.30)\qquad \lim_{n\to\infty}\frac{1}{n+1}\sum_{k=1}^{n}\sum_{\lambda=1}^{k}\varphi_\lambda(x)\int_a^b f(t)w(t)\varphi_\lambda(t)dt = f(x) \qquad (a < x < b).$$

Wir haben also zugleich in (a, b)

$$(5.31)\qquad \sum_{\lambda=1}^{\infty}\gamma_\lambda\varphi_\lambda(x) = f(x),$$

w. z. b. w.

[55] Da $u = \int_a^x \left(\frac{w(t)}{g(t)}\right)^{\frac{1}{2}}dt$ und $x = V(u)$ in $[a, b]$ bzw. $[0, H]$ zur Klasse Lip 1 gehören, ersieht man leicht, daß Nullmengen auf der x- und u-Achse ineinander übergehen.

[56] Vgl. [19], S. 262.

[57] Dies ist bekanntlich gestattet, weil der Limes in (5. 30) nach den vorigen in jedem Inneren Teilintervalle von (a, b) gleichmäßig existiert.

5. 2. Wie man sieht, entspricht der soeben bewiesene Satz der Differentiation gewöhnlicher Fourierschen Sinusreihen. Es können offenbar ähnliche Sätze aufgestellt werden, in welchen ein anderes Paar homogen-linearer Randbedingungen, z. B. $\varphi_\lambda(a) = \varphi_\lambda(b)$, $\varphi'_\lambda(a) = \varphi'_\lambda(b)$ $(\lambda = 1, 2, \ldots)$ auftritt;[58] letzteres umfaßt auch den Fall der allgemeinen trigonometrischen Fourierreihe.

(Eingegangen am 20. März 1954.)

Literaturverzeichnis

[1] J. Aczél, Eine Bemerkung über die Charakterisierung der klassischen orthogonalen Polynome, *Acta Math. Acad. Sci. Hung.*, **4** (1953), S. 315—321.
[2] S. Bochner, Über Sturm—Liouvillesche Polynomsysteme, *Math. Zeitschrift*, **29** (1929), S. 730—736.
[3] H. Cramér, *Mathematical methods of statistics* (Princeton, 1946).
[4] R. Courant und D. Hilbert, *Methoden der mathematischen Physik*, Bd. I (Berlin, 1924).
[5] J. Egerváry, Über die charakteristischen geometrischen Eigenschaften der Legendreschen und Tschebyscheffschen Polynome, *Archiv Math. u. Physik* (3), **27** (1918), S. 23—24.
[6] L. Fejér, Sur la différentiation de la série de Fourier, *Comptes Rendus Acad. Sci. Paris*, **134** (1902), S. 762—765.
[7] L. Fejér, Untersuchungen über Fouriersche Reihen, *Math. Annalen*, **58** (1904), S. 51—69.
[8] L. Fejér, Sur les singularités de la série de Fourier des fonctions continues, *Ann. École Normale Sup.*, **28** (1911), S. 64—103.
[9] Ph. Frank und R. Mises, *Die Differential- und Integralgleichungen der Mechanik und Physik*, Bd. I (Braunschweig, 1925).
[10] B. Gagaeff, Sur l'unicité de fonctions orthogonales invariant relativement à la dérivation, *Comptes Rendus Acad. Sci. Paris*, **188** (1929), S. 222—225.
[11] Б. Гагаев, О некоторых классах ортогональных функций, Известия Академии Наук СССР, Сер. матем., **10** (1946), S. 197—206.
[12] Б. Гнеденко, О единственности системы ортогональных функций, инвариантной относительно дифференцирования, Доклады Академии Наук СССР, **14** (1939), S. 159—161.
[13] A. Haar, Zur Theorie der orthogonalen Funktionensysteme (Habilitationsschrift), 1. und 2. Mitteil., *Math. Annalen*, **69** (1910), S. 331—371, bzw. ibid. **71** (1911), S. 38—53.
[14] W. Hahn, Über die Jacobischen Polynome und zwei verwandte Polynomklassen, *Math. Zeitschrift*, **39** (1935), S. 634—638.
[15] W. Hahn, Über höhere Ableitungen von Orthogonalpolynomen, ibid., **43** (1937), S. 101.
[16] G. H. Hardy and W. W. Rogosinski, *Fourier series* (Cambridge, 1944).
[17] E. W. Hobson, On a general convergence theorem and the theory of a function by series of normal functions, *Proc. London Math. Soc.* (2), **6** (1908) S. 349—395.

[58] Über die Asymptotik der Orthogonalfunktionen in den Fällen allgemeinerer Randbedingungen vgl. [31].

[18] D. Jackson, *Fourier-series and orthogonal polynomials* (Oberlin—Ohio, 1941).
[19] E. Kamke, *Differentialgleichungen, Lösungsmethoden und Lösungen*, Bd. I (Leipzig, 1951), 4. Auflage.
[20] E. Kamke, *Differentialgleichungen reeller Funktionen* (Leipzig, 1952), 2. Auflage.
[21] H. L. Krall, On derivatives of orthogonal polynomials, *Bull. Amer. Math. Soc.*, **42** (1936), S. 423—428.
[22] H. L. Krall, On higher derivatives of orthogonal polynomials, ibid., **42** (1936), S. 867—870.
[23] H. L. Krall, On higher derivatives of orthogonal polynomials. II, ibid., **47** (1941), S. 261—264.
[24] М. А. Лаврентьев и Л. А. Люстерник, Курс вариационного исчисления (Москва—Ленинград, 1950).
[25] D. C. Lewis, Orthogonal functions whose derivatives are also orthogonal, *Rend. Palermo* (2), **11** (1953), S. 159—168.
[26] F. Riesz et B. Sz.-Nagy, *Leçons d'analyse fonctionnelle* (Budapest, 1953), 2. Auflage.
[27] W. Rogosinsky, *Fouriersche Reihen* (Berlin—Leipzig, 1930), S. Göschen.
[28] J. Shohat, E. Hille and J. Walsh, *A bibliography on orthogonal polynomials*, Bull Nat. Res. Council, **103** (Washington, 1940).
[29] G. Szegő, *Orthogonal polynomials* (New York, 1939).
[30] M. S. Webster, Orthogonal polynomials with orthogonal derivatives, *Bull. Amer. Math. Soc.*, **44** (1938), S. 880—888.
[31] A. C. Zaanen, On some orthogonal systems of functions, *Compositio Math.*, **7** (1940) S. 253—282.
[32] A. Zygmund, *Trigonometrical series* (Warszawa—Lwów, 1935).

О НЕКОТОРЫХ СВОЙСТВАХ ОРТОГОНАЛЬНЫХ СИСТЕМ И О СОБСТВЕННЫХ ЗНАЧЕНИЯХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ ТИПА ШТУРМА—ЛИУВИЛЛЯ

М. Миколаш (Будапешт)

(Резюме)

С точки зрения применений не лишен интереса факт, что тригонометрическая ортогональная система и системы „классических" ортогональных полиномов при дифференцировании и при интегрировании (выполненном после умножения на весовую функцию) переходят опять в ортогональные системы, ведь это является причиной того, что ортогональные ряды по элементам указанных систем, в аналогии с степенными рядами, сохраняют свой „характер" при почленном дифференцировании и интегрировании.

Настоящая работа посвящена изучении тех ортогональных системах пространства L^2, которые обладают указанными свойствами важнейших ортогональных систем. Более подробно исследуются следующие проблемы:

I. Пусть $\varphi_\lambda(x) \in L^2_w(a, b)$ $(\lambda = 1, 2, \ldots)$ — ортогональная система, а $w(x)$ — неотрицательная весовая функция. Пусть даны некоторые постоянные $c_1, c_2, \ldots$ и другая весовая функция $Q(x) \geqq O$ $(a \leqq x \leqq b)$. Спрашивается необходимое и достаточное условие для того, чтобы — при обозначении $\Phi_\lambda(x) = \int_a^x w(t)\,\varphi_\lambda(t)\,dt$ $(\lambda = 1, 2, \ldots)$ — система $\{\Phi_\lambda(x) + c_\lambda\}$ или $1, \{\Phi_\lambda(x) + c_\lambda\}$ также была ортогональной системой на (a, b) относительно $Q(x)$.

II. Пусть $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ система функций дифференцируемых на (a, b), ортогональная на этом интервале относительно весовой функции $w(x)$ (w-ортогональная). Зададимся также другой весовой функцией $g(x) \geqq 0$ $(a \leqq x \leqq b)$. Спрашивается условие необходимое и достаточное для того, чтобы система $\{\varphi'_\lambda(x)\}$ была g-ортогональна на (a, b).

III. Требуется задать и изучать по возможности широкий класс ортогональных систем, элементы которого обладают свойствами I и II.

IV. На основе изложенного следует охарактеризовать „классические" ортогональные системы, в частности многочлены Якоби, Лагерра и Эрмита, а также собственные функции Штурма—Лиувилля.

V. Что можно сказать о сходимости или о суммируемости (общего) ряда Фурье и его производного ряда, если известно, что ряд полученный почленным дифференцированием также ортогонален?

Эти вопросы кажутся новыми, однако следует упоминать о многочисленных частных результатах, дающих различные от общепринятых определения классически ортогональных полиномов (см. [28]). Сюда относится и одно замечание Гагаева о системе $1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots$. Исследования, отчасти примыкающие к II, можно найти и в новейшей работе Д. С. Льюиса [25], о которой автор узнал при корректуре.

Главные результаты настоящей работы состоят в следующем:

1. Полная ортогональная система обладает свойствами I и II тогда и только тогда, если ее элементы удовлетворяют почти всюду некоторому интегральному уравнению, или (что в случае непрерывных функций этому эквивалентно) дифференциальному уравнению типа Штурма—Лиувилля

$$[g(x)y']' = \varrho_\lambda w(x)y,$$

связанному удобными краевыми условиями. (См. теоремы 1, 2 и 3.)

2. Таким образом легко найти некоторый общий класс ортогональных систем, из которых дифференцированием, также как и интегрированием, получается опять ортогональная система. Итак, класс этих „обыкновенных систем SL" содержит „классические" ортогональные системы. (Теоремы 4 и 5.)

3. В качестве применения автору удается дать общую характеристику тригонометрической системы, собственных функций Штурма—Лиувилля первого и второго рядов и классических ортогональных полиномов, использующую только свойства общей природы (ограниченность, непрерывность, дифференцируемость, ортогональность, полнота) без требования выполненности дифференциального уравнения или другого специального соотношения. (Теоремы 6, 7, 8.) Пример: С точностью до перестановок и до тривиальных линейных преобразований, система $1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots$ является единственной полной ортогональной системой в интервале $(0, 2\pi)$, для которой система первых и вторых производных, дополненная единицей, также полная ортогональная система на том же интервале.

4. Наконец, в связи с проблемой V, доказывается теорема, которая может рассматриваться как обобщение результата, полученного Фейером для обыкновенных рядов Фурье ([6], [7]). В этой теореме утверждается, по существу, следующее: если для некоторой функции $f(x) \in \mathrm{Lip}\, 1$ $(a \leqq x \leqq b)$ ее ряд Фурье по системе $\{\varphi_\lambda(x)\}$ $(\varphi_\lambda(a) = \varphi_\lambda(b) = 0,\ \lambda = 1, 2, \ldots)$ w-ортогональной и w-полной на (a, b) переходит при почленном дифференцировании в ряд также ортогональный на (a, b), то ряд Фурье сходится на (a, b) и его сумма равна $f(x)$, производный же ряд почти всюду является суммируемым C_1 и его C_1-сумма равна $f'(x)$; суммируемость является равномерной во всяком замкнутом интервале непрерывности $f'(x)$.

ÜBER TRIGONOMETRISCHE INTEGRALE, WELCHE GANZE FUNKTIONEN MIT NUR REELLEN NULLSTELLEN DARSTELLEN

Von
LJUBOMIR ILIEFF (Sofia)
(*Vorgelegt von* P. TURÁN)

1. Es sei $E(a)$ die Gesamtheit der Funktionen $f(t)$, die im Intervall $(0, a)$, $a > 0$, nicht negativ, R-integrierbar sind, und der Bedingung genügen, daß die ganze Funktion

$$F(z) = \int_0^a f(t) \cos tz \, dt \tag{F}$$

nur reelle Nullstellen besitzt.

Es sei $x = x(t)$ ein reelles, gerades Polynom von t oder auch eine reelle, gerade, ganze Funktion, so daß

a) $x(a) = 0$,

b) $x'(it)$ ein Polynom mit nur reellen Nullstellen oder eine ganze Funktion, die eine Grenze solcher Polynome ist.

Da $x'(t)$ eine reelle Funktion ist, die außer $t = 0$ keine reellen Nullstellen besitzt, so folgt, daß sie in $-\infty < t < 0$ und in $0 < t < +\infty$ ihr Vorzeichen nicht wechselt. $x(t)$ genügt also auch der Bedingung:

c) In jedem der Intervalle $(-\infty, 0)$, $(0, +\infty)$ ist die Funktion $x(t)$ entweder streng wachsend oder streng abnehmend.

Da, wegen der Bedingung b), $x'(a) \neq 0$ ist, so folgt aus c), daß $x(t)$ auch der Bedingung genügt:

d) $x(t)$ besitzt nur zwei reelle Nullstellen: $t_1 = a$ und $t_2 = -a$. Diese Nullstellen sind einfach.

Es sei $A(a)$ die Klasse der reellen, geraden Funktionen $x(t)$, die den Bedingungen a) und b) genügen.

Gemäß der Bedingung c) hat jede Funktion von der Klasse $A(a)$ im Intervall $[0, a]$ eine inverse Funktion $t = t(x)$.

Es sei $f_0 \in E(a)$, $x(t) \in A(a)$ und $t = t(x)$ sei die inverse Funktion von $x(t)$ im Intervall $[0, a]$.

Man setze

$$\varphi_0(x) = f_0(t), \quad f_1(t) = \varphi_1(x) = \int_0^x \varphi_0(v)\,dv$$

und bezeichne, der Kürze wegen,

$$f_1(t) = L(f_0, x).$$

SATZ I. *Wenn* $f_0(t) \in E(a)$, $x(t) \in A(a)$ *und* $f_1(t) = L(f_0, x)$, *so haben wir* $f_1(t) \in E(a)$.

BEWEIS. Da $f_0(t) \in E(a)$, so besitzt die ganze Funktion

$$F_0(z) = \int_0^a f_0(t) \cos tz\,dt$$

nur reelle Nullstellen. Man setze $f_0(-t) = f_0(t)$ für $0 \leqq t \leqq a$. Dann ist

$$F_0(z) = \frac{1}{2}\int_{-a}^{a} f_0(t) e^{itz}\,dt.$$

Gemäß der Bedingung b) und einem Ergebnis von G. PÓLYA [1] hat auch die ganze Funktion

$$\int_{-a}^{a} f_0(t)\,x'(t) e^{itz}\,dt = 2i\int_0^a f_0(t) x'(t) \sin tz\,dt$$

auter reelle Nullstellen.

Es ist

$$\int_0^a f_0(t)x'(t) \sin tz\,dt = \int_{t=0}^{t=a} \varphi_0(x) \sin tz\,dx = \int_{t=0}^{t=a} \sin tz\,d\varphi_1(x) =$$

$$= \varphi_1(x) \sin tz \Big|_{t=0}^{t=a} - z\int_0^a \varphi_1(x) \cos tz\,dt = -z\int_0^a f_1(t) \cos tz\,dt,$$

da $\varphi_1[x(a)] = \varphi_1(0) = 0$. Der Satz ist bewiesen.

2. Mit Hilfe des Satzes I bekommt man also aus jeder Funktion von der Form (F) mit nur reellen Nullstellen eine neue Funktion derselben Art.

Die Funktion $x(t) \in A(a)$ genüge noch der Bedingung: a') $x(0) > 0$. In diesem Falle nimmt die Funktion $x(t)$ im Intervall $[0, a]$ zu, so daß $x^{-\alpha}$, wo $0 \leqq \alpha < 1$ ist, eine wachsende Funktion in demselben Intervall darstellt. Nach wohlbekannten Ergebnissen von G. PÓLYA [2] folgt, daß die Funktion $f_0(t) = x^{-\alpha}$ zu der Klasse $E(a)$ gehört. Aus dem Satze I folgt in diesem Falle, daß die Funktionen $f_0(t) = \varphi_0(x) = x^{-\alpha}$, $(1-\alpha)f_1(t) = (1-\alpha)\varphi_1(x) = x^{1-\alpha}, \ldots,$ $(1-\alpha)(2-\alpha)\ldots(n-\alpha)f_n(t) = x^{n-\alpha}, \ldots$ zu der Klasse $E(a)$ gehören. Da zu jedem $\lambda > -1$ immer ein ganzes n und eine Zahl α mit $0 \leqq \alpha < 1$ gehört, so daß $\lambda = n - \alpha$ ist, bekommt man die

FOLGERUNG. *Wenn* $x(t) \in A(a)$, $x(0) > 0$ *und* $\lambda > -1$ *besteht, so hat die ganze Funktion*

$$\int_0^a x^\lambda \cos tz\, dt$$

nur reelle Nullstellen.

So gehört z. B. die Funktion $x = 1 - t^{2q}$, wo q eine positive ganze Zahl bezeichnet, zu der Klasse $A(1)$, wobei $x(0) > 0$. Aus der Folgerung bekommt man das bekannte Ergebnis von G. PÓLYA [1], nach dem die ganze Funktion

$$\int_0^1 (1 - t^{2q})^\lambda \cos tz\, dt,$$

falls $\lambda > -1$ und q eine ganze positive Zahl bezeichnet, lauter reelle Nullstellen besitzt.

3. Es sei $f(t)$ ein reelles, gerades Polynom, so daß $f'(it)$ lauter reelle Nullstellen besitzt. In diesem Falle hat $f'(t)$ im Intervall $(0, +\infty)$ ein und dasselbe Vorzeichen, so daß $f(t)$ daselbst streng monoton ist. Da aber auch $f''(t) \neq 0$ für $0 < t < +\infty$ ist, haben wir entweder $\lim_{t\to\infty} f(t) = +\infty$ oder $\lim_{t\to\infty} f(t) = -\infty$. Im ersten Falle, wenn man eine passende Konstante zufügt, kann man annehmen, daß $f(t) \geqq 0$; und im zweiten Falle, daß $f(t) \leqq 0$. Es sei $f(t) \geqq 0$. Dann stellt $f(t)$ eine wachsende Funktion im Intervall $(0, +\infty)$ dar. Ist also $n > f(0)$ eine beliebige Zahl, so existiert wegen $\lim_{t\to\infty} f(t) = +\infty$ eine Zahl $a_n > 0$, so daß $f(a_n) = n$. Wegen der Monotonie von $f(t)$ gibt es nur ein einziges a_n, wo $\lim_{n\to+\infty} a_n = +\infty$. Es ist also $1 - \frac{f(t)}{n} \in A(a_n)$. Jetzt liefert unsere Folgerung das Ergebnis, daß die ganze Funktion

$$\int_0^{a_n} \left(1 - \frac{f(t)}{n}\right)^n \cos tz\, dt$$

nur reelle Nullstellen besitzt. Da im Intervall $(0, a_n)$ die Ungleichungen $0 < f(t) < n$, $0 < \left(1 - \frac{f(t)}{n}\right)^n < e^{-f(t)}$ bestehen, so bekommt man, daß die ganze Funktion $\int_0^\infty e^{-f(t)} \cos tz\, dt$ lauter reelle Nullstellen besitzt. Durch Grenzübergang bekommt man:

SATZ II. *Wenn* $f(t)$ *eine nichtnegative, gerade Funktion bezeichnet, so daß* $f'(it)$ *ein Polynom mit nur reellen Nullstellen oder eine Grenze solcher Polynome ist, so besitzt die ganze Funktion*

$$\int_0^\infty e^{-f(t)} \cos tz\, dt$$

nur reelle Nullstellen.

Die Funktion $f(t) = a \operatorname{ch} t = \frac{a}{2}(e^t + e^{-t})$, $a > 0$, genügt den Bedingungen von Satz II. So bekommt man aus demselben Satz das bekannte Ergebnis von G. PÓLYA [3], nach dem die Funktion

$$\int_0^\infty e^{-a \operatorname{ch} t} \cos tz \, dt$$

nur reelle Nullstellen besitzt.

(Eingegangen am 12. November 1954.)

Literaturverzeichnis

[1] G. PÓLYA, Über trigonometrische Integrale mit nur reellen Nullstellen, *Journal f. reine u. angew. Math.*, **158** (1927), S. 6—18.

[2] G. PÓLYA, Über die Nullstellen gewisser ganzer Funktionen, *Math. Zeitschr.*, **2** (1918), S. 352—383.

[3] G. PÓLYA, Bemerkung über die Integraldarstellung der Riemannschen ξ-Funktion, *Acta Math.*, **48** (1926), S. 305—317.

О ТРИГОНОМЕТРИЧЕСКИХ ИНТЕГРАЛАХ, ПРЕДСТАВЛЯЮЩИХ ЦЕЛЫЕ ФУНКЦИИ, ИМЕЮЩИЕ ТОЛЬКО ВЕЩЕСТВЕННЫХ КОРНЕЙ

Л. Илиев (София)

(Резюме)

В настоящей работе дается простое доказательство двух теорем.

Теорема I. Если $x(t)$ — четная целая функция, для некоторого положительного a $x(a) = 0$, а $x'(it)$ является пределом многочленов, все корни которых вещественны, то из того, что все корни интеграла

$$\int_0^a f_0(t) \cos tz \, dt$$

вещественны, следует, что

$$\int_0^a f_1(t) \cos tz \, dt$$

также обладает этим свойством, где

$$\varphi_0(x) = f_0(t), \qquad f_1(t) = \int_0^x \varphi_0(v) \, dv.$$

Теорема II. Если $f(t) \geqq 0$ — четная функция и $f'(it)$ является пределом многочленов, имеющих только вещественные корни, то все корни интеграла

$$\int_0^\infty e^{-f(t)} \cos tz \, dt$$

вещественны.

PROJECTIONS OF PROBABILITY DISTRIBUTIONS

By
WALTER M. GILBERT (Pullman, USA)
(*Presented by* A. RÉNYI)

§ 1. Introduction

A two dimensional probability distribution generates a family of one dimensional distributions by (orthogonal) projection upon straight lines in the plane of the distribution. Since projections on parallel lines are the same it is sufficient to consider only lines through the origin. It is of interest to see what properties of the "parent" distribution can be determined by an examination of some or all of the projection distributions. If all of the projections are known the parent is completely determined, by a theorem of H. CRAMÉR and H. WOLD [1]. A. RÉNYI [3] has given a condition under which a two dimensional distribution is determined by any infinite set of projections. His condition is that the two dimensional characteristic function be an analytic function of θ for every r, where r and θ are polar coordinates. In this note we show that in general a distribution need not be determined by an infinite set of projections, and we give another condition which insures determination. We also show that if two projections of a distribution have analytic characteristic functions, then so have all the others, and the two dimensional characteristic function is itself analytic.

§ 2. Determination of the parent

The fundamental theorem on the subject is due to CRAMÉR and WOLD [1]. It states that any distribution is determined by the set of all of its projections. This is established by showing that the two dimensional characteristic function, considered on the line through the origin of inclination θ, coincides with the characteristic function of the projection on a line of inclination θ. RÉNYI's problem is thus equivalent to finding conditions under which a two dimensional characteristic function is determined by its values on an infinite set of lines through the origin.

That an infinite number of projections need not determine a two dimensional distribution, is clear from a slight modification of a well-known example of two distinct one dimensional characteristic functions which coincide on a finite interval. Such an example, attributed to KHINTCHINE, is given in LÉVY [2]. Let $\varphi_1(t)$ equal $1-|t|$ for $|t| \leqq 1$ and zero for $|t| \geqq 1$, and let $\varphi_2(t)$ coincide with $\varphi_1(t)$ for $|t| \leqq 1$ but be periodic with period 2. (The distributions of which these are the characteristic functions are given in [2].) To form our example we consider the two dimensional characteristic functions $\psi_1(t_1, t_2) = \varphi_1(t_1)\varphi_1(t_2)$ and $\psi_2(t_1, t_2) = \varphi_1(t_1)\varphi_2(t_2)$. Then ψ_1 and ψ_2 coincide along all lines L_θ through the origin for which $\pi/4 \leqq \theta \leqq 3\pi/4$, but differ along all other lines. It may be noted that neither of the distributions associated with ψ_1 and ψ_2 has any moments.

On the other hand, we may prove the following theorem.

THEOREM 1. *If the moments of a distribution $F(x_1, x_2)$ are those of a determined moment problem, then $F(x_1, x_2)$ is determined by its projections on any infinite set of distinct lines through the origin.*

PROOF. It is sufficient to show that the values at the origin of all of the derivatives of the characteristic function $\psi(t_1, t_2)$ of $F(x_1, x_2)$ may be calculated from the characteristic functions $\varphi_i(t)$ of a denumerable set of the projection distributions $F_i(x)$.

Consider then characteristic functions $\varphi_i(t)$ $(i = 1, 2, \ldots, n+1)$ of $n+1$ projections on $n+1$ distinct lines, where the line L_i makes an angle $0 \leqq \theta_i < \pi$ with the x_1 axis. By the theorem of CRAMÉR and WOLD, $\varphi_i(t)$ is the value of $\psi(t_1, t_2)$ along the line through the origin making an angle θ_i with the t_1 axis. We may differentiate each of the $\varphi_i(t)$ n times and put $t = 0$. This yields $n+1$ linear equations:

$$\varphi_i^{(n)}(0) = \sum_{j=0}^{n} C_j^n \cos^{n-j} \theta_i \sin^j \theta_i \, . \, \psi^{(n-j, j)}(0, 0).$$

These may be solved for $\psi^{(n-j, j)}(0, 0)$ provided the determinant of the coefficients is not zero. To show that the determinant is not zero, suppose for the moment that no θ_i is $\pi/2$. We may then divide the $j+1$-st column by C_j^n and the i-th row by $\cos^n \theta_j$. Letting $y_i = \operatorname{tg} \theta_i$ we have the Vandermonde determinant, which will be zero only for $\operatorname{tg} \theta_i = \operatorname{tg} \theta_j$ with $i \neq j$. But since $0 \leqq \theta_i < \pi$ this will occur only for $\theta_i = \theta_j$, and we have supposed the θ_i's distinct. If some $\theta_i = \pi/2$, all the elements of the i-th row are zero save the $n+1$-st, which is equal to one. We may then apply the above argument to the cofactor of this element.

A distribution may satisfy RÉNYI's condition without satisfying ours. For example, the characteristic function associated with the distribution with density $\dfrac{1}{\pi(1+r^2)^2}$ is constant for fixed r, but the distribution has no proper moments. An example of a distribution whose moment problem is determined

but whose characteristic function is not analytic (if such exists) would serve to show the converse.

A condition which insures that a two dimensional moment problem is determined has been given by CRAMÉR and WOLD [1], p. 291. We may apply it to obtain a corollary which imposes requirements on the projections rather than the two dimensional distribution.

COROLLARY. Suppose given the projections of a two dimensional distribution on a countable set of distinct lines through the origin. Let μ_n^i be the n-th moment of the i-th projection. Suppose that the j-th and k-th projections are such that

$$\sum_{n=1}^{\infty} (\mu_{2n}^j + \mu_{2n}^k)^{-\frac{1}{2n}} = \infty .$$

Then the two dimensional distribution is uniquely determined by the set of projections.

§ 3. Projections with analytic characteristic functions

The proof of the CRAMÉR—WOLD theorem suggests the following theorem, which, however, does not lead to unique determination of the original distribution.

THEOREM 2. *If two projections of a two dimensional distribution have analytic characteristic functions, then the same is true for any projection, and the characteristic function of the parent distribution is an analytic function of two complex variables in a neighbourhood of the origin.*

PROOF. In WINTNER [4] it is shown that a characteristic function is analytic for $t < q$ if and only if $\mu_{2n} = O[(2n)! p^{-2n}]$ as $n \to \infty$ for every $0 < p < q$. Following [4] and using Stirling's formula, $(\mu_{2n})^{\frac{1}{2n}} \leqq kn$ is equivalent to $\mu_{2n} = O[(2n)! (ke)^{2n} 2^{-2n}]$. Thus $(\mu_{2n})^{\frac{1}{2n}} \leqq kn$ is necessary and sufficient for the analyticity of the characteristic function, and the k determines a minimum radius of convergence of the Taylor series about the origin.

It is no restriction to suppose that the two lines on which $\psi(t_1, t_2)$ is analytic are the coordinate axes, since a linear transformation will make this so. We have then $(\mu_{2n}^1)^{\frac{1}{2n}} \leqq k_1 n$ and $(\mu_{2n}^2)^{\frac{1}{2n}} \leqq k_2 n$. Let $\mu_{2n}^1 + \mu_{2n}^2 = \lambda_{2n}$. Then $\lambda_{2n} < [(k_1)^{2n} + (k_2)^{2n}] n^{2n} \leqq (k_1 + k_2)^{2n} n^{2n}$ so that $\lambda_{2n}^{\frac{1}{2n}} < kn$ also. Let F_c be the distribution resulting from a projection on a line with direction cosines c_1 and c_2. Then

$$\mu_{2n}(F_c) = \int\int (c_1 x_1 + c_2 x_2)^{2n} dF(x_1, x_2) \leqq 2^{2n} \int\int (x_1^{2n} + x_2^{2n}) dF(x_1, x_2) = 2^{2n} \lambda_{2n}$$

and so there is a K such that for any projection $[\mu_{2n}(F_c)]^{\frac{1}{2n}} < Kn$, and a corresponding uniform minimum radius of convergence. It follows that the multiple Taylor series converges for any point within a circle about the origin in the real (t_1, t_2) plane and therefore within a polycylinder in the space of two complex variables.

The argument of the last paragraph is easily applied to show also that if any two projections have n moments, then so has any other projection and so has the parent distribution.

STATE COLLEGE OF WASHINGTON

(*Received 12 November 1954*)

References

[1] H. Cramér and H. Wold, Some theorems on distribution functions, *Journal of the London Mathematical Society*, **11** (1936), pp. 290—294.
[2] P. Lévy, *Théorie de l'addition des variables aléatoires* (Paris, 1937).
[3] A. Rényi, On projections of probability distributions, *Acta Math. Acad. Sci. Hung.*, **3** (1952), pp. 131—142.
[4] A. Wintner, *The Fourier transforms of probability distributions* (Baltimore, 1947).

О ПРОЕКЦИЯХ РАСПРЕДЕЛЕНИЙ ВЕРОЯТНОСТЕЙ

В. М. Джилберт (Пульман, США)

(Резюме)

А. Реньи в работе [3] ставил вопрос об условиях, при которых двумерное распределение вероятностей определяется его проекциями на бесконечное множество различных прямых, проходящих через начало координат; он же дал достаточное условие этого. В настоящей работе приводится, с одной стороны, простой пример распределения, не являющегося определенным бесконечным множеством его различных проекций, а с другой стороны доказывается, что если для двумерного распределения (двумерная) проблема моментов является определенным, то распределение определяется любым бесконечным множеством его различных проекций. Доказывается, наконец, что если характеристические функции двух проекций двумерного распределения аналитичны на интервале содержащем точку O, то этим же свойством обладают все проекции данного распределения, а двумерная характеристическая функция распределения также аналитична в области содержащей начало координат.

REFINEMENT OF A PROBABILITY LIMIT THEOREM AND ITS APPLICATION TO BESSEL FUNCTIONS

By
M. FISZ (Warszawa)
(*Presented by* A. RÉNYI)

1. We consider two independent Poisson random variables X_1 and X_2 distributed according to the law

$$P(X_i = k) = e^{-w_i} \frac{w_i^k}{k!}, \tag{1}$$

where $w_i > 0$; $i = 1, 2$; $k = 0, 1, 2, \ldots$.

The difference $X_2 - X_1$ is, as $w_1 \to \infty$ and $w_2 \to \infty$, asymptotically normal. This has been shown by J. O. IRWIN [1] for the special case $w_1 = w_2$, and for the general case by M. FISZ [2] and [3].

The problem of the refinement of these limit theorems is dealt with in this paper. The results obtained are given in formulae (7) and (11) and as a byproduct the asymptotical formula (13) for Bessel functions with purely imaginary argument has been obtained.

2. Let us consider the Poisson random variables mentioned above and let us suppose that

$$w_i = n d_i \qquad (i = 1, 2), \tag{2}$$

where d_1 and d_2 are some real and positive constants. Let us denote

$$z = \frac{k - (w_2 - w_1)}{\sqrt{w_1 + w_2}}, \tag{3}$$

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right), \tag{4}$$

$$\mu_{r+2} = \begin{cases} (w_1 + w_2)^{-\frac{r}{2}} & (r = 2, 4, 6, \ldots), \\ (w_2 - w_1)(w_1 + w_2)^{-\frac{r+2}{2}} & (r = 1, 3, 5, \ldots). \end{cases} \tag{5}$$

Finally let $F_r[-\varphi(z)]$ denote the function which can be obtained for each value of $r=1, 2, \ldots$ from the identity:

$$\exp\left\{\sum_{r=1}^{\infty}\frac{\mu_{r+2}}{(r+2)!}\left[-\varphi(z)\right]^{(r+2)}\left(\frac{1}{\sqrt{n}}\right)^r\right\}=1+\sum_{r=1}^{\infty}F_r[-\varphi(z)]\left(\frac{1}{\sqrt{n}}\right)^r \tag{6}$$

by comparing the coefficients of $\frac{1}{\sqrt{n}}$ on both sides of this identity. In the formula (6) we have

$$[-\varphi(z)]^{(r+2)}=(-1)^{r+1}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right)H_{r+2}(z),$$

where $H_{r+2}(z)$ is the Hermite polynomial of order $r+2$.

The following theorem will be proved.

THEOREM. *Let X_1 and X_2 be two independent random variables distributed according to* (1) *and let the equalities* (2) *hold. Then, as $w_i\to\infty$ $(i=1, 2)$, the asymptotical formula*

$$P_{w_1w_2}(k)=P(X_2-X_1=k)=$$

$$=\frac{1}{\sqrt{w_1+w_2}}\left\{\varphi(z)+\sum_{r=1}^{u}F_r[-\varphi(z)]\right\}+o\left[\frac{1}{(w_1+w_2)^{\frac{u+1}{2}}}\right] \tag{7}$$

holds for each integer $u\geqq 1$.

PROOF. We deduce the assertion of our Theorem from the following theorem given by C. G. ESSEEN [4]:

Let $\xi_1, \xi_2, \ldots$ be independent and equally distributed, integer-valued random variables having the mean value m and the standard deviation $\sigma\neq 0$ and finite moments of order $v\geqq 3$ and let the greatest common divisor of the differences of pairs of all possible values of ξ be equal to 1. Then the asymptotical formula

$$P(\xi_1+\xi_2+\cdots+\xi_n=k)=$$

$$=\frac{1}{\sigma\sqrt{n}}\left\{\varphi(z)+\sum_{r=1}^{v-2}\frac{1}{n^{\frac{r}{2}}}U_r[-\varphi(z)]\right\}+o\left(\frac{1}{n^{\frac{v-1}{2}}}\right) \tag{8}$$

holds, where

$$z=\frac{k-mn}{\sigma\sqrt{n}},$$

$\varphi(z)$ is given by (4) and the functions U_r can be determined for each r from the formula

$$\exp\left\{\sum_{r=1}^{\infty}\frac{\lambda_{r+2}}{(r+2)!}[-\varphi(z)]^{(r+2)}\left(\frac{1}{\sqrt{n}}\right)^r\right\}=1+\sum_{r=1}^{\infty}U_r[-\varphi(z)]\left(\frac{1}{\sqrt{n}}\right)^r, \tag{9}$$

in which λ_{r+1} is the $(r+2)$-th semiinvariant of the variable $\frac{\xi}{\sigma}$.

The applicability of the theorem of ESSEEN to our Theorem follows from the fact that the addition theorem is satisfied for Poisson variables and their differences. Taking into account the relations (2), we can write

$$X_i = \sum_{k=1}^{n} \eta_{ik} \qquad (i = 1, 2),$$

where η_{ik} $(k = 1, 2, \ldots)$ are independent and equally distributed Poisson variables having the mean value d_i $(i = 1, 2)$. The variables

$$\xi_k = \eta_{2k} - \eta_{1k} \qquad (k = 1, 2, \ldots)$$

satisfy all the assumptions of the theorem of ESSEEN and have finite moments of arbitrary order. The semiinvariants of ξ_k are given by the formulae

$$\lambda_{r+2} = \begin{cases} n^{\frac{r}{2}} (w_1 + w_2)^{-\frac{r}{2}} & (r = 2, 4, \ldots), \\ n^{\frac{r}{2}} (w_2 - w_1)(w_1 + w_2)^{-\frac{r+2}{2}} & (r = 1, 3, \ldots). \end{cases} \tag{10}$$

Taking into account the formula (10), we obtain from (8) and (9) the formula (7).

Our Theorem is thus proved.

In the case when X_1 and X_2 are equally distributed, thus when $w_1 = w_2$, all coefficients μ_{r+2} given by (5) vanish for odd values of r. It is then convenient to write the formula (7) in the following form:

$$P_w(k) = P(X_2 - X_1 = k) = \frac{1}{\sqrt{2w}} \left\{ \varphi(z) + \sum_{r=1}^{u} F_{2r}[-\varphi(z)] \right\} + o\left(\frac{1}{w^{u+1}}\right). \tag{11}$$

Let us observe that the exact value of $P_w(k)$ is given by the formula

$$P_w(k) = e^{-2w} \sum_{s=0}^{\infty} \frac{w^{2s+k}}{s!\,(s+k)!} = e^{-2w} I_k(2w), \tag{12}$$

where $I_k(2w)$ is a Bessel function with purely imaginary argument. From (11) and (12) we obtain the following asymptotical formula:

$$I_k(2w) = e^{2w} \left[\frac{1}{\sqrt{2w}} \left\{ \varphi(z) + \sum_{r=1}^{u} F_{2r}[-\varphi(z)] \right\} + o\left(\frac{1}{w^{u+1}}\right) \right]. \tag{13}$$

Let us observe that the formula (13) should be used for large k and large w, namely when k is of order $\sqrt{2w}$.

We shall finally give two short expansions of the formulae (7) and (12). We find namely

$$F_1 = \frac{\mu_3}{3!} [-\varphi(z)]^{(3)} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{w_2 - w_1}{(w_1 + w_2)^{\frac{3}{2}}} H_3(z),$$

$$F_2 = \frac{\mu_4}{4!} [-\varphi(z)]^{(4)} + \frac{1}{2!} \frac{\mu_3^2}{(3!)^2} [-\varphi(z)]^{(6)} =$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \left[\frac{-H_4(z)}{24(w_1 + w_2)} - \frac{(w_2 - w_1)^2 H_6(z)}{72(w_1 + w_2)^3} \right],$$

where $H_j(z)$ is the Hermite polynomial of order j. We obtain then

$$P_{w_1 w_2}(k) = \frac{1}{\sqrt{2\pi(w_1+w_2)}} \exp\left(-\frac{z^2}{2}\right)\left\{1+\frac{1}{6\sqrt{w_1+w_2}}\frac{w_2-w_1}{w_1+w_2}(z^3-3z)+\right.$$

$$\left.+\frac{1}{24(w_1+w_2)}\left[-z^4+6z^2+3+\frac{1}{3}\left(\frac{w_2-w_1}{w_1+w_2}\right)^2(-z^6+15z^4-45z^2+15)\right]\right\}+$$

$$+o\left[\frac{1}{(w_1+w_2)^{\frac{3}{2}}}\right],$$

$$P_w(k) = \frac{1}{2\sqrt{\pi w}} \exp\left(-\frac{z^2}{2}\right)\left[1+\frac{1}{48w}(-z^4+6z^2+3)\right]+o\left(\frac{1}{w^2}\right).$$

I express my gratitude to A. RÉNYI for his valuable suggestions.

(Received 7 December 1954)

References

[1] J. O. IRWIN, The frequency distribution of the difference between two independent variates following the same Poisson distribution, *Journal of the Royal Statistical Society*, (1937).

[2] M. FISZ, Rozkład graniczny zmiennej losowej będącej różnicą dwóch niezależnych zmiennych losowych o rozkładzie Poissona, *Zastosowania Matematyki*, **1** (1953), pp. 41 – 45.

[3] M. FISZ, The limiting distribution of some function of two independent random variables and its statistical application, *Colloquium Mathematicum*, **3** (1955), pp. 138—146.

[4] C. G. ESSEEN, Fourier analysis of distribution functions. A mathematical study of the Laplace—Gaussian law, *Acta Math.*, **77** (1945), pp. 1—125.

УТОЧНЕНИЕ ОДНОЙ ВЕРОЯТНОСТНОЙ ПРЕДЕЛЬНОЙ ТЕОРЕМЫ И ЕЕ ПРИМЕНЕНИЕ К БЕССЕЛЕВЫМ ФУНКЦИЯМ

М. Фиш (Варшава)

(Резюме)

Применяя общую теорему Эссеена (см [4]), автор дает асимптотическое разложение (7) для распределения вероятностей разности двух независимых пуассоновских случайных величин. С помощью этого разложения получается асимптотическая формула (13) для бесселевой функции порядка k с чисто мнимым аргументом.

BEMERKUNGEN ZUR APPROXIMATION EINER REELLEN ZAHL DURCH BRÜCHE

Von
P. SZÜSZ (Budapest)
(*Vorgelegt von* P. TURÁN)

Es sei α eine Irrationalzahl. Zu jeder α gibt es bekanntlich unendlich viele Zahlenpaare p, q, für die

$$|q\alpha - p| < \frac{1}{q}, \tag{1}$$

und sogar

$$|q\alpha - p| < \frac{1}{\sqrt{5}q} \tag{1'}$$

gilt; der Faktor $5^{-\frac{1}{2}}$ auf der rechten Seite von (1') läßt sich durch keinen kleineren ersetzen.

Während der Abfassung eines Lehrbuches hat Prof. TURÁN folgende Frage gestellt: Es sei δ eine reelle Zahl mit

$$0 \leqq \delta \leqq 1$$

und man betrachte die Zahlenpaare p, q für die

$$|q\alpha - p| < q^{\delta - 1} \tag{2}$$

gilt. Was läßt sich dann über die „Dichtigkeit" der q aussagen? Wenn also $M(\alpha, \delta, n)$ die Anzahl der $q \leqq n$ bedeutet, für die (2) gültig ist, so soll das Verhalten des Quotienten $\frac{M(\alpha, \delta, n)}{n}$ untersucht werden. Es soll ferner untersucht werden, was sich über die „Lokalisation" dieser q aussprechen läßt; es soll also für jedes n eine möglichst kleine Zahl $m = m(n)$ derart angegeben werden, daß zwischen n und $n+m$ mindestens ein (2) erfüllendes q existiert.

Prof. TURÁN hat bewiesen, daß stets

$$M(\alpha, \delta, n) > n^{\delta} \qquad (\delta > 0) \tag{3}$$

gilt, ferner, daß ebenfalls für $\delta > 0$, zwischen n und $n^{\frac{1}{\delta}}$ mindestens ein (2) erfüllendes q liegt.

Prof. TURÁN hat seine obigen Resultate P. ERDŐS und V. JARNIK brieflich mitgeteilt. In seinem Antwortsbriefe hat P. ERDŐS folgendes bewiesen: Es gilt

$$M(\alpha, 0, n) > K \log n \qquad (K \text{ konstant}), \tag{4}$$

sogar etwas schärfer: bezeichnet $M_c(\alpha, 0, n)$ die Anzahl der natürlichen $q \leqq n$ mit $|q\alpha - p| < cq^{-1}$, so gilt für $c \leqq 5^{-\frac{1}{2}}$ stets

$$M_c(\alpha, 0, n) > K_c \log n, \tag{4'}$$

wobei zu beachten ist, daß K_c nicht einmal von α abhängt.

V. JARNIK hat für die Sätze von TURÁN und auch für (4') einen Beweis gefunden und die Frage aufgeworfen, ob sich (3) bzw. (4) nicht verbessern lassen.

In der vorliegenden Note wiedergebe ich mit der freundlichen Erlaubnis von Prof. TURÁN seinen Beweis für (3), zeige, daß (3) im allgemeinen nicht verbessert werden kann und gebe die möglichst beste Lokalisation der (2) erfüllenden q an. Daß (4) scharf ist, kann z. B. bei $\alpha = \frac{\sqrt{5}-1}{2}$ durch eine leichte Rechnung bestätigt werden. Ich beweise die folgenden Sätze:

SATZ 1.[1] *Für $\delta > 0$ gilt* (3); *andererseits gibt es ein α, für welches mit unendlich vielen n*

$$M(\alpha, \delta, n) < K(\delta) n^{\delta} \tag{5}$$

gilt, wobei $K(\delta)$ nur von δ abhängt.

SATZ 2. *Es sei $\varepsilon > 0$ und $0 \leqq \delta \leqq 1$ gegeben. Dann gibt es ein nur von ε und δ abhängiges $N_0 = N_0(\varepsilon, \delta)$ so, daß zu einem reellen α zwischen N und $\varepsilon N^{1+\delta}$ stets ein n mit*

$$|n\alpha - m| < n^{-\delta} \qquad (m \text{ eine geeignete ganze Zahl})$$

liegt, falls nur $N > N_0$.

SATZ 3. *Es sei $g(x)$ eine positive, für $x \to \infty$ monoton gegen Null strebende Funktion. Dann existiert eine Irrationalzahl α und unendlich viele $N_r (r = 1, 2, \ldots)$ für die zwischen*

$$N_r \quad \text{und} \quad g(N_r) N^{1+\delta}$$

kein n existiert, für welches mit irgendeinem ganzen n

$$|n\alpha - m| < n^{-\delta} \tag{2'}$$

gilt.

Dieser Satz besagt, daß Satz 2 nicht mehr verschärft werden kann.

Bevor ich zum Beweis der ausgesprochenen Tatsachen übergehe, möchte ich bemerken, daß nichts ausgesprochen werden kann, wenn man außer (2)

[1] Wie mir Prof. TURÁN mitteilte, hat diese Ungleichung P. ERDŐS auch gefunden.

bzw. $|n\alpha - m| < n^{-\delta}$ noch $(p, q) = 1$ bzw. $(n, m) = 1$ verlangt. Dies folgt leicht aus den elementaren Tatsachen der Kettenbruchlehre.[2]

Nun beweise ich die ausgesprochenen Sätze.

BEWEIS DES SATZES 1.

Den folgenden Beweis von (3) hat mir Prof. TURÁN mitgeteilt.

Man setze $k = \left[n^{\frac{1}{1+\delta}}\right]$, $r = \left[n^{\frac{\delta}{1+\delta}}\right]$ und man betrachte die Zahlen

$$m_\varkappa = \varkappa\alpha - [\varkappa\alpha] \qquad (\varkappa = 0, 1, \ldots, rk).$$

Nach dem Schubfachprinzip gibt es ein Intervall

$$I_l = \left(\frac{l}{k}, \frac{l+1}{k}\right)$$

derart, daß in I_l mindestens $r+1$ Zahlen $m_{\varkappa_\varrho}$ $(\varrho = 1, 2, \ldots, r+1)$ liegen; es gilt also

$$|m_{\varkappa_\varrho} - m_{\varkappa_1}| = |y\alpha - x| < \frac{1}{k} \qquad (\varrho = 2, 3, \ldots, r+1), \tag{6}$$

wobei y eine natürliche Zahl $\leqq kr$ bedeutet, während x ganz ist. Aus (6) und aus der Definition von k und r folgt, daß für mindestens $\left[n^{\frac{\delta}{1+\delta}}\right]$ Zahlen unter kr, also erst recht unter n die Relation

$$|y\alpha - x| < \frac{1}{k} \leqq \left[y^{\frac{1}{1+\delta}}\right]^{-1} \tag{7}$$

gilt. Bestimmt man δ' aus

$$\frac{1}{1+\delta} = 1 - \delta',$$

so erhält man (3) mit δ' statt δ, was zu beweisen war.

Nun beweise ich (5). Es bezeichnen (hier und im Folgenden) $a_0, a_1, \ldots$ die Kettenbruchnenner von α, $\frac{A_0}{B_0}, \frac{A_1}{B_1}, \ldots$ die Näherungsbrüche von α. Ferner bezeichne $N(\alpha, I, n)$ die Anzahl der $\nu \leqq n$ mit $(\nu\alpha) \in I$, wobei I ein gegebenes Teilintervall von (0,1) bezeichnet, während ( ) den Bruchteil der in der Klammer stehenden Zahl bedeutet. Im Folgenden bezeichne I_n der Kürze halber die Vereinigung der Intervalle $(0, n^{\delta-1})$ und $(1-n^{\delta-1}, 1)$. Dies wird, da wir δ $(\delta > 0)$ festgelegt haben, kein Mißverständnis verursachen. Dann gilt offenbar

$$M(\alpha, \delta, n) < \sum_{\nu=1}^{r+1} \{N(\alpha, I_{B_{\nu-1}}, B_\nu) - N(\alpha, I_{B_{\nu-1}}, B_{\nu-1})\}, \tag{8}$$

[2] Prof. TURÁN hat die entsprechenden Fragen auch im Falle simultaner Approximation mehrerer reellen Zahlen aufgeworfen. Die entsprechenden Resultate haben auch Anwendungen. Auf diese Fragen beabsichtige ich in einer späteren Note zurückzukommen.

wobei r denjenigen Index bedeutet, für welchen $B_r \leqq n < B_{r+1}$ gilt. Die Rolle der B_ν kann in (8) jede streng zunehmende Folge ganzer Zahlen übernehmen. Daß wir eben die Folge der Näherungsnenner von α gewählt haben, wird später ausgenutzt werden, vgl. Formel (9).

Aus

$$|l\alpha - k| < l^{\delta-1} \qquad (l \text{ natürlich, } k \text{ ganz})$$

und

$$B_{\nu-1} \leqq l < B_\nu$$

folgt nämlich jedenfalls

$$(l\alpha) \in I_{B_{\nu-1}}.$$

Man setze insbesondere $a_0 = a_1 = \cdots = 1$ $\left(\text{also } \alpha = \dfrac{1+\sqrt{5}}{2}\right)$, und $n = B_r$.

Dann folgt aus (8)

$$M(\alpha, \delta, B_r) < \sum_{\nu=1}^{r+1} \{N(\alpha, I_{B_{\nu-1}}, B_\nu) - N(\alpha, I_{B_{\nu-1}}, B_{\nu-1})\}. \tag{8'}$$

Nach einem Satz von BEHNKE[3] gilt

$$\begin{cases} N(\alpha, I_{B_{\nu-1}}, B_\nu) = 2B_\nu B_{\nu-1}^{\delta-1} + O(1), \\ N(\alpha, I_{B_{\nu-1}}, B_{\nu-1}) = 2B_{\nu-1}^{\delta} + O(1). \end{cases} \tag{9}$$

Daher ist die rechte Seite von (8') gleich

$$2\sum_{\nu=1}^{r+1} B_{\nu-1}^{\delta-1}(B_\nu - B_{\nu-1}) + O(r) = 2\sum_{\nu=1}^{r+1} B_{\nu-1}^{\delta}\left(\frac{B_\nu}{B_{\nu-1}} - 1\right) + O(r),$$

also $\left(\text{wegen } \alpha = \dfrac{\sqrt{5}+1}{2},\ \dfrac{B_\nu}{B_{\nu-1}} < 2\right)$ kleiner als

$$2\sum_{\nu=1}^{r+1} B_{\nu-1}^{\delta} + O(\log B_{r+1}).$$

Da (für jedes α) bekanntlich

$$B_{r+1} > 2B_{r-1} \qquad (r = 1, 2, \ldots)$$

gilt, folgt hieraus

$$M(\alpha, \delta, B_r) < 4B_r^{\delta}(1 + 2^{-\delta} + 2^{-2\delta} + \cdots) + O(\log B_{r+1}) = \frac{4B_r^{\delta}}{1 - 2^{-\delta}} + O(\log B_r),$$

womit (5) bewiesen ist. Die Frage nach dem kleinsten Wert $K_0(\delta)$ von $K(\delta)$, d. h. nach derjenigen Zahl, für die

$$M(\alpha, \delta, n) > K_0(\delta) n^{\delta},$$

[3] H. BEHNKE, *Hamburger Abhandlungen*, **3** (1924), S. 261—318.

aber bei beliebigem $\varepsilon > 0$ für geeignetes α und unendlich viele n

$$M(\alpha, \delta, n) < (K_0(\delta) + \varepsilon) n^{\delta}$$

gilt, bleibt offen. Sowohl der Turánsche Beweis von (3), als mein Ansatz, der zu (5) führte, scheint in dieser Hinsicht zu roh zu sein.

BEWEIS DES SATZES 2. $a_0, a_1, \ldots$ bzw. $\frac{A_0}{B_0}, \frac{A_1}{B_1}, \ldots$ habe dieselbe Bedeutung, wie bei dem Beweis von (5). Bekanntlich gilt für jedes natürliche r

$$|B_r \alpha - A_r| < \frac{1}{B_{r+1}}, \tag{10}$$

also für jedes natürliche c

$$|cB_r \alpha - cA_r| < \frac{c}{B_{r+1}}. \tag{10'}$$

Für $c \leqq \left(\frac{B_{r+1}}{B_r^{\delta}}\right)^{\frac{1}{1+\delta}}$ gilt $\frac{c}{B_{r+1}} \leqq \frac{1}{(cB_r)^{\delta}}$, wegen (10') ist also für natürliches

$$c \leqq \left(\frac{B_{r+1}}{B_r^{\delta}}\right)^{\frac{1}{1+\delta}}$$

erst recht

$$|cB_r \alpha - cA_r| < (cB_r)^{-\delta}. \tag{11}$$

Es sei nun n gegeben, und es bezeichne $r = r(n)$ denjenigen Index, für welchen

$$B_r \leqq n < B_{r+1} \tag{12}$$

gilt.

Es darf ohne Beschränkung der Allgemeinheit

$$B_{r+1} > 4B_r, \tag{13}$$

also $a_{r+1} \geqq 4$ angenommen werden; sonst wäre nämlich B_{r+1} eine Zahl, die (2) genügt und für die $n < B_{r+1} \leqq 4n$.

Wir unterscheiden nun zwei Fälle.

1) Es sei zuerst

$$n \leqq (B_r B_{r+1})^{\frac{1}{1+\delta}} - B_r. \tag{14}$$

Für jedes natürliche $c \leqq \left(\frac{B_{r+1}}{B_r^{\delta}}\right)^{\frac{1}{1+\delta}}$ gilt nach (11)

$$|cB_r \alpha - cA_r| < (cB_r)^{-\delta}. \tag{15}$$

Je zwei konsekutive der Zahlen cB_r $\left(c=1,2,\ldots,\left[\left(\frac{B_{r+1}}{B_r^{\delta}}\right)^{\frac{1}{1+\delta}}\right]\right)$ haben voneinander den Abstand B_r; wegen (14) ist jedenfalls

$$n \leqq B_r\left[\left(\frac{B_{r+1}}{B_r^{\delta}}\right)^{\frac{1}{1+\delta}}\right].$$

Daher gibt es ein cB_r mit (15), für welches $0 \leqq cB_r - n < B_r$, womit unser Satz 2 im Falle 1) bewiesen ist.

2) Es sei nun

$$n > (B_r B_{r+1})^{\frac{1}{1+\delta}} - B_r, \tag{16}$$

also wegen (13) und $\delta \leqq 1$ gewiß

$$n > \left(\frac{B_{r+1}B_r}{4}\right)^{\frac{1}{1+\delta}}. \tag{16'}$$

Dann folgt aus (12) und (16')

$$n < B_{r+1} < \frac{4}{B_r} n^{1+\delta};$$

da B_{r+1}, als Näherungsnenner von α der Relation $|B_{r+1}\alpha - A_{r+1}| < B_{r+1}^{-1}$ genügt und $B_r \to \infty$ für $r \to \infty$, ist unser Satz 2 auch im Falle 2) bewiesen.

Beweis des Satzes 3. Es bezeichnen $a_0, a_1, \ldots$ die Kettenbruchnenner, $\frac{A_0}{B_0}, \frac{A_1}{B_1}, \ldots$ die Näherungsbrüche von α.

Bekanntlich[4] gilt

$$|y\alpha - x| \geqq |B_{r-1}\alpha - A_{r-1}|, \tag{17}$$

falls y natürlich, x ganz ist und $y < B_r$ besteht.

Es sei n gegeben; r bedeutet denjenigen Index, für den (12) gilt. Dann läßt n eine Darstellung von der Form

$$n = qB_r + s \qquad (1 \leqq q \leqq a_{r+1},\ 0 \leqq s < B_r) \tag{18}$$

zu. Ist dann m eine beliebige ganze Zahl, so gilt

$$|n\alpha - m| = |qB_r\alpha + s\alpha - m| = |q(B_r\alpha - A_r) + s\alpha + qA_r - m|.$$

Ist insbesondere $s \geqq 1$ und $q \leqq \frac{a_{r+1}}{2}$, so folgt aus der bekannten Formel[5]

$$\left|\frac{B_r\alpha - A_r}{B_{r-1}\alpha - A_{r-1}}\right| = \frac{1}{\zeta_{r+1}} < \frac{1}{a_{r+1}}$$

[4] Vgl. z. B. Perron, *Die Lehre von den Kettenbrüchen*, II. Aufl. (Berlin—Leipzig, 1929), S. 52.

[5] Vgl. Perron, loc. cit., S. 42, Formel (4).

(hier bedeuten $\zeta_0, \zeta_1, \ldots$ die vollständigen Quotienten der Kettenbruchentwicklung von α)

$$(19)\quad \begin{cases} |n\alpha - m| \geqq |s\alpha - m + qA_r| - q|B_r\alpha - A_r| \geqq |B_{r-1}\alpha - A_{r-1}| - q|B_r\alpha - A_r| > \\ > |B_{r-1}\alpha - A_{r-1}| - \dfrac{q}{a_{r+1}}|B_{r-1}\alpha - A_{r-1}| = \\ = \left(1 - \dfrac{q}{a_{r+1}}\right)|B_{r-1}\alpha - A_{r-1}| \geqq \dfrac{1}{2}|B_{r-1}\alpha - A_{r-1}|. \end{cases}$$

Alles bisher bewiesene gilt allgemein. Wir versuchen nun, die Kettenbruchnenner $a_0, a_1, \ldots$ von α geeignet zu wählen.

Ich nehme an, daß die a_r $(r = 0, 1, \ldots)$ so stark anwachsen, daß stets

$$(20)\qquad B_{r+1} > 4^{\frac{1+\delta}{\delta}} B_r^{\frac{1}{\delta}}$$

gilt. Dann behaupte ich, daß für $r = 1, 2, \ldots$ zwischen

$$2(B_r B_{r+1})^{\frac{1}{1+\delta}} \quad \text{und} \quad \frac{B_{r+1}}{2}$$

keine solche natürliche Zahl n liegt, für welche

$$|n\alpha - m| < n^{-\delta} \qquad (m \text{ ganz})$$

gilt.

Man betrachte nämlich für $2(B_r B_{r+2})^{\frac{1}{1+\delta}} < n < \dfrac{B_{r+1}}{2}$ erfüllendes n und ganzes m den Ausdruck

$$(21)\qquad |n\alpha - m|.$$

Wir stellen n in der Form (18) dar; es ist also eine untere Schranke für den Ausdruck

$$(21')\qquad |qB_r\alpha + s\alpha - m|$$

anzugeben, wobei m sämtliche ganze Zahlen durchläuft, während

$$(22)\qquad 2(B_r B_{r+1})^{\frac{1}{1+\delta}} < qB_r + s < \frac{B_{r+1}}{2}$$

besteht. Wir unterscheiden nun zwei Fälle, je nach dem $s = 0$ oder $s \geqq 1$ gilt.

1) Es sei zuerst $s = 0$.

Dann ist wegen (22)

$$(23)\qquad 2\left(\frac{B_{r+1}}{B_r^{\delta}}\right)^{\frac{1}{1+\delta}} < q < \frac{B_{r+1}}{2B_r} < \frac{a_{r+1}}{2} + \frac{1}{2}.$$

Ich zeige, daß für $q \leqq \frac{a_{r+1}}{2}$

$$|qB_r\alpha - m| > \frac{2}{n^\delta}$$

gilt. Dies kann folgendermaßen geschehen: Ist $|qB_r\alpha - m| < 1$, so ist jedenfalls entweder

$$m = [qB_r\alpha]$$

oder

$$m = [qB_r\alpha] + 1.$$

Daher gilt

$$|qB_r\alpha - m| > \min\{(qB_r\alpha), 1-(qB_r\alpha)\}.$$

Nun ist aber für $q \leqq \frac{a_{r+1}}{2}$

$$\min\{(qB_r\alpha), 1-(qB_r\alpha)\} = q|B_r\alpha - A_r|.$$

Daher ist wegen der bekannten Relation $|B_r\alpha - A_r| > \frac{1}{2}\frac{1}{B_{r+1}}$ und wegen (23)

$$|qB_r\alpha - m| \geqq q|B_r\alpha - A_r| > \frac{1}{2}\frac{q}{B_{r+1}} \geqq \left(\frac{B_{r+1}}{B_r^\delta}\right)^{\frac{1}{1+\delta}} \frac{1}{B_{r+1}} = (B_rB_{r+1})^{-\frac{\delta}{1+\delta}} > \frac{2}{n^\delta},$$

was zu beweisen war.

2) Es sei jetzt in (21′)

$$1 \leqq s < B_r.$$

Die zweite Ungleichung der Formel (23) gilt auch in diesem Falle. Daher folgt aus (19) und (20)

$$|n\alpha - m| > \frac{1}{2}|B_{r-1}\alpha - A_{r-1}| > \frac{1}{4B_r} > (B_rB_{r+1})^{-\frac{\delta}{1+\delta}} > \frac{2}{n^\delta}, \tag{24}$$

womit wir unsere Behauptung auch im Falle 2) bewiesen haben.

Wir haben also gezeigt, daß falls die Kettenbruchnenner $a_0, a_1, \ldots$ von α genügend stark anwachsen, zwischen

$$2(B_rB_{r+1})^{\frac{1}{1+\delta}}$$

und

$$\frac{1}{2}B_{r+1}$$

keine Zahl n liegt, die mit irgendeinem ganzen m

$$|n\alpha - m| < n^{-\delta} \tag{25}$$

leistet.

Setzt man $2(B_r B_{r+1})^{\frac{1}{1+\delta}} = N_r$, so besagt unser Satz, daß zwischen

$$N_r \quad \text{und} \quad \frac{N_r^{1+\delta}}{2^{2+\delta} B_r} \tag{26}$$

kein n mit (25) liegt.

Schreibt man

$$f(x) = \frac{1}{B_r} \quad \text{für} \quad B_r \leqq x < B_{r+1}$$

und ist $g(x)$ eine positive, monoton gegen Null strebende, aber sonst beliebige gegebene Funktion, so kann durch genügend starkes Anwachsen der a_r offenbar erreicht werden, daß für unendlich viele r

$$f(N_r) > g(N_r)$$

gilt. Hieraus und aus (26) folgt, daß für unendlich viele r zwischen

$$N_r \quad \text{und} \quad g(N_r) N_r^{1+\delta}$$

kein n mit (25) liegt. Damit haben wir auch Satz 3 bewiesen.

Zum Schluß spreche ich Herrn Professor TURÁN meinen Dank aus dafür, daß er mich auf diese Probleme aufmerksam gemacht hat.

BUDAPEST, INSTITUT FÜR ANGEWANDTE MATHEMATIK
DER UNGARISCHEN AKADEMIE DER WISSENSCHAFTEN

(Eingegangen am 29. Januar 1955.)

ЗАМЕЧАНИЯ К АППРОКСИМАЦИИ ВЕЩЕСТВЕННЫХ ЧИСЕЛ ДРОБЯМИ

П. Сюс (Будапешт)

(Резюме)

Пусть α — вещественное число, а δ вещественное число, удовлетворяющее условию $0 \leqq \delta < 1$. Рассмотрим натуральные числа q, для которых при некотором целом p имеет место

$$|q\alpha - p| < q^{-\delta}. \tag{1}$$

Обозначим через $M(\alpha, \delta, n)$ число тех значений q, не превышающих n, для которых выполняется (1). Тогда справедлива следующая

Теорема 1. Для всякого вещественного α имеет место

$$M(\alpha, \delta, n) > n^{1-\delta}, \tag{2}$$

и, с другой стороны, существует значение α, для которого

$$M(\alpha, \delta, n) < K(\delta) n^{1-\delta}, \tag{3}$$

где число $K(\delta)$ зависит только от δ. Его наименьшее значение неизвестно.

О „локализации“ натуральных чисел q, удовлетворяющих (1), можно сказать следующее:

Теорема 2. Для данного $\varepsilon > 0$ существует такое число $N = N(\varepsilon)$, что между

$$n \quad \text{и} \quad \varepsilon n^{1+\delta}$$

попадает хотя бы одно значение q, удовлетворяющее (1), если только $n > N$.

Теорема 3. Пусть $f(x)$ — положительная функция, стремящаяся к нулю для $x \to \infty$. Тогда существует такое число α, что для бесконечного множества значений n между

$$n \quad \text{и} \quad f(n) n^{1+\delta}$$

нет ни одного значения q, удовлетворяющего (1), так что теорема 2 не допускает улучшения.

Приведенное доказательство от (2) принадлежит профессору П. Туран. Все другие утверждения доказываются с помощью теории цепных дробей.

EXTENSION OF FUNCTIONS SATISFYING A LIPSCHITZ CONDITION

By

J. CZIPSZER (Budapest) and L. GEHÉR (Miskolc)

(*Presented by* G. ALEXITS)

It is a simple fact that a function $f(x)$ defined on a set of the real line and satisfying a Lipschitz condition

$$|f(x)-f(y)| \leqq K|x-y|$$

may be extended to the whole real line so that the Lipschitz condition with the same constant K remains valid for the extended function. The present paper deals with the generalizations of this theorem for functions defined on a subset of a given metric space.

§ 1

In this section we shall prove the following theorem:

THEOREM I. *Let X be a metric space. Let $f(x)$ be a real function defined on a subset S of X and satisfying a Lipschitz condition of the form*

$$|f(x)-f(y)| \leqq K\overline{xy}$$

($K \geqq 0$, $\overline{xy}$ denotes the distance of x and y), i. e.

$$f(x) \in \mathrm{Lip}_S(1, K).$$

Then there is a real function $F(x)$ defined on X, equal to $f(x)$ on S and satisfying the same Lipschitz condition, i. e.

$$F(x) \in \mathrm{Lip}_X(1, K).$$

If X is the n-dimensional Euclidean space R_n, then this theorem is contained in a theorem of M. D. KIRSZBRAUN,[1] which asserts that if $f(x)$ is

[1] M. D. KIRSZBRAUN, Über die zusammenziehende und Lipschitzsche Transformation, *Fundamenta Math.*, **22** (1934), pp. 77—108 (esp. p. 94, Hauptsatz I).

defined in a subspace of R_n, takes values in R_n and satisfies the Lipschitz condition

$$\overline{f(x)f(y)} \leqq K\overline{xy},$$

then it may be extended to the whole space R_n preserving the same Lipschitz condition.

From the other hand, Theorem I guarantees the possibility to extend every function, defined on a subset of a metric space, ranging over R_n (or more generally over any n-dimensional Banach space) and satisfying a Lipschitz condition with constant K, to a function which is defined throughout the space and satisfies a Lipschitz condition with constant $\sqrt{n}K$ (or with some other constant).

We shall show that the function

$$F(x) = \inf_{y \in S} [f(y) + K\overline{xy}]$$

is the desired extension of $f(x)$.

Before starting the proof, we should like to show how one can hit upon this formula. If x is any point outside S, then $f(x)$ must be defined so that the inequalities

$$(1) \qquad f(y) - K\overline{xy} \leqq f(x) \leqq f(y) + K\overline{xy}$$

shall hold for every $y \in S$. Now for any two points y, z of S

$$f(y) - f(z) \leqq K\overline{yz} \leqq K\overline{xy} + K\overline{xz}$$

and by this

$$\sup_{y \in S} [f(y) - K\overline{xy}] \leqq \inf_{z \in S} [f(z) + K\overline{xz}].$$

This inequality shows that with suitable choice of $f(x)$ we can satisfy (1): with

$$(2) \qquad \sup_{y \in S} [f(y) - K\overline{xy}] \leqq f(x) \leqq \inf_{y \in S} [f(y) + K\overline{xy}]$$

this is already attained. Now repeating this procedure we can extend $f(x)$ by every step to one new point, and with the help of the well-ordering theorem we can define $f(x)$ at last everywhere. This procedure is nearly the same as the one used for the proof of the theorem of HAHN—BANACH.[2] But fortunately this successive construction can be replaced by a much more simple construction: if we define $f(x)$ for every x to be equal to the left or right end-point of the interval (2), we get a function which satisfies the prescribed Lipschitz condition not only "between" two points one being in S and the other outside S, but "between" any two points of the space.

PROOF OF THEOREM I. A) $F(x)$ is everywhere defined in X. For a fixed $z \in X$ we have the inequality

$$f(y) + K\overline{xy} = f(z) + K\overline{xy} + [f(y) - f(z)] \geqq f(z) + K\overline{xy} - K\overline{yz} \geqq f(z) - K\overline{xz}$$

which shows that the set of values

$$f(y) + K\overline{xy} \quad y \in S$$

for any fixed x is bounded from below and that gives the assertion.

[2] S. BANACH, *Théorie des opérations linéaires* (Warszawa, 1932), Théorème I, Chap II and Théorème 2, Chap. IV.

B) If $x \in S$ and $y \in S$, then

$$f(x) + K\overline{xx} = f(x)$$

and

$$f(y) + K\overline{xy} \geqq f(x).$$

Consequently,

$$F(x) = f(x) \quad \text{for} \quad x \in S$$

and so $F(x)$ is in fact an extension of $f(x)$.

C) We shall now prove that $F(x) \in \mathrm{Lip}_X(1, K)$. Let x_1 and x_2 be two points of X. For a given $\varepsilon > 0$ we may choose a $y \in S$ so that

$$F(x_1) \geqq f(y) + K\overline{x_1 y} - \varepsilon$$

holds. Then we have

$$F(x_2) - F(x_1) \leqq f(y) + K\overline{x_2 y} - [f(y) + K\overline{x_1 y} - \varepsilon] =$$
$$= K(\overline{x_2 y} - \overline{x_1 y}) + \varepsilon \leqq K\overline{x_1 x_2} + \varepsilon.$$

We get similarly

$$F(x_1) - F(x_2) \leqq K\overline{x_1 x_2} + \varepsilon.$$

Hence

$$|F(x_1) - F(x_2)| \leqq K\overline{x_1 x_2} + \varepsilon.$$

Since ε is arbitrary, we see that indeed

$$F(x) \in \mathrm{Lip}_X(1, K).$$

Thus the proof of Theorem I is complete.

§ 2

We shall now give a sort of generalization of Theorem I.

THEOREM II. *Let X be a metric space. Let $f(x)$ be a real function defined on a closed subset S of X which satisfies in every point x of S a Lipschitz condition of the form*

$$|f(x) - f(y)| \leqq K\overline{xy} \quad \text{for} \quad \overline{xy} < \delta \qquad (K \geqq 0, \delta > 0),$$

where K and δ depend on x. (We shall denote this property by saying that $f(x)$ belongs to LIP_S *or briefly to* LIP*). Then there is a real function $F(x)$ defined on X, equal to $f(x)$ on S and which belongs to* LIP_X.

PROOF. For the time being let us suppose that $f(x)$ is bounded

$$|f(x)| \leqq M \quad \text{for} \quad x \in S.$$

It is obvious that by a suitable change of K the Lipschitz condition

$$|f(x) - f(y)| \leqq K(x)\overline{xy}$$

shall hold for every y. The notation $K(x)$ sets in evidence that K depends

on x. We shall show that the function

$$F(x) = \inf_{y \in S} [f(y) + K(y)\overline{xy}]$$

is the desired extension of $f(x)$.

It is obvious that — since $f(x)$ is bounded — $F(x)$ is defined everywhere.

In the same way as in the previous §, we see that

$$F(x) = f(x) \quad \text{if} \quad x \in S.$$

We have still to prove that $F(x)$ satisfies a Lipschitz condition in every point x.

We choose a point x_0. First we suppose that x_0 belongs to S. Then we have

$$F(x) \leqq f(x_0) + K(x_0)\overline{xx_0}$$

or

$$F(x) - f(x_0) = F(x) - F(x_0) \leqq K(x_0)\overline{xx_0}.$$

If y belongs to S and $y \neq x_0$, then

$$f(y) - f(x_0) \geqq -K(y)\overline{x_0 y}$$

and

$$f(y) - f(x_0) \geqq -K(x_0)\overline{x_0 y}.$$

Multiplying the first inequality with $\dfrac{\overline{xy}}{\overline{xy} + \overline{xx_0}}$, the second with $\dfrac{\overline{xx_0}}{\overline{xy} + \overline{xx_0}}$ we get

$$f(y) - f(x_0) \geqq -[K(y)\overline{xy} + K(x_0)\overline{xx_0}] \frac{\overline{x_0 y}}{\overline{xy} + \overline{xx_0}} \geqq -[K(y)\overline{xy} + K(x_0)\overline{xx_0}].$$

Hence

$$f(y) + K(y)\overline{xy} \geqq f(x_0) - K(x_0)\overline{xx_0},$$

and this last inequality is true also in the case $y = x_0$.

We see at once that

$$F(x) = \inf_{y \in S} [f(y) + K(y)\overline{xy}] \geqq f(x_0) - K(x_0)\overline{xx_0}$$

or

$$F(x) - f(x_0) \geqq -K(x_0)\overline{xx_0}.$$

So the inequality

$$|F(x) - f(x_0)| = |F(x) - F(x_0)| \leqq K(x_0)\overline{xx_0}$$

holds, i. e. $F(x)$ satisfies a Lipschitz condition in every point of S.

Now let us suppose that $x_0 \in X - S$. We denote by $d = d(x_0, S)$ the (positive) distance of x_0 from S and consider the open neighbourhood U of x_0 with center x_0 and radius $\dfrac{1}{2}d$.

For $x \in U$ and for any fixed $z \in S$

$$F(x) \leqq f(z) + K(z)\overline{xz} \leqq f(z) + K(z)\, d(z, U).$$

From the other side

$$F(x) \geqq \inf_{y \in S} f(y) \geqq -M.$$

The last inequalities show that $|F(x)|$ is bounded in U by some N:

$$|F(x)| \leqq N \qquad (x \in U).$$

Now let us choose a point x_1 from U. For a given $\varepsilon > 0$ we can find a $y \in S$ so that the inequality

(1) $$F(x_1) \geqq f(y) + K(y)\overline{x_1 y} - \varepsilon$$

shall hold. If we subtract this inequality from

$$F(x_0) \leqq f(y) + K(y)\overline{x_0 y},$$

we get

(2) $$F(x_0) - F(x_1) \leqq K(y)(\overline{x_0 y} - \overline{x_1 y}) + \varepsilon \leqq K(y)\overline{x_0 x_1} + \varepsilon.$$

From (1) it follows

$$K(y) \leqq \frac{F(x_1) - f(y) + \varepsilon}{\overline{x_1 y}} \leqq \frac{N + M + \varepsilon}{\frac{d}{2}}.$$

Hence and by (2)

$$F(x_0) - F(x_1) \leqq \frac{N + M + \varepsilon}{\frac{d}{2}} \overline{x_0 x_1} + \varepsilon,$$

and since ε is arbitrary, we have

$$F(x_0) - F(x_1) \leqq 2\frac{N + M}{d} \overline{x_0 x_1}.$$

The analogous inequality

$$F(x_1) - F(x_0) \leqq 2\frac{N + M}{d} \overline{x_0 x_1}$$

may be deduced in the same way, so that together we have

$$|F(x_1) - F(x_0)| \leqq 2\frac{N + M}{d(x_0, S)} \overline{x_0 x_1}$$

for

$$\overline{x_0 x_1} < \frac{1}{2} d(x_0, S),$$

i. e. $F(x)$ satisfies a Lipschitz condition in every point of $X - S$.

Thus Theorem II is proved under the hypothesis of the boundedness of $f(x)$.

In the case the function which is to be extended is not necessarily bounded we shall need the following remark:

If $|f(x)| < M$ on S, then the extension of $f(x)$ may be chosen so that this inequality will hold throughout X. We can assume that the Lipschitz constants $K(x)$ are all larger then 1. Then for $x \in X - S$

$$F(x) \geqq -M + d(x, S) > -M.$$

The function

$$F_1(x) = \min [F(x), M]$$

is obviously also an extension of $f(x)$ which belongs to LIP_X; moreover $F_1(x)$ satisfies the inequality

$$-M < F_1(x) \leqq M.$$

In the same way we can get a function $F_2(x)$ defined on X which is an extension of $-f(x)$, belongs to LIP_X and satisfies the inequality

$$-M < F_2(x) \leqq M.$$

Now the function $\Phi(x) = \frac{1}{2}[F_1(x) - F_2(x)]$ belongs also to LIP_X, is also an extension of $f(x)$ and satisfies the inequality

$$-M < \Phi(x) < M.$$

Hence our statement is proved.

Now let us depart of the hypothesis of the boundedness of $f(x)$. Instead of $f(x)$ let us consider the function

$$g(x) = \operatorname{arctg} f(x)$$

which obviously also belongs to LIP_S.

By what we have already proved, we know that there exists a function $G(x)$ identical with $g(x)$ on S and which belongs to LIP_X. Moreover, we know that since $|g(x)| < \frac{\pi}{2}$, $G(x)$ may be chosen so that the inequality

$$|G(x)| < \frac{\pi}{2}$$

shall hold everywhere. Now if we put

$$F(x) = \operatorname{tg} G(x),$$

then we have the desired extension of $f(x)$ since $\operatorname{tg} t$ is differentiable in $-\frac{\pi}{2} < t < \frac{\pi}{2}$.

§ 3

We should like to point out that Theorems I and II remain valid if the Lipschitz condition is replaced by the more general condition

$$|f(x)-f(y)| \leqq K\omega(\overline{xy}) \quad \text{resp.} \quad |f(x)-f(y)| \leqq K(x)\omega(\overline{xy}) \quad \text{for} \quad \overline{xy} < \delta(x) \quad (\delta(x) > 0),$$

where $\omega(t)$ is a function defined in $0 \leqq t < \infty$ with the following properties:

$\omega(t)$ is continuous and monotone increasing,
$\omega(0) = 0$, $\omega(t) > 0$ if $t > 0$,
$\omega(x+y) \leqq \omega(x) + \omega(y)$.

(An important example of such an $\omega(t)$ is t^{α} $(0 < \alpha \leqq 1)$.) To see this, we have only to redefine the metric of X by putting

$$\varrho(x, y) = \omega(\overline{xy}) \qquad (x, y \in X),$$

where $\varrho(x, y)$ designes the distance of x and y according to the new metric. The prescribed properties of $\omega(t)$ guarantee that $\varrho(x, y)$ satisfy all the axioms valid for a metric in a metric space. In this new metric $f(x)$ satisfies a Lipschitz condition

$$f(x)-f(y)| \leqq K\varrho(x, y) \text{ resp. } |f(x)-f(y)| \leqq K(x)\varrho(x, y) \text{ for } \varrho(x, y) < \omega[\delta(x)].$$

Thus we have reduced this general case to the special one discussed above. We have still to remark that in case of Theorem II the set S remains closed in the new metric.

(Received 8 February 1955)

Note added in proof. (10 June 1955.) Theorem I and its generalization to a Lipschitz condition of arbitrary exponent p $(0 < p \leqq 1)$ with almost the same proof as in the present paper is to be found in a book of STEFAN BANACH, *Wstep do teorii funkcji rzeczywistych* (Warszawa—Wrocław, 1951), pp. 121—122. We were informed of this fact by an oral communication of Prof. STANISŁAW MAZUR late after the proof had been read.

О РАСПРОСТРАНЕНИИ ФУНКЦИЙ, УДОВЛЕТВОРЯЮЩИХ УСЛОВИЯМ ЛИПШИЦА

Я. Ципсер (Будапешт) и Л. Гехер (Мишкольц)

(Резюме)

Пусть $\omega(t)$ $(0 \leqq t < \infty)$ — непрерывная монотонно возрастающая функция, удовлетворяющая условиям

$$\omega(0) = 0, \quad \omega(t) > 0 \quad (t > 0), \quad \omega(x+y) \leqq \omega(x) + \omega(y).$$

(Например, $\omega(t) = t^{\alpha}$ $(0 < \alpha \leqq 1)$.)

В настоящей работе доказываются следующие две теоремы:

Теорема 1. Пусть $f(x)$ — вещественная функция, определенная на подмножестве S метрического пространства X и удовлетворяющая условию

$$|f(x) - f(y)| \leqq \omega(\overline{xy}),$$

где x и y две произвольные точки S, а $\overline{xy}$ означает расстояние этих двух точек. Тогда существует функция $F(x)$, определенная на всем пространстве X и совпадающая на S с $f x$, для которой

$$|F(x) - F(y)| \leqq \omega(\overline{xy}),$$

где x и y две произвольные точки X.

Теорема 2. Пусть $f(x)$ — вещественная функция, определенная на замкнутом подмножестве S метрического пространства X, и удовлетворяющая во всякой точке x множества S условию

$$|f(y) - f(x)| \leqq k(x)\,\omega(\overline{xy}) \qquad (\overline{xy} < \delta(x)),$$

где y произвольная точка S, а $k(x)$ и $\delta(x)$ положительные числа, зависящие от x. Тогда можно найти функцию $F(x)$ определенную на X и совпадающую на S с $f(x)$, которая во всякой точке x пространства X удовлетворяет условию

$$|F(y) - F(x)| \leqq K(x)\,\omega(\overline{xy}) \qquad (\overline{xy} < \Delta(x)),$$

где y произвольная точка X, а $K(x)$ и $\Delta(x)$ положительные числа, зависящие от x.

ÜBER DAS GLIEDWEISE DIFFERENZIEREN EINER ORTHOGONALEN POLYNOMREIHE

Von
G. FREUD (Budapest)
(*Vorgelegt von* G. ALEXITS)

Es sei $p_n(x)$ die Folge der normierten Orthogonalpolynome mit der nichtnegativen Gewichtsfunktion $w(x) \in L$ in $(-1, +1)$. Wir bilden die Orthogonalentwicklung

$$f(x) \sim \sum_{\nu=0}^{\infty} c_\nu p_\nu(x) \tag{1}$$

einer Funktion $f(x)$, welche in $[-1, +1]$ k-mal stetig differenzierbar ist. Wir wollen in vorliegender Arbeit die absolute Konvergenz der Reihe

$$f^{(k)}(x) \sim \sum_{\nu=0}^{\infty} c_\nu p_\nu^{(k)}(x) \tag{2}$$

untersuchen, die durch formales gliedweise Differenzieren aus (1) entsteht.

Zuerst einige Bemerkungen. Betrachten wir die Fourierreihe

$$F(x) \sim \sum_{\nu=0}^{\infty} (a_\nu \cos \nu x + b_\nu \sin \nu x)$$

einer k-mal stetig differenzierbaren Funktion $F(x)$, dann erhält man durch k-maliges Differenzieren die Fourierreihe der Derivierten $F^{(k)}(x)$:

$$F^{(k)}(x) \sim \sum_{\nu=0}^{\infty} \nu^k \left[a_\nu \cos \nu \left(x + \frac{k\pi}{2} \right) + b_\nu \sin \nu \left(x + \frac{k\pi}{2} \right) \right].$$

Die Frage der absoluten Konvergenz der derivierten Reihe ist also identisch mit der Frage der absoluten Konvergenz der Fourierreihe der Derivierten, die schon ausführlich behandelt wurde. Jedoch ist die Polynomreihe (2) (mit ev. Ausnahme bei speizieller Wahl von $w(x)$) keine Orthogonalreihe, also kann man ohne weiteres die bei der Untersuchung von Orthogonalreihen üblichen Methoden (vgl. z. B. STEČKIN [5]), ja sogar nicht einmal die Resultate des Verfassers (G. FREUD [3]) bezüglich der absoluten Konvergenz von Orthogonalpolynomreihen, auf die Reihe (2) nicht übertragen. Allerdings stellt es sich heraus, daß (2) dem Wesen nach mit denselben Methoden behandelt werden

kann, mit welchen Verfasser die absolute Konvergenz der Entwicklung nach Orthogonalpolynomen untersuchte. Unser Resultat kann in folgenden beiden Sätzen zusammengefaßt werden:

SATZ I. *Es sei* $(a, b) \subset [-1, +1]$ *und*

$$w(x) \geqq m > 0 \quad \text{für} \quad x \in (a, b); \tag{3}$$

ferner sei $f(x)$ *eine in* $[-1, +1]$ *k-mal stetig differenzierbare Funktion, für welche*

$$\sum_{n=1}^{\infty} \frac{\omega\left(\frac{1}{n}, f^{(k)}\right)}{\sqrt{n}} < \infty \tag{4}$$

besteht, wobei $\omega(\delta, f^{(k)})$ *den Stetigkeitsmodul von* $f^{(k)}(x)$ *bedeutet; dann ist die Reihe* (2) *gleichmäßig in jedem inneren Teilintervall von* (a, b) *absolut konvergent und stellt dort die Funktion* $f^{(k)}(x)$ *dar.*

SATZ II. *Es genüge* $w(x)$ *denselben Bedingungen, wie in Satz I und es sei außerdem* $w(x) \leqq W(1-x^2)^{-\frac{1}{2}}$. *Sei ferner* $f(\cos\vartheta) = g(\vartheta)$ *und bezeichne* $\omega_k^{(2)*}(\delta, f)$ *den Stetigkeitsmodul von* $\frac{d^k g}{d\vartheta^k}$ *bezüglich der Metrik* L^2, *also*

$$\omega_k^{(2)*}(\delta, f) = \left\{ \operatorname*{Max}_{|h| \leqq \delta} \int_0^{\pi} [g^{(k)}(\vartheta + h) - g^{(k)}(\vartheta)]^2 d\vartheta \right\}^{\frac{1}{2}}, \tag{5}$$

und es bestehe

$$\sum_{n=1}^{\infty} \frac{\omega_k^{(2)*}\left(\frac{1}{n}, f\right)}{\sqrt{n}} < \infty, \tag{6}$$

dann ist die Reihe (2) *gleichmäßig in jedem inneren Teilintervall von* (a, b) *absolut konvergent und stellt dort die Funktion* $f^{(k)}(x)$ *dar.*

Wir bemerken, daß (4) für $f^{(k)} \in \operatorname{Lip} \alpha$ mit $\alpha > \frac{1}{2}$ bestimmt befriedigt ist.

HILFSSATZ I. *Aus* (3) *folgt*

$$\sigma_n^{(k)}(x) \equiv \sum_{\nu=0}^{n} [p_\nu^{(k)}(x)]^2 = O(n^{2k+1}) \tag{7}$$

gleichmäßig in x *für jedes innere Teilintervall von* (a, b).

BEWEIS. $\Pi_n(x)$ durchlaufe alle Polynome höchstens n-ten Grades mit

$$\int_{-1}^{+1} [\Pi_n(t)]^2 w(t)\, dt \leqq 1. \tag{8}$$

Wir suchen das Maximum von $\Pi_n^{(k)}(x)$. Es ist offenbar

$$\Pi_n(t) \equiv \sum_{\nu=0}^{n} \gamma_\nu p_\nu(t)$$

mit

$$\sum_{\nu=0}^{n} \gamma_\nu^2 \leqq 1.$$

Also ist

$$\Pi_n^{(k)}(x) = \sum_{\nu=0}^{n} \gamma_\nu p_\nu^{(k)}(x) \leqq \left\{ \sum_{\nu=0}^{n} [p_\nu^{(k)}(x)]^2 \right\}^{\frac{1}{2}},$$

und das Gleichheitszeichen findet statt, wenn man

$$\Pi_n(t) = \left\{ \sum_{\nu=0}^{n} [p_\nu^{(k)}(x)]^2 \right\}^{-\frac{1}{2}} \sum_{\nu=0}^{n} p_\nu^{(k)}(x) p_\nu(t) \tag{9}$$

wählt.

Wir vergleichen nun $\sigma_n^{(k)}(x)$ mit dem für eine Gewichtsfunktion $\overline{w}(t) \leqq w(t)$ ähnlich gebauten Ausdruck $\overline{\sigma}_n^{(k)}(x)$. Dann ist die Bedingungsungleichung (8) für das Polynom (9) auch mit der Gewichtsfunktion $\overline{w}(t)$ befriedigt, und infolgedessen ist das entsprechende Maximum mit $\overline{w}$, das gleich $\{\overline{\sigma}_n^{(k)}\}^{\frac{1}{2}}$ ist, sicher nicht kleiner als $\{\sigma_n^{(k)}\}^{\frac{1}{2}}$. Es ist also

$$\sum_{\nu=0}^{n} [p_\nu^{(k)}(x)]^2 \leqq \sum_{\nu=0}^{n} [\overline{p}_\nu^{(k)}(x)]^2. \tag{10}$$

Es sei nun speziell

$$\overline{w}(t) = \begin{cases} m & \text{für} \quad t \in (a, b), \\ 0 & \text{für} \quad t \overline{\in} (a, b), \end{cases}$$

die entsprechenden normierten Orthogonalpolynome sind

$$\overline{p}_\nu(x) = \sqrt{\frac{(2\nu+1)m}{b-a}}\, P_\nu\left(-1 + 2\frac{x-a}{b-a}\right),$$

wobei $P_\nu(x)$, wie üblich, das ν-te Legendresche Polynom bedeutet. Hieraus folgt $\overline{p}_\nu(x) = O(1)$ gleichmäßig in jedem inneren Teilintervall von (a, b) und unter Anvendung des Satzes von BERNSTEIN

$$\overline{p}_\nu^{(k)}(x) = O(\nu^k) \tag{11}$$

ebenfalls gleichmäßig in jedem inneren Teilintervall von (a, b). Aus (10) und (11) erhalten wir unmittelbar die Abschätzung (7), w. z. b. w.

Bezeichne nun

$$E_n^{(2)}(f; w) = \left\{ \sum_{\nu=n+1}^{\infty} c_\nu^2 \right\}^{\frac{1}{2}}. \tag{12}$$

HILFSSATZ II. *Die Gewichtsfunktion sei so beschaffen, daß die Derivierten der Orthogonalpolynome an der Stelle* x (7) *befriedigen; ferner gelte für die Funktion* $f(x)$

$$\sum_{\nu=1}^{\infty} \nu^{k-\frac{1}{2}} E_\nu^{(2)}(f;w) < \infty; \tag{13}$$

dann ist die Reihe (2) *an der Stelle* x *absolut konvergent.*

BEWEIS. Es ist, da $E_n^{(2)}(f;w)$ eine monoton abnehmende Funktion von n ist,

$$\sum_{\nu=2^m}^{2^{m+1}-1} |c_\nu p_\nu^{(k)}(x)| \leqq \left\{ \sum_{\nu=2^m}^{2^{m+1}-1} c_\nu^2 \right\}^{\frac{1}{2}} \left\{ \sum_{\nu=0}^{2^{m+1}} [p_\nu^{(k)}(x)]^2 \right\}^{\frac{1}{2}} = E_{2^m}^{(2)}(f;w) \cdot O\left(2^{\left(k+\frac{1}{2}\right)m}\right) =$$

$$= O\left(2^{\left(k-\frac{1}{2}\right)m}\right) \sum_{\nu=2^{m-1}+1}^{2^m} E_n^{(2)}(f;w) = O\left\{ \sum_{\nu=2^{m-1}+1}^{2^m} \nu^{k-\frac{1}{2}} E_n^{(2)}(f;w) \right\},$$

und man addiere diese Ungleichung vom $m=1$ bis $m=\infty$.

HILFSSATZ III. *Es besteht die Ungleichung*

$$E_n^{(2)}(f;w) \leqq c_1 n^{-k} \omega\left(\frac{1}{n}; f^{(k)}\right), \tag{14}$$

wobei die Zahl c_1 *von* n *unabhängig ist.*

BEWEIS. Bekanntlich ist

$$E_n^{(2)}(f;w) = \left\{ \operatorname{Min} \int_{-1}^{+1} [f(x) - \Pi_n(x)]^2 w(x)\, dx \right\}^{\frac{1}{2}},$$

wobei $\Pi_n(x)$ alle Polynome höchstens n-ten Grades durchläuft. Als Konkurrenzpolynom wählen wir das Polynom, welche die bestmögliche Approximation von $f(x)$ im Tschebyscheffschen Sinne liefert; die Abweichung sei $E_n(f;w)$. Nach dem Satze von JACKSON ist

$$E_n(f;w) \leqq c_2 n^{-k} \omega\left(\frac{1}{n}; f^{(k)}\right),$$

und somit besteht

$$E_n^{(2)}(f;w) \leqq \left\{ \int_{-1}^{+1} w(x)\, dx \right\}^{\frac{1}{2}} c_2 n^{-k} \omega\left(\frac{1}{n}; f^{(k)}\right),$$

w. z. b. w.

Aus den Hilfssätzen I, II und III folgt unmittelbar die absolute Konvergenz von (2). Es bleibt nur übrig zu beweisen, daß die Summe der in jedem inneren Teilintervall von (a, b) gleichmäßig absolut konvergenten Reihe (2) die Funktion $f^{(k)}(x)$ darstellt.

Nach k-maliger Integration erhalten wir aber die orthogonale Polynomreihe (1), die ebenfalls in jedem inneren Teilintervall gleichmäßig absolut konvergent ist. Nach einem Satze des Verfassers [4] konvergieren die $(C, 1)$

Summen der Reihe (1) gegen $f(x)$, also ist auch die gewöhnliche Summe dieser Reihe gleich $f(x)$. Infolge der gleichmäßigen Konvergenz der differenzierten Reihe (2) ist die gliedweise Differentiation der Reihe (1) erlaubt, und hieraus folgt, daß die Summe der Reihe (2) gleich $f^{(k)}(x)$ sein muß. Um auch den Beweis des Satzes II zu Ende führen, muß Hilfssatz III durch den folgenden ersetzt werden:

HILFSSATZ IV. *Ist*

$$w(x) \leqq W(1-x^2)^{-\frac{1}{2}}, \tag{15}$$

dann gilt

$$E_n^{(2)}(f; w) \leqq c_3 n^{-k} \omega_k^{(2)*}\left(\frac{1}{n}; f\right). \tag{16}$$

Wir bezeichnen mit

$$a_n = \frac{2}{\pi}\int_0^{\pi} f(\cos\theta)\cos n\theta\, d\theta$$

die Koeffizienten der Cosinusentwicklung von $f(\cos\theta)$. Nach einem Lemma von G. ALEXITS [1] besteht für jede monoton nicht abnehmende Folge $\{\lambda_\nu\}$ positiver Zahlen

$$\sum_{\nu=0}^{\infty}\lambda_\nu c_\nu^2 \leqq \frac{W\pi}{2}\sum_{\nu=0}^{\infty}\lambda_\nu a_\nu^2. \tag{17}$$

Wir betrachten die Identität

$$\sum_{\nu=1}^{\infty}\nu^{2k}a_\nu^2\left(1-\frac{\sin\frac{k}{n}}{\frac{k}{n}}\right) = \frac{n}{\pi}\int_0^{\frac{1}{n}}\int_0^{\pi}\left[g^{(k)}\left(\vartheta+\frac{h}{2}\right)-g^{(k)}\left(\vartheta-\frac{h}{2}\right)\right]^2 d\vartheta.$$

Die rechte Seite ist höchstens gleich $\omega_k^{(2)*}\left(\frac{1}{n}\right)$; beachten wir noch, daß für $k \geqq n$ der Faktor $1-\frac{\sin\frac{k}{n}}{\frac{k}{n}}$ oberhalb einer positiven Zahl c_4^{-1} bleibt, so erhält man

$$\sum_{\nu=n+1}^{\infty}\nu^{2k}a_\nu^2 \leqq c_4\left[\omega_k^{(2)*}\left(\frac{1}{n}\right)\right]^2,$$

also wegen (17)

$$\sum_{\nu=n+1}^{\infty}c_\nu^2 \leqq n^{-2k}\sum_{\nu=n+1}^{\infty}\nu^{2k}c_\nu^2 \leqq \frac{W\pi}{2}n^{-2k}\sum_{\nu=n+1}^{\infty}\nu^{2k}a_\nu^2 \leqq \frac{W\pi}{2}c_4 n^{-2k}\left[\omega_k^{(2)*}\left(\frac{1}{n}\right)\right]^2$$

und damit haben wir (16) bewiesen.

Satz II folgt aus den Hilfssätzen I, II und IV.

(Eingegangen am 18. März 1955.)

Literaturverzeichnis

[1] G. ALEXITS, Sur la convergence et la sommabilité presque partout des series de polynomes orthogonaux, *Acta Sci. Math.*, **12** (1950), S. 223—225.

[2] S. N. BERNSTEIN, Sur la convergence absolue des series trigonometriques, *Comptes Rendus Acad. Sci. Paris*, **199** (1934), S. 397—400.

[3] G. FREUD, Über die absolute Konvergenz von orthogonalen Polynomreihen, *Acta Math. Acad. Sci. Hung.*, **4** (1953), S. 127—135.

[4] G. FREUD, Über die starke $(C, 1)$-Summierbarkeit von orthogonalen Polynomreihen, *Acta Math. Acad. Sci. Hung.*, **3** (1952), S. 83—88.

[5] С. Б. Стечкин, Об абсолютной сходимости ортогональных рядов, Математический Сборник, **29** (1951), S. 225—232.

О ПОЧЛЕННОМ ДИФФЕРЕНЦИРОВАНИИ РЯДОВ ПО ОРТОГОНАЛЬНЫМ МНОГОЧЛЕНАМ

Г. Фрайд (Будапешт)

(Резюме)

Пусть $w(x) \in L$ $(-1 \leqq x \leqq 1)$ — неотрицате льная весовая функция; обозначим соответствующие нормированные ортогональные многочлены через $p_0(x), p_1(x), \ldots$.

Пусть

$$f(x) \sim \sum c_\nu p_\nu(x) \tag{1}$$

разложения по системе $\{p_\nu(x)\}$ k-раз непрерывно дифференцируемой функций $f(x)$. k-кратным формальным дифференцированием отсюда получаем

$$f^{(k)}(x) \sim \sum c_\nu p_\nu^{(k)}(x). \tag{2}$$

Имеет место следующая

Теорема. Пусть

$$w(x) \geqq m > 0 \qquad (-1 \leqq a < x < b \leqq 1) \tag{3}$$

и, обозначая модуль непрерывности функции $f^{(k)}(x)$ через $\omega(\delta, f^{(k)})$, пусть

$$\sum_{n=1}^{\infty} \frac{\omega\left(\frac{1}{n}, f^{(k)}\right)}{\sqrt{n}} < \infty; \tag{4}$$

тогда (2) во всяком внутреннем подинтервале (a, b) сходится абсолютно и равномерно, и его сумма равно $f^{(k)}(x)$.

Условие (4) можно заменять следующим: Пусть

$$w(x) \leqq w(1 - x^2)^{-1/2}$$

и для $g(\vartheta) = f(\cos \vartheta)$

$$\sum_{n=1}^{\infty} \frac{\omega^{(2)}\left(\frac{1}{n}, g^{(k)}\right)}{\sqrt{n}} < \infty.$$

Les Acta Mathematica paraissent en russe, français, anglais et allemand et publient des mémoires du domaine des sciences mathématiques.

Les Acta Mathematica sont publiés sous forme de fascicules qui seront réunis en un volume.

On est prié d'envoyer les manuscrits destinés à la rédaction et écrits à la machine à l'adresse suivante:

Acta Mathematica, Budapest 62, Postafiók 440.

Toute correspondance doit être envoyée à cette même adresse.

Le prix de l'abonnement est de 110 forints par volume.

On peut s'abonner à l'Entreprise pour le Commerce Extérieur des Livres et Journaux „Kultúra" (Budapest, VI., Sztálin út 21. Compte-courant No. 43-790-05-1817) ou à l'étranger chez tous les représentants ou dépositaires

---

The Acta Mathematica publish papers on mathematics in Russian, French, English and German.

The Acta Mathematica appear in parts of various size, making up one volume.

Manuscripts should be typed and addressed to:

Acta Mathematica, Budapest 62, Postafiók 440.

Correspondence with the editors and publishers should be sent to the same address.

The rate of subscription to the Acta Mathematica is 110 forints a volume. Orders may be placed with „Kultúra" Foreign Trade Company for Books and Newspapers (Budapest, VI., Sztálin út 21. Account No. 43-790-057-181) or with representatives abroad.

Die Acta Mathematica veröffentlichen Abhandlungen aus dem Bereiche der mathematischen Wissenschaften in russischer, französicher, englischer und deutscher Sprache

Die Acta Mathematica erscheinen in Heften wechselnden Umfanges. Mehrere Hefte bilden einen Band.

Die zur Veröffentlichung bestimmten Manuskripte sind, mit Maschine geschrieben, an folgende Adresse zu senden:

Acta Mathematica, Budapest 62, Postafiók 440.

An die gleiche Anschrift ist auch jede für die Redaktion und den Verlag bestimmte Korrespondenz zu richten.

Abonnementspreis pro Band 110 Forints Bestellbar bei dem Buch- und Zeitungs-Außenhandels-Unternehmen „Kultúra" (Budapest, VI., Sztálin út 21. Bankkonto Nr. 43-790-057-181) oder bei seinen Auslandsvertretungen und Kommissionären.

56,— Ft

INDEX

Alexits, G., Über die Konvergenz der Orthogonalpolynomentwicklungen 1
Rédei, L., Zetafunktionen in der Algebra 5
Rédei, L., Neuer Beweis des Hajósschen Satzes über die endlichen Abelschen Gruppen 27
Alexits, G., Sur la caractérisation de certaines classes de fonctions au sens de la théorie constructive des fonctions 41
Erdős, P. and *Turán, P.*, On the role of the Lebesgue functions in the theory of the Lagrange interpolation 47
Surányi, J. and *Turán, P.*, Notes on interpolation. I. (On some interpolatorical properties of the ultraspherical polynomials) 67
Takács, L., On processes of happenings generated by means of a Poisson process . 81
Takács, L., Investigation of waiting time problems by reduction to Markov processes 101
Aczél, J., Lösung der Vektor-Funktionalgleichung der homogenen und inhomogenen n-dimensionalen einparametrigen „Translation", der erzeugenden Funktion von Kettenreaktionen und des stationären und nicht-stationären Bewegungsintegrals 131
Fejes Tóth, L., Extremum properties of the regular polytopes 143
Mikolás, M., Über gewisse Eigenschaften orthogonaler Systeme der Klasse L^2 und die Eigenfunktionen Sturm—Liouvillescher Differentialgleichungen 147
Ilieff, L., Über trigonometrische Integrale, welche ganze Funktionen mit nur reellen Nullstellen darstellen 191
Gilbert, W. M., Projections of probability distributions 195
Fisz, M., Refinement of a probability limit theorem and its application to Bessel functions 199
Szüsz, P., Bemerkungen zur Approximation einer reellen Zahl durch Brüche 203
Czipszer, J. and *Gehér, L.*, Extension of functions satisfying a Lipschitz condition . . 213
Freud, G., Über das gliedweise Differenzieren einer orthogonalen Polynomreihe . . 221

Technikai szerkesztő: Molnár Ferenc

A kiadásért felelős: az Akadémiai Kiadó igazgatója Műszaki felelős: Farkas Sándor
A kézirat beérkezett: 1955. IV. 12. — Terjedelem: 19½ (A/5) ív, 2 ábra.

Szegedi Nyomda Vállalat, Szeged, 55-1876 Felelős vezető: Vincze György

ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

ADIUVANTIBUS

G. ALEXITS, E. EGERVÁRY, L. FEJÉR, CH. JORDAN,
L. KALMÁR, L. RÉDEI, A. RÉNYI, F. RIESZ,
B. SZ.-NAGY, P. TURÁN, O. VARGA

REDIGIT

G. HAJÓS

TOMUS VI FASCICULI 3—4

MAGYAR TUDOMÁNYOS AKADÉMIA
BUDAPEST, 1955

ACTA MATH. HUNG.

ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

A MAGYAR TUDOMÁNYOS AKADÉMIA III. OSZTÁLYÁNAK MATEMATIKAI KÖZLEMÉNYEI

SZERKESZTŐSÉG ÉS KIADÓHIVATAL: BUDAPEST, V., ALKOTMÁNY U. 21

Az Acta Mathematica orosz, francia, angol és német nyelven közöl értekezéseket a matematika köréből.

Az Acta Mathematica változó terjedelmű füzetekben jelenik meg, több füzet alkot egy kötetet.

A közlésre szánt kéziratok géppel írva, a következő címre küldendők:

Acta Mathematica, Budapest 62, Postafiók 440.

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Az Acta Mathematica előfizetési ára kötetenként belföldre 80 forint, külföldre 110 forint. Megrendelhető a belföld számára az „Akadémiai Kiadó"-nál (Budapest, V., Alkotmány utca 21. Bankszámla 04-878-111-46), a külföld számára pedig a „Kultúra" Könyv- és Hírlap Külkereskedelmi Vállalatnál (Budapest, VI., Sztálin út 21. Bankszámla 43-790-057-181) vagy külföldi képviseleteinél és bizományosainál.

„Acta Mathematica" публикует трактаты из области математических наук на русском, французском, английском и немецком языках.

„Acta Mathematica" выходит отдельными выпусками разного объема. Несколько выпусков составляют один том.

Предназначенные для публикации рукописи (в напечатанном на машинке виде) следует направлять по адресу:

Acta Mathematica, Budapest 62, Postafiók 440.

По этому же адресу направлять всякую корреспонденцию для редакции и администрации.

Подписная цена *„Acta Mathematica"* — 110 форинтов за том. Заказы принимает предприятие по внешней торговле книг и газет *„Kultúra"* (*Budapest, VI., Sztálin út 21.* Текущий счет № 43-790-057-181), или его заграничные представительства и уполномоченные.

VERSCHIEDENE BEMERKUNGEN ELEMENTARER NATUR ÜBER DIE GRUNDPOLYNOME, DIE BEI DEN PARABOLISCHEN INTERPOLATIONEN AUFTRETEN

Von

LEOPOLD FEJÉR (Budapest), Mitglied der Akademie

GÁBOR SZEGŐ *gewidmet zu seinem 60-sten Geburtstage*

Einleitung

1. Die Aufgabe der *parabolischen* Interpolation besteht bekanntlich in der Bestimmung einer ganzen rationalen Funktion der unabhängigen Variabel x, deren Wert an den beliebig vorgeschriebenen Stellen $x_1, x_2, \ldots, x_n$, die aber voneinander verschieden sind, der Reihe nach, die beliebig vorgeschriebenen Werte $y_1, y_2, \ldots, y_n$ annimmt; weiter können, an den Stellen $x_1, x_2, \ldots, x_n$, für die gesuchte ganze rationale Funktion eventuell noch einige Ableitungswerte bis zu einer gegebenen Ordnung vorgeschrieben werden.

2. In dieser Note beschäftige ich mich mit den beiden elementarsten Fällen parabolischer Interpolation: mit der „Lagrange"-schen und mit der „Hermite"-schen Interpolation. Bei der Lagrangeschen Interpolation sind für die gesuchte ganze rationale Funktion nur ihre Werte $y_1, y_2, \ldots, y_n$ an den Interpolationsstellen $x_1, x_2, \ldots, x_n$ vorgeschrieben, und es wird gefordert, daß sie von höchstens $(n-1)$-tem Grade sei. Hingegen werden bei der Hermiteschen Interpolation für die gesuchte ganze rationale Funktion außer den Funktionswerten $y_1, y_2, \ldots, y_n$ noch sämtliche erste Ableitungswerte $y_1', y_2', \ldots, y_n'$ vorgeschrieben, und es wird weiter gefordert, daß das Interpolationspolynom von höchstens $(2n-1)$-tem Grade sei.

Für beide Aufgaben gibt es tatsächlich ein den Forderungen entsprechendes Polynom, u. zw. nur ein einziges. Übrigens lautet die Lagrangesche Interpolationsformel

$$L(x) = y_1 l_1(x) + y_2 l_2(x) + \cdots + y_n l_n(x) \tag{1}$$

und die Hermitesche

$$H(x) = y_1 h_1(x) + y_2 h_2(x) + \cdots + y_n h_n(x) + y_1' \mathfrak{h}_1(x) + y_2' \mathfrak{h}_2(x) + \cdots + y_n' \mathfrak{h}_n(x). \tag{2}$$

Hier wird $l_k(x)$ das k-te Grundpolynom (oder Grundfunktion) von Lagrange genannt und ist diejenige (einzige) ganze rationale Funktion von genau $(n-1)$-tem Grade, die an der k-ten Interpolationsstelle x_k den Wert 1 annimmt, an allen übrigen Interpolationsstellen aber verschwindet. Weiter ist $h_k(x)$ die k-te Grundfunktion *erster* Art der Hermiteschen Interpolation, die im allgemeinen von höchstens $(2n-1)$-tem Grade sein soll, aber eventuell

auch von $(2n-2)$-tem Grade sein kann, die an der Stelle x_k ebenfalls den Wert 1 annimmt, an allen übrigen Interpolationsstellen verschwindet, und deren Ableitung an allen Interpolationsstellen verschwindet. Es gibt nur eine einzige g. r. Funktion, die diesen Forderungen genüge leistet.

Obzwar ich mich in dieser Einleitung auf die Grundpolynome *zweiter* Art $\mathfrak{h}_k(x)$ nicht ausbreiten werde, möchte ich doch schon hier erwähnen, daß sie diejenige (einzige) g. r. Funktion von genau $(2n-1)$-tem Grade bedeutet, die an sämtlichen Interpolationsstellen $x_1, x_2, \ldots, x_n$ verschwindet, deren erste Ableitung an der Stelle x_k den Wert 1 hat, und an allen übrigen Interpolationsstellen verschwindet.

3. Fragen der Approximierbarkeit einer *beliebigen stetigen Funktion* am Wege der Interpolation haben dazu geführt, den Verlauf dieser soeben erwähnten Grundpolynome etwas näher zu untersuchen. Wenn es nun ein Intervall $a \leqq x \leqq b$ gibt, das $x_1, x_2, \ldots, x_n$ einschließt und in dessen Inneren $a < x < b$ keines der Grundpolynome erster Art $h_k(x)$ negativ wird (in welchem Falle man die Verteilung der Interpolationsstellen im Intervalle $a \leqq x \leqq b$ als *„normal"* bezeichnet), so kann man, in angedeuteter Weise, mit Hilfe der Identität

$$h_1(x) + h_2(x) + \cdots + h_n(x) = 1 \tag{3}$$

(die man erhält, wenn man $y_1 = y_2 = \cdots y_n = 1$ wählt) gleich einen großen Schritt weiterkommen. Man erhält nämlich aus ihr, daß irgendein $h_k(x)$, das also, wie gesagt, an der Stelle $x = x_k$ den Wert 1 annimmt, *im ganzen Intervalle $a \leqq x \leqq b$ den Wert 1 nicht überschreiten kann*, ja man erhält in diesem speziellen Falle, daß sogar *die Summe aller $h_k(x)$ und an jeder Stelle x des ganzen Intervalles $a \leqq x \leqq b$ den Wert 1 nicht überschreiten kann.* $h_k(x_k) = 1$ ist also jedenfalls ein Maximum von $h_k(x)$ für das ganze Intervall $a \leqq x \leqq b$.

4. Lassen wir jetzt wieder fallen die Beschränkung der „Normalität" der Verteilung. Wir fragen: ist dann für $h_k(x)$ wenigstens soviel wahr, daß $h_k(x)$ für $x = x_k$ ein *Umgebungsmaximum* hat? Die Frage ist zu *bejahen*. Es gilt folgender allgemeiner

SATZ 1. *Sind $x_1, x_2, \ldots, x_n$ beliebige voneinander verschiedene reelle Zahlen (Interpolationsstellen), und bezeichnet $h_k(x)$ diejenige ganze rationale Funktion von $(2n-1)$-tem oder $(2n-2)$-tem Grade, die an der Stelle x_k den Wert 1 besitzt, sonst an allen übrigen Interpolationsstellen verschwindet, und deren erste Ableitung an sämtlichen Interpolationsstellen verschwindet, dann hat $h_k(x)$ an der Stelle $x = x_k$ ein „wirkliches" Umgebungsmaximum.*

Wenn

$$(x-x_1)(x-x_2)\cdots(x-x_n) = \omega(x) \tag{4}$$

gesetzt wird, und wenn

$$\xi_1, \xi_2, \ldots, \xi_{n-1}$$

die Nullstellen von $\frac{d\omega(x)}{dx}$ bezeichnen, d. h. diejenigen $n-1$, notwendigerweise voneinander verschiedenen Abszissen, an welchen das Polynom $\omega(x)$ ein Extremum hat, so lautet die nach steigenden Potenzen von $x-x_k$ laufende Entwickelung der ganzen rationalen Funktion $h_k(x)$[1]

$$(5) \qquad h_k(x)=1-\frac{1}{12}\left\{5\left(\sum_{m=1}^{n-1}\frac{1}{x_k-\xi_m}\right)^2+4\sum_{m=1}^{n-1}\frac{1}{(x_k-\xi_m)^2}\right\}(x-x_k)^2+\cdots.$$

Da

$$(6) \qquad \left(\sum_{m=1}^{n-1}\frac{1}{x_k-\xi_m}\right)^2$$

niemals negativ wird, und da

$$(7) \qquad \sum_{m=1}^{n-1}\frac{1}{(x_k-\xi_m)^2}$$

entschieden positiv ist, so liefert die Formel (5) den Beweis dafür, daß $h_k(x)$ an der Stelle $x=x_k$ ein „wirkliches" (d. h. isoliertes) Maximum hat.[2]

Zu diesem Satze 1 bemerke ich zunächst, daß der entsprechende Satz für das Lagrangesche $l_k(x)$ im allgemeinen *nicht* gültig ist. Weiter erwähne ich hier, daß zum Beweise von Satz 1 auch ein anderer Weg gangbar ist: die s. g. *„Methode der Wurzelabzählung"* führt ganz kurz zum Ziele. Der Grund, weshalb ich doch die auch sonst beachtenswerte Formel (5) an die Spitze dieser Note stellte, ist der, daß in ihr einerseits die *Nullstellen* $x_1, x_2, \ldots, x_n$ und die *Extremumstellen* $\xi_1, \xi_2, \ldots, \xi_{n-1}$ des *Stammpolynoms* $\omega(x)=(x-x_1)(x-x_2)\cdots(x-x_n)$ figurieren, und andererseits der Umstand,

[1] In der älteren Sprache der Analysten lautet dieses Resultat, daß die zur Stelle $x=x_k$ gehörige Grundfunktion $h_k(x)$ der Hermiteschen Interpolation in der „Umgebung" von x_k „gleich"

$$1-\frac{1}{12}\left\{5\left(\sum_{m=1}^{n-1}\frac{1}{x_k-\xi_m}\right)^2+4\sum_{m=1}^{n-1}\frac{1}{(x_k-\xi_m)^2}\right\}(x-x_k)^2$$

ist „bis auf Größen dritter Ordnung in Bezug auf $x-x_k$". Vielleicht darf ich noch zusammenfassend hinzufügen, daß der Wert der Ordinate von $y=h_k(x)$ an der Stelle $x=x_k$ den Wert 1 hat, die Tangente an dieser Stelle $x=x_k, y=1$ der Abszissenaxe parallel läuft, und daß die *Krümmung* von $y=h_k(x)$ an dieser Stelle gleich

$$\frac{1}{6}\left\{5\left(\sum_{m=1}^{n-1}\frac{1}{x_k-\xi_m}\right)^2+4\sum_{m=1}^{n-1}\frac{1}{(x_k-\xi_m)^2}\right\}$$

ist.

[2] Ich meine natürlich, daß für $h_k(x)$ eine zweiseitige Umgebung der Stelle $x=x_k$ existiert, in welcher, x_k ausgenommen, überall $h_k(x)<h_k(x_k)=1$ gültig ist. Ich setze hier voraus, daß mindestens zwei Interpolationsstellen gegeben sind, also $n=2, 3, \ldots$. Übrigens ist der Satz auch im Falle einer einzigen Interpolationsstelle x_1 gültig, da doch dann $h_1(x)\equiv 1$ ist, aber natürlich mit der Modifikation, daß das Maximum von $h_1(x)$ für $x=x_1$ *nicht* isoliert ist.

daß die Nichtnegativität von (6) und die Positivität von (7) die Wahrheit von Satz 1 in so anschaulicher Weise zur Evidenz erheben.

Endlich wollte ich durch Anführung von Satz 1 in der Einleitung den Charakter der elementaren Bemerkungen andeuten, die in dieser Note enthalten sind und die sich älteren, bekannten anreihen.

Es ist nicht zu erwarten, daß eine so kurze Note wie diese* auch die erforderlichen literarischen Daten des ausgedehnten Gebietes der Interpolationstheorie enthalte. Doch möchte ich erwähnen, daß SZEGŐ zu den *ersten* gehört, dem die Interpolationstheorie (und auch die Theorie der mechanischen Quadratur in der hier in Rede stehenden Richtung) bedeutende Fortschritte verdankt, die teilweise in seinem Standardwerk „*Orthogonal Polynomials*" (American Mathematical Society Colloquium Publications, Volume XXIII, 1939) entwickelt sind.

§ 1. Beweis für die Gültigkeit der Reihenentwickelung des Grundpolynoms erster Art $h_k(x)$:

$$h_k(x) = 1 - \frac{1}{12}\left\{5\left(\sum_{m=1}^{n-1}\frac{1}{x_k-\xi_m}\right)^2 + 4\sum_{m=1}^{n-1}\frac{1}{(x_k-\xi_m)^2}\right\}(x-x_k)^2 + \cdots$$

1. Es seien

$$x_1, x_2, \ldots, x_n \tag{1}$$

voneinander verschiedene reelle Werte (Interpolationsabszissen), und

$$\omega(x) = (x-x_1)(x-x_2)\cdots(x-x_n) \tag{2}$$

das zur Abszissengruppe (1) gehörige „*Stammpolynom*". Dann gilt bekanntlich

$$l_k(x) = \frac{\omega(x)}{\omega'(x_k)(x-x_k)}, \tag{3}$$

$$h_k(x) = \left(1 - \frac{\omega''(x_k)}{\omega'(x_k)}(x-x_k)\right)(l_k(x))^2. \tag{4}$$

Wenn wir nun das Polynom $\omega(x)$ um den Wert $x = x_k$ in Taylorsche Reihe entwickeln, so erhalten wir

$$\omega(x) = 0 + \omega'(x_k)(x-x_k) + \frac{1}{2}\omega''(x_k)(x-x_k)^2 + \frac{1}{6}\omega'''(x_k)(x-x_k)^3 + \cdots, \tag{5}$$

und mit Rücksicht auf (3) und (4) nacheinander

$$l_k(x) = 1 + \frac{1}{2}\frac{\omega''(x_k)}{\omega'(x_k)}(x-x_k) + \frac{1}{6}\frac{\omega'''(x_k)}{\omega'(x_k)}(x-x_k)^2 + \cdots, \tag{6}$$

* Bei der Drucklegung, resp. bei Anfertigung der Figuren waren mir J. BALÁZS, KATALIN HORVAY, M. MIKOLÁS, F. MOLNÁR behilflich. Es sei mir gestattet für ihre Mühe auch hier meinen herzlichen Dank auszusprechen.

$$(7)\qquad [l_k(x)]^2 = 1 + \frac{\omega''(x_k)}{\omega'(x_k)}(x-x_k) + \left[\frac{1}{3}\frac{\omega'''(x_k)}{\omega'(x_k)} + \frac{1}{4}\left(\frac{\omega''(x_k)}{\omega'(x_k)}\right)^2\right](x-x_k)^2 + \cdots,$$

$$(8)\qquad h_k(x) = 1 - \frac{A_k}{12}(x-x_k)^2 + \cdots,$$

wo

$$(9)\qquad A_k = 9\left(\frac{\omega''(x_k)}{\omega'(x_k)}\right)^2 - 4\frac{\omega'''(x_k)}{\omega'(x_k)}.$$

Zu beweisen habe ich, daß A_k positiv ist. Ich werde mehr beweisen, indem ich nämlich zeige, daß die aus dem Stammpolynom $\omega(x)$ entstandene gebrochene rationale Funktion von x

$$(10)\qquad A(x) = 9\left(\frac{\omega''(x)}{\omega'(x)}\right)^2 - 4\frac{\omega'''(x)}{\omega'(x)}$$

überhaupt für *jeden* reellen Wert von x einen endlichen positiven Wert besitzt und nur an den Nullstellen von $\omega'(x)$, die wir mit $\xi_1, \xi_2, \ldots, \xi_{n-1}$ bezeichnen, und die mit der Gesamtheit der Extremum-Stellen von $\omega(x)$ identisch sind, positiv unendlich wird. Ist nämlich x von $\xi_1, \xi_2, \ldots, \xi_{n-1}$ verschieden, so ist

$$(11)\qquad \frac{d}{dx}\frac{\omega''(x)}{\omega'(x)} = \frac{\omega'''(x)\omega'(x) - (\omega''(x))^2}{(\omega'(x))^2} = \frac{\omega'''(x)}{\omega'(x)} - \left(\frac{\omega''(x)}{\omega'(x)}\right)^2,$$

so daß

$$(12)\qquad \frac{\omega'''(x)}{\omega'(x)} = \left(\frac{\omega''(x)}{\omega'(x)}\right)^2 + \frac{d}{dx}\frac{\omega''(x)}{\omega'(x)}.$$

Jetzt ergibt (10)

$$(13)\qquad A(x) = 5\left(\frac{\omega''(x)}{\omega'(x)}\right)^2 - 4\frac{d}{dx}\frac{\omega''(x)}{\omega'(x)}.$$

Da weiter $\xi_1, \xi_2, \ldots, \xi_{n-1}$ die voneinander verschiedenen Wurzeln von $\omega'(x) = (x-\xi_1)(x-\xi_2)\cdots(x-\xi_{n-1})$ bezeichnen, so erhalten wir nacheinander

$$(14)\qquad \frac{\omega''(x)}{\omega'(x)} = \sum_{\nu=1}^{n-1}\frac{1}{x-\xi_\nu},$$

$$(15)\qquad \frac{d}{dx}\frac{\omega''(x)}{\omega'(x)} = -\sum_{\nu=1}^{n-1}\frac{1}{(x-\xi_\nu)^2},$$

so daß also, mit Rücksicht auf (13), (14), (15),

$$(16)\qquad A(x) = 5\left(\sum_{\nu=1}^{n-1}\frac{1}{x-\xi_\nu}\right)^2 + 4\sum_{\nu=1}^{n-1}\frac{1}{(x-\xi_\nu)^2}.$$

Hiermit ist der Satz 1 der Einleitung bewiesen.

2. Zur Illustration diene der Spezialfall $n=3$, $x_1=-1$, $x_2=0$, $x_3=1$. Jetzt ist $\omega(x) = x^3 - x$, $\omega'(x) = 3x^2 - 1$, $\xi_1 = -\frac{1}{\sqrt{3}}$, $\xi_2 = \frac{1}{\sqrt{3}}$, $\omega''(x) = 6x$,

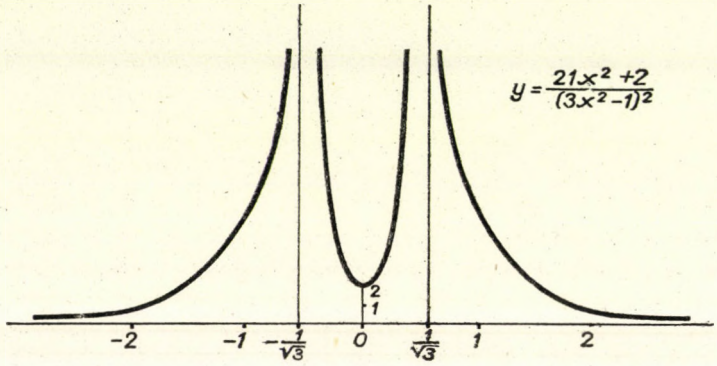
$\omega'''(x)=6$, so daß also die Formel (10)

$$(17)\quad A(x)=\frac{9(\omega''(x))^2-4\omega'(x)\,\omega'''(x)}{(\omega'(x))^2}=\frac{252x^2+24}{(3x^2-1)^2}=12\,\frac{21x^2+2}{(3x^2-1)^2},$$

d. h.

$$\frac{A(x)}{12}=\frac{21x^2+2}{(3x^2-1)^2}$$

liefert. (S. Figur 1.)



Figur 1.

§ 2. Über die Summe der Grundpolynome zweiter Art

1. Es seien $x_1, x_2, \ldots, x_n$ wieder voneinander verschiedene reelle Werte, die zu einer Hermiteschen Interpolation dienen. Wir haben in der Einleitung auf die einfache Tatsache hingewiesen, daß die Summe der Grundpolynome erster Art $h_1(x)+h_2(x)+\cdots+h_n(x)$ identisch gleich 1 ist. Mit der Summe der Grundpolynome *zweiter* Art $\mathfrak{h}_1(x)+\mathfrak{h}_2(x)+\cdots+\mathfrak{h}_n(x)$ habe ich mich zum erstenmal im § 10 meiner Arbeit „Über Weierstraßsche Approximation, besonders durch Hermitesche Interpolation" beschäftigt (*Mathematische Annalen*, **102** (1930), S. 707—725). Da werden ausschließlich die s. g. Tschebyscheff-schen Abszissen betrachtet, und die dort befindliche Formel (II), die durch leichteste Rechnung erhältlich ist, lautet[3]

$$(1)\quad \sum_{k=1}^{n}\mathfrak{h}_k(x)=\mathfrak{H}(x)=\mathfrak{H}(\cos\theta)=$$

$$=\frac{1}{n}\cos n\theta\cos(n-1)\theta=\frac{1}{2n}(\cos\theta+\cos(2n-1)\theta).$$

[3] Übrigens enthält dieser § 10, der die Überschrift trägt: „Über die wichtigsten Eigenschaften der Grundfunktionen beider Art im Falle Tschebyscheffscher Abszissen", auf S. 721, u. A. einen nicht mehr so leicht beweisbaren Satz, der sich auf $\sum_{k=1}^{n}|\mathfrak{h}_k(x)|$ bezieht,

2. Weiter wird in meiner Arbeit „Die Abschätzung eines Polynoms in einem Intervalle, wenn Schranken für seine Werte und ersten Ableitungswerte in einzelnen Punkten des Intervalles gegeben sind, und ihre Anwendung auf die Konvergenz Hermitescher Interpolationsreihen“ (*Mathematische Zeitschrift*, **32** (1930), S. 426—457) mit Hilfe der Formel (1) ein Extremumproblem *exakt* gelöst. Auf S. 441 steht nämlich der Satz 4: „Genügt eine ganze rationale Funktion $g(x)$ von höchstens $(2n-1)$-tem Grade an den Tschebyscheffschen Abszissen $x_1, x_2, \ldots, x_n$ den Ungleichungen

$$|g(x_k)| \leqq A, \quad |g'(x_k)| \leqq B \qquad (k = 1, 2, \ldots, n), \tag{2}$$

dann ist

$$|g(\pm 1)| \leqq A + \frac{B}{n}. \tag{3}$$

Hier ist das Zeichen der Gleichheit (bei gegebenen n, A, B) dann und nur dann gültig, wenn

$$\begin{aligned} g(x) = g(\cos\theta) &= A + B\frac{\cos n\theta \cos(n-1)\theta}{n} = \\ &= A + B\frac{\cos\theta + \cos(2n-1)\theta}{2n} \end{aligned} \tag{4}$$

ist, oder dasselbe Polynom mit entgegengesetztem Vorzeichen.“

3. Endlich möchte ich noch auf eine zweite Stelle derselben Math. Zeitschriftarbeit hinweisen, die sich auf Seite 456 befindet, und sich wieder nicht auf $\sum_{k=1}^{n} \mathfrak{h}_k(x)$, die uns jetzt beschäftigt, sondern auf $\sum_{k=1}^{n} |\mathfrak{h}_k(x)|$ bezieht. Es werden dort unter Nummer (19) und (20) die beiden Sätze ausgesprochen, daß in Tschebyscheffschem Falle

$$\sum_{k=1}^{n} |h_k(x)| = \sum_{k=1}^{n} h_k(x) = 1 \qquad (-1 < x < +1), \tag{5}$$

$$\sum_{k=1}^{n} |\mathfrak{h}_k(x)| \leqq 1 \qquad (-1 \leqq x \leqq +1). \tag{6}$$

4. Vor nicht langer Zeit habe ich die fast triviale Tatsache bemerkt, daß die Summe der Grundpolynome zweiter Art, d. h., kurz ausgedrückt, $\mathfrak{H}(x)$ „immer“ voneinander verschiedene reelle Nullstellen besitzt. Ich darf vielleicht den Satz formulieren:

und später von verschiedenen Autoren kunstvoll geeignet verallgemeinert wurde, für den Fall, in welchem die Abszissen nicht mehr die Tschebyscheffschen, sondern allgemeinere, in diesem oder jenem Sinne „normal“ verteilte Abszissen des Interpolationsintervalls $(-1, +1)$ sind. Mein Satz lautet: „Die Summe der absoluten Beträge der Grundpolynome zweiter Art $\mathfrak{h}_k(x)$, d. h. die Summe $\sum_{k=1}^{n} |\mathfrak{h}_k(x)|$, konvergiert im Falle Tschebyscheffscher Abszissen zu Null, wenn $n \to \infty$, u. zw. gleichmäßig im Intervalle $-1 \leqq x \leqq +1$.“

SATZ 2. *Sind $x_1, x_2, \ldots, x_n$ beliebige voneinander verschiedene reelle Abszissenwerte, dann sind die Grundpolynome zweiter Art*

$$\mathfrak{h}_1(x), \mathfrak{h}_2(x), \ldots, \mathfrak{h}_n(x)$$

der Hermiteschen Interpolation durchweg ganze rationale Funktionen von genau $(2n-1)$-tem Grade. Jedes dieser Grundpolynome hat lauter reelle Nullstellen, und zwar hat $\mathfrak{h}_k(x)$ an der Stelle x_k eine einfache Nullstelle, während sie an den übrigen Stellen $x_1, x_2, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n$ eine doppelte Nullstelle besitzt. Die Summe

$$\mathfrak{h}_1(x) + \mathfrak{h}_2(x) + \cdots + \mathfrak{h}_n(x) = \mathfrak{H}(x)$$

aller Grundfunktionen zweiter Art ist ebenfalls eine ganze rationale Funktion von genau $(2n-1)$-tem Grade, die nun ebenfalls lauter, voneinander verschiedene reelle Nullstellen hat. n dieser einfachen Nullstellen liegen in den Interpolationsstellen $x_1, x_2, \ldots, x_n$ selbst, die übrigen $n-1$ einfachen Nullstellen $\eta_1, \eta_2, \ldots, \eta_{n-1}$ befinden sich (wenn wir etwa $x_1 < x_2 < \cdots < x_n$ voraussetzen) in den Intervallen $(x_1, x_2), (x_2, x_3), \ldots, (x_{n-1}, x_n)$, so daß also

$$x_1 < \eta_1 < x_2 < \eta_2 < x_3 < \cdots < x_{n-1} < \eta_{n-1} < x_n. \tag{7}$$

5. Der Beweis dieses Satzes läßt sich auf verschiedenen Wegen leisten. Er folgt natürlich unmittelbar aus der Tatsache, daß, wenn

$$\sum_{k=1}^{n} \mathfrak{h}_k(x) = \omega(x)\,\Omega(x) \tag{8}$$

gesetzt wird, $\Omega(x)$ an den Stellen $x_1, x_2, \ldots, x_n$, der Reihe nach, die von Null verschiedenen endlichen Werte

$$\frac{1}{\omega'(x_1)}, \frac{1}{\omega'(x_2)}, \ldots, \frac{1}{\omega'(x_n)}$$

annimmt,[4] die ihr Vorzeichen immer ändern, wenn man von einer Nullstelle x_k des Stammpolynoms $\omega(x)$ zur folgenden x_{k+1} schreitet.

Übrigens läßt sich dieser Satz auch noch einfacher beweisen. Ist doch

$$\begin{aligned}\mathfrak{H}(x) = \sum_{k=1}^{n} \mathfrak{h}_k(x) &= \sum_{k=1}^{n} (x-x_k)\,(l_k(x))^2 = \sum_{k=1}^{n} (x-x_k)\left(\frac{\omega(x)}{\omega'(x_k)\,(x-x_k)}\right)^2 = \\ &= \omega(x) \sum_{k=1}^{n} \frac{1}{\omega'(x_k)} \frac{\omega(x)}{\omega'(x_k)\,(x-x_k)} = \omega(x) \sum_{k=1}^{n} \frac{1}{\omega'(x_k)}\, l_k(x),\end{aligned} \tag{9}$$

d. h. $\sum_{k=1}^{n} \mathfrak{h}_k(x)$ ist gleich dem Stammpolynom $\omega(x)$, multipliziert mit dem Polynome $(n-1)$-ten Grades, das (mit Rücksicht auf die Lagrangesche Inter-

[4] Dies folgt aus der Gleichung (8):

$$1 = \left(\sum_{k=1}^{n} \mathfrak{h}_k(x)\right)'_{x=x_k} = \omega(x_k)\,\Omega'(x_k) + \omega'(x_k)\,\Omega(x_k) = \omega'(x_k)\,\Omega(x_k).$$

polationsformel) an den Stellen $x_1, x_2, \ldots, x_n$, der Reihe nach, die Wertenfolge alternierenden Vorzeichens $\frac{1}{\omega'(x_1)}, \frac{1}{\omega'(x_2)}, \ldots, \frac{1}{\omega'(x_n)}$ annimmt.

Ein dritter, vielleicht einfachster Beweis dieses so naheliegenden Satzes lautet: da

$$\mathfrak{H}(x_1) = \mathfrak{H}(x_2) = \cdots = \mathfrak{H}(x_n) = 0,$$
$$\mathfrak{H}'(x_1) = \mathfrak{H}'(x_2) = \cdots = \mathfrak{H}'(x_n) = 1,$$

so *wächst* $\mathfrak{H}(x)$ an ihren (einfachen) Nullstellen $x_1 < x_2 < \cdots < x_n$ nach *vorwärts* und *nimmt ab* nach *rückwärts* usw.

§ 3. Über die Summe der Grundpolynome zweiter Art. Fortsetzung

1. Betrachten wir die Differenz

$$k(x) = x - \sum_{k=1}^{n} \mathfrak{h}_k(x) = x - \mathfrak{H}(x). \tag{1}$$

Da $\mathfrak{H}'(x_1) = \mathfrak{H}'(x_2) = \cdots = \mathfrak{H}'(x_n) = 1$, so gilt für die Ableitung $k'(x)$ der ganzen rationalen Funktion $(2n-1)$-ten Grades $k(x)$:

$$k'(x_1) = k'(x_2) = \cdots = k'(x_n) = 0.$$

Also ist $k(x)$ eine zu den Stellen $x_1, x_2, \ldots, x_n$ gehörige, sogenannte *Treppenparabel.* Daher liefert die Formel (2) der Einleitung, mit Rücksicht auf $k(x_1) = x_1, k(x_2) = x_2, \ldots, k(x_n) = x_n$,

$$k(x) = x - \mathfrak{H}(x) = \sum_{k=1}^{n} x_k h_k(x), \tag{2}$$

woraus, mit Rücksicht auf $\sum_{k=1}^{n} h_k(x) = 1$,

$$\mathfrak{H}(x) = x - \sum_{k=1}^{n} x_k h_k(x) = x \sum_{k=1}^{n} h_k(x) - \sum_{k=1}^{n} x_k h_k(x) = \sum_{k=1}^{n} (x - x_k) h_k(x) \tag{3}$$

folgt. Wir haben also für die Summe der Grundpolynome zweiter Art die Formel

$$\sum_{k=1}^{n} \mathfrak{h}_k(x) = \sum_{k=1}^{n} (x - x_k) h_k(x) \tag{4}$$

gefunden, die, ich wiederhole es, für beliebige Stellen $x_1, x_2, \ldots, x_n$ gültig ist.

§ 4. Über die Summe der Grundpolynome zweiter Art, wenn die Abszissen $x_1, x_2, \ldots, x_n$ im Intervalle $a \leqq x \leqq b$, in welchem sie eingeschlossen sind, dort normal verteilt sind

1. Wenn $H'(x)$ die Ableitung nach der reellen Variabel x einer ganzen rationalen Funktion $H(x)$ mit reellen Koeffizienten und von höchstens $(2n-1)$-tem Grade bezeichnet, und diese Ableitung $H'(x)$ an n verschiedenen reellen Stellen $x_1, x_2, \ldots, x_n$ verschwindet (also $H'(x_1) = H'(x_2) = \cdots = H'(x_n) = 0$), so nenne ich die Kurve $y = H(x)$, oder auch das Polynom $H(x)$ selbst, eine Treppenparabel. Ist weiter $a \leqq x \leqq b$ irgend ein abgeschlossenes Intervall, das die Stellen $x_1, x_2, \ldots, x_n$ ausnahmslos enthält, während keine einzige der Stellen $X_1, X_2, \ldots, X_n$, die, mit Hilfe der Stammfunktion

$$\omega(x) = (x-x_1)(x-x_2)\cdots(x-x_n)$$

und *nur* mit ihrer Hilfe, durch die Formel

$$X_k = x_k + \frac{\omega'(x_k)}{\omega''(x_k)} \qquad (k = 1, 2, \ldots, n) \tag{1}$$

definiert sind, in das Innere des Intervalles $a \leqq x \leqq b$ fällt (d. h. ist entweder $X_k > b$ oder $X_k < a$ $(k = 1, 2, \ldots, n)$), dann sage ich, daß die Stellen $x_1, x_2, \ldots, x_n$ im Intervalle $a \leqq x \leqq b$ *normal verteilt* sind.[5]

Die Formel (2) der Einleitung liefert für eine Treppenparabel $H(x)$ folgende Darstellung:

$$H(x) = H(x_1)h_1(x) + H(x_2)h_2(x) + \cdots + H(x_n)h_n(x), \tag{2}$$

die natürlich für jeden reellen (oder komplexen) Wert von x gültig ist.

Da für jedes x die Identität $h_1(x) + h_2(x) + \cdots + h_n(x) = 1$ gültig ist und für $a < x < b$ die Ungleichungen $h_1(x) \geqq 0, h_2(x) \geqq 0, \ldots, h_n(x) \geqq 0$ stattfinden, so entnimmt man aus der Formel (2) folgenden Satz:

Wenn die Stellen $x_1, x_2, \ldots, x_n$ des Intervalls $a \leqq x \leqq b$ in diesem Intervalle normal verteilt sind, so kann eine zu den Stellen $x_1, x_2, \ldots, x_n$ gehörige Treppenparabel $H(x)$ im *ganzen* Intervalle $a \leqq x \leqq b$ keinen größeren bzw. keinen kleineren Wert annehmen als den größten der n Werte

$$H(x_1), H(x_2), \ldots, H(x_n) \tag{3}$$

bzw. den kleinsten dieser n Werte.

(Man sucht also das Maximum und das Minimum der Treppenparabel $(2n-1)$-ten Grades im *Kontinuum* $a \leqq x \leqq b$ und findet diese schon *unter den n Werten* (3).)

[5] Ist $\omega''(x_k) = 0$, d. h. hat die Stammparabel $\omega(x)$ an der Stelle x_k, wo sie die x-Axe schneidet, einen Inflexionspunkt, so ist die Forderung (1) für $x = x_k$ natürlich „erfüllt". Wir sagen in diesem Falle, daß $X_k = \infty$.

2. Wir wollen nun wieder die Summe der Grundpolynome $\mathfrak{H}(x) = \sum_{k=1}^{n} \mathfrak{h}_k(x)$ betrachten, diesmal aber die Stellen x_k der Beschränkung unterwerfen, daß sie sämtlich in ein Intervall $a \leqq x \leqq b$ fallen, in welchem sie *normal* verteilt sind.

Betrachten wir das Polynom

$$\mathfrak{H}(x) - x. \tag{4}$$

Wir können etwa $a = -1$, $b = 1$ annehmen. Dieses Polynom nimmt, nach der Definition von $\mathfrak{H}(x)$, an den Stellen $x_1, x_2, \ldots, x_n$, der Reihe nach, die Werte $-x_1, -x_2, \ldots, -x_n$ an, und ihre Ableitung ist an denselben Stellen gleich 0. Also ist $\mathfrak{H}(x) - x$ ein Treppenpolynom, die zu einer Abszissengruppe $x_1, x_2, \ldots, x_n$ gehört, die im Intervalle $-1 \leqq x \leqq +1$ normal verteilt ist. Das soeben erwähnte allgemeine Lemma über die Treppenpolynome liefert also in diesem speziellen Falle:

$$\operatorname*{Min}_{k=1,2,\ldots,n} \{-x_k\} \leqq \operatorname*{Min}_{-1 \leqq x \leqq +1} \{\mathfrak{H}(x) - x\} \leqq \operatorname*{Max}_{-1 \leqq x \leqq +1} \{\mathfrak{H}(x) - x\} \leqq \operatorname*{Max}_{k=1,2,\ldots,n} \{-x_k\},$$

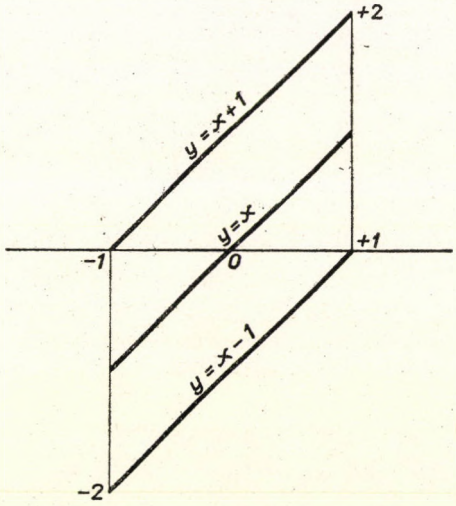
also

$$-1 \leqq \mathfrak{H}(x) - x \leqq 1,$$

und endlich

$$x - 1 \leqq \mathfrak{H}(x) \leqq x + 1,$$

somit den



Figur 2.

SATZ 3. *Wenn $x_1, x_2, \ldots, x_n$ im Intervalle $-1 \leqq x \leqq 1$ normal verteilt sind, so ist für die Summe der Grundpolynome zweiter Art:*

$$x - 1 \leqq \mathfrak{h}_1(x) + \mathfrak{h}_2(x) + \cdots + \mathfrak{h}_n(x) \leqq x + 1 \tag{5}$$

im ganzen Intervalle $-1 \leqq x \leqq +1$ gültig.

Aus (5) folgt natürlich

$$\left|\sum_{k=1}^{n} \mathfrak{h}_k(x)\right| \leqq 2 \quad \text{für} \quad -1 \leqq x \leqq +1. \tag{6}$$

Diese Ungleichung ist für *jede normale* Interpolationsgruppe und für *jeden* Wert von n ($n = 1, 2, 3, \ldots$) gültig. Sie läßt sich nicht verbessern, wie das folgende Beispiel zeigt. Es sei $a = -1$, $b = +1$, $n = 1$, und die einzige Interpolationsstelle, die ich mit x_1 bezeichne, möge in -1 liegen, d. h. $x_1 = -1$. Da jetzt $h_1(x) = 1$, $\mathfrak{h}_1(x) = x + 1$, so kann ich also in diesem Falle behaupten, daß einerseits $x_1 = -1$ eine normale Verteilung repräsentiert, daß andererseits $|\mathfrak{h}_1(1)| = \mathfrak{h}_1(1) = 1 + 1 = 2$. (S. Figur 2.)

§ 5. Einige Formeln für die Grundpolynome, tabellarisch zusammengestellt

In diesem Paragraphen möchte ich einige, auf die Grundpolynome der Lagrangeschen und Hermiteschen Interpolation bezüglichen Formeln tabellarisch zusammenstellen. Nicht von allen Formeln habe ich in dieser Note Gebrauch gemacht, andererseits fehlen einige, die im Texte zur Anwendung kamen. Dem Gegenstande etwas ferner stehenden kann sie vielleicht doch nützlich sein.

Tabelle 1

$$x_1, x_2, \ldots, x_n \tag{1}$$
$$\omega(x) = (x - x_1)(x - x_2) \cdots (x - x_n)$$

$$l_k(x) = \frac{\omega(x)}{\omega'(x_k)\,(x - x_k)} \tag{2}$$

$$h_k(x) = (1 - 2\,l_k'(x_k)\,(x - x_k))\,(l_k(x))^2 \tag{3}$$

$$\mathfrak{h}_k(x) = (x - x_k)\,(l_k(x))^2 \tag{4}$$

Tabelle 2

$$h_k(x) = \left(1 - \frac{\omega''(x_k)}{\omega'(x_k)}\,(x - x_k)\right)(l_k(x))^2 \tag{3'}$$

$$\mathfrak{h}_k(x) = (x - x_k)\,(l_k(x))^2 \tag{4'}$$

Tabelle 3

$$(1'') \quad \omega(x) = 0 + \omega'(x_k)(x-x_k) + \frac{1}{2}\omega''(x_k)(x-x_k)^2 + \frac{1}{6}\omega'''(x_k)(x-x_k)^3 + \cdots$$

$$(2'') \quad l_k(x) = 1 + \frac{1}{2}\frac{\omega''(x_k)}{\omega'(x_k)}(x-x_k) + \frac{1}{6}\frac{\omega'''(x_k)}{\omega'(x_k)}(x-x_k)^2 + \cdots$$

$$(2''') \quad (l_k(x))^2 = 1 + \frac{\omega''(x_k)}{\omega'(x_k)}(x-x_k) + \left[\frac{1}{3}\frac{\omega'''(x_k)}{\omega'(x_k)} + \frac{1}{4}\left(\frac{\omega''(x_k)}{\omega'(x_k)}\right)^2\right](x-x_k)^2 + \cdots$$

$$(3'') \quad h_k(x) = 1 + 0 - \frac{A_k}{12}(x-x_k)^2 + \cdots,$$

wo

$$A_k = 9\left(\frac{\omega''(x_k)}{\omega'(x_k)}\right)^2 - 4\frac{\omega'''(x_k)}{\omega'(x_k)}.$$

Bezeichnen

$$\xi_1, \xi_2, \ldots, \xi_{n-1}$$

die Nullstellen von $\frac{d\omega(x)}{dx}$, so ist auch

$$(3''') \quad h_k(x) = 1 - \frac{1}{12}\left[5\left(\sum_{m=1}^{n-1}\frac{1}{x_k-\xi_m}\right)^2 + 4\sum_{m=1}^{n-1}\frac{1}{(x_k-\xi_m)^2}\right](x-x_k)^2 + \cdots$$

gültig.

(Eingegangen am 1. Juli 1955.)

НЕСКОЛЬКО ЭЛЕМЕНТАРНЫХ ЗАМЕЧАНИЙ О ФУНДАМЕНТАЛЬНЫХ МНОГОЧЛЕНАХ ПАРАБОЛИЧЕСКОГО ИНТЕРПОЛИРОВАНИЯ

Л. Фейер (Будапешт)

(Резюме)

В настоящей статье сообщаются некоторые данные о поведении фундаментальных многочленов лагранжева и эрмитова интерполирования. В введении и в первых трех параграфах абсциссы интерполирования $x_1, x_2, \ldots, x_n$ считаются произвольными, в четвертом параграфе — нормальными относительно некоторого отрезка (a, b) $(a \leqq x_1 < x_2 < \cdots < x_n \leqq b)$. В пятом параграфе приведена таблица формул. Из результатов упоминаются здесь лишь три.

1. Если $h_k(x)$ обозначает какой-нибудь фундаментальный многочлен первого порядка эрмитова интерполирования, то разложение этого многочлена степени не выше $2n-1$ по степеням $x-x_k$ начинается так:

$$h_k(x)=\left[1-\frac{\omega''(x_k)}{\omega'(x_k)}(x-x_k)\right](l_k(x))^2=$$

$$=1-\frac{1}{12}\left[5\left(\sum_{\nu=1}^{n-1}\frac{1}{x_k-\xi_\nu}\right)^2+4\sum_{\nu=1}^{n-1}\frac{1}{(x_k-\xi_\nu)^2}\right](x-x_k)^2+\cdots,$$

где $x_1, x_2, \ldots, x_n$ суть не совпадающие друг с другом, вещественные, а во всем остальном произвольные узлы интерполирования, $\omega(x)=(x-x_1)(x-x_2)\cdots(x-x_n)$ и $\xi_1, \xi_2, \ldots, \xi_{n-1}$ — корни уравнения $\omega'(x)=0$. Отсюда следует, что фундаментальный многочлен первого порядка $h_k(x)$ имеет локальный максимум в точке $x=x_k$, в которой он равен единице ($k=1,2,\ldots,n$).

2. Известно, что фундаментальные многочлены второго порядка эрмитова интерполирования $\mathfrak{h}_1(x), \mathfrak{h}_2(x), \ldots, \mathfrak{h}_n(x)$ равны нулю во всех узлах интерполирования $x_1, x_2, \ldots, x_n$. Следовательно, там исчезает и их сумма

$$\mathfrak{h}_1(x)+\mathfrak{h}_2(x)+\cdots+\mathfrak{h}_n(x)=\mathfrak{H}(x).$$

В статье отмечается тот факт, что все корни этого многочлена (в точности) $2n-1$-ой степени — вещественны.

3. Если в 1 и 2 речь шла о любых простых вещественных абсциссах, то пусть теперь $-1\leqq x_1<x_2<\cdots<x_n\leqq 1$ и пусть эти абсциссы нормально распределяются на отрезке $-1\leqq x\leqq 1$. Тогда на этом отрезке выполняется неравенство

$$x-1\leqq\mathfrak{H}(x)\leqq x+1$$

и, следовательно,

$$|\mathfrak{H}(x)|\leqq 2 \qquad (-1\leqq x\leqq 1).$$

В этом неравенстве 2 не может быть заменено меньшим числом, если мы не хотим ущемлять общности.

ON SOME NEW THEOREMS IN THE THEORY OF DIOPHANTINE APPROXIMATIONS

By

VERA T. SÓS (Budapest) and P. TURÁN (Budapest), member of the Academy

1. In a recent book[1] one of us based a series of applications on the following three theorems.

I. If

$$|z_1| \geqq |z_2| \geqq \cdots \geqq |z_n|,$$

m is a non-negative integer, the b_j's are arbitrary complex numbers, there is then an integer ν_1 such that

$$m+1 \leqq \nu_1 \leqq m+n$$

and

$$|b_1 z_1^{\nu_1} + b_2 z_2^{\nu_1} + \cdots + b_n z_n^{\nu_1}| \geqq |z_n|^{\nu_1} \left(\frac{n}{2e(m+n)}\right)^n |b_1 + \cdots + b_n|.$$

II. With the above notations there is an integer ν_2 such that

$$m+1 \leqq \nu_2 \leqq m+n$$

and

$$|b_1 z_1^{\nu_2} + b_2 z_2^{\nu_2} + \cdots + b_n z_n^{\nu_2}| \geqq |z_1|^{\nu_2} \left(\frac{n}{24e^2(m+2n)}\right)^n \min_{j=1,\ldots,n} |b_1 + \cdots + b_j|.$$

III. With the above notations and

$$\omega(z) = \prod_{\nu=1}^{n} (z - z_\nu)$$

there is an integer ν_3 such that

$$m+1 \leqq \nu_3 \leqq m+n$$

and

$$|b_1 z_1^{\nu_3} + b_2 z_2^{\nu_3} + \cdots + b_n z_n^{\nu_3}| \geqq$$

$$\geqq \left(\sum_{j=1}^{n} |b_j| |z_j|^{\nu_3}\right) \cdot \prod_{l=1}^{n} \frac{1}{1+|z_l|} \cdot \frac{1}{\max\limits_{k=0,1,\ldots,(n-1)} \sum\limits_{l=1}^{n} \frac{|z_l|^k}{|\omega'(z_l)|(1+|z_l|)}}.$$

In mentioned book[1] **III** was discussed mainly as a matter of orientation and had only one application in the investigation of integral functions of type

$$\sum_{\nu=1}^{n} a_\nu F(c_\nu z)$$

[1] P. TURÁN, *Eine neue Methode in der Analysis und deren Anwendungen* (Budapest, 1953), Akadémiai Kiadó.

where $F(z)$ is an integral function. Since that time the second of us found some more applications of it. To show the essence of these three Theorems **I, II, III** we call the quantities

$$(1.1) \qquad \sum_{j=1}^{n} |b_j| |z_j|^r, \qquad |z_n|^r, \qquad |z_1|^r$$

$N_l(r,f)$ norms of $f(r) = \sum_{j=1}^{n} b_j z_j^r$ $(l = 1, 2, 3)$. Then the Theorems **I, II** and **III** can be expressed by saying that, for a suitable integer r, $\left| \sum_{j=1}^{n} b_j z_j^r \right|$ is estimated from below by the $N_l(r,f)$ norms (1.1) so, that the lower estimation of their quotient should be independent

(1.2) a) of the z_j-values

or

b) of the b_j coefficients.

Theorems **I** and **II** are of a)-type, Theorem **III** is of b)-type. This formulation of the theory is more symmetrical than that given in [1], where only problems of a)-type were *systematically* treated. In connection with an application [2] the necessity of dual theorems emerged where *non-trivial upper* estimations of

$$\min_{\substack{m+1 \leq r \leq m+n \\ r \text{ integer}}} \frac{|f(r)|}{N_l(r,f)}$$

are needed even at the rate of simple geometrical restrictions on the z_j's. In [1] one can find detailed motivation, on which way Theorems **I, II** and **III** can be considered as generalisations of KRONECKER's and DIRICHLET's classical theorems in the theory of diophantine approximations.

2. In [1] the emphasis was laid upon the applicability of these theorems and no care was taken to *best-possible* inequalities, though these have a significance for some applications, too. One can show this e. g. on the estimation [3] of $N(\alpha, T)$, the number of zeros of $\zeta(s)$ $(s = \sigma + it)$ in the parallelogram $\sigma \geq \alpha$, $0 < t \leq T$

$$(2.1) \qquad N(\alpha, T) = O(T^{2(1-\alpha)+(1-\alpha)^{1,11}}),$$

which is uniformly valid for

$$(2.2) \qquad 1 - \delta \leq \alpha \leq 1$$

with a (small) numerical positive δ. This constitutes the best-known estima-

[2] P. TURÁN, On Lindelöf's conjecture, *Acta Math. Acad. Sci. Hung.*, **5** (1954), pp. 145—163.

[3] This is an unpublished sharpening of Theorem XXXVIII of [1].

tion in this range today. The proof of this is largely based on the case

(2. 3) $$b_1 = b_2 = \cdots = b_n = 1$$

of Theorem **II**; thereby the decrease of the numerical factor $24e^2$, since n is "large" in this case, would result an increase of δ in (2. 2) which in turn would result a decrease of the smallest known Θ with the property

$$p_{n+1} - p_n = O(p_n^{\Theta})$$

where p_n denotes the nth prime. The main aim of this paper is to review the a)-type results in the first part of [1] from this point of view, in particular in the case (2. 3), and to study certain $(z_1, z_2, \ldots, z_n)$-systems, which will play a role in these questions. We suppose without loss of generality

(2. 4) $$z_n = 1, \qquad |z_j| \geqq 1 \qquad (j = 1, \ldots, n)$$

in **I** and

(2. 5) $$z_1 = 1, \qquad |z_j| \leqq 1 \qquad (j = 1, \ldots, n)$$

in **II**. In the case (2. 5) we ask for the "smallest" numerical positive value A_1, for which

(2. 6) $$\max_{\substack{m+1 \leqq \nu \leqq m+n \\ \nu \text{ integer}}} |z_1^{\nu} + z_2^{\nu} + \cdots + z_n^{\nu}| \geqq \left(\frac{n}{A_1(m+n)}\right)^n$$

holds for an arbitrary non-negative integer m and positive integer n. We shall show in **6—9** the following

THEOREM. *We have for the* A_1 *defined in* (2. 6) *the inequality*

$$1{,}321 < A_1 < 2e^{1+\frac{4}{e}} (< 24).$$

The gap is still large but the upper bound is much better than the previous $24e^2 \sim 177$. That $A_1 > 1{,}1$, can already be shown taking

$$m = 0, \quad n = 2, \quad z_1 = 1, \quad z_2 = \frac{1}{2} e^{\frac{2\pi i}{3}};$$

in this case we have

$$|z_1 + z_2| = \frac{\sqrt{3}}{2}, \qquad |z_1^2 + z_2^2| = \frac{\sqrt{13}}{4},$$

i. e.

$$\frac{\sqrt{13}}{4} \geqq \frac{1}{A_1^2}, \quad A_1 \geqq \frac{2}{\sqrt[4]{13}} > 1{,}1.$$

An interesting feature of the proof of the upper bound is the avoiding of the use of H. CARTAN's theorem and replacing it by a lemma of CHEBYSEV-type. The new proof furnishes mutatis mutandis the following improvement of **II**.

Under the conditions of **II** we have

$$(2.7)\quad |b_1 z_1^{\nu_2} + \cdots + b_n z_n^{\nu_2}| \geqq |z_1|^{\nu_2} \left(\frac{n}{2e^{1+\frac{4}{e}}(m+n)} \right)^n \cdot \min_{j=1,\ldots,n} |b_1 + \cdots + b_j|.$$

It will be sufficient to indicate only the changes, the proof of (2. 7) needs compared to that of the theorem.

3. A further refinement of the upper estimation in the Theorem would be given if the constant $2e$ in **I** could be diminished, even only in the special case $b_1 = \cdots = b_n = 1$. Asking for the "smallest" numerical positive value A_2, for which in the case (2. 4)

$$(3.1)\quad \max_{\substack{m+1 \leqq \nu \leqq m+n \\ \nu \text{ integer}}} |z_1^\nu + z_2^\nu + \cdots + z_n^\nu| \geqq \left(\frac{n}{A_2(m+n)} \right)^n$$

holds for any positive integer n and non-negative integer m, nothing better than

$$(3.2)\quad 1 \leqq A_2 \leqq 2e$$

can be asserted at the present. Something better can be said on the "smallest" numerical positive value A_3 for which in the case (2. 4)

$$(3.3)\quad \max_{\substack{m+1 \leqq \nu \leqq m+n \\ \nu \text{ integer}}} |b_1 z_1^\nu + \cdots + b_n z_n^\nu| \geqq \left(\frac{n}{A_3(m+n)} \right)^n |b_1 + \cdots + b_n|$$

holds for any complex b_ν's, positive integer n and non-negative integer m. We shall show in **10** that

$$(3.4)\quad 1,27 \sim \frac{4}{\pi} \leqq A_3 \leqq 2e \sim 5,44.$$

4. The lower limitation of A_1 in the Theorem will be proved in **9** by refining an idea of P. ERDŐS, i. e. considering $(z_1, z_2, \ldots, z_n)$-systems with the property

$$(4.1)\quad s_2 = s_3 = \cdots = s_{n-1} = 0, \quad z_1 = 1$$

where s_ν stands for $z_1^\nu + z_2^\nu + \cdots + z_n^\nu$. This suggests for the sake of counter examples the usefulness of the study of all $(z_1, \ldots, z_n)$-systems with the property

$$s_2 = s_3 = \cdots = s_{n-1} = s_n = 0,$$

or more generally with a prescribed non-negative integer m the determination of all those with

$$(4.2)\quad s_{m+1} = s_{m+2} = \cdots = s_{m+n-1} = 0.$$

We mention another reason why (4. 2) is interesting. The whole theory emerged from the necessity to diminish the interval for ν_j in **I**, **II** and **III** as much as

possible; the question arises now, whether or not for *some* integer values m there is a non-trivial lower limitation for

$$\max_{\substack{m+1 \leqq \nu \leqq m+n-1 \\ \nu \text{ integer}}} \frac{|s_\nu|}{|z_1|^\nu}$$

or for

$$\max_{\substack{m+1 \leqq \nu \leqq m+n-1 \\ \nu \text{ integer}}} \frac{|s_\nu|}{|z_n|^\nu}.$$

That such a reduction of the interval for ν is *generally* impossible, is trivial, since in the case $m \equiv 0 \bmod n$ we have for the nth roots of unity

$$s_{m+1} = s_{m+2} = \cdots = s_{m+n-1} = 0.$$

Probably the same holds for *all* non-negative integer values m. As the Newton—Girard formulae show at once, for $m = 0$ *all the* $(z_1, \ldots, z_n)$-systems with

$$s_1 = s_2 = \cdots = s_{n-1} = 0 \tag{4.3}$$

are given by the zeros of an equation

$$z^n + a = 0 \qquad (a \text{ arbitrary complex}). \tag{4.4}$$

We can determine all systems satisfying (4.2) with $m = 1$ and $m = 2$. For $m = 1$ we assert that *all* $(z_1, z_2, \ldots, z_n)$*-systems with the property*

$$s_2 = s_3 = \cdots = s_n = 0 \tag{4.5}$$

are formed by the zeros of an equation

$$\varphi_n(z, a) = z^n + \frac{a}{1!} z^{n-1} + \cdots + \frac{a^n}{n!} = 0 \qquad (a \text{ arbitrary complex}). \tag{4.6}$$

An asymptotical determination of these systems for fixed a and $n \to \infty$ follows at once from SZEGŐ's[4] results. For $m = 2$ we shall see in **11** that *all the* $(z_1, z_2, \ldots, z_n)$*-systems with*

$$s_3 = s_4 = \cdots = s_n = s_{n+1} = 0 \tag{4.7}$$

are formed by the zeros of an equation

$$f_n(z, a, \lambda) = z^n + \frac{H_1(\lambda)}{1!} a z^{n-1} + \cdots + \frac{H_n(\lambda)}{n!} a^n = 0, \tag{4.8}$$

where $H_\nu(y)$ *stands for the* ν^{th} *Hermite polynomial defined by*

$$H_\nu(y) = (-1)^\nu e^{y^2} \frac{d^\nu}{dy^\nu} e^{-y^2},$$

λ *denotes any zero of* $H_{n+1}(y) = 0$ *and* a *is an arbitrary complex number.* An asymptotical determination of the zeros of $f_n(z, a, \lambda)$ is not known. The

[4] G. SZEGŐ, Über eine Eigenschaft der Exponentialreihe, *Sitzungsber. der Berl. Math. Ges.*, **23** (1924), pp. 50—63.

result (4. 5)—(4. 6) can be easily[5] proved by Newton—Girard formulae so we omit the details. It is interesting to remark a characteristic difference between the cases $m=1$ and $m=2$. All solutions of (4. 5) can be derived from the single equation

$$1+\frac{z}{1!}+\frac{z^2}{2!}+\cdots+\frac{z^n}{n!}=0,$$

all solutions of (4. 7), however, can be derived from the $(n+1)$ equations

$$\sum_{\nu=0}^{n}\frac{H_\nu(\lambda)}{\nu!}z^\nu=0 \quad \text{where} \quad H_{n+1}(\lambda)=0.$$

5. Owing to the applicability to the approximative solution of algebraic equations it is of interest to study **II** in the special case

(5. 1) $$m=0, \qquad b_1=b_2=\cdots=b_n=1.$$

In this case we have to determine

(5. 2) $$M_n=\min_{\substack{|z_j|\leqq 1\\ j=1,\ldots,n}}\ \max_{\nu=1,\ldots,n}|1+z_2^\nu+\cdots+z_n^\nu|.$$

In[1] we have shown

(5. 3) $$M_n\geqq\frac{\log 2}{\frac{1}{1}+\frac{1}{2}+\cdots+\frac{1}{n}};$$

according to a written communication of DE BRUIJN this can be replaced by

(5. 4) $$M_n>c\,\frac{\log\log n}{\log n} \qquad (c \text{ numerical positive constant}),$$

what is for large n somewhat better. This makes still more probable the conjecture that $M_n\geqq d$, independently of n. As to the upper estimation of M_n MR. HYLTÉN-CAVALLIUS showed $M_2=\frac{\sqrt{5}-1}{\sqrt{2}}\sim 0{,}86$ and found by considering $z_2=0{,}1295+i\,0{,}7063$, $z_3=-0{,}5128+i\ 0{,}1508$ the estimation

$$M_3<0{,}831.$$

6. Before turning to the proof of the Theorem we need the following simple

LEMMA. *Let be* $0\leqq\delta\leqq 1$ *and*

(6. 1) $$f(z)=z^n+\cdots=\prod_{\nu=1}^{n}(z-z_\nu).$$

[5] Another very elegant verification is due to E. EGERVÁRY. Using Euler's formulae $\sum_{j=1}^{n}\frac{z_j^\nu}{\omega'(z_j)}=0$ $(\nu=0,1,\ldots,n-2)$ valid for any polynomial $\omega(z)$ with simple z_j-zeros of degree n and using $\varphi_n'(z_j,a)=(n!\,z_j^2)^{-1}a^{n+1}$ the result follows at once.

Then, there is a circle $|z|=r_0$ *with* $\delta \leqq r_0 \leqq 1$ *such that on the whole periphery*

$$|f(z)| \geqq 2\left(\frac{1-\delta}{4}\right)^n.$$

For the proof we introduce

$$f^*(z)=\prod_{\nu=1}^{n}(z-|z_\nu|). \tag{6.2}$$

Owing to the classical theorem of CHEBYSEV there is a ξ with $\delta \leqq \xi \leqq 1$ and

$$|f^*(\xi)| \geqq 2\left(\frac{1-\delta}{4}\right)^n. \tag{6.3}$$

But on the circle $|z|=\xi$ we have

$$|z-z_j| \geqq \big||z|-|z_j|\big| = |\xi-|z_j|| \qquad (j=1,2,\ldots,n),$$

i. e. multiplying, further using (6.2) and (6.3) we get

$$|f(z)| \geqq |f^*(\xi)| \geqq 2\left(\frac{1-\delta}{4}\right)^n.$$

Hence the above ξ can be chosen as r_0. Q. e. d.

7. Next we turn to the proof of our Theorem. We may suppose that $n \geqq 2$. Let be, with our numbers z_j,

$$\omega(z)=\prod_{j=1}^{n}(z-z_j) \tag{7.1}$$

and let δ be a positive number $0 \leqq \delta \leqq 1$ to be determined later. To this $\omega(z)$ and δ there is, according to the Lemma, an r_0 with $\delta \leqq r_0 \leqq 1$ and such that for $|z|=r_0$

$$|\omega(z)| \geqq 2\left(\frac{1-\delta}{4}\right)^n. \tag{7.2}$$

Since each factor $|z-z_j|$ is at most 2, we have from (7.2) for $|z|=r_0$ and for any choice of the indices $(1\leqq)\ i_1<i_2<\cdots<i_k \leqq n$ $(1 \leqq k \leqq n)$

$$\prod_{\nu=1}^{k}|z-z_{i_\nu}| \geqq 2\left(\frac{1-\delta}{4}\right)^n \frac{1}{2^{n-k}} = 2\left(\frac{1-\delta}{8}\right)^n 2^k. \tag{7.3}$$

We investigate two cases.

Case a) All z_j's are absolutely $\geqq r_0$. Then, owing to **I**, there is an integer ν_1 with $m+1 \leqq \nu_1 \leqq m+n$ and

$$|z_1^{\nu_1}+z_2^{\nu_1}+\cdots+z_n^{\nu_1}| \geqq r_0^{\nu_1}\left(\frac{n}{2e(m+n)}\right)^n \geqq \delta^{m+n}\left(\frac{n}{2e(m+n)}\right)^n. \tag{7.4}$$

Case b) There is an index l with $1 \leqq l < n$ and

$$1=|z_1| \geqq |z_2| \geqq \cdots \geqq |z_l| > r_0 > |z_{l+1}| \geqq \cdots \geqq |z_n|. \tag{7.5}$$

In the treatment of this case we shall suppose first that the z_j's are all different.

8. Let first be

$$f_1(z)=\prod_{j=l+1}^{n}(z-z_j)=\sum_{j=0}^{n-l}c_j^{(1)}z^{n-l-j},\qquad c_0^{(1)}=1. \tag{8.1}$$

We have obviously

$$|c_j^{(1)}|=\left|\sum_{l+1\leqq i_1<\cdots<i_j\leqq n}z_{i_1}z_{i_2}\cdots z_{i_j}\right|\leqq\binom{n-l}{j}. \tag{8.2}$$

Next let $f_2(z)$ be that polynomial of degree $\leqq l-1$, which assumes for $z=z_1, z_2, \ldots, z_l$ the values

$$\frac{1}{z_1^{m+1}f_1(z_1)},\quad \frac{1}{z_2^{m+1}f_1(z_2)},\quad \ldots,\quad \frac{1}{z_l^{m+1}f_1(z_l)},$$

respectively. If $l=1$, then

$$f_2(z)\equiv\frac{1}{z_1^{m+1}f_1(z_1)}\equiv c_0^{(2)};$$

if $1<l<n$, then we can represent $f_2(z)$ as Newton-interpolatorical polynomial

$$f_2(z)=c_0^{(2)}+c_1^{(2)}(z-z_1)+c_2^{(2)}(z-z_1)(z-z_2)+\cdots+c_{l-1}^{(2)}(z-z_1)(z-z_2)\cdots(z-z_{l-1}).$$

Since the function $\dfrac{1}{z^{m+1}f_1(z)}$ is regular for $|z|>r_0$ and vanishes for $z=\infty$, we have according to NÖRLUND's representation

$$c_j^{(2)}=\frac{1}{2\pi i}\int_{|w|=r_0}\frac{dw}{w^{m+1}f_1(w)(w-z_1)(w-z_2)\cdots(w-z_{j+1})} \tag{8.3}$$

$$(j=0,1,\ldots,(l-1)).$$

But $f_1(w)(w-z_1)\cdots(w-z_{j+1})$ is of type (7.3) with $k=n-l+j+1$ and thus owing to (7.3) we have for $|w|=r_0$

$$|f_1(w)(w-z_1)(w-z_2)\cdots(w-z_{j+1})|\geqq 2\left(\frac{1-\delta}{8}\right)^n 2^{n-l+j+1}=2\left(\frac{1-\delta}{4}\right)^n\frac{1}{2^{l-j-1}},$$

i. e. from (8.3)

$$|c_j^{(2)}|\leqq\frac{1}{2r_0^m}\left(\frac{4}{1-\delta}\right)^n 2^{l-j-1}\leqq\frac{1}{2\delta^m}\left(\frac{4}{1-\delta}\right)^n 2^{l-j-1}. \tag{8.4}$$

But we need also $f_2(z)$ in the form

$$f_2(z)=\sum_{j=0}^{l-1}c_j^{(3)}z^j$$

and we have to estimate the coefficients $c_j^{(3)}$. We have

$$c_{l-1}^{(3)}=c_{l-1}^{(2)}$$

and if $l>1$, for $j=0, 1, \ldots, (l-2)$

$$c_j^{(3)} = c_j^{(2)} - c_{j+1}^{(2)} \sum_{1 \leqq r_1 \leqq j+1} z_{r_1} + c_{j+2}^{(2)} \sum_{1 \leqq r_1 < r_2 \leqq j+2} z_{r_1} z_{r_2} - \cdots +$$

$$+ (-1)^{l-j-1} c_{l-1}^{(2)} \sum_{1 \leqq x_1 < r_2 < \ldots < r_{l-j-1} \leqq l-1} z_{r_1} z_{r_2} \cdots z_{r_{l-j-1}},$$

i. e.

$$|c_j^{(3)}| \leqq |c_j^{(2)}| + |c_{j+1}^{(2)}| \binom{j+1}{1} + |c_{j+2}^{(2)}| \binom{j+2}{2} + \cdots + |c_{l-1}^{(2)}| \binom{l-1}{l-j-1}.$$

Thus from (8. 4)

$$|c_j^{(3)}| \leqq \frac{2^{l-j-1}}{2\delta^m} \left(\frac{4}{1-\delta}\right)^n \left\{ 1 + \binom{j+1}{1} \frac{1}{2} + \binom{j+2}{2} \frac{1}{2^2} + \cdots + \binom{l-1}{l-1-j} \frac{1}{2^{l-j-1}} \right\} <$$

$$(8.5) \qquad < \frac{2^{l-j-1}}{2\delta^m} \left(\frac{4}{1-\delta}\right)^n \sum_{d=0}^{\infty} \binom{j+d}{d} \frac{1}{2^d} =$$

$$= \frac{2^{l-j-1}}{2\delta^m} \left(\frac{4}{1-\delta}\right)^n \sum_{d=0}^{\infty} (-1)^d \binom{-j-1}{d} \frac{1}{2^d} = \frac{2^{l-1}}{\delta^m} \left(\frac{4}{1-\delta}\right)^n.$$

Let finally be

$$(8.6) \qquad f_3(z) = z^{m+1} f_1(z) f_2(z) = \sum_{j=m+1}^{m+n} c_j^{(4)} z^j.$$

It follows from the definition of $f_1(z)$ and $f_2(z)$ that

$$(8.7) \qquad f_3(z_1) = f_3(z_2) = \cdots = f_3(z_l) = 1, \quad f_3(z_{l+1}) = \cdots = f_3(z_n) = 0.$$

Replacing z in (8. 6) by z_j and [6] summing for $j=1, 2, \ldots, n$ we get, using (8. 7) and writing

$$(8.8) \qquad s_\nu = z_1^\nu + z_2^\nu + \cdots + z_n^\nu,$$

the identity

$$\left| \sum_{\nu=m+1}^{m+n} c_\nu^{(4)} s_\nu \right| = \sum_{\nu=m+1}^{m+n} c_\nu^{(4)} s_\nu = l,$$

i. e.

$$(8.9) \qquad l \leqq \max_{\substack{m+1 \leqq \nu \leqq m+n \\ \nu \text{ integer}}} |s_\nu| \left(\sum_{j=m+1}^{m+n} |c_j^{(4)}| \right).$$

But from (8. 6) we have

$$c_j^{(4)} = \sum_{n-l-j_1+j_2=j-m-1} c_{j_1}^{(1)} c_{j_2}^{(3)},$$

i. e.

$$\sum_{j=m+1}^{m+n} |c_j^{(4)}| \leqq \left(\sum_{j_1=0}^{n-l} |c_{j_1}^{(1)}| \right) \left(\sum_{j_2=0}^{l-1} |c_{j_2}^{(3)}| \right);$$

[6] If we want to prove (2. 7) instead of the upper estimation of the Theorem, we have also to multiply by b_j and then to sum for j.

thus, using (8. 2) and (8. 5), we get

$$\sum_{j=m+1}^{m+n} |c_j^{(4)}| \leqq 2^{n-l} l \frac{2^{l-1}}{\delta^m} \left(\frac{4}{1-\delta}\right)^n = \frac{l}{2\delta^m} \left(\frac{8}{1-\delta}\right)^n.$$

Then we have from (8. 9)

$$\max_{\substack{m+1 \leqq \nu \geqq m+n \\ \nu \text{ integer}}} |s_\nu| \geqq 2\delta^m \left(\frac{1-\delta}{8}\right)^n. \tag{8. 10}$$

Now we choose δ so, that

$$\delta^{m+n} \left(\frac{n}{2e(m+n)}\right)^n = \delta^m \left(\frac{1-\delta}{8}\right)^n,$$

i. e.

$$\delta = \frac{1}{1+\frac{4}{e}\frac{n}{m+n}} \equiv \delta_0;$$

then from (7. 4) and (8. 10) we got

$$\max_{\nu=m+1,\, m+2,\, \ldots,\, m+n} |z_1^\nu + z_2^\nu + \cdots + z_n^\nu| \geqq 2\delta_0^{m+n} \left(\frac{n}{2e(m+n)}\right)^n.$$

But

$$\left(1+\frac{4n}{e(m+n)}\right)^{m+n} \leqq e^{\frac{4n}{e}}, \quad \text{i. e.} \quad \delta_0^{m+n} \geqq e^{-\frac{4n}{e}}$$

whence

$$\max_{\nu=m+1,\, m+2,\, \ldots,\, m+n} |z_1^\nu + \cdots + z_n^\nu| \geqq 2\left(\frac{n}{2e^{1+\frac{4}{e}}(m+n)}\right)^n,$$

i. e. our Theorem is for different z_j's proved. We can get rid of the restriction $z_\mu \neq z_\nu$ $(\mu \neq \nu)$ exactly on the same way as in the previous proof in [1] and we do not detail it.

9. Now we turn to the lower estimation in the Theorem. Let ϑ be a positive constant, less than 1 and to be determined later, let $z_1 = 1$ and $z_2, z_3, \ldots, z_n$ be determined by the conditions (with the notation (8. 8))

$$s_1 = \vartheta n, s_2 = s_3 = \cdots = s_{n-1} = 0. \tag{9. 1}$$

It is well known that these conditions determine uniquely the numbers z_j $(j = 2, 3, \ldots, n)$; let us denote them by ζ_ν $(\zeta_1 = 1,\ \nu = 2, 3, \ldots, n)$. Let further be

$$f(z) = z^n + a_1 z^{n-1} + \cdots + a_n$$

the polynomial with the zeros ζ_n $(\nu = 1, \ldots, n)$. Then the Newton—Girard formulae give successively

$$a_1 = -s_1, a_2 = \frac{s_1^2}{2!}, a_3 = -\frac{s_1^3}{3!}, \ldots, a_{n-1} = (-1)^{n-1} \frac{s_1^{n-1}}{(n-1)!}. \tag{9. 2}$$

Since $\zeta_1 = 1$ is a zero of $f(z)$, we have

$$a_n = -(1 + a_1 + \cdots + a_{n-1}) = -\left(1 - \frac{s_1}{1!} + \frac{s_1^2}{2!} - \cdots + (-1)^{n-1}\frac{s_1^{n-1}}{(n-1)!}\right) =$$

$$= -e^{-s_1} + \sum_{j=n}^{\infty}(-1)^j \frac{1}{j!} s_1^j. \tag{9.3}$$

Further, we have from Newton—Girard formulae

$$s_n = -a_{n-1}s_1 - n a_n = (-1)^n \frac{s_1^n}{(n-1)!} + n e^{-s_1} - \sum_{j=n}^{\infty} n(-1)^j \frac{s_1^j}{j!} =$$

$$= n\left(e^{-s_1} - \sum_{j=n+1}^{\infty}(-1)^j \frac{s_1^j}{j!}\right) = n\sum_{j=0}^{n}(-1)^j \frac{s_1^j}{j!} \tag{9.4}$$

and from (9.3) and (9.4)

$$s_{n+1} = -a_1 s_n - a_n s_1 = s_1(s_n - a_n) =$$

$$= s_1\left\{(n+1)e^{-s_1} + (-1)^{n+1}\frac{s_1^n}{n!} - (n+1)\sum_{j=n+1}^{\infty}(-1)^j \frac{s_1^j}{j!}\right\}. \tag{9.5}$$

Since $0 < \vartheta < 1$, the terms $\frac{s_1^j}{j!}$ decrease monotonically if $j \geqq n$, it follows

$$\left|\sum_{j=n+1}^{\infty}(-1)^{j+1}\frac{s_1^j}{j!}\right| < \frac{s_1^{n+1}}{(n+1)!},$$

i. e. from (9.4) and (9.5)

$$|s_n| \leqq n\left(e^{-\vartheta n} + \frac{(\vartheta n)^{n+1}}{(n+1)!}\right), \qquad |s_{n+1}| \leqq \vartheta n\left\{(n+1)e^{-\vartheta n} + \frac{(\vartheta n)^n}{n!} + \frac{(\vartheta n)^{n+1}}{n!}\right\}.$$

Since $k! > \left(\frac{k}{e}\right)^k e$, we get for $n > 10$

$$|s_n| \leqq n\left\{e^{-\vartheta n} + \frac{1}{e}(e\vartheta)^{n+1}\right\}, \qquad |s_{n+1}| \leqq (n+1)^2(e^{-\vartheta n} + 4(e\vartheta)^n)$$

whence with $m = 1$

$$\max_{m+1 \leqq \nu \leqq m+n} |s_\nu| < 4(n+1)^2\{e^{-\vartheta n} + (e\vartheta)^n\}.$$

If ϑ is the (only) positive zero α_0 of the transcendental equation

$$x = e^{-x-1}, \tag{9.6}$$

then we have $0{,}2784 \leqq \vartheta \leqq 0{,}2785$ and for our ζ_ν's

$$\max_{\substack{2 \leqq \nu \leqq n+1 \\ \nu \text{ integer}}} |s_\nu| \leqq 8(n+1)^2 e^{-\alpha_0 n}. \tag{9.7}$$

It may occur that (2.5) is not fulfilled by the ζ_ν-system, i. e. some of them are absolutely $\geqq 1$. If so, we can construct simply by contraction a ζ_ν^*-system

satisfying (2. 5) and for which (9. 7) holds a fortiori. Thus for an arbitrary small $\varepsilon > 0$ and for $n > n_0(\varepsilon)$ there is a ζ_ν-system satisfying (2. 5) such that

$$\max_{\substack{2 \leqq \nu \leqq n+1 \\ \nu \text{ integer}}} |s_\nu| \leqq e^{-(\alpha_0-\varepsilon)n}.$$

Hence, if for each integer $n \geqq 1$ and $m \geqq 0$ the estimation (2. 6) is valid, then

$$\left(\frac{n}{A_1(1+n)}\right)^n \leqq e^{(-\alpha_0+\varepsilon)n}, \quad \text{i. e.} \quad A_1 \geqq e^{\alpha_0} \sim 1{,}321,$$

indeed.

10. We are going to prove (3. 4). Obviously we need only to show

$$A_3 \geqq \frac{4}{\pi}.$$

We choose e. g. n big odd, $m = n^3$ and with a slight modification of idea of P. STEIN [7]

$$(10.\,1)\quad z_j = e^{\frac{\pi i}{2(m+n)}(2j-n-1)}, \quad b_j = \frac{1}{2^{n-1}}\binom{n-1}{j-1} e^{\frac{\pi i n}{4(m+n)}(2j-n-1)} \qquad (j = 1, 2, \ldots, n).$$

Then we get

$$f_0(y) = \sum_{j=1}^{n} b_j z_j^y = \frac{1}{2^{n-1}} \sum_{j=1}^{n} \binom{n-1}{j-1} e^{\frac{\pi i}{2(m+n)}(2j-n-1)\left(y+\frac{n}{2}\right)} =$$

$$= \left(\frac{e^{\frac{\pi i}{2(m+n)}\left(y+\frac{n}{2}\right)} + e^{-\frac{\pi i}{2(m+1)}\left(y+\frac{n}{2}\right)}}{2}\right)^{n-1} = \cos^{n-1}\left\{\frac{\pi}{2(m+n)}\left(y+\frac{n}{2}\right)\right\}.$$

Let us observe that for all sufficiently large n's

$$\left|\sum_{j=1}^{n} b_j\right| = \cos^{n-1}\frac{\pi n}{4(m+n)} = \cos^{n-1}\frac{\pi}{4(n^2+1)} > \cos^{n-1}\frac{\pi}{4n^2} > \frac{1}{1+\frac{1}{n}}.$$

Hence for $m \leqq y \leqq m+n$ we obtain

$$|f_0(y)| \leqq \left(1+\frac{1}{n}\right)\left|\sum_{j=1}^{n} b_j\right| \sin^{n-1}\frac{\pi n}{4(m+n)} < \left(1+\frac{1}{n}\right)\left(\frac{\pi}{4}\cdot\frac{n}{m+n}\right)^{n-1}\left|\sum_{j=1}^{n} b_j\right| =$$

$$= \left(1+\frac{1}{n}\right)\frac{4}{\pi}(n^2+1)\left(\frac{n}{\frac{4}{\pi}(m+n)}\right)^n \left|\sum_{j=1}^{n} b_j\right|.$$

Since for an arbitrary small $\varepsilon > 0$ we have for $n > n_0(\varepsilon)$

$$\left(1+\frac{1}{n}\right)\frac{4}{\pi}(n^2+1) < \frac{1}{(1-\varepsilon)^n},$$

we got

$$A_3 \geqq \frac{4}{\pi}(1-\varepsilon). \qquad \text{Q. e. d.}$$

[7] See J. E. LITTLEWOOD, Math. Notes (12), An inequality for a sum of cosines, *Journ. of London Math. Soc.*, **12** (1937), pp. 217—222.

11. In order to show (4. 7)—(4. 8) we take

$$s_1 = -2u, \quad s_2 = 2v^2 \qquad (v \neq 0). \tag{11. 1}$$

Then Newton—Girard's formulae give

$$a_1 = 2u = vH_1\left(\frac{u}{v}\right),$$

$$2a_2 = -s_2 - a_1 s_1 = -2v^2 + 4u^2 = v^2\left\{4\left(\frac{u}{u}\right)^2 - 2\right\} = v^2 H_2\left(\frac{u}{v}\right),$$

$$a_2 = \frac{v^2}{2!} H_2\left(\frac{u}{v}\right),$$

denoting by $H_r(y)$ the rth Hermite polynomial defined in (4. 9). Suppose we showed already

$$a_m = \frac{v^m}{m!} H_m\left(\frac{u}{v}\right) \qquad (m = 1, 2, \ldots, k) \tag{11. 2}$$

for a $k < n$. Then

$$(k+1)a_{k+1} = -(s_{k+1} + a_1 s_k + \cdots + a_k s_1)$$

and, using (4. 7) and (11. 2),

$$(k+1)a_{k+1} = -(a_{k-1}s_2 + a_k s_1) = -\frac{v^{k+1}}{(k-1)!} H_{k-1}\left(\frac{u}{v}\right)2 + \frac{v^k}{k!} H_k\left(\frac{u}{v}\right)2u =$$

$$= \frac{v^{k+1}}{k!}\left\{-2kH_{k-1}\left(\frac{u}{v}\right) + 2\frac{u}{v} H_k\left(\frac{u}{v}\right)\right\}. \tag{11. 3}$$

But as well known

$$2xH_k(x) - 2kH_{k-1}(x) = H_{k+1}(x), \tag{11. 4}$$

i. e. from (11. 3)

$$a_{k+1} = \frac{v^{k+1}}{(k+1)!} H_{k+1}\left(\frac{u}{v}\right),$$

what shows that (11. 2) is true for $m = 1, 2, \ldots, n$. Conversely these values a_r assure that $s_3 = s_4 = \cdots = s_n = 0$. For s_{n+1} we have

$$s_{n+1} = -(a_1 s_n + \cdots + a_n s_1) = -(a_{n-1}s_2 + a_n s_1),$$

i. e. from (11. 1) and (11. 2)

$$s_{n+1} = -\frac{v^{n+1}}{(n-1)!} H_{n-1}\left(\frac{u}{v}\right)2 + \frac{v^n}{n!} H_n\left(\frac{u}{v}\right)2u =$$

$$= \frac{v^{n+1}}{n!}\left\{2\frac{u}{v} H_n\left(\frac{u}{v}\right) - 2nH_{n-1}\left(\frac{u}{v}\right)\right\},$$

or from (11. 4)

$$s_{n+1} = \frac{v^{n+1}}{n!} H_{n+1}\left(\frac{u}{v}\right).$$

Since $v \neq 0$, $s_{n+1} = 0$ implies $u = \lambda v$ where

$$H_{n+1}(\lambda) = 0.$$

This proves already (4. 7)—(4. 8).

For the value s_{n+2} we have in our case

$$s_{n+2} = -(a_1 s_{n+1} + a_2 s_n + \cdots + a_n s_2) = -a_n s_2 = -2\frac{v^{n+2}}{n!} H_n(\lambda),$$

i. e.

$$\max_{\nu=3,\,4,\,\ldots,\,(n+2)} |s_\nu| = \frac{2}{n!}|v|^{n+2}|H_n(\lambda)|. \tag{11.5}$$

If the minimal absolute value of the zeros of

$$\sum_{\nu=0}^{n} \frac{H_\nu(\lambda)}{\nu!} z^\nu = 0 \tag{11.6}$$

is denoted by $A_n(\lambda)$, then the maximal one of

$$\sum_{\nu=0}^{n} \frac{H_\nu(\lambda)}{\nu!} v^\nu z^{n-\nu}$$

is $\dfrac{|v|}{A_n(\lambda)}$. Hence, if we choose $v = A_n(\lambda)$, we obtained a $(z_1^*, \ldots, z_n^*)$-system with

$$\max_{i=1,\,\ldots,\,n} |z_j^*| = 1$$

and

$$\max_{\nu=3,\,4,\,\ldots,\,(n+2)} |z_1^{*\nu} + \cdots + z_n^{*\nu}| = \frac{2}{n!}|H_n(\lambda)|A_n(\lambda)^{n+2}. \tag{11.7}$$

An asymptotical determination of $A_n(\lambda)$ (or even a good upper estimation of it) and a suitable choice of λ would probably result a better lower bound for A_1 in the Theorem.

(Received 29 January 1955)

О НЕКОТОРЫХ НОВЫХ ТЕОРЕМАХ ТЕОРИИ ДИОФАНТОВЫХ ПРИБЛИЖЕНИЙ

Вера Т. Шош и П. Туран (Будапешт)

(Резюме)

Второй из авторов настоящей работы во своей недавно вышедшей книге дал целый ряд применений диофантовых неравенств I, II, III, относящихся к различным ветвьям анализа и аналитической теории чисел. Улучшение этих неравенств важно и с точки зрения применений. В настоящей работе авторы значительно улучшают теорему II. В одном из самых важных для применений случаев это улучшение состоит в следующем: Пусть $\max_{j=1,\ldots,n} |z_j| = 1$, и пусть A означает наименьшую числовую постоянную,

для которой

$$\max_{m+1 \leqq \nu \leqq m+n} |s_\nu| \equiv \max_{m+1 \leqq \nu \leqq m+n} |z_1^\nu + \cdots + z_n^\nu| \geqq \left(\frac{n}{A(m+n)}\right)^n$$

(ν принимает целые значения), при всех целых неотрицательных m и целых n. Тогда

$$1{,}321 < A < 2e^{1+\frac{4}{e}}.$$

В связи с улучшением оценки снизу возникает вопрос о всех системах $(z_1, \ldots, z_n)$, для которых

a) $s_2 = s_3 = \cdots = s_n = 0$,

b) $s_3 = s_4 = \cdots = s_{n+1} = 0$.

Легко доказать, что с точностью до растяжения и вращения единственная система $(z_1, \ldots, z_n)$, удовлетворяющая условию a) состоит из корней уравнения $\sum_{\nu=0}^{n} \frac{z^{n-\nu}}{\nu!} = 0$. Несколько труднее доказать, что — опять с точностью до растяжения и вращения — все системы $(z_1, \ldots, z_n)$, удовлетворяющие условию b) состоит из корней $n+1$ уравнений

$$\sum_{\nu=0}^{n} \frac{H_\nu(\lambda)}{\nu!} z^{n-\nu} = 0,$$

где $H_\nu(y)$ есть многочлен Эрмита степени ν, а λ любой корень уравнения $H_{n+1}(y) = 0$.

ON THE INSTABILITY OF SYSTEMS OF DIFFERENTIAL EQUATIONS

By

P. TURÁN (Budapest), member of the Academy

1. We consider systems of differential equations in the normalform

$$\frac{dX_\nu(t)}{dt} = \sum_{j=1}^{n} a_{\nu j} X_j(t) + w_\nu(X_1, \ldots, X_n, t) \qquad (\nu = 1, 2, \ldots, n) \tag{1.1}$$

where the $a_{\nu j}$'s are complex numbers and

$$w_\nu(0, \ldots, 0, t) \equiv 0 \qquad (\nu = 1, \ldots, n). \tag{1.2}$$

This means that

$$X_\nu(t) \equiv 0 \qquad (\nu = 1, \ldots, n) \tag{1.3}$$

is a solution of (1.1). The stability theory of the system (1.1) asks what can be said on a solution for $t > 0$, when the $|X_\nu(0)|$ numbers are small. Among the many questions of this theory we consider here the question of the complete instability. For the system (1.1) the solution (1.3) is called by PERRON [1] completely instable, when there is a $c > 0$ such that for each solution of (1.1) with

$$0 < \sum_{\nu=1}^{n} |X_\nu(0)|^2 < c$$

either the solution ceases to exist for a $t_0 > 0$ or for a $t_1 > 0$

$$\sum_{\nu=1}^{n} |X_\nu(t_1)|^2 \geqq c$$

holds. He proved that if for some a and A the functions $w_\nu(y_1, \ldots, y_n, t)$ are continuous for

$$t \geqq 0, \quad |y_1| \leqq a, \ldots, |y_n| \leqq a \tag{1.4}$$

with complex variables y_ν and satisfy here the inequality

$$\frac{\sum_{\nu=1}^{n} |w_\nu(y_1, \ldots, y_n, t)|}{\sum_{\nu=1}^{n} |y_\nu|} \leqq A, \tag{1.5}$$

[1] O. PERRON, Über Stabilität und asymptotisches Verhalten der Integrale von Differentialgleichungssysteme, *Math. Zeitschr.*, **29** (1929), pp. 129—160.

further for an arbitrary small $\varepsilon > 0$ two positive numbers $\delta(\varepsilon)$ and $T_0(\varepsilon)$ can be found so that for

$$t \geqq T_0(\varepsilon), \quad |y_1| \leqq \delta(\varepsilon), \ldots, |y_n| \leqq \delta(\varepsilon) \tag{1.6}$$

the inequality

$$\frac{\sum_{\nu=1}^{n} |w_\nu(y_1, \ldots, y_n, t)|}{\sum_{\nu=1}^{n} |y_n|} \leqq \varepsilon \tag{1.7}$$

holds and finally the zeros λ_j of the equations

$$\Phi(\lambda) = \begin{vmatrix} a_{11} - \lambda & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} - \lambda \end{vmatrix} = 0 \tag{1.8}$$

are all in the right open half-plane, then complete instability occurs.

2. In what follows we shall give first another criterion for the complete instability which weakens (1. 7). We consider throughout this paper systems (1. 1) where

$$\min_{\nu=1,\ldots,n} \Re \lambda_\nu = L > 0 \tag{2.1}$$

and

$$\max_{1 \leqq \nu, j \leqq n} |a_{\nu j}| = C \tag{2.2}$$

are fixed. Let D be such that

$$D = \frac{n \log(4C) + \log(16n)}{2L}; \tag{2.3}$$

in this case we have evidently

$$e^{2LD} = 16n(4e)^n. \tag{2.4}$$

Then we shall prove the

THEOREM I. *If for some $a > 0$ the functions $w_\nu(y_1, \ldots, y_n, t)$ are in the domain*

$$t \geqq 0, \quad |y_1| \leqq a, \ldots, |y_n| \leqq a \tag{2.5}$$

continuous and here with above D

$$\frac{\sum_{k=1}^{n} |w_k(y_1, \ldots, y_n, t)|^2}{\sum_{k=1}^{n} |y_k|^2} \leqq \frac{1}{30nD^2} e^{-4nCD} \tag{2.6}$$

holds, then the solution $X_\nu(t) \equiv 0$ is completely instable.[2]

[2] Of course, it would be sufficient to require everything for $t > c$ only.

Surprisingly, on the right side of (2.6) we have a positive constant, dependent only upon easily calculable parameters of the system (1.1). PERRON's proof of his theorem was based on an explicit lower estimation of the norm of the solutions for *all* values t, valid, however, only in the special case

$$a_{rj} = 0 \qquad (j > r). \tag{2.7}$$

Our improvement was made possible by starting from an explicit lower estimation of the maximum of the norm in finite intervals. This inequality will be an easy deduction from Theorem VIII of my book.[3]

3. If we replace the domain (2.5) by the domain

$$t \geqq 0, \quad y_1, y_2, \ldots, y_n \text{ arbitrary} \tag{3.1}$$

and Δ is so large that

$$\Delta > 1 \tag{3.2}$$

and

$$e^{2L\Delta} \geqq 3n(2e(1+\Delta))^{2n}, \tag{3.3}$$

we assert the following

THEOREM II. *If the functions* $w_r(y_1, \ldots, y_n, t)$ *are in the domain* (3.1) *continuous and here*

$$\frac{\sum_{k=1}^{n} w_k(y_1, \ldots, y_n, t)|^2}{\sum_{k=1}^{n} |y_k|^2} \leqq \left\{ \frac{1}{4n(\Delta^2+\Delta)} \cdot \frac{e^{-nC(\Delta^2+\Delta)}}{(2e(\Delta+1))^n} \right\}^2 \tag{3.4}$$

holds, then each interval of the length $(\Delta^2+\Delta)$ *lying in* $t \geqq 0$ *contains a* t^* *with*

$$\sum_{k=1}^{n} |X_k(t^*)|^2 \geqq e^{2Lt^*\left(1-\frac{2}{\Delta+1}\right)} \sum_{k=1}^{n} |X_k(0)|^2. \tag{3.5}$$

This inequality is in a certain sense best-possible, even for a large subclass for the systems for which Theorem II refers. We consider namely systems

$$\frac{dY_r(t)}{dt} = \sum_{j=1}^{n} f_{rj}(t) Y_j(t) \qquad (r = 1, 2, \ldots, n) \tag{3.6}$$

where for $1 \leqq r, j \leqq n$ the functions $f_{rj}(t)$ are for $t \geqq 0$ continuous, for each $1 \leqq r, j \leqq n$

$$\lim_{t \to +\infty} f_{rj}(t) = a_{rj}$$

[3] P. TURÁN, *Eine neue Methode in der Analysis und deren Anwendungen* (Budapest, 1953), p. 42. This book contains already a weak result in this direction, rather only to indicate the possibility of application of the method also in this field. The results of the present paper are, however, much stronger and also a gap in the former proof is now eliminated.

exists and $L>0$. It is easy to see that condition (3.4) is in this case for $t>c_0=c_0(\Delta)$ fulfilled and also the continuity requirement (the w_ν's are now linear in all y_ν's). Hence Theorem II gives at once that each interval of length $(\Delta^2+\Delta)$ in $t\geqq 0$ contains a t^* with

$$\sum_{\nu=1}^{n}|X_\nu(t^*+c_0)|^2\geqq e^{2Lt^*\left(1-\frac{2}{\Delta+1}\right)}\sum_{\nu=1}^{n}|X_\nu(c_0)|^2. \tag{3.7}$$

But according to a part of the well-known theorem of POINCARÉ—PERRON, if

$$\Re\lambda_1=\min_j\Re\lambda_j=L>0,$$

then the system (3.6) has a $(Y_1,\ldots,Y_n)$-solution with

$$\overline{\lim_{t\to+\infty}}\frac{1}{t}\log\left(|Y_1|+\cdots+|Y_n|\right)=L. \tag{3.8}$$

Since

$$\sum_{\nu=1}^{n}|Y_\nu(t)|^2\leqq\left(\sum_{\nu=1}^{n}|Y_\nu(t)|\right)^2\leqq n\sum_{\nu=1}^{n}|Y_\nu(t)|^2,$$

(3.8) implies

$$\overline{\lim_{t\to+\infty}}\frac{1}{t}\log\left(\sum_{\nu=1}^{n}|Y_\nu(t)|^2\right)=2L. \tag{3.9}$$

Writing (3.7) in the form

$$\sum_{\nu=1}^{n}|X_\nu(\tau^*)|^2\geqq e^{2L(\tau^*-c_0)\left(1-\frac{2}{\Delta+1}\right)}\sum_{\nu=1}^{n}|X_\nu(c_0)|^2,$$

we see at once from (3.9) that the inequality

$$\sum_{\nu=1}^{n}|X_\nu(\tau^*)|^2>e^{(2+\eta)L\tau^*\left(1-\frac{2}{\Delta+1}\right)}\sum_{\nu=1}^{n}|X_\nu(0)|^2,$$

instead of (3.5), with a universal but arbitrarily small $\eta>0$ would be certainly false.

The Theorems I, II deal exclusively with that case, when all the zeros of the characteristic equation lie in the right half-plane. It would be of interest to obtain similar results also for the more general case, when only *some* zeros are in the right half-plane, when in PERRON's terminology conditional stability occurs. In such investigations instead of Theorem VIII the Theorem IX of my book can play a role.

The fact that in Theorem I instability could have been refined to an explicit inequality seems to me of some importance in the applications, e. g. it can mean a new approach to astronomical age-estimations.

The Theorems I and II can be extended also to systems of difference equations. We shall not go here into details of this. I shall return elsewhere to an algebraic application.

4. Next we turn to the proof of the theorems. As previously told they are largely based upon the following theorem of my book. If

$$\Re \alpha_\nu \geqq 0 \qquad (\nu = 1, 2, \ldots, n) \tag{4.1}$$

and β_ν are arbitrary complex numbers, then for arbitrary positive numbers a and d the estimation

$$\max_{a \leqq t \leqq a+d} \left| \sum_{j=1}^{n} \beta_j e^{\alpha_j t} \right| \geqq \left(\frac{d}{2e(a+d)} \right)^n \left| \sum_{j=1}^{n} \beta_j \right| \tag{4.2}$$

holds. Since the proofs of Theorem I and II run parallel we perform them parallel in order to avoid unnecessary repetitions.

Let $b \geqq 0$ be fixed and beside the system (1.1) we consider the system

$$\frac{dx_\nu(t)}{dt} = \sum_{j=1}^{n} a_{\nu j} x_j(t) \qquad (\nu = 1, \ldots, n) \tag{4.3}$$

with the initial conditions

$$x_\nu(b) = X_\nu(b) \qquad (\nu = 1, \ldots, n). \tag{4.4}$$

Let us first suppose that the zeros of $\Phi(\lambda) = 0$ in (1.8) are all different. Then the required solution of (4.3)—(4.4) is with suitable $c_{\nu j}$'s

$$x_\nu(t) = \sum_{j=1}^{n} c_{\nu j} e^{\lambda_j t} \qquad (\nu = 1, \ldots, n).$$

We put

$$\psi_\nu(t) = x_\nu(t) e^{-Lt} = \sum_{j=1}^{n} c_{\nu j} e^{(\lambda_j - L)t} \qquad (\nu = 1, \ldots, n). \tag{4.5}$$

We apply (4.1)—(4.2) to $\psi_\nu(t+b)$, first (for Theorem I) with

$$a = D, \quad d = D \tag{4.6}$$

and then (for Theorem II) with

$$a = \Delta^2, \quad d = \Delta. \tag{4.7}$$

Then we get in the first case for $\nu = 1, 2, \ldots, n$

$$\max_{b+D \leqq t \leqq b+2D} |x_\nu(t)| e^{-Lt} = \max_{D \leqq t \leqq 2D} |\psi_\nu(t+b)| \geqq \frac{1}{(4e)^n} |\psi_\nu(b)| = \frac{|x_\nu(b)|}{(4e)^n} e^{-Lb},$$

i. e.

$$\max_{b+D \leqq t \leqq b+2D} |x_\nu(t)| \geqq \frac{e^{LD}}{(4e)^n} |x_\nu(b)|. \tag{4.8}$$

In the second case we get

$$\max_{b+\Delta^2 \leqq t \leqq b+\Delta^2+\Delta} |x_\nu(t)| e^{-Lt} = \max_{\Delta^2 \leqq t \leqq \Delta^2+\Delta} |\psi_\nu(t+b)| \geqq$$

$$\geqq \left(\frac{\Delta}{2e(\Delta^2+\Delta)} \right)^n |\psi_\nu(b)| = \left(\frac{1}{2e(\Delta+1)} \right)^n |x_\nu(b)| e^{-Lb},$$

i. e.

$$\max_{b+\Delta^2\leqq t\leqq b+\Delta^2+\Delta}|x_\nu(t)|\geqq\frac{e^{L\Delta^2}}{(2e(\Delta+1))^n}|x_\nu(b)| \qquad (\nu=1,\dots,n). \tag{4.9}$$

Since in (4.8) and (4.9) nothing depends upon the configuration of the λ_j's and a polynomial can be approximated, uniformly in all finite intervals, by polynomials with exlusively simple zeros, a trivial limiting process results the validity of (4.8) and (4.9) without any simplicity requirement on the λ_j's. If further j_0 is an index with

$$|x_{j_0}(b)|^2\geqq\frac{1}{n}\sum_{j=1}^{n}|x_j(b)|^2,$$

we obtain from (4.8) and (4.9)

$$\begin{aligned}\max_{b+D\leqq t\leqq b+2D}\sum_{j=1}^{n}|x_j(t)|^2&\geqq\max_{b+D\leqq t\leqq b+2D}|x_{j_0}(t)|^2\geqq\frac{e^{2LD}}{(4e)^{2n}}|x_{j_0}(b)|^2\geqq\\ &\geqq\frac{e^{2LD}}{(4e)^{2n}n}\sum_{j=1}^{n}|x_j(b)|^2=\frac{e^{2LD}}{n(4e)^{2n}}\sum_{j=1}^{n}|X_j(b)|^2,\end{aligned} \tag{4.10}$$

$$\max_{b+\Delta^2\leqq t\leqq b+\Delta^2+\Delta}\sum_{j=1}^{n}|x_j(t)|^2\geqq\frac{e^{2L\Delta^2}}{n\{2e(\Delta+1)\}^{2n}}\sum_{j=1}^{n}|X_i(b)|^2, \tag{4.11}$$

respectively.

5. Further we shall need some lemmas.

LEMMA I. *For the function*

$$g(t)=\sum_{l=1}^{n}\gamma_l e^{\delta_l t} \tag{5.1}$$

with

$$\Re\delta_l\geqq 0 \qquad (l=1,2,\dots,n) \tag{5.2}$$

and

$$0<A_1<B_1<C_1<D_1 \tag{5.3}$$

we have[4]

$$\max_{A_1\leqq t\leqq D_1}|g(t)|\leqq\left(2e\frac{D_1-A_1}{C_1-B_1}\right)^n\max_{B_1\leqq t\leqq C_1}|g(t)|. \tag{5.4}$$

The proof follows easily from (4.1)—(4.2). Putting

$$\sum_{j=1}^{n}\beta_j e^{\alpha_j t}=h(t)$$

(4.2) means

$$\max_{a\leqq t\leqq a+d}|h(t)|\geqq\left(\frac{d}{2e(a+d)}\right)^n|h(0)|. \tag{5.5}$$

[4] This Lemma occurs in my book, p. 70 for the case $\Re\delta_j=0$ $(j=1,2,\dots,n)$ since I used it only for this case.

If $|g(t)|$ assumes its absolute maximum with respect to $[A_1, D_1]$ at $t = t_2$ and $B_1 \leqq t_2 \leqq C_1$, then (5. 4) is trivial. Consequently, we may assume without loss of generality

$$A_1 \leqq t_2 < B_1.$$

We have simply apply (5. 5) with

$$h(t) = g(t+t_2), \quad a = B_1 - t_2, \quad d = C_1 - B_1.$$

This gives

$$\max_{B_1 \leqq t \leqq C_1} |g(t)| = \max_{B_1 - t_2 \leqq t \leqq C_1 - t_2} |g(t+t_2)| \geqq \left(\frac{C_1 - B_1}{2e(C_1 - t_2)} \right)^n |g(t_2)| \geqq$$

$$\geqq \left(\frac{C_1 - B_2}{2e(D_1 - A_1)} \right)^n \max_{A_1 \leqq t \leqq D_1} |g(t)|. \qquad \text{Q. e. d.}$$

6. Next we reproduce for the sake of completeness the (short) proofs of two well-known lemmas; the first goes back to LIOUVILLE, the second one to LIAPUNOFF and PERRON. We consider beside the systems (1. 1) and (4. 3)—(4. 4) for $k = 1, \ldots, n$ the functions

$$x_{k1}(t), \ldots, x_{kn}(t),$$

satisfying again the system (4. 3) but with the initial-conditions

$$x_{kj}(b) = \begin{cases} 1 & \text{for } j = k, \\ 0 & \text{for } j \neq k. \end{cases} \tag{6. 1}$$

Then we have the

LEMMA II. *We have for* $\nu = 1, 2, \ldots, n$ *and* $t \geqq b$

$$X_\nu(t) = x_\nu(t) + \int_b^t \left\{ \sum_{k=1}^n x_{k\nu}(t+b-\tau) w_k(X_1(\tau), \ldots, X_n(\tau), \tau) \right\} d\tau.$$

For the proof let

$$x_\nu(t) + \int_b^t \left\{ \sum_{k=1}^n x_{k\nu}(t+b-\tau) w_k(X_1(\tau), \ldots, X_n(\tau), \tau) \right\} d\tau = Z_\nu(t) \tag{6. 2}$$

$$(\nu = 1, \ldots, n)$$

where the functions $x_\nu(t)$ and $X_\nu(t)$ are considered as given. We have obviously

$$Z_\nu(b) = x_\nu(b) = X_\nu(b) \qquad (\nu = 1, \ldots, n). \tag{6. 3}$$

Further

$$\frac{dZ_\nu(t)}{dt}=\frac{dx_\nu(t)}{dt}+w_\nu(X_1(t),\dots,X_n(t),t)+$$

$$+\int_b^t\left\{\sum_{k=1}^n\frac{dx_{k\nu}(t+b-\tau)}{dt}\cdot w_k(X_1(\tau),\dots,X_n(\tau),\tau)\right\}d\tau=w_\nu(X_1(t),\dots,X_n(t),t)+$$

$$+\sum_{j=1}^n a_{\nu j}\left(x_j(t)+\int_b^t\left\{\sum_{k=1}^n x_{kj}(t+b-\tau)w_k(X_1(\tau),\dots,X_n(\tau),\tau)\right\}d\tau\right)=$$

$$=\sum_{j=1}^n a_{\nu j}Z_j(t)+w_\nu(X_1,\dots,X_n,t),$$

i. e. $(Z_1,\dots,Z_n)$ constitutes a solution of the system

$$\frac{dZ_\nu(t)}{dt}=\sum_{j=1}^n a_{\nu j}Z_j(t)+w_\nu(X_1(t),\dots,X_n(t),t)\qquad(\nu=1,2,\dots,n).\tag{6.4}$$

But owing to (1.1) the same holds for $(X_1,\dots,X_n)$; owing to (6.3) and the uniqueness of the system (6.4) we have

$$Z_\nu(t)\equiv X_\nu(t)\qquad(\nu=1,\dots,n),$$

indeed.

Finally we need the

LEMMA III. *When for the coefficient functions $b_{\nu j}(t)$ of the system*

$$\frac{dz_\nu(t)}{dt}=\sum_{j=1}^n G_{\nu j}(t)z_j(t)\qquad(\nu=1,\dots,n)\tag{6.5}$$

for $b\leqq t\leqq T$ the estimation

$$|G_{\nu j}(t)|\leqq G\qquad(1\leqq\nu,j\leqq n)\tag{6.6}$$

holds, then we have for $b\leqq t\leqq T$

$$\sum_{j=1}^n|z_j(t)|^2\leqq e^{2nG(t-b)}\sum_{j=1}^n|z_j(b)|^2.\tag{6.7}$$

and

$$\sum_{j=1}^n|z_j(t)|^2\geqq e^{-2nG(t-b)}\sum_{j=1}^n|z_j(b)|^2\tag{6.8}$$

PROOF. We have for $\nu=1,2,\dots,n$

$$\frac{1}{2}\frac{d}{dt}|z_\nu|^2=\Re\left(\bar z_\nu\frac{dz_\nu}{dt}\right)\begin{cases}\leqq G|z_\nu|\sum_{j=1}^n|z_j|\leqq G\sum_{j=1}^n\frac{|z_\nu|^2+|z_j|^2}{2},\\ \geqq -G|z_\nu|\sum_{j=1}^n|z_j|\geqq -G\sum_{j=1}^n\frac{|z_\nu|^2+|z_j|}{2}.\end{cases}$$

Summation with respect to ν gives with $\sum_{\nu=1}^{n} |z_\nu|^2 \equiv \varphi(t)$

$$\frac{1}{2}\varphi'(t) \leqq Gn\varphi(t), \qquad \frac{1}{2}\varphi'(t) \geqq -Gn\varphi(t),$$

i. e.

$$\frac{d}{dt}(e^{-2nGt}\varphi(t)) \leqq 0, \qquad \frac{d}{dt}(e^{2nGt}\varphi(t)) \geqq 0.$$

Integration from b to t gives already the required inequalities.

Applying Lemma III to the functions

$$x_{k1}(t), \ldots, x_{kn}(t)$$

we obtain for $k = 1, \ldots, n$ and $t \geqq b$

$$\sum_{j=1}^{n} |x_{kj}(t)|^2 \leqq e^{2nC(t-b)} \tag{6.9}$$

and

$$\sum_{j=1}^{n} |x_{kj}(t)|^2 \geqq e^{-2nC(t-b)}. \tag{6.10}$$

7. Next we turn to the proof of Theorem I. We suppose the theorem to be false, i. e. there is a solution $(X_1, \ldots, X_n)$ not identically vanishing for which

$$\sum_{\nu=1}^{n} |X_\nu(t)|^2 < a \tag{7.1}$$

for $t \geqq b_0$; we may suppose without loss of generality that

$$\sum_{\nu=1}^{n} |X_\nu(b_0)|^2 > 0. \tag{7.2}$$

We may apply Lemma II with $b = b_0$ for $\nu = 1, \ldots, n$, if $b_0 \leqq t \leqq b_0 + 2D$. This gives

$$|X_\nu(t) - x_\nu(t)| \leqq \int_{b_0}^{t} \left\{ \sum_{k=1}^{n} |x_{k\nu}(t + b_0 - \tau)| \, |w_k(X_1(\tau), \ldots, X_n(\tau), \tau)| \right\} d\tau$$

or

$$|X_\nu(t) - x_\nu(t)|^2 \leqq 2D \int_{b_0}^{t} \left\{ \sum_{k=1}^{n} |x_{k\nu}(t + b_0 - \tau)|^2 \right\} \left\{ \sum_{k=1}^{n} w_k(X_1, \ldots, X_n, \tau)|^2 \right\} d\tau,$$

i. e.

$$\sum_{\nu=1}^{n} |X_\nu(t) - x_\nu(t)|^2 \leqq$$

$$\leqq 2D \int_{b_0}^{t} \left\{ \sum_{k=1}^{n} \sum_{\nu=1}^{n} |x_{k\nu}(t + b_0 - \tau)|^2 \right\} \left\{ \sum_{k=1}^{n} w_k(X_1(\tau), \ldots, X_n(\tau), \tau)|^2 \right\} d\tau. \tag{7.3}$$

Taking in account (6. 9) we have in the range of integration

$$\sum_{k=1}^{n}\sum_{r=1}^{n}|x_{kr}(t+b_0-\tau)|^2 \leqq ne^{2nC(t-\tau)} \leqq ne^{2nC(t-b_0)} \leqq ne^{4nCD},$$

i. e. from (7. 3)

$$(7.4)\quad \sum_{r=1}^{n}|X_r(t)-x_r(t)|^2 \leqq 2nDe^{4nCD}\int_{b_0}^{b_0+2D}\left\{\sum_{k=1}^{n}|w_k(X_1(\tau),\ldots,X_n(\tau),\tau)|^2\right\}d\tau.$$

Making use of the condition (2. 6) (which is justified by (7. 1)), (7. 4) gives

$$(7.5)\quad \begin{aligned}\sum_{r=1}^{n}|X_r(t)-x_r(t)|^2 &\leqq \frac{1}{15D}\int_{b_0}^{b_0+2D}\left\{\sum_{k=1}^{n}|X_k(\tau)|^2\right\}d\tau \leqq \\ &\leqq \frac{2}{15D}\left(\int_{b_0}^{b_0+2D}\left\{\sum_{k=1}^{n}|x_k(\tau)|^2\right\}d\tau+\int_{b_0}^{b_0+2D}\left\{\sum_{k=1}^{n}|X_k(\tau)-x_k(\tau)|^2\right\}d\tau\right).\end{aligned}$$

If

$$U(t)=\sum_{k=1}^{n}|x_k(t)|^2$$

and

$$V(t)=\sum_{k=1}^{n}|X_k(t)-x_k(t)|^2$$

assume for $t=\tau^*$ and $t=\tau'$, respectively, their absolut maximum referring to $[b_0, b_0+2D]$, then the mean-value theorem gives from (7. 5), when t on the left is replaced by τ',

$$V(\tau') \leqq \frac{4}{15}U(\tau^*)+\frac{4}{15}V(\tau')$$

or

$$(7.6)\quad V(\tau') \leqq \frac{4}{11}U(\tau^*).$$

Hence, if

$$W(t)=\sum_{k=1}^{n}|X_k(t)|^2$$

assumes its absolute maximum within (b_0, b_0+2D) at $t=\tau''$, then, taking in account that from

$$|u_1|^2=|v_1+(u_1-v_1)|^2 \leqq 2(|v_1|^2+|u_1-v_1|^2)$$

we have for arbitrary u_1 and v_1

$$(7.7)\quad |v_1|^2 \geqq \frac{1}{2}|u_1|^2-|u_1-v_1|^2,$$

we get by (7. 6)

$$\sum_{k=1}^{n}|X_k(\tau'')|^2 \geqq \sum_{k=1}^{n}|X_k(\tau^*)|^2 \geqq \frac{1}{2}\sum_{k=1}^{n}|x_k(\tau^*)|^2 - \sum_{k=1}^{n}|X_k(\tau^*)-x_k(\tau^*)|^2 =$$

$$= \frac{1}{2}U(\tau^*) - V(\tau^*) \geqq \frac{1}{2}U(\tau^*) - V(\tau') \geqq \frac{1}{2}U(\tau^*) - \frac{4}{11}U(\tau^*) > \frac{1}{8}U(\tau^*).$$

Hence

$$\max_{b_0 \leqq t \leqq b_0+2D}\sum_{k=1}^{n}|X_k(t)|^2 \geqq \frac{1}{8}\max_{b_0 \leqq t \leqq b_0+2D}\sum_{k=1}^{n}|x_k(t)|^2 \geqq \frac{1}{8}\max_{b_0+D \leqq t \leqq b_0+2D}\sum_{k=1}^{n}|x_k(t)|^2$$

and by (4. 10)

$$\sum_{k=1}^{n}|X_k(\tau'')|^2 \geqq \frac{e^{2LD}}{8n(4e)^{2n}}\sum_{k=1}^{n}|X_k(b_0)|^2 \geqq 2\sum_{k=1}^{n}|X_k(b_0)|^2,$$

taking in account (2. 4). Replacing b_0 by this τ'' we obtain that for a suitable τ''' with

$$(b_0 \leqq)\tau'' \leqq \tau''' \leqq \tau'' + 2D$$

we have

$$\sum_{k=1}^{n}|X_k(\tau''')|^2 \geqq 2\sum_{k=1}^{n}|X_k(\tau'')|^2 \geqq 2^2\sum_{k=1}^{n}|X_k(b_0)|^2.$$

Continuing this process we get the existence of a $\zeta > b_0$ with

$$\sum_{k=1}^{n}|X_k(\zeta)|^2 > a$$

which contradicts to (7. 1).

8. Next we turn to the proof of Theorem II. Let $b_1 \geqq 0$. We apply Lemma II with $b = b_1$ for $\nu = 1, \ldots, n$ and $b_1 \leqq t \leqq b_1 + \Delta^2 + \Delta$. Proceeding exactly as at the proof of Theorem I we obtain

$$\sum_{\nu=1}^{n}|X_\nu(t)-x_\nu(t)|^2 \leqq (\Delta^2+\Delta)ne^{2nC(\Delta^2+\Delta)}\int_{b_1}^{b_1+\Delta^2+\Delta}\left\{\sum_{k=1}^{n}|w_k(X_1, \ldots, X_n, \tau)|^2\right\}d\tau.$$

Applying now (3. 4) it follows

$$\sum_{\nu=1}^{n}|X_\nu(t)-x_\nu(t)|^2 \leqq \frac{\{2e(\Delta+1)\}^{-2n}}{16n(\Delta^2+\Delta)}\int_{b_1}^{b_1+\Delta^2+\Delta}\left\{\sum_{k=1}^{n}|X_k(\tau)|^2\right\}d\tau. \tag{8. 1}$$

When τ^* and τ' stand now for the place of the absolute maximum of $U(t)$ and $V(t)$, respectively, with respect to $(b_1, b_1+\Delta^2+\Delta)$, we obtain from (8. 1), analogously as in (7. 6),

$$V(\tau') \leqq \frac{\{2e(\Delta+1)\}^{-2n}}{8n}U(\tau^*) + \frac{1}{8}V(\tau'),$$

$$V(\tau') \leqq \frac{\{2e(\Delta+1)\}^{-2n}}{7n}U(\tau^*). \tag{8. 2}$$

$U(\tau^*)$ is the maximum with respect to $(b_1, b_1+\Delta^2+\Delta)$; Lemma I gives possibility[5] to estimate $U(\tau^*)$ from above by

$$\max_{b_1+\Delta^2\leqq t\leqq b_1+\Delta^2+\Delta} U(t) = U(t_3)$$

with

$$b_1+\Delta^2 \leqq t_3 \leqq b_1+\Delta^2+\Delta. \tag{8.3}$$

Let j_1 be such an index for which

$$|x_{j_1}(\tau^*)|^2 \geqq \frac{1}{n}\sum_{j=1}^{n}|x_j(\tau^*)|^2 = \frac{1}{n}U(\tau^*);$$

then

$$U(\tau^*) \leqq n|x_{j_1}(\tau^*)|^2 \leqq n \max_{b_1\leqq t\leqq b_1+\Delta^2+\Delta}|x_{j_1}(t)|^2. \tag{8.4}$$

Applying Lemma I to $|x_{j_1}(t)|^2$ we obtain, taking in account the remarks after (4.9),

$$\max_{b_1\leqq t\leqq b_1+\Delta^2+\Delta}|x_{j_1}(t)|^2 \leqq \left(2e\frac{\Delta^2+\Delta}{\Delta}\right)^{2n}\max_{b_1+\Delta^2\leqq t\leqq b_1+\Delta^2+\Delta}|x_{j_1}(t)|^2 \leqq$$

$$\leqq (2e(\Delta+1))^{2n}\max_{b_1+\Delta^2\leqq t\leqq b_1+\Delta^2+\Delta}\sum_{j=1}^{n}|x_j(t)|^2 = \{2e(\Delta+1)\}^{2n}U(t_3).$$

Thus (8.4) gives

$$U(\tau^*) \leqq n\{2e(\Delta+1)\}^{2n}U(t_3).$$

This, together with (8.2), gives

$$V(\tau') \leqq \frac{1}{7}U(t_3). \tag{8.5}$$

Then

$$\max_{b_1+\Delta^2\leqq t\leqq b_1+\Delta^2+\Delta}\sum_{r=1}^{n}|X_r(t)|^2 \geqq \sum_{r=1}^{n}|X_r(t_3)|^2 \geqq \frac{1}{2}U(t_3) - V(t_3) \geqq$$

$$\geqq \frac{1}{2}U(t_3) - V(\tau') \geqq \frac{1}{3}U(t_3),$$

using (7.7) and (8.5). The inequality (4.11) gives then

$$\max_{b_1+\Delta^2\leqq t\leqq b_1+\Delta^2+\Delta}\sum_{k=1}^{n}|X_k(t)|^2 \geqq \frac{e^{2L\Delta^2}}{3n\{2e(\Delta+1)\}^{2n}}\sum_{k=1}^{n}|X_k(b_1)|^2 \geqq$$

$$\geqq e^{2L(\Delta^2-\Delta)}\sum_{k=1}^{n}|X_k(b_1)|^2, \tag{8.6}$$

using (3.3).

First we apply (8.6) with $b_1 = 0$. This gives the existence of a τ_1 with

$$\Delta^2 \leqq \tau_1 \leqq \Delta^2+\Delta$$

[5] As VERA T. SÓS remarked, a similar but weaker inequality can be deduced without Lemma I, using (6.8).

and

$$\sum_{k=1}^{n} |X_k(\tau_1)|^2 \geqq e^{2L(\Delta^2-\Delta)} \sum_{k=1}^{n} |X_k(0)|^2.$$

Next we use (8.6) with $b_1 = \tau_1$. This gives the existence of a τ_2 with

$$\tau_1 + \Delta^2 \leqq \tau_2 \leqq \tau_1 + \Delta^2 + \Delta$$

and such that

$$\sum_{k=1}^{n} |X_k(\tau_2)|^2 \geqq e^{2L(\Delta^2-\Delta)} \sum_{k=1}^{n} |X_k(\tau_1)|^2 \geqq e^{4L(\Delta^2-\Delta)} \sum_{k=1}^{n} |X_k(0)|^2.$$

Continuing this process we obtain the existence of a sequence

$$(0 = \tau_0 <) \tau_1 < \tau_2 < \cdots$$

such that for $m = 1, 2, \ldots$

(8.7) $$\tau_{m-1} + \Delta^2 \leqq \tau_m \leqq \tau_{m-1} + \Delta^2 + \Delta$$

and

(8.8) $$\sum_{k=1}^{n} |X_k(\tau_m)|^2 \geqq e^{2mL(\Delta^2-\Delta)} \sum_{k=1}^{n} |X_k(0)|^2.$$

Since from (8.7)

$$\tau_m \leqq m(\Delta^2 + \Delta),$$

we have

$$2L(\Delta^2-\Delta)m \geqq 2L\tau_m\left(1 - \frac{2}{\Delta+1}\right);$$

taking in account that owing to (8.7) each interval in $t \geqq 0$ of length $(\Delta^2+\Delta)$ contains at least one τ_j, Theorem II is proved.

(Received 8 August 1955)

О НЕУСТОЙЧИВОСТИ СИСТЕМ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

П. Туран (Будапешт)

(Резюме)

Как известно,[1] О. Перрон в случае системы дифференциальных уравнений в нормальной форме

(1) $$\frac{dX_\nu}{dt} = \sum_{j=1}^{n} a_{\nu j} X_j + w_\nu(X_1, \ldots, X_n, t),$$

$$w_\nu(0, \ldots, 0, t) \equiv 0 \qquad (\nu = 1, \ldots, n)$$

называет решение

(2) $$X_\nu(t) \equiv 0 \qquad (\nu = 1, \ldots, n)$$

„вполне неустойчивым“, если существует $c > 0$ такое, что для всякого решения (1)

выполняющего условие

$$0 < \sum_{\nu=1}^{n} |X_\nu(0)|^2 < c$$

можно найти либо такое $t_0 > 0$, что для $t = t_0$ решение не существует, либо такое $t_1 > 0$, что

$$\sum_{\nu=1}^{n} |X_\nu(t_1)|^2 \geqq c.$$

Улучшая в некотором смысле достаточное условие Перрона, автор доказывает две теоремы, опираясь на теорему VIII своей книги цитированной на стр. 259.

Пусть для корней $\lambda_1, \ldots, \lambda_n$ уравнения

$$\begin{vmatrix} a_{11}-\lambda & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0$$

имеет место

$$\min_{\nu=1,\ldots,n} \Re\lambda_\nu = L > 0,$$

и пусть

$$D = 1 + \frac{3n + \log(16n)}{2L}.$$

Пусть далее

$$\max_{\nu, j} |a_{\nu j}| = C.$$

Тогда имеет место следующая

Теорема I. Если для некоторого $a > 0$ через всякую точку области

$$t \geqq 0, \quad |y_1| \leqq a, \ldots, |y_n| \leqq a$$

проходит единственное решение (1) и

$$\frac{\sum_{k=1}^{n} |w_k(y_1, \ldots, y_n, t)|^2}{\sum_{k=1}^{n} |y_k|^2} \leqq \frac{1}{30nD^2} e^{-4nCD},$$

то (2) является вполне неустойчивым решением (1).

Пусть теперь число Δ столь велико, что справедливо неравенство

$$e^{2L\Delta} \geqq 3n(2e(\Delta+1))^{2n}. \tag{3}$$

Тогда имеет место

Теорема II. Если через всякую точку $n+1$-мерной области, определяемой $t \geqq 0$ и произвольными комплексными $y_1, \ldots, y_n$, проходит единственное решение (1) и в этой области

$$\frac{\sum_{k=1}^{n} |w_k(y_1, \ldots, y_n, t)|^2}{\sum_{k=1}^{n} |y_k|^2} \leqq \left\{ \frac{1}{4n(\Delta^2+\Delta)} \cdot \frac{e^{-nC(\Delta^2+\Delta)}}{(2e(\Delta+1))^n} \right\}^2,$$

то во всяком интервале длины $\Delta^2 + \Delta$, лежащем в $t \geqq 0$, существует точка t^* такая, что

$$\sum_{k=1}^{n} |X_k(t^*)|^2 \geqq e^{2Lt^*\left(1-\frac{2}{\Delta+1}\right)} \sum_{k=1}^{n} |X_k(0)|^2.$$

Автор намерен указать одно алгебраическое следствие теоремы II в другой работе.

DIE GRUPPENTHEORETISCHEN ZETAFUNKTIONEN UND DER SATZ VON HAJÓS

Von

LADISLAUS RÉDEI (Szeged), Mitglied der Akademie

§ 1. Einleitung

Der Satz von HAJÓS über die endlichen Abelschen Gruppen ist einer der bewunderswertesten in der Mathematik. Die beiden Hauptcharakterzüge des Satzes sind seine Vielseitigkeit und die überraschenden Schwierigkeiten seines Beweises. Unter seiner Vielseitigkeit verstehe ich, daß der Satz eine ganze Reihe äquivalente Umformulierungen zuläßt, die teils sehr wichtig sind und in verschiedene Gebiete der Mathematik hineingreifen. Als solche Umformulierungen nenne ich die berühmte Minkowskische Vermutung über den „Grenzfall" des homogen linearen Ungleichungssystems und über die raumzerlegenden Würfelgitter (vgl. HAJÓS [1]).[1] Auf viele weitere, allerdings weniger wichtige Umformulierungen stößt man, wenn man versucht für den Satz einen neuen Beweis zu gewinnen. Auf solchem Wege bin ich neulich durch den Satz von HAJÓS zur Entdeckung meiner Zetafunktionen (vgl. RÉDEI [2]) geleitet worden, auch entstand dabei eine neue Äquivalente des Satzes, die der einzige Gegenstand der vorliegenden Arbeit sein wird. Es wird sich zeigen, daß der Satz von HAJÓS einer Poleigenschaft von gewissen solchen Zetafunktionen gleichkommt. Der Beweis wird durch den ebenfalls mit Hilfe meiner Zetafunktionen entstandenen tiefliegenden Trägheitssatz der endlichen Abelschen Gruppen (vgl. RÉDEI [2], Satz 9) ermöglicht. Mit Hilfe meiner Umformulierung des Satzes von HAJÓS einen neuen Beweis für diesen Satz zu finden, ist mir nicht gelungen. Zum Erfolg würde vielleicht verhelfen, wenn man die Theorie meiner Zetafunktionen weiter ausbaut. Ich bemerke noch, daß für den Satz von HAJÓS außer seinem Originalbeweis und seinen Varianten bisher nur noch ein zweiter, wesentlich verschiedener, von mir kürzlich veröffentlichter Beweis vorliegt. (S. HAJÓS [1], RÉDEI [3] und die Literaturangaben in der letzteren Arbeit.)

[1] Mit [ ] verweisen wir auf das Literaturverzeichnis am Ende unserer Arbeit.

§ 2. Die Umformulierung des Satzes von Hajós mittels Zetafunktionen

Durchgängig werden G eine endliche Abelsche Gruppe, n eine natürliche Zahl, $p_1, \ldots, p_n$ (nicht notwendig verschiedene) Primzahlen, $a_1, \ldots, a_n$ Elemente von G, ferner $o(a)$ $(a \in G)$ die Ordnung von a bezeichnen.

Der Satz von HAJÓS lautet: *Wenn die* $p_1 \ldots p_n$ *Produkte*

$$a_1^{x_1} \ldots a_n^{x_n} \qquad (x_i = 0, \ldots, p_i - 1;\ i = 1, \ldots, n) \tag{1}$$

eine Gruppe $p_1 \ldots p_n$*-ter Ordnung bilden, so gilt* $o(a_i) = p_i$ *für mindestens ein* i.

(Der Satz gilt sogar auch dann, wenn man an Stelle von $p_1, \ldots, p_n$ beliebige von 1 verschiedene natürliche Zahlen nimmt, jedoch ist diese Verallgemeinerung nach HAJÓS [1] unwesentlich. Für einen kurzen Beweis s. man RÉDEI [3], Hilfssatz 2.)

Wir führen noch die folgenden, ebenfalls *durchgängigen Bezeichnungen* ein:

$\mathfrak{N}$ die Menge der Zahlen $1, \ldots, n$.

$\mathbf{0}$ die leere Menge.

$(\ldots)$ die Anzahl der Elemente einer endlichen Menge, insbesondere die Ordnung einer endlichen Gruppe.

$\{\cdots\}$ die durch die eingeklammerten Elemente erzeugte Gruppe.

ε das Einselement von G.

$A_1, \ldots, A_n$ zyklische Untergruppen von G.

$B_1 = A_1^{p_1}, \ldots, B_n = A_n^{p_n}$ (d. h. B_i die Gruppe der p_i-ten Potenzen der Elemente von A_i; $i = 1, \ldots, n$).

$\mathfrak{M}$ eine Untermenge von $\mathfrak{N}$ (in Zeichen $\mathfrak{M} \subseteq \mathfrak{N}$).

$p_{\mathfrak{M}}$ das Produkt der p_i mit $i \in \mathfrak{M}$.[2]

$A_{\mathfrak{M}}$ das Produkt der A_i mit $i \in \mathfrak{M}$.

$B_{\mathfrak{M}}$ das Produkt der B_i mit $i \in \mathfrak{M}$.

Wir setzen (wie in der Arbeit RÉDEI [2] (1))

$$\varrho(z) = \varrho(z; A_1, \ldots, A_n) = \sum_{\mathfrak{M} \subseteq \mathfrak{N}} (-1)^{(\mathfrak{M})} (A_{\mathfrak{M}})^{-z} \tag{2}$$

mit einer komplexen Variablen z und nennen $\zeta(z) = \varrho(z)^{-1}$ die (zu den Untergruppen $A_1, \ldots, A_n$ gehörende) *Zetafunktion*.[3] Unter den zwei Funktionen $\varrho(z), \zeta(z)$ werden wir nur die erste verwenden, aber alle unseren auf $\varrho(z)$ bezüglichen Aussagen lassen sich auch auf $\zeta(z)$ beziehen. Insbesondere ist

[2] Im Fall $\mathfrak{M} = \mathbf{0}$ ist freilich $p_{\mathfrak{M}} = 1$ zu verstehen, wenn ferner $\mathfrak{M}$ aus dem einzigen Element i besteht, so ist $p_{\mathfrak{M}} = p_i$ zu verstehen. Diese Bemerkungen sind sinnmäßig auch auf andere ähnliche Fälle zu beziehen.

[3] In RÉDEI [2] wurde $\zeta(z)$ für viel allgemeinere Fälle definiert, hier wird aber nur der obige Fall und zwar nur für die (reellen) ganzen z in Betracht kommen. Man bemerke, daß in (2) die Reihenfolge der $A_1, \ldots, A_n$ gleichgültig ist.

für $z_0 \neq \infty$ „$\varrho(z_0) = 0$" gleichwertig mit „$z = z_0$ ist Polstelle der Zetafunktion $\zeta(z)$".[4]

In dieser Arbeit werden unter allen Funktionen $\varrho(z)$ nur die

$$\varrho_{\mathfrak{M}}(z) = \varrho(z; B_{i_1} A_{\mathfrak{M}}/A_{\mathfrak{M}}, \ldots, B_{i_m} A_{\mathfrak{M}}/A_{\mathfrak{M}}) \qquad (\mathfrak{M} \subset \mathfrak{N}) \tag{3}$$

zu Wort kommen, wobei $i_1, \ldots, i_m$ die sämtlichen verschiedenen Elemente von $\mathfrak{N} - \mathfrak{M}$ bezeichnen. Insbesondere ist also für $\mathfrak{M} = \mathbf{0}$

$$\varrho_0(z) = \varrho(z; B_1, \ldots, B_n). \tag{4}$$

Gleich bemerken wir, daß sich (3) wegen der für alle Untergruppen U, V von G gültigen Formel $(UV/U) = (U)^{-1}(UV)$ und wegen (2) als

$$\varrho_{\mathfrak{M}}(z) = (A_{\mathfrak{M}})^{-z} \sum_{\mathfrak{L} \subseteq \mathfrak{N} - \mathfrak{M}} (-1)^{(\mathfrak{L})} (A_{\mathfrak{M}} B_{\mathfrak{L}})^{-z} \qquad (\mathfrak{M} \subset \mathfrak{N}) \tag{5}$$

schreiben läßt.

Die gemeinte Äquivalente des Satzes von HAJÓS ist der folgende:

SATZ 1. *Wenn*

$$(A_{\mathfrak{N}}) = p_{\mathfrak{N}}, \quad \varrho_{\mathfrak{M}}(1) = 0 \qquad (\mathfrak{M} \subset \mathfrak{N}) \tag{6}$$

gelten, so gilt

$$\varrho_0(2) = \varrho_0(3) = \cdots = 0. \tag{7}$$

Hiernach ist der Satz von HAJÓS in der Tat äquivalent mit einer Poleigenschaft von Zetafunktionen. Satz 1 können wir in einer einfacheren Form aussprechen, die auch für den Beweis geeigneter wird. Für eine beliebige Funktion $\varrho(z)$ in (2) ist nämlich $\varrho(2) = \varrho(3) = \cdots = 0$ (vgl. RÉDEI [2], Korollare der Sätze 10, 2) gleichbedeutend damit, daß $\varrho(z)$ konstant, d. h. mindestens ein A_i gleich ε ist. Wegen (3) ist also (7) gleichbedeutend damit, daß mindestens ein B_i gleich ε ist. Indem wir auch noch (6) wenig umformen, können wir nunmehr Satz 1 so aussprechen:

SATZ 1′. *Wenn*

$$\frac{1}{p_{\mathfrak{N}}} = \frac{1}{(A_{\mathfrak{N}})}, \quad \varrho_{\mathfrak{M}}(1) = 0 \qquad (\mathfrak{M} \subset \mathfrak{N}) \tag{8}$$

ist, so ist mindestens ein (B_i) *gleich* 1 $(i = 1, \ldots, n)$.

Es genügt, wenn wir die Äquivalenz vom Hajósschen Satz und vom Satz 1′ ausweisen.

§ 3. Beweis der Äquivalenz vom Hajósschen Satz und vom Satz 1

Zu (5) ähnlich definieren wir die Funktionen

$$\Delta_{\mathfrak{M}}(z) = \sum_{\mathfrak{L} \subseteq \mathfrak{M}} (-1)^{(\mathfrak{L})} p_{\mathfrak{L}}^{-z} (A_{\mathfrak{M}-\mathfrak{L}} B_{\mathfrak{L}})^{-z} \qquad (\mathbf{0} \subset \mathfrak{M} \subseteq \mathfrak{N}), \tag{9}$$

[4] Wenn eine (komplexe) Funktion identisch gleich ∞ ist, so sagen wir, daß sie überall einen Pol hat.

die übrigens nur für $z=1$ verwendet werden. Für diesen Fall lautet (9) so:

$$\Delta_{\mathfrak{M}}(1)=\sum_{\mathfrak{L}\subseteq\mathfrak{M}}(-1)^{(\mathfrak{L})}p_{\mathfrak{L}}^{-1}(A_{\mathfrak{M}-\mathfrak{L}}B_{\mathfrak{L}})^{-1} \qquad (\mathbf{0}\subset\mathfrak{M}\subseteq\mathfrak{N}). \tag{10}$$

Es wird sich zeigen, daß sich die Prämisse des Satzes von HAJÓS mit Hilfe dieser $\Delta_{\mathfrak{M}}(1)$ ausdrücken läßt.

Zu diesem Zweck benötigen wir den Ring $\mathfrak{J}(G)$ der Gruppe G über dem Ring $\mathfrak{J}$ der ganzen Zahlen. Die Elemente von $\mathfrak{J}(G)$ lassen sich eindeutig als

$$\sigma=c_1\varepsilon+c_2\gamma_2+\cdots+c_v\gamma_v$$

schreiben, wobei $\gamma_2,\ldots,\gamma_v$ die von ε verschiedenen Elemente von G und $c_1,\ldots,c_v$ ganze Zahlen sind. Wir setzen

$$|\sigma|_\varepsilon=c_1 \tag{11}$$

und nennen diese Zahl den *Wert* von σ. Offenbar gilt

$$|s\sigma+t\tau|_\varepsilon=s|\sigma|_\varepsilon+t|\tau|_\varepsilon \qquad (s,t\in\mathfrak{J};\ \sigma,\tau\in\mathfrak{J}(G)). \tag{12}$$

Für eine Untermenge H von G, bestehend aus den Elementen $\omega_1,\ldots,\omega_k$, setzen wir

$$[H]=\omega_1+\cdots+\omega_k\ (\in\mathfrak{J}(G)) \tag{13}$$

und nennen es die *Summe* (der Elemente) von H.

Nunmehr besteht zunächst der folgende:

HILFSSATZ 1. *Gilt die Prämisse des Satzes von* HAJÓS *für* $\alpha_1,\ldots,\alpha_n$, *so gilt sie auch für*

$$\alpha_1^{k_1},\ldots,\alpha_n^{k_n} \qquad (p_i\nmid k_i;\ i=1,\ldots,n), \tag{14}$$

ferner gilt dann auch

$$p_i|o(\alpha_i) \qquad (i=1,\ldots,n). \tag{15}$$

Der leichte Beweis findet sich in RÉDEI [3], Hilfssatz 3.

Wir nennen

$$\begin{pmatrix}\alpha_1 & \ldots & \alpha_n\\ p_1 & \ldots & p_n\end{pmatrix} \tag{16}$$

kurz ein *System*. Werden höchstens $n-1$ (eventuell keine) Spalten von (16) gestrichen, so ensteht ein *Teilsystem* von (16). Ist kein Produkt

$$\alpha_1^{k_1}\ldots\alpha_n^{k_n} \qquad (p_i\nmid k_i;\ 0<k_i<o(\alpha_i);\ i=1,\ldots,n) \tag{17}$$

gleich ε, so nennen wir das System (16) *gut*. Wir nennen es *vollständig gut*, wenn alle Teilsysteme gut sind.

HILFSSATZ 2. *Die Prämisse des Satzes von* HAJÓS *ist gleichbedeutend damit, daß*

$$(\{\alpha_1,\ldots,\alpha_n\})=p_1\ldots p_n \tag{18}$$

gilt und das System (16) *vollständig gut ist.*

Denn nehmen wir zuerst die Erfülltheit der Prämisse des Satzes von HAJÓS an. Dann machen die Elemente (1) eben die Gruppe $\{\alpha_1, \ldots, \alpha_n\}$ aus, woraus (18) folgt. Auch folgt nach Hilfssatz 1, daß für die $k_1, \ldots, k_n$ in (14) die $p_1 \ldots p_n$ Produkte

$$\alpha_1^{k_1 x_1} \ldots \alpha_1^{k_n x_n} \qquad (x_1 = 0, \ldots, p_i - 1;\ i = 1, \ldots, n) \tag{19}$$

eine Gruppe bilden. Da (19) für $x_1 = \cdots = x_n = 0$ gleich ε ist, so folgt, daß (19) von ε verschieden ist, wenn jedes $x_1, \ldots, x_n$ eine der Zahlen 0, 1 bedeutet aber nicht alle gleich 0 sind. Da das insbesondere für die k_i mit $p_i \nmid k_i$, $0 < k_i < o(\alpha_i)$ $(i = 1, \ldots, n)$ gilt, so folgt, daß das System (16) vollständig gut ist.

Umgekehrt wenn (18) gilt und (16) vollständig gut ist, so folgt aus letzterem, daß die Produkte in (1) verschieden sind, ferner folgt aus (18), daß diese eben die Gruppe $\{\alpha_1, \ldots, \alpha_n\}$ von der Ordnung $p_1 \ldots p_n$ ausmachen, also die Prämisse des Satzes von HAJÓS erfüllt ist. Somit haben wir Hilfssatz 2 bewiesen.

HILFSSATZ 3. *Im Fall* (15) *ist das System* (16) *dann und nur dann gut, wenn für*

$$A_i = \{\alpha_i\} \qquad (i = 1, \ldots, n) \tag{20}$$

die Bedingung

$$\left| \prod_{i=1}^{n} ([A_i] - [B_i]) \right|_{\varepsilon} = 0 \tag{21}$$

erfüllt ist.

Aus (15), (20) folgt nämlich, daß $B_i = A_i^{p_i} = \{\alpha_i^{p_i}\}$ aus den Elementen

$$\alpha_i^{k_i} \qquad (p_i \mid k_i;\ 0 \leqq k_i < o(\alpha_i))$$

besteht $(i = 1, \ldots, n)$. Weiter folgt hieraus, daß jetzt das Produkt in (21) eben die Summe der Elemente (17) ist. Dies bedeutet die Richtigkeit vom Hilfssatz 3.

HILFSSATZ 4. *Wenn*

$$p_i \mid (A_i) \qquad (i = 1, \ldots, n) \tag{22}$$

ist, so gilt

$$\left| \prod_{i=1}^{n} ([A_i] - [B_i]) \right|_{\varepsilon} = (A_1) \cdots (A_n) \Delta_{\mathfrak{N}}(1). \tag{23}$$

Offenbar gilt nämlich für das „Anfangsglied" des (ausmultiplizierten) Produktes auf der linken Seite von (23):

$$\prod_{i=1}^{n} [A_i] = \frac{(A_1) \cdots (A_n)}{(A_1 \cdots A_n)} [A_1 \cdots A_n].$$

(Dies wäre ja sogar für beliebige, nicht nur für zyklische Untergruppen von

G richtig.) Wegen $A_1 \cdots A_n = A_{\mathfrak{N}}$ und $|[A_{\mathfrak{N}}]|_\varepsilon = 1$ gilt also

$$\left| \prod_{i=1}^{n} [A_i] \right|_\varepsilon = \frac{(A_1) \cdots (A_n)}{(A_{\mathfrak{N}})}. \tag{24}$$

Andererseits berücksichtige man, daß (22) und

$$(B_i) = \frac{1}{p_i}(A_i) \qquad (i = 1, \ldots, n) \tag{25}$$

gleichbedeutend sind. Wenn man das Produkt auf der linken Seite von (23) ausmultipliziert und auf die so entstandenen 2^n Glieder die Formel (24) anwendet, so bekommt man für die linke Seite von (23) wegen (12) und (25) den Wert

$$\sum_{\mathfrak{M} \subseteq \mathfrak{N}} (-1)^{(\mathfrak{M})} \frac{(A_1) \cdots (A_n)}{p_{\mathfrak{M}} (A_{\mathfrak{N}-\mathfrak{M}} B_{\mathfrak{M}})}.$$

Dies stimmt nach (10) mit der rechten Seite von (23) überein. Das beweist Hilfssatz 4.

HILFSSATZ 5. *Im Fall* (15) *ist das System* (16) *dann und nur dann vollständig gut, wenn für* (20) *die Bedingung*

$$\Delta_{\mathfrak{M}}(1) = 0 \qquad (\mathbf{0} \subset \mathfrak{M} \subseteq \mathfrak{N})$$

erfüllt ist.

Da nämlich aus (15) und (20) das Bestehen von (22) folgt und in diesem Falle nach Hilfssatz 4 die Gleichung (21) mit $\Delta_{\mathfrak{N}}(1) = 0$ gleichbedeutend ist, so folgt aus Hilfssatz 3 die Richtigkeit vom Hilfssatz 5.

HILFSSATZ 6. *Die Prämisse des Satzes von* HAJÓS *ist äquivalent damit, daß für* (20)

$$\frac{1}{p_{\mathfrak{N}}} - \frac{1}{(A_{\mathfrak{N}})} = 0, \quad \Delta_{\mathfrak{M}}(1) = 0 \qquad (\mathbf{0} \subset \mathfrak{M} \subseteq \mathfrak{N}) \tag{26}$$

gelten.

Dem Beweis schicken wir die Bemerkung voran, daß nach (10) insbesondere für $\mathfrak{M} = i$

$$\Delta_i(1) = \frac{1}{(A_i)} - \frac{1}{p_i(B_i)} \qquad (i = 1, \ldots, n) \tag{27}$$

gilt, also in (26_2) die Bedingung (22) enthalten ist. Andererseits ist in der Prämisse des Satzes von HAJÓS die Bedingung (15) enthalten. Man bemerke auch, daß (22), (15) im Fall (20) gleichbedeutend sind. Hieraus und aus den Hilfssätzen 2, 5 folgt die Richtigkeit vom Hilfssatz 6.

Wir verwenden für $\Delta_{\mathfrak{N}}(z)$ auch die ausführlichere Bezeichnung

$$\Delta_{\mathfrak{N}}(z) = \Delta(z; A_1, \ldots, A_n). \tag{28}$$

(In der Wirklichkeit hängt $\Delta_{\mathfrak{N}}(z)$ auch von $p_1, \ldots, p_n$ ab.)

HILFSSATZ 7. *Ist*

$$p_n \nmid (A_n), \tag{29}$$

so gilt

$$\Delta_{\mathfrak{N}}(z) = (1-p_n^{-z})\,(A_n)^{-z}\,\Delta(z; A_1A_n/A_n, \ldots, A_{n-1}A_n/A_n). \tag{30}$$

Nach (9) ist nämlich

$$\Delta_{\mathfrak{N}}(z) = \sum_{\mathfrak{M}\subseteq\mathfrak{N}} (-1)^{(\mathfrak{M})} p_{\mathfrak{M}}^{-z} (A_{\mathfrak{N}-\mathfrak{M}} B_{\mathfrak{M}})^{-z}. \tag{31}$$

Wir bezeichnen mit $\mathfrak{N}'$ die Menge der Zahlen $1, \ldots, n-1$. Durchläuft $\mathfrak{M}$ die Untermengen von $\mathfrak{N}'$ und bezeichnet $\mathfrak{M}_1$ die Vereinigungsmenge $\mathfrak{M} \cup n$, so durchlaufen $\mathfrak{M}, \mathfrak{M}_1$ insgesamt alle verschiedenen Untermengen von $\mathfrak{N}$. Da (29) mit $B_n = A_n$ gleichbedeutend ist, so gilt dabei stets

$$A_{\mathfrak{N}-\mathfrak{M}_1} B_{\mathfrak{M}_1} = A_{\mathfrak{N}-\mathfrak{M}} B_{\mathfrak{M}}.$$

Somit folgt aus (31) wegen $(\mathfrak{M}_1) = (\mathfrak{M}) + 1$, $p_{\mathfrak{M}} = p_n p_{\mathfrak{M}}$

$$\Delta_{\mathfrak{N}}(z) = (1-p_n^{-z}) \sum_{\mathfrak{M}\subseteq\mathfrak{N}'} (-1)^{(\mathfrak{M})} p_{\mathfrak{M}}^{-z} (A_{\mathfrak{N}-\mathfrak{M}} B_{\mathfrak{M}})^{-z}. \tag{32}$$

Nun gilt auch

$$(A_{\mathfrak{N}-\mathfrak{M}} B_{\mathfrak{M}}) = (A_{\mathfrak{N}'-\mathfrak{M}} B_{\mathfrak{M}} A_n) = (A_n)\,(A_{\mathfrak{N}'-\mathfrak{M}} B_{\mathfrak{M}} A_n/A_n).$$

Wird dies in (32) eingesetzt, so entsteht eben (30). Das beweist Hilfssatz 7.

HILFSSATZ 8. *Stets gilt*

$$\Delta_{\mathfrak{M}}(1) \geqq 0 \qquad (\mathbf{0} \subset \mathfrak{M} \subseteq \mathfrak{N}). \tag{33}$$

Es genügt $\Delta_{\mathfrak{N}}(1) \geqq 0$ zu beweisen.

Im Fall $n = 1$ ist nach (27)

$$\Delta_1(1) = \frac{1}{(A_1)} - \frac{1}{p_1(B_1)}.$$

Da stets $(A_1) \leqq p_1(B_1)$ ist, so ist die Behauptung für diesen Fall richtig. Im Fall $n \geqq 2$ setzen wir die Richtigkeit der Behauptung für $n-1$ voraus. Zuerst nehmen wir an, daß auch (22) gilt. Da die linke Seite von (23) wegen $B_i \subset A_i$ $(i = 1, \ldots, n)$ nichtnegativ ist, so folgt aus Hilfssatz 4 jedenfalls $\Delta_{\mathfrak{N}}(1) \geqq 0$. Wenn dagegen (22) falsch ist, so dürfen wir annehmen, daß eben $p_n \nmid (A_n)$ ist. Jetzt folgt die Behauptung aus der Induktionsvoraussetzung und aus Hilfssatz 7. Somit haben wir Hilfssatz 8 bewiesen.

HILFSSATZ 9. *Es gilt die Indentität:*

$$\frac{1}{p_{\mathfrak{N}}^z} - \frac{1}{(A_{\mathfrak{N}})^z} + \sum_{\mathbf{0}\subset\mathfrak{M}\subseteq\mathfrak{N}} \frac{1}{p_{\mathfrak{N}-\mathfrak{M}}^z}\,\Delta_{\mathfrak{M}}(z) = \sum_{\mathfrak{M}\subseteq\mathfrak{N}} \frac{1}{p_{\mathfrak{N}-\mathfrak{M}}^z (A_{\mathfrak{M}})^z}\,\varrho_{\mathfrak{M}}(z). \tag{34}$$

Es genügt den Beweis für den Fall $z = 1$ auszuführen, da sich der allgemeine Fall ebenso beweisen läßt. (Übrigens werden wir (34) nur mit

$z=1$ anwenden.) Für $z=1$ lautet (34) nach (5) und (10) so:

$$\frac{1}{p_{\mathfrak{N}}}-\frac{1}{(A_{\mathfrak{N}})}+\sum_{\mathbf{0}\subset\mathfrak{M}\subseteq\mathfrak{N}}\sum_{\mathfrak{L}\subseteq\mathfrak{M}}(-1)^{(\mathfrak{L})}\frac{1}{p_{\mathfrak{N}-\mathfrak{M}}p_{\mathfrak{L}}A_{\mathfrak{M}-\mathfrak{L}}B_{\mathfrak{L}})}= \tag{35}$$

$$=\sum_{\mathfrak{M}\subset\mathfrak{N}}\sum_{\mathfrak{L}\subseteq\mathfrak{N}-\mathfrak{M}}(-1)^{(\mathfrak{L})}\frac{1}{p_{\mathfrak{N}-\mathfrak{M}}(A_{\mathfrak{M}}B_{\mathfrak{L}})}.$$

Links und rechts durchläuft $\mathfrak{L}$ alle Untermengen von $\mathfrak{N}$. Nach leichter Umordnung der Glieder und Berücksichtigung von $p_{\mathfrak{N}-\mathfrak{M}}p_{\mathfrak{L}}=p_{\mathfrak{N}-(\mathfrak{M}-\mathfrak{L})}$ lautet (35) so:

$$\frac{1}{p_{\mathfrak{N}}}+\sum_{\mathfrak{L}\subseteq\mathfrak{N}}\sum_{\substack{\mathfrak{L}\subseteq\mathfrak{M}\subseteq\mathfrak{N}\\ \mathfrak{M}\neq\mathbf{0}}}(-1)^{(\mathfrak{L})}\frac{1}{p_{\mathfrak{N}-(\mathfrak{M}-\mathfrak{L})}(A_{\mathfrak{M}-\mathfrak{L}}B_{\mathfrak{L}})}= \tag{36}$$

$$=\frac{1}{(A_{\mathfrak{N}})}\sum_{\mathfrak{L}\subseteq\mathfrak{N}}\sum_{\substack{\mathfrak{M}\subseteq\mathfrak{N}-\mathfrak{L}\\ \mathfrak{M}\neq\mathfrak{N}}}(-1)^{(\mathfrak{L})}\frac{1}{p_{\mathfrak{N}-\mathfrak{M}}(A_{\mathfrak{M}}B_{\mathfrak{L}})},$$

wo man links und rechts in der inneren Summe über $\mathfrak{M}$ zu summieren hat. Man kann auf beiden Seiten das erste Glied in die Summe einschmelzen, so daß man links auch $\mathfrak{M}=\mathbf{0}$ (für $\mathfrak{L}=\mathbf{0}$) bzw. rechts auch $\mathfrak{M}=\mathfrak{N}$ (für $\mathfrak{L}=\mathbf{0}$) zuläßt. Dann wird die Richtigkeit von (36) augenscheinlich, womit Hilfssatz 9 bewiesen ist.

HILFSSATZ 10. *Im Fall*

$$(A_{\mathfrak{N}})=p_{\mathfrak{N}} \tag{37}$$

gelten dann und nur dann alle Gleichungen

$$\Delta_{\mathfrak{M}}(1)=0 \qquad (\mathbf{0}\subset\mathfrak{M}\subseteq\mathfrak{N}), \tag{38}$$

wenn alle Gleichungen

$$\varrho_{\mathfrak{M}}(1)=0 \qquad (\mathfrak{M}\subset\mathfrak{N}) \tag{39}$$

gelten.

Nach meinem Trägheitssatz (RÉDEI [2], Satz 9) gilt nämlich stets $0\leqq\varrho(z)$ $(z=1,2,\ldots)$. Insbesondere folgt hieraus

$$\varrho_{\mathfrak{M}}(1)\geqq 0 \qquad (\mathfrak{M}\subset\mathfrak{N}). \tag{40}$$

Wenn nun (37) gilt, so fallen in (34) die ersten zwei Glieder heraus. Wird also Hilfssatz 9 mit $z=1$ angewendet, so folgt aus ihm wegen Hilfssatz 8 und (40) die Richtigkeit vom Hilfssatz 10.

Nunmehr beweisen wir die Äquivalenz vom Satz 1′ mit dem von HAJÓS wie folgt. Die Prämisse des Satzes von HAJÓS ist nach den Hilfssätzen 6, 10 äquivalent mit der vom Satz 1′ für die $A_1,\ldots,A_n$ in (20).[5] Dabei folgt nach Hilfssatz 1 aus der Prämisse des Satzes von HAJÓS das Bestehen von (22)

[5] Man beachte auch, daß durch (20) alle überhaupt möglichen Fälle für die $A_1,\ldots,A_n$ erschöpft sind.

für die gesagten $A_1, \ldots, A_n$, somit läßt sich dann die Konklusion dieses Satzes auch so aussprechen, daß mindestens ein B_i $(i = 1, \ldots, n)$ gleich ε ist. Somit ist die Äquivalenz vom Satz 1′ mit dem von HAJÓS bewiesen.

(Eingegangen am 23. September 1955.)

Literaturverzeichnis

[1] G. HAJÓS, Über einfache und mehrfache Bedeckung des n-dimensionalen Raumes mit einem Würfelgitter, *Math. Zeitschrift*, **47** (1942), S. 427—467.

[2] L. RÉDEI, Zetafunktionen in der Algebra, *Acta Math. Acad. Sci. Hung.*, **6** (1955), S. 5—25.

[3] L. RÉDEI, Neuer Beweis des Hajósschen Satzes über die endlichen Abelschen Gruppen, Ebenda, S. 27—40.

ζ-ФУНКЦИИ В ТЕОРИИ ГРУПП И ТЕОРЕМА ХАЙОША

Л. Редеи (Сегед)

(Резюме)

Теоретико-групповая теорема Хайоша равносильна теореме, относящейся к выдвинутым недавно автором ζ-функциям (см. Редеи [2]). Согласно этой теореме если некоторые ζ-функции имеют полюс в точке $z = 1$, то одна определенная из них тождественно бесконечна. Доказательство этого основывается на теореме инерции конечных абелевых групп, которая была доказана в упомянутой статье тоже с помощью ζ-функций. Этим путем можно ожидать новое доказательство теоремы Хайоша, быть может путем дальнейшего развития теории упомянутых ζ-функций.

GENERALIZATION OF AN INEQUALITY OF KOLMOGOROV

By

J. HÁJEK (Prague) and A. RÉNYI (Budapest), corresponding member of the Academy

In what follows $\mathbf{P}(A)$ denotes the probability of the event A, $\mathbf{M}(\xi)$ the mean value, and $\mathbf{M}(\xi|A)$ the conditional mean value, under the condition A, of the random variable ξ.

The inequality of A. N. KOLMOGOROV [1] in question states that if $\xi_1, \xi_2, \ldots, \xi_k, \ldots$ is a sequence of mutually independent random variables with mean values $\mathbf{M}(\xi_k) = 0$ and finite variances $\mathbf{M}(\xi_k^2) = D_k^2 \quad (k = 1, 2, \ldots)$, we have, for any $\varepsilon > 0$,

$$\mathbf{P}\Big(\max_{1 \leqq k \leqq m} |\xi_1 + \xi_2 + \cdots + \xi_k| \geqq \varepsilon\Big) \leqq \frac{1}{\varepsilon^2} \sum_{k=1}^{m} D_k^2. \tag{1}$$

This inequality is extremely useful in proving the strong law of large numbers and related theorems. In what follows the inequality (2) will be proved which contains (1) as a special case. The use of (2) instead of (1) makes it possible to simplify the proofs mentioned.

The inequality (2) has been found in 1953 by the first named author; the proof of the inequality given below is due to the second named author.

The following theorem will be proved:

THEOREM. *If $\xi_1, \xi_2, \ldots, \xi_k, \ldots$ is a sequence of mutually independent random variables with mean values $\mathbf{M}(\xi_k) = 0$ and finite variances $\mathbf{M}(\xi_k^2) = D_k^2$ $(k = 1, 2, \ldots)$ and $c_k \quad (k = 1, 2, \ldots)$ is a non-increasing sequence of positive numbers, we have for any $\varepsilon > 0$ and any positive integers n and m $(n < m)$*

$$\mathbf{P}\Big(\max_{n \leqq k \leqq m} c_k |\xi_1 + \xi_2 + \cdots + \xi_k| \geqq \varepsilon\Big) \leqq \frac{1}{\varepsilon^2}\left(c_n^2 \sum_{k=1}^{n} D_k^2 + \sum_{k=n+1}^{m} c_k^2 D_k^2\right). \tag{2}$$

PROOF. Let us put

$$\zeta = \sum_{k=n}^{m-1} (\xi_1 + \xi_2 + \cdots + \xi_k)^2 (c_k^2 - c_{k+1}^2) + c_m^2 (\xi_1 + \cdots + \xi_m)^2. \tag{3}$$

It follows

$$\mathbf{M}(\zeta) = c_n^2 \sum_{k=1}^{n} D_k^2 + \sum_{k=n+1}^{m} c_k^2 D_k^2. \tag{4}$$

Denoting by A_r $(r = n, n+1, \ldots, m)$ the event consisting in the simultaneous

validity of the inequalities[1]

$$c_s|\xi_1+\cdots+\xi_s|<\varepsilon \quad (n\leqq s<r) \quad \text{and} \quad c_r|\xi_1+\cdots+\xi_r|\geqq\varepsilon,$$

the inequality (2) can be written in the form

$$\sum_{r=n}^{m}\mathbf{P}(A_r)\leqq\frac{1}{\varepsilon^2}\mathbf{M}(\zeta). \tag{5}$$

The inequality (5) is the consequence of (6a), (6b) and (6c) of which the first two are evident:

$$\mathbf{M}(\zeta)\geqq\sum_{r=n}^{m}\mathbf{M}(\zeta|A_r)\,\mathbf{P}(A_r), \tag{6a}$$

$$\mathbf{M}(\zeta|A_r)\geqq\sum_{k=r}^{m-1}\mathbf{M}((\xi_1+\cdots+\xi_k)^2|A_r)\,(c_k^2-c_{k+1}^2)+c_m^2\,\mathbf{M}((\zeta_1+\cdots+\zeta_m)^2|A_r), \tag{6b}$$

$$\mathbf{M}((\xi_1+\cdots+\xi_k)^2|A_r)\geqq\mathbf{M}((\xi_1+\cdots+\xi_r)^2|A_r)\geqq\frac{\varepsilon^2}{c_r^2} \qquad (r\leqq k\leqq m). \tag{6c}$$

In verifying the inequality (6c), one has to use the fact that according to the definition of the event A_r the random variables ξ_k $(k>r)$ are mutually independent of each other also under the condition A_r, further that they are independent also under the condition A_r of the variables $\xi_1,\xi_2,\ldots,\xi_r$. It should be mentioned that the same fact is utilised in the proof of (1) due to KOLMOGOROV [1].

REMARK 1. If we choose $n=1$ and $c_1=c_2=\cdots=c_m=1$, we obtain from (2) as a special case the inequality (1). If we choose $c_k=\frac{1}{k}$ $(k=n,n+1,\ldots,m)$, we obtain the inequality

$$\mathbf{P}\left(\max_{n\leqq k\leqq m}\frac{|\xi_1+\xi_2+\cdots+\xi_k|}{k}\geqq\varepsilon\right)\leqq\frac{1}{\varepsilon^2}\left(\frac{\sum_{k=1}^{n}D_k^2}{n^2}+\sum_{k=n+1}^{m}\frac{D_k^2}{k^2}\right). \tag{7}$$

REMARK 2. By means of passing to the limit $m\to\infty$, it is easy to deduce from (2) the following inequality:

$$\mathbf{P}(\sup_{n\leqq k}c_k|\xi_1+\cdots+\xi_k|\geqq\varepsilon)\leqq\frac{1}{\varepsilon^2}\left(c_n^2\sum_{k=1}^{n}D_k^2+\sum_{k=n+1}^{\infty}c_k^2D_k^2\right). \tag{8}$$

Especially, if $c_k=\frac{1}{k}$ $(k=1,2,\ldots)$, we obtain the inequality

$$\mathbf{P}\left(\sup_{n\leqq k}\frac{|\xi_1+\xi_1+\cdots+\xi_k|}{k}\geqq\varepsilon\right)\leqq\frac{1}{\varepsilon^2}\left(\frac{\sum_{k=1}^{n}D_k^2}{n^2}+\sum_{k=n+1}^{\infty}\frac{D_k^2}{k^2}\right). \tag{9}$$

[1] If $r=n$, only the second inequality is supposed.

REMARK 3. It follows from (9) immediately, that the strong law of large numbers holds for the sequence of mutually independent random variables $\xi_1, \xi_2, \ldots, \xi_k, \ldots$ if the ξ_k's have mean values 0, finite variances $D_k^2 = M(\xi_k^2)$ and

$$\sum_{k=1}^{\infty} \frac{D_k^2}{k^2} \tag{10}$$

converges ([2]).

As a matter of fact, it follows from (9) and (10) that for any $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbf{P}\left(\sup_{n \leqq k} \frac{|\xi_1 + \cdots + \xi_k|}{k} \geqq \varepsilon \right) = 0 \tag{11}$$

and therefore we have

$$\mathbf{P}\left(\lim_{n\to\infty} \frac{\xi_1 + \xi_2 + \cdots + \xi_n}{n} = 0 \right) = 1. \tag{12}$$

(Received 25 July 1955)

References

[1] A. KOLMOGOROFF Über die Summen durch den Zufall bestimmten unabhängigen Größen, *Math Ann.*, **99** (1928), pp. 309—319. (See also the corrections, ibid., **102** (1929), pp. 484—488.)

[2] A. KOLMOGOROFF, Sur la loi forte des grands nombres, *Comptes Rendus Acad. Sci. Paris* **191** (1930), pp. 910—912. (See also A. KOLMOGOROFF, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergebnisse der Math. 2., no. 3 (Berlin, 1933).)

ОБОБЩЕНИЕ ОДНОГО НЕРАВЕНСТВА А. Н. КОЛМОГОРОВА

Я. Хаек (Прага) и А. Реньи (Будапешт)

(Резюме)

В работе доказывается следующее неравенство:

Пусть $\xi_1, \xi_2, \ldots, \xi_m$ независимые случайные величины, математические ожидания которых равны нулю ($\mathbf{M}(\xi_k) = 0$; $k = 1, 2, \ldots, m$), а дисперсии конечни ($\mathbf{D}^2(\xi_k) = D_k^2$; $k = 1, 2, \ldots, m$); пусть, далее, c_k не возрастающая последовательность положительных чисел. Тогда если $1 \leqq n \leqq m$ и $\varepsilon > 0$, имеем

$$\mathbf{P}\left(\max_{n \leqq k \leqq m} c_k |\xi_1 + \cdots + \xi_k| \geqq \varepsilon \right) \leqq \frac{1}{\varepsilon^2} \left(c_n^2 \sum_{k=1}^{n} D_k^2 + \sum_{k=n+1}^{m} c_k^2 D_k^2 \right). \tag{2}$$

Если $n = 1$ и $c_k \equiv 1$, то в качестве частного случая получаем хорошо известное неравенство (1) А. Н. Колмогорова. Воспользовавшись неравенством (2), можно упростить доказательство усиленных законов больших чисел.

ON A NEW AXIOMATIC THEORY OF PROBABILITY

By

ALFRÉD RÉNYI (Budapest), corresponding member of the Academy

Dedicated to Professors L. FEJÉR *and* F. RIESZ *on their 75th birthday*

Introduction

The axiomatic foundation of probability theory given by A. N. KOLMOGOROV [1] in the year 1933 has been the starting point of a new and brilliant period in the development of probability theory. According to this theory to every situation (experiment, observation etc.) in which chance plays a role, there corresponds a *probability space* $[S, \mathcal{A}, \mathbf{P}]$, i. e. an abstract space S (the space of elementary events), a σ-algebra $\mathcal{A}$ (the set of events) of subsets of S, and a measure $\mathbf{P}(A)$ (the probability of the event A) defined for $A \in \mathcal{A}$ and satisfying $\mathbf{P}(S) = 1$. The theory of KOLMOGOROV furnished an appropriate and mathematically exact basis for the rapid development of probability theory, which took place in the last 30 years, as well as for its fruitfull application in a great number of branches of science, including other chapters of mathematics too. Nevertheless in the course of development there arose some problems concerning probability which can not be fitted into the frames of the theory of KOLMOGOROV.

The common feature of these problems is that in them *unbounded measures* occur, while in the theory of KOLMOGOROV probability is a bounded measure normed by the condition $\mathbf{P}(S) = 1$. Unbounded measures occur in statistical mechanics, quantum mechanics, in some problems of mathematical statistics (for example in connection with the application of the theorem of Bayes) as limiting distributions of Markov chains and Markov processes, in integral geometry, in connection with the applications of probability concepts in number theory etc.

In the theory of KOLMOGOROV it has, for instance, no sense to speak about a probability distribution which is uniform on the whole real axis or on the whole plane, further it has no sense to say that we choose an integer in such a way that all integers (or all non-negative integers) are equiprobable.

At the first glance it seems that unbounded measures can play no role in probability theory, because, in view of the connection between probability and relative frequency, probability clearly can not take any value greater than 1. But if we observe more attentively how unbounded measures are really used in all cases mentioned above, we see that unbounded measures are

used only to calculate conditional probabilities as the quotient of the values of the unbounded measure of two sets (the first being contained in the second) and in this way reasonable values (not exceeding 1) are obtained. This is the reason why unbounded measures can be used with success in calculating (conditional) probabilities. But as the use of unbounded measures can not be justified in the theory of KOLMOGOROV, the necessity arises to generalize this theory. In the present paper such an attempt is made.

Clearly in a theory in which unbounded measures are allowed, *conditional probability must be taken as the fundamental concept.* This is natural also from an other point of view. In fact, the probability of an event depends essentially on the circumstances under which the event possibly occurs, and it is a commonplace to say that in reality every probability is conditional.

This has been realized by several authors; I mention here without aiming at completeness only H. JEFFREYS [2], H. REICHENBACH [3], J. KEYNES [4], R. KOOPMAN [5], A. COPELAND [6], G. A. BARNARD [7] and I. J. GOOD [8].

But none of the mentioned authors developed his theory on a measure-theoretic basic. The axiomatic theory developed in the present paper combines the measure-theoretic treatment of KOLMOGOROV with the idea proposed by the authors mentioned (and also by others) to consider conditional probability as the fundamental concept.

The novelty of the theory lies only in this combination. It follows from what has been said that in developing the theory proposed in this paper we follow the same way as KOLMOGOROV, only we try to go one step further. Thus, this new theory should be considered as a generalization of that of KOLMOGOROV. In fact, it contains the theory of KOLMOGOROV as a special case, but includes also cases which can not be fitted into the theory of KOLMOGOROV, namely cases in which conditional probabilities are calculated by means of unbounded measures.

Among the authors mentioned above only H. JEFFREYS uses explicitely unbounded probability distributions (especially random variables ξ for which $\log \xi$ is uniformly distributed on the whole real axis) but he does not give an exact mathematical meaning to such distributions, and restricts himself to the remark that it is a mere convention that to a certain event there corresponds the probability 1; he says that in some cases instead of 1 the value $+\infty$ can also be taken.

The attempt to extend the scope of the mathematical theory of probability with the aim to give a well founded basis for such calculating procedures which were successfully used in many fields of applications without being mathematically rigorously justified, has many analogues in the history of mathematics. I mention only, as one of the latest such successful attempts, that of S. L. SOBOLEV and L. SCHWARTZ, concerning the generalization of the

notion of a function, to include, for instance, the function of DIRAC, together with its derivatives etc.[1]

The theory as presented in this paper is still far from being fully developed; we restrict ourselves to give here only the axioms and their immediate consequences (§ 1), further to discuss some examples (§ 2) and applications (§ 3), and finally to develop some "conditional" laws of large numbers (§ 4). As an application of the conditional strong law of large numbers, a generalization of BOREL's theorem on normal decimals for CANTOR's series is given.

I hope to return to the questions left unsolved in a forthcoming publication.

I had the occasion to lecture on the present theory in 1954 at several congresses: in Budapest [9] at the General Assembly of the Hungarian Academy in May 1954, at the Conference on probability and statistics in Prague in June 1954, at the International Congress in Amsterdam [10] in September 1954, at the Conference on probability and statistics in Berlin [11] in October 1954 and at the Conference on stochastic processes in Wrocław in December 1954. At these (and other) occasions many valuable remarks have been made. I mention only the following ones: B. V. GNEDENKO kindly informed me in Prague in June 1954 that in a lecture held some years ago in Moscow A. N. KOLMOGOROV himself has put forward the idea to develop his theory in such a manner that conditional probability should be taken as the fundamental concept, but he never published his ideas regarding this question. I was glad to learn from this information that my attempt, besides of being a continuation of the work of KOLMOGOROV, follows the lines which have been pointed out by himself. In § 3 an inequality of J. HÁJEK is used which he communicated to me in Prague in June 1954. A proof of this inequality is published in a joint paper of the author and J. HÁJEK [12] in this volume. Some interesting measure-theoretic problems, which arose in connection with the present paper, have been solved by Á. CSÁSZÁR: his results, which he

[1] The Dirac δ-"function" and the uniform distribution on the whole real axis (or the whole space etc.) arise in the same way in quantum mechanics and they are in a certain sense dual to each other. As a matter of fact, it is easy to verify the following fact connected with Heisenberg's uncertainty relation (for the sake of simplicity we restrict ourselves to the one dimensional case): if the wave function $\psi(x)$ of the position of a particle degenerates into a Dirac δ-"function", the wave function $\varphi(p)$ of the impulse of the particle degenerates into a function for which $|\varphi(p)|^2 \equiv \text{const.}$ is the "density" of a "uniform probability distribution in the whole phase space". This can be shown as follows: as it is known, denoting by $\varphi(p)$ the wave function of the position and by $\varphi(p)$ the wave function of the impulse of a particle, we have $\varphi(p) = \frac{1}{\sqrt{2\pi h}} \int_{-\infty}^{+\infty} \psi(x) e^{-\frac{ipx}{h}} dx$. This formula shows that if $\psi(x)$ degenerates into the Dirac function $\delta(x-x_0)$, we have $\varphi(p) = \frac{1}{\sqrt{2\pi h}} e^{-\frac{ipx_0}{h}}$ for which $|\varphi(p)|^2 = \text{const.}$

exposed at the first time as comments to my lecture [9], are published in his paper [13] in this volume and settle the question under what conditions can the conditional probability, introduced in the present paper as a set function of two set variables, be expressed in quotient form by means of (one or more) set functions of one variable. D. VAN DANTZIG called my attention to the work of KEYNES, COPELAND and KOOPMAN. JU. V. LINNIK called my attention to a paper of J. NEYMAN [14], where the foundations of the theory of probability are sketched in a form which has some points of contact with the point of view of the present paper. In Berlin A. N. KOLMOGOROV called my attention to those conditional probability spaces which I call "Cavalieri spaces" (§ 2). I hope to return to the thorough investigation of such spaces in a forthcoming paper. I owe to a discussion with E. MARCZEWSKI Theorem 14 of § 2. The result of 3. 5 has been suggested to the author by a remark of K. SARKADI. Á. CSÁSZÁR and J. CZIPSZER kindly read the manuscript of the present paper and made some valuable remarks which I have utilized in preparing the final form of this paper.

I am thankful to all those mentioned for their remarks and suggestions.

§ 1. The axioms and their immediate consequences

1. 1. Notations. In what follows if A and B are sets, we denote by $A+B$ the sum of the sets A and B (i. e. the set of those elements which belong at least to one of the sets A and B); AB denotes the product of the sets A and B (i. e. the set of those elements which belong to both of the sets A and B); to denote the sum resp. the product of a finite or infinite family of sets, we use also the notation Σ resp. Π. The empty set will be denoted by $\bigcirc$; $A \subseteq B$ expresses the fact that A is a subset of B; the subset of B consisting of those elements of B which do not belong to A will be denoted by $B-A$. If a is an element of the set A, this will be denoted by $a \in A$. If a does not belong to the set A, this will be denoted by $a \notin A$.

1. 2. Definitions and axioms. Let us be given an arbitrary set S; the elements of S which will be denoted by small letters $a, b, \ldots$ will be called elementary events. Let $\mathcal{A}$ denote a σ-algebra of subsets of S; the subsets of S which are elements of $\mathcal{A}$ will be denoted by capital letters $A, B, C, \ldots$ and called random events, or simply events. [The supposition that $\mathcal{A}$ is a σ-algebra means (see [15], p. 28) that 1. if $A_n \in \mathcal{A}$ $(n = 1, 2, \ldots)$, we have $\sum_{n=1}^{\infty} A_n \in \mathcal{A}$; 2. if $A \in \mathcal{A}$, we have $S - A \in \mathcal{A}$; 3. $\mathcal{A}$ is not empty. This implies that $\bigcirc \in \mathcal{A}$ and $S \in \mathcal{A}$.] Let us suppose further that a non empty subset $\mathfrak{B}$

of $\mathfrak{A}$ is given; we do not suppose any restrictions regarding the set $\mathfrak{B}$.[2] We suppose finally that a set function $\mathbf{P}(A|B)$ of two set variables is defined for $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$; $\mathbf{P}(A|B)$ will be called the conditional probability of the event A with respect to the event B. As the conditional probability of the event $A \in \mathfrak{A}$ with respect to the event B is defined if and only if B belongs to $\mathfrak{B}$, $\mathfrak{B}$ may be called the set of possible conditions. We suppose that the set function $\mathbf{P}(A|B)$ satisfies the following axioms:

AXIOM I. $\mathbf{P}(A|B) \geqq 0$, *if* $A \in \mathfrak{A}$ *and* $B \in \mathfrak{B}$; *further* $\mathbf{P}(B|B) = 1$, *if* $B \in \mathfrak{B}$.

AXIOM II. *For any fixed* $B \in \mathfrak{B}$, $\mathbf{P}(A|B)$ *is a measure, i. e. a countably additive set function of* $A \in \mathfrak{A}$, *i. e. if* $A_n \in \mathfrak{A}$ $(n = 1, 2, \ldots)$ *and* $A_j A_k = O$ *for* $j \neq k$ $(j, k = 1, 2, \ldots)$, *we have*

$$\mathbf{P}\left(\sum_{n=1}^{\infty} A_n \,\middle|\, B\right) = \sum_{n=1}^{\infty} \mathbf{P}(A_n|B).$$

AXIOM III.[3] *If* $A \in \mathfrak{A}$, $B \in \mathfrak{A}$, $C \in \mathfrak{B}$, *and* $BC \in \mathfrak{B}$, *we have*

$$\mathbf{P}(A|BC) \cdot \mathbf{P}(B|C) = \mathbf{P}(AB|C).$$

If the Axioms I—III are satisfied, we shall call the set S, together with the σ-algebra $\mathfrak{A}$ of subsets of S, the subset $\mathfrak{B}$ of $\mathfrak{A}$ and the set function $\mathbf{P}(A|B)$ defined for $A \in \mathfrak{A}$, $B \in \mathfrak{B}$, a *conditional probability space* and denote it for the sake of brevity by $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$.

1. 3. Connection with Kolmogorov's theory. If $\mathbf{P}(A)$ is a measure (i. e. a countably additive and non-negative set function) defined on the σ-algebra $\mathfrak{A}$ of subsets of the set S, if further $\mathbf{P}(S) = 1$, then the triple $[S, \mathfrak{A}, \mathbf{P}(A)]$ is called a probability space in the sense of KOLMOGOROV, and if we define $\mathfrak{A}^*$ as the set of those sets $B \in A$ for which $\mathbf{P}(B) > 0$ and put $\mathbf{P}(A|B) = \dfrac{\mathbf{P}(AB)}{\mathbf{P}(B)}$ for $A \in \mathfrak{A}$, $B \in \mathfrak{A}^*$, clearly $[S, \mathfrak{A}, \mathfrak{A}^*, \mathbf{P}(A|B)]$ is a conditional probability space which will be called the conditional probability space generated by the probability space $[S, \mathfrak{A}, \mathbf{P}(A)]$.

Conversely, if $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ is a conditional probability space and C is an arbitrary element of $\mathfrak{B}$, putting $\mathrm{P}_C(A) = \mathbf{P}(A|C)$, $[S, A, \mathrm{P}_C(A)]$ will be a probability space in the sense of KOLMOGOROV. Thus a conditional probability space is nothing else than a set of ordinary probability spaces which

[2] It will be seen that our axioms imply that $O \notin \mathfrak{B}$, but it is possible that $\mathfrak{B}$ contains all elements of $\mathfrak{A}$ except O; on the other hand, it is possible that $\mathfrak{B}$ contains only one set. The theory is somewhat less general, but considerably simpler if it is supposed that $\mathfrak{B}$ is an additive class of sets, i. e. if from $B_1 \in \mathfrak{B}$ and $B_2 \in \mathfrak{B}$ it follows $B_1 + B_2 \in \mathfrak{B}$. Concerning the consequences of this supposition see [13].

[3] See 1. 7 where an equivalent axiom (Axiom III') is discussed, and 1. 8 where a stronger form of this axiom is mentioned.

are connected with each other by Axiom III. This connection is such that it is in conformity with the usual definition of conditional probability. Namely, if we put $\mathbf{P}_C(A) = \mathbf{P}(A|C)$ for $A \in \mathcal{A}$ with $C \in \mathcal{B}$ fixed, and define the conditional probability $\mathbf{P}^*(A|B)$ for a $B \in \mathcal{B}$ for which $\mathbf{P}_C(B) > 0$, as usual in the theory of KOLMOGOROV, by $\mathbf{P}^*(A|B) = \frac{\mathbf{P}_C(AB)}{\mathbf{P}_C(B)}$, we have by Axiom III $\mathbf{P}^*(A|B) = \frac{\mathbf{P}(AB|C)}{\mathbf{P}(B|C)} = \mathbf{P}(A|BC)$. In case $S \in \mathcal{B}$, clearly $[S, \mathcal{A}, \mathbf{P}_S(A)]$ is a probability space in the sense of KOLMOGOROV. It must be mentioned that in this case $[S, \mathcal{A}, \mathcal{B}, \mathbf{P}(A|B)]$ may not be identical with the conditional probability space generated by $[S, \mathcal{A}, \mathbf{P}_S(A)]$ because $\mathcal{B}$ may contain sets B for which $\mathbf{P}(B|S) = 0$ and at the same time need not contain every set B for which $\mathbf{P}(B|S) > 0$, i. e. the system $\mathcal{B}_S$ consisting of all sets $B \in \mathcal{A}$ for which $\mathbf{P}(B|S) > 0$ need not be identical with $\mathcal{B}$. However, if $\mathbf{P}^*(A|B)$ is defined by $\mathbf{P}^*(A|B) = \frac{\mathbf{P}(AB|S)}{\mathbf{P}(B|S)}$ for $B \in \mathcal{B}_S$, we have $\mathbf{P}^*(A|B) = \mathbf{P}(A|B)$, provided that $B \in \mathcal{B}$.

1. 4. Immediate consequences of the axioms. We shall prove some simple theorems which follow from our axioms. In what follows if $\mathbf{P}(A|B)$ occurs, it is always tacitly assumed that $A \in \mathcal{A}$ and $B \in \mathcal{B}$. We denote the set $S - A$ by $\bar{A}$.

THEOREM 1. $\mathbf{P}(A|B) = \mathbf{P}(AB|B)$.

PROOF. If in Axiom III $C = B$, we have $\mathbf{P}(A|B)\mathbf{P}(B|B) = \mathbf{P}(AB|B)$. Taking Axiom I into account Theorem 1 follows.

REMARK. It follows from Theorem 1 that $\mathbf{P}(S|B) = 1$; namely, by Theorem 1 $\mathbf{P}(S|B) = \mathbf{P}(SB|B) = \mathbf{P}(B|B)$ and thus, by Axiom I, $\mathbf{P}(S|B) = 1$.

THEOREM 2. $\mathbf{P}(A|B) \leqq 1$.

PROOF. According to Axiom II we have $\mathbf{P}(AB|B) + \mathbf{P}(\bar{A}B|B) = \mathbf{P}(B|B)$. As by Axiom I $\mathbf{P}(B|B) = 1$ and $\mathbf{P}(\bar{A}B|B) \geqq 0$, it follows $\mathbf{P}(AB|B) \leqq 1$ and thus by Theorem 1 we obtain $\mathbf{P}(A|B) \leqq 1$.

THEOREM 3. $\mathbf{P}(O|B) = 0$.

PROOF. According to Axiom II $\mathbf{P}(O|B) = \mathbf{P}(O + O|B) = 2\mathbf{P}(O|B)$ and thus $\mathbf{P}(O|B) = 0$.

REMARK 1. It follows from Theorem 3 that $O \notin \mathcal{B}$, because if O would belong to $\mathcal{B}$, we should have $\mathbf{P}(O|O) = 1$ by Axiom I and $\mathbf{P}(O|O) = 0$ by Theorem 3; thus the assumption $O \in \mathcal{B}$ leads to a contradiction.

REMARK 2. It follows from Theorem 1 and 3 that if $AB = O$, then $\mathbf{P}(A|B) = 0$.

THEOREM 4. $\mathbf{P}(A|BC)\,\mathbf{P}(B|C) = \mathbf{P}(B|AC)\,\mathbf{P}(A|C)$.

PROOF. By Axiom III both expressions are equal to $\mathbf{P}(AB|C)$ and thus to each other.

THEOREM 5. *If* $A \subseteq A' \subseteq B \subseteq B'$, *we have*

$$\mathbf{P}(A|B') \leqq \mathbf{P}(A'|B).$$

PROOF. We have

$$\mathbf{P}(A|B') = \mathbf{P}(AA'B|B') = \mathbf{P}(AA'|BB')\,\mathbf{P}(B|B') \leqq$$
$$\leqq \mathbf{P}(AA'|B) = \mathbf{P}(A'|B) - \mathbf{P}(\bar{A}A'|B) \leqq \mathbf{P}(A'|B).$$

REMARK 1. If $A = A'$, we obtain the following special case of Theorem 5: if $A \subseteq B \subseteq B'$, we have $\mathbf{P}(A|B') \leqq \mathbf{P}(A|B)$.

REMARK 2. If $B = B'$, we obtain the following special case of Theorem 5: if $A \subseteq A'$, we have (without supposing that $A' \subseteq B$) $\mathbf{P}(A|B) \leqq \mathbf{P}(A'|B)$.[4] As a matter of fact, $\mathbf{P}(A|B) = \mathbf{P}(AB|B)$ and $\mathbf{P}(A'|B) = \mathbf{P}(A'B|B)$ by Theorem 1; if $A \subseteq A'$, we have $AB \subseteq A'B \subseteq B$, and Theorem 5 can be applied.

THEOREM 6. *If* $A_1 + A_2 \subseteq B_1 B_2 \in \mathfrak{B}$, *further* $\mathbf{P}(A_2|B_1)\,\mathbf{P}(A_2|B_2) > 0$, *we have*

$$\frac{\mathbf{P}(A_1|B_1)}{\mathbf{P}(A_2|B_1)} = \frac{\mathbf{P}(A_1|B_2)}{\mathbf{P}(A_2|B_2)}.$$

PROOF. According to Axiom III

$$\mathbf{P}(A_1|B_1B_2)\,\mathbf{P}(B_1|B_2) = \mathbf{P}(A_1B_1|B_2) = \mathbf{P}(A_1|B_2) \tag{1}$$

and similarly

$$\mathbf{P}(A_2|B_1B_2)\,\mathbf{P}(B_1|B_2) = \mathbf{P}(A_2B_1|B_2) = \mathbf{P}(A_2|B_2). \tag{2}$$

Dividing (1) by (2) we obtain

$$\frac{\mathbf{P}(A_1|B_1B_2)}{\mathbf{P}(A_2|B_1B_2)} = \frac{\mathbf{P}(A_1|B_2)}{\mathbf{P}(A_2|B_2)}. \tag{3}$$

Interchanging B_1 and B_2 in (3) we obtain

$$\frac{\mathbf{P}(A_1|B_1B_2)}{\mathbf{P}(A_2|B_1B_2)} = \frac{\mathbf{P}(A_1|B_1)}{\mathbf{P}(A_2|B_1)}. \tag{4}$$

Comparing (3) with (4), Theorem 6 follows.

THEOREM 7. *If* $C \subseteq B = \sum_{k=1}^{\infty} B_k$ *and* $AB_jB_kC = O$ *for* $j \neq k$ $(j, k = 1, 2, \ldots)$, *we have*

$$\mathbf{P}(A|C) = \sum_{k=1}^{\infty} \mathbf{P}(A|B_kC)\,\mathbf{P}(B_k|C).$$

[4] It should be mentioned that this special case of Theorem 5 follows from Axioms I and II; Axiom III is not needed.

PROOF. By Axiom III $\mathbf{P}(A|B_kC)\,\mathbf{P}(B_k|C)=\mathbf{P}(AB_k|C)$ and thus by Axiom II and applying Theorem 1 twice, it follows

$$\sum_{k=1}^{\infty}\mathbf{P}(A|B_kC)\,\mathbf{P}(B_k|C)=\sum_{k=1}^{\infty}\mathbf{P}(AB_k|C)=\mathbf{P}(AB|C)=\mathbf{P}(ABC|C)=$$
$$=\mathbf{P}(AC|C)=\mathbf{P}(A|C).$$

REMARK. We mention the following consequences of Theorem 7: Let us suppose $B_k\in\mathfrak{B}$ and $B_jB_k=O$ for $j\neq k$, $B=\sum_{k=1}^{\infty}B_k$ and $B\in\mathfrak{B}$; if $\mathbf{P}(A|B_k)\leqq\lambda\mathbf{P}(A'|B_k)$ for $k=1,2,\ldots$ where $\lambda\geqq 0$, we have $\mathbf{P}(A|B)\leqq\lambda\mathbf{P}(A'|B)$.

Especially if $\mathbf{P}(A|B_k)=\lambda\mathbf{P}(A'|B_k)$ for $k=1,2,\ldots$, we have $\mathbf{P}(A|B)=\lambda\mathbf{P}(A'|B)$, further if $\mathbf{P}(A|B_k)=\lambda$ for $k=1,2,\ldots$, we have $\mathbf{P}(A|B)=\lambda$.

PROOF. Applying Theorem 7 twice with $C=B$ we obtain

$$\mathbf{P}(A|B)=\sum_{k=1}^{\infty}\mathbf{P}(A|B_k)\,\mathbf{P}(B_k|B)\leqq\lambda\sum_{k=1}^{\infty}\mathbf{P}(A'|B_k)\,\mathbf{P}(B_k|B)=\lambda\mathbf{P}(A'|B).$$

The two other assertions are evident consequences.

1.5. Representation of the conditional probability as a quotient. We shall give a sufficient condition under which the set function $\mathbf{P}(A|B)$ of two set variables can be represented in "quotient form", i. e. in the form $\mathbf{P}(A|B)=\dfrac{\mathbf{Q}(AB)}{\mathbf{Q}(B)}$ where the set function $\mathbf{Q}(A)$ is a measure on $\mathfrak{A}$ and satisfies $\mathbf{Q}(B)>0$ if $B\in\mathfrak{B}$.

THEOREM 8. *Let* $[S,\mathfrak{A},\mathfrak{B},\mathbf{P}(A|B)]$ *be a conditional probability space. Let us suppose that there exists a sequence of sets* $B_n(n=0,1,\ldots)$ *for which the following properties hold:*

a) $B_n\subseteq B_{n+1}\qquad(n=0,1,\ldots)$,

b) $\mathbf{P}(B_0|B_n)>0\qquad(n=1,2,\ldots)$,

c) *For any* $B\in\mathfrak{B}$ *there can be found a* B_n *for which* $B\subseteq B_n$ *and* $\mathbf{P}(B|B_n)>0$.

Then there exists a finite measure $\mathbf{Q}(C)$ *defined for* $C\in\mathfrak{A}^*$ *where* $\mathfrak{A}^*$ *is the ring of those sets* $C\in\mathfrak{A}$ *for which there can be found a* B_n *with* $C\subseteq B_n$, *and this measure* $\mathbf{Q}(C)$ *has the following properties:*

α) $\mathbf{Q}(B)>0$ *if* $B\in\mathfrak{B}$,

β) $\mathbf{P}(A|B)=\dfrac{\mathbf{Q}(AB)}{\mathbf{Q}(B)}$.

If the sequence B_n *satisfies besides* a), b), c) *also the following condition:*

d) $\lim_{n\to\infty}P(B_0|B_n)>0$,

then $\mathbf{Q}(C)$ *can be defined for all* $C \in \mathcal{A}$ *and is a bounded measure on* $\mathcal{A}$, *and thus, putting* $\mathbf{P}(C) = \frac{\mathbf{Q}(C)}{\mathbf{Q}(S)}$, *we have* $\mathbf{P}(S) = 1$. *Denoting by* $\mathfrak{B}^*$ *the set of those sets* $B \in \mathcal{A}$ *for which* $\mathbf{P}(B) > 0$, *if* $\mathfrak{B}^*$ *is not identical with* $\mathfrak{B}$, *we may extend the definition* $\mathbf{P}(A|B)$ *to all* $B \in \mathfrak{B}^*$ *putting*

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(AB)}{\mathbf{P}(B)};$$

the conditional probability space $[S, \mathcal{A}, \mathfrak{B}^*, \mathbf{P}(A|B)]$ *obtained in this way will be identical with the conditional probability space generated by the ordinary probability space* $[S, \mathcal{A}, \mathbf{P}(A)]$.

PROOF. First we suppose only that the sequence B_n has the properties a), b) and c). Clearly $\mathcal{A}^*$ is a ring and we have $\mathfrak{B} \subseteq \mathcal{A}^* \subseteq \mathcal{A}$. Let us consider a set $A \in \mathcal{A}^*$, choose an index n for which $A \subseteq B_n$ and define $\mathbf{Q}(A)$ as follows:

$$\mathbf{Q}(A) = \frac{\mathbf{P}(A|B_n)}{\mathbf{P}(B_0|B_n)}. \tag{5}$$

It is clear that the value of $\mathbf{Q}(A)$ does not depend on the choice of n. As a matter of fact, if $A \subseteq B_n$ and $A \subseteq B_m$ where $n < m$, we have by Theorem 6

$$\frac{\mathbf{P}(A|B_n)}{\mathbf{P}(B_0|B_n)} = \frac{\mathbf{P}(A|B_m)}{\mathbf{P}(B_0|B_m)}.$$

Now, if $B \in \mathfrak{B}$, $\mathbf{Q}(B) > 0$, we have

$$\mathbf{P}(A|B) = \frac{\mathbf{Q}(AB)}{\mathbf{Q}(B)}. \tag{6}$$

This can be shown as follows: if $B \subseteq B_n$ and $\mathbf{P}(B|B_n) > 0$, we have

$$\frac{\mathbf{Q}(AB)}{\mathbf{Q}(B)} = \frac{\mathbf{P}(AB|B_n)}{\mathbf{P}(B_0|B_n)} \cdot \frac{\mathbf{P}(B_0|B_n)}{\mathbf{P}(B|B_n)} = \frac{\mathbf{P}(AB|B_n)}{\mathbf{P}(B|B_n)}.$$

Applying Axiom III we obtain

$$\frac{\mathbf{Q}(AB)}{\mathbf{Q}(B)} = \mathbf{P}(A|BB_n).$$

As $BB_n = B$, (6) is proved.

From (5) it can be seen that $\mathbf{Q}(A)$ is non-negative. To show that $\mathbf{Q}(A)$ is a measure, we have to prove that $\mathbf{Q}(A)$ is countably additive on $\mathcal{A}^*$, i. e. if $A_k \in \mathcal{A}^*$ $(k = 1, 2, \ldots)$ and $A_j A_k = O$ for $j \neq k$ and $\sum_{k=1}^{\infty} A_k = A \in \mathcal{A}^*$, then $\mathbf{Q}(A) = \sum_{k=1}^{\infty} \mathbf{Q}(A_k)$. This follows simply from the remark that if $A \subseteq B_n$, we have $A_k \subseteq B_n$ for $1, 2, \ldots$ and thus in the relations

$$\mathbf{Q}(A_k) = \frac{\mathbf{P}(A_k|B_n)}{\mathbf{P}(B_0|B_n)} \quad (k = 1, 2, \ldots); \qquad \mathbf{Q}(A) = \frac{\mathbf{P}(A|B_n)}{\mathbf{P}(B_0|B_n)}$$

the same B_n may be used, and therefore the countable additivity of $\mathbf{Q}(A)$ follows from that of $\mathbf{P}(A|B)$ for fixed $B \in \mathfrak{B}$ (Axiom II).

Thus the first part of Theorem 8 is proved. Now we suppose that the sequence B_n has besides a)—c) also the property d). By a well-known theorem (see p. e. [15], § 13, Theorem A) the definition of $\mathbf{Q}(A)$ can be extended to the smallest σ-ring $\mathcal{A}^{**}$ containing $\mathcal{A}^*$ in such a manner that $\mathbf{Q}(A)$ remains countably additive on $\mathcal{A}^{**}$. Let us put $S^* = \sum_{n=0}^{\infty} B_n$; we shall show that $\mathcal{A}^{**}$ is identical with the σ-algebra $\mathcal{A}S^*$, i. e. with the set of all sets of the form AS^* where $A \in \mathcal{A}$. As a matter of fact, $\mathcal{A}^* \subseteq \mathcal{A}S^*$ and $\mathcal{A}S^*$ is a σ-ring, thus we have clearly $\mathcal{A}^{**} \subseteq \mathcal{A}S^*$. On the other hand, if $A \subseteq \mathcal{A}$, we have

$$AS^* = \sum_{n=0}^{\infty} AB_n .$$

Now $AB_n \subseteq B_n$, and thus $AB_n \in \mathcal{A}^*$, and therefore $AS^* \in \mathcal{A}^{**}$; thus $\mathcal{A}S^* \subseteq \mathcal{A}^{**}$ which implies $\mathcal{A}S^* = \mathcal{A}^{**}$. Thus the definition of $\mathbf{Q}(A)$ can be extended to all $A \in \mathcal{A}S^*$. We prove now that $\mathbf{Q}(A)$ is bounded on $\mathcal{A}S^*$. To show this it is sufficient to prove that $\mathbf{Q}(S)$ is finite. But $S^* = \lim_{n\to\infty} B_n$ and $B_n \subseteq B_{n+1}$, and thus $\mathbf{Q}(S^*) = \lim_{n\to\infty} \mathbf{Q}(B_n)$ where $\mathbf{Q}(B_n)$ is non-decreasing. On the other hand,

$$\mathbf{Q}(B_n) = \frac{\mathbf{P}(B_n|B_n)}{\mathbf{P}(B_0|B_n)} = \frac{1}{\mathbf{P}(B_0|B_n)} .$$

Thus e) implies that $\mathbf{Q}(S^*) < +\infty$. Defining $\mathbf{Q}(A)$ by $\mathbf{Q}(A) = \mathbf{Q}(AS^*)$ for $A \in \mathcal{A}$, $A \notin \mathcal{A}S^*$, the definition of $\mathbf{Q}(A)$ is extended to the whole σ-algebra A. The final part of Theorem 8 concerning $\mathbf{P}(A)$ is obvious. Thus Theorem 8 is proved.

A necessary and sufficient condition for the existence of the quotient representation $\mathbf{P}(A|B) = \frac{\mathbf{Q}(AB)}{\mathbf{Q}(B)}$ is contained in the paper [13] of Á. CSÁSZÁR.

1.6. Random variables on a conditional probability space. Let $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ be a conditional probability space. If $\xi = \xi(a)$ denotes a real-valued function defined for $a \in S$ which is measurable with respect to $\mathcal{A}$, i. e. if A_x denotes the set of those $a \in S$ for which $\xi(a) < x$, we have $A_x \in \mathcal{A}$ for all real x, we shall call ξ a random variable on $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}(A|B)]$. Vector-valued random variables are defined similarly. The (ordinary) conditional probability distribution function of a random variable ξ with respect to an event $B \in \mathfrak{B}$ is defined by $F(x|B) = \mathbf{P}(A_x|B)$; if $F(x|B)$ is absolutely continuous, $F'(x|B) = f(x|B)$ is called the (ordinary) conditional probability density function of ξ with respect to B. The conditional mean value $\mathbf{M}(\xi|B)$ of ξ with respect to an event $B \in \mathfrak{B}$ is defined as the abstract Lebesgue integral

$$\mathbf{M}(\xi|B) = \int_S \xi(a) d\mathbf{P}(A|B)$$

of ξ with respect to the measure defined on S by $\mathbf{P}(A|B)$ with B fixed. Higher conditional moments, the conditional characteristic function etc. are defined similarly. The random variables ξ and η are called independent with respect to an event C, if denoting by A_x the set of those $a \in S$ for which $\xi(a) < x$ and by B_y the set of those $a \in S$ for which $\eta(a) < y$, we have $\mathbf{P}(A_x B_y | C) = \mathbf{P}(A_x | C)\mathbf{P}(B_y | C)$ for every real x and y.

As $[S, \mathfrak{A}, \mathbf{P}(A|B)]$ is for any fixed $B \in \mathfrak{B}$ a probability field in the sense of the theory of KOLMOGOROV, any theorem of ordinary probability theory remains valid when ordinary probabilities, distributions, mean values, independence etc. are replaced by conditional probabilities, conditional distributions, conditional mean values, conditional independence etc. with respect to the same $B \in \mathfrak{B}$. Problems which are peculiar to the theory presented in this paper are those in which conditional probabilities with respect to different conditions are figuring. We shall see in § 3 some examples of such problems.

Let us mention that if ξ is a random variable, and A_α^β denotes the set consisting of those elements $a \in S$ for which $\alpha \leqq \xi(a) < \beta$, and if $A_\alpha^\beta \in \mathfrak{B}$ for a set X of intervals (α, β), then the conditional probabilities $\mathbf{P}(A_x^y | A_\alpha^\beta)$ can be considered for $(\alpha, \beta) \in X$ and thus ξ generates a conditional probability space on the real axis $\mathfrak{R}$, as the space of elementary events, the σ-algebra $\mathfrak{A}$ being the set of Borel subsets of $\mathfrak{R}$ and $\mathfrak{B}$ consisting of the intervals $(\alpha, \beta) \in X$.

This conditional probability space will be called the conditional probability distribution generated by ξ on the real axis.

Let $F(x)$ denote a non-decreasing function of x which is continuous to the left for $-\infty < x < +\infty$. If the set A_α^β belongs to $\mathfrak{B}$ whenever $F(\beta) - F(\alpha) > 0$, and we have for any subinterval (x, y) of such an interval (α, β)

$$\mathbf{P}(A_x^y | A_\alpha^\beta) = \frac{F(y) - F(x)}{F(\beta) - F(\alpha)}, \tag{7a}$$

we shall call $F(x)$ the generalized distribution function[5] of ξ; the function $F(x)$ is not uniquely determined, as together with $F(x)$ $G(x) = cF(x) + d$ where $c > 0$ is also a distribution function of ξ; but as $F(x)$ will be used only to calculate the conditional probabilities (7a), this will never lead to a misunderstanding. If the distribution function $F(x)$ of ξ is absolutely continuous, and $F'(x) = f(x)$, we shall call $f(x)$ the generalized density function of ξ; clearly $f(x)$ is determined only up to a positive constant factor. If $F(x) = x$ (i. e. $f(x) = 1$) for $-\infty < x < +\infty$, we shall say that the distribution of ξ is uniform in $(-\infty, +\infty)$.

[5] Conditions ensuring the existence of a generalized distribution function of a random variable ξ can be formulated by using the results of the paper [13].

If $f(x)$ is the generalized density function of ξ, we have

$$\mathbf{P}(A_x^y \mid A_\alpha^\beta) = \frac{\int_x^y f(u)\,du}{\int_\alpha^\beta f(u)\,du}. \tag{7b}$$

The generalized distribution function resp. density function of a random vector is defined similarly.

1. 7. An alternative form of Axiom III. We have already pointed out that our system of axioms can be characterised in the following manner: the set S, the σ-algebra $\mathcal{A}$ of subsets of S, the subset $\mathcal{B}$ of $\mathcal{A}$ and the set function of two set variables $\mathbf{P}(A|B)$ defined for $A \in \mathcal{A}$ and $B \in \mathcal{B}$ form a conditional probability space if $\mathcal{S}_B = [S, \mathcal{A}, \mathbf{P}(A|B)]$ is an ordinary probability space for every fixed $B \in \mathcal{B}$ and if the probability spaces $\mathcal{S}_C$ and $\mathcal{S}_{BC}$ are connected by Axiom III if $C \in \mathcal{B}$ and $BC \in \mathcal{B}$. Thus different probability fields can be combined to form a conditional probability field only if they are "compatible" in that sense that they satisfy Axiom III which can be considered as the condition of compatibility.

Now it is easy to prove

THEOREM 9. *Axiom III can be brought to the following equivalent form:*

AXIOM III′. *If* $B \in \mathcal{B}$, $C \in \mathcal{B}$, $B \subseteq C$ *and* $\mathbf{P}(B|C) > 0$, *we have for any* $A \in \mathcal{A}$

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(AB|C)}{\mathbf{P}(B|C)}.$$

PROOF. Clearly Axiom III′ is a special case of Axiom III. Conversely, if Axiom III′ is valid, Axiom III follows. This can be shown as follows: if $A \in \mathcal{A}$, $B \in \mathcal{A}$, $C \in \mathcal{B}$ and $BC \in \mathcal{B}$, two cases are possible: either $\mathbf{P}(B|C) = 0$ or $\mathbf{P}(B|C) > 0$. In the first case we have also[6] $\mathbf{P}(AB|C) = 0$ and thus $\mathbf{P}(A|BC)\mathbf{P}(B|C) = \mathbf{P}(AB|C)$ reduces to $0 = 0$. Now let us suppose $\mathbf{P}(B|C) > 0$. It is easy to see that Theorem 1 follows already from Axioms I—III′, and thus it can be applied in the present proof. But this means that $\mathbf{P}(BC|C) = \mathbf{P}(B|C) > 0$ and thus the conditions of Axiom III′ are satisfied with BC instead of B, and it follows from Axiom III′

$$\mathbf{P}(A|BC)\mathbf{P}(BC|C) = \mathbf{P}(ABC|C). \tag{8}$$

As $\mathbf{P}(BC|C) = \mathbf{P}(B|C)$ and $\mathbf{P}(ABC|C) = \mathbf{P}(AB|C)$, it follows from (8)

$$\mathbf{P}(A|BC)\mathbf{P}(B|C) = \mathbf{P}(AB|C).$$

Thus Axiom III follows from Axiom III′.

[6] See the footnote [4] to Remark 2 to Theorem 5.

REMARK. Theorem 9 means that Axiom III contains a compatibility condition for $\mathfrak{S}_B$ and $\mathfrak{S}_C$ where $B \subseteq C$ if and only if $\mathbf{P}(B|C) > 0$; if $\mathbf{P}(B|C) = 0$, $\mathfrak{S}_B$ and $\mathfrak{S}_C$ are compatible without any restriction. This fact is the basis of a general principle by use of which conditional probability spaces can be constructed. This principle is expressed in Theorem 14 of § 2.

1. 8. Extensions of a conditional probability space. If $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ is a conditional probability space, it is natural to ask how could this space be extended, by including into $\mathfrak{B}$ sets $A \in \mathfrak{A}$ which are not contained in $\mathfrak{B}$. The most simple way is suggested by Axiom III, and is contained in the following

THEOREM 10. *Let B_1 denote a set for which $B_1 \in \mathfrak{A}$ and $B_1 \notin \mathfrak{B}$. If there exists at least one set B_2 with the following three properties: α) $B_2 \in \mathfrak{B}$, β) $B_1 \subseteq B_2$, γ) $\mathbf{P}(B_1|B_2) > 0$, further if for any other set B_3 which also has the properties α), β), γ), we have $B_2 B_3 \in \mathfrak{B}$, the definition of $\mathbf{P}(A|B)$ can be extended for $B = B_1$ by putting*

$$\mathbf{P}(A|B_1) = \frac{\mathbf{P}(AB_1|B_2)}{\mathbf{P}(B_1|B_2)}. \tag{9}$$

PROOF. If $\mathbf{P}(A|B_1)$ is defined by (9), Axioms I and II are clearly satisfied, therefore we have to verify only Axiom III. Three cases must be distinguished. a) If we put $B_1 = C'$ and B' is a set for which $B'C' \in \mathfrak{B}$, we must verify

$$\mathbf{P}(A|B'C')\mathbf{P}(B'|C') = \mathbf{P}(AB'|C').$$

But this means by (9)

$$\mathbf{P}(A|B_1B')\mathbf{P}(B_1B'|B_2) = \mathbf{P}(AB_1B'|B_2)$$

what is true by force of Axiom III, in view of $B_1B' = B'C' \in \mathfrak{B}$ and $B_2 \in \mathfrak{B}$. b) If $B_1 = B'C'$ where $C' \in \mathfrak{B}$, we must verify $\mathbf{P}(A|B_1)\mathbf{P}(B'|C') = \mathbf{P}(AB'|C')$. But this means by (9)

$$\mathbf{P}(B'|C')\mathbf{P}(AB_1|B_2) = \mathbf{P}(AB'|C')\mathbf{P}(B_1|B_2)$$

what reduces to $0 = 0$ if $\mathbf{P}(B'|C') = 0$ and to

$$\frac{\mathbf{P}(AB_1|B_2)}{\mathbf{P}(B_1|B_2)} = \frac{\mathbf{P}(AB'|C')}{\mathbf{P}(B'|C')} \tag{10}$$

if $\mathbf{P}(B'|C') > 0$. But as $B_1 = B'C'$, by applying Theorem 1 to both conditional probabilities on the right of (10) we find that (10) is equivalent to

$$\frac{\mathbf{P}(AB_1|B_2)}{\mathbf{P}(B_1|B_2)} = \frac{\mathbf{P}(AB_1|C')}{\mathbf{P}(B_1|C')}. \tag{11}$$

But (11) follows from Theorem 6, taking into account that

$$AB_1 + B_1 = B_1 \subseteq B_2 \quad \text{and} \quad AB_1 + B_1 = B_1 = B'C' \subseteq C',$$

and that clearly $\mathbf{P}(B_1|C') = \mathbf{P}(B'|C') > 0$ and thus C' has the properties α), β) and γ) and therefore $B_2C' \in \mathfrak{B}$.

If $B'C' = B_1$ and $C' = B_1$ where $B' \in \mathfrak{A}$, we have to verify $\mathbf{P}(A|B'C')\mathbf{P}(B'|C') = \mathbf{P}(AB'|C')$. As $\mathbf{P}(AB'|C') = \mathbf{P}(AB'C'|C')$, this relation is trivially equivalent to $\mathbf{P}(A|B_1) = \mathbf{P}(AB_1|B_1)$, which is satisfied by force of (9).

It is easy to see that the definition of $\mathbf{P}(A|B_1)$ does not depend on the choice of B_2; as a matter of fact, if both B_2 and B_3 have properties α), β) and γ), it follows by Theorem 6 that

$$\frac{\mathbf{P}(AB_1|B_2)}{\mathbf{P}(B_1|B_2)} = \frac{\mathbf{P}(AB_1|B_3)}{\mathbf{P}(B_1|B_3)}.$$

It is also clear that $\mathbf{P}(AB_1|B_3)$ can not be defined otherwise as by (9) because if B_1 is included into $\mathfrak{B}$, (9) must hold by force of Axiom III.

Another possibility for including new sets into $\mathfrak{B}$ is yielded by passing to the limit; this procedure is described by the following

THEOREM 11. *Let us suppose that* $B_n \in \mathfrak{B}$, $B_n \subseteq B_{n+1}$, *further* $\mathbf{P}(B_n|B_{n+1}) > 0$ $(n = 0, 1, \ldots)$ *and that* $\prod_{n=0}^{\infty} \mathbf{P}(B_n|B_{n+1})$ *converges. If* $B_\infty = \sum_{n=0}^{\infty} B_n$ *does not belong to* $\mathfrak{B}$, *the definition of* $\mathbf{P}(A|B)$ *can be extended for* $B = B_\infty$, *by putting*

$$\mathbf{P}(A|B_\infty) = \lim_{n\to\infty} \mathbf{P}(A|B_n) \quad \textit{for any} \quad A \in \mathfrak{A}, \tag{12}$$

provided that the following condition is satisfied: if $B \in \mathfrak{B}$ *and* $\mathbf{P}(B|B_N) > 0$ *for some* N, *we have* $BB_N \in \mathfrak{B}$.

PROOF. For an arbitrary $A \in \mathfrak{A}$ we put $A^{(0)} = AB_0$, $A^{(k)} = AB_k\bar{B}_{k-1}$ $(k = 1, 2, \ldots)$. As $A^{(k)} \subseteq B_n$ for $n \geqq k$, the sequence $\mathbf{P}(A^{(k)}|B_n)$ is by Theorem 5, Remark 1 monotonically decreasing for $n = k, k+1, \ldots$ and thus

$$\lim_{n\to\infty} \mathbf{P}(A^{(k)}|B_n) = \mathbf{P}(A^{(k)}|B_\infty)$$

exists.

Putting

$$\mathbf{P}^*(A|B_\infty) = \sum_{k=0}^{\infty} \mathbf{P}(A^{(k)}|B_\infty) \tag{13}$$

we have defined $\mathbf{P}^*(A|B_\infty)$ for every $A \in \mathfrak{A}$. To prove Theorem 11, we have to show that if $\mathbf{P}^*(A|B_\infty)$ is defined by (13), Axioms I—III remain valid, further that $\lim_{n\to\infty} \mathbf{P}(A|B_n) = \mathbf{P}(A|B_\infty)$ exists for all $A \in \mathfrak{A}$ and $\mathbf{P}(A|B_\infty) = \mathbf{P}^*(A|B_\infty)$, i. e. that (12) and (13) are equivalent. As regards Axiom I, it is clear that $\mathbf{P}^*(A|B_\infty) \geqq 0$; the validity of $\mathbf{P}^*(B_\infty|B_\infty) = 1$ can be shown as

follows: if $A = B_\infty$, $A^{(0)} = B_0$ and $A^{(k)} = B_k \bar{B}_{k-1}$ $(k = 1, 2, \ldots)$ which implies

$$\mathbf{P}^*(B_\infty | B_\infty) = \mathbf{P}(B_0 | B_\infty) + \sum_{k=1}^{\infty} \mathbf{P}(B_k \bar{B}_{k-1} | B_\infty) = \lim_{N \to \infty} \mathbf{P}(B_N | B_\infty).$$

But

$$\mathbf{P}(B_N | B_\infty) = \lim_{n \to \infty} \mathbf{P}(B_N | B_n) = \lim_{n \to \infty} \prod_{k=N}^{n-1} \mathbf{P}(B_k | B_{k+1}) = \prod_{k=N}^{\infty} \mathbf{P}(B_k | B_{k+1})$$

and thus

$$\mathbf{P}^*(B_\infty | B_\infty) = \lim_{N \to \infty} \prod_{k=N}^{\infty} \mathbf{P}(B_k | B_{k+1}) = 1$$

because the remainder-products of a convergent infinite product are always tending to 1. Now let us verify Axiom II, i. e. that $\mathbf{P}^*(A | B_\infty)$ defined by (13) is countably additive. Let us suppose that $A_k \in \mathfrak{A}$ and $A_j A_k = O$, if $j \neq k$ $(j, k = 1, 2, \ldots)$, and $A_\infty = \sum_{n=1}^{\infty} A_n$. Let us put $A_n^{(0)} = A_n B_0$ and $A_n^{(k)} = A_n B_k \bar{B}_{k-1}$ for $k = 1, 2, \ldots$, further $A_\infty^{(0)} = A_\infty B_0$ and $A_\infty^{(k)} = A_\infty B_k \bar{B}_{k-1}$ for $k = 1, 2, \ldots$. As for $A \subseteq B_k$ we have

$$\mathbf{P}(A | B_\infty) = \mathbf{P}(A | B_k) \prod_{n=k}^{\infty} \mathbf{P}(B_n | B_{n+1}) \tag{14}$$

and $\mathbf{P}(A | B_k)$ is countably additive, further $A_\infty^{(k)} = \sum_{n=1}^{\infty} A_n^{(k)}$ and $A_n^{(k)} \cdot A_m^{(k)} = O$ for $n \neq m$, it follows

$$\mathbf{P}(A_\infty^{(k)} | B_\infty) = \sum_{n=1}^{\infty} \mathbf{P}(A_n^{(k)} | B_\infty). \tag{15}$$

But

$$\mathbf{P}^*(A_\infty | B_\infty) = \sum_{k=0}^{\infty} \mathbf{P}(A_\infty^{(k)} | B_\infty). \tag{16}$$

From (15) and (16) it follows

$$\mathbf{P}^*(A_\infty | B_\infty) = \sum_{k=0}^{\infty} \sum_{n=1}^{\infty} \mathbf{P}(A_n^{(k)} | B_\infty) = \sum_{n=1}^{\infty} \sum_{k=0}^{\infty} \mathbf{P}(A_n^{(k)} | B_\infty). \tag{17}$$

As we have from (13)

$$\sum_{k=0}^{\infty} \mathbf{P}(A_n^{(k)} | B_\infty) = \mathbf{P}^*(A_n | B_\infty), \tag{18}$$

it follows from (16) and (17) that

$$\mathbf{P}^*(A_\infty | B_\infty) = \sum_{n=1}^{\infty} \mathbf{P}^*(A_n | B_\infty), \tag{19}$$

i. e. that $\mathbf{P}^*(A | B_\infty)$ is countably additive.

Now we prove that

$$\mathbf{P}^*(A | B_\infty) = \lim_{N \to \infty} \mathbf{P}(A | B_N) \tag{20}$$

for any $A \in \mathfrak{A}$. As a matter of fact, by (14)

$$\mathbf{P}(A|B_N) = \sum_{k=0}^{\infty} \mathbf{P}(A^{(k)}|B_N) = \sum_{k=0}^{N} \mathbf{P}(A^{(k)}|B_N) = \frac{\sum_{k=0}^{N} \mathbf{P}(A^{(k)}|B_\infty)}{\prod_{n=N}^{\infty} \mathbf{P}(B_n|B_{n+1})}. \tag{21}$$

It follows from (21) that

$$\lim_{N\to\infty} \mathbf{P}(A|B_N) = \sum_{k=0}^{\infty} \mathbf{P}(A^{(k)}|B_\infty) \tag{22}$$

and thus, by (13), (20) follows.

Now let us prove that Axiom III′ is valid. We distinguish three cases. The first case is when $C \in \mathfrak{B}$, $B_\infty \subseteq C$ and $\mathbf{P}(B_\infty|C) > 0$; then we have $\mathbf{P}(B_N|C) > 0$ if N is sufficiently large and thus

$$\mathbf{P}(A|B_N) = \frac{\mathbf{P}(AB_N|C)}{\mathbf{P}(B_N|C)}. \tag{23}$$

Passing to the limit $N \to \infty$ in (23), and using (20), it follows that

$$\mathbf{P}(A|B_\infty) = \frac{\mathbf{P}(AB_\infty|C)}{\mathbf{P}(B_\infty|C)}. \tag{24}$$

The second case is when $B \in \mathfrak{B}$, $B \subseteq B_\infty$ and $\mathbf{P}(B|B_\infty) > 0$; we have to prove that

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(AB|B_\infty)}{\mathbf{P}(B|B_\infty)}. \tag{25}$$

If there exists an index N for which $B \subseteq B_N$, we have for sufficiently large N

$$\frac{\mathbf{P}(AB|B_N)}{\mathbf{P}(B|B_N)} = \mathbf{P}(A|B)$$

and thus, passing to the limit $N \to \infty$, (25) follows.

In the general case, let us consider the two measures $\mu_1(A) = \mathbf{P}(A|B)$ and $\mu_2(A) = \mathbf{P}(A|B_\infty)$. Let $A \in \mathfrak{A}$ denote an arbitrary set, for which $A \subseteq B_N B$ for some N. Then we have by Theorem 6, taking N sufficiently large, to ensure $\mathbf{P}(B|B_N) > 0$,

$$\frac{\mathbf{P}(A|B_N)}{\mathbf{P}(B|B_N)} = \frac{\mathbf{P}(A|B)}{\mathbf{P}(B|B)} \quad \text{and thus} \quad \mu_2(A) = \mathbf{P}(B|B_\infty)\mu_1(A),$$

taking into account that $\lim_{N\to\infty} \mathbf{P}(B|B_N) = \mathbf{P}(B|B_\infty) = 1$ and thus $\mathbf{P}(B|B_N) > 0$ for sufficiently large N. Thus $\mu_2(A) = C\mu_1(A)$ where $C = \mathbf{P}(B|B_\infty)$ does not depend on A if A runs over all sets for which there exists an index N for which $A \subseteq B_N B$.

Clearly these sets A form a ring. As $\mu_2(A)$ and $C\mu_1(A)$ are finite measures, it follows that they coincide for the least σ-ring which contains all

mentioned sets A. As $B_\infty = \sum_{N=0}^{\infty} B_N$, it follows that $\mu_2(A) = C\mu_1(A)$ if $A \subseteq B_\infty B$ and $A \in \mathfrak{A}$. As $B \subseteq B_\infty$, this means that $\mu_2(A) = C\mu_1(A)$ if $A \subseteq B$, i. e. we have proved that

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A|B_\infty)}{\mathbf{P}(B|B_\infty)} \qquad \text{if } A \subseteq B.$$

As $AB \subseteq B$ for any $A \in \mathfrak{A}$, this means that for any $A \in \mathfrak{A}$ we have

$$\mathbf{P}(A|B) = \mathbf{P}(AB|B) = \frac{\mathbf{P}(AB|B_\infty)}{\mathbf{P}(B|B_\infty)}.$$

Thus (25) is proved. The third case is, when $B = B_\infty$ and $C = B_\infty$; in this case we have to prove that $\mathbf{P}(A|B_\infty) = \frac{\mathbf{P}(AB_\infty|B_\infty)}{\mathbf{P}(B_\infty|B_\infty)}$; but this is evident because we have shown that $\mathbf{P}(B_\infty|B_\infty) = 1$ and by (15) $\mathbf{P}(AB_\infty|B_\infty) = \mathbf{P}(A|B_\infty)$. Thereby the proof of Theorem 11 is accomplished.

REMARK. The assertion of Theorem 6 for the case $A_1 + A_2 \subseteq B_1 B_2$, but without the supposition $B_1 B_2 \in \mathfrak{B}$, can be considered as a stronger form of Axiom III, it shall be called Axiom III*.

AXIOM III*. *If* $A_1 \in \mathfrak{A}$, $A_2 \in \mathfrak{A}$, $B_1 \in \mathfrak{B}$ *and* $B_2 \in \mathfrak{B}$, *further* $A_1 + A_2 \subseteq B_1 B_2$ *and* $\mathbf{P}(A_2|B_1)\mathbf{P}(A_2|B_2) > 0$, *we have*

$$\frac{\mathbf{P}(A_1|B_1)}{\mathbf{P}(A_2|B_1)} = \frac{\mathbf{P}(A_1|B_2)}{\mathbf{P}(A_2|B_2)}.$$

As Theorem 1 is not a consequence of Axiom III*, in case we replace Axiom III by Axiom III*, we must suppose the validity of Theorem 1 as

AXIOM III**. $\mathbf{P}(A|B) = \mathbf{P}(AB|B)$ *for* $A \in \mathfrak{A}$ *and* $B \in \mathfrak{B}$.

Axiom III* and Axiom III** together imply Axiom III′ and thus Axiom III. As a matter of fact, choosing $A_1 = AB$, $A_2 = B$, $B_1 = B$ and $B_2 = C$ where $A \in \mathfrak{A}$, $B \in \mathfrak{B}$, $C \in \mathfrak{B}$, $B \subseteq C$ and $\mathbf{P}(B|C) > 0$, the conditions of Axiom III* are satisfied, and as by Axiom III** $\mathbf{P}(A|B) = \mathbf{P}(AB|B)$ it follows that

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(AB|B)}{\mathbf{P}(B|B)} = \frac{\mathbf{P}(AB|C)}{\mathbf{P}(B|C)},$$

i. e. Axiom III′ is really a consequence of Axiom III* and Axiom III**.

As Axiom III and Axiom III′ are equivalent, Axiom III is also a consequence of Axiom III*. Conversely, Axiom III* does not follow from Axiom III, only in the special case when $B_1 B_2 \in \mathfrak{B}$. If we would suppose Axioms III* and III** instead of Axiom III, the conditions of the extension theorems Theorem 10 and Theorem 11 could be reduced: in Theorem 10 the condition that for two sets B_2, B_3 with properties α), β) and γ), $B_2 B_3 \in \mathfrak{B}$ could be omitted, in Theorem 11 the condition that if $B \in \mathfrak{B}$ and $\mathbf{P}(B|B_N) > 0$, then

$BB_N \in \mathfrak{B}$ could be omitted. In what follows we do not suppose the validity of Axiom III*; it should be mentioned, however, that Axiom III* is contained as a special case in Axiom IV_2 introduced by Á. CSÁSZÁR [13].

If $\mathfrak{S} = [S, \mathcal{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ and $\mathfrak{S}' = [S', \mathcal{A}', \mathfrak{B}', \mathbf{P}'(A|B)]$ are two such conditional probability spaces that $S \subseteq S'$, $\mathcal{A} \subseteq \mathcal{A}'$, $\mathfrak{B} \subseteq \mathfrak{B}'$ and $\mathbf{P}'(A|B) = \mathbf{P}(A|B)$ if $A \in \mathcal{A}$ and $B \in \mathfrak{B}$, we shall say that the conditional probability space $\mathfrak{S}$ is imbedded into the conditional probability space $\mathfrak{S}'$. In the foregoing section we considered only such special imbeddings of a conditional probability space $\mathfrak{S}$ into a conditional probability space $\mathfrak{S}'$ for which $S' = S$, $\mathcal{A}' = \mathcal{A}$ and $\mathfrak{B} \subseteq \mathfrak{B}'$. These imbeddings of $\mathfrak{S}$ were obtained by extension of $\mathfrak{B}$ in several manners. If $\mathfrak{S}$ is imbedded into $\mathfrak{S}'$, we shall write $\mathfrak{S} \ll \mathfrak{S}'$.

1.9. Continuity properties of conditional probability. It follows from Axiom II that $\mathbf{P}(A|B)$ is for fixed $B \in \mathfrak{B}$ a bounded measure in A, and thus, of course, continuous in A, i. e. if $A_n \in A$ and $A_n \subseteq A_{n+1}$ (or $A_n \supseteq A_{n+1}$) for $n = 1, 2, \ldots$, we have for $B \in \mathfrak{B}$

$$\lim_{n\to\infty} \mathbf{P}(A_n|B) = \mathbf{P}(\lim_{n\to\infty} A_n|B).$$

As regards the continuity of $\mathbf{P}(A|B)$ as a function of B, we have

THEOREM 12. *If $B_n \in \mathfrak{B}$ and $B_n \subseteq B_{n+1}$ $(n = 1, 2, \ldots)$, further $\sum_{n=1}^{\infty} B_n = B \in \mathfrak{B}$, we have for $A \in \mathcal{A}$*

$$\lim_{n\to\infty} \mathbf{P}(A|B_n) = \mathbf{P}(A|B).$$

PROOF. We have by Axiom III

$$\mathbf{P}(A|B_n) = \frac{\mathbf{P}(AB_n|B)}{\mathbf{P}(B_n|B)}$$

if $\mathbf{P}(B_n|B) > 0$. As $\lim_{n\to\infty} \mathbf{P}(B_n|B) = \mathbf{P}(B|B) = 1$, the last condition is satisfied for sufficiently large n, and thus

$$\lim_{n\to\infty} \mathbf{P}(A|B_n) = \frac{\lim_{n\to\infty} \mathbf{P}(AB_n|B)}{\lim_{n\to\infty} \mathbf{P}(B_n|B)} = \frac{\mathbf{P}(AB|B)}{\mathbf{P}(B|B)} = \mathbf{P}(A|B).$$

Thus Theorem 12 is proved.

The situation is somewhat more complicated if we consider a decreasing sequence of conditions. In this case we have

THEOREM 13. *If $B \in \mathfrak{B}$, $BC_n \in \mathfrak{B}$ and $C_n \supseteq C_{n+1}$ $(n = 1, 2, \ldots)$ further if putting $C = \prod_{n=1}^{\infty} C_n$ we have $BC \in \mathfrak{B}$ and $\mathbf{P}(C|B) > 0$, it follows that*

$$\lim_{n\to\infty} \mathbf{P}(A|BC_n) = \mathbf{P}(A|BC).$$

PROOF. We have by Axiom III

$$\mathbf{P}(A|BC_n) = \frac{\mathbf{P}(AC_n|B)}{\mathbf{P}(C_n|B)} \quad \text{if} \quad \mathbf{P}(C_n|B) > 0.$$

As $\mathbf{P}(C_n|B) \geqq \mathbf{P}(C|B) > 0$, the condition $\mathbf{P}(C_n|B) > 0$ is satisfied by every n, and thus it follows

$$\lim_{n\to\infty} \mathbf{P}(A|BC_n) = \frac{\lim\limits_{n\to\infty} \mathbf{P}(AC_n|B)}{\lim\limits_{n\to\infty} \mathbf{P}(C_n|B)} = \frac{\mathbf{P}(AC|B)}{\mathbf{P}(C|B)} = \mathbf{P}(A|BC)$$

what proves Theorem 13.

1.10. Products of conditional probability spaces. The product of two conditional probability spaces $(S^{(k)}, \mathcal{A}^{(k)}, \mathfrak{B}^{(k)}, \mathbf{P}^{(k)}) = \mathfrak{S}^{(k)}$ $(k = 1, 2)$ is defined as follows: let $S = S^{(1)} * S^{(2)}$ denote the Cartesian product of the sets $S^{(k)}$ $(k = 1, 2)$, i. e. let S denote the set of all ordered pairs $(a^{(1)}, a^{(2)})$ where $a^{(1)} \in S^{(1)}$ and $a^{(2)} \in S^{(2)}$. Let $\mathfrak{B}$ denote the set of all subsets $B_1 = B_1 * B_2$ of S where $B_1 \in \mathfrak{B}^{(1)}$ and $B_2 \in \mathfrak{B}^{(2)}$. Let $\mathcal{A}$ denote the set of all subsets $A = A_1 * A_2$ where $A_1 \in \mathcal{A}^{(1)}$ and $A_2 \in \mathcal{A}^{(2)}$, and let $\overline{\mathcal{A}}$ denote the least σ-algebra containing $\mathcal{A}$. Let us define $\mathbf{P}(A|B)$ for $A = A_1 * A_2$ and $B = B_1 * B_2$, where $A_1 \in \mathcal{A}^{(1)}$, $A_2 \in \mathcal{A}^{(2)}$, $B_1 \in \mathfrak{B}^{(1)}$ and $B_2 \in \mathfrak{B}^{(2)}$ by

$$\mathbf{P}(A|B) = \mathbf{P}^{(1)}(A^{(1)}|B^{(1)})\,\mathbf{P}^{(2)}(A^{(2)}|B^{(2)}),$$

and let us extend the definition of $\mathbf{P}(A|B)$ for every fixed $B \in \mathfrak{B}$ to all $A \in \overline{\mathcal{A}}$ in the usual way (see [15], § 13). Thus we obtain a conditional probability space $\mathfrak{S} = [S, \overline{\mathcal{A}}, \mathfrak{B}, \mathbf{P}]$ which will be called the Cartesian product of the conditional probability spaces $\mathfrak{S}^{(1)}$ and $\mathfrak{S}^{(2)}$ and denoted by $\mathfrak{S} = \mathfrak{S}^{(1)} * \mathfrak{S}^{(2)}$.

The Cartesian product of a finite number of conditional probability spaces is defined similarly. The Cartesian product of a denumerable sequence of conditional probability spaces $P^{(k)} = [S^{(k)}, \mathcal{A}^{(k)}, \mathfrak{B}^{(k)}, \mathbf{P}^{(k)}]$ $(k = 1, 2, \ldots)$ is defined as follows: we denote by $S = S_1 * \cdots * S_n * \cdots$ the Cartesian product of the sets $S^{(k)}$ $(k = 1, 2, \ldots)$ by $\mathfrak{B}$ the set of all sets $B = B_1 * \cdots * B_n * \cdots$, where $B_n \in \mathfrak{B}^{(n)}$ $(n = 1, 2, \ldots)$ and by $\mathcal{A}$ the set of all sets A of the form $A = A_1 * A_2 * \cdots A_n * S^{(n+1)} * \cdots$, i. e. $\mathcal{A}$ is the set of all "cylinders" of S. We define $\mathbf{P}(A|B)$ for $A \in \mathcal{A}$ and $B \in \mathfrak{B}$, where $A = A_1 \cdots A_n * S^{(n+1)} \cdots$ and $B = B_1 \cdots B_n$, by

$$\mathbf{P}(A|B) = \prod_{k=1}^{n} \mathbf{P}^{(k)}(A_k|B_k),$$

and extend the definition of $\mathbf{P}(A|B)$ in the usual way for any fixed $B \in \mathfrak{B}$ to all sets A belonging to the least σ-algebra $\overline{\mathcal{A}}$ containing $\mathcal{A}$. In this way we obtain a conditional probability space $\mathbf{P} = [S, \overline{\mathcal{A}}, \mathfrak{B}, \mathbf{P}(A|B)]$ which will be called the Cartesian product of the conditional probability spaces $\mathfrak{S}^{(k)}$ and denoted by $\mathfrak{S} = \mathfrak{S}^{(1)} * \cdots * \mathfrak{S}^{(k)} * \cdots$. To prove that $\mathfrak{S}$ is really a conditional probability space, we have only to verify the validity of Axiom III,

as Axioms I and II are clearly valid for $\mathfrak{S}$. As regards the validity of $\mathbf{P}(AB|B)=\mathbf{P}(A|B)$ (i. e. of Theorem 1) for $\mathfrak{S}$, it suffices to prove that

$$\mathbf{P}(A|BC)\,\mathbf{P}(BC|C)=\mathbf{P}(ABC|C) \tag{26}$$

for $A\in\mathfrak{A}$, $B\in\overline{\mathfrak{A}}$, $C\in\mathfrak{B}$ and $BC\in\mathfrak{B}$.

As both sides of (26) are measures as functions of A for fixed B and C, it suffices to verify (26) for the case when A is a cylinder, $A=A^{(1)}*\cdots\cdots*A^{(n)}*S^{(n+1)}\cdots$.

In this case

$$\mathbf{P}(A|BC)=\prod_{1}^{n}\mathbf{P}^{(k)}(A^{(k)}|(BC)^{(k)}).$$

Putting

$$P_N=\prod_{k=1}^{N}\mathbf{P}^{(k)}((BC)^{(k)}|C^{(k)}),$$

the sequence P_N is monotonically non-increasing, and thus $\lim\limits_{N\to\infty}P_N=p$ exists. Two cases are to be distinguished. If $p=0$, we have $\mathbf{P}(BC|C)=0$ and thus $P(ABC|C)=0$ and therefore (26) is satisfied. If $p>0$, we have

$$\mathbf{P}(ABC|C)=\prod_{1}^{N}\mathbf{P}^{(k)}(A^{(k)}(BC)^{(k)}|C^{(k)})\prod_{N+1}^{\infty}\mathbf{P}^{(k)}((BC)^{(k)}|C^{(k)})$$

and

$$\mathbf{P}(BC|C)=\prod_{k=1}^{\infty}\mathbf{P}^{(k)}((BC)^{(k)}|C^{(k)}),$$

and thus

$$\mathbf{P}(A|BC)\,\mathbf{P}(BC|C)=\prod_{k=1}^{N}\mathbf{P}^{(k)}(A^{(k)}|(BC)^{(k)})\,\mathbf{P}^{(k)}((BC)^{(k)}|C^{(k)})\prod_{N+1}^{\infty}\mathbf{P}^{(k)}((BC)^{(k)}|C^{(k)})$$

and thus, as $\mathbf{P}^{(k)}$ satisfies Axiom III, we have

$$\mathbf{P}^{(k)}(A^{(k)}|(BC)^{(k)})\,\mathbf{P}^{(k)}((BC)^{(k)}|C^{(k)})=\mathbf{P}^{(k)}(A^{(k)}(BC)^{(k)}|C^{(k)})$$

and thus (26) is valid.

§ 2. Examples of conditional probability spaces

2. 1. Conditional probability spaces of simple quotient type. An important class of conditional probability spaces is obtained as follows:

Let S denote a set, $\mathfrak{A}$ a σ-algebra of subsets of S, $\mu(A)$ a measure on $\mathfrak{A}$, $\mathfrak{B}$ the set of those sets $B\in\mathfrak{A}$ for which $\mu(B)$ is positive and finite and put $\mathbf{P}(A|B)=\dfrac{\mu(AB)}{\mu(B)}$ for $A\in\mathfrak{A}$, $B\in\mathfrak{B}$. Then $[S,\mathfrak{A},\mathfrak{B},\mathbf{P}(A|B)]$ is a conditional probability space which will be called a "conditional probability space of simple quotient-type". We mention two special cases.

a) Let us choose for S the n-dimensional Euclidean space E_n, for A the set of all measurable subsets of E_n, let $f(x_1, x_2, \ldots, x_n)$ be a non-negative measurable function on E_n. For $\mathfrak{B}$ we choose the set of those subsets $B \in \mathfrak{A}$ for which

$$\Phi(B) = \int_B f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n$$

is finite and positive and we suppose that $\mathfrak{B}$ is not empty.

$$\mathbf{P}(A|B) = \frac{\Phi(AB)}{\Phi(B)}$$

for $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$. Clearly $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ is a conditional probability space of simple quotient-type, provided that B is not empty.

Especially, when $f(x_1, x_2, \ldots, x_n) \equiv 1$, we obtain a conditional probability space, for which $\mathbf{P}(A|B) = \dfrac{m(AB)}{m(B)}$, where $m(C)$ denotes the Lebesgue measure of the measurable set C; in this case $\mathfrak{B}$ consists of all measurable sets which have a positive and finite Lebesgue measure. The conditional probability space thus obtained will be called the uniform conditional probability space in E_n.

b) Let S denote a finite or denumerable set, with elements $a_1, a_2, \ldots, a_k, \ldots$; let p_k $(k = 1, 2, \ldots)$ denote an arbitrary sequence of non-negative numbers. Let $\mathfrak{A}$ denote the set of all subsets of S and $\mathfrak{B}$ the set of those subsets B of S for which $\sum_{a_k \in B} p_k = \Pi(B)$ is positive and finite. Let us suppose that $\mathfrak{B}$ is not empty and put

$$\mathbf{P}(A|B) = \frac{\Pi(AB)}{\Pi(B)} \quad \text{for} \quad A \in \mathfrak{A} \quad \text{and} \quad B \in \mathfrak{B}.$$

Clearly $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ is a conditional probability space of simple quotient-type. Especially, if $p_k \equiv 1$ $(k = 1, 2, \ldots)$, we obtain a conditional probability space which will be called uniform on the denumerable set S.

2. 2. Conditional probability spaces of alternative quotient-type. A general type of conditional probability spaces is described by the following

THEOREM 14. *Let S denote a set, $\mathfrak{A}_\gamma$ a σ-ring of subsets of S if $\gamma \in \Gamma$ where Γ is an arbitrary ordered set of "indices", and suppose that $\mathfrak{A}_\beta \subseteq \mathfrak{A}_\gamma$ if $\beta < \gamma$ according to the order relation of Γ. Let us put $\mathfrak{A} = \sum_{\gamma \in \Gamma} \mathfrak{A}_\gamma$. Let $\mathfrak{B}_\gamma$ denote a subset of $\mathfrak{A}_\gamma$ and put $\mathfrak{B} = \sum_{\gamma \in \Gamma} \mathfrak{B}_\gamma$; we suppose that $\mathfrak{B}_\gamma$ is not empty for any $\gamma \in \Gamma$ and that $\mathfrak{B}_\beta$ and $\mathfrak{B}_\gamma$ are disjoint for $\beta \neq \gamma$. Let us suppose that for any $\gamma \in \Gamma$ a measure $\mu_\gamma(A)$ is given for $A \in \mathfrak{A}$ (which can take also the value $+\infty$), and that if $B \in \mathfrak{B}_\gamma$, we have $0 < \mu_\gamma(B) < +\infty$. Let us sup-*

pose that if $\beta \in \Gamma$, $\gamma \in \Gamma$, $\beta < \gamma$ *(according to the order relation defined in* Γ*) and* $\mu_\beta(A) < +\infty$, *then* $\mu_\gamma(A) = 0$.

Let $\mathbf{P}(A|B)$ *be defined for* $A \in \mathfrak{A}$ *and* $B \in \mathfrak{B}$ *as follows:*

$$\mathbf{P}(A|B) = \frac{\mu_\gamma(AB)}{\mu_\gamma(B)},$$

where γ *is uniquely determined by the condition* $B \in \mathfrak{B}_\gamma$. *Then* $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ *is a conditional probability space.*

PROOF. Axioms I and II are clearly satisfied. Instead of Axiom III we verify the validity of Axiom III′ which has been shown to be equivalent to Axiom III. Clearly Axiom III′ is satisfied, because if $B \in \mathfrak{B}_\beta$ and $C \in \mathfrak{B}_\gamma$ and $B \subseteq C$, only three cases are possible: $\beta = \gamma$, $\beta < \gamma$ or $\beta > \gamma$. The case $\beta = \gamma$ is evident. If $\beta < \gamma$, $\mathbf{P}(B|C) = \frac{\mu_\gamma(BC)}{\mu_\gamma(C)} = \frac{\mu_\gamma(B)}{\mu_\gamma(C)}$ and as $\mu_\beta(B) < \infty$ and $\beta < \gamma$, we have by supposition $\mu_\gamma(B) = 0$ and therefore $\mathbf{P}(B|C) = 0$; and thus there is nothing to prove.

Finally $\gamma < \beta$ is impossible, because from $C \in \mathfrak{B}\gamma$ it follows $\mu_\gamma(C) < +\infty$ and thus $\mu_\beta(C) = 0$, what implies $\mu_\beta(B) = 0$ because of $B \subseteq C$; but this contradicts $B \in \mathfrak{B}_\beta$. Thus Theorem 14 is proved.

The conditional probability spaces described by Theorem 14 will be called "conditional probability spaces of alternative quotient-type". If $B \in \mathfrak{B}_\gamma$, we may call γ the "dimension" of B, and say that the measures μ_γ are "dimensionally ordered".

We mention two special cases:

a) A special case of a conditional probability space of alternative quotient-type is obtained as follows: let $f(x_1, x_2, \ldots, x_n)$ denote a non-negative Borel measurable function in $S = E_n$; let $\mathfrak{A}$ denote the set of all measurable subsets of E_n, and $\mathfrak{B}$ the set of those measurable sets B, for which either $\Phi(B) = \int_B f(x_1, \ldots, x_n)\, dx_1\, dx_2 \cdots dx_n$ is finite and positive, or B is contained in a k-dimensional subspace $E_{i_1, i_2, \ldots, i_{n-k}}^{c_1, c_2, \ldots, c_{n-k}}$ of E_n $(k = 1, 2, \ldots, n-1)$ defined by $x_{i_1} = c_1, x_{i_2} = c_2, \ldots, x_{i_{n-k}} = c_{i_{n-k}}$, and the integral

$$\Phi_{i_1, i_2, \ldots, i_{n-k}}^{c_1, c_2, \ldots, c_{n-k}} = \int_B f(x_1, x_2, \ldots, x_n)\, dx_{j_1}\, dx_{j_2} \cdots dx_{j_k} \tag{27}$$

is finite and positive, where $j_1, j_2, \ldots, j_k$ denote those of the indices $1, 2, \ldots, n$ which are different from $i_1, i_2, \ldots, i_{n-k}$ and in (26) $x_{i_1} = c_1, x_{i_2} = c_2, \ldots, x_{i_{n-k}} = c_{n-k}$ are substituted; we suppose that $\mathfrak{B}$ is not empty; in the first case we put

$$\mathbf{P}(A|B) = \frac{\Phi(AB)}{\Phi(B)};$$

if $B \in \mathfrak{B}$ and B is contained in $E_{i_1, i_2, \ldots, i_{n-k}}^{c_1, c_2, \ldots, c_{n-k}}$, we put

$$\mathbf{P}(A|B) = \frac{\Phi_{i_1, \ldots, i_{n-k}}^{c_1, \ldots, c_{n-k}}(AB)}{\Phi_{i_1, \ldots, i_{n-k}}^{c_1, \ldots, c_{n-k}}(B)}$$

where $\Phi_{i_1, \ldots, i_{n-k}}^{c_1, \ldots, c_{n-k}}$ is defined by (27).

b) Another special case of a conditional probability space of alternative quotient-type is the following: let $S = E_n$ be the n-dimensional Euclidean space, $\mathfrak{A}$ the set of all Borel measurable subsets of E_n, $\mathfrak{B}_k$ the set of those Borel measurable subsets of E_n, which have finite positive k-dimensional measure[7] $(1 \leqq k \leqq n)$. Let us put $\mathfrak{B} = \sum_{k=1}^{n} \mathfrak{B}_k$ and $\mathbf{P}(A|B) = \frac{m_k(AB)}{m_k(B)}$, if $B \in \mathfrak{B}_k$, where $m_k(B)$ denotes the k-dimensional Lebesgue measure of the set B. Clearly this conditional probability space is an extension of the uniform conditional probability space mentioned in 2. 1, a).

2. 3. Cavalieri spaces. Now we introduce a general concept, that of a "Cavalieri space". Let $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ be a conditional probability space and let us make the following assumptions:

C 1. To every real number t $(\alpha \leqq t < \beta)$ there corresponds a set $C_t \in \mathfrak{B}$.

C 2. For any numbers a and b for which $\alpha \leqq a < b \leqq \beta$

$$C_a^b = \sum_{a \leqq t < b} C_t \in \mathfrak{B}.$$

C 3. If $\alpha \leqq s < t < \beta$, we have $C_s \cdot C_t = O$.

C 4. If for $A \in \mathfrak{A}$, $A' \in \mathfrak{A}$, we have $\mathbf{P}(A|C_t) \leqq \lambda \mathbf{P}(A'|C_t)$, where $\lambda \geqq 0$, for every t lying in the interval $[a, b)$ $(\alpha \leqq a < b \leqq \beta)$, it follows $\mathbf{P}(A|C_a^b) \leqq \leqq \lambda \mathbf{P}(A'|C_a^b)$.

If C 1—C 4 hold, we shall say that $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ is a Cavalieri space with respect to the family of sets $\{C_t; \alpha \leqq t < \beta\}$.

Let us mention that if instead of a family C_t of the power of the continuum we consider a denumerable set $\{C_n\}$ satisfying properties C 1, C 2, and C 3, then (as it has been mentioned in the remark to Theorem 7) C 4 is always satisfied. On the other hand, Theorem 14 shows that conditional probability spaces in general are not Cavalieri spaces with respect to a family $\{C_t\}$ of the power of the continuum, satisfying C 1, C 2 and C 3, because if $\mathbf{P}(C_t|C_a^b) = 0$ for $a \leqq t < b$, $\mathbf{P}(A|C_t)$ can be replaced by any other measure for every $t \in [a, b)$, by leaving $\mathbf{P}(A|C_a^b)$ unchanged.

The following simple consequences of the definition of a Cavalieri space may be mentioned: if $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ is a Cavalieri space with respect to the family of sets $\{C_t; \alpha \leqq t < \beta\}$, then

[7] See [26], p. 108.

I.[8] If $\mathbf{P}(A|C_t) = \lambda \mathbf{P}(A'|C_t)$ for $a \leqq t < b$, we have $\mathbf{P}(A|C_a^b) = \lambda \mathbf{P}(A'|C_a^b)$.

II. If $\mathbf{P}(A|C_t) = \lambda$ for $a \leqq t < b$, we have $\mathbf{P}(A|C_a^b) = \lambda$.

As a matter of fact, if $\mathbf{P}(A|C_t) = \lambda \mathbf{P}(A'|C_t)$ for $a \leqq t < b$, we have $\mathbf{P}(A|C_t) \leqq \lambda \mathbf{P}(A'|C_t)$ and thus by C 4 $\mathbf{P}(A|C_a^b) \leqq \lambda \mathbf{P}(A'|C_a^b)$, further (if $\lambda > 0$) we have $\mathbf{P}(A'|C_t) \leqq \frac{1}{\lambda} \mathbf{P}(A|C_t)$ and thus $\mathbf{P}(A'|C_a^b) \leqq \frac{1}{\lambda} \mathbf{P}(A|C_a^b)$, i. e. $\lambda \mathbf{P}(A'|C_a^b) \leqq \mathbf{P}(A|C_a^b) \leqq \lambda \mathbf{P}(A'|C_a^b)$ (also for $\lambda = 0$) and therefore $\mathbf{P}(A|C_a^b) = \lambda \mathbf{P}(A'|C_a^b)$.

On the other hand, II follows from I by choosing $A' = C_a^b$. The notion of a Cavalieri space can be extended by replacing the one-parametric family $\{C_t\}$ by a k-parametric family $\{C_{t_1, t_2, \ldots, t_k};\ \alpha_i \leqq t_i < \beta_i;\ i = 1, 2, \ldots, k\}$.

2.4. Regular spaces. Let $\mathfrak{S} = [S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ denote a conditional probability space with the following properties:

R 1. To every real number t $(\alpha \leqq t \leqq \beta)$ there corresponds a set $R_t \in \mathfrak{B}$.

R 2. For any pair of numbers a, b $(\alpha \leqq a < b \leqq \beta)$ $R_a^b = \sum_{a \leqq t < b} R_t \in \mathfrak{B}$.

R 3. If $\alpha \leqq s < t < \beta$, we have $R_s \cdot R_t = O$.

R 4. If $A \in \mathfrak{A}$, $\mathbf{P}(A|R_t)$ is a Borel measurable function of t for $\alpha \leqq t < \beta$.

R 5. There exists a strictly increasing and continuous function $F(t)$ in (α, β) with the property that if $\alpha \leqq a < b \leqq \beta$ we have for any $A \in \mathfrak{A}$

$$\mathbf{P}(A|R_a^b) = \frac{\int_a^b \mathbf{P}(A|R_t)\, dF(t)}{\int_a^b dF(t)}. \tag{28}$$

In this case we shall say that $\mathfrak{S}$ is *regular* with respect to the family $\{R_t;\ \alpha \leqq t < \beta\}$. If $F(t)$ is absolutely continuous, we shall say that $\mathfrak{S}$ is *absolutely regular.*

Clearly if $\mathfrak{S}$ is regular with respect to the family $\{R_t\}$, it is also a Cavalieri space with respect to $\{R_t\}$.[9] As a matter of fact, C 1, C 2 and C 3 are identical with R 1, R 2 and R 3, and C 4 follows from R 4 and R 5.

It is easy to prove by using FUBINI's theorem, that if $0 < f(x_1, x_2, \ldots, x_n) < M$, the conditional probability space defined in 2. 2, a) is regular with respect to

[8] This consequence of the definition of a Cavalieri space (for $\lambda = 1$) is analogous with the well-known "principle of Cavalieri" (according to which if the areas of the intersections of two solids with every horizontal plane are equal, the two solids have the same volume) introduced by BONAVENTURA CAVALIERI (1598—1647) in 1635. This justifies the name "Cavalieri space".

[9] J. CZIPSZER has shown that there exist Cavalieri spaces with respect to a family $\{R_t\}$ of sets which are not regular with respect to the same family $\{R_t\}$.

the family $\{R_t\}$ defined as follows: R_t is the set of points $(x_1, x_2, \ldots, x_n)$ for which $a_i \leqq x_i < b_i$ for $i \neq k$, $i = 1, 2, \ldots, n$, and $x_k = t$ $(\alpha \leqq t < \beta)$.

It is clear that if $\mathfrak{S}$ is regular with respect to the family $\{R_t; \alpha \leqq t < \beta\}$, we have

$$\lim_{h>0,\, h\to 0} \mathbf{P}(A|R_t^{t+h}) = \mathbf{P}(A|R_t) \tag{29}$$

for almost every t in (α, β) with respect to the measure generated by $F(t)$ on (α, β) for any fixed A.

It is also evident that if $\mathfrak{S}$ is regular with respect to the family $\{R_t; \alpha \leqq t < \beta\}$, we have $\mathbf{P}(R_t|R_\alpha^b) = 0$ for $\alpha \leqq a \leqq t < b \leqq \beta$.

Let $\xi = \xi(a)$ $(a \in S)$ be a random variable on a conditional probability space $\mathfrak{S} = [S, \mathcal{A}, \mathfrak{B}, \mathbf{P}(A|B)]$, further let us denote by R_t the set of those $a \in S$ for which $\xi(a) = t$ $(-\infty \leqq t < +\infty)$. Let us suppose that $\mathfrak{S}$ is regular with respect to the family $\{R_t\}$, i. e. we have for $A \in \mathcal{A}$

$$\mathbf{P}(A|R_\alpha^b) = \frac{\int_a^b \mathbf{P}(A|R_t)\, dF(t)}{\int_a^b dF(t)} \tag{30}$$

for $-\infty < a < b < +\infty$. Evidently the conditional distribution generated by ξ on the real axis is the distribution

$$P(a \leqq \xi < b | c \leqq \xi < d) = \mathbf{P}(R_a^b|R_c^d) = \frac{F(b) - F(a)}{F(d) - F(c)} \quad \text{for} \quad c \leqq a < b \leqq d,$$

i. e. $F(x)$ is the (generalized) distribution function of ξ. If (30) holds, we shall say that ξ has a regular conditional distribution with the distribution function $F(x)$; if $F(x) = x$, we shall say that ξ has a regular uniform distribution.

The notion of a regular space can be extended by replacing the one-parametric family $\{R_t\}$ by a k-parametric family $\{R_{t_1, t_2, \ldots, t_k}\}$. We consider only the case when the following postulates are satisfied:

R 1. To every point $T = (t_1, t_2, \ldots, t_k)$ of a k-dimensional interval $I = (\alpha_i \leqq t_i < \beta_i;\ i = 1, 2, \ldots, k)$ there corresponds a set $R_t \in \mathfrak{B}$.

R 2. For any subinterval $J = (a_i \leqq t_i < b_i;\quad i = 1, 2, \ldots, k)$ where $\alpha_i \leqq a_i < b_i \leqq \beta_i$ $(i = 1, 2, \ldots, k)$ we have $R_J = \sum_{T \in J} R_T \in \mathfrak{B}$.

R 3. If $T_1 \neq T_2$, we have $R_{T_1} \cdot R_{T_2} = O$.

R 4. If $A \in \mathcal{A}$, $\mathbf{P}(A|R_T)$ is a Borel-measurable function of the variables $t_1, t_2, \ldots, t_k$.

R 5. There exists a function $f(t_1, t_2, \ldots, t_k)$ which is positive and integrable on I, and if J is the interval $a_i \leqq t_i < b_i$ $(i = 1, 2, \ldots, k)$,

$(\alpha_i \leqq a_i < b_i \leqq \beta_i;\ i = 1, 2, \ldots, k)$, we have

$$\mathbf{P}(A|R_J) = \frac{\int_J \mathbf{P}(A|R_T) f(t_1, t_2, \ldots, t_k)\, dt_1 \ldots dt_k}{\int_J f(t_1, \ldots, t_k)\, dt_1 \ldots dt_k}.$$

If R 1—R 5 hold, and $f(t_1, \ldots, t_k) = \prod_{j=1}^{k} f_j(t_j)$, we say that *the space is absolutely regular with respect to the family* $\{R_T\}$ *with independent parameters.*

Especially if R_T is the set defined by $\xi_1 = t_1, \ldots, \xi_k = t_k$ where $\xi_1, \xi_2, \ldots, \xi_k$ are random variables defined on $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ and $f(t_1, \ldots, t_n) = \prod_{j=1}^{k} f_j(t_j)$, the random variables $\xi_1, \ldots, \xi_k$ are independent with respect to any R_J, further ξ_j has a regular conditional distribution with the generalized density function $f_j(t)$ $(j = 1, 2, \ldots, k)$.

2. 5. Remark on Borel's paradox. The notions of Cavalieri spaces resp. regular spaces introduced in the preceding sections, are connected with the well-known paradox of BOREL. (Cf. [1], p. 44.) Let S denote the surface of the unit sphere $x^2 + y^2 + z^2 = 1$; let us introduce spherical coordinates φ and ϑ, defined by $x = \cos\varphi \sin\vartheta$, $y = \sin\varphi \sin\vartheta$, $z = \cos\vartheta$ $(0 \leqq \varphi < 2\pi;\ 0 \leqq \vartheta < \pi)$. Let us denote by $\mathcal{A}$ the set of all Borel subsets of S (i. e. sets A for which the set of all pairs of numbers (ϑ, φ) for which the corresponding point belongs to A, is a plane Borel set). Let us denote by $\mathfrak{B}$ the set of all sets $B_{a,c}^{b,d}$ defined by the inequalities $a \leqq \varphi \leqq b$, $c \leqq \vartheta \leqq d$, where $0 \leqq a \leqq b \leqq 2\pi$ and $0 \leqq c < d \leqq \pi$. Let us put

$$\mathbf{P}(A|B_{a,c}^{b,d}) = \frac{\int_A \int_{B_{a,c}^{b,d}} \sin\vartheta\, d\vartheta\, d\varphi}{\int_a^b \int_c^d \sin\vartheta\, d\vartheta\, d\varphi} \qquad \text{if} \quad a < b,$$

i. e. let $\mathbf{P}(A|B)$ for fixed B be proportional to the area of AB. Let us denote by B_t the set $\varphi = t, 0 \leqq \vartheta \leqq \pi$ (i. e. put $B_t = B_{t,0}^{t,\pi}$); let us choose a bounded measure $\mu_t(B_t) = 1$ defined on the Borel subsets of B_t and normed by $\mu_t(B_t) = 1$, and put

$$\mathbf{P}(A|B_{t,c}^{t,d}) = \frac{\mu_t(AB_{t,c}^{t,d})}{\mu_t(B_{t,c}^{t,d})} \qquad (0 \leqq c < d \leqq \pi),$$

and let us suppose that $\mathbf{P}(B_{0,0}^{2\pi,x}|B_t) = F(x,t)$ is a continuous function of x and t, strictly increasing as a function of x. Clearly $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}(A|B)] = \mathcal{S}$ is a conditional probability space by any choice of the measures $\mu_t(A)$; we

shall prove that it is a Cavalieri space with respect to the family $\{B_t;\ 0 \leqq t < 2\pi\}$ if and only if

$$\mu_t(A) = \frac{1}{2}\int_{AB_t} \sin\vartheta\, d\vartheta. \tag{31}$$

This can be shown as follows: let us suppose that $\mathfrak{S}$ is a Cavalieri space with respect to $\{B_t;\ 0 \leqq t < 2\pi\}$; let us choose a number λ $(0 \leqq \lambda < 1)$ and define $x(t, \lambda)$ as the number for which, if $A_t(\lambda)$ denotes the interval $0 \leqq \vartheta \leqq x(t, \lambda)$, we have $\mathbf{P}(A_t(\lambda)|B_t) = \lambda$, i. e. $F(x(t, \lambda), t) = \lambda$. Now it follows from our supposition, that $x(t, \lambda)$ is a continuous function of t and thus if $A(\lambda) = \sum_{0 \leqq t < 2\pi} A_t(\lambda)$, we have $A(\lambda) \in \mathfrak{A}$, further $\mathbf{P}(A(\lambda)|B_t) = \lambda$ for $0 \leqq t < 2\pi$. It follows by II of 2.3 that putting $B_{a,0}^{b,\pi} = B_a^b$, we have $\mathbf{P}(A(\lambda)|B_a^b) = \lambda$ for $0 \leqq a < b < 2\pi$, i. e. $\frac{1}{2}\int_a^b \int_0^{x(t,\lambda)} \sin\vartheta\, d\vartheta\, dt = \lambda(b-a)$. By taking the derivative of both sides, with respect to b, by virtue of the continuity of $x(t, \lambda)$, we obtain that $\frac{1}{2}\int_0^{x(t,\lambda)} \sin\vartheta\, d\vartheta = \lambda$ for every t $(0 \leqq t < 2\pi)$. Thus, it follows easily that (31) holds.

Our result can be stated as follows: the uniform probability distribution on the surface of the sphere combined with arbitrary distributions on the meridians $\varphi = t$ gives always a conditional probability space, but this space is a Cavalieri space with respect to the set of meridians (under suitable continuity restrictions) if and only if the density of the distribution is on every meridian the same, and equal to $\frac{1}{2}\sin\vartheta$. In this case the conditional probability space is clearly also regular with respect to the set of meridians, as we have

$$\mathbf{P}(A|B_a^b) = \frac{1}{b-a}\int_a^b \left(\int_{AB_t} \frac{1}{2}\sin\vartheta\, d\vartheta \right) dt = \frac{\int_a^b \mathbf{P}(A|B_t)\, dt}{\int_a^b dt}.$$

2.6. Composition of conditional probability distributions. Let us suppose that $\xi_j = \xi_j(a)$ $(j = 1, 2, \ldots, k;\ a \in S)$ are positive random variables on the conditional probability space $\mathfrak{S} = [S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$. Let us denote by $R_{t_1, t_2, \ldots, t_k}$ the set of those $a \in S$ for which $\xi_1(a) = t_1, \ldots, \xi_k(a) = t_k$ and suppose that the space $\mathfrak{S}$ is absolutely regular with respect to the family $\{R_{t_1, \ldots, t_k};\ 0 \leqq t_i < +\infty;\ i = 1, 2, \ldots, k\}$, with independent parameters and thus the functions $F_j(T_j)$ figuring in R 5 are absolutely continuous and $F_j'(t) = f_j(t)$

is almost everywhere positive $(0 < t < +\infty)$. Let us suppose further that if M is a Borel subset of the k-dimensional Euclidean space, $\sum_{(t_1, \ldots, t_k) \in M} R_{t_1, \ldots, t_k}$ belongs to $\mathfrak{B}$. In this case if I and $J \subseteq I$ are k-dimensional intervals,

$$I = (\alpha_i \leqq t_i < \beta_i;\ i = 1, 2, \ldots, k) \quad \text{and} \quad J = (a_i \leqq t_i < b_i;\ i = 1, 2, \ldots, k)$$

where $0 \leqq \alpha_i \leqq a_i < b_i \leqq \beta_i$ $(i = 1, 2, \ldots, k)$, we have, putting $\vec{\xi} = (\xi_1, \ldots, \xi_k)$,

$$\mathbf{P}(\vec{\xi} \in J \mid \vec{\xi} \in I) = \prod_{j=1}^{k} \frac{\int_{a_j}^{b_j} f_j(t_j)\,dt}{\int_{\alpha_j}^{\beta_j} f_j(t)\,dt}. \tag{32}$$

Thus the random variables $\xi_1, \xi_2, \ldots, \xi_k$ are independent with respect to any R_I.

Let us calculate the conditional probability distribution of $\zeta_k = \xi_1 + \xi_2 + \cdots + \xi_k$. Let B_T be the interval $0 \leqq t_j < T$ $(j = 1, 2, \ldots, k)$ and A_z the set defined by $\zeta_k < z$. Clearly, $A_z \in \mathfrak{B}$, as $A_z = \sum_{t_1 + t_2 + \ldots + t_k < z} B_{t_1, \ldots, t_k}$; we have by R 5

$$\mathbf{P}(A_z \mid B_T) = \frac{\int_{\sum_{j=1}^{k} t_j < z,\ 0 \leqq t_j < T} \prod_{j=1}^{k} f_j(t_j)\,dt_j}{\prod_{j=1}^{k} \left(\int_0^T f_j(t)\,dt \right)}.$$

Now let us suppose that $0 \leqq z_1 < z_2 < T$; in this case $A_{z_1} \subseteq A_{z_2} \subseteq B_T$ and thus

$$\mathbf{P}(A_{z_i} \mid B_T) = \frac{\int_{\sum_{j=1}^{k} t_j < z_i} \prod_{j=1}^{k} f_j(t_j)\,dt_j}{\prod_{j=1}^{k} \left(\int_0^T f_j(t)\,dt \right)}$$

and

$$\mathbf{P}(A_{z_1} \mid A_{z_2}) = \mathbf{P}(A_{z_1} \mid A_{z_2} B_T) = \frac{\mathbf{P}(A_{z_1} A_{z_2} \mid B_T)}{\mathbf{P}(A_{z_2} \mid B_T)} = \frac{\mathbf{P}(A_{z_1} \mid B_T)}{\mathbf{P}(A_{z_2} \mid B_T)}$$

and thus

$$\mathbf{P}(A_{z_1} \mid A_{z_2}) = \frac{\int_{\sum_{j=1}^{k} t_j < z_1} \prod_{j=1}^{k} f_j(t_j)\,dt_j}{\int_{\sum_{j=1}^{k} t_j < z_2} \prod_{j=1}^{k} f_j(t_j)\,dt_j}. \tag{33}$$

It follows that if we put

$$g_1(t)=f_1(t) \quad \text{and} \quad g_j(t)=\int_0^t g_{j-1}(t-u)f_j(u)du \quad \text{if} \quad t\geqq 0$$

and

$$g_j(t)=0 \quad \text{if} \quad t<0 \qquad (j=2,3,\ldots,k),$$

we have by (33)

$$\mathbf{P}(A_{z_1}|A_{z_2})=\frac{\int_0^{z_1} g_k(t)dt}{\int_0^{z_2} g_k(t)dt}. \tag{34}$$

Thus $g_k(t)$ is the generalized density function of ζ_k.

If $\int_0^\infty f_j(t)dt<+\infty$ $(j=1,2,\ldots,k)$, we obtain as a special case the well-known law of composition of independent probability distributions, but clearly some or all of the integrals $\int_0^\infty f_j(t)dt$ may be divergent.

The method of Laplace transforms can be applied also in the case when $\int_0^\infty f_j(t)dt=+\infty$, but $\int_0^\infty e^{-st}f_j(t)dt=\varphi_j(s)$ exists for some $s>0$ $(j=1,2,\ldots,k)$ and in this case we have, putting $\psi_j(s)=\int_0^\infty e^{-st}g_j(t)dt$, evidently $\psi_k(s)=\prod_{j=1}^k \varphi_j(s)$.

If e. g. ξ and η are independent, regularly distributed random variables which have the (generalized) density functions $x^{\alpha-1}$ resp. $x^{\beta-1}$ in $(0,+\infty)$, where $\alpha>0$ and $\beta>0$, the (generalized) density function of the random variable $\xi+\eta$ is $x^{\alpha+\beta-1}$. A striking property of the random variables with density functions $x^{\alpha-1}$ $(0<x<+\infty)$ is the following: if ξ has the density function $x^{\alpha-1}$ in $(0,+\infty)$ and $C>0$ is a constant, the random variable $C\xi$ has the same density function.

The problem of composition of distributions in the general case (i. e. if it is not supposed that the random variables considered are positive) is more intricate.[10] If ξ and η are independent random variables with regular joint distribution and density functions $f(x)$, $g(y)$, respectively, and the limit

$$\lim_{A\to\infty}\frac{\int_{-A}^{+A} f(x-y)g(y)dy}{\int_{-A}^{+A} f(x_0-y)g(y)dy}=h(x)$$

[10] The theory of "distributions" of L. SCHWARTZ has to be applied.

exists uniformly in x and is positive on a set of positive measure, then $h(x)$ is the density function of $\xi+\eta$. This implies for example that if the distribution of η is uniform on the real axis, the distribution of $\xi+\eta$ is also uniform, whatever the density function of ξ should be.

§ 3. Some applications of the notion of a conditional probability space

In this § we consider simple applications of the notions introduced in § 1 and § 2 to some problems of ordinary probability theory. The random variables considered in this § are thus random variables defined on an ordinary probability space $[S, \mathcal{A}, \mathbf{P}]$, except when it is explicitly stated that they are defined on a conditional probability space.

3. 1. The limit of the Poisson distribution for $\lambda \to +\infty$. Let ξ_λ denote a random variable which is distributed according to Poisson's law with mean value $\lambda > 0$. The random variable ξ_λ generates the probability distribution on the set $\mathfrak{I}_+$ of non-negative integers

$$p_k(\lambda) = \mathbf{P}(\xi_\lambda = k) = \frac{\lambda^k e^{-\lambda}}{k!} \qquad (k = 0, 1, \ldots).$$

It follows easily from Stirling's formula that

$$\lim_{\lambda \to \infty} \frac{p_{[\lambda]+j}(\lambda)}{p_{[\lambda]+k}(\lambda)} = 1 \tag{35}$$

for $j, k = 0, \pm 1, \pm 2, \ldots$; thus if $\mathfrak{I}$ is the set of all integers, A and B are finite subsets of $\mathfrak{I}$ and B is not empty, then we have

$$\lim_{\lambda \to \infty} \mathbf{P}(\xi_\lambda - [\lambda] \in A \mid \xi_\lambda - [\lambda] \in B) = \frac{n(AB)}{n(B)},$$

where $n(B)$ denotes the number of elements of B. The result can be stated as follows: the conditional probability distribution generated by $\xi_\lambda - [\lambda]$ on $\mathfrak{I}$ tends to the uniform conditional probability distribution on $\mathfrak{I}$ if $\lambda \to +\infty$.

3. 2. The number of prime divisors. An interesting application of the result in 3. 1 to number theory is as follows: let $U(n)$ denote the number of different prime divisors of the integer n; let $\nu_k(N)$ denote the number of those natural numbers $n \leqq N$ for which $U(n) = k$. By a theorem of P. ERDŐS [16] we have

$$\frac{\nu_k(N)}{N} = \frac{(\log\log N)^k}{k!} e^{-\log\log N} \cdot (1 + o(1)) \tag{36}$$

for $|k - \log\log N| < c\sqrt{\log\log N}$, where $c > 0$ is constant, and $o(1)$ tends uniformly to 0 for $N \to \infty$, if c is fixed. It follows that if $D(1) = 0$, $D(2) = 1$ and $D(n) = U(n) - [\log\log n]$ for $n = 3, 4, \ldots$ and $P_k(N)$ denotes the number

of positive integers $n \leqq N$ for which $D(n) = k$, we have by (35) and (36)

$$\lim_{N\to\infty} \frac{P_j(N)}{P_k(N)} = 1$$

for $j, k = 0, \pm 1, \pm 2, \ldots$; thus the conditional probability distribution of the number-theoretical function $D(n)$ $(1 \leqq n \leqq N)$ tends for $N \to \infty$ to the uniform conditional probability distribution on $\mathfrak{I}$.

3.3. The normal distribution for $\sigma \to +\infty$. Let ξ_σ be a random variable, normally distributed with mean value 0 and variance σ^2. Then we have for $c \leqq a < b \leqq d$

$$\mathbf{P}(a \leqq \xi_\sigma < b \mid c \leqq \xi_\sigma < d) = \frac{\int_{\frac{a}{\sigma}}^{\frac{b}{\sigma}} e^{-\frac{x^2}{2}} dx}{\int_{\frac{c}{\sigma}}^{\frac{d}{\sigma}} e^{-\frac{x^2}{2}} dx}$$

and thus

$$\lim_{\sigma\to+\infty} \mathbf{P}(a \leqq \xi_\sigma < b \mid c \leqq \xi_\sigma < d) = \frac{b-a}{d-c},$$

i. e. the conditional probability distribution, generated by ξ_σ on the real axis $\mathfrak{R}$ tends for $\sigma \to +\infty$ to the uniform conditional probability distribution on $\mathfrak{R}$.

3.4. Limiting distribution of the sum of independent random variables. Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ denote independent random variables; let us suppose that the variables ξ_n have the same distribution with the density function $f(x)$ for which $\int_{-\infty}^{+\infty} f^2(x)dx$ exists.[11] Let us suppose that $\mathbf{M}(\xi_n) = 0$ and

$$\mathbf{M}(\xi_n^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx = D^2 < +\infty.$$

Let us put $\zeta_n = \xi_1 + \xi_2 + \cdots + \xi_n$. Putting

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{ixt} f(x)dx \qquad (-\infty < t < +\infty),$$

we have $\int_{-\infty}^{+\infty} |\varphi(t)|^n dt < +\infty$ for $n \geqq 2$ (see [17]), and thus denoting by $f_n(x)$ the density function of ζ_n, we have

$$f_n(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} (\varphi(t))^n e^{-itx} dt. \tag{37}$$

[11] These conditions can be replaced by weaker ones.

As by supposition $|\varphi(t)|+1$ for $t+0$ and $\lim_{t\to\pm\infty}|\varphi(t)|=0$, it follows easily by the method of LAPLACE (see [17]) that

$$f_n(x)=\frac{1}{D\sqrt{2\pi n}}(1+o(1)), \tag{38}$$

where $o(1)\to 0$ if $n\to\infty$, uniformly for $|x|\leqq C$, where C does not depend on n; thus for $c\leqq a<b\leqq d$ we have

$$\lim_{n\to\infty}\mathbf{P}(a\leqq\zeta_n<b|c\leqq\zeta_n<d)=\frac{b-a}{d-c}, \tag{39}$$

i. e. the conditional probability distribution generated by ζ_n on $\mathfrak{R}$ tends to the uniform conditional probability distribution on $\mathfrak{R}$ for $n\to\infty$.

3.5. A connection between the uniform conditional probability distribution on $\mathfrak{R}_+$ and on $\mathfrak{I}_+$. Let $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ denote a conditional probability field, $\xi=\xi(a)$ $(a\in S)$ a random variable, which takes only non-negative integer values; let us denote by A_k the set of those $a\in S$ for which $\xi(a)=k$, put $B_n=\sum_{k=0}^{n}A_k$, and let us suppose that $A_k\in\mathfrak{B}$ and $B_n\in\mathfrak{B}$ $(k, n=0, 1, 2, \ldots)$. Let us suppose that $\lambda=\lambda(a)$ is an other random variable which generates a regular uniform conditional probability distribution on the positive real axis $\mathfrak{R}_+$; let us denote by C_t the set of those $a\in S$ for which $\lambda(a)=t$ and put $D_x=\sum_{0\leqq y<x}C_y$; let us suppose that $C_x\in\mathfrak{B}$, $D_x\in\mathfrak{B}$, further that $B_nD_x\in\mathfrak{B}$ $(0<x<+\infty,\ n=0, 1, 2, \ldots)$. We suppose

$$\mathbf{P}(A_k|C_t)=\frac{t^k e^{-t}}{k!}\qquad (k=0, 1, \ldots), \tag{40}$$

i. e. that ξ has a Poisson distribution with mean value t under the hypothesis that $\lambda=t$. We shall prove that in this case the conditional probability distribution generated by ξ on the set of non-negative integers $\mathfrak{I}_+$ is the uniform distribution, i. e.

$$\mathbf{P}(A_k|B_n)=\mathbf{P}(\xi=k|\xi\leqq n)=\frac{1}{n+1}\quad\text{for}\quad k=0, 1, \ldots, n;\ n=0, 1, \ldots.$$

This can be shown as follows: We have by Theorem 12

$$\mathbf{P}(A_k|B_n)=\lim_{x\to\infty}\mathbf{P}(A_k|B_nD_x)=\lim_{x\to\infty}\frac{\mathbf{P}(A_k|D_x)}{\mathbf{P}(B_n|D_x)}. \tag{41}$$

As we have supposed that the distribution of λ is not only uniform on $\mathfrak{R}_+$ but also regular, we have

$$\mathbf{P}(A|D_x)=\frac{1}{x}\int_0^x\mathbf{P}(A|C_t)dt \tag{42}$$

and thus by (41) and (42)

$$\mathbf{P}(A_k|B_n)=\frac{\int_0^\infty \mathbf{P}(A_k|C_t)\,dt}{\sum_{k=0}^{n}\int_0^\infty \mathbf{P}(A_k|C_t)\,dt}. \tag{43}$$

As by (40)

$$\int_0^\infty \mathbf{P}(A_k|C_t)\,dt=\int_0^\infty \frac{t^k e^{-t}}{k!}\,dt=1,$$

it follows from (43) that

$$\mathbf{P}(A_k|B_n)=\frac{1}{n+1}$$

which was to be proved.

A realisation of what has been said can be given as follows: let S denote the strip $0\leqq x<+\infty$, $0\leqq y<1$ of the (x,y) plane, $\mathfrak{A}$ the set of all measurable subsets of S, $\mathfrak{B}$ the set of all Borel-measurable subsets of S with

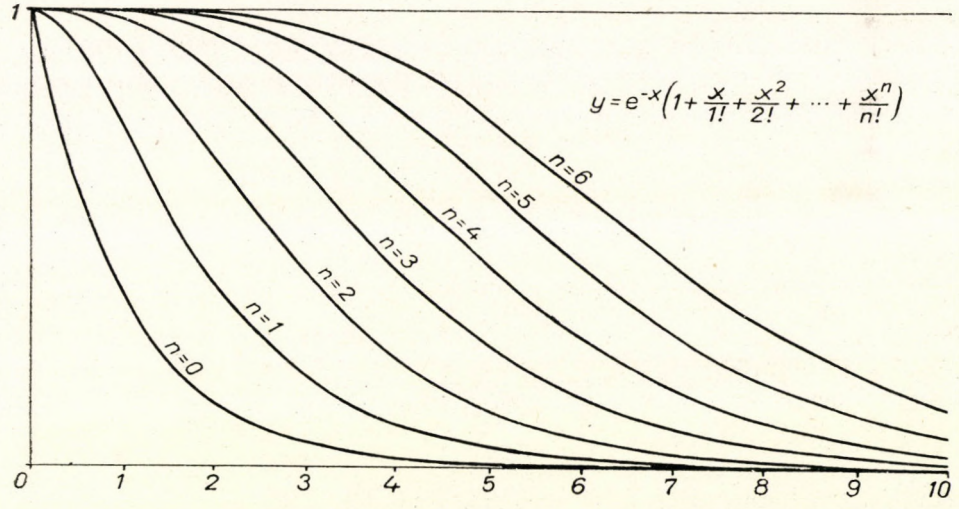


Fig. 1

positive and finite measure, and of all sets of positive linear measure lying on some line $x=t$ $(0\leqq t<+\infty)$, and let us put $\mathbf{P}(A|B)=\frac{m_2(AB)}{m_2(B)}$, if $m_2(B)>0$, where $m_2(B)$ and $m_2(AB)$ denote the plane Lebesgue measure of B and AB, respectively, and put $\mathbf{P}(A|B)=\frac{m_1(AB)}{m_1(B)}$, if B is lying on a line $x=t$, where

$m_1(B)$ denotes the linear Lebesgue measure of B. Define λ as follows: $\lambda = \lambda(x, y) = x$, and define $\xi = \xi(x, y)$ as follows: for a fixed value of x let us have $\xi = 0$ for $0 \leqq y < e^{-x}$, and $\xi = k$ for $\sum_{j=0}^{k-1} \frac{x^j e^{-x}}{j!} \leqq y < \sum_{j=0}^{k} \frac{x^j e^{-x}}{j!}$ $(k = 1, 2, \ldots)$.

The sets A_k are shown on Fig. 1.

Clearly each set A_k $(k = 0, 1, \ldots)$ has the area 1. Our result can be stated in the following suggestive form: if ξ has the Poisson distribution with mean value t under the hypothesis $\lambda = t$ and the conditional distribution of λ is regular and uniform on $\mathfrak{R}_+$, ξ is uniformly distributed on $\mathfrak{I}_+$. Clearly the result obtained is a consequence of the following remarkable property of the Poisson distribution: each member $\frac{\lambda^k e^{-\lambda}}{k!}$ of the Poisson distribution is as a function of λ for fixed k a probability density function in the interval $(0 \leqq \lambda < +\infty)$.

3.6. Remarks on the method of Bayes. Let $\mathfrak{S} = [S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ denote a conditional probability space. Let $\xi = \xi(a)$ and $\eta = \eta(a)$ $(a \in S)$ be random variables on $\mathfrak{S}$ and suppose that $\mathfrak{S}$ is regular with respect to the family of sets R_{xy} defined by $\xi = x, \eta = y$ for $-\infty < x < +\infty$; $-\infty < y < +\infty$ and let us suppose that $R_{c,\gamma}^{d,\delta} = \sum_{c<x<d,\, \gamma \leqq y < \delta} R_{xy} \in \mathfrak{B}$ where $-\infty < c < d < +\infty$, $-\infty < \gamma < \delta < +\infty$, further if $-\infty \leqq c < d \leqq +\infty$ and $-\infty < \gamma < \delta < +\infty$, we have

$$\mathbf{P}(A|R_{c,\gamma}^{d,\delta}) = \frac{\int_c^d \int_\gamma^\delta \mathbf{P}(A|R_{xy}) h(x, y)\, dx\, dy}{\int_c^d \int_\gamma^\delta h(x, y)\, dx\, dy}, \tag{44}$$

where $h(x, y)$ is positive and integrable on any finite rectangle of the (x, y) plane.

Let us suppose further that $R_{c,y}^{d,y} \in \mathfrak{B}$ if $c < d$ and $R_{x,\gamma}^{x,\delta} \in \mathfrak{B}$ if $\gamma < \delta$, further that we have

$$\mathbf{P}(A|R_{c,y}^{d,y}) = \frac{\int_c^d \mathbf{P}(A|R_{xy}) h(x, y)\, dx}{\int_c^d h(x, y)\, dx} \qquad (c < d) \tag{45}$$

and

$$\mathbf{P}(A|R_{\gamma,x}^{\delta,x}) = \frac{\int_\gamma^\delta \mathbf{P}(A|R_{xy}) h(x, y)\, dy}{\int_\gamma^\delta h(x, y)\, dy} \qquad (\gamma < \delta). \tag{46}$$

In this case clearly for any fixed y, $h(x, y)$ is, as a function of x, the conditional density function of ξ under hypothesis $\eta = y$ and, for any fixed x, $h(x, y)$ is, as a function of y, the conditional density function of η under the hypothesis $\xi = x$. As a matter of fact, choosing $\gamma = \delta = y$ we have for $c \leqq a < b \leqq d$ from (45)

$$\mathbf{P}(a \leqq \xi < b | c \leqq \xi < d, \eta = y) = \frac{\int_a^b h(x, y)dx}{\int_c^d h(x, y)dx} \tag{47}$$

and choosing $\alpha = \beta = x$ we have for $\gamma \leqq \alpha < \beta \leqq \delta$ from (44)

$$\mathbf{P}(\alpha \leqq \eta < \beta | \gamma \leqq \eta < \delta, \xi = x) = \frac{\int_\alpha^\beta h(x, y)dy}{\int_\gamma^\delta h(x, y)dy}. \tag{48}$$

It should be mentioned that the set R_x defined by $\xi = x$ is not supposed to belong to $\mathfrak{B}$, and thus $h(x, y)$ is not an ordinary conditional density function, but only a generalized conditional density function, in the sense defined in 1.6. By other words, $h(x, y)$ is the generalized density function of the random variable η on the subspace $(R_x, \mathfrak{A}_x, \mathfrak{B}_x, \mathbf{P}_x)$ of $(S, A, B, \mathbf{P})$ where $\mathfrak{A}_x$ is the set of all sets of the form AB_x with $A \in \mathfrak{A}$, $\mathfrak{B}_x$ the set of all sets $B \in \mathfrak{B}$ which are subsets of R_x and $\mathbf{P}_x(A|B)$ is defined by

$$\mathbf{P}_x(A|B) = \mathbf{P}(A|B)$$

for $A \in \mathfrak{A}_x$ and $B \in \mathfrak{B}_x$. If $R_x \in \mathfrak{B}$, clearly $h(x, y)$ is up to a constant factor the ordinary conditional density function of η under the condition R_x and in this case $f(x) = \int h(x, y)dy$ is finite.

If $f(x) = \int_{-\infty}^{+\infty} h(x, y)dy$ is finite for every x, then $f(x)$ is the generalized density function of ξ. As a matter of fact, we have by Theorem 12

$$\mathbf{P}(a \leqq \xi < b | c \leqq \xi < d) = \lim_{\substack{\gamma \to -\infty \\ \delta \to +\infty}} \mathbf{P}(a \leqq \xi < b | c \leqq \xi < d, \gamma \leqq \eta < \delta) = \frac{\int_a^b f(x)dx}{\int_c^d f(x)dx}.$$

Putting

$$g(y|x) = \frac{h(x, y)}{f(x)} \tag{49}$$

besides $h(x, y)$, $g(y|x)$ is the normed (i. e. ordinary) conditional density func-

tion of η under condition $\xi = x$. Similarly, if we denote by $f(x|y)$ the generalized conditional density function of ξ under the condition $\eta = y$, we may put $f(x|y) = h(x, y)$ and thus we have

$$f(x|y) = g(y|x)f(x). \tag{50}$$

If $\int_{-\infty}^{+\infty} h(x, y)\,dx$ is also finite,

$$f^*(x|y) = \frac{g(y|x)f(x)}{\int_{-\infty}^{+\infty} g(y|x)f(x)\,dx} \tag{51}$$

is the normed (ordinary) conditional density function of ξ under the hypothesis $\eta = y$.

The result obtained in this § is a generalization of the well-known method of BAYES. The generalization consists in the possibility that $\int_{-\infty}^{+\infty} f(x)\,dx$ and $\int_{-\infty}^{+\infty} f(x|y)\,dx$ may be divergent.

Clearly if $h(x, y)$ vanishes in some points of the plane, all what we have said remains valid, only we may use as conditions only such sets $c \leqq \xi < d$, $\gamma \leqq \eta < \delta$ for which $\int_c^d \int_\gamma^\delta h(x, y)\,dx\,dy > 0$.

3.7. Conditional ergodicity of Markov chains. Let us consider a temporally homogeneous Markov chain with a denumerable set of possible states; let us denote the states by the numbers $0, \pm 1, \pm 2, \ldots$ and put $\xi_n = k$ if the system is in state k at time n $(n = 0, 1, 2, \ldots)$. Let us denote by P_{jk} the probability of passing from state j to state k in one step, i. e. $P_{jk} = \mathbf{P}(\xi_{n+1} = k | \xi_n = j)$. We introduce the following notion: if there exists a sequence P_k of positive numbers, with the property that, putting $\mathbf{P}(\xi_n = k | \xi_0 = j) = P_{jk}^{(n)}$, we have

$$\lim_{n \to \infty} \frac{P_{hi}^{(n)}}{P_{jk}^{(n)}} = \frac{P_i}{P_k} \tag{52}$$

for all $h, i, j, k = 0, \pm 1, \pm 2, \ldots$, we shall say that the Markov chain is conditionally ergodic. If the chain is ergodic in the usual sense (see [18]), i. e. if

$$\lim_{n \to \infty} P_{ik}^{(n)} = P_k > 0 \qquad (i, k = 0, \pm 1, \pm 2, \ldots), \tag{53}$$

then (52) holds and we have $\sum_{k=-\infty}^{+\infty} P_k = 1$.

It is possible, however, that (52) holds without (53) taking place; especially, this is the case if (52) holds and $\sum_{k=-\infty}^{+\infty} P_k = +\infty$. If a Markov

chain is conditionally ergodic, this implies that the conditional probability distribution generated by ξ_n on the set of integers converges to the conditional probability distribution defined by $\mathbf{P}(A|B) = \frac{\sum_{k \in AB} P_k}{\sum_{k \in B} P_k}$ where A and B are subsets of the set of integers and B is finite and not empty.

Let us suppose that the Markov chain considered is irreducible and all its states are recurrent null states (see [18]) in which case $\lim_{n \to \infty} P_{jk}^{(n)} = 0$. If the numbers P_k in (52) satisfy the system of equations

$$P_k = \sum_{j=-\infty}^{\infty} P_j P_{jk} \qquad (k = 0, \pm 1, \pm 2, \ldots), \tag{54}$$

we say that the conditional limiting distribution defined by the numbers P_k is a stationary distribution. If the conditional limiting distribution exists, it is always a stationary distribution, i. e. (52) implies (54). As a matter of fact, as we have supposed that all states are recurrent, the series $\sum_{n=1}^{\infty} P_{kk}^{(n)} (k = 0, \pm 1, \ldots)$ diverges, and thus, it follows from (52) that

$$V_k = \lim_{n \to \infty} \frac{\sum_{r=1}^{n} P_{kk}^{(r)}}{\sum_{r=1}^{n} P_{00}^{(r)}} = \frac{P_k}{P_0}. \tag{55}$$

But it has been proved by C. DERMAN [19] that if all states of an irreducible chain are recurrent null states, the only solution $P_k = V_k$ of the system of equations (54) with $V_0 = 1$ is given by the limits in (55) which always exist.

The existence of the limits (52) has been proved by P. ERDŐS and K. L. CHUNG [20] in the special case of additive Markov chains. Let us consider an additive Markov chain, i. e. put $\xi_n = \xi_0 + \delta_1 + \delta_2 + \cdots + \delta_n$ where the δ_k's are independent equidistributed random variables which take only integer values, further suppose that the following conditions are satisfied:

Let us put $\mathbf{P}(\delta_k = r) = W_r$ $(r = 0, \pm 1, \pm 2, \ldots)$. We suppose that the greatest common divisor of the differences $r - s$ where r and s are such integers, for which $W_r > 0$ and $W_s > 0$, is equal to 1, further that $\sum_{r=-\infty}^{+\infty} |r| W_r < +\infty$ and $\sum_{r=-\infty}^{+\infty} r W_r = 0$. It has been shown by K. L. CHUNG and P. ERDŐS [20] that in this case the chain is (in our terminology) conditionally ergodic, and the limiting conditional distribution is the uniform distribution over the set of all integers, i. e. $P_k \equiv 1$. If the variance $D^2 = \sum_{r=-\infty}^{+\infty} r^2 W_r$ of the variables δ_k exists, this follows easily by LAPLACE's

method; as a matter of fact in this case, putting $\varphi(t) = \sum_{r=-\infty}^{+\infty} W_r e^{irt}$, we have

$$P_{jk}^{(n)} = \frac{1}{2\pi} \int_{-\pi}^{+\pi} (\varphi(t))^2 e^{it(j-k)}\, dt \tag{56}$$

and thus, by LAPLACE's method,

$$P_{ik}^{(n)} \sim \frac{1}{D\sqrt{2\pi n}} \tag{57}$$

what implies

$$\lim_{n \to \infty} \frac{P_{hj}^{(n)}}{P_{ik}^{(n)}} = 1. \tag{58}$$

If $D = +\infty$, the proof is more intricate.

The existence of the limits (52) in the general case is still open, as has been pointed out by CHUNG [21].

3. 8. On a paradox of the renewal theory. Renewal theory (see [22]) deals with stochastic processes of the following type ("recurrent processes"): events occur at the moments $0 < \tau_1 < \tau_2 < \cdots < \tau_n < \cdots$ where the differences $\tau_n - \tau_{n-1} = \xi_n$ are independent positive random variables with the same distribution function $F(x) = \mathbf{P}(\xi_n < x)$; for any $t > 0$ let $\tau_{n(t)}$ denote the moment when the first event after time t occurs, i. e. we suppose $\tau_{n(t)-1} \leqq t < \tau_{n(t)}$ and put $\vartheta(t) = \tau_{n(t)} - t$; $\vartheta(t)$ is the waiting time at t, i. e. the time somebody arriving at time t has to wait until the next event occurs. Let us now suppose that somebody arrives at random in the time interval $(0, T)$ at time τ, where τ is uniformly distributed in $(0, T)$ and let us put $\delta(T) = \vartheta(\tau)$. It has been proved by W. FELLER [22] that if $m = \int_0^\infty x\, dF(x)$ (the mean value of the time interval between consecutive events) exists, and $F(x)$ is not a lattice distribution function, the distribution of $\vartheta(T)$ tends for $T \to \infty$ to the limiting distribution

$$H(x) = \lim_{T \to \infty} \mathbf{P}(\vartheta(T) < x) = \frac{1}{m} \int_0^x (1 - F(u))\, du. \tag{59}$$

It has been shown by L. TAKÁCS [23] that if we consider $\delta(T)$ instead of $\vartheta(T)$, the limiting distribution exists for any $F(x)$ with finite mean value, and is the same as in the case of $\vartheta(T)$, i. e. we have

$$H(x) = \lim_{T \to \infty} \mathbf{P}(\delta(T) < x) = \frac{1}{m} \int_0^x (1 - F(u))\, du. \tag{60}$$

It has been pointed out by L. TAKÁCS [23] that the fact that the mean value

of the distribution function $H(x)$ is $\frac{m^2+\sigma^2}{2m}$ where σ^2 denotes the variance of ξ_n, i. e. $\sigma^2=\int_0^\infty (x-m)^2\,dF(x)$, implies that if $\sigma=+\infty$, the mean value of $\delta(T)$ tends to $+\infty$ in spite of the fact that $\delta(T)$ is the part of some time interval ξ_n having the finite mean value m.

The situation is still more paradox, if $m=\infty$; it can be shown (see [9]) that in this case

$$\lim_{T\to\infty} \mathbf{P}(\delta(T)<x)=0 \tag{61}$$

for all finite x; nevertheless, if putting $G(x)=\int_0^x (1-F(t))\,dt$ the condition

$$\lim_{x\to\infty} \frac{xG'(x)}{G(x)}=0 \tag{62}$$

is satisfied, there exists the limit of the conditional distribution of $\delta(T)$, and we have

$$\lim_{T\to\infty} \mathbf{P}(\delta(T)<x\,|\,\delta(T)<y)=\frac{G(x)}{G(y)} \quad \text{for} \quad 0<x\leqq y<+\infty. \tag{63}$$

3.9. Deduction of Maxwell's law of velocity distribution. Let us consider an ideal gas consisting of N particles with equal masses m, let ξ_k, η_k, ζ_k denote the x, y, z-component of the velocity of the k-th particle; we suppose that ξ_k, η_k, ζ_k $(k=1, 2, \ldots, N)$ are random variables defined on a conditional probability space $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}(A|B)]$. Let us suppose that the conditional probability distribution generated by the variables ξ_k, η_k, ζ_k $(k=1, 2, \ldots, N)$ in the $3N$-dimensional Euclidean space is uniform, further that the space is regular with respect to the sets defined by $\xi_k=x_k$, $\eta_k=y_k$, $\zeta_k=z_k$ $(-\infty<x_k<+\infty, -\infty<y_k<+\infty, -\infty<z_k<+\infty;\ k=1, 2, \ldots, N)$, further that the sets $\sum_{(t_1, \ldots, t_{3N})\in M} R_{t_1, \ldots, t_{3N}}$ belong to $\mathfrak{B}$, where $R_{t_1, \ldots, t_{3N}}$ is the set defined by $\xi_1=t_1$, $\eta_1=t_2$, $\zeta_1=t_3$, $\xi_2=t_4, \ldots, \zeta_n=t_{3N}$ and M is a Borel subset of the $3N$-dimensional Euclidean space, further that the space is also regular with respect to the sets defined by $\xi_k^2+\eta_k^2+\zeta_k^2=z_k$ $(k=1, 2, \ldots, N;\ 0<z_k<+\infty)$. All these conditions are satisfied, if S is the $3N$-dimensional Euclidean space, $\mathcal{A}$ the set of all Borel subsets of S, $\mathfrak{B}_k$ the set of all Borel subsets of S which have positive and finite k-dimensional measure $(k=1, 2, \ldots, 3N)$, $\mathfrak{B}_0$ the set of all points of S and $\mathfrak{B}=\sum_{k=0}^{3N}\mathfrak{B}_k$ and we put $\mathbf{P}(A|B)=1$ if $B\in\mathfrak{B}_0$ and $B\subseteq A$, $\mathbf{P}(A|B)=0$ if $B\in\mathfrak{B}_0$ and $B\subseteq\bar{A}$, further $\mathbf{P}(A|B)=\frac{m_k(AB)}{m_k(B)}$ if $B\in\mathfrak{B}_k$, where $m_k(C)$ is the k-dimensional measure of C $(k=1, 2, \ldots, 3N)$, further if denoting

by $P=(x_1, y_1, z_1, \ldots, x_N, y_N, z_N)$ a point of S, we put $\xi_1(P)=x_1$, $\eta_1(P)=y_1, \ldots, \zeta_N(P)=z_N$.

It follows that the kinetic energy $\varepsilon_k=\frac{1}{2}m(\xi_k^2+\eta_k^2+\zeta_k^2)$ of the k-th particle has the density function $\sqrt{x}$ for $0<x<+\infty$ and thus by 2.6 that the kinetic energy $\varepsilon=\sum_{k=1}^{N}\varepsilon_k$ of the whole gas has the density function $g_N(x)$ defined by the recursion

$$g_1(x)=\sqrt{x}, \quad g_k(x)=\int_0^x g_{k-1}(x-y)\sqrt{y}\,dy \quad \text{for} \quad x\geqq 0.$$

Thus

$$g_N(x)=\frac{\left(\frac{\sqrt{\pi}}{2}\right)^N}{\Gamma\left(\frac{3N}{2}\right)}x^{\frac{3N}{2}-1} \qquad (x\geqq 0),$$

and therefore, as any constant factor is irrelevant, we may take $x^{\frac{3N}{2}-1}$ $(x>0)$ for the density function of ε.

Similarly, the conditional density function of ε under condition $\varepsilon_k=t$ is clearly $(x-t)^{\frac{3N-5}{2}}$ for $x\geqq t$. It can be easily verified that the conditions for applying formula (48) are fulfilled, and thus we obtain that the conditional density function of ε_k with respect to the condition $\varepsilon=E$ is

$$f(x|E)=\frac{(E-x)^{\frac{3N-5}{2}}\sqrt{x}}{\int_0^E (E-x)^{\frac{3N-5}{2}}\sqrt{x}\,dx}=\frac{\left(1-\frac{x}{E}\right)^{\frac{3N-5}{2}}\sqrt{x}}{\int_0^E\left(1-\frac{x}{E}\right)^{\frac{3N-5}{2}}\sqrt{x}\,dx}.$$

If N is a large number, we have, putting $\beta=\frac{3N}{2E}$, approximately $\left(1-\frac{x}{E}\right)^{\frac{3N-5}{2}}\sim e^{-\beta x}$ and thus

$$f(x|E)\sim\frac{2\beta^{3/2}}{\sqrt{\pi}}\sqrt{x}\cdot e^{-\beta x}.$$

Thus the conditional density function $g(x|E)=mxf\left(\frac{m}{2}x^2|E\right)$ of the velocity $v_k=\sqrt{\xi_k^2+\eta_k^2+\zeta_k^2}$ of the k-th particle under condition $\varepsilon=E$ is approximately

$$g(x|E)\sim\sqrt{\frac{2}{\pi}}\frac{1}{\sigma^{3/2}}x^2e^{-\frac{x^2}{2\sigma^2}}$$

where $\sigma=\frac{1}{\sqrt{m\beta}}$. As it is known that $\frac{E}{N}=\frac{3}{2}kT$ where T denotes the absolute temperature of the gas and k the constant of Boltzmann, we have $\beta=\frac{1}{kT}$ and $\sigma=\sqrt{\frac{kT}{m}}$. Thus finally we obtain

$$g(x|E)\sim\sqrt{\frac{2}{\pi}}\left(\frac{m}{kT}\right)^{\frac{3}{2}}x^2e^{-\frac{mx^2}{2kT}}\qquad(0<x<+\infty),\tag{64}$$

i. e. Maxwell's law of velocity distribution.

Summarising, we may state our result as follows: if the components of the velocity of each of N particles are independent of each other and of the components of the velocity of the other particles, further if all these random variables have regular uniform (a priori) distributions on the whole real axis, then the (a posteriori) distribution of the velocity of each particle under the condition that the kinetic energy of the whole system is given, is approximately, if N is large, the Maxwell distribution (64).

The proof given above shows clearly that the law of Maxwell is an approximation, which is valid only for a large number of particles. For a small number N we obtain the exact formula

$$g_N(x|E)=\sqrt{\frac{2}{\pi}}\left(\frac{m}{E}\right)^{3/2}\frac{\Gamma\left(\frac{3N}{2}\right)}{\Gamma\left(\frac{3N-3}{2}\right)}x^2\left(1-\frac{mx^2}{2E}\right)^{\frac{3N-5}{2}}.\tag{65}$$

It can also be shown by the same way that the conditional density function of each of the variables ξ_k, η_k, ζ_k under the condition $\varepsilon=E$ is exactly

$$h(x|E)=\frac{1}{\sqrt{2\pi}}\sqrt{\frac{m}{E}}\cdot\frac{\Gamma\left(\frac{3N}{2}\right)}{\Gamma\left(\frac{3N-1}{2}\right)}\left(1-\frac{mx^2}{2E}\right)^{\frac{3N-3}{2}}\sim\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}\tag{66}$$

where $\sigma=\sqrt{\frac{kT}{m}}$. Thus under the condition $\varepsilon=E$ all these variables are approximately normally distributed. This is the reason why the velocities have a Maxwell distribution, as the Maxwell distribution can be defined as the distribution of a variable $\sqrt{\xi^2+\eta^2+\zeta^2}$ where ξ, η, ζ are independent and normally distributed random variables with mean value 0 and variance σ^2.

§ 4. Conditional laws of large numbers

Conditional probability is in the same relation to conditional relative frequency as ordinary probability to ordinary relative frequency. This relation, which is well known as an empirical fact from everyday experience, is described mathematically by the laws of large numbers.

The laws of large numbers concerning the behaviour in the limit of the conditional relative frequency (and generalizations concerning conditional means of observations) shall be called "conditional laws of large numbers". These laws will be investigated in this §. Before going into details it should be emphasized that the conditional probability $\mathbf{P}(A|B)$ is considered in this paper as an objective characteristic of the random event A, under an objective condition B, and its value as that number in the near neighbourhood of which the corresponding conditional relative frequency $\frac{k_{AB}}{k_B}$ will be found in general, if a sufficiently great number n of observations is made, where k_B denotes the number of those under the n observations with respect to which the condition B has been realised, and k_{AB} the number of those observations which gave such results that, besides the condition B being realised, the event A occurred also, and it is supposed that $k_B > 0$.

4. 1. A strong law of large numbers. In this section we consider random variables defined on an ordinary probability space. We shall need the following

LEMMA 1. *Let* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *be mutually independent random variables with mean values* $\mathbf{M}(\xi_n) = M_n \geqq 0$ *and finite variances* $D_n^2 = \mathbf{D}^2(\xi_n)$. *Let us put* $\zeta_n = \sum_{k=1}^{n} \xi_k$ *and* $A_n = \mathbf{M}(\zeta_n) = M_1 + M_2 + \cdots + M_n$, *and suppose that the following conditions are fulfilled:*

a) $\lim_{n\to\infty} A_n = +\infty$,

b) $\sum_{n=1}^{\infty} \frac{D_n^2}{A_n^2} \leqq +\infty$.

It follows that

$$\mathbf{P}\left(\lim_{n\to\infty} \frac{\zeta_n}{A_n} = 1\right) = 1. \tag{67}$$

PROOF. Lemma 1 (see e. g. [27], p. 238) is a consequence of the well-known theorem of KOLMOGOROV, according to which if $\eta_1, \eta_2, \ldots, \eta_n, \ldots$ are mutually independent random variables and $\mathbf{M}(\eta_k) = 0$ $(k = 1, 2, \ldots)$, then $\sum_{k=1}^{\infty} \eta_k$ converges with probability 1, provided that the series $\sum \mathbf{D}^2(\eta_k)$ con-

verges. Lemma 1 can be proved also directly by applying the following inequality which has been found recently by J. HÁJEK (see [12]):

LEMMA 2. *Let $\eta_1, \eta_2, \ldots, \eta_n, \ldots$ denote mutually independent random variables with mean values $\mathbf{M}(\eta_k)=0$ and variances $\mathbf{D}^2(\eta_k)=D_k^2$ $(k=1,2,\ldots)$. Let c_k denote a decreasing sequence of positive numbers, $c_k \leqq c_{k+1}$ $(k \geqq n)$; then we have for any $\varepsilon>0$*

$$\mathbf{P}(\max_{n \leqq k} c_k |\eta_1+\eta_2+\cdots+\eta_k| \geqq \varepsilon) \leqq \frac{1}{\varepsilon^2}\left(c_n^2 \sum_{k=1}^{n} D_k^2 + \sum_{k=n+1}^{\infty} c_k^2 D_k^2\right). \tag{68}$$

Applying Lemma 2, Lemma 1 can be proved as follows: it follows from (68), applied for $\eta_k=\xi_k-M_k$, $c_k=\frac{1}{A_k}$ that

$$\mathbf{P}\left(\sup_{n \leqq k}\left|\frac{\zeta_k}{A_k}-1\right| \geqq \varepsilon\right) \leqq \frac{1}{\varepsilon^2}\left(\frac{\sum_{k=1}^{n} D_k^2}{A_n^2} + \sum_{k=n+1}^{\infty} \frac{D_k^2}{A_k^2}\right). \tag{69}$$

As a) and b) clearly imply that

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} D_k^2}{A_n^2} = 0, \tag{70}$$

it follows from (69) that

$$\lim_{n\to\infty} \mathbf{P}\left(\sup_{n \leqq k}\left|\frac{\zeta_k}{A_k}-1\right| \geqq \varepsilon\right) = 0 \quad \text{for any} \quad \varepsilon>0. \tag{71}$$

and (71) is evidently equivalent to (67).

4.2. A conditional law of large numbers.

THEOREM 15. *Let $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}(A|B)]$ denote a conditional probability space and $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ random variables on S which are mutually independent with respect to $C \in \mathfrak{B}$. Let $\mathfrak{I}$ denote the interval $a \leqq x < b$ $(a<b)$ of the real axis. Let B_n denote the set of those $a \in S$ for which $\xi_n(a) \in \mathfrak{I}$, and suppose that $B_n \subseteq C$ and $B_n \in \mathfrak{B}$; let us suppose that $\mathbf{M}(\xi_n|B_n)=M_n>0$ and $\mathbf{D}^2(\xi_n|B_n)=D_n^2$ exists $(n=1,2,\ldots)$. Let us put $\mathbf{P}(B_n|C)=p_n$ and suppose that the following conditions are satisfied:*

$$\sum_{n=1}^{\infty} p_n = +\infty \quad \text{and} \quad \sum_{n=1}^{\infty} p_n M_n = +\infty; \tag{72}$$

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} p_k M_k}{\sum_{k=1}^{n} p_k} = M \tag{73}$$

exists;

$$\sum_{k=1}^{\infty} \frac{p_k(D_k^2+(1-p_k)M_k^2)}{\left(\sum_{j=1}^{k} p_j M_j\right)^2} < +\infty. \tag{74}$$

Then we have

$$\mathbf{P}\left(\lim_{n\to\infty} \frac{\sum\limits_{\substack{\xi_k \in \mathfrak{J} \\ 1\leq k\leq n}} \xi_k}{\sum\limits_{\substack{\xi_k \in \mathfrak{J} \\ 1\leq k\leq n}} 1} = M \,\middle|\, C \right) = 1. \tag{75}$$

PROOF. Let us define the random variable ε_k as follows: $\varepsilon_k = 1$ if $\xi_k \in \mathfrak{J}$ and $\varepsilon_k = 0$ if $\xi_k \notin \mathfrak{J}$; let us put $\xi_k^* = \xi_k \varepsilon_k$. Then we have

$$\mathbf{M}(\xi_k^* | C) = p_k M_k \quad \text{and} \quad \mathbf{M}(\xi_k^{*2} | C) = p_k(D_k^2 + M_k^2)$$

and thus

$$\mathbf{D}^2(\xi_k^* | C) = p_k(D_k^2 + (1-p_k)M_k^2).$$

Applying Lemma 1 to the sequence ξ_k^* of random variables on C, it follows that

$$\mathbf{P}\left(\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \xi_k^*}{\sum_{k=1}^{n} p_k M_k} = 1 \,\middle|\, C \right) = 1. \tag{76}$$

On the other hand, let us apply Lemma 1 to the sequence of random variables ε_k on C. As

$$\mathbf{M}(\varepsilon_k | C) = p_k \quad \text{and} \quad \mathbf{D}^2(\varepsilon_k | C) = p_k(1-p_k),$$

it follows that

$$\mathbf{P}\left(\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \varepsilon_k}{\sum_{k=1}^{n} p_k} = 1 \,\middle|\, C \right) = 1. \tag{77}$$

Combining (76), (77) and condition b) of the theorem, and taking into account that

$$\sum_{\substack{\xi_k \in \mathfrak{J} \\ 1\leq k\leq n}} \xi_k = \sum_{k=1}^{n} \xi_k^* \quad \text{and} \quad \sum_{\substack{\xi_k \in \mathfrak{J} \\ 1\leq k\leq n}} 1 = \sum_{k=1}^{n} \varepsilon_k,$$

(75) follows. Thus Theorem 15 is proved.

The statement of Theorem 15 can be expressed in words as follows: the conditional empirical mean value of those of the variables $\xi_1, \xi_2, \ldots, \xi_n$ which take on values lying in the interval $\mathfrak{J}$, converges with conditional probability 1 with respect to C to the limit M defined by (73). In the special

case, when $M_n = M > 0$ and $D_n = D > 0$ do not depend on n, the conditions (72)—(74) reduce to the single condition that the series $\sum_{n=1}^{\infty} p_n$ diverges. Especially let us suppose further that $\mathfrak{J}$ is the closed interval [0, 1] and the only values in $\mathfrak{J}$ which the variables ξ_n can take on B_n are the values 0 and 1 (of course ξ_n can take also other values outside $\mathfrak{J}$). Suppose that A_n is the set on which $\xi_n = 1$, and $p = \mathbf{P}(A_n | B_n) > 0$ does not depend on n.

In this case the events A_n and B_n can be interpreted as the events consisting in the realisation of some events A and B, respectively, at the n-th experiment in a sequence of independent experiments, and we may put $p = \mathbf{P}(A|B)$; in this case

$$f_n(A|B) = \frac{\sum_{\substack{\xi_k \in \mathfrak{J} \\ 1 \leqq k \leqq n}} \xi_k}{\sum_{\substack{\xi_k \in \mathfrak{J} \\ 1 \leqq k \leqq n}} 1} = \frac{\sum_{\xi_k = 1,\ 1 \leqq k \leqq n} 1}{\sum_{\xi_k = 0 \text{ or } 1,\ 1 \leqq k \leqq n} 1}$$

is the conditional relative frequency of the event A with respect to the event B in course of the first n observations.[12] The statement of Theorem 15 gives for this special case

$$\mathbf{P}(\lim_{n \to \infty} f_n(A|B) = \mathbf{P}(A|B) | C) = 1, \tag{78}$$

if $\mathbf{P}(B|C) > 0$, i. e. *the conditional relative frequency of the event A with respect to the event B converges with the conditional probability* 1 *(with respect to C) to the conditional probability of A with respect to B.*

4. 3. Generalization of the theorem of Borel on normal decimals. The theorem of BOREL [25] states that if $f_n(k; x)$ denotes the relative frequency of the number k between the first n digits of the decimal expansion of the real number x $(0 \leqq x < 1)$, we have

$$\lim_{n \to \infty} f_n(k; x) = \frac{1}{10} \qquad (k = 0, 1, 2, \ldots, 9) \tag{79}$$

for almost every x; similarly all possible digits occur in the limit with equal frequency in the expansion, in the number system with any basis q of almost all real numbers.

We consider a more general representation of real numbers; namely expansion of real numbers into Cantor's series. Let q_n $(n = 1, 2, \ldots)$ denote an arbitrary sequence of integers, $q_n \geqq 2$. It is easy to prove (see [24]) that

[12] This interpretation suggests itself in the following special case: let us form the Cartesian product of a denumerable sequence of isomorphic conditional probability spaces $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ and denote by A_n resp. B_n the sets defined by $\overset{1}{S} * \overset{2}{S} * \cdots * \overset{n}{A} * S * \cdots$ resp. $\overset{1}{S} * \overset{2}{S} * \cdots * \overset{n}{B} * S * \cdots$ (see section 1. 11).

any real number x $(0 \leqq x < 1)$ can be represented in the form

$$x = \sum_{n=1}^{\infty} \frac{\varepsilon_n(x)}{q_1 q_2 \cdots q_n},$$

where $\varepsilon_n(x)$ may take on the values $0, 1, \ldots, q_n - 1$; the digits $\varepsilon_n(x)$ are uniquely determined except for those rational numbers x which can be written in the form $x = \sum_{n=1}^{N} \frac{\varepsilon_n(x)}{q_1 q_2 \cdots q_n}$ $(\varepsilon_N(x) > 0)$, where besides this finite representation the alternative infinite representation

$$x = \sum_{n=1}^{N-1} \frac{\varepsilon_n(x)}{q_1 q_2 \ldots q_n} + \frac{\varepsilon_N(x) - 1}{q_1 q_2 \ldots q_N} + \sum_{n=N+1}^{\infty} \frac{q_n - 1}{q_1 q_2 \ldots q_n}$$

is also possible. We shall prove that in the case $q_n \to +\infty$, when all non-negative integers are possible "digits", all these digits occur in the limit with the same (conditional) relative frequency, provided that $\sum_{n=1}^{\infty} \frac{1}{q_n}$ diverges. This is contained in the following theorem which is a consequence of Theorem 15.

THEOREM 16. *Let the numbers q_n satisfy the following conditions:*

a) *q_n is an integer, $q_n \geqq 2$ for $n = 1, 2, \ldots$,*

b) *$q_n \geqq N$ for $n \geqq n_0$,*

c) $\sum_{n=1}^{\infty} \frac{1}{q_n} = +\infty$.

Let us consider the expansion of real numbers x $(0 \leqq x < 1)$ into Cantor's series with quotients q_n, i. e. the expansion

$$x = \sum_{n=1}^{\infty} \frac{\varepsilon_n(x)}{q_1 q_2 \ldots q_n} \tag{80}$$

where $\varepsilon_n(x)$ is one of the numbers $0, 1, \ldots, q_n - 1$.

Let $F_n(k; x)$ denote how many of the "digits" $\varepsilon_1(x), \varepsilon_2(x), \ldots, \varepsilon_n(x)$ are equal to k. Then for almost every x in $(0, 1)$ we have

$$\lim_{n \to \infty} \frac{F_n(j; x)}{F_n(k; x)} = 1 \quad \text{for} \quad 0 \leqq j \leqq k \leqq N - 1. \tag{81}$$

REMARK 1. If $q_n \to \infty$, condition b) is fulfilled for every N, and thus (81) holds for any pair of non-negative integers j and k, i. e. *all digits $0, 1, 2, \ldots$ occur in the limit asymptotically with the same frequency, in the Cantor expansion of almost all real numbers.* This is the case, for instance, if $q_n = n + 1$, i. e. if we consider the expansion

$$x = \sum_{k=2}^{\infty} \frac{\varepsilon_k(x)}{k!} \tag{82}$$

where $\varepsilon_k(x)$ can have the values $0, 1, \ldots, k - 1$.

REMARK 2. It follows from (81) that for almost all x

$$\lim_{M\to\infty} \frac{F_M(k;x)}{\sum_{j=0}^{l} F_M(j;x)} = \frac{1}{l+1} \quad \text{for} \quad k=0,1,\ldots,l \quad \text{and} \quad l=1,2,\ldots,N-1. \tag{83}$$

This is an other way of expressing that the frequencies of all digits $0, 1, \ldots, N-1$ are in the limit equal to another.

If $q_n \to \infty$, we have

$$\lim_{M\to\infty} \frac{F_M(k;x)}{M} = 0 \quad \text{for} \quad k=0,1,\ldots \tag{84}$$

for almost all x. This can be shown to be a consequence of Lemma 1 (see (86) below).

REMARK 3. Theorem 16 clearly reduces to the theorem of Borel if $q_n = 10$ for $n = 1, 2, \ldots$.

PROOF. Theorem 16 is a special case of Theorem 15, but it is more simple to prove it directly by means of Lemma 1 of section 4. 1. Let us consider the ordinary probability space $[S, \mathcal{A}, \mathbf{P}(A)]$ which we obtain if we choose for S the interval $(0, 1)$, for $\mathcal{A}$ the set of all measurable subsets of S and put $P(A) = m(A)$ where $m(A)$ is the Lebesgue measure of the set A. Let us define the random variables $\xi_{nk} = \xi_{nk}(x)$ $(0 \leqq x < 1)$ for $0 \leqq k \leqq n-1$ $(n = 1, 2, \ldots)$ as follows:

$$\xi_{nk}(x) = \begin{cases} 1 & \text{if} \quad \varepsilon_n(x) = k, \\ 0 & \text{otherwise.} \end{cases} \tag{85}$$

The variables ξ_{nk} $(n = 1, 2, \ldots)$ are clearly mutually independent, further we have $M_{nk} = \mathbf{M}(\xi_{nk}(x)) = \frac{1}{q_n}$ and $D_{nk}^2 = \mathbf{D}^2(\xi_{nk}(x)) = \frac{1}{q_n}\left(1 - \frac{1}{q_n}\right)$ for $q_n \geqq k$, i. e. for $n \geqq n_0$ if $k < N$; the values of M_{nk} resp. D_{nk} for $n < n_0$ are irrelevant. Condition a) of Lemma 1 follows from supposition b) of Theorem 16, and condition b) of Lemma 1 follows also from supposition b) of Theorem 16. Thus we have

$$\mathbf{P}\left(\lim_{M\to\infty} \frac{\sum_{n=1}^{M} \xi_{nk}}{\sum_{n=n_0}^{M} \frac{1}{q_n}} = 1 \right) = 1. \tag{86}$$

As (86) holds for all $k = 0, 1, \ldots, N-1$ and $\sum_{n=1}^{M} \xi_{nk} = F_M(k;x)$ the assertion of Theorem 16 follows.

The results of this section (as well as those of section 3. 2) show that there exist sequences $x_1, x_2, \ldots, x_n$, consisting of the numbers $0, 1, 2, \ldots$, in which the relative frequency $\frac{F_n(k)}{n}$ (where $F_n(k)$ denotes how many of the

numbers $x_1, x_2, \ldots, x_n$ are equal to k) of every number $k = 0, 1, \ldots$ tends to 0 for $n \to \infty$, but the conditional relative frequencies converge to definite limits, i. e.

$$\lim_{n\to\infty} \frac{F_n(h)}{\sum_{j=0}^{k} F_n(j)} = \frac{P_h}{\sum_{j=0}^{k} P_j} \qquad (h = 0, 1, \ldots, k;\ k = 1, 2, \ldots)$$

where $P_k > 0$ for $k = 0, 1, \ldots$ and $\sum_{k=0}^{\infty} P_k = +\infty$. Such sequences may be considered as mathematical models of sequences of observations on a random variable ξ, defined on a conditional probability field, and having the conditional distribution

$$\mathbf{P}(\xi = h \mid 0 \leq \xi \leq k) = \frac{P_h}{\sum_{j=0}^{k} P_j} \qquad (h = 0, 1, \ldots, k;\ k = 1, 2, \ldots).$$

(Received 25 July 1955)

Bibliography

[1] A. N. KOLMOGOROFF, *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Berlin, 1933).

[2] H. JEFFREYS, *Theory of probability* (London, 1943).

[3] H. REICHENBACH, *Wahrscheinlichkeitslehre* (Leiden, 1935).

[4] M. I. KEYNES, *A treatise on probability theory.*

[5] B. O. KOOPMAN, The axioms and algebra of intuitive probability, *Annals of Math.*, **41** (1940), pp. 269—292.

[6] A. H. COPELAND, Postulates for the theory of probability, *Amer. Journal of Math.*, **63** (1941), pp. 741—762.

[7] G. A. BARNARD, Statistical inference, *Journal of the Royal Statistical Soc.*, Ser. B, **11** (1949), pp. 115—139.

[8] I. J. GOOD, *Probability and the weighing of evidence* (London, 1950).

[9] A. RÉNYI, A valószínűségszámítás új axiomatikus felépítése, *MTA Mat. és Fiz. Oszt. Közl.*, **4** (1954), pp. 369—427.

[10] A. RÉNYI, On a new axiomatic foundation of the theory of probability. Under press in the Volume 1 of the *Proceedings of the International Mathematical Congress in Amsterdam, 1954.*

[11] A. RÉNYI, Über die axiomatische Begründung der Wahrscheinlichkeitsrechnung. Under press in the *Proceedings of the Conference on Probability Theory in Berlin, 1954.*

[12] J. HÁJEK and A. RÉNYI, Generalization of an inequality of Kolmogorov, *Acta Math. Acad. Sci. Hung.*, **6** (1955), pp. 281—283.

[13] Á. CSÁSZÁR, Sur la structure des espaces de probabilité conditionnelle, *Acta Math. Acad. Sci. Hung.*, **6** (1955), pp. 337—361.

[14] J. NEYMAN, L'estimation statistique traitée comme un problème classique de probabilité, *Act. Sci. et Ind.*, 739 (Paris, 1938), pp. 25—57.

[15] P. HALMOS, *Measure theory* (New-York, 1951).

[16] P. Erdös, On the integers having exactly k prime factors, *Annals of Math.*, **49** (1948), pp. 53—66.
[17] B. V. Gnedenko, Über ein lokales Grenzwerttheorem für gleichverteilte unabhängige Summanden, *Wiss. Zeitschrift der Humboldt-Universität Berlin*, **3** (1954), pp. 287—293.
[18]. J. L. Doob, *Stochastic processes* (New-York, 1953).
[19] C. Derman, A solution to a set of fundamental equations in Markov chains, *Proc. Amer. Math. Soc.*, **5** (1954), pp. 332—334.
[20] Erdös P. and K. L. Chung, Probability limit theorems assuming only the first moment. I, *Memoirs of the Amer. Math. Soc.*, **6** (1952).
[21] K. L. Chung, Contributions to the theory of Markov chains, *Trans. Amer. Math. Soc.*, **76** (1954), pp. 397—419. See also K. L. Chung, An ergodic theorem for stationary Markov chains with a countable number of states, *Proc. Internat. Congress of Math. Cambridge (USA), 1950*, (Amer. Math. Soc., 1952), I, p. 568.
[22] W. Feller, On the integral equation of renewal theory, *Ann. Math. Stat.*, **12** (1941), pp. 243—267. See also S. Täcklind, Elementare Behandlung vom Erneuerungsproblem, *Skandinavisk Aktuarietidskrift*, (1944), pp. 1—15, and Fourier-analytische Behandlung vom Erneuerungsproblem, *Skandinavisk Aktuarietidskrift*, (1945), pp. 101—105; and J. L. Doob, Renewal theory from the point of view of the theory of probability, *Trans. Amer. Math. Soc.*, **63** (1948), pp. 422—438.
[23] L. Takács, Egy új módszer rekurrens sztochasztikus folyamatok tárgyalásánál, *MTA Alk. Mat. Int. Közleményei*, **2** (1953), pp. 135—163.
[24] O. Perron, *Irrationalzahlen* (Berlin, 1921), pp. 111—116.
[25] E. Borel, Les probabilités dénombrables et leurs applications arithmétiques, *Rend. Circ. Math. Palermo*, **27** (1909), pp. 247—271.
[26] H. Hahn and A. Rosenthal, Set functions (New Mexico, 1948).
[27] M. Loève, *Probability theory* (New-York, 1955).

НОВОЕ АКСИОМАТИЧЕСКОЕ ПОСТРОЕНИЕ ТЕОРИИ ВЕРОЯТНОСТЕЙ

А. Реньи (Будапешт)

(Резюме)

Работа содержит новое аксиоматическое построение теории вероятностей; основное понятие в этой новой теории, являющейся обобщением теории А. Н. Колмогорова, — понятие условной вероятности. Потребность в разработке новой теории была вызвана тем, что в теории А. Н. Колмогорова неограниченные меры исключены, так, например, не имеет смысла говорить о равномерном распределении, во всем n-мерном евклидовом пространстве, в то время, как приложения теории вероятности к физике, к интегральной геометрии, теории чисел и т. д. требуют рассмотрения таких не нормируемых распределений. Выбирая в качестве основного понятия теории вероятностей понятие условной вероятности, упомянутым не нормируемым распределениям можно дать точный математический смысл. Новая теория делает возможным обобщение целого ряда теорем теории вероятностей и новые приложения теории; работа занимается разработкой новой теории и знакомит с некоторыми ее типичными приложениями. Теория исходит из следующих предположений и аксиом:

Пусть S есть любое множество, которое мы будем называть пространством событий; пусть $\mathcal{A}$ есть σ-алгебра подмножеств пространства S, элементы $\mathcal{A}$ мы будем

называть событиями; пусть $\mathfrak{B}$ есть некоторое не пустое подмножество от $\mathfrak{A}$ ($\mathfrak{B}$ есть множество всех допустимых условии), а $\mathbf{P}(A|B)$ функция множеств от двух переменных, определенная, если $A \in \mathfrak{A}$ и $B \in \mathfrak{B}$; число $\mathbf{P}(A|B)$ мы будем называть условной вероятностью события A относительно условия B. Предположим, что выполняются следующие 3 аксиомы:

Аксиома I. $\mathbf{P}(A|B) \geqq 0$, если $A \in \mathfrak{A}$ и $B \in \mathfrak{B}$, и $\mathbf{P}(B|B) = 1$, если $B \in \mathfrak{B}$.

Аксиома II. Если $B \in \mathfrak{B}$ фиксирован, $\mathbf{P}(A|B)$ есть мера на σ-алгебре $\mathfrak{A}$.

Аксиома III. Если $A \in \mathfrak{A}$, $B \in \mathfrak{A}$, $C \in \mathfrak{B}$ и $BC \in \mathfrak{B}$, то

$$\mathbf{P}(A|BC)\,\mathbf{P}(B|C) = \mathbf{P}(AB|C).$$

Если аксиомы I—III выполненны, то совокупность множества S, σ-алгебры $\mathfrak{A}$ системы множеств $\mathfrak{B}$ и функции множеств $\mathbf{P}(A|B)$ назовем пространством условной вероятности и обозначим через $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$.

Очевидно, что если B фиксированно, то $S, \mathfrak{A}$ и $\mathbf{P}(A|B)$ образуют по теории А. Н. Колмогорова пространство вероятности. Поэтому пространство условной вероятности является ни чем иным, как совокупностью обыкновенных пространств вероятности, связанных аксиомой III.

В качестве непосредственного следствия аксиом получается равенство $\mathbf{P}(A|B) = \mathbf{P}(AB|B)$, а поэтому $\mathbf{P}(A|B) \leqq 1$. Раздел 1.4 занимается дальнейшими простыми следствиями аксиом. В разделе 1.5 исследуется при каких условиях условная вероятность $\mathbf{P}(A|B)$ может быть представлена в виде $\mathbf{P}(A|B) = \dfrac{\mathbf{Q}(AB)}{\mathbf{Q}(B)}$, где $\mathbf{Q}$ мера (не обязательно ограниченная) на σ-алгебре $\mathfrak{A}$ и $\mathbf{Q}(B) > 0$ для $B \in \mathfrak{B}$. Исчерпывающее исследование этого вопроса, а также следующего более общего вопроса: в каком случае условная вероятность $\mathbf{P}(A|B)$ может быть представлена в виде $\mathbf{P}(A|B) = \dfrac{\mathbf{Q}_B(AB)}{\mathbf{Q}_B(B)}$, где мера $\mathbf{Q}_B$ выбирается из некоторого множества мер в зависимости от B, можно найти в работе А. Чацара [13]. В разделе 1.6 исследуются случайные величины, определенные в пространстве условной вероятности, и содержит определение их функции распределения и функции плотности.

Функцию $\xi = \xi(s)$ $(s \in S)$ мы называем случайной величиной, если она измерима относительно $\mathfrak{A}$. Монотонно неубывающую и слева непрерывную функцию $F(x)$ мы называем (обобщенной) функцией распределения случайной величины ξ, если множество, где $a \leqq \xi(s) < b$, принадлежит $\mathfrak{B}$, предполагая, что $F(b) - F(a) > 0$, и в этом случае

$$\mathbf{P}(c \leqq \xi < d \,|\, a \leqq \xi < b) = \frac{F(d) - F(c)}{F(b) - F(a)}, \text{ если } a \leqq c < d \leqq b.$$

Функция распределения $F(x)$ не определена однозначно, потому что вместе с $F(x)$ определению удовлетворяет и $\lambda F(x) + \mu$, где $\lambda > 0$ и μ вещественно; функция распределения $F(x)$ может принимать любые вещественные (значит и отрицательные) значения. Если $F(x)$ абсолютно непрерывна, функцию $f(x) = F'(x)$ назовем обобщенной функцией плотности случайной величины ξ; функция плотности определена лишь с точностью до положительного постоянного сомножителя и $\int_{-\infty}^{+\infty} f(x)\,dx$ не обязательно существует. Если, в частности, $f(x) \equiv 1$ $(-\infty < x < +\infty)$, то мы говорим, что случайная величина ξ равномерно распределена на всей вещественной оси. Раздел 1.7 занимается одной переформулировкой аксиомы III. В разделе 1.8 рассматривается расширение пространств условной вероятности, в 1.9 — непрерывность условной вероятности, в 1.10 — определение произведения пространств условной вероятности. Разделы 2.1 и

2.2 содержат примеры построения пространств условной вероятности, в разделах 2.3 и 2.4 изучаются пространства условной вероятности, обладающие некоторыми дополнительными свойствами: так называемые пространства Кавальери и регулярные пространства. Раздел 2.5 рассматривает парадокс Б о р е л я, 2.6 — распределение суммы независимых случайных величин и композицию введенных в разделе 1.6 обобщенных функций распределений. В § 3 на нескольких примерах показывается, как новая теория приводит к открытию некоторых новых соотношений, которые в рамках обычной теории вероятности не могут быть сформулированны. Раздел 3.6 содержит обобщение метода Б э й с а. Раздел 3.7 содержит определение условной эргодичности относительно цепей Маркова. В разделе 3.8 рассматривается один парадокс теории возобновления. Раздел 3.9 содержит простое и исходящее из более простых чем обычно предположений доказательство закона Максвелла о распределении скоростей, используя данное в разделе 3.6 обобщение метода Б э й с а. В § 4 речь идет об условных законах больших чисел, относящихся к сходимости с вероятностью 1 условного среднего наблюдений. В качестве приложения раздел 4.3 содержит обобщение теоремы Б о р е л я, относящейся к нормальным разложениям в десятичные дроби, на ряды Кантора.

SUR LA STRUCTURE DES ESPACES DE PROBABILITÉ CONDITIONNELLE

Par
ÁKOS CSÁSZÁR (Budapest)
(*Présenté par* A. RÉNYI)

1. Comme généralisation de la notion d'espace de probabilité, notion due à A. KOLMOGOROFF, M. A. RÉNYI vient de définir les espaces de probabilité conditionnelle [1, 2] par les axiomes suivants:

Soit S un ensemble donné, soit $\mathfrak{A}$ un σ-corps des sous-ensembles de S qui contient l'ensemble S (c'est-à-dire une famille des sous-ensembles de S qui contient S et, en outre, la différence de deux ensembles quelconques et la réunion d'une suite dénombrable quelconque d'ensembles qui appartiennent à elle), et soit $\mathfrak{B}$ une famille non vide qui fait partie de $\mathfrak{A}$. Supposons qu'à tout couple d'ensembles A, B, où $A \in \mathfrak{A}$ et $B \in \mathfrak{B}$, correspond un nombre réel $\mathbf{P}(A|B)$ et que les conditions suivantes sont remplies:

I. $\mathbf{P}(A|B) \geqq 0$ *et* $\mathbf{P}(B|B) = 1$ *pour* $A \in \mathfrak{A}$, $B \in \mathfrak{B}$.

II. *Pour* $B \in \mathfrak{B}$ *fixé,* $\mathbf{P}(A|B)$ *est une fonction* σ*-additive de l'ensemble* $A \in \mathfrak{A}$, *c'est-à-dire que l'on a pour* $A = \sum_{k=1}^{\infty} A_k$, $A_k \in \mathfrak{A}$ $(k = 1, 2, \ldots)$, $A_i A_k = 0$ $(i \neq k)$

$$\mathbf{P}(A|B) = \sum_{k=1}^{\infty} \mathbf{P}(A_k|B).$$

III. *On a pour* $A \in \mathfrak{A}$, $B \in \mathfrak{A}$, $C \in \mathfrak{B}$, $BC \in \mathfrak{B}$

$$\mathbf{P}(A|BC)\,\mathbf{P}(B|C) = \mathbf{P}(AB|C). \tag{1.1}$$

Il est aisé de constater qu'on arrive à un espace de probabilité conditionnelle en considérant une mesure μ définie sur $\mathfrak{A}$ (c'est-à-dire une fonction non-négative et σ-additive de l'ensemble $A \in \mathfrak{A}$, qui peut aussi prendre des valeurs infinies) telle que tout ensemble $B \in \mathfrak{B}$ ait une mesure positive et finie et en posant

$$\mathbf{P}(A|B) = \frac{\mu(AB)}{\mu(B)}.$$

De façon plus générale, soit $\{\mu_\alpha\}$ une famille de mesures définies sur $\mathfrak{A}$ et dont les indices forment un ensemble ordonné, et supposons que, pour $\alpha < \beta$, $\mu_\alpha(A) < +\infty$ implique $\mu_\beta(A) = 0$. S'il correspond à tout ensemble $B \in \mathfrak{B}$ une mesure μ_α telle que $0 < \mu_\alpha(B) < +\infty$ (il ne peut y avoir qu'une seule mesure

de ce type) et si l'on pose

$$\mathbf{P}(A|B) = \frac{\mu_\alpha(AB)}{\mu_\alpha(B)},$$

on arrive à un espace de probabilité conditionnelle (pour la démonstration, voir [1], p. 381 ou [2], p. 305).

En considérant les exemples mentionnés, la question se pose, quelles sont les conditions qu'un espace de probabilité conditionnelle doit remplir pour qu'il possède une structure plus ou moins semblable à celle des exemples en question. Pour pouvoir formuler d'une façon précise ces questions, convenons de la terminologie suivante: nous dirons qu'une famille $\{\mu_\alpha\}$ de mesures *engendre* un espace de probabilité conditionnelle donné, si a) chacune des mesures μ_α est définie sur $\mathcal{A}$, b) à tout ensemble $B \in \mathcal{B}$ correspond au moins un indice α tel que $0 < \mu_\alpha(B) < +\infty$, c) on a pour $A \in \mathcal{A}$, $B \in \mathcal{B}$,

$$0 < \mu_\alpha(B) < +\infty \tag{1.2}$$

l'égalité

$$\mathbf{P}(A|B) = \frac{\mu_\alpha(AB)}{\mu_\alpha(B)} \tag{1.3}$$

(l'inégalité (1.2) peut être remplie pour plusieurs indices α, mais l'égalité (1.3) doit être valable pour chacun de ces indices). Nous dirons qu'une famille $\{\mu_\alpha\}$ de mesures définies sur un σ-corps $\mathcal{A}$ est *ordonnée suivant la dimension*, si les indices $\{\alpha\}$ forment un ensemble ordonné et si $A \in \mathcal{A}$, $\mu_\alpha(A) < +\infty$, $\alpha < \beta$ implique $\mu_\beta(A) = 0$.

En employant la terminologie que nous venons d'introduire, le but de cet ouvrage peut se formuler de la façon suivante: examiner les conditions nécessaires et suffisantes de ce qu'un espace de probabilité conditionnelle puisse être engendré par une famille de mesures quelconque, de ce qu'il puisse être engendré par une famille de mesures ordonnée suivant la dimension, et de ce qu'il puisse être engendré par une famille de mesures composée d'une seule mesure (en ce cas nous dirons que l'espace est engendré par cette mesure). Nous allons donner la réponse à ces questions principales et nous trouverons quelques résultats complémentaires concernant la structure des espaces de probabilité conditionnelle.

2. Dans ce qui suit, nous supposerons toujours qu'un ensemble S, un σ-corps $\mathcal{A}$ des sous-ensembles de S qui contient l'ensemble S et une famille non vide $\mathcal{B} \subset \mathcal{A}$ sont donnés et qu'une fonction réelle de deux ensembles $\mathbf{P}(A|B)$ est définie pour $A \in \mathcal{A}$, $B \in \mathcal{B}$. Un système $[S, \mathcal{A}, \mathcal{B}, \mathbf{P}]$ de ce type et satisfaisant aux axiomes I, II et III sera appelé *espace de probabilité conditionnelle*.

On a la proposition suivante:

(2.1) *En admettant les axiomes* I *et* II, III *équivaut à l'ensemble des conditions*

III_1. $\mathbf{P}(A|B) = \mathbf{P}(AB|B)$ *pour* $A \in \mathcal{A}, B \in \mathcal{B}$

et

III_2. *On a pour* $A \subset B \subset C$, $A \in \mathcal{A}$, $B \in \mathcal{B}$, $C \in \mathcal{B}$

$$\mathbf{P}(A|B)\,\mathbf{P}(B|C) = \mathbf{P}(A|C). \tag{2.2}$$

DÉMONSTRATION. a) Supposons que III est valable et posons $B = C$ dans (1.1) pour $A \in \mathcal{A}, B \in \mathcal{B}$. On obtient $\mathbf{P}(A|B)\,\mathbf{P}(B|B) = \mathbf{P}(AB|B)$, ce qui donne d'après I la condition III_1. Considérons ensuite des ensembles $A \subset B \subset C$, $A \in \mathcal{A}$, $B \in \mathcal{B}$, $C \in \mathcal{B}$, et appliquons la formule (1.1), on arrive à III_2. (Cf. [1], p. 373 et [2], p. 291.)

b) En supposant que les conditions III_1 et III_2 sont remplies, considérons des ensembles A, B et C tels que $A \in \mathcal{A}, B \in \mathcal{A}, C \in \mathcal{B}$ et $BC \in \mathcal{B}$; on aura en vertu de $ABC \subset BC \subset C$

$$\mathbf{P}(A|BC)\,\mathbf{P}(B|C) = \mathbf{P}(ABC|BC)\,\mathbf{P}(BC|C) = \mathbf{P}(ABC|C) = \mathbf{P}(AB|C),$$

donc III est valable.

En généralisant la notion d'espace de probabilité conditionnelle, nous appellerons *espace général de probabilité conditionnelle* un système $[S, \mathcal{A}, \mathcal{B}, \mathbf{P}]$ du type envisagé plus haut et satisfaisant aux conditions I, II et III_1. Nous dirons qu'un espace général de probabilité conditionnelle $[S, \mathcal{A}, \mathcal{B}, \mathbf{P}]$ est *engendré* par une famille $\{\mu_\alpha\}$ de mesures, si

a) chacune des mesures μ_α est définie sur $\mathcal{A}$,

b) à tout ensemble $B \in \mathcal{B}$ correspond au moins un indice α tel que $0 < \mu_\alpha(B) < +\infty$,

c) on a pour $A \in \mathcal{A}$, $B \in \mathcal{B}$, $0 < \mu_\alpha(B) < +\infty$ l'égalité

$$\mathbf{P}(A|B) = \frac{\mu_\alpha(AB)}{\mu_\alpha(B)}.$$

(2.3) *Un ensemble S, un σ-corps $\mathcal{A}$ des sous-ensembles de S tel que $S \in \mathcal{A}$, une famille non vide $\mathcal{B} \subset \mathcal{A}$ et une famille $\{\mu_\alpha\}$ de mesures définies sur $\mathcal{A}$ étant donnés, pour que la famille $\{\mu_\alpha\}$ engendre un espace général de probabilité conditionnelle $[S, \mathcal{A}, \mathcal{B}, \mathbf{P}]$, il faut et il suffit que les conditions* a), b) *et* d) *soient remplies, où* d) *désigne la condition suivante:*

d) *Pour $A \in \mathcal{A}$, $B \in \mathcal{B}$, $A \subset B$, $\mu_\alpha(B) < +\infty$, $\mu_\beta(B) < +\infty$ on a*

$$\mu_\alpha(A)\,\mu_\beta(B) = \mu_\beta(A)\,\mu_\alpha(B). \tag{2.4}$$

DÉMONSTRATION. Les conditions a) et b) étant nécessaires par définition, pour que la famille $\{\mu_\alpha\}$ engendre un espace général de probabilité conditionnelle, il faut encore que la formule

$$\mathbf{P}(A|B) = \frac{\mu_\alpha(AB)}{\mu_\alpha(B)}$$

définisse la fonction $\mathbf{P}(A|B)$ univoquement, c'est-à-dire qu'on ait pour $A \in \mathcal{A}$,

$B \in \mathfrak{B}$, $0 < \mu_\alpha(B) < +\infty$, $0 < \mu_\beta(B) < +\infty$

$$\frac{\mu_\alpha(AB)}{\mu_\alpha(B)} = \frac{\mu_\beta(AB)}{\mu_\beta(B)}. \tag{2.5}$$

(2.4) est la conséquence de (2.5) pour $A \in \mathfrak{A}$, $B \in \mathfrak{B}$, $A \subset B$, si $0 < \mu_\alpha(B) < +\infty$, $0 < \mu_\beta(B) < +\infty$. Or, si $\mu_\alpha(B) = 0$ ou si $\mu_\beta(B) = 0$, on a encore $\mu_\alpha(A) = 0$ ou $\mu_\beta(A) = 0$, selon le cas, donc (2.4) est valable en ce cas aussi et la nécessité de la condition d) est établie.

Réciproquement, si a), b) et d) se vérifient, l'égalité (2.5) étant la conséquence de (2.4), la fonction $\mathbf{P}(A|B)$ est définie univoquement par la formule de la condition c) et on constate sans peine qu'elle satisfait aux axiomes I, II et III_1.

(2.6) *Si l'espace général de probabilité conditionnelle* $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$ *est engendré par une famille* $\{\mu_\alpha\}$ *de mesures, il est un espace de probabilité conditionnelle (c'est-à-dire il remplit l'axiome* III_2).

DÉMONSTRATION. Il correspond par hypothèse à tout ensemble $B \in \mathfrak{B}$ au moins un indice α tel que $0 < \mu_\alpha(B) < +\infty$, et on a pour $A \in \mathfrak{A}$

$$\mathbf{P}(A|B) = \frac{\mu_\alpha(AB)}{\mu_\alpha(B)}.$$

Soit $A \subset B \subset C$, $A \in \mathfrak{A}$, $B \in \mathfrak{B}$, $C \in \mathfrak{B}$. Si $\mathbf{P}(B|C) = 0$, on a d'après I et II $0 \leqq \mathbf{P}(A|C) \leqq \mathbf{P}(B|C) = 0$, donc (2.2) est valable. Si $\mathbf{P}(B|C) > 0$, soit α un indice tel que $0 < \mu_\alpha(C) < +\infty$; on a alors

$$\mathbf{P}(B|C) = \frac{\mu_\alpha(BC)}{\mu_\alpha(C)} = \frac{\mu_\alpha(B)}{\mu_\alpha(C)} > 0,$$

donc $0 < \mu_\alpha(B) < +\infty$, et

$$\mathbf{P}(A|B)\,\mathbf{P}(B|C) = \frac{\mu_\alpha(AB)}{\mu_\alpha(B)} \frac{\mu_\alpha(B)}{\mu_\alpha(C)} = \frac{\mu_\alpha(A)}{\mu_\alpha(C)} = \frac{\mu_\alpha(AC)}{\mu_\alpha(C)} = \mathbf{P}(A|C),$$

donc (2.2) est valable en ce cas aussi.

On a la réciproque suivante de (2.6):

(2.7) *Tout espace de probabilité conditionnelle* $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$ *peut être engendré par une famille de mesures.*

DÉMONSTRATION. Définissons la mesure $\mu_B(A)$ pour $A \in \mathfrak{A}$, $B \in \mathfrak{B}$ de la façon suivante:

$$\mu_B(A) = \begin{cases} \mathbf{P}(A|B) & \text{pour} \quad A \subset B, \\ +\infty & \text{pour} \quad A - B \neq 0. \end{cases}$$

On voit aisément d'après I et II que $\mu_B(A)$ est une mesure sur $\mathfrak{A}$ et que $\mu_B(B) = \mathbf{P}(B|B) = 1$ pour $B \in \mathfrak{B}$, donc que les conditions a) et b) sont remplies. Pour démontrer la validité de la condition c), soit $B \in \mathfrak{B}$, $B' \in \mathfrak{B}$, $0 < \mu_{B'}(B) <$

$< +\infty$. On a alors $B \subset B'$ et $\mathbf{P}(B|B') = \mu_{B'}(B)$, donc d'après III_1 et III_2

$$\mathbf{P}(A|B) = \mathbf{P}(AB|B) = \frac{\mathbf{P}(AB|B')}{\mathbf{P}(B|B')} = \frac{\mu_{B'}(AB)}{\mu_{B'}(B)}$$

pour $A \in \mathcal{A}$ quelconque, ce qu'il fallait démontrer.

En résumant (2. 5) et (2. 6), on peut donner la réponse suivante à la première question principale :

(2. 8) *Pour qu'un espace général de probabilité conditionnelle puisse être engendré par une famille de mesures, il faut et il suffit qu'il soit un espace de probabilité conditionnelle (c'est-à-dire qu'il remplisse l'axiome* III_2*).*

3. Nos résultats précédents montrent que tout espace de probabilité conditionnelle $[S, \mathcal{A}, \mathcal{B}, \mathbf{P}]$ peut être engendré par une famille de mesures, mais l'analyse de la démonstration montre que la puissance de cette famille peut être très élevée, jusqu'à être égale à la puissance de la famille $\mathcal{B}$. La question se pose, quelles sont les conditions qui permettent de diminuer le nombre des mesures qui sont nécessaires à engendrer un espace de probabilité conditionnelle donné, ou au moins de diminuer le nombre des différentes valeurs des mesures d'un même ensemble $A \in \mathcal{A}$. Dans cet ordre d'idées, nous allons déterminer la condition nécessaire et suffisante de ce qu'un espace de probabilité conditionnelle $[S, \mathcal{A}, \mathcal{B}, \mathbf{P}]$ puisse être engendré par une famille $\{\mu_\alpha\}$ de mesures de façon que, pour un ensemble $A \in \mathcal{A}$ donné, la valeur $\mu_\alpha(A)$ ne prenne au plus que les valeurs $0, +\infty$ et une seule valeur finie et positive. Dans cette condition, et dans plusieurs autres conditions qui vont suivre, le rôle principal sera joué par la formule

$$\prod_{i=1}^{n} \mathbf{P}(A_i|B_i) = \prod_{i=1}^{n} \mathbf{P}(A_i|B_{i+1}) \tag{3. 1}$$

pour

$$A_i \in \mathcal{A},\ B_i \in \mathcal{B},\ A_i \subset B_i B_{i+1} \qquad (i = 1, \ldots, n;\ B_{n+1} = B_1). \tag{3. 2}$$

Introduisons l'axiome suivant :

IV_1. (3. 2) *entraîne* (3. 1) *pourvu que l'un et l'autre membres de* (3. 1) *soient positifs.*

On a alors la proposition suivante :

(3. 3) *Si l'espace de probabilité conditionnelle* $[S, \mathcal{A}, \mathcal{B}, \mathbf{P}]$ *peut être engendré par une famille* $\{\mu_\alpha\}$ *de mesures de façon que* $A \in \mathcal{A}$, $0 < \mu_\alpha(A) < +\infty$, $0 < \mu_\beta(A) < +\infty$ *entraîne* $\mu_\alpha(A) = \mu_\beta(A)$, *cet espace satisfait à* IV_1.

DÉMONSTRATION. En supposant que les ensembles A_i et B_i satisfont à (3. 2) et que l'un et l'autre membres de (3. 1) sont positifs, et en choisissant les indices α_i de façon que l'on ait $0 < \mu_{\alpha_i}(B_i) < +\infty$, on a

$$\mathbf{P}(A_i|B_i) = \frac{\mu_{\alpha_i}(A_i)}{\mu_{\alpha_i}(B_i)}, \quad \mathbf{P}(A_i|B_{i+1}) = \frac{\mu_{\alpha_{i+1}}(A_i)}{\mu_{\alpha_{i+1}}(B_{i+1})},$$

donc $0 < \mu_{\alpha_i}(A_i) < +\infty$, $0 < \mu_{\alpha_{i+1}}(A_i) < +\infty$ pour $i = 1, \ldots, n$, $\alpha_{n+1} = \alpha_1$. Il s'ensuit $\mu_{\alpha_i}(A_i) = \mu_{\alpha_{i+1}}(A_i)$, on a par conséquent

$$(3.4) \qquad \prod_{i=1}^{n} \mathbf{P}(A_i|B_i) = \frac{\prod_{i=1}^{n} \mu_{\alpha_i}(A_i)}{\prod_{i=1}^{n} \mu_{\alpha_i}(B_i)} = \frac{\prod_{i=1}^{n} \mu_{\alpha_{i+1}}(A_i)}{\prod_{i=1}^{n} \mu_{\alpha_{i+1}}(B_{i+1})} = \prod_{i=1}^{n} \mathbf{P}(A_i|B_{i+1}).$$

donc (3.1) est valable.

Dans les démonstrations qui vont suivre, les notions suivantes nous serons utiles: $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$ étant un espace général de probabilité conditionnelle, nous dirons que la suite finie $B_1, \ldots, B_n$ $(B_i \in \mathfrak{B})$ forme une *chaîne*, si on a

$$\mathbf{P}(B_iB_{i+1}|B_i) > 0, \quad \mathbf{P}(B_iB_{i+1}|B_{i+1}) > 0 \qquad (i = 1, \ldots, n-1).$$

Nous définissons pour les ensembles de la famille $\mathfrak{B}$ une relation d'équivalence $B \sim B'$ en appelant *équivalents* les ensembles $B \in \mathfrak{B}$ et $B' \in \mathfrak{B}$, s'il existe une chaîne $B_1, \ldots, B_n$ telle que $B_1 = B$ et $B_n = B'$. Le caractère symétrique et transitif de cette relation est évident et sa réflexivité s'ensuit de l'axiome I, en particulier de la relation $\mathbf{P}(B|B) = 1 > 0$.

La proposition suivante fournit la réciproque de (3.3):

(3.5) *Si l'espace de probabilité conditionnelle* $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$ *satisfait à l'axiome* IV_1, *il peut être engendré par une famille* $\{\mu_\alpha\}$ *de mesures de façon que* $A \in \mathfrak{A}$, $0 < \mu_\alpha(A) < +\infty$, $0 < \mu_\beta(A) < +\infty$ *entraîne* $\mu_\alpha(A) = \mu_\beta(A)$.

DÉMONSTRATION. Rangeons les ensembles de $\mathfrak{B}$ dans des classes d'après la relation d'équivalence introduite et formons une famille $\mathfrak{S} \subset \mathfrak{B}$ qui a avec chacune de ces classes un élément commun et un seul.

Nous définissons pour $B \in \mathfrak{B}$ la mesure $\mu_B(A)$ $(A \in \mathfrak{A})$ comme suit: pour $A - B \neq 0$ posons $\mu_B(A) = +\infty$, et pour $A \subset B$, soit $E \in \mathfrak{S}$, $E \sim B$, soit $E = B_1, \ldots, B_n = B$ une chaîne et posons

$$(3.6) \qquad \mu_B(A) = \mathbf{P}(A|B) \prod_{i=1}^{n-1} \frac{\mathbf{P}(B_iB_{i+1}|B_i)}{\mathbf{P}(B_iB_{i+1}|B_{i+1})}.$$

Pour justifier cette définition, il faut démontrer que la valeur de $\mu_B(A)$ est indépendante du choix particulier de la chaîne qui relie E à B. Ceci étant évident si $\mathbf{P}(A|B) = 0$, on peut supposer que $\mathbf{P}(A|B) > 0$. Nous démontrons en ce cas un peu plus, à savoir que si $E = B'_1, \ldots, B'_m = B' \supset A$ est une autre chaîne et que si $\mathbf{P}(A|B') > 0$, on a $\mu_B(A) = \mu_{B'}(A)$. En effet, en ce cas $B' = B'_m, B'_{m-1}, \ldots, B'_1 = E = B_1, B_2, \ldots, B_n = B$ est une chaîne elle aussi, donc en vertu de IV_1

$$(3.7) \qquad \begin{aligned} &\mathbf{P}(A|B') \prod_{k=m-1}^{1} \mathbf{P}(B'_kB'_{k+1}|B'_k) \prod_{i=1}^{n-1} \mathbf{P}(B_iB_{i+1}|B_{i+1}) = \\ &= \prod_{k=m-1}^{1} \mathbf{P}(B'_kB'_{k+1}|B'_{k+1}) \prod_{i=1}^{n-1} \mathbf{P}(B_iB_{i+1}|B_i) \cdot \mathbf{P}(A|B), \end{aligned}$$

l'un et l'autre membres étant positifs par hypothèse; il s'ensuit

$$(3.8)\qquad \mathbf{P}(A|B')\prod_{k=1}^{m-1}\frac{\mathbf{P}(B'_kB'_{k+1}|B'_k)}{\mathbf{P}(B'_kB'_{k+1}|B'_{k+1})}=\mathbf{P}(A|B)\prod_{i=1}^{n-1}\frac{\mathbf{P}(B_iB_{i+1}|B_i)}{\mathbf{P}(B_iB_{i+1}|B_{i+1})}.$$

Ceci établi, μ_B est évidemment une mesure sur $\mathfrak{A}$. Si $A\in\mathfrak{A}$, $0<\mu_B(A)<+\infty$, $0<\mu_{B'}(A)<+\infty$, on a $A\subset B$, $A\subset B'$ et $\mathbf{P}(A|B)>0$, $\mathbf{P}(A|B')>0$, donc a fortiori $\mathbf{P}(BB'|B)>0$, $\mathbf{P}(BB'|B')>0$, de sorte que B, B' est une chaîne, $B\sim B'$, et d'après ce que nous venons de démontrer, on a $\mu_B(A)=\mu_{B'}(A)$.

La famille $\{\mu_B\}$ jouit donc de la propriété affirmée dans l'énoncé. Il faut encore démontrer qu'elle engendre l'espace $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$. Or, on a pour $B\in\mathfrak{B}$

$$\mu_B(B)=\mathbf{P}(B|B)\prod_{i=1}^{n-1}\frac{\mathbf{P}(B_iB_{i+1}|B_i)}{\mathbf{P}(B_iB_{i+1}|B_{i+1})}=\prod_{i=1}^{n-1}\frac{\mathbf{P}(B_iB_{i+1}|B_i)}{\mathbf{P}(B_iB_{i+1}|B_{i+1})},$$

donc $0<\mu_B(B)<+\infty$ et, pour $A\in\mathfrak{A}$ quelconque,

$$\mathbf{P}(A|B)=\mathbf{P}(AB|B)=\frac{\mu_B(AB)}{\mu_B(B)}.$$

Enfin, si $0<\mu_{B'}(B)<+\infty$, $B\in\mathfrak{B}$, on a

$$\frac{\mu_{B'}(AB)}{\mu_{B'}(B)}=\frac{\mu_{B'}(AB)}{\mu_{B'}(B')}\cdot\frac{\mu_{B'}(B')}{\mu_{B'}(B)}=\mathbf{P}(AB|B')\frac{1}{\mathbf{P}(B|B')}=\mathbf{P}(AB|B),$$

en vertu de III_2, ce qui était encore à démontrer.

En résumant (3.3) et (3.5), on peut dire:

(3.9) *L'axiome* IV_1 *fournit la condition nécessaire et suffisante de ce qu'un espace de probabilité conditionnelle* $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$ *puisse être engendré par une famille* $\{\mu_\alpha\}$ *de mesures de façon que* $A\in\mathfrak{A}$, $0<\mu_\alpha(A)<+\infty$, $0<\mu_\beta(A)<+\infty$ *entraîne* $\mu_\alpha(A)=\mu_\beta(A)$.

4. L'axiome IV_1 assure donc l'existence d'une famille $\{\mu_\alpha\}$ de mesures telle qu'elle engendre l'espace de probabilité conditionnelle $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$ et que $A\in\mathfrak{A}$, $0<\mu_\alpha(A)<+\infty$, $0<\mu_\beta(A)<+\infty$ entraîne $\mu_\alpha(A)=\mu_\beta(A)$. On voit cependant (cf. § 7) que les conditions en question n'impliquent pas que μ_α et μ_β coïncident sur tout ensemble $A'\in\mathfrak{A}$, $A'\subset A$. Pour pouvoir formuler la condition qui entraîne cette proposition plus forte, convenons de la définition suivante: $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$ étant un espace général de probabilité conditionnelle, nous dirons que la suite finie $B_1, \ldots, B_n$ $(B_i\in B)$ forme un *cycle*, si l'on a

$$\mathbf{P}(B_iB_{i+1}|B_i)>0,\quad \mathbf{P}(B_iB_{i+1}|B_{i+1})>0\qquad (i=1,\ldots,n;\ B_{n+1}=B_1).$$

Introduisons l'axiome suivant:

IV_2. (3.2) *entraîne* (3.1) *pourvu que* $B_1, \ldots, B_n$ *soit un cycle.*

On a la proposition suivante:

(4.1) *En admettant les axiomes* I, II *et* III_1, IV_2 *entraîne* III_2 *et* IV_1.

DÉMONSTRATION. Soit $A \in \mathfrak{A}$, $B \in \mathfrak{B}$, $C \in \mathfrak{B}$, $A \subset B \subset C$. L'égalité

$$\mathbf{P}(A|B)\,\mathbf{P}(B|C) = \mathbf{P}(A|C)$$

étant valable au cas où l'on a $\mathbf{P}(B|C) = 0$ (donc a fortiori $\mathbf{P}(A|C) = 0$ d'après I et II), on peut supposer que $\mathbf{P}(B|C) > 0$. En ce cas, les ensembles B et C forment un cycle, on a donc en vertu de IV_2 et I

$$\mathbf{P}(A|B)\,\mathbf{P}(B|C) = \mathbf{P}(A|C)\,\mathbf{P}(B|B) = \mathbf{P}(A|C),$$

l'axiome III_2 est donc valable.

Soit ensuite $A_i \in \mathfrak{A}$, $B_i \in \mathfrak{B}$, $A_i \subset B_i B_{i+1}$ $(i = 1, \ldots, n;\ B_{n+1} = B_1)$ et supposons que les nombres

$$\prod_{i=1}^{n} \mathbf{P}(A_i|B_i) \quad \text{et} \quad \prod_{i=1}^{n} \mathbf{P}(A_i|B_{i+1})$$

sont positifs. Il s'ensuit que $\mathbf{P}(B_i B_{i+1}|B_i) > 0$ et que $\mathbf{P}(B_i B_{i+1}|B_{i+1}) > 0$ $(i = 1, \ldots, n)$, donc que $B_1, \ldots, B_n$ est un cycle, d'où l'on a d'après IV_2 l'égalité (3 1), de sorte que IV_1 est, lui aussi, valable.

Le rôle de l'axiome IV_2 est mis en évidence par la proposition suivante:

(4. 2) *L'axiome* IV_2 *fournit la condition nécessaire et suffisante de ce que l'espace de probabilité conditionnelle* $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$ *puisse être engendré par une famille* $\{\mu_\alpha\}$ *de mesures de façon que* $A \in \mathfrak{A}$, $0 < \mu_\alpha(A) < +\infty$, $0 < \mu_\beta(A) < +\infty$ *entraîne* $\mu_\alpha(A') = \mu_\beta(A')$ *pour* $A' \in \mathfrak{A}$, $A' \subset A$.

DÉMONSTRATION. *Nécessité.* Si $\{\mu_\alpha\}$ est une famille de mesures du type en question qui engendre l'espace $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$, soit $B_1, \ldots, B_n$ un cycle et soit $A_i \in \mathfrak{A}$, $A_i \subset B_i B_{i+1}$ $(i = 1, \ldots, n;\ B_{n+1} = B_1)$. On a alors pour des indices α_i convenables $0 < \mu_{\alpha_i}(B_i) < +\infty$, donc

$$\mathbf{P}(B_i B_{i+1}|B_i) = \frac{\mu_{\alpha_i}(B_i B_{i+1})}{\mu_{\alpha_i}(B_i)} > 0, \quad \mathbf{P}(B_i B_{i+1}|B_{i+1}) = \frac{\mu_{\alpha_{i+1}}(B_i B_{i+1})}{\mu_{\alpha_{i+1}}(B_{i+1})} > 0,$$

où $\alpha_{n+1} = \alpha_1$. Il s'ensuit $\mu_{\alpha_i}(A_i) = \mu_{\alpha_{i+1}}(A_i)$ et par conséquent l'égalité (3. 4) est valable, ce qui montre la validité de IV_2.

Suffisance. Supposons que l'axiome IV_2 est rempli dans l'espace de probabilité conditionnelle $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$. D'après (4. 1), il en est de même quant à l'axiome IV_1, de sorte qu'en définissant $\mu_B(A)$ comme dans la démonstration de (3. 5), on obtient une famille $\{\mu_B\}$ de mesures qui engendre l'espace $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$. Si $A \in \mathfrak{A}$, $0 < \mu_B(A) < +\infty$, $0 < \mu_{B'}(A) < +\infty$, $A' \in \mathfrak{A}$, $A' \subset A$, on a $A \subset BB'$ et

$$\mathbf{P}(BB'|B) \geqq \mathbf{P}(A|B) = \frac{\mu_B(A)}{\mu_B(B)} > 0, \quad \mathbf{P}(BB'|B') \geqq \mathbf{P}(A|B') = \frac{\mu_{B'}(A)}{\mu_{B'}(B')} > 0,$$

donc les ensembles B et B' forment un cycle, de sorte que

$$\mathbf{P}(A'|B)\,\mathbf{P}(BB'|B') = \mathbf{P}(A'|B')\,\mathbf{P}(BB'|B)$$

en vertu de IV_2. On en conclut que les nombres $\mathbf{P}(A'|B)$ et $\mathbf{P}(A'|B')$ ou bien s'annulent l'un et l'autre, ou bien sont positifs. On a dans le premier cas

$$\mu_B(A') = \mathbf{P}(A'|B)\ \mu_B(B) = 0, \quad \mu_{B'}(A') = \mathbf{P}(A'|B')\ \mu_{B'}(B') = 0,$$

dans le second $\mu_B(A') > 0$, $\mu_{B'}(A') > 0$ et d'après la démonstration de (3. 5), $\mu_B(A') = \mu_{B'}(A')$. En tout cas, la famille $\{\mu_B\}$ possède la propriété affirmée.

Les propositions (2. 1), (2. 6), (4. 1) et (4. 2) se résument comme suit:

(4. 3) *Pour qu'un espace général de probabilité conditionnelle* $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *puisse être engendré par une famille* $\{\mu_\alpha\}$ *de mesures de façon que* $A \in \mathcal{A}$, $0 < \mu_\alpha(A) < +\infty$, $0 < \mu_\beta(A) < +\infty$ *entraîne* $\mu_\alpha(A') = \mu_\beta(A')$ *pour* $A' \in \mathcal{A}$, $A' \subset A$, *il faut et il suffit que cet espace remplisse l'axiome* IV_2.

5. Passons à la question des espaces de probabilité conditionnelle engendrés par des familles de mesures ordonnées suivant la dimension. L'existence d'une famille génératrice de ce type sera assurée par l'axiome suivant:

IV_3. *Les conditions* (3. 2) *entraînent l'égalité* (3. 1).

L'axiome IV_3 implique évidemment IV_2, donc (en admettant I, II et III_1) III_2 et IV_1.

La nécessité de IV_3 est exprimée par la proposition suivante:

(5. 1) *Si l'espace de probabilité conditionnelle* $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *peut être engendré par une famille* $\{\mu_\alpha\}$ *de mesures ordonnée suivant la dimension, cet espace satisfait à* IV_3.

DÉMONSTRATION. Soit $B_i \in \mathfrak{B}$, $A_i \in \mathcal{A}$, $A_i \subset B_i B_{i+1}$, $0 < \mu_{\alpha_i}(B_i) < +\infty$ $(i = 1, \ldots, n;\ B_{n+1} = B_1)$. Il faut démontrer qu'on a

$$\prod_{i=1}^{n} \frac{\mu_{\alpha_i}(A_i)}{\mu_{\alpha_i}(B_i)} = \prod_{i=1}^{n} \frac{\mu_{\alpha_{i+1}}(A_i)}{\mu_{\alpha_{i+1}}(B_{i+1})} \qquad (\alpha_{n+1} = \alpha_1) \tag{5. 2}$$

Cette égalité étant évidemment valable si l'un et l'autre numérateurs contiennent chacun au moins un facteur qui s'annule, supposons qu'on a par exemple $\mu_{\alpha_i}(A_i) > 0$ $(i = 1, \ldots, n)$. On a alors $\alpha_i \leqq \alpha_{i+1}$ pour $i = 1, \ldots, n$, puisqu'autrement on aurait $\mu_{\alpha_{i+1}}(A_i) = +\infty$ pour au moins une valeur de l'indice i. Il s'ensuit que $\alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_n \leqq \alpha_1$, donc que $\alpha_1 = \alpha_2 = \cdots = \alpha_n$ et que par conséquent $\mu_{\alpha_i}(A_i) = \mu_{\alpha_{i+1}}(A_i)$, de sorte que (5. 2) est valable.

La démonstration de la réciproque de la proposition (5. 1) sera un peu plus compliquée. Elle fera usage de la notion et du résultat suivants: E et E' étant deux ensembles partiellement ordonnés, considérons une transformation $f(x)$ qui fait correspondre à tout élément $x \in E$ un élément $x' = f(x) \in E'$. La transformation $f(x)$ sera dite *strictement isotone*, si l'inégalité $x < y$ entraîne toujours celle $f(x) < f(y)$. En faisant usage du théorème du bon ordre, on peut démontrer que *tout ensemble partiellement ordonné peut être transformé par une transformation strictement isotone en un ensemble* (complètement)

ordonné. En effet, on sait qu'on peut définir dans tout ensemble partiellement ordonné un ordre complet de façon que si la relation $x < y$ était valable d'après l'ordre partiel, elle reste valable d'après l'ordre complet (v. p. ex. [3]). En désignant par E l'ensemble partiellement ordonné en question et par E' le même ensemble pourvu de l'ordre complet dont nous venons de parler, la transformation $f(x) = x$ fournit une transformation du type envisagé.

Nous pouvons maintenant démontrer la réciproque suivante de (5. 1):

(5. 3) *Si l'espace de probabilité conditionnelle* $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *satisfait à l'axiome* IV_3, *il peut être engendré par une famille de mesures ordonnée suivant la dimension.*

DÉMONSTRATION. La suite finie $B_1, \ldots, B_n$ $(B_i \in \mathfrak{B})$ sera appelée *chaîne ascendante* si l'on a $\mathbf{P}(B_iB_{i+1}|B_i) > 0$ pour $i = 1, \ldots, n-1$, mais l'égalité $\mathbf{P}(B_iB_{i+1}|B_{i+1}) = 0$ est valable pour au moins un indice $i < n$. Nous introduisons un ordre partiel dans la famille $\mathfrak{B}$ en convenant de ce que la relation $B < B'$ est valable pour deux ensembles $B \in \mathfrak{B}$, $B' \in \mathfrak{B}$, si et seulement s'il existe une chaîne ascendante $B_1, \ldots, B_n$ telle que $B = B_1, B' = B_n$. Les inégalités $B < B'$ et $B' < B''$ entraînent évidemment celle $B < B''$ et $B < B$ est impossible, car si $B = B_1, \ldots, B_n = B$ était une chaîne ascendante, le produit $\prod_{i=1}^{n} \mathbf{P}(B_iB_{i+1}|B_i)$ (où $B_{n+1} = B_1$) devrait être positif (puisque $\mathbf{P}(B_nB_{n+1}|B_n) = \mathbf{P}(B|B) > 0$), tandis que celui $\prod_{i=1}^{n} \mathbf{P}(B_iB_{i+1}|B_{i+1})$ serait zéro, se qui contredit à IV_3. On obtient donc en réalité un ordre partiel dans $\mathfrak{B}$.

Remarquons que

(5. 4) *les hypothèses* $A \in \mathcal{A}$, $B \in \mathfrak{B}$, $B' \in \mathfrak{B}$, $A \subset BB'$, $\mathbf{P}(A|B) > 0$, $\mathbf{P}(A|B') = 0$ *impliquent* $B < B'$.

En effet, on a d'après IV_3

$$\mathbf{P}(A|B)\,\mathbf{P}(BB'|B') = \mathbf{P}(A|B')\,\mathbf{P}(BB'|B),$$

donc $\mathbf{P}(BB'|B') = 0$, tandis que $\mathbf{P}(BB'|B) \geqq \mathbf{P}(A|B) > 0$, de sorte que B, B' est une chaîne ascendante.

L'ordre partiel que nous venons d'introduire dans $\mathfrak{B}$ est compatible avec la relation d'équivalence introduite plus haut. En effet, si l'on a $B < B'$, $B^* \sim B, B^{**} \sim B'$, soit $B = B_1, \ldots, B_n = B'$ une chaîne ascendante et soient $B^* = B_1', \ldots, B_m' = B$ et $B' = B_1'', \ldots, B_p'' = B^{**}$ des chaînes, la suite finie $B^* = B_1', \ldots, B_m' = B = B_1, \ldots, B_n = B' = B_1'', \ldots, B_p'' = B^{**}$ sera évidemment une chaîne ascendante, on aura donc $B^* < B^{**}$.

En désignant donc par $\{\mathbb{S}_a\}$ la famille des classes d'équivalence correspondant à la relation que nous avons introduite, on peut définir un ordre partiel dans l'ensemble $\{\mathbb{S}_a\}$ ou, ce qui veut dire la même chose, dans l'ensemble $\{a\}$ des indices correspondants, en posant $a < b$ si et seulement

s'il y a des ensembles $B \in \mathfrak{S}_a$ et $B' \in \mathfrak{S}_b$ tels que $B < B'$. L'inégalité $a < b$ impliquera alors $B_1 < B_1'$ pour deux ensembles quelconques $B_1 \in \mathfrak{S}_a$ et $B_1' \in \mathfrak{S}_b$.

Soit $\alpha = f(a)$ une transformation strictement isotone de l'ensemble $\{a\}$ en un ensemble ordonné $\{\alpha\}$ (nous avons parlé plus haut de la question d'existence d'une transformation de ce type). Désignons par $\mathfrak{Q}_\alpha$ la famille des ensembles $A \in \mathfrak{A}$ tels qu'il y a un ensemble B pour lequel on a $A \subset B$, $B \in \mathfrak{S}_a$, $f(a) \leqq \alpha$.

Puisque l'axiome IV_3 entraîne IV_1, on peut définir pour $B \in \mathfrak{B}$ la mesure μ_B comme dans la démonstration de (3.5); on aura

$$\mathbf{P}(A|B) = \frac{\mu_B(AB)}{\mu_B(B)}$$

pour $A \in \mathfrak{A}$, $B \in \mathfrak{B}$, et si $A \in \mathfrak{A}$, $B \in \mathfrak{B}$, $B' \in \mathfrak{B}$, $B \sim B'$, $A \subset BB'$, avec les mêmes notations que dans la démonstration que nous venons de mentionner, l'axiome IV_3 implique la validité de l'égalité (3.7), donc de celle (3.8). Par conséquent

(5.5) *les hypothèses* $A \in \mathfrak{A}$, $A \subset BB'$, $B \in \mathfrak{S}_a$, $B' \in \mathfrak{S}_a$ *impliquent* $\mu_B(A) = \mu_{B'}(A)$.

Nous allons définir pour $A \in \mathfrak{Q}_\alpha$ une fonction d'ensemble $\varphi_\alpha(A)$ comme suit:

(5.6) $\qquad \varphi_\alpha(A) = \mu_B(A) \quad \text{si} \quad A \subset B,\ B \in \mathfrak{S}_a,\ f(a) = \alpha,$

(5.7) $\qquad \varphi_\alpha(A) = 0 \qquad \text{si} \quad A \subset B',\ B' \in \mathfrak{S}_b,\ f(b) = \beta < \alpha.$

Pour justifier cette définition, remarquons d'abord que, pour tout ensemble $A \in \mathfrak{Q}_\alpha$, au moins une des possibilités (5.6) et (5.7) se réalise. Supposons ensuite que les hypothèses de (5.6) sont réalisées de deux façons différentes:

$$A \subset B,\ B \in \mathfrak{S}_a,\ f(a) = \alpha \quad \text{et} \quad A \subset B',\ B' \in \mathfrak{S}_{a'},\ f(a') = \alpha.$$

En ce cas, ou bien on a $\mathbf{P}(BB'|B) = \mathbf{P}(BB'|B') = 0$, donc $\mathbf{P}(A|B) = \mathbf{P}(A|B') = 0$ et par conséquent $\mu_B(A) = \mu_{B'}(A) = 0$, ou bien on a $\mathbf{P}(BB'|B) > 0$, $\mathbf{P}(BB'|B') > 0$, donc $B \sim B'$, $a = a'$, et par conséquent, d'après (5.5), $\mu_B(A) = \mu_{B'}(A)$, ou bien on a, disons, $\mathbf{P}(BB'|B) > 0$, $\mathbf{P}(BB'|B') = 0$, donc, d'après (5.4), $B < B'$, $a < a'$, $\alpha = f(a) < f(a') = \alpha$, ce qui est impossible. Il s'ensuit que la valeur de $\varphi_\alpha(A)$ fournie par (5.6) ne dépend pas du choix de l'ensemble B. Supposons enfin que les hypothèses de (5.6) et de (5.7) sont réalisées en même temps:

$$A \subset B,\ B \in \mathfrak{S}_a,\ f(a) = \alpha \quad \text{et} \quad A \subset B',\ B' \in \mathfrak{S}_b,\ f(b) = \beta < \alpha.$$

On a en ce cas $\mu_B(A) = 0$; en effet, si $\mu_B(A) = \mathbf{P}(A|B)\mu_B(B)$ était positif, l'inégalité $\mathbf{P}(A|B') > 0$ entraînerait $B \sim B'$, $a = b$, $\alpha = f(a) = f(b) = \beta < \alpha$, ce qui est impossible, tandis que de l'égalité $\mathbf{P}(A|B') = 0$ s'ensuivrait d'après (5.4) $B < B'$, $a < b$, $\alpha = f(a) < f(b) = \beta < \alpha$, ce qui est également impossible. Les formules (5.6) et (5.7) fournissent donc la même valeur à $\varphi_\alpha(A)$ au cas où les hypothèses de l'une et de l'autre sont remplies.

On constate sans peine que les conditions $A \in \mathcal{A}$, $A' \in \mathcal{Q}_\alpha$, $A \subset A'$ entraînent $A \in \mathcal{Q}_\alpha$ et $\varphi_\alpha(A) \leqq \varphi_\alpha(A')$, et que la fonction d'ensemble φ_α est σ-additive sur $\mathcal{Q}_\alpha$, c'est-à-dire que les hypothèses $A = \sum_1^\infty A_i$, $A \in \mathcal{Q}_\alpha$, $A_i \in \mathcal{Q}_\alpha$ $(i = 1, 2, \ldots)$, $A_i A_j = 0$ $(i \neq j)$ impliquent $\varphi_\alpha(A) = \sum_1^\infty \varphi_\alpha(A_i)$.

Définissons pour $X \subset S$ la mesure extérieure ν_α^* par la formule suivante :

$$\nu_\alpha^*(X) = \inf_{\substack{X \subset \sum_1^\infty A_i \\ A_i \in Q_\alpha}} \sum_1^\infty \varphi_\alpha(A_i), \qquad \nu_\alpha^*(0) = 0,$$

en convenant d'attribuer la valeur $+\infty$ à la borne inférieure d'un ensemble vide, c'est-à-dire en posant $\nu_\alpha^*(X) = +\infty$ si l'on ne peut pas couvrir l'ensemble X avec une suite dénombrable d'ensembles $A_i \in \mathcal{Q}_\alpha$. On démontre comme d'habitude que ν_α^* est une mesure extérieure, c'est-à-dire qu'on a $\nu_\alpha^*(X) \geqq 0$ pour $X \subset S$ et que $\nu_\alpha^*(X) \leqq \sum_1^\infty \nu_\alpha^*(X_i)$ pour $X \subset \sum_1^\infty X_i$.

Pour $A \in \mathcal{Q}_\alpha$, on a $\nu_\alpha^*(A) = \varphi_\alpha(A)$. En effet, en couvrant A avec soi-même, on obtient $\nu_\alpha^*(A) \leqq \varphi_\alpha(A)$; d'autre part, si $A \subset \sum_1^\infty A_i$, $A_i \in \mathcal{Q}_\alpha$, soit $A_i' = A A_i - \sum_{k<i} A A_k$. On aura $A_i' \in \mathcal{A}$ et $A_i' \subset A_i$, donc $A_i' \in \mathcal{Q}_\alpha$, $A = \sum_1^\infty A_i'$ et $A_i' A_j' = 0$ $(i \neq j)$, ce qui entraîne

$$\varphi_\alpha(A) = \sum_1^\infty \varphi_\alpha(A_i') \leqq \sum_1^\infty \varphi_\alpha(A_i).$$

Ceci étant valable pour toute suite dénombrable $\{A_i\}$ qui couvre A, on a $\varphi_\alpha(A) \leqq \nu_\alpha^*(A)$, ce qui était encore à démontrer.

En appelant ν_α^*-mesurable un ensemble $A \subset S$ si et seulement si l'on a pour tout ensemble $X \subset S$

$$\nu_\alpha^*(X) = \nu_\alpha^*(XA) + \nu_\alpha^*(X - A), \tag{5.8}$$

on démontre (v. p. ex. [4], p. 41 à 47) que les ensembles ν_α^*-mesurables forment un σ-corps et que la fonction ν_α^* est une mesure sur ce σ-corps.

Tout ensemble $A \in \mathcal{A}$ est ν_α^*-mesurable. En effet, le signe $\leqq$ étant toujours valable entre les deux membres de (5.8), soit $X \subset \sum_1^\infty A_i$, $A_i \in \mathcal{Q}_\alpha$. On a $AX \subset \sum_1^\infty A A_i$, $X - A \subset \sum_1^\infty (A_i - A)$, $A A_i \in \mathcal{A}$, $A A_i \subset A_i$, $A_i - A \in \mathcal{A}$, $A_i - A \subset A_i$,

donc $AA_i \in \mathcal{Q}_\alpha$, $A_i - A \in \mathcal{Q}_\alpha$, et par conséquent, en vertu de l'égalité $A_i = AA_i + (A_i - A)$,

$$\nu_\alpha^*(XA) + \nu_\alpha^*(X-A) \leqq \sum_1^\infty \varphi_\alpha(AA_i) + \sum_1^\infty \varphi_\alpha(A_i - A) = \sum_1^\infty \varphi_\alpha(A_i),$$

ce qui étant valable pout toute suite dénombrable $\{A_i\}$ qui couvre l'ensemble X,

$$\nu_\alpha^*(XA) + \nu_\alpha^*(X-A) \leqq \nu_\alpha^*(X),$$

d'où s'ensuit la proposition à démontrer.

Désignons par ν_α la mesure définie sur $\mathcal{A}$ et qui coïncide pour $A \in \mathcal{A}$ avec $\nu_\alpha^*(A)$. La famille $\{\nu_\alpha\}$ de mesures est ordonnée suivant la dimension. En effet, soit $\alpha < \beta$, $A \in \mathcal{A}$ et $\nu_\alpha(A) < +\infty$. On a $A \subset \sum_1^\infty A_i$, $A_i \in \mathcal{Q}_\alpha$, donc $A_i \subset B_i$, $B_i \in \mathcal{S}_{\alpha_i}$, $f(\alpha_i) = \alpha_i \leqq \alpha < \beta$. Mais les conditions

$$AA_i \subset B_i,\ B_i \in \mathcal{S}_{\alpha_i},\ f(\alpha_i) < \beta$$

entraînent d'après (5. 7) que $\varphi_\beta(AA_i) = 0$ $(i = 1, 2, \ldots)$, donc que $\nu_\beta^*(AA_i) = 0$ et que

$$0 \leqq \nu_\beta(A) = \nu_\beta^*(A) \leqq \sum_1^\infty \nu_\beta^*(AA_i) = 0.$$

La famille $\{\nu_\alpha\}$ engendre l'espace de probabilité conditionnelle $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$. Soit en effet $B \in \mathfrak{B}$, $0 < \nu_\alpha(B) < +\infty$. Si $B \in \mathcal{S}_b$, $f(b) = \beta$, on a $\nu_\beta(B) = \nu_\beta^*(B) = \varphi_\beta(B) = \mu_B(B)$, donc $0 < \nu_\beta(B) < +\infty$, ce qui n'est possible que si $\alpha = \beta$ (la famille $\{\nu_\alpha\}$ étant ordonnée suivant la dimension), on a par conséquent pour tout ensemble $A \in \mathcal{A}$

$$\mathbf{P}(A|B) = \frac{\mu_B(AB)}{\mu_B(B)} = \frac{\varphi_\alpha(AB)}{\varphi_\alpha(B)} = \frac{\nu_\alpha^*(AB)}{\nu_\alpha^*(B)} = \frac{\nu_\alpha(AB)}{\nu_\alpha(B)}.$$

Nous avons par là démontré le théorème (5. 3).

En résumant (5. 1) et (5. 3), on peut donner la réponse suivante à la seconde question principale :

(5. 9) *Pour qu'un espace de probabilité conditionnelle puisse être engendré par une famille de mesures ordonnée suivant la dimension, il faut et il suffit que cet espace remplisse l'axiome* IV_3.

En faisant usage des propositions (2. 1), (2. 6), (4. 1) et (5. 9) et en observant que IV_3 entraîne IV_2, on peut encore dire :

(5. 10) *Pour qu'un espace général de probabilité conditionnelle puisse être engendré par une famille de mesures ordonnée suivant la dimension, il faut et il suffit que cet espace remplisse l'axiome* IV_3.

6. L'analyse de la démonstration de (5. 3) montre que si l'on veut faire engendrer un espace de probabilité conditionnelle donné par une famille de mesures ordonnée suivant la dimension et se composant d'un nombre aussi

petit que possible de mesures différentes, il faut trouver une transformation strictement isotone de l'ensemble partiellement ordonné des classes d'équivalence définies dans $\mathfrak{B}$ en un ensemble ordonné de puissance aussi petite que possible. Dans cet ordre d'idées, il est aisé de démontrer la proposition suivante:

(6. 1) *Soit X un ensemble partiellement ordonné et supposons qu'il y a un nombre naturel N tel que $x_1 < x_2 < \cdots < x_n$, $x_i \in X$ entraîne $n \leqq N$. En désignant par N_0 le plus petit des nombres N de ce type, l'ensemble X peut être transformé de façon strictement isotone en un ensemble ordonné ayant la puissance N_0.*

DÉMONSTRATION. Nous dirons qu'un élément $x \in X$ a le *rang* $n(x)$, si $n(x)$ est le plus grand nombre naturel n pour lequel il existe des éléments $x_i \in X$ tels que $x_1 < x_2 < \cdots < x_n = x$. Par hypothèse tout élément $x \in X$ possède un rang $1 \leqq n(x) \leqq N_0$. En faisant correspondre à $x \in X$ le nombre $y = n(x)$, on parvient à une transformation strictement isotone de l'ensemble X en Y, où Y désigne l'ensemble $\{1, 2, \ldots, N_0\}$ pourvu de l'ordre naturel. En effet, si $x < x'$, $x \in X$, $x' \in X$, $n(x) = y$, il y a des éléments $x_i \in X$ tels que $x_1 < x_2 < \cdots < x_y = x$; par conséquent $x_1 < x_2 < \cdots < x_y < x'$, donc $n(x') \geqq \geqq y + 1 > n(x)$.

La proposition que nous venons de démontrer permet de donner une condition nécessaire et suffisante pour qu'un espace de probabilité conditionnelle puisse être engendré par une famille de mesures ordonnée suivant la dimension de puissance finie qui ne dépasse pas un nombre naturel donné. Pour pouvoir mieux formuler cette condition, employons la terminologie suivante: $B_1, \ldots, B_n$ étant une chaîne ascendante dans l'espace de probabilité conditionnelle $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$ (c'est-à-dire si l'on a $\mathbf{P}(B_i B_{i+1} | B_i) > 0$ pour $i = 1, \ldots, n-1$, mais si l'égalité $\mathbf{P}(B_i B_{i+1} | B_{i+1}) = 0$ est valable pour au moins un indice i), nous dirons qu'il y a un *saut* entre B_i et B_{i+1}, si l'on a $\mathbf{P}(B_i B_{i+1} | B_{i+1}) = 0$; toute chaîne ascendante contient donc au moins un saut.

Voici la condition dont nous venons de parler:

(6. 2) *Pour qu'un espace de probabilité conditionnelle $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$ puisse être engendré par une famille de mesures ordonnée suivant la dimension et composée d'un nombre de mesures qui ne dépasse pas le nombre naturel N, il faut et il suffit que cet espace remplisse l'axiome IV_3 et que le nombre des sauts ne dépasse le nombre $N-1$ dans aucune chaîne ascendante de l'espace.*

DÉMONSTRATION. *Nécessité.* La nécessité de la validité de IV_3 étant la conséquence de (5. 9), remarquons que si la famille $\{\mu_\alpha\}$ de mesures est ordonnée suivant la dimension et si elle engendre l'espace $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$, les conditions $B \in \mathfrak{B}$, $B' \in \mathfrak{B}$, $\mathbf{P}(BB'|B) > 0$, $\mathbf{P}(BB'|B') = 0$, $0 < \mu_\alpha(B) < +\infty$ entraînent $\mu_\alpha(BB') > 0$, $\mu_\beta(BB') = 0$, donc $\alpha < \beta$. Par conséquent s'il y avait au moins N sauts dans une chaîne ascendante $B_1, \ldots, B_n$, en faisant corres-

pondre à tout indice i ($i = 1, \ldots, n$) une mesure μ_{α_i} telle que $0 < \mu_{\alpha_i}(B_i) < +\infty$, on aurait $\alpha_1 \leqq \alpha_2 \cdots \leqq \alpha_n$ et le signe d'inégalité serait valable au moins N fois, de sorte que le nombre des indices différents, donc celui des mesures différentes, serait au moins $N+1$.

Suffisance. En répétant le raisonnement de la démonstration de (5. 3), on constate que, dans l'ensemble partiellement ordonnée $\{\mathfrak{S}_\alpha\}$ des classes d'équivalence, le rang d'aucun élément ne dépasse le nombre N. Donc, d'après (6. 1), l'espace $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ peut être engendré par une famille de mesures ordonnée suivant la dimension et ne contenant que N mesures au plus.

Pour pouvoir répondre à la troisième question principale, il ne faut qu'appliquer la proposition (6. 2) au cas $N = 1$. La condition qu'aucune chaîne ascendante ne peut contenir que $N-1$ sauts, veut dire en ce cas que, dans l'espace $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$, il n'y a point de chaîne ascendante, ou ce qui est la même chose, que les relations $\mathbf{P}(BB'|B) > 0$, $\mathbf{P}(BB'|B') = 0$ ne peuvent pas être valables simultanément pour $B \in \mathfrak{B}$, $B' \in \mathfrak{B}$. Toute famille de mesures se composant d'une seule mesure étant ordonnée suivant la dimension, on parvient à la proposition suivante qui répond à la troisième question principale :

(6. 3) *Pour qu'un espace de probabilité conditionnelle* $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *puisse être engendré par une seule mesure, il faut et il suffit qu'il remplisse l'axiome* IV_3 *et que, pour* $B \in \mathfrak{B}$, $B' \in \mathfrak{B}$, *on ait ou bien* $\mathbf{P}(BB'|B) = \mathbf{P}(BB'|B') = 0$, *ou bien* $\mathbf{P}(BB'|B) > 0$, $\mathbf{P}(BB'|B') > 0$.

Il est aisé de voir qu'en écrivant, dans (6. 2) et (6. 3), espace général de probabilité conditionnelle au lieu d'espace de probabilité conditionnelle, les propositions en question restent valables.

Il faut encore remarquer que l'axiome IV_3 peut être remplacé par IV_2 dans (6. 3). En effet, d'une part IV_3 entraîne IV_2, de l'autre si, pour $B \in \mathfrak{B}$, $B' \in \mathfrak{B}$, on a ou bien $\mathbf{P}(BB'|B) = \mathbf{P}(BB'|B') = 0$, ou bien $\mathbf{P}(BB'|B) > 0$, $\mathbf{P}(BB'|B') > 0$, les ensembles $B_1, \ldots, B_n$ ($B_i \in \mathfrak{B}$) forment ou bien un cycle, ou bien on a pour au moins un indice i $\mathbf{P}(B_iB_{i+1}|B_i) = \mathbf{P}(B_iB_{i+1}|B_{i+1}) = 0$ ($B_{n+1} = B_1$) et en ce cas (3. 2) entraîne évidemment (3. 1).

7. Dans les raisonnements qui précèdent, nous avons introduit en dehors des axiomes I, II et III_1 qui servent de fondement les axiomes III_2, IV_1, IV_2 et IV_3. Nous savons qu'entre ces axiomes les relations logiques suivantes subsistent (en admettant toujours I, II et III_1) :

$$\mathrm{IV}_3 \rightarrow \mathrm{IV}_2 \begin{array}{l} \nearrow \mathrm{IV}_1 \\ \searrow \mathrm{III}_2 \end{array}.$$

Nous allons démontrer par des exemples très simples qu'en dehors des implications énumérées aucune n'existe entre les axiomes en question. Dans

les exemples qui vont suivre, le symbole $\mathfrak{L}(a, b)$ désignera la famille des sous-ensembles mesurables au sens de Lebesgue de l'intervalle (a, b), et $|X|$ désignera la mesure de X.

a) IV_2 est valable sans que IV_3 le soit dans l'exemple suivant: $S=(0, 3)$, $\mathcal{A}=\mathfrak{L}(0, 3)$, $\mathfrak{B}=\{B_1, B_2, B_3\}$ où $B_1=(0, 2)$, $B_2=(1, 3)$, $B_3=(0, 1)+(2, 3)$; $\mathbf{P}(A|B_i)=\frac{1}{2}|AB_i|$ pour $A\in\mathcal{A}$ et $i=1, 2$; $\mathbf{P}(A|B_3)=|A\cdot(2, 3)|$.

IV_2 est valable puisque cet espace est engendré par la famille $\{\mu_1, \mu_2, \mu_3\}$ de mesures où

$$\mu_i(A)=\begin{cases}|A| & \text{pour } A\subset B_i,\\ +\infty & \text{pour } A-B_i\neq 0\end{cases}\qquad (i=1, 2)$$

et

$$\mu_3(A)=\begin{cases}|A\cdot(2, 3)| & \text{pour } A\subset B_3,\\ +\infty & \text{pour } A-B_3\neq 0,\end{cases}$$

famille qui satisfait évidemment aux conditions posées dans le théorème (4. 2). IV_3 est en défaut puisque

$$\mathbf{P}(B_1B_2|B_1)\,\mathbf{P}(B_2B_3|B_2)\,\mathbf{P}(B_3B_1|B_3)=\frac{1}{2}\cdot\frac{1}{2}\cdot 0$$

tandis que

$$\mathbf{P}(B_1B_2|B_2)\,\mathbf{P}(B_2B_3|B_3)\,\mathbf{P}(B_3B_1|B_1)=\frac{1}{2}\cdot 1\cdot\frac{1}{2}.$$

b) IV_1 est valable sans que IV_2 le soit dans l'exemple suivant: $S=\{a, b, c\}$, $\mathcal{A}$ est la classe de tous les sous-ensembles de S, $\mathfrak{B}=\{C_1, C_2\}$, où $C_1=\{a, b\}$ et $C_2=S$; $\mathbf{P}(A|C_1)=p(A\cdot\{a\})$, $\mathbf{P}(A|C_2)=\frac{1}{2}p(A\cdot\{b, c\})$, où $p(X)$ désigne le nombre des éléments de l'ensemble X.

IV_1 est valable, car si (3. 2) est rempli et si l'un et l'autre membres de (3. 1) sont positifs, on a nécessairement, pour tout indice i tel que $\alpha)$ $B_i=C_1$, $B_{i+1}=C_2$ ou que $\beta)$ $B_i=C_2$, $B_{i+1}=C_1$, l'égalité $A_i=C_1$, tandis que pour $\gamma)$ $B_i=B_{i+1}$ on a $\mathbf{P}(A_i|B_i)=\mathbf{P}(A_i|B_{i+1})$, de sorte qu'on peut supprimer des deux membres de (3. 1) les facteurs qui correspondent aux indices i du type $\gamma)$ ce qui donne l'égalité

$$\mathbf{P}(C_1|C_1)^\alpha\,\mathbf{P}(C_1|C_2)^\beta=\mathbf{P}(C_1|C_2)^\alpha\,\mathbf{P}(C_1|C_1)^\beta,$$

où α et β désignent le nombre des indices du type $\alpha)$ et de celui $\beta)$ respectivement, égalité valable en conséquence de la relation évidente $\alpha=\beta$. Cependant IV_2 est en défaut puisque

$$\mathbf{P}(\{a\}|C_1)\,\mathbf{P}(\{b\}|C_2)=1\cdot\frac{1}{2}$$

et que

$$\mathbf{P}(\{a\}|C_2)\,\mathbf{P}(\{b\}|C_1)=0\cdot 0.$$

Le même exemple b) montre que IV_1 n'entraîne pas III_2:

$$1 \cdot \frac{1}{2} = \mathbf{P}(\{a\}|C_1)\,\mathbf{P}(C_1|C_2) + \mathbf{P}(\{a\}|C_2) = 0.$$

c) III_2 est valable sans que IV_1 le soit dans l'exemple suivant: $S=(0,3)$, $\mathfrak{A}=\mathfrak{L}(0,3)$, $\mathfrak{B}=\{B_1, B_2\}$ où $B_1=(0,2)$, $B_2=(1,3)$; $\mathbf{P}(A|B_1)=\frac{1}{2}|AB_1|$, $\mathbf{P}(A|B_2)=\frac{1}{4}\int\limits_{AB_2} x\,dx$.

La validité de III_2 est la conséquence de ce que cet espace est engendré par la famille $\{\mu_1, \mu_2\}$ de mesures où

$$\mu_1(A) = \begin{cases} |A| & \text{pour } A \subset B_1, \\ +\infty & \text{pour } A - B_1 \neq 0 \end{cases}$$

et

$$\mu_2(A) = \begin{cases} \int\limits_A x\,dx & \text{pour } A \subset B_2, \\ +\infty & \text{pour } A - B_2 \neq 0. \end{cases}$$

On a cependant pour $A_1=\left(1, \frac{3}{2}\right)$, $A_2=(1,2)$

$$\frac{1}{4}\cdot\frac{3}{8} = \mathbf{P}(A_1|B_1)\ \mathbf{P}(A_2|B_2) + \mathbf{P}(A_1|B_2)\ \mathbf{P}(A_2|B_1) = \frac{5}{32}\cdot\frac{1}{2}.$$

d) III_2 et IV_1 sont valables sans que IV_2 le soit dans l'exemple suivant: $S=\{a,b,c,d\}$; $\mathfrak{A}$ se compose de tous les sous-ensembles de S; $\mathfrak{B}=\{D_1, D_2\}$ où $D_1=\{a,b,c\}$ et $D_2=\{b,c,d\}$; $\mathbf{P}(A|D_1)=\frac{1}{2}p(A\cdot\{a,b\})$, $\mathbf{P}(A|D_2)=\frac{1}{2}p(A\cdot\{c,d\})$ où $p(X)$ désigne de nouveau le nombre des éléments de X.

III_2 est valable puisque cet espace est engendré par la famille $\{\mu_1, \mu_2\}$ de mesures telles que

$$\mu_1(a)=1,\ \mu_1(b)=1,\ \mu_1(c)=0,\ \mu_1(d)=+\infty,$$
$$\mu_2(a)=+\infty,\ \mu_2(b)=0,\ \mu_2(c)=1,\ \mu_2(d)=1.$$

On constate aisément en appliquant (3.3) que IV_1 est également valable. Tout de même IV_2 est en défaut:

$$\frac{1}{2}\cdot\frac{1}{2} = \mathbf{P}(\{b\}|D_1)\ \mathbf{P}(\{b,c\}|D_2) + \mathbf{P}(\{b\}|D_2)\ \mathbf{P}(\{b,c\}|D_1) = 0\cdot\frac{1}{2}.$$

Les axiomes IV_1, IV_2 et IV_3 montrent la particularité qu'ils se rapportent à un nombre indéterminé d'ensembles de $\mathfrak{B}$, tandis que III_2 ne concerne qu'un nombre déterminé d'ensembles. On pourrait se figurer que, peut-être, il suffirait de postuler les axiomes IV_1, IV_2 et IV_3 pour un nombre déterminé

(suffisamment élevé) d'ensembles. Pour être plus précis, désignons par $IV_k^{(n)}$ ($k = 1, 2, 3$; $n = 2, 3, \ldots$) l'axiome qui exige que les conditions (3. 2) (et les conditions supplémentaires correspondant aux axiomes IV_1 ou IV_2, si $k = 1$ ou 2) entraînent l'égalité (3. 1) *pour un nombre naturel n donné.* Nous allons étudier dans ce qui suit les relations d'implication entre les axiomes III_2, IV_k et $IV_k^{(n)}$ (en admettant toujours I, II et III_1).

On a évidemment

$$IV_k \to IV_k^{(n)} \qquad (k = 1, 2, 3;\ n = 2, 3, \ldots)$$

et

$$IV_3^{(n)} \to IV_2^{(n)} \qquad (n = 2, 3, \ldots),$$

tandis que la démonstration de la proposition (4. 1) fournit les implications

$$IV_2^{(2)} \to III_2 \quad \text{et} \quad IV_2^{(n)} \to IV_1^{(n)} \qquad (n = 2, 3, \ldots).$$

Il est aisé de voir que

$$IV_k^{(n)} \to IV_k^{(n-1)} \qquad (k = 1, 2, 3;\ n = 3, 4, \ldots).$$

En effet, si l'on a $B_i \in \mathfrak{B}$, $A_i \in \mathfrak{A}$, $A_i \subset B_i B_{i+1}$ ($i = 1, \ldots, n-1$; $B_n = B_1$) (et si, pour $k = 1$, les conditions $\mathbf{P}(A_i|B_i) > 0$, $\mathbf{P}(A_i|B_{i+1}) > 0$ sont valables pour $i = 1, \ldots, n-1$, ou, pour $k = 2$, si $B_1, \ldots, B_{n-1}$ est un cycle), on peut poser $B_i' = B_i$ pour $i = 1, \ldots, n-1$, $B_n' = B_{n+1}' = B_1$, $A_i' = A_i$ pour $i = 1, \ldots, n-1$ et $A_n' = B_1$. On aura en ce cas en vertu de $IV_k^{(n)}$ (même pour $k = 1$ ou 2)

$$\prod_{i=1}^{n} \mathbf{P}(A_i'|B_i') = \prod_{i=1}^{n} \mathbf{P}(A_i'|B_{i+1}'),$$

ou ce qui veut dire la même chose

$$\prod_{i=1}^{n-1} \mathbf{P}(A_i|B_i) \cdot \mathbf{P}(B_1|B_1) = \prod_{i=1}^{n-1} \mathbf{P}A_i|B_{i+1}) \cdot \mathbf{P}(B_1|B_1),$$

par conséquent $IV_k^{(n-1)}$ est valable.

On a enfin

$$IV_2^{(2)} \to IV_3^{(2)},$$

car si B_1 et B_2 ($B_i \in \mathfrak{B}$, $i = 1, 2$) ne forment pas un cycle, on a $\mathbf{P}(B_1 B_2|B_1) = 0$ ou $\mathbf{P}(B_1 B_2|B_2) = 0$, donc pour $A_i \in \mathfrak{A}$, $A_i \subset B_1 B_2$ ($i = 1, 2$)

$$\mathbf{P}(A_1|B_1)\,\mathbf{P}(A_2|B_2) = \mathbf{P}(A_1|B_2)\,\mathbf{P}(A_2|B_1) = 0.$$

D'après ce que nous venons d'établir, le tableau suivant est valable:

$$\begin{array}{ccccccc}
IV_3 & \to & \cdots & \to & IV_3^{(3)} & \to & IV_3^{(2)} \\
\downarrow & & & & \downarrow & & \updownarrow \\
IV_2 & \to & \cdots & \to & IV_2^{(3)} & \to & IV_2^{(2)} \to III_2. \\
\downarrow & & & & \downarrow & & \downarrow \\
IV_1 & \to & \cdots & \to & IV_1^{(3)} & \to & IV_1^{(2)}
\end{array}$$

Nous allons montrer qu'il n'existe pas d'autres relations d'implication entre les axiomes en question. En effet, l'exemple a) montre que IV_2 n'entraîne

pas $IV_3^{(3)}$, l'exemple b) montre que IV_1 n'implique pas $IV_2^{(2)}$, l'exemple c) montre que III_2 peut être valable sans que $IV_1^{(2)}$ le soit, et celui d) montre que III_2 et IV_1 peuvent être valables sans que $IV_2^{(2)}$ le soit. Pour démontrer notre proposition, il faut encore montrer que $IV_3^{(n)}$ n'entraîne pas $IV_1^{(n+1)}$ $(n=2,3,\ldots)$, ce qui s'ensuit de l'exemple suivant:

e) $S=(0,\,n+1)$, $\mathfrak{A}=\mathfrak{L}(0,\,n+1)$, $\mathfrak{B}=\{C_1,\ldots,C_{n+1}\}$ où $C_i=$ $=(i-1,\,i+1)$ $(i=1,\ldots,n)$, $C_{n+1}=(0,1)+(n,\,n+1)$; $\mathbf{P}(A|C_i)=\frac{1}{2}|A\,C_i|$ pour $A\in\mathfrak{A}$ et $i=1,\ldots,n$, $\mathbf{P}(A|C_{n+1})=\frac{1}{3}|A\cdot(0,1)|+\frac{2}{3}|A\cdot(n,n+1)$ pour $A\in\mathfrak{A}$.

$IV_1^{(n+1)}$ est en défaut:

$$\left(\frac{1}{2}\right)^n\cdot\frac{1}{3}=\mathbf{P}(C_1C_2|C_1)\cdots\mathbf{P}(C_nC_{n+1}|C_n)\,\mathbf{P}(C_{n+1}C_1|C_{n+1})\neq$$

$$\neq\mathbf{P}(C_1C_2|C_2)\cdots\mathbf{P}(C_nC_{n+1}|C_{n+1})\,\mathbf{P}C_{n+1}C_1|C_1)=\left(\frac{1}{2}\right)^{n-1}\cdot\frac{2}{3}\cdot\frac{1}{2}.$$

Pour démontrer que cet espace satisfait à $IV_3^{(n)}$, considérons des ensembles $A_i\in\mathfrak{A}$, $B_i\in\mathfrak{B}$, $A_i\subset B_iB_{i+1}$ $(i=1,\ldots,n;\,B_{n+1}=B_1)$. La formule (3.1) étant évidemment valable s'il y a un indice i tel que $B_iB_{i+1}=0$, on peut supposer qu'on a $B_iB_{i+1}\neq 0$, donc que

(7.1) $\qquad B_i=C_m$ *entraîne* $B_{i+1}=C_m$, *ou* C_{m+1}, *ou* C_{m-1} $(i=1,\ldots,n;\,C_0=C_{n+1},\,C_{n+2}=C_1)$.

Si $B_i\neq C_{n+1}$ pour $i=1,\ldots,n$, la formule (3.1) se réduit à

$$\left(\frac{1}{2}\right)^n\prod_{i=1}^{n}|A_i|=\left(\frac{1}{2}\right)^n\prod_{i=1}^{n}|A_i|$$

ce qui est valable, de sorte qu'on peut supposer qu'on a $B_i=C_{n+1}$ pour au moins un indice i. Soient $1\leqq i_1<i_2<\cdots<i_r\leqq n$ les indices de ce type. En posant $B_0=B_n$, on a en vertu de (7.1)

$$B_{i_k\pm1}=C_{n+1},\ \text{ou}\ C_1,\ \text{ou}\ C_n\qquad(k=1,\ldots,r).$$

(7.2) *Si* $B_{i_k+1}=C_m$ $(m=1$ *ou* $n)$, *on a* $B_{i_{k+1}-1}=C_m$ $(i_{r+1}=i_1)$.

En effet, $B_{i_{k+1}-1}=C_{n+1}$ étant impossible d'après la définition des i_k, on a $B_{i_{k+1}-1}=C_1$ ou C_n; mais l'hypothèse $B_{i_{k+1}-1}\neq B_{i_k+1}$ est impossible, car si, disons, $B_{i_k+1}=C_1$ et $k<r$, on a d'après (7.1) $B_{i_{k+1}-1}=C_m$ où $m\leqq i_{k+1}-$ $-1-i_k\leqq n-1$ et un raisonnement analogue s'applique au cas où $B_{i_k+1}=$ $=C_n$ et pour $k=r$. En tenant compte de ce qu'aux indices i tels que $B_i=$ $=B_{i+1}=C_{n+1}$, correspondent des facteurs égaux dans les produits

(7.3) $$\prod_{i=1}^{n}\mathbf{P}(A_i|B_i)\quad\text{et}\quad\prod_{i=1}^{n}\mathbf{P}(A_i|B_{i+1}),$$

on peut supprimer ces facteurs de l'un et de l'autre membres de (3. 1). En désignant par α le nombre des indices i tels que $B_i = C_1$, $B_{i+1} = C_{n+1}$, par β celui des indices tels que $B_i = C_n$, $B_{i+1} = C_{n+1}$, par α' celui des indices tels que $B_i = C_{n+1}$, $B_{i+1} = C_1$, par β' celui des indices tels que $B_i = C_{n+1}$, $B_{n+1} = C_n$ et par γ le nombre des indices tels que $B_i \neq C_{n+1}$, $B_{i+1} \neq C_{n+1}$, les produits (7. 3) prennent, après avoir supprimé les facteurs dont nous avons parlé au commencement, la forme

$$(7.4) \qquad \prod' |A_i| \cdot \left(\frac{1}{2}\right)^{\gamma} \cdot \left(\frac{1}{2}\right)^{\alpha} \cdot \left(\frac{1}{2}\right)^{\beta} \cdot \left(\frac{1}{3}\right)^{\alpha'} \cdot \left(\frac{2}{3}\right)^{\beta'}$$

et

$$(7.5) \qquad \prod' |A_i| \cdot \left(\frac{1}{2}\right)^{\gamma} \cdot \left(\frac{1}{3}\right)^{\alpha} \cdot \left(\frac{2}{3}\right)^{\beta} \cdot \left(\frac{1}{2}\right)^{\alpha'} \cdot \left(\frac{1}{2}\right)^{\beta'},$$

où on a désigné par $\prod'$ le produit extendu sur les indices que nous n'avons pas supprimés. La proposition (7. 2) a pour conséquence que $\alpha = \alpha'$ et que $\beta = \beta'$, de sorte que les produits (7. 4) et (7. 5) sont égaux, ce qui était encore à démontrer.

8. Jusqu'ici nous n'avons fait aucune hypothèse concernant la structure de la famille $\mathfrak{B}$. En introduisant des hypothèses de cette sorte, on peut réduire considérablement les conditions d'autre nature. On doit, en particulier, à M. A. RÉNYI la remarque que *si la famille* $\mathfrak{B}$ *est additive* (c'est-à-dire si elle contient avec B_1 et B_2 l'ensemble $B_1 + B_2$), *l'axiome* IV_3 *est la conséquence de* I, II, III_1 *et* III_2. De façon plus générale, soit $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$ un espace de probabilité conditionnelle ; nous dirons que la famille $\mathfrak{B}$ est *quasi-additive*, si à deux ensembles quelconques $B_1 \in \mathfrak{B}$ et $B_2 \in \mathfrak{B}$ correspond au moins un ensemble $B \in \mathfrak{B}$ tel que

$$(8.1) \qquad B_1 + B_2 \subset B \quad \text{et} \quad \mathbf{P}(B_1|B) + \mathbf{P}(B_2|B) > 0;$$

un ensemble B de ce type sera désigné par $B_1 \circ B_2$. Si la famille $\mathfrak{B}$ est additive, elle est quasi-additive puisque, en posant $B_1 \circ B_2 = B_1 + B_2$, les conditions (8. 1) sont valables (on tient compte de ce que $\mathbf{P}(B_1|B_1 + B_2) + \mathbf{P}(B_2|B_1 + B_2) \geqq \mathbf{P}(B_1 + B_2|B_1 + B_2) = 1$).

On démontre alors par un raisonnement analogue à celui de M. A. RÉNYI la proposition

(8. 2) $[S, \mathfrak{A}, \mathfrak{B}, \mathbf{P}]$ *étant un espace de probabilité conditionnelle, si la famille* $\mathfrak{B}$ *est quasi-additive, cet espace satisfait à l'axiome* IV_3.

DÉMONSTRATION. Soient $B_1, \ldots, B_n$ des ensembles de la famille $\mathfrak{B}$ et posons

$$B^1 = B_1, \quad B^{i+1} = B^i \circ B_{i+1} \quad \text{pour} \quad i = 1, \ldots, n-1$$

et

$$B = B^n.$$

Eu égard à (8. 1), on a $B^i \subset B^k$ et $B_i \subset B^k$ pour $i, k = 1, \ldots, n$ et $i \leqq k$. Également d'après (8. 1), on a ou bien $\mathbf{P}(B_n|B) > 0$, ou bien $\mathbf{P}(B^{n-1}|B) > 0$; en ce cas on a ou bien $\mathbf{P}(B_{n-1}|B^{n-1}) > 0$, donc en vertu de III_2 $\mathbf{P}(B_{n-1}|B) = \mathbf{P}(B_{n-1}|B^{n-1})\,\mathbf{P}(B^{n-1}|B) > 0$, ou bien $\mathbf{P}(B^{n-2}|B^{n-1}) > 0$, donc $\mathbf{P}(B^{n-2}|B) = \mathbf{P}(B^{n-2}|B^{n-1})\,\mathbf{P}(B^{n-1}|B) > 0$. En poursuivant ce raisonnement, on aboutit au résultat qu'au moins une des valeurs $\mathbf{P}(B_i|B)$ $(i = 1, \ldots, n)$ doit être positive.

Soient maintenant A_i $(i = 1, \ldots, n)$ des ensembles appartenant à $\mathcal{A}$ et tels que $A_i \subset B_i B_{i+1}$ $(B_{n+1} = B_1)$. Si au moins un des nombres $\mathbf{P}(B_i|B)$ s'annule, on a au moins un indice i_1 tel que $\mathbf{P}(B_{i_1}|B) = 0$, $\mathbf{P}(B_{i_1+1}|B) > 0$ et aussi au moins un indice i_2 tel que $\mathbf{P}(B_{i_2}|B) > 0$, $\mathbf{P}(B_{i_2+1}|B) = 0$, en tenant toujours compte de la convention $B_{n+1} = B_1$. On a donc en ce cas d'après III_2

$$\mathbf{P}(A_{i_1}|B_{i_1+1}) \leqq \mathbf{P}(B_{i_1} B_{i_1+1}|B_{i_1+1}) = \frac{\mathbf{P}(B_{i_1} B_{i_1+1}|B)}{\mathbf{P}(B_{i_1+1}|B)} \leqq \frac{\mathbf{P}(B_{i_1}|B)}{\mathbf{P}(B_{i_1+1}|B)} = 0,$$

donc $\mathbf{P}(A_{i_1}|B_{i_1+1}) = 0$ et on établit de la même façon l'égalité $\mathbf{P}(A_{i_2}|B_{i_2}) = 0$. L'un et l'autre membres de l'égalité

$$\prod_{i=1}^{n} \mathbf{P}(A_i|B_i) = \prod_{i=1}^{n} \mathbf{P}(A_i|B_{i+1}) \tag{8. 3}$$

s'annulent donc en ce cas et (8. 3) est valable.

D'autre part si $\mathbf{P}(B_i|B) > 0$ pour $i = 1, \ldots, n$, on a en vertu de III_2

$$\mathbf{P}(A_i|B_i) = \frac{\mathbf{P}(A_i|B)}{\mathbf{P}(B_i|B)} \quad \text{et} \quad \mathbf{P}(A_i|B_{i+1}) = \frac{\mathbf{P}(A_i|B)}{\mathbf{P}(B_{i+1}|B)},$$

donc (8. 3) se réduit à l'égalité évidente

$$\prod_{i=1}^{n} \frac{\mathbf{P}(A_i|B)}{\mathbf{P}(B_i|B)} = \prod_{i=1}^{n} \frac{\mathbf{P}(A_i|B)}{\mathbf{P}(B_{i+1}|B)}.$$

Il s'ensuit de (8. 2) et de (5. 3) le théorème suivant:

(8. 4) $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *étant un espace de probabilité conditionnelle, si la famille* $\mathfrak{B}$ *est quasi-additive (en particulier, si elle est additive), cet espace peut être engendré par une famille de mesures ordonnée suivant la dimension.*[1]

Pour mieux formuler le théorème suivant, convenons de la terminologie suivante: $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ et $[S, \mathcal{A}, \mathfrak{B}^*, \mathbf{P}^*]$ étant deux espaces de probabilité conditionnelle, on dira que le second est un *prolongement* du premier, si $\mathfrak{B}^* \supset \mathfrak{B}$ et si l'on a pour $A \in \mathcal{A}$, $B \in \mathfrak{B}$ l'égalité $\mathbf{P}^*(A|B) = \mathbf{P}(A|B)$.

On démontre alors le théorème

(8. 5) $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *étant un espace de probabilité conditionnelle, les propositions suivantes sont équivalentes:*

[1] M. A. Rényi a donné une démonstration directe de ce théorème (pour le cas d'une famille $\mathfrak{B}$ additive) bien plus simple que celle du théorème général (5. 3).

i) $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *satisfait à l'axiome* IV_3;

ii) $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *peut être engendré par une famille de mesures ordonnée suivant la dimension;*

iii) $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *possède un prolongement* $[S, \mathcal{A}, \mathfrak{B}^*, \mathbf{P}^*]$ *tel que* $\mathfrak{B}^*$ *est additive;*

iv) $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *possède un prolongement* $[S, \mathcal{A}, \mathfrak{B}^*, \mathbf{P}^*]$ *tel que* $\mathfrak{B}^*$ *est quasi-additive.*

DÉMONSTRATION. i) entraîne ii) d'après (5. 3); iii) implique iv) puisque toute famille additive est quasi-additive; iv) entraîne i) car en vertu de (8. 3) il s'ensuit de iv) que $[S, \mathcal{A}, \mathfrak{B}^*, \mathbf{P}^*]$, et alors a fortiori $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ lui aussi, satisfait à IV_3. Il ne reste donc qu'à démontrer que la proposition iii) est la conséquence de ii).

Pour s'en convaincre, considérons une famille $\{\mu_\alpha\}$ de mesures ordonnée suivant la dimension et qui engendre $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$. On peut définir $\mathfrak{B}^*$ comme la famille de tous les ensembles $B \in \mathcal{A}$ tels qu'il y a au moins un indice α pour lequel on a $0 < \mu_\alpha(B) < +\infty$ (la famille $\{\mu_\alpha\}$ étant ordonnée suivant la dimension, il n'y a qu'un seul indice de ce type) et on peut poser pour $A \in \mathcal{A}, B \in \mathfrak{B}^*$

$$\mathbf{P}^*(A|B) = \frac{\mu_\alpha(AB)}{\mu_\alpha(B)},$$

où on a désigné par α un indice tel que $0 < \mu_\alpha(B) < +\infty$. En effet, on constate sans peine que les conditions a), b) et d) du théorème (2. 3) sont remplies pour l'ensemble S, le σ-corps $\mathcal{A}$, la famille d'ensembles $\mathfrak{B}^*$ et la famille de mesures $\{\mu_\alpha\}$. Celles a) et b) étant évidentes, pour vérifier la condition d), on remarque que (2. 4) est valable pour $\alpha = \beta$, tandis que pour $\alpha < \beta$ l'hypothèse $\mu_\alpha(B) < +\infty$ entraîne $\mu_\beta(B) = 0$, donc $\mu_\beta(A) = 0$ de sorte que (2. 4) est valable en ce cas aussi. D'après (2. 3), la famille $\{\mu_\alpha\}$ de mesures engendre donc un espace général de probabilité conditionnelle $[S, \mathcal{A}, \mathfrak{B}^*, \mathbf{P}^*]$ qui est en vertu de (2. 6) un espace de probabilité conditionnelle et qui est évidemment un prolongement de $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$. Or, si $B \in \mathfrak{B}^*$, $B' \in \mathfrak{B}^*$, choisissons les indices α et β de façon qu'on ait $0 < \mu_\alpha(B) < +\infty$, $0 < \mu_\beta(B') < +\infty$ et supposons pour fixer les idées que $\alpha \leqq \beta$. On a alors $\mu_\beta(B) < +\infty$, donc

$$0 < \mu_\beta(B') \leqq \mu_\beta(B + B') \leqq \mu_\beta(B) + \mu_\beta(B') < +\infty,$$

de sorte que $B + B' \in \mathfrak{B}^*$. La famille $\mathfrak{B}^*$ est donc additive, ce qui était encore à démontrer.

Considérons maintenant la condition suivante par rapport à l'espace de probabilité conditionnelle $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$:

(8. 6) *A deux ensembles quelconques* $B_1 \in \mathfrak{B}$, $B_2 \in \mathfrak{B}$, *correspond au moins un ensemble* $B \in \mathfrak{B}$ *tel que* $B_1 + B_2 \subset B$ *et que*

$$\mathbf{P}(B_1|B) > 0, \quad \mathbf{P}(B_2|B) > 0.$$

On a alors la proposition suivante :

(8. 7) $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *étant un espace de probabilité conditionnelle qui satisfait à la condition* (8. 6), *cet espace peut être engendré par une seule mesure.*

DÉMONSTRATION. La condition (8. 6) entraîne évidemment la quasi-additivité de la famille $\mathfrak{B}$, l'axiome IV_3 se vérifie donc en vertu de (8. 2). B_1 et B_2 étant des ensembles quelconques de $\mathfrak{B}$ et B étant l'ensemble qui leur correspond d'après (8. 6), on a en vertu de III_2

$$\mathbf{P}(B_1B_2|B_1) = \frac{\mathbf{P}(B_1B_2|B)}{\mathbf{P}(B_1|B)} \quad \text{et} \quad \mathbf{P}(B_1B_2|B_2) = \frac{\mathbf{P}(B_1B_2|B)}{\mathbf{P}(B_2|B)},$$

de sorte qu'ou bien $\mathbf{P}(B_1B_2|B_1) = \mathbf{P}(B_1B_2|B_2) = 0$, ou bien $\mathbf{P}(B_1B_2|B_1) > 0$, $\mathbf{P}(B_1B_2|B_2) > 0$. La proposition à établir s'ensuit donc de (6. 3).

Le théorème (8. 7) a pour conséquence presque immédiate le théorème suivant dû à M. A. RÉNYI ([1], p. 375 et [2], p. 292) :

(8. 8) $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *étant un espace de probabilité conditionnelle, s'il existe une suite* $C_0, C_1, C_2, \ldots$ *telle que*

1. $C_0 \subset C_1 \subset C_2 \subset \ldots$, $C_r \in \mathfrak{B}$ $(r = 0, 1, 2, \ldots)$,
2. $\mathbf{P}(C_0|C_r) > 0$ $(r = 0, 1, 2, \ldots)$,
3. *à tout ensemble* $B \in \mathfrak{B}$ *correspond au moins un* C_r *tel que* $B \subset C_r$ *et que* $\mathbf{P}(B|C_r) > 0$,

cet espace peut être engendré par une seule mesure.

DÉMONSTRATION. On a d'après III_2 pour $r < s$

$$\mathbf{P}(C_0|C_r)\,\mathbf{P}(C_r|C_s) = \mathbf{P}(C_0|C_s),$$

donc, en vertu de la condition 2, $\mathbf{P}(C_r|C_s) > 0$. Il s'ensuit également d'après III_2 que si $B \subset C_r$ et si $\mathbf{P}(B|C_r) > 0$, on a pour $r < s$

$$\mathbf{P}(B|C_r)\,\mathbf{P}(C_r|C_s) = \mathbf{P}(B|C_s),$$

donc $\mathbf{P}(B|C_s) > 0$. On constate alors en faisant usage de 1 et de 3 que la condition (8. 6) se vérifie en posant $B = C_r$ pour un indice r suffisamment élevé, de sorte que (8. 7) fournit la proposition à démontrer.

On démontre enfin le théorème suivant :

(8. 9) $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *étant un espace de probabilité conditionnelle, les deux propositions suivantes sont équivalentes :*

i) $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *peut être engendré par une seule mesure ;*

ii) $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ *possède un prolongement* $[S, \mathcal{A}, \mathfrak{B}^*, \mathbf{P}^*]$ *qui satisfait à la condition* (8. 6) (en y substituant $\mathfrak{B}^*$ à $\mathfrak{B}$ et $\mathbf{P}^*$ à $\mathbf{P}$).

DÉMONSTRATION. ii) implique d'après (8. 7) que $[S, \mathcal{A}, \mathfrak{B}^*, \mathbf{P}^*]$, et alors $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ lui aussi, peut être engendré par une seule mesure. Réciproquement si $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$ est engendré par la mesure μ, en définissant $\mathfrak{B}^*$ comme

a famille des ensembles $B \in \mathcal{A}$ tels que $0 < \mu(B) < +\infty$ et en posant pour $A \in \mathcal{A}$, $B \in \mathfrak{B}^*$

$$\mathbf{P}^*(A|B) = \frac{\mu(AB)}{\mu(B)},$$

on constate tout comme dans la démonstration de (8.5) que $[S, \mathcal{A}, \mathfrak{B}^*, \mathbf{P}^*]$ est un espace de probabilité conditionnelle et qu'il est un prolongement de $[S, \mathcal{A}, \mathfrak{B}, \mathbf{P}]$. Si $B_1 \in \mathfrak{B}^*$, $B_2 \in \mathfrak{B}^*$, on a en posant $B = B_1 + B_2$

$$0 < \mu(B_1) \leqq \mu(B_1 + B_2) \leqq \mu(B_1) + \mu(B_2) < +\infty,$$

donc $B \in \mathfrak{B}^*$ et

$$\mathbf{P}^*(B_1|B) = \frac{\mu(B_1)}{\mu(B)} > 0, \qquad \mathbf{P}^*(B_2|B) = \frac{\mu(B_2)}{\mu(B)} > 0,$$

de sorte que $[S, \mathcal{A}, \mathfrak{B}^*, \mathbf{P}^*]$ satisfait à la condition (8.6), ce qu'il fallait démontrer.

9. Remarquons encore que si l'on remplace l'hypothèse que $\mathcal{A}$ est un σ-corps par celle que $\mathcal{A}$ est un corps (c'est-à-dire une famille d'ensembles qui contient la réunion et la différence de deux ensembles quelconques qui appartiennent à elle) et que si l'on substitue à l'axiome II l'axiome II_0 suivant :

II_0. *Pour $B \in \mathfrak{B}$ fixé, $\mathbf{P}(A|B)$ est une fonction additive de l'ensemble $A \in \mathcal{A}$, c'est-à-dire que l'on a pour $A = A_1 + A_2$, $A_k \in \mathcal{A}$ $(k = 1, 2)$, $A_1 A_2 = 0$*

$$\mathbf{P}(A|B) = \mathbf{P}(A_1|B) + \mathbf{P}(A_2|B),$$

tous les résultats du présent article restent valables avec des modifications évidentes. En particulier, il faut alors remplacer les mesures par des fonctions non-négatives et additives d'ensemble.

(Reçu le 17 août 1955.)

Ouvrages cités

[1] A. Rényi, A valószínűségszámítás új axiomatikus felépítése, *A MTA Mat. és Fiz. Oszt. Közl.*, **4** (1954), p. 369—427.

[2] A. Rényi, On a new axiomatic theory of probability, *Acta Math. Acad. Sci. Hung.*, **6** (1955), p. 285—335.

[3] E. Szpilrajn—Marczewski, Sur l'extension de l'ordre partiel, *Fund. Math.*, **16** (1930), p. 386—389.

[4] P. R. Halmos, *Measure theory* (New York, 1950).

О СТРУКТУРЕ ПРОСТРАНСТВ УСЛОВНОЙ ВЕРОЯТНОСТИ

А. Часар (Будапешт)

(Резюме)

Пусть S некоторое множество, $\mathcal{A}$ — σ-алгебра, состоящее из некоторых подмножеств множества S, далее пусть $\mathcal{B}$ некоторое непустое подмножество от $\mathcal{A}$ и $\mathbf{P}(A|B)$—функция множеств, определенная для всех $A \in \mathcal{A}$, $B \in \mathcal{B}$; для которого выполняются следующие условия:

I. Если $A \in \mathcal{A}$, $B \in \mathcal{B}$, то $\mathbf{P}(A|B) \geqq 0$ и $\mathbf{P}(B|B) = 1$.

II. При фиксированном $B \in \mathcal{B}$ $\mathbf{P}(A|B)$ счетно-аддитивно на σ-алгебре $\mathcal{A}$.

III. Если $A \in \mathcal{A}$, $B \in \mathcal{A}$, $C \in \mathcal{B}$, $BC \in \mathcal{B}$, то $\mathbf{P}(A|BC)\,\mathbf{P}(B|C) = \mathbf{P}(AB|C)$. В этом случае система $[S, \mathcal{A}, \mathcal{B}, \mathbf{P}]$ называется пространством условной вероятности.

Пусть $\{\mu_\alpha\}$ некоторая система мер, определенная на σ-алгебре $\mathcal{A}$ и предположим что для каждого множества $B \in \mathcal{B}$ найдется по крайней мере один индекс α такой, что $0 < \mu_\alpha(B) < +\infty$ и для всех таких α и для каждого множества $A \in \mathcal{A}$ имеет место равенство

$$\mathbf{P}(A|B) = \frac{\mu_\alpha(AB)}{\mu_\alpha(B)}.$$

В этом случае говорят, что система $\{\mu_\alpha\}$ представляет пространство условной вероятности $[S, \mathcal{A}, \mathcal{B}, \mathbf{P}]$.

Настоящая работа дает необходимое и достаточное условие для того, чтобы пространство условной вероятности $[S, \mathcal{A}, \mathcal{B}, \mathbf{P}]$ может быть представлено системами мер определенного типа. В частности, работа занимается такими системами $\{\mu_\alpha\}$, у которых $\mu_\alpha(A)$ при фиксированном $A \in \mathcal{A}$ может принимать — кроме значения 0 и $+\infty$ — только единственное конечное положительное значение. Далее рассматривается случай димензионально упорядоченных систем мер; система $\{\mu_\alpha\}$ называется димензионально упорядоченной, если множество индексов α—упорядоченное множество и из $\alpha < \beta$ $\mu_\alpha(A) < +\infty$ следует, что $\mu_\beta(A) = 0$. Работа дает также условия для представимости пространства условной вероятности димензионально упорядоченной конечной системы мер, в частности условие представимости единственной мерой. Подробно исследуется и вопрос независимости употребленных условий.

ON STOCHASTIC PROCESSES CONNECTED WITH CERTAIN PHYSICAL RECORDING APPARATUSES

By

L. TAKÁCS (Budapest)

(*Presented by* A. RÉNYI)

§ 1. Introduction

In our paper [4] we have considered the following stochastic process:

Let us consider a Poisson process in the moments t $(0 \leqq t < \infty)$. Let the temporal density of the events be a non-negative, continuous and bounded function $\lambda(t)$ of t. We suppose that each event of the Poisson process generates a signal (a random pulse) with the temporal progress

$$f(u, \chi) = \begin{cases} \chi e^{-\alpha u} & \text{if} \quad u \geqq 0, \\ 0 & \text{if} \quad u < 0, \end{cases} \tag{1}$$

where u denotes the time passed since the beginning of the signal; χ, the amplitude of the signal, is a positive random variable and α is a positive constant. Let the random variable $\eta(t)$ denote, in the moment t, the sum of the values of the signals beginning in the time interval $(0, t)$, i. e.

$$\eta(t) = \sum_{0 \leqq t_n \leqq t} f(t - t_n, \chi_n), \tag{2}$$

where $0 \leqq t_1 < t_2 < \cdots < t_n < \cdots$ are the moments in which signals are beginning in the interval $0 \leqq t < \infty$ and $\chi_1, \chi_2, \ldots, \chi_n, \ldots$ are the amplitudes of these signals. We suppose that $\chi_1, \chi_2, \ldots$ are equidistributed independent random variables with the common distribution function $P(\chi_n \leqq x) = H(x)$. In [4] we have proved the following theorem:

$\Phi(t, \omega)$, the characteristic function of the random variable $\eta(t)$, has the form

$$\Phi(t, \omega) = E\{e^{i\omega\eta(t)}\} = \exp\left\{ -\int_0^t \lambda(t-u)[1 - \varphi(\omega e^{-\alpha u})]\, du \right\} \tag{3}$$

where $\varphi(\omega)$ denotes the characteristic function of the random variable χ_n $(n = 1, 2, \ldots)$, i. e.

$$\varphi(\omega) = E\{e^{i\omega\chi_n}\} = \int_0^\infty e^{i\omega x}\, dH(x). \tag{4}$$

Moreover, in the case of temporal homogeneity of the underlying Poisson process (i. e. if $\lambda(t) \equiv \lambda$ constant) we have shown that the following theorem holds:

If the random variables χ_n $(n = 1, 2, \ldots)$ *have a finite expectation, then the limiting distribution* $\lim_{t\to\infty} P(\eta(t) \leqq x) = F^*(x)$ *exists. The characteristic function* $\Phi^*(\omega)$ *of the distribution function* $F^*(x)$ *is given by*

$$\Phi^*(\omega) = \exp\left(-\frac{\lambda}{\alpha}\int_0^1 \frac{1-\varphi(\omega u)}{u}\,du\right). \tag{5}$$

In the present paper we shall deal with a more general case. We suppose that the time differences $t_n - t_{n-1} = \xi_{n-1}$ $(n = 1, 2, \ldots; t_0 = 0)$ *between the consecutive events in the underlying process are mutually independent non-negative random variables with the same distribution function* $P(\xi_n \leqq x) = G(x)$ *where* $G(x)$ *is arbitrary;* $G(0) < 1$. *Again we suppose that the temporal course of the generated signals is described by the function* (1). *Let* $\eta(t)$ *denote the sum of the values of the signals beginning in the interval* $(0, t)$ *at the moment* t; *this is given again by* (2). *Let* $P(\eta(t) \leqq x) = F(t, x)$ *denote the distribution function of* $\eta(t)$. *In our present paper we shall not deal with the explicit determination of* $F(t, x)$; *we are interested merely in the behaviour of the distribution function* $F(t, x)$ *when* $t \to \infty$.

The process $\eta(t)$ describes a Markov process only in the special case when the sequence $\{t_n\}$ forms a Poisson process. However, in the general case it is true that the moments t_n are the regeneration points of the process. Let us write $\eta(t_n - 0) = \eta_n$ and $P(\eta_n \leqq x) = F_n(x)$. Then the sequence $\{\eta_n\}$ forms a Markov process also in the general case.

In particular, if $G(x) = 1 - e^{-\lambda x}$, i. e. the underlying process is a Poisson process homogeneous in time, then $\lim_{t\to\infty} F(t, x) = F^*(x)$ exists as mentioned above. In case of an arbitrary distribution function $G(x)$, the existence of a limiting distribution $\lim_{t\to\infty} F(t, x) = F^*(x)$ and the determination of $F^*(x)$ will be treated in § 3. First we shall examine the convergence of the sequence $\{F_n(x)\}$ of distribution functions. § 2 contains the determination of the limiting distribution function $\lim_{n\to\infty} F_n(x) = F(x)$. Then the determination of the limiting distribution function $F^*(x)$ will be reduced to that of $F(x)$.

It is possible to give an other definition of the limiting distribution function $F^*(x)$. Suppose that the investigated process $\eta(t)$ lasts since an infinitely long time. In this process we choose a moment t^* "at random" and ask the value of $P(\eta(t^*) \leqq x)$. Precisely, by this we understand, that we select in the interval $(0, t)$ a uniformly distributed random point, then determine the probability in question for this random point and take the limit for $t \to \infty$. We shall establish the equality $P(\eta(t^*) \leqq x) = F^*(x)$. However, this latter interpretation of $F^*(x)$ seems to be more general than the preceding one; actually, the distribution function $F^*(x)$ so defined may exist also in cases

when $\lim_{t\to\infty} F(t,x)$ does not. $F^*(x)$ may be regarded as the distribution function of the equilibrium state.

Finally, in § 4 and § 5 we shall deal with the application of the above processes to some problems occurring in the theory of counters.

§ 2. Determination of the limiting distribution

As it has been said, $\eta(t)$ means in the moment t, the sum of the signals beginning in the time interval $(0, t)$; the course of $\eta(t)$ may be seen in Fig. 1. Let $\eta(t_n-0)=\eta_n$, i. e. the value of the sum of the signals immediately before the beginning of the n-th signal. If the amplitude of the n-th signal equals χ_n, then $\eta(t_n)=\eta_n+\chi_n$; henceforth, $\eta(t)$ decreases exponentially until the moment when the next signal occurs, and we have

$$\eta_{n+1}=(\eta_n+\chi_n)e^{-\alpha\xi_n} \tag{6}$$

where the random variables χ_n and ξ_n are independent of each other and of η_n. Consequently, the sequence $\{\eta_n\}$ of random variables forms a Markov process. In other words, the moments t_n are the regeneration points of the investigated process of non-Markov type.

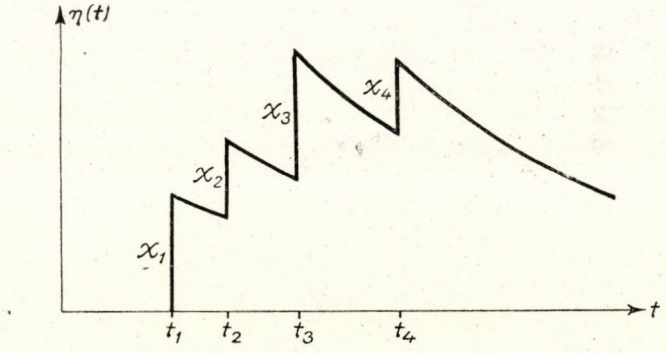


Fig. 1

By (6), for the Markov process $\{\eta_n\}$, we have the following transition probabilities:

$$P(\eta_{n+1}\leqq x\,|\,\eta_n=y)=\int_0^\infty H(xe^{\alpha u}-y)\,dG(u) \qquad (n=1,2,\ldots) \tag{7}$$

as $P(\chi_n\leqq x)=H(x)$ and $P(\xi_n\leqq x)=G(x)$.

With the aid of (7), starting from $P(\eta_1\leqq x)=F_1(x)$, (where $F_1(x)=1$ if $x\geqq 0$ and $F_1(x)=0$ if $x<0$) we can determine the distribution functions $P(\eta_n\leqq x)=F_n(x)$ successively. The computations may be made convenient

with the aid of the following recurrence formulae:

$$K_n(x) = \int_0^\infty F_n(x-y)\,dH(y) \tag{8}$$

and

$$F_{n+1}(x) = \int_0^\infty K_n(xe^{\alpha u})\,dG(u). \tag{9}$$

Here we have $K_n(x) = P(\eta(t_n+0) \leqq x)$, i. e. the distribution function of the sum of the values of the signals immediately after the beginning of the n-th signal.

Relation (6) seems to be extremely suitable for the determination of the moments of the random variables η_n, provided that they exist. By (6) we have

$$E\{\eta_{n+1}^k\} = E\{e^{-\alpha k\xi_n}\}E\{(\eta_n+\chi_n)^k\} = \beta_k \sum_{j=0}^{k}\binom{k}{j} E\{\eta_n^j\}\mu_{k-j} \tag{10}$$

where

$$\beta_k = E\{e^{-\alpha k\xi_n}\} = \int_0^\infty e^{-\alpha kx}\,dG(x) \tag{11}$$

and

$$\mu_j = E\{\chi_n^j\} = \int_0^\infty x^j\,dH(x). \tag{12}$$

Taking the equations (10) for $k = 1, 2, \ldots$ and $n = 1, 2, \ldots$, we obtain a recurrence formula for the determination of the sequence of moments, $E\{\eta_n^k\}$. In particular, for $M_n = E\{\eta_n\}$ we have

$$M_{n+1} = (M_n+\mu)\int_0^\infty e^{-\alpha x}\,dG(x), \tag{13}$$

provided that $\mu = \mu_1 < \infty$.

As $\int_0^\infty e^{-\alpha x}\,dG(x) < 1$, the sequence $\{M_n\}$ converges and

$$M = \lim_{n\to\infty} M_n = \frac{\mu\int_0^\infty e^{-\alpha x}\,dG(x)}{1-\int_0^\infty e^{-\alpha x}\,dG(x)}. \tag{14}$$

REMARK. If the limits $\lim_{n\to\infty} E\{\eta_n^j\} = E_j$ $(j = 1, 2, \ldots)$ exist, then they can be successively determined by the equations

$$E_j = \left[\sum_{k=0}^{j}\binom{j}{k} E_k\mu_{j-k}\right]\beta_j \tag{15}$$

obtained from (10).

Especially, $E_1 = M$ and

$$M = \frac{\mu\beta_1}{1-\beta_1},$$

as we have seen. Let $\sigma^2 = \mu_2 - \mu_1^2$, then we have

$$E_2 = \frac{\mu^2(1+\beta_1)+\sigma^2(1-\beta_1)}{(1-\beta_1)(1-\beta_2)}\beta_2,$$

and the moments of higher order can be determined in a similar way.

Now let us turn our attention to the following question: Under what conditions does the limiting distribution $\lim_{n\to\infty} F_n(x) = F(x)$ exist and how can it be obtained? Here we prove the following

THEOREM 1. *If χ_n $(n = 1, 2, \ldots)$ have a finite expectation $E\{\chi_n\} = \mu$ and $G(0) < 1$, then the sequence $\{F_n(x)\}$ of distribution functions converges to the same limiting distribution function $F(x)$, independently of the starting distribution function $F_1(x)$, as n tends to infinity. This $F(x)$ is the unique solution of the integral equation*

$$F(x) = \int_0^\infty K(x, y)\,dF(y) \tag{16}$$

where

$$K(x, y) = \int_0^\infty H(xe^{\alpha u} - y)\,dG(u) \tag{17}$$

PROOF. First, as before, let $\eta_1 = 0$, i. e. $F_1(x) = 1$ if $x \geqq 0$ and $F_1(x) = 0$ if $x < 0$. We state that the sequence $\{F_n(x)\}$ decreases monotonically if x is fixed. By a repeated application of (6) we easily obtain

$$\eta_n = \chi_1 e^{-\alpha(\xi_1+\ldots+\xi_{n-1})} + \chi_2 e^{-\alpha(\xi_2+\ldots+\xi_{n-1})} + \cdots + \chi_{n-1}e^{-\alpha\xi_{n-1}} \qquad (n = 2, 3, \ldots). \tag{18}$$

Applying (18) for η_{n+1}, we are led to

$$\eta_{n+1} = \chi_1 e^{-\alpha(\xi_1+\ldots+\xi_n)} + \eta_n^* \tag{19}$$

where η_n^* represents the sum of the other terms; it is easy to see that the distribution functions of the random variables η_n and η_n^* are the same. The first term in the right-hand side of (19) is a positive random variable, hence we have $P(\eta_{n+1} \leqq x) \leqq P(\eta_n \leqq x)$, i. e. $F_{n+1}(x) \leqq F_n(x)$ $(n = 2, 3, \ldots)$, as stated. The sequence $\{F_n(x)\}$ is bounded from below $(F_n(x) \geqq 0)$ and is monotonically decreasing, therefore the function $\lim_{n\to\infty} F_n(x) = F(x)$ exists for all x. $F(x)$ is a distribution function. In fact, since $F(0) = 0$, and it is a non-decreasing function of x, we have only to prove $F(\infty) = 1$. By (13) the sequence $M_1 = 0, M_2, M_3, \ldots$ of the expectations $E\{\eta_n\}$ is monotonically increasing and converges to a limit M defined by (15). Let $x > M/\varepsilon$ where $\varepsilon > 0$ is arbitrary. Then we have also $x > M_n/\varepsilon$, and, by the well-known Markov inequality

concerning non-negative random variables, we obtain $F_n(x) \geqq 1-\varepsilon$ for all n. Hence, for $x > M/\varepsilon$ we have $F(x) \geqq 1-\varepsilon$, i. e. $\lim_{x\to\infty} F(x) = 1$, as stated.

Next we prove that the distribution function $F(x)$ is independent of the starting distribution. Up to now we have taken $\eta_1 = 0$; now, if we suppose η_1 is an arbitrary random variable, then the sum of the values of the signals immediately before the occurrence of the n-th event will be given by $\eta_n + \eta_1 \exp(-\alpha(\xi_1 + \cdots + \xi_{n-1}))$ instead of η_n. If n increases, the random variable $\eta_1 \exp(-\alpha(\xi_1 + \cdots + \xi_{n-1}))$ converges in probability to 0; therefore a limiting distribution function exists and agrees with $F(x)$ obtained previously.

In this case the distribution function $F(x)$ has to satisfy the following integral equation:

$$F(x) = \int_0^x P(\eta_{n+1} \leqq x \mid \eta_n = y)\, dF(y) \tag{20}$$

which is equivalent to (16). This fact can be easily established.

Moreover, the integral equation (16) has a unique solution which is a distribution function and is the sought limiting distribution function $F(x)$. Indeed, if there were another solution, also a distribution function in x, then we could take it as the distribution function of η_1. By virtue of equation (6) and the integral equation (16), this distribution function reproduces itself step by step; and so the limiting distribution itself agrees with it. However, we have seen that any distribution function leads always to the same limiting distribution function $F(x)$ independently of the initial distribution function. This contradiction proves the assertion.

We note that the integral equation (16) is equivalent to the following system of integral equations:

$$K(x) = \int_0^\infty F(x-y)\, dH(y) \tag{21}$$

and

$$F(x) = \int_0^\infty K(xe^{\alpha y})\, dG(y). \tag{22}$$

Here $K(x) = \lim_{n\to\infty} P(\eta_n + \chi_n \leqq x)$, i. e. the limiting distribution function of $\eta(t)$, if t tends to infinity through the values $t = t_n + 0$.

Moreover, we note that if at least one of the random variables χ_n and ξ_n has a non-degenerated distribution, then also the limiting distribution function $F(x)$ is a non-degenerated one.

§ 3. Determination of the distribution function $F^*(x)$

Here we shall answer the following question: What are the conditions of the existence of the limiting distribution function $\lim_{t\to\infty} F(t, x) = F^*(x)$ and how can it be obtained? We prove the following theorem:

THEOREM 2. *If χ_n and ξ_n have finite expectations $E(\chi_n) = \mu$ and $E(\xi_n) = \vartheta$, furthermore, for the distribution function $G(x)$ of ξ_n we have $G(0) < 1$, and $G(x)$ is not equal to a distribution function which increases only in jumps at some integer multiples of a certain number, — then there exists a uniquely determined limiting distribution function $\lim_{t\to\infty} F(t, x) = F^*(x)$. This is independent of the starting distribution $F(0, x)$ and can be obtained from*

$$F^*(x) = \frac{1}{\vartheta} \int_0^\infty K(xe^{\alpha u})[1 - G(u)]\, du \tag{23}$$

where $K(x)$ is given by (21).

We remark that by definition (2) we have $\eta_i(0) = 0$, i. e. $F(0, x) = 1$ if $x \geqq 0$ and $F(0, x) = 0$ if $x < 0$. If we suppose that $\eta_i(0) = \eta_{i0}$ where η_{i0} is an arbitrary random variable, then in the definition (2) of $\eta_i(t)$ we must have also an additive term $\eta_{i0}e^{-\alpha t}$.

We introduce a new random variable $\zeta(t)$ as follows: If in the interval $(0, t)$ some event occurs at all, then let $\zeta(t)$ be the difference between the moment t and the moment of the preceding event; in the contrary case let $\zeta(t) = t$. The process $\zeta(t)$ so defined agrees with a special case of the renewal process investigated by J. L. DOOB [2]. Under the conditions of our theorem we can make use of Theorem 12 of DOOB stating that the limiting distribution function

$$G^*(x) = \lim_{t\to\infty} P(\zeta(t) \leqq x) = \frac{1}{\vartheta} \int_0^x [1 - G(u)]\, du \tag{24}$$

exists.

If in the interval $(0, t)$ there occur n events and $t - t_n = u$, then under these conditions let $K_n^*(xe^{\alpha u})$ be the distribution function of $\eta_i(t)$. Introducing the notation $P(\zeta(t) \leqq u) = G(t, u)$, the conditional distribution function of $\eta_i(t)$ relative to the occurrence of n events in the interval $(0, t)$ may be given by

$$\int_0^t K_n^*(xe^{\alpha u})\, d_u G(t, u). \tag{25}$$

Remark that if $t \to \infty$, then $G(t, u) \to G^*(u)$ and $n \to \infty$ with probability 1. In this case, under the conditions expressed, $K_n^*(xe^{\alpha u}) \to K(xe^{\alpha u})$. Using the well-known theorem of E. HELLY we obtain

$$F^*(x) = \lim_{t\to\infty} F(t, x) = \int_0^\infty K(xe^{\alpha u})\, dG^*(u) \tag{26}$$

with probability 1, and this is what we intended to prove. (It agrees with (23).) The uniqueness and the independence from the starting distribution is a consequence of Theorem 1.

If we interprete $F^*(x)$ as the distribution function of the random variable $\eta(t^*)$, when the process $\eta(t)$ is going on since an infinitely long time and t^* is a moment chosen "at random", then $F^*(x)$ is given again by (23). However, we must emphasize that, in view of the latter interpretation, for ξ_n we have only the restriction that the expected value $E(\xi_n) = \vartheta$ be positive and finite. If the limiting distribution $\lim_{t \to \infty} F(t, x)$ exists, then it evidently coincides with $F^*(x)$. However, $F^*(x)$ may also exist when $F(t, x)$ has no limit.

In order to prove our latter statement, first we have to determine the probability that the distance between a point, chosen "at random" in the process going on since an infinitely long time, and the moment of the preceding event is $\leqq x$. This probability is given by the distribution function $G^*(x)$ defined in (24); in fact, this is trivially true if the conditions of Theorem 2 are fulfilled, however, by the aid of the strong law of large numbers, we can establish its validity under the assumption $0 < \vartheta < \infty$. The probability mentioned above will be determined as follows. First, let us consider the process in the time interval (t_1, t_{n+1}). In this interval we choose a point and suppose that it is uniformly distributed in (t_1, t_{n+1}). We determine the probability sought for this interval, then by taking the limit $n \to \infty$ we obtain the solution of the problem. But it is easily seen that we obtain the same result if we take arbitrary intervals $(0, t)$ instead of (t_1, t_{n+1}) and then let $t \to \infty$. In fact, if n events occur in the interval $(0, t)$, then $(t_{n+1} - t_1)/t \to 1$ with probability 1 as $t \to \infty$.

Now let us consider the interval (t_1, t_{n+1}) and let $\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_m}$ be those among the distances $\xi_1, \xi_2, \ldots, \xi_n$ which exceed x. A point chosen at random will have a distance less than x from the preceding event when it lies either in an interval of length $\xi_k < x$ or in the left part of the length x of an interval longer than x. Under these conditions the sought probability will be given by

$$(27) \qquad \frac{\sum_{k=1}^{n} \xi_k - \sum_{k=1}^{m} (\xi_{i_k} - x)}{\sum_{k=1}^{n} \xi_k} = 1 - \frac{m}{n} \frac{\dfrac{\xi_{i_1} + \cdots + \xi_{i_m}}{m} - x}{\dfrac{\xi_1 + \cdots + \xi_n}{n}},$$

since a uniform distribution was assumed.

For values x with $G(x) = 1$ (i. e. the length of each interval is $\leqq x$) this probability equals 1; consequently, $G^*(x) = 1$, in accordance with (24). Suppose now $G(x) < 1$. If $\vartheta = E(\xi_k)$ $(k = 0, 1, 2, \ldots)$, i. e.

$$(28) \qquad \vartheta = \int_0^\infty x \, dG(x) = \int_0^\infty [1 - G(x)] \, dx$$

is finite, then by the strong law of large numbers we have

$$\lim_{n\to\infty}\frac{\xi_1+\xi_2+\cdots+\xi_n}{n}=\vartheta \tag{29}$$

with probability 1. Similarly with probability 1 we have

$$\lim_{m\to\infty}\frac{\xi_{i_1}+\xi_{i_2}+\cdots+\xi_{i_m}}{m}=E(\xi_k\,|\,\xi_k>x)=\frac{1}{1-G(x)}\int_x^\infty u\,dG(u) \tag{30}$$

as

$$P(\xi_k\leqq u\,|\,\xi_k>x)=\frac{G(u)-G(x)}{1-G(x)} \tag{31}$$

and finally

$$\lim_{n\to\infty}\frac{m}{n}=1-G(x). \tag{32}$$

By making use of these relations, we obtain from (27)

$$G^*(x)=\frac{1}{\vartheta}\int_0^x[1-G(u)]\,du. \tag{33}$$

As the formula (33) gives a correct result also for x with $G(x)=1$, therefore this completes the proof.

In case of a process going on since an infinitely long time, a uniformely distributed point chosen "at random" will be with probability 1 farther from the origin than any previously given long time. Thus, in the moment t_n of the event preceding the random point, the probability $P(\eta(t_n-0)\leqq x)$ will be the limiting distribution $F(x)$ with probability 1. Since the distribution function of the distance between this t_n and the random point is equal to $G^*(x)$, consequently, the probability that in a random point $\eta(t^*)\leqq x$ holds, will be given by

$$F^*(x)=\frac{1}{\vartheta}\int_0^\infty K(xe^{\alpha u})[1-G(u)]\,du \tag{34}$$

where, knowing $F(x)$, $K(x)$ may be got by (21). $F^*(x)$ agrees with the limiting distribution $\lim_{t\to\infty}F(t,x)$ provided this exists at all.

Let M^* denote the mean value of $F^*(x)$. By (29) we have

$$M^*=\int_0^\infty x\,dF^*(x)=\frac{M+\mu}{\vartheta}\int_0^\infty e^{-\alpha u}[1-G(u)]\,du=\frac{\mu}{\alpha\vartheta}. \tag{35}$$

Here M is given by (15).

REMARK 1. Let E_j^* denote the j-th moment of $F^*(x)$. If E_j^* exists, then by (23) we have

$$E_j^*=\frac{E_j}{\beta_j}\,\frac{1-\beta_j}{j\alpha\vartheta} \tag{36}$$

where E_j was defined earlier and we used the following relation

$$\frac{1}{\vartheta}\int_0^\infty e^{-\alpha x}[1-G(x)]\,dx = \frac{1-\beta_j}{j\alpha\vartheta}. \tag{37}$$

In particular, for $E_1^* = M^*$ we have

$$M^* = \frac{M}{\beta_1}\frac{1-\beta_1}{\alpha\vartheta} = \frac{\mu}{\alpha\vartheta}$$

and for E_2^*

$$E_2^* = \frac{\sigma^2}{2\alpha\vartheta} + \frac{\mu^2}{2\alpha\vartheta}\frac{1+\beta_1}{1-\beta_2}.$$

REMARK. 2. If $G(x) = 1-e^{-x/\vartheta}$, i. e. t_n's are the moments of the events occurring in a Poisson process characterized by the density $\lambda = 1/\vartheta$, then

$$\frac{1-G(x)}{\vartheta} = G'(x)$$

and so by (22) and (23) we get

$$F^*(x) = F(x).$$

§ 4. Determination of some expectations

Applying the stochastic process $\eta(t)$ investigated by us for physical problems, the following problem arises. Suppose that the process is going on since an infinitely long time, i. e. we have an equilibrium state. We ask in an interval of length t chosen "at random" the average number, of the moments in which $\eta(u)$ passes a fixed threshold a, i. e. $\eta(u-0) \leqq a$, but $\eta(u+0) > a$ — further, for the mean value of the time duration in which $\eta(u) > a$. Let $m^*(t, a)$ and $\tau^*(t, a)$ denote the former and the latter expectations, respectively.

A passing through the threshold a occurs only in moments where a signal is beginning. An easy calculation for the expected number of the signals occurring during the time t gives t/ϑ. A signal starting in the moment u represents also a threshold-passing when before the signal $\eta(u-0) \leqq a$ holds and after it $\eta(u+0) > a$. The probability of this event is $F(a)-K(a)$, hence

$$m^*(t, a) = \frac{[F(a)-K(a)]t}{\vartheta}. \tag{38}$$

The mean value of the sum of the lengths of the intervals lying in the interval of length t in which $\eta(u) > a$, is easily obtained by

$$\tau^*(t, a) = [1-F^*(a)]t. \tag{39}$$

The expectations (38) and (39) surely exist provided we suppose that $0 < \vartheta < \infty$ and μ is finite.

If we define these expectations for an arbitrary interval of length t of a process going on for an infinitely long time, then we must suppose also the fulfilment of the conditions of Theorem 2, namely $\lim_{t\to\infty} \eta(t)$ exists with probability 1 if these conditions are satisfied.

§ 5. Some examples from experimental physics

In our paper [4] we have mentioned as example the case of particle-counting by means of an electron multiplier. If we are counting particles arriving according to a Poisson process with the density λ, then the moments of the voltage impulses served by the electron multiplier form also a Poisson process. So we can apply the discussions given in [4]. But if we lead the pulses into a scaling circuit which registrates every $m+1$-st pulse only, then the starting points of the voltage pulses obtained in this way form no longer a Poisson process. Here the distances between the starting points of two consecutive pulses will have the distribution function

$$G(x) = \int_0^x e^{-\lambda y} \frac{(\lambda y)^m}{m!} \lambda \, dy. \tag{40}$$

In this case we can apply the discussion given above with $\vartheta = (m+1)/\lambda$.

The following case is also interesting: we count particles arriving according to a Poisson process by a Geiger—Müller counter. Here the moments of the voltage impulses generated by the Geiger—Müller counter do not form a Poisson process. If the inoperative time τ of the Geiger—Müller tube is constant, then the distribution function of the distances between the consecutive voltage impulses is given by

$$G(x) = \begin{cases} 0 & \text{if } x \leqq \tau, \\ 1 - e^{-\lambda(x-\tau)} & \text{if } x \geqq \tau. \end{cases} \tag{41}$$

Here $\vartheta = \tau + 1/\lambda$. In this case we can again apply the discussion given previously.

If we know the distribution function $H(x)$ of the values of voltage pulses, and the distribution function $G(x)$ of the distances between the moments of the pulses, furthermore, if we can suppose that the pulses are decreasing in time exponentially with a time constant RC (i. e. $\alpha = 1/RC$), then the present discussion may be applied.

In [4] we have considered the case when the voltage pulses have been led into an amplifier with the threshold voltage a and we have got impulses only when the input voltage jumped to a value exceeding a. If we call the density of the voltage impulses in the amplifier as the *real pulse density* (in our case it is $A = 1/\vartheta$), then we can ask: what will the counter connected

with the amplifier indicate? Let us call it the *apparent density of the events* and denote by $A'(a)$. It depends on the value of the threshold voltage a. By (38) we obtain

$$A'(a)/A = F(a) - K(a) \tag{42}$$

where $K(x)$ is given by (21). Then we must determine only $F(x)$. It can be obtained by solving the integral equation (16).

In the sequel we give some examples for the determination of $F(x)$ by which the apparent density of events can be calculated.

We limit ourselves to the special case when $\chi_n = \mu$ (constant). Then by (21) we have

$$K(x) = F(x-\mu) \tag{43}$$

whence

$$A'(a)/A = F(a) - F(a-\mu). \tag{44}$$

The integral equation to be solved is

$$F(x) = \int_0^\infty F(xe^{\alpha y} - \mu)\, dG(y). \tag{45}$$

Or, taking into account the equation

$$K(x, y) = 1 - G\left(\frac{1}{\alpha} \log \frac{\mu + y}{x}\right), \tag{46}$$

we can write it in the following form equivalent to the preceeding one:

$$F(x) = \int_0^\infty K(x, y)\, dF(y). \tag{47}$$

Case 1. If

$$G(x) = \int_0^x e^{-\lambda y} \frac{(\lambda y)^m}{m!} \lambda\, dy, \tag{48}$$

then by (40)

$$F(x) = \int_0^\infty F(xe^{\alpha y} - \mu) e^{-\lambda y} \frac{(\lambda y)^m}{m!} \lambda\, dy = \frac{\lambda^{m+1} x^{\lambda/\alpha}}{m!\, \alpha^{m+1}} \int_x^\infty \frac{F(t-\mu)}{t^{\lambda/\alpha+1}} \left(\log \frac{t}{x}\right)^m dt = \tag{49}$$

$$= \frac{\lambda^{m+1} x^{\lambda/\alpha}}{m!\, \alpha^{m+1}} \sum_{j=0}^{m} (-1)^{m-j} \binom{m}{j} (\log x)^{m-j} \int_x^\infty \frac{F(t-\mu)}{t^{\lambda/\alpha+1}} (\log t)^j\, dt.$$

Let

$$C_j = \int_\mu^\infty \frac{F(t-\mu)}{t^{\lambda/\alpha+1}} (\log t)^j\, dt \qquad (j = 0, 1, \ldots, m), \tag{50}$$

then (49) can be written as

$$(51)\quad F(x)=\frac{\lambda^{m+1}x^{\lambda/\alpha}}{m!\,\alpha^{m+1}}\sum_{j=0}^{m}(-1)^{m-j}\binom{m}{j}(\log x)^{m-j}\left[C_j-\int_0^{x-\mu}F(t)\frac{(\log(t+\mu))^j}{(t+\mu)^{\lambda/\alpha+1}}\,dt\right].$$

By the aid of this formula we can obtain $F(x)$ for the intervals $(k\mu,(k+1)\mu)$ $(k=0,1,2,\ldots)$ successively. The constants $C_0, C_1, \ldots, C_m$ may be determined from the equations (50) on using the requirement $F(\infty)=1$.

Case 2. Especially, if $m=0$, then $G(x)=1-e^{-\lambda x}$ and we have a process generated by a Poisson one dealt with in [4] in details. Here we obtain the recurrence formula

$$(52)\quad F(x)=\frac{\lambda x^{\lambda/\alpha}}{\alpha}\left[C_0-\int_0^{x-\mu}\frac{F(t)}{(t+\mu)^{\lambda/\alpha+1}}\,dt\right]$$

for the determination of $F(x)$. The constant C_0 may be determined from the requirement $F(\infty)=1$.

If we know the constant C_0 in advance, then by (52) starting from the representation

$$(53)\quad F(x)=\frac{C_0\lambda}{\alpha}x^{\lambda/\alpha}\qquad(0\leqq x\leqq\mu)$$

we should be able to present $F(x)$ in an explicit form for the intervals $(\mu,2\mu),(2\mu,3\mu),\ldots$ successively. Here we give the following determination of C_0.

In our work [4] we have determined the characteristic function $\Phi^*(\omega)$ of $F(x)$ in the general case. Here it is given by formula (5). Applying it to the special case $\chi_n=\mu$ (constant) and passing to the Laplace transform, we obtain

$$(54)\quad \int_0^{\infty}e^{-sx}F(x)\,dx=\frac{1}{s}\exp\left(-\frac{\lambda}{\alpha}\int_0^{1}\frac{1-e^{-s\mu v}}{v}\,dv\right).$$

Suppose $\Re(z)>0$, then by the well-known integral representation of Euler's constant $C=0{,}577215\ldots$ we obtain by a simple modification

$$(55)\quad C=\int_0^{1}\frac{1-e^{-zt}}{t}\,dt-\int_1^{\infty}\frac{e^{-zt}}{t}-\log z.$$

Using it, (54) may be transformed into

$$(56)\quad \int_0^{\infty}e^{-sx}F(x)\,dx=\frac{e^{-\frac{\lambda C}{\alpha}}\mu^{-\frac{\lambda}{\alpha}}}{s^{\lambda/\alpha+1}}\exp\left(-\frac{\lambda}{\alpha}\int_{\alpha}^{\infty}\frac{e^{-st}}{t}\,dt\right)\qquad(\Re(s)>0).$$

Hence, introducing the notation $\gamma = e^C = 1{,}781072\ldots$, we have

$$\int_0^\infty e^{-sx} F(x)\,dx \sim \frac{1}{(\mu\gamma)^{\lambda/\alpha} s^{\lambda/\alpha+1}} \qquad (s \to \infty). \tag{57}$$

Now, from a well-known theorem of Abel type (see [1], p. 503, Theorem 3) it follows that for $x \to 0$

$$F(x) \sim \frac{x^{\lambda/\alpha}}{\Gamma\left(1+\frac{\lambda}{\alpha}\right)(\mu\gamma)^{\lambda/\alpha}}. \tag{58}$$

Comparing (58) with (49), we get the value of C_0, namely

$$\frac{C_0 \lambda}{\alpha} = \frac{1}{\Gamma\left(1+\frac{\lambda}{\alpha}\right)(\mu\gamma)^{\lambda/\alpha}}. \tag{59}$$

We note that if we want to deduce (58), then it is sufficient to show (57) e. g. for all real s. Actually, the truth of (53) implies that of (47) (see [1], p. 473, Theorem 1) and for the determination of the numerical value of the constant factor it suffices to investigate the limit only for special values s.

Thus for the values $0 \leqq x \leqq \mu$ we obtain

$$F(x) = \frac{x^{\lambda/\alpha}}{\Gamma\left(1+\frac{\lambda}{\alpha}\right)(\mu\gamma)^{\lambda/\alpha}}. \tag{60}$$

Starting from (60) we can determine $F(x)$ for the intervals $(\mu, 2\mu), (2\mu, 3\mu), \ldots$ successively by (52). E. g. for $\mu \leqq x \leqq 2\mu$

$$F(x) = \frac{x^{\lambda/\alpha}}{\Gamma\left(1+\frac{\lambda}{\alpha}\right)(\mu\gamma)^{\lambda/\alpha}} \left[1 - \frac{\lambda}{\alpha} \int_0^{x-\mu} \frac{t^{\lambda/\alpha}}{(t+\mu)^{\lambda/\alpha+1}}\,dt\right]. \tag{61}$$

So it is possible to solve the coincidence problem given in our paper [4] for exponentially distributed amplitudes also for constant amplitudes which is very important in practice.

We note that, inverting the Laplace transform (56), $F(x)$ may be obtained in the form

$$F(x) = \frac{1}{\Gamma\left(1+\frac{\lambda}{\alpha}\right)(\mu\gamma)^{\lambda/\alpha}} \left[x^{\lambda/\alpha} + \sum_{n=1}^{\left[\frac{x}{\mu}\right]} \frac{(-1)^n (\lambda/\alpha)^n}{n!} \int_{n\mu}^{x} (x-y)^{\lambda/\alpha} f_n(y)\,dy\right] \tag{62}$$

where

$$f_1(x) = \begin{cases} 0 & \text{if } x < \mu, \\ 1/x & \text{if } x \geqq \mu, \end{cases} \tag{63}$$

and $f_n(x)$ denotes the n-fold convolution of (63) which may be obtained by the following recurrence formula:

$$f_n(x) = \int_{\mu}^{x-(n-1)\mu} \frac{f_{n-1}(x-y)}{y}\, dy \qquad (n = 2, 3, \ldots). \tag{64}$$

From (52) we see that $F(x)$ is absolutely continuous. Hence it may be differentiated for $x > 0$, and for $F'(x)$ we have

$$xF'(x) = \frac{\lambda}{\alpha}[F(x) - F(x-\mu)]. \tag{65}$$

This equation agrees with a recent result of L. PÁL [3] obtained in another way.

For $m = 0$ the integral equation may be written in an other form too. In the special case $G(x) = 1 - e^{-\lambda x}$ by (46) we have

$$K(x, y) = \begin{cases} \left(\dfrac{\mu + y}{x}\right)^{-\lambda/\alpha} & \text{if} \quad x < \mu + y\,, \\ 1 & \text{if} \quad x \geqq \mu + y, \end{cases} \tag{66}$$

hence, for $F(x)$ we obtain the integral equation

$$F(x) = \int_{x-\mu}^{x} \left(\frac{\mu + y}{x}\right)^{-\lambda/\alpha} dF(y) + F(x-\mu). \tag{67}$$

Case 3. If

$$G(x) = \begin{cases} 0 & \text{if} \quad x \leqq \tau, \\ 1 - e^{-\lambda(x-\tau)} & \text{if} \quad x \geqq \tau, \end{cases} \tag{68}$$

then by (46)

$$K(x, y) = \begin{cases} \left(\dfrac{\mu + y}{x}\right)^{-\lambda/\alpha} e^{\lambda\tau} & \text{if} \quad xe^{\alpha\tau} \leqq \mu + y, \\ 1 & \text{if} \quad xe^{\alpha\tau} \geqq \mu + y \end{cases} \tag{69}$$

and the distribution function $F(x)$ may be obtained by solving the integral equation

$$F(x) = e^{\lambda\tau} \int_{xe^{\alpha\tau}-\mu}^{x} \left(\frac{\mu + y}{x}\right)^{-\lambda/\alpha} dF(y) + F(xe^{\alpha\tau} - \mu). \tag{70}$$

On the other hand, from (40) we get

$$F(x) = e^{\lambda\tau} \int_{\tau}^{\infty} F(xe^{\alpha y} - \mu) e^{-\lambda y} \lambda\, dy = \frac{\lambda e^{\lambda\tau} x^{\lambda/\alpha}}{\alpha} \int_{xe^{\alpha\tau}}^{\infty} \frac{F(t-\mu)}{t^{\lambda/\alpha+1}}\, dt \tag{71}$$

or, introducing the constant

$$C_0 = \int_{\mu}^{\infty} \frac{F(t-\mu)}{t^{\lambda/\alpha+1}}\, dt, \tag{72}$$

we have

$$F(x)=\frac{\lambda e^{\lambda t}x^{\lambda/\alpha}}{\alpha}\left[C_0-\int_{\mu}^{xe^{\alpha t}}\frac{F(t-\mu)}{t^{\lambda/\alpha+1}}\,dt\right] \tag{73}$$

for the determination of $F(x)$. Hence we see that $F(x)$ is continuous and by (73) it is also differentiable for $x>0$. By (73), for $F'(x)$ we have

$$xF'(x)=\frac{\lambda}{\alpha}[F(x)-F(xe^{\alpha t}-\mu)]. \tag{74}$$

From (73) we obtain for the values $0\leqq x\leqq \mu e^{-\alpha t}$

$$F(x)=\frac{C_0\lambda e^{\lambda t}x^{\lambda/\alpha}}{\alpha}. \tag{75}$$

Knowing (75), $F(x)$ may be determined successively for the intervals $(\mu e^{-\alpha t}+\cdots+\mu e^{-k\alpha t},\ \mu e^{-\alpha t}+\cdots+\mu e^{-(k+1)\alpha t})$ $(k=1,2,\ldots)$. Consequently, by (73) we can determine $F(x)$ in the whole interval $(0,\ \mu/(e^{\alpha t}-1))$. It is easy to show that $F(x)=1$ if $x\geqq \mu/(e^{\alpha t}-1)$. The constant C_0, unknown previously, may be determined from the latter requirement.

If we pass from the equation (74) to the Laplace transform of $F(x)$, then, using the theorem of Abel type already mentioned, we can show that

$$\frac{C_0\lambda}{\alpha}=\frac{1}{\Gamma\left(1+\frac{\lambda}{\alpha}\right)(\mu\gamma)^{\lambda/\alpha}} \tag{76}$$

where $\gamma=e^C=1{,}781072\ldots$; hence we have

$$F(x)=\frac{e^{\lambda t}x^{\lambda/\alpha}}{\Gamma\left(1+\frac{\lambda}{\alpha}\right)(\mu\gamma)^{\lambda/\alpha}} \tag{77}$$

and

$$F(x)=\frac{e^{\lambda t}x^{\lambda/\alpha}}{\Gamma\left(1+\frac{\lambda}{\alpha}\right)(\mu\gamma)^{\lambda/\alpha}}\left[1-\frac{\lambda e^{\lambda t}}{\alpha}\int_{0}^{xe^{\alpha t}}\frac{(t-\mu)^{\lambda/\alpha}}{t^{\lambda/\alpha+1}}\,dt\right] \tag{78}$$

in the intervals $(0,\ \mu e^{-\alpha t})$ and $(\mu e^{-\alpha t},\ \mu e^{-\alpha t}+\mu e^{-2\alpha t})$, respectively. For the other intervals, $F(x)$ may be also obtained in a form similar to (62).

For the determination of the constant C_0, put

$$\Phi(s)=\int_{0}^{\infty}e^{-sx}\,dF(x)\qquad(\Re(s)\geqq 0). \tag{79}$$

Then, by (74) we have

$$s\Phi'(s)=\frac{\lambda}{\alpha}[e^{-s\mu e^{-\alpha t}}\Phi(se^{-\alpha t})-\Phi(s)]. \tag{80}$$

Hence, integrating $\Phi'(s)/\Phi(s)$, we obtain for a real s

$$\Phi(s) = \exp\left[-\frac{\lambda}{\alpha}\int_0^1 \frac{1-e^{-s\mu e^{-\alpha t}x}}{x}\,dx - \frac{\lambda}{\alpha}\int_0^1 e^{-s\mu e^{-\alpha t}x}\,\frac{1-\dfrac{\Phi(se^{-\alpha t}x)}{\Phi(sx)}}{x}\,dx\right]. \tag{81}$$

It is easy to prove that the second integral figuring on the right-hand side tends to 0 when $s\to\infty$. Hence we have an expression similar to that of (57):

$$\Phi(s) \sim \frac{e^{\lambda t}}{(\mu\gamma)^{\lambda/\alpha}s^{\lambda/\alpha}} \qquad (s\to\infty). \tag{82}$$

Hence, using (75), we can establish the validity of the equation (76) for the constant C_0.

INSTITUTE FOR APPLIED MATHEMATICS
OF THE HUNGARIAN ACADEMY OF SCIENCES

(*Received 12 January 1953*)

Bibliography

[1] G. Doetsch, *Handbuch der Laplace-Transformation,* Bd. 1 (Basel, 1950).
[2] J. L. Doob, Renewal theory from the point of view of the theory of probability, *Transactions Amer. Math. Soc.*, **63** (1948), pp. 422—438.
[3] L. Pál, Néhány megjegyzés véletlen koincidenciák számításával kapcsolatban (to appear in *MTA Alk. Mat. Int. Közl.*).
[4] L. Takács, On secondary processes generated by a Poisson process and their applications in physics, *Acta Math. Acad. Sci. Hung.*, **5** (1954), pp. 203—236.

О СТОХАСТИЧЕСКИХ ПРОЦЕССАХ, СВЯЗАННЫХ С НЕКОТОРЫМИ ФИЗИЧЕСКИМИ РЕГИСТРИРУЮЩИМИ ПРИБОРАМИ

Л. Такач (Будапешт)

(Резюме)

Рассмотрим случайную величину

$$\eta(t) = \sum_{0\leqq t_n\leqq t} f(t-t_n, \chi_n)$$

определенную для $t\geqq 0$. Здесь

$$f(t,x) = \begin{cases} xe^{-\alpha t}, & \text{если } x\geqq 0 \text{ и } t\geqq 0,\\ 0 \text{ в других случаях,}\end{cases}$$

а α положительное постоянное. Предположим, что для последовательности $t_1, t_2, \ldots, t_n, \ldots$ разности времен $t_{n+1}-t_n$ ($n=0,1,2,\ldots$; $t_0=0$) одинаково распределенные независимые случайные величины с общей функцией распределения $G(x)$, а параметры $\chi_1, \chi_2, \ldots, \chi_n, \ldots$

одинаково распределенные независимые случайные величины с общей функцией распределения $H(x)$.

Вышеописанный процесс $\eta(t)$ является марковским процессом только в частном случае, когда последовательность $\{t_n\}$ представляет собой процесс Пуассона. Однако и в общем случае моменты времени t_n представляют собой марковские точки процесса. Пусть $\eta(t_n - 0) = \eta_n$.

Тогда последовательность переменных $\{\eta_n\}$ является марковской цепью.

Определим сначала распределения переменных $\eta(t)$ и η_n. Пусть $P(\eta(t) \leqq x) = F(t, x)$ и $P(\eta_n \leqq x) = F_n(x)$.

В нашем случае $\eta_1 \equiv 0$, но если взять за η_1 произвольную случайную величину, то имеет место следующая теорема:

Теорема 1. Если $E(\chi_n) = \mu < \infty$ и $G(0) < 1$, то независимо от начальной функции распределения существует предельное распределение $\lim_{n \to \infty} F_n(x) = F(x)$, являющаяся однозначно определенным решением интегрального уравнения

$$F(x) = \int_0^\infty K(x, y)\, dF(y),$$

где

$$K(x, y) = \int_0^\infty H(xe^{\alpha u} - y)\, dG(u).$$

В нашем случае $\eta(0) \equiv 0$; если $\eta(0) = \eta_0$, где η_0 произвольная случайная величина, то в вышеприведенном выражении для $\eta(t)$ следует учесть и сумманд вида $\eta_0 e^{-\alpha t}$. В этом случае имеет место следующая теорема:

Теорема 2. Если $E(\chi_n) = \mu < \infty$, $G(x)$ не решетчатое распределение со среднем равным конечному положительному числу ϑ, и $G(0) < 1$, то независимо от начальной функции распределения $F(0, x)$ существует однозначно определенное распределение $\lim_{t \to \infty} F(t, x) = F^*(x)$ представимая в виде

$$F^*(x) = \frac{1}{\vartheta} \int_0^\infty K(xe^{\alpha u})\, [1 - G(u)]\, du,$$

где

$$K(x) = \int_0^x F(x - y)\, dH(y).$$

После этого установим математическое ожидание для числа прохождений $\eta(t)$ через данное значение a, во времени t, в стационарном случае.

В качестве примера применим полученные результаты к решению проблем, связанных с электронным умножителем и со счетчиком Гейгера—Мюллера.

ZUR THEORIE DER GAMMAFUNKTION, DER RIEMANNSCHEN ZETAFUNKTION UND VERWANDTER FUNKTIONEN. I

Von
MIKLÓS MIKOLÁS (Budapest)
(*Vorgelegt von* A. RÉNYI)

Meinem Professor, LEOPOLD FEJÉR *anläßlich seines 75. Geburtstages zugeeignet*

§ 1. Einleitung

Die Gammafunktion und die Riemannsche Zetafunktion, ferner die mit ihnen verknüpften Funktionen zählen wohl zu denjenigen Transzendenten, deren ausgedehnte Anwendung von der Zahlentheorie bis zur theoretischen Physik eine besondere Wichtigkeit verleiht.[1] Es ist also kaum zu bewundern, daß man in der einschlägigen Literatur bis heute nicht nur ganz neue Ergebnisse, sondern auch solche Aufsätze findet, welche die Theorie der genannten Funktionen von einem oder anderen Standpunkt ausgehend wieder ausbauen, ohne auf mehr als methodologische Neuheit Anspruch zu erheben.

Was die Gammafunktion betrifft, so sind in dieser Hinsicht vor allem gewisse Arbeiten von ARTIN, GRONWALL, JENSEN, LERCH, PRINGSHEIM zu erwähnen. PRINGSHEIM war meines Wissens der erste, der sich bei der Herleitung der Haupteigenschaften (im Reellen) konsequent auf didaktische Gesichtspunkte und zwar auf einen Vergleich der Eulerschen und Gaußschen Erklärung von $\Gamma(z)$ stützte: „Wägt man ... die Vorzüge jener beiden Definitionen gegen einander ab, so dürfte ... es sich am meisten empfehlen, zwar von der Integraldefinition auszugehen, sodann aber vor allem die Produktentwicklung daraus abzuleiten und nunmehr die übrige Theorie auf den gewonnenen Produktausdruck zu gründen."[2] Von ihm stammt eine unmittelbare und strenge Methode für die Verwandlung der beiden genannten Ausdrücke ineinander, welche aber durchaus nicht einfach anzusehen ist. JENSEN[3] gab eine ziemlich elementare, auf Funktionalgleichungen und Reihen- bzw. Produktdarstellungen einschränkte Behandlung des Stoffes; diesen schönen Aufsatz entwickelte GRONWALL[4] fort, indem er, nach Vollständigkeit bestrebt, auch die Beziehungen der Gammafunktion zur Integralrechnung (ohne komplexe Kurvenintegrale zu betrachten) systematisch verarbeitete. LERCH beschäf-

[1] Wir wollen gleich z. B. auf die wohlbekannten zusammenfassenden Arbeiten bzw. ausgezeichneten Monographien [42], [26], [8], [39] hinweisen. Für Literaturangaben und geschichtliche Bemerkungen s. auch [30].

[2] Vgl. [34], S. 457.

[3] [12].

[4] [11].

tigte sich in zahlreichen, meistens böhmisch publizierten Arbeiten mit verschiedenen Teilen der Theorie und fügte interessante neue Ergebnisse hinzu. Man verdankt ihm auch solche Aufsätze, in denen die Theorie der Gammafunktion von Anfang an mit originellen Methoden entwickelt wurde. Besonders sind gewisse Lerchsche Entdeckungen über den Zusammenhang der aus $\sum_{\lambda=0}^{\infty}{}' \frac{1}{(z+\lambda)^s}$ durch analytische Fortsetzung entspringende Funktion $\zeta(s, z)$ (die man heute als Hurwitzsche Zetafunktion zu bezeichnen pflegt) und von $\Gamma(z)$, die daher fließende Formel: $\Gamma(z) = \exp\left[\frac{\partial}{\partial s}(\zeta(s, z) - \zeta(s))\right]_{s=0}$, ferner die fundamentalen Leistungen von LERCH betreffs der unvollständigen Gammafunktion hervorzuheben.[5] Etwa vor zwanzig Jahren ist eine Abhandlung von ARTIN[6] erschienen, welche vom bekannten H. Bohr—Mollerupschen Satze, d. h. von der logarithmischer Konvexität der Gammafunktion ausgeht; daher werden die Grundeigenschaften der positiven Funktion $y = \Gamma(x)$ $(x > 0)$ abgeleitet. — Eng schließen sich mehrere umfangreiche Arbeiten von ALEXEJEWSKY, BARNES, BENDERSKY, MELLIN heran, welche die Verallgemeinerung von $\Gamma(z)$ zum Gegenstand haben; sie greifen teilweise tief in die Funktionentheorie hinein (MELLIN), die zum Grund gelegten Definitionen erscheinen aber nicht immer so einfach bzw. natürlich, daß man bedeutsame Anwendungen hoffen könnte.[7]

Das Ziel der vorliegenden Abhandlung ist nun meine gewissen unlängst publizierten[8] und seitdem wesentlich erweiterten neuen Resultate über die Gammafunktion und verwandte Funktionen in einen möglichst übersichtlichen, konzisen Durchschnitt der ganzen Theorie hineinzubauen. Es sollen zum Hintergrund folgende Bemerkungen dienen: I. Sucht man eine einfache Verallgemeinerung $\Gamma(z)$ von $n!$ ins Komplexe mit $\Gamma(z+1) = z\Gamma(z)$ (*), so ist es naheliegend, das Produkt $z(z+1)\cdots(z+n-1)$ bzw. seine (etwas handlichere) logarithmische Ableitung $\beta_{-1}(z, n) = \frac{1}{z} + \frac{1}{z+1} + \cdots + \frac{1}{z+n-1}$ mit $\beta_{-1}(z+1, n) - \beta_{-1}(z, n) = \frac{1}{z+n} - \frac{1}{z}$ (**) zu betrachten; um daher eine Lösung von (*) zu erhalten, bedürfen wir offenbar des Grenzübergangs $n \to \infty$; da aber für feste z $\beta_{-1}(z, n) \sim \log n$ besteht, so gelangt man natürlicherweise zum

[5] Vgl. z. B. [17]—[25], ferner [33]. Wichtige Angaben über mehrere schöne Resultate von LERCH findet man in [10]. — Professor BORŮVKA (Brno) war so freundlich, meine Aufmerksamkeit auf diese Abhandlungen zu lenken.

[6] [2].

[7] Vgl. [1], [3], [4], [5], [27]. — Einige Erweiterungen sind übrigens nur scheinbar, indem man sie durch passende Transformationen auf $\Gamma(z)$ zurückführen kann; ähnlich ist der Sachverhalt mit den früher untersuchten sog. „analytischen Fakultäten". (WEIERSTRASS.)

[8] [29].

Ausdruck $\Psi(z) = \lim_{n\to\infty}[\log n - \beta_{-1}(z,n)]$ mit $\Psi(z+1) - \Psi(z) = \frac{1}{z}$ und durch $\Gamma(1) = 1$ zum Ansatz

$$\Gamma(z) = \exp\int_1^z \Psi(t)\,dt = \lim_{n\to\infty}\left[n^{z-1}e^{-\int_1^z \beta_{-1}(t,n)\,dt}\right] = \lim_{n\to\infty}\frac{n!\,n^z}{z(z+1)\cdots(z+n)}$$
$$(z \neq 0, -1, \ldots),$$

d. h. zur *Gaußschen Erklärung* der Gammafunktion. — II. Setzt man allgemein $\beta_p(z,n) = z^p + (z+1)^p + \cdots + (z+n-1)^p$ (p ganz), so lassen sich die Ableitungen von $\Psi(z)$ folgendermaßen ausdrücken: $\Psi^{(r-1)}(z) = (-1)^r(r-1)!\cdot \lim_{n\to\infty}\beta_{-r}(z,n) = (-1)^r(r-1)!\sum_{\lambda=0}^{\infty}\frac{1}{(z+\lambda)^r}$ ($r = 2, 3, \ldots$). — III. Die (für jedes ganze $k \geqq 0$ bestehenden) Identitäten: $\frac{1}{z+k} = \int_0^1 u^{z+k-1}\,du$ mit $\Re(z) > 0$ und $\frac{1}{z+k} = \frac{(-1)^k \pi}{\sin \pi z}\int_0^1 \cos(z+k)\pi u\,du$ mit $z \neq 0, \pm 1, \pm 2, \ldots$ ermöglichen durch Summierung offenbar zweierlei Integraldarstellungen für $\Psi(z)$ und verwandte Funktionen, nämlich solche, die in der rechten Halbebene $\Re(z) > 0$ bzw. solche, die für jedes nicht-ganze z gelten.[9]

Hiermit ist der folgende Aufbau motiviert: wir untersuchen im § 2 die *Potenzsummen* (rationalen Funktionen) $\beta_p(z,n) = \sum_{k=0}^{n-1}(z+k)^p$ mit beliebigen ganzen Exponenten p, im § 3 die Funktion $\Psi(z)$ und ihre Ableitungen, welche für $n \to \infty$, endlich im § 4 $\Gamma(z)$ bzw. $\log\Gamma(z)$, welche durch Integration entspringen; jeder Abschnitt beginnt mit den entsprechenden Differenzengleichungen und den daher fließenden anderen Funktionalgleichungen, auf welche die verschiedenen Reihenentwicklungen, Integraldarstellungen, die einfachsten Charakterisierungen, gleichwie asymptotische Entwicklungen folgen. Die sehr leicht sich ergebende Erzeugungsfunktion

$$e^{zw}\frac{e^{nw}-1}{e^w-1} = \sum_{p=0}^{\infty}\beta_p(z,n)\frac{w^p}{p!} \tag{1.1}$$

legt im § 1 ganz nahe, die Bernoullischen Polynome als Hilfsfunktionen durch

$$\frac{we^{zw}}{e^w-1} = B_0(z) + B_1(z)w + \cdots + B_\lambda(z)w^\lambda + \cdots \qquad (|w| < 2\pi) \tag{1.2}$$

[9] Alle klassischen Integraldarstellungen der Theorie, bis auf ein, das sog. Hankelsche Integral, sind bekanntlich nur in einer Halbebene gültig. — Wir bezeichnen mit u stets eine *reelle* Variable.

einzuführen; dann ist $\beta_p(z, n) = p![B_{p+1}(z+n) - B_{p+1}(z)]$ $(p = 0, 1, 2, \ldots)$ und man erhält nebenbei bequem alle wesentlichen Beziehungen über $B_\lambda(z)$ und die Bernoullischen Zahlen $B_\lambda = \lambda! B_\lambda(0)$, ferner die Euler—Maclaurinsche Summenformel, welche in den folgenden (insbesondere in asymptotischen Formeln und Integraldarstellungen) oft benutzt werden.

Betreffs der Vorteile der skizzierten, auf Summendefinitionen und zwei Identitäten gegründeten Behandlungsmethode seien die folgenden erwähnt: 1. dieser Weg erlaubt die auftretenden Reihen- und Produktentwicklungen mit *Restgliedern* zu versehen;[10] die erhaltenen verschärften Beziehungen enthalten speziell — im Fall der Teilbruchzerlegungen — auch die klassischen Integralformeln $(n = 0)$, so daß es sich zugleich um die *Verallgemeinerung* derselben handelt; 2. die asymptotischen Entwicklungen (insbesondere die Stirlingsche Formel) entnimmt man durch geringe Rechnung den entsprechenden für $\beta_p(z, n)$, welche aber nach Euler—Maclaurin sofort aufgeschrieben werden können; 3. man gewinnt auch solche eigentlichen, auf (0, 1) erstreckten *neuen Integraldarstellungen*, deren Integrand trigonometrische Funktionen enthält und welche — gegen die bisher bekannten — in der ganzen z-Ebene gelten; 4. die letzten gehen durch einfache Transformationen in komplexe *Kreis-* bzw. *Schleifenintegrale* über, wobei sich die Verwendung einer bekannten Riemannschen Idee[11] als entbehrlich erweist; 5. es leuchtet durchwegs unschwer ein, welcher Wert des (unendlichvieldeutigen) komplexen Logarithmus, bzw. welcher Zweig von $\operatorname{Log}\Gamma(z)$ zu verstehen ist;[12] 6. als *Nebenprodukte* ergeben sich leicht die (mit Restgliedern versehene) Teilbruch- bzw. Produktzerlegung des Kotangens und Sinus, die von EULER bzw. SONIN herrührenden geschlossenen Ausdrücke für $\zeta(\nu) = \sum_{\lambda=1}^{\infty} \frac{1}{\lambda^\nu}$ $(\nu = 2, 3, \ldots)$, in welchen Bernoullische Zahlen und Polynome vorkommen, eine neuartige Integraldarstellung für die Eulersche Konstante γ, eine neue, sehr einfache Herleitung der sog. Lerchschen trigonometrischen Entwicklung und somit der Kummerschen Reihe usw.

[10] Dies scheint in der Literatur überhaupt und im Fall eines komplexen Veränderlichen noch nicht geschehen zu sein.

[11] Man denkt des Ansatzes: ist der untere oder obere Grenze eines (uneigentlichen) Integrals eine singuläre Stelle des Integranden, so mag es nützlich sein, einen die singulären Punkte umgehenden neuen Integrationsweg zu betrachten. — Die von RIEMANN in 1859 für $\zeta(s)$ benutzte Methode diente wohl HANKEL als Muster in seiner Habilitationsschrift über $\Gamma(z)$ (1863). Vgl. z. B. [7], S. 159.

[12] Mehrere einschlägige Arbeiten fixieren die Bedeutung des log-Zeichens nicht; andere betrachten überall den Hauptwert, was aber, wie z. B. auf Grund der Ginzelschen Untersuchungen [9] mühelos nachweisen kann, nicht stichhaltig ist. — Wir weisen gleich auf einen interessanten Aufsatz von O BORŮVKA ([6]) hin, welcher auf einen diesbezüglichen Fehler im Lehrbuch: SERRET—SCHEFFERS, *Diff. u. Integralrechnung*, 6–7. Auflage (1921), Bd. II, S. 220, aufmerksam macht.

Man findet z. B. die Beziehungen für $\Psi(z)=\frac{d\operatorname{Log}\Gamma(z)}{dz}=\frac{\Gamma'(z)}{\Gamma(z)}$ (vgl. § 3, VII ff.):

$$\Psi(z)+\gamma=\sum_{\lambda=0}^{n-1}\left(\frac{1}{\lambda+1}-\frac{1}{z+\lambda}\right)+Q_n(z) \qquad (n=0,1,2,\ldots)^{13} \tag{1.3}$$

mit

$$Q_n(z)=\begin{cases}\displaystyle\int_0^1 u^n\frac{1-u^{z-1}}{1-u}\,du \text{ für } \Re(z)>0,\\ \displaystyle(-1)^n\frac{\pi}{\sin\pi z}\int_0^1\sec\frac{\pi u}{2}\left[(\cos\pi z u-\cos\pi z)\cos(2n+1)\frac{\pi u}{2}+\right.\\ \left.+(u\sin\pi z-\sin\pi z u)\sin(2n+1)\frac{\pi u}{2}\right]du-\frac{1}{z+n}\end{cases} \tag{1.4}$$

für jedes nicht-ganze z.

Daher entspringen folgende Ausdrücke für die Hurwitzsche Zetafunktion $\zeta(s,z)=\sum_{\lambda=0}^{\infty}\frac{1}{(z+\lambda)^s}$ mit $s=\nu+1$ (ν positiv-ganz):

$$\zeta(\nu+1,z+1)=\begin{cases}\displaystyle\frac{1}{\nu!}\int_0^1 u^z\left(\log\frac{1}{u}\right)^\nu\frac{1}{1-u}\,du \qquad (\Re(z)>-1),\\ \displaystyle\frac{(-1)^{\nu+1}}{\nu!}\left[\frac{d^\nu}{dz^\nu}\left(\frac{1}{2z}-\frac{\pi}{2}\operatorname{ctg}\pi z\right)-\right.\\ \left.\displaystyle-\frac{\pi}{2}\int_0^1\operatorname{tg}\frac{\pi u}{2}\frac{\partial^{\nu-1}}{\partial z^{\nu-1}}\left(\frac{\sin\pi z u}{\sin\pi z}\right)du\right] \quad (z\neq 0,\pm1,\ldots).\end{cases} \tag{1.5}$$

(1.3) liefert mit $n=0$ die Legendresche Formel:

$$\Psi(z)+\gamma=\int_0^1\frac{1-u^{z-1}}{1-u}\,du \qquad (\Re(z)>0) \tag{1.6}$$

bzw. die neuartige Darstellung:

$$\Psi(z)+(\gamma+2\log 2)+\frac{\pi}{2}\operatorname{ctg}\pi z=\frac{\pi}{2}\int_0^1\sec\frac{\pi u}{2}\left[1-\frac{\cos\pi u\left(z-\frac{1}{2}\right)}{\sin\pi z}\right]du \tag{1.7}$$

$$(z\neq 0,\pm1,\ldots),$$

für $n\to\infty$ aber die Teilbruchzerlegung von $\Psi(z)$:

$$\Psi(z)+\gamma=\sum_{\lambda=0}^{\infty}\left(\frac{1}{\lambda+1}-\frac{1}{z+\lambda}\right) \quad (z\neq 0,-1,-2,\ldots). \tag{1.8}$$

[13] Die rechtsstehende Summe ist für $n=0$ mit Null zu ersetzen.

Mittels Integration bekommt man aus (1.6) und (1.7)

$$\gamma = \begin{cases} \int_0^1 \left(\frac{1}{1-u} + \frac{1}{\log u} \right) du, \\ \pi \int_0^1 \operatorname{tg} \frac{\pi u}{2} \left(\frac{1}{2} - \int_0^{\frac{1}{2}} \frac{\sin \pi x u}{\sin \pi x} dx \right) du, \end{cases} \tag{1.9}$$

während aus (1.5) insbesondere für $z \to 0$ die Ausdrücke

$$\zeta(r+1) = \frac{1}{r!} \int_0^1 \left(\log \frac{1}{u} \right)^r \frac{1}{1-u} du^{14} \tag{1.10}$$

bzw.

$$\zeta(r+1) = \begin{cases} (-1)^{\mu-1} \frac{B_{2\mu}}{2(2\mu)!} (2\pi)^{2\mu} & (r = 2\mu+1), \\ (-1)^{\mu+1} (2\pi)^{2\mu+1} \int_0^{\frac{1}{2}} B_{2\mu+1}(u) \operatorname{ctg} \pi u \, du & (r = 2\mu) \end{cases} \tag{1.11}$$

folgen.

Die rechte Seite von (1.7) läßt sich leicht in ein komplexes Kurvenintegral umformen; man erhält

$$\Psi(z) + (\gamma + 2\log 2) + \frac{\pi}{2} \operatorname{ctg} \pi z = \frac{1}{2i} \oint \frac{1}{1+t} \left(t^{-\frac{1}{2}} - \frac{t^{z-1}}{\sin \pi z} \right) dt \tag{1.12}$$

$$(z \neq 0, \pm 1, \ldots),$$

wo $t^\alpha = e^{\alpha \log \alpha}$ (mit dem Hauptwert des Logarithmus) und der in positivem Sinn durchlaufene Einheitskreis $t = e^{i\pi u}$ $(-1 \leqq u \leqq 1)$ als Integrationsweg zu wählen ist.

Ein anderes Beispiel bietet die in § 4, IX ff. gewonnene Darstellung:

$$\Gamma(z) = g(z) + \sum_{\lambda=0}^{n} \frac{(-1)^\lambda}{\lambda!} \frac{1}{z+\lambda} + q_n(z) \qquad (n = 0, 1, 2, \ldots), \tag{1.13}$$

wobei $g(z)$ eine ganze Funktion, nämlich $\int_1^\infty u^{z-1} e^{-u} du$ bedeutet, während

$$q_n(z) = \begin{cases} \frac{(-1)^{n+1}}{n!} \int_0^1 \int_0^1 (1-x)^n u^{z+n} e^{-ux} dx \, du & (\Re(z) > 0), \\ \frac{\pi}{n! \sin \pi z} \int_0^1 \int_0^1 (1-x)^n e^{x \cos \pi u} \cos(\pi u(z+n+1) + \\ \qquad + x \sin \pi u) \, dx \, du & (z \neq 0, \pm 1, \pm 2, \ldots). \end{cases} \tag{1.14}$$

[14] Durch die Substitution $t = e^{-v}$ erhält man daher die Riemannsche Integraldarstellung von $\zeta(r)$.

Setzt man hierin $n=0$, so ergibt sich

$$(1.15)\qquad \Gamma(z)=\int_1^\infty u^{z-1}e^{-u}\,du+\int_0^1 u^{z-1}e^{-u}\,du=\int_0^\infty u^{z-1}e^{-u}\,du \qquad (\Re(z)>0),$$

d. h. das *Eulersche Integral* zweiter Gattung für die Gammafunktion bzw. die neue Integralformel:

$$(1.16)\qquad \Gamma(z)=\int_1^\infty u^{z-1}e^{-u}\,du+\frac{\pi}{\sin \pi z}\int_0^1 e^{\cos \pi u}\cos(\pi u z+\sin \pi u)\,du \qquad (z\neq 0,\pm 1,\pm 2,\ldots);$$

läßt man dagegen $n\to\infty$ streben, so entsteht

$$(1.17)\qquad \Gamma(z)=g(z)+\sum_{\lambda=0}^{\infty}\frac{(-1)^\lambda}{\lambda!}\frac{1}{z+\lambda} \qquad (z\neq 0,-1,-2,\ldots),$$

die sog. Prymsche Teilbruchzerlegung von $\Gamma(z)$,[15] welche die *Residuen* an den einfachen Polen $z=0,-1,-2,\ldots$ $\left(\text{nämlich } \operatorname*{res}_{z=-\lambda}\Gamma(z)=\frac{(-1)^\lambda}{\lambda!}\right)$ in Evidenz setzt.

Aus (1.16) gelangt man durch die Transformation $t=e^{i\pi u}$ gleich zu

$$(1.18)\qquad \Gamma(z)=\frac{1}{2i\sin \pi z}\left[\int_{-\infty}^{-1}(t^{z-1})_-e^t\,dt+\oint t^{z-1}e^t\,dt+\int_{-1}^{-\infty}(t^{z-1})_+e^t\,dt\right],$$

wo die Integration im mittleren Glied auf den (von -1 bis -1 in positivem Sinn durchlaufenen) Einheitskreis, im ersten und dritten aber auf die negativ-reelle Achse zu erstrecken, ferner $t^{z-1}=e^{(z-1)\log t}$, $(t^{z-1})_\pm=e^{(z-1)(\log|t|\pm i\pi)}$ (mit den Hauptlogarithmen) zu verstehen ist; die geradlinigen Integrale beziehen sich also auf das obere bzw. untere Ufer des Schnittes längs $t\leqq 0$ auf der Riemannschen log-Fläche.

Die Anwendung des Cauchyschen Hauptsatzes führt nun mühelos zur allgemeineren Darstellung:

$$(1.19)\qquad \Gamma(z)=\frac{1}{2i\sin \pi z}\left[\int_{\mathfrak{h}_{-\delta}}t^{z-1}e^t\,dt+\int_{\mathfrak{k}_R}t^{z-1}e^t\,dt+\int_{\mathfrak{h}_\delta}t^{z-1}e^t\,dt\right]$$
$$(z\neq 0,\pm 1,\ldots),$$

wo $\mathfrak{h}_{\pm\delta}$ die nach wachsenden bzw. abnehmenden u zu durchlaufenden Halbgeraden $t=ue^{\pm i\delta\pi}$ $(R\leqq u\leqq\infty)$, $\mathfrak{k}_R$ aber den Kreisbogen $t=Re^{i\pi\vartheta}$ $(-\delta\leqq\vartheta\leqq\delta)$ bedeutet; ja sogar, man gewinnt

$$(1.20)\qquad \Gamma(z)=\frac{1}{2i\sin \pi z}\int_{\mathfrak{C}}t^{z-1}e^t\,dt \qquad (z\neq 0,\pm 1,\pm 2,\ldots)$$

[15] Sie entspricht, wie man sieht, genau dem wohlbekannten Mittag-Lefflerschen Teilbruchsatz über meromorphe Funktionen.

mit einem beliebigen Integrationsweg $\mathfrak{C}$, welcher von ∞ ausgehend bis zu einem Punkt $\omega > 0$ in der unteren, dann nach ∞ zurück in der oberen Halbebene verläuft und zwar so, daß er gänzlich in einem Bereich $|t| \leqq \varrho$, $|\arg t| \geqq \delta\pi > \frac{\pi}{2}$ (ϱ, δ fest) liegt. — (1.20) ist als eine *Erweiterung des Hankelschen Schleifenintegrals* für $\Gamma(z)$ (wobei eine wesentlich speziellere „Schleife" in Betracht kommt) anzusehen.

Von den übrigen neuen Darstellungen sei noch die für $0 < z < 1$ gültige

$$(1.21)\quad \Gamma(z) = 2^{-z}\left(\frac{\pi}{\sin \pi z}\right)^{\frac{1}{2}} \exp\left[\pi z \int_0^1 \int_0^{\frac{1}{2}} \operatorname{tg}\frac{\pi u}{2}\left(\frac{\sin \pi x u}{\sin \pi x} - \frac{\sin 2\pi x z u}{\sin 2\pi x z}\right) dx\, du\right]$$

hervorgehoben. (Vgl. § 4, VII.)

Über die naturgemäß sich aufwerfende Frage, unter welchen Zusatzbedingungen die (bei der Definition benutzten) Differenzengleichungen[16] von $\Gamma(z), \Psi(z), \Psi'(z), \ldots$ diese Funktionen *allein* zur Lösung haben, ergeben sich bekanntlich leicht gewisse Limesrelationen mit $\Re(z) \to \infty$. (Vgl. §§ 3, 4, VI. 1.) Wir zeigen nun im Fall von $\Psi(z)$ und ihrer Ableitungen auf Grund des Liouvilleschen und Mittag-Lefflerschen Satzes, daß jene Bedingungen durch Regularitäts- und Beschränktheitsforderungen inbezug auf den Streifen $1 \leqq \Re(z) < 2$ ersetzt werden können (vgl. § 3, VI. 2); $\Gamma(z)$ läßt sich laut eines kaum bekannten Scheeffer—Wielandtschen Satzes (vgl. § 4, VI. 2) ähnlicherweise *charakterisieren*. Die letzten Sätze sind für die Identifizierung irgendeiner vermuteten Integraldarstellung mit der Summen- bzw. Produktdefinition einer betrachteten Funktion offenbar handlicher, als die vorher erwähnten.[17]

Endlich sei bemerkt, daß ich — in Anbetracht des Charakters dieser Abhandlung — an allen Stellen, wo es mir möglich ist, Autoren und Literaturangaben zitieren werde, obgleich es sich teilweise um klassisch gewordene Resultate handelt.

Ich hoffe in der zweiten Mitteilung auf eine verallgemeinerte Theorie der Riemannschen und Hurwitzschen Zetafunktion zurückzukommen.

[16] Es leuchtet ein, daß die Lösung einer Differenzengleichung der Form $F(z+1) - F(z) = \varphi(z)$ nur bis auf eine additive, mit 1 periodische Funktion, die von $F(z+1) = \varphi(z)F(z)$ aber bis auf einen (beliebigen) mit 1 periodischen Faktor bestimmt ist.

[17] Vgl. KNOPP [16], S. 47. — Auf solche Art gelangt man am kürzesten zum Integralausdruck von $g(z)$ in (1.17) (§ 4, X).

§ 2. Eigenschaften der Funktionen $\beta_p(z, n)$

Als eine einfache Verallgemeinerung der Potenzsummen sukzessiver natürlicher Zahlen, betrachten wir die *summatorische Funktion*

$$\beta_p(z, n) = z^p + (z+1)^p + \cdots + (z+n-1)^p, \tag{2. 1}$$

wobei z eine *komplexe* Variable, während der Parameter p eine *ganze* Zahl bezeichnet.

Speziell ist also $\beta_0(z, n) \equiv n \quad (z \neq 0)$; $\beta_p(z, n)$ ist eine *ganze rationale* Funktion von z für jedes fixe $p > 0$, dagegen erhält man bei $p = -\nu$ $(\nu = 1, 2, \ldots)$

$$\beta_{-\nu}(z, n) = \frac{1}{z^\nu} + \frac{1}{(z+1)^\nu} + \cdots + \frac{1}{(z+n-1)^\nu}, \tag{2. 2}$$

d. h. eine *gebrochene rationale* Funktion von z. Wie man sieht, liefert (2. 2) zugleich die Partialbruchzerlegung von $\beta_{-\nu}(z, n)$; sie setzt in Evidenz, daß die Punkte $z = 0, -1, \ldots, -(n-1)$ sämtliche Pole (ν-ter Ordnung) von $\beta_{-\nu}(z, n)$ sind.

Es seien gleich die Werte

$$\beta_p(1, n) = 1^p + 2^p + \cdots + n^p, \tag{2. 3}$$

$$\beta_p\left(\frac{1}{2}, n\right) = \frac{1}{2^p}[1^p + 3^p + \cdots + (2n-1)^p] = \frac{1}{2^p}\beta_p(1, 2n) - \beta_p(1, n), \tag{2. 4}$$

ferner

$$\beta_p(0, n) = 1^p + 2^p + \cdots + (n-1)^p \qquad (p > 0) \tag{2. 5}$$

erwähnt.

Wir gehen zur systematischen Untersuchung der Funktionen $\beta_p(z, n)$ über. — Zunächst setzt man die folgenden, für *jedes* p geltenden Grundeigenschaften von $\beta_p(z, n)$ fest:[18]

I. Differenzengleichungen:

$$\beta_p(z, n+1) - \beta_p(z, n) = (z+n)^p, \tag{2. 6}$$

$$\beta_p(z+1, n) - \beta_p(z, n) = (z+n)^p - z^p. \tag{2. 7}$$

II. Additionstheorem:

$$\beta_p(z, m+n) = \beta_p(z+n, m) + \beta_p(z, n) = \beta_p(z+m, n) + \beta_p(z, m) \quad (m = 1, 2, \ldots). \tag{2. 8}$$

III. Ergänzungssätze:

$$\beta_p(1-z, n) = (-1)^p \beta_p(z-n, n), \tag{2. 9}$$

$$\beta_p(z, n) + (-1)^p \beta_p(1-z, n) = \beta_p(z-n, 2n). \tag{2. 10}$$

[18] Da und in den folgenden Zeilen kommen alle Werte von z in Betracht, für welche die hierauftretenden Funktionen erklärt sind. Die Benennungen entsprechen meistens denjenigen der Differenzenrechnung.

Fast ebenso unschwer erkennt man, daß die Summe der Funktionswerte an gewissen m äquidistanten Stellen der z-Ebene durch *einen* Funktionswert mit m-fachen Argumenten ausdrückbar ist, genauer

IV. Multiplikationstheorem:

$$\sum_{\lambda=0}^{m-1} \beta_p\left(z+\frac{\lambda}{m},\ n\right)=\frac{1}{m^p}\beta_p(mz,\ mn) \qquad (m=1,2,\ldots). \tag{2.11}$$

BEWEIS. Es ist

$$\sum_{\lambda=0}^{m-1} \beta_p\left(z+\frac{\lambda}{m},\ n\right)=\sum_{\lambda=0}^{m-1}\sum_{k=0}^{n-1}\left(z+\frac{\lambda}{m}+k\right)^p=$$

$$=\frac{1}{m^p}\left\{\sum_{k=0}^{n-1}(mz+km)^p+\sum_{k=0}^{n-1}[mz+(km+1)]^p+\cdots+\sum_{k=0}^{n-1}[mz+(km+m-1)]^p\right\}=$$

$$=\frac{1}{m^p}\sum_{l=0}^{mn-1}(mz+l)^p=\frac{1}{m^p}\beta_p(mz,\ mn).$$

Die gliedweise Differentiation bzw. Integration von $\beta_p(z, n)$ (p, n fest) zeigt, daß diese Operationen — abgesehen vom wichtigen Spezialfall $p=-1$ — im wesentlichen nicht dem Kreis unserer Funktionen hinausführen.

V. Ableitungen und Stammfunktionen. *Für jedes p ist*

$$\frac{d^\lambda}{dz^\lambda}\beta_p(z, n)=p(p-1)\cdots(p-\lambda+1)\beta_{p-\lambda}(z, n), \tag{2.12}$$

insbesondere für $p=-1$

$$\frac{d^\lambda}{dz^\lambda}\beta_{-1}(z, n)=(-1)^\lambda\lambda!\,\beta_{-(\lambda+1)}(z, n). \tag{2.13}$$

Es gilt ferner für $p\neq -1$

$$\int\beta_p(z, n)dz=\frac{1}{p+1}\beta_{p+1}(z, n)+const. \tag{2.14}$$

und[19]

$$\int\beta_{-1}(z, n)dz=\log z+\log(z+1)+\cdots+\log(z+n-1)+const. \tag{2.15}$$

in jedem Bereich der z-Ebene, welcher keinen Punkt der negativ-reellen Achse (einschl. 0) enthält.

Auf Grund von (2.14) und (2.15) ist ein beliebiges, längs eines geeigneten Weges genommenes Kurvenintegral von $\beta_p(z, n)$ sofort berechenbar.

[19] Da und durchwegs bezeichnet „log" den *Hauptwert* des komplexen natürlichen Logarithmus und „arg" dasjenige Argument, welches zwischen $-\pi$ (ausschl.) und $+\pi$ (einschl.) gelegen ist.

Besonders einfach sind die folgenden *Integralrelationen:*

$$\int_z^{z+1} \beta_p(t,n)\,dt = \frac{1}{p+1}\left[\beta_{p+1}(z+1,n)-\beta_p(z,n)\right] = \tag{2.16}$$

$$= \frac{1}{p+1}\left[(z+n)^{p+1}-z^{p+1}\right] \qquad (p \neq -1),$$

$$\int_z^{z+1} \beta_{-1}(t,n)\,dt = \log\frac{z+n}{z} = \log(z+n)-\log z, \tag{2.17}$$

wobei $|\arg z| < \pi$ und der Integrationsweg ganz im Inneren der längs $t \leqq 0$ aufgeschlitzten Ebene verlaufen soll. (Wir wollen sie als *Spannenintegrale* bezeichnen.) Man braucht noch[20]

$$\int_{\frac{1}{2}}^{1} \beta_{-1}(u,n)\,du = \log\frac{(2n)!!}{(2n-1)!!} = -\log\left|\binom{-\frac{1}{2}}{n}\right|. \tag{2.18}$$

VI. Generatorfunktionen. *Es gelten die Entwicklungen*

$$e^{zw}\frac{e^{nw}-1}{e^{w}-1} = \sum_{p=0}^{\infty}\beta_p(z,n)\frac{w^p}{p!}, \tag{2.19}$$

$$\beta_{-1}(z-w,n) = \sum_{\nu=1}^{\infty}\beta_{-\nu}(z,n)\,w^{\nu-1}; \tag{2.20}$$

und zwar die erste für jedes z und w,[21] *die zweite für $z \neq 0, -1, \ldots, -n+1$ und $|w| < \min_{k=0,1,\ldots,n-1}|z+k|$. — Somit mögen die linksstehenden Ausdrücke als Generatorfunktionen der Funktionenfolgen $\beta_p(z,n)$ $(p=0,1,2,\ldots)$ bzw. $\beta_{-\nu}(z,n)$ $(\nu=1,2,\ldots)$ angesehen werden.*

Beweis. Man hat einerseits für beliebige x und z

$$\sum_{p=0}^{\infty}\beta_p(z,n)\frac{w^p}{p!} = \sum_{k=0}^{n-1}\sum_{p=0}^{\infty}(z+k)^p\frac{w^p}{p!} = \sum_{k=0}^{n-1}e^{w(z+k)} = e^{zw}\frac{e^{nw}-1}{e^{w}-1}.$$

Wendet man anderseits auf $\beta_{-1}(z-w,n)$ den Cauchy—Taylorschen Entwicklungssatz an (vgl. (2.12)), so ergibt sich gerade (2.20) samt dem genannten Konvergenzradius.

Die Relation (vgl. (2.13))

$$\beta_{-\nu}(z,n) = \frac{(-1)^{\nu-1}}{(\nu-1)!}\frac{d^{\nu-1}\beta_{-1}(z,n)}{dz^{\nu-1}} \qquad (\nu=2,3,\ldots) \tag{2.21}$$

[20] u sei stets zur Bezeichnung einer *reellen* Variablen aufgehalten; x soll vielmehr auch in diesem Sinn zu verstehen, wenn nicht eine andere Festsetzung getroffen wird.

[21] Unter dem zu $w=0$ gehörigen Funktionswert der regulären linken Seite von (2.19) ist natürlich der entsprechende (offenbar existierende) Grenzwert n zu verstehen

und (2. 20) zeigen die ausgezeichnete Rolle, die $\beta_{-1}(z, n)$ unter allen Funktionen $\beta_{-\nu}(z, n)$ spielt: $\beta_{-2}(z, n)$, $\beta_{-3}(z, n)$, ... lassen sich einfach mittels $\beta_{-1}(z, n)$ erzeugen.

Was (2. 19) betrifft, so liegt es nahe, die Potenzreihenentwicklung

$$\frac{we^{zw}}{e^w - 1} = B_0(z) + B_1(z)w + \cdots + B_\lambda(z)w^\lambda + \cdots \tag{2. 22}$$

als Hilfsmittel heranzuziehen. Sie gilt offenbar für beliebige z und $|w| < 2\pi$; schreibt man (2. 22) in der Form:

$$\sum_{p=0}^{\infty} \frac{(zw)^p}{p!} = \left(\sum_{\lambda=0}^{\infty} B_\lambda(z)w^\lambda\right)\left(\sum_{\mu=0}^{\infty} \frac{1}{(\mu+1)!}w^\mu\right) \qquad (|w| < 2\pi),$$

so folgt (nach dem Identitätssatz für Potenzreihen) die Gleichheit der entsprechenden Koeffizienten, d. h. identisch

$$\frac{1}{(p+1)!}B_0(z) + \frac{1}{p!}B_1(z) + \cdots + \frac{1}{1!}B_p(z) = \frac{z^p}{p!} \quad (p = 0, 1, 2, \ldots). \tag{2. 23}$$

Auf Grund dieser Rekursionsformel sind $B_0(z), B_1(z), \ldots$ sukzessiv leicht zu berechnen: man erhält

$$\left\{\begin{array}{l} B_0(z) = 1, \quad B_1(z) = z - \frac{1}{2}, \quad B_2(z) = \frac{1}{2}z(z-1) + \frac{1}{12}, \\ B_3(z) = \frac{1}{6}z(z-1)\left(z - \frac{1}{2}\right), \quad B_4(z) = \frac{1}{24}z^2(z-1)^2 - \frac{1}{720}, \\ B_5(z) = \frac{1}{120}z(z-1)\left(z - \frac{1}{2}\right)\left(z^2 - z - \frac{1}{3}\right) \qquad \text{usw.} \end{array}\right. \tag{2. 24}$$

Man sieht zunächst, daß allgemein $B_p(z)$ sich als ein Polynom (p-ten Grades) mit rationalen Koeffizienten erweist; $B_0(z), B_1(z), \ldots$ heißen bekanntlich *Bernoullische Polynome* und treten sehr oft in der Analysis auf.

Die Werte

$$B_\lambda = \lambda!\, B_\lambda(0) \qquad (\lambda = 0, 1, 2, \ldots) \tag{2. 25}$$

nennt man üblich *Bernoullische Zahlen*; sie sind also rational und genügen der Formel (vgl. (2. 23) für $z = 0$)

$$\binom{p+1}{0}B_0 + \binom{p+1}{1}B_1 + \cdots + \binom{p+1}{p}B_p = 0 \qquad (p = 0, 1, 2, \ldots), \tag{2. 26}$$

woher man ohne Mühe z. B.[22]

[22] Da die linke Seite der Entwicklung [vgl. (2. 22) für $z = 0$]: $\frac{w}{e^w - 1} + \frac{1}{2}w = 1 + \frac{B_2}{2!}w^2 + \frac{B_3}{3!}w^3 + \cdots$ eine gerade Funktion von w ist, so muß die rechte Seite nur gerade Potenzen von w enthalten, d. h. *jedes B_λ mit einem ungeraden $\lambda > 1$ verschwinden.*

$$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_4 = -\frac{1}{30}, \tag{2.27}$$

$$B_6 = \frac{1}{42}, \ldots, B_3 = B_5 = B_7 = \cdots = 0$$

findet.

Die Ähnlichkeit der Erzeugungsfunktionen in (2. 19) und (2. 22) zeigt, daß die Eigenschaften der Bernoullischen Polynome analog denjenigen von $\beta_p(z, n)$ sein sollen. In der Tat, durch die Methode der Koeffizientenvergleichung ergeben sich aus (2. 22) leicht die (für jedes z geltenden) Relationen:

$$B_{p+1}(z+1) - B_{p+1}(z) = \frac{z^p}{p!}, \tag{2.28}$$

$$B_p(1-z) = (-1)^p B_p(z), \tag{2.29}$$

$$\sum_{\lambda=0}^{m-1} B_p\left(z + \frac{\lambda}{m}\right) = \frac{1}{m^{p-1}} B_p(mz), \tag{2.30}$$

$$B'_{p+1}(z) = B_p(z) \tag{2.31}$$

$(p = 0, 1, 2, \ldots).$

Die wiederholte Anwendung von (2. 31) und der Taylorsche Satz liefern $(p = 0, 1, 2, \ldots)$

$$B_p(z+h) = B_p(h) + \frac{1}{1!} B_{p-1}(h) z + \frac{1}{2!} B_{p-2}(h) z^2 + \cdots + \frac{1}{p!} B_0(h) z^p;$$

speziell für $h = 0$ bekommt man (vgl. (2. 25))

$$B_p(z) = \frac{1}{p!}\left[B_p + \binom{p}{1} B_{p-1} \cdot z + \binom{p}{2} B_{p-2} \cdot z^2 + \cdots + \binom{p}{p} B_0 \cdot z^p\right], \tag{2.32}$$

also die explizite Form von $B_p(z)$ mit den Bernoullischen Zahlen in den Koeffizienten.

(2. 31) und (2. 28) ergeben sofort das Spannenintegral:

$$\int_z^{z+1} B_p(t)\,dt = \frac{z^p}{p!} \qquad (p = 1, 2, \ldots), \tag{2.33}$$

wobei der Integrationsweg zwischen z und $z+1$ beliebig gewählt werden kann; man hat den Spezialfall

$$\int_0^1 B_p(t)\,dt = 0 \qquad (p = 1, 2, \ldots) \tag{2.34}$$

für sich zu erwähnen.[23]

[23] Die Bernoullischen Polynome sind offenbar durch $B_0(z) = 1$, (2. 31) und (2. 34) vollständig determiniert, können also durch diese Bedingungen definiert werden. — Es zeigt sich auch leicht, daß $B_{p+1}(z)$, von additiven Konstanten abgesehen, die einzige Polynomlösung von (2. 28) ist.

Es sei noch bemerkt, daß man von der wohlbekannten trigonometrischen Entwicklung

$$B_1(x) = x - \frac{1}{2} = -\sum_{\lambda=1}^{\infty} \frac{\sin 2\lambda\pi x}{\lambda\pi} \qquad (0 < x < 1) \tag{2. 35}$$

ausgehend, durch wiederholte Integrationen und unter Beachtung von (2. 31) allgemein ($\mu = 1, 2, \ldots$)

$$\left.\begin{aligned} B_{2\mu}(x) &= (-1)^{\mu-1} \sum_{\lambda=1}^{\infty} \frac{2\cos 2\lambda\pi x}{(2\lambda\pi)^{2\mu}}, \\ B_{2\mu+1}(x) &= (-1)^{\mu+1} \sum_{\lambda=\nu}^{\infty} \frac{2\sin 2\lambda\pi x}{(2\lambda\pi)^{2\mu+1}} \end{aligned}\right\} \qquad (0 \leqq x \leqq 1) \tag{2. 36}$$

findet.

Bezeichnet man nun die (mit 1 periodischen) rechtsstehenden Reihensummen in (2. 35), (2. 36) mit $\bar{B}_1(x)$ bzw. $\bar{B}_{2\mu}(x)$, $\bar{B}_{2\mu+1}(x)$,[24] so hat man die leicht nachweisbare Identität

$$\sum_{k=0}^{n-1} f(k) = \int_0^n f(x)dx - \frac{1}{2}[f(n) - f(0)] + \int_0^n \bar{B}_1(x) f'(x)dx$$

und die daraus mittels Teilintegration fließende *Euler—Maclaurinsche Summenformel:*

$$\begin{aligned} \sum_{k=0}^{n-1} f(k) = \int_0^n f(x)dx + \sum_{\lambda=1}^{N} \frac{B^{\lambda}}{\lambda!}[f^{(\lambda-1)}(n) - f^{(\lambda-1)}(0)] + \\ + (-1)^{N-1} \int_0^n \bar{B}_N(x) f^{(N)}(x)dx \qquad (N = 1, 2, \ldots); \end{aligned} \tag{2. 37}$$

sie gilt für eine beliebige Funktion $f(x)$, deren hierauftretende Ableitungen in $[0, n]$ vorhanden und stetig sind.[25]

Von den soeben angegebenen (bekannten) Tatsachen bezüglich der Bernoullischen Zahlen und Polynome werden wir in den folgenden oft Gebrauch machen; sie sind uns nützlich vor allem bei der Untersuchung der natürlichen Frage, wie jede summatorische Funktion $\beta_p(z, n)$ sich in einer geschlossenen Form (d. h. als ein Ausdruck mit einer von n unabhängigen Gliederanzahl bzw. als ein Integral) aufschreiben läßt. Was (2. 37) betrifft, so ist dies eine Quelle asymptotischer Beziehungen.

[24] Also $\bar{B}_{p+1}(x) \equiv B_{p+1}(x)$ für jedes p in $0 < x < 1$.

[25] Für Anwendungen und Verallgemeinerungen vgl. z. B. WIRTINGER [41], KNOPP [15], NÖRLUND [31] bzw. MIKOLÁS [28], KOSCHLIAKOFF [13], [14]. — Letztere Arbeiten enthalten auch erweiterte Beziehungen über die Gammafunktion.

VII. Geschlossene Darstellungen. *Es ist*

$$\beta_p(z,n) = p!\,[B_{p+1}(z+n) - B_{p+1}(z)] \qquad (p = 0, 1, 2, \ldots) \tag{2.38}$$

für jedes z und

$$\beta_{-\nu}(z,n) = \frac{1}{(\nu-1)!}\int_0^1 u^{z-1}\left(\log\frac{1}{u}\right)^{\nu-1}\frac{1-u^n}{1-u}\,du \qquad (\nu = 1, 2, \ldots) \tag{2.39}$$

für $\Re(z) > 0$.[26] — *Ferner*

$$\beta_{-\nu}(z,n) = \frac{(-1)^{\nu-1}}{(\nu-1)!}\frac{\pi}{2}\int_0^1 \sec\frac{\pi u}{2}\,\frac{\partial^{\nu-1}}{\partial z^{\nu-1}}H_n(z,u)\,du \quad (\nu = 1, 2, \ldots) \tag{2.40}$$

mit

$$H_n(z,u) = \frac{1}{\sin \pi z}\left[\cos\left(z-\frac{1}{2}\right)\pi u + (-1)^{n+1}\cos\left(z+n-\frac{1}{2}\right)\pi u\right] \tag{2.41}$$

für jedes nicht-ganze z und $(\mu = 1, 2, \ldots)$

$$\beta_{-\nu}(1,n) = \begin{cases} (-1)^{\mu}(2\pi)^{2\mu-1}\displaystyle\int_0^{\frac{1}{2}} B_{2\mu-1}(u)\left[\operatorname{ctg}\pi u - \frac{\cos(2n+1)\pi u}{\sin \pi u}\right]du, \\ \qquad\qquad \textit{falls } \nu = 2\mu-1, \\ (-1)^{\mu-1}(2\pi)^{2\mu}\displaystyle\int_0^{\frac{1}{2}} B_{2\mu}(u)\,\frac{\sin(2n+1)\pi u}{\sin \pi u}\,du, \\ \qquad\qquad \textit{wenn } \nu = 2\mu.^{27} \end{cases} \tag{2.42}$$

BEWEIS. 1° Setzt man $z+k$ statt x in (2.28) und summiert für $k = 0, 1, \ldots, n-1$, so entsteht (vgl. (2.1)) unmittelbar (2.38).

2° Man betrachte die Identitäten ($k \geqq 0$, ganz)

$$\frac{1}{z+k} = \int_0^1 u^{z+k-1}\,du \qquad (\Re(z) > 0) \tag{2.43}$$

und

$$\frac{1}{z+k} = \frac{(-1)^k \pi}{\sin \pi z}\int_0^1 \cos(z+k)\pi u\,du \qquad (z \neq 0, \pm 1, \ldots). \tag{2.44}$$

[26] Unter u^z $(u > 0)$ ist stets $e^{z\log u}$ mit dem reellen Wert des Logarithmus zu verstehen.

[27] (2.42) mag auch als ein Spezialfall von (2.40) aufgefaßt werden, indem die rechte Seite von (2.40) (wie die Potenzreihenentwicklung von $H_n(z,u)$ mittels (2.22) zeigt) für $z \to 1$ eben in diejenige von (2.42) übergeht.

Da man laut bekannter Sätze der Funktionentheorie[28] die wiederholten Differentiationen nach z in (2. 43) und (2. 44) unter dem Integralzeichen ausführen darf, gelangt man auf Grund von (2. 21) durch Ableitung und Summation gerade zu den Formeln (2. 39) bzw. (2. 40).

3° (2. 42) fließt am bequemsten aus

$$\left.\begin{aligned} \text{(2. 45)} \qquad \frac{1}{\lambda^{2\mu}} &= (-1)^{\mu-1}\, 2\,(2\pi)^{2\mu} \int_0^{\frac{1}{2}} B_{2\mu}(u) \cos 2\lambda\pi u \, du, \\ \text{(2. 46)} \qquad \frac{1}{\lambda^{2\mu-1}} &= (-1)^{\mu}\, 2\,(2\pi)^{2\mu-1} \int_0^{\frac{1}{2}} B_{2\mu-1}(u) \sin 2\lambda\pi \, du \end{aligned}\right\} \quad (\lambda = 1, 2, \ldots)$$

(vgl. (2. 35), (2. 36)) unter Benutzung üblicher trigonometrischer Summationsformeln.

Es sei gleich hervorgehoben, daß man aus (2. 40) mit $\nu = 1$ leicht ein bemerkenswertes, *über den Einheitskreis erstrecktes Kurvenintegral für* $\beta_{-1}(z, n)$ gewinnen kann.

Wegen der Eulerschen Beziehungen findet man nämlich

$$\frac{\cos\left(z-\frac{1}{2}\right)\pi u + (-1)^{n+1}\cos\left(z+n-\frac{1}{2}\right)\pi u}{\cos\frac{\pi u}{2}} =$$

$$= e^{i\pi z u}\,\frac{1+(-1)^{n+1}e^{in\pi u}}{1+e^{i\pi u}} + e^{-i\pi z u}\,\frac{1+(-1)^{n+1}e^{-in\pi u}}{1+e^{-i\pi u}},$$

also durch die Substitution $t = e^{i\pi u}$ die Darstellung

$$\text{(2. 47)} \qquad \begin{aligned} \beta_{-1}(z, n) &= \frac{\pi}{2\sin\pi z}\int_{-1}^{1} e^{i\pi z u}\,\frac{1+(-1)^{n+1}e^{in\pi u}}{1+e^{i\pi u}}\,du = \\ &= \frac{1}{2i\sin\pi z}\oint t^{z-1}\,\frac{1+(-1)^{n+1}t^{n}}{1+t}\,dt \qquad (z = 0, \pm 1, \pm 2, \ldots); \end{aligned}$$

[28] Im Fall von (2. 44) ist ein Korollar des Moreraschen Satzes anwendbar (vgl. z. B. Osgood [32], S. 307), während man bei (2. 43) sich auf die folgende (leicht direkt verifizierbare) Bemerkung stützen kann:

Es sei ein für $\Re(z) > 0$ *existierendes Integral der Form* $J(z) = \int_0^1 [g(u)u^{z-1} + h(u)]\,du$ *betrachtet. Gilt hierbei* $\lim_{u\to+0} g(u)u^{\vartheta} = 0$ *für jedes* $\vartheta > 0$, *so ist* $J(z)$ *in der Halbebene* $\Re(z) > 0$ *analytisch und können ihre alle Ableitungen durch formales Differenzieren unter dem Integralzeichen gebildet werden.*

Wir werden uns weiterhin auf beide Sätze mehrmals berufen.

da und durchwegs ist $t^z = e^{z \log t}$ (mit dem Hauptwert des Logarithmus) zu verstehen, während das Integralzeichen mit dem Reifen auf den in positivem Sinn zu durchlaufenden Einheitskreis als Integrationsweg verweist.

VIII. Asymptotische Formeln. *Es gilt für jedes positiv-ganze N nebst arg* $z \neq \pi$

$$\beta_{-1}(z,n) = \log(z+n) - \log z + \sum_{\lambda=1}^{N} (-1)^{\lambda-1} \frac{B_\lambda}{\lambda} \left[\frac{1}{(z+n)^\lambda} - \frac{1}{z^\lambda} \right] - N! \int_0^n \frac{\bar{B}_N(x)}{(z+x)^{N+1}}\, dx, \tag{2.48}$$

ferner für $\nu \geq 2$

$$\beta_{-\nu}(z,n) = \frac{1}{1-\nu} \sum_{\lambda=0}^{N} (-1)^\lambda B_\lambda \cdot \binom{\lambda+\nu-2}{\lambda} \left[\frac{1}{(z+n)^{\lambda+\nu-1}} - \frac{1}{z^{\lambda+\nu-1}} \right] - \tag{2.49}$$
$$- N! \binom{N+\nu-1}{N} \int_0^n \frac{\bar{B}_N(x)}{(z+x)^{N+\nu}}\, dx.$$

BEWEIS. Setzt man in (2.37) $f(x) = (z+x)^p$, so ergibt sich (für geeignete z)

$$\beta_p(z,n) = \int_0^n (z+x)^p\, dx + \sum_{\lambda=1}^{N} \frac{B_\lambda}{\lambda} \binom{p}{\lambda-1} \left[(z+n)^{p-\lambda+1} - z^{p-\lambda+1} \right] +$$
$$+ (-1)^{N-1} N! \binom{p}{N} \int_0^n \bar{B}_N(x) (z+x)^{p-N}\, dx,$$

was aber für $p = 0, 1, 2, \ldots$ und $N = p+1$ in (2.38) (vgl. (2.32)), für negative $p = -\nu$ in (2.48) bzw. (2.49) übergeht.

§ 3. Der Grenzübergang $n \to \infty$. Die Funktion $\Psi(z)$ und ihre Ableitungen

Die erhaltenen Beziehungen für $\beta_{-\nu}(z,n)$ legen ganz nahe, den Grenzübergang $n \to \infty$ zu versuchen, wobei mehrere unsere Ergebnisse sich offenbar einfacher gestalten werden. (2.21) zeigt, daß man vor allem auf den Fall $\nu = 1$ Rücksicht nehmen muß; da aber $\beta_{-1}(z,n)$ bei festem z divergiert (und zwar ersichtlich so, wie eine Teilsumme der harmonischen Reihe), setzen wir

$$\Psi(z) = \lim_{n \to \infty} [\log n - \beta_{-1}(z,n)]^{29} \tag{3.1}$$

[29] Wie man weiß (vgl. den nächsten Abschnitt), ist $\Psi(z)$ die logarithmische Ableitung der Gammafunktion. Sie heißt auch „Digammafunktion", ihre Derivierte „Trigammafunktion" usw.

und beweisen, daß dieser Limes in jedem abgeschlossenen, beschränkten Gebiet $\mathfrak{G}$ der z-Ebene, welches keinen der Punkte $0, -1, -2, \ldots$ enthält, gleichmäßig existiert.

Man hat nämlich

$$\log n - \beta_{-1}(z, n) = \sum_{\lambda=1}^{n-1}\left[\log\left(1+\frac{1}{\lambda}\right)-\frac{1}{z+\lambda}\right]-\frac{1}{z}$$

und die Abschätzung

$$(3.2)\qquad \left|\log\left(1+\frac{1}{\lambda}\right)-\frac{1}{z+\lambda}\right| \leqq \frac{1}{2\lambda^2}+\left|\frac{1}{\lambda}-\frac{1}{z+\lambda}\right| < \frac{1}{\lambda^2}\left(\frac{1}{2}+2\max_{(\mathfrak{G})}|z|\right),$$

falls λ genug groß genommen wird.

Das Kriterium von WEIERSTRASS ergibt zugleich, daß die Reihenentwicklungen $(z \neq 0, -1, -2, \ldots)$

$$(3.3)\qquad \Psi(z) = -\frac{1}{z}+\sum_{\lambda=1}^{\infty}\left[\log\left(1+\frac{1}{\lambda}\right)-\frac{1}{z+\lambda}\right]$$

oder

$$(3.4)\qquad \Psi(z) = -\gamma+\sum_{\lambda=0}^{\infty}\left(\frac{1}{\lambda+1}-\frac{1}{z+\lambda}\right)$$

mit

$$(3.5)\qquad \gamma = \lim_{n\to\infty}\left[\beta_{-1}(1, n)-\log n\right] = \sum_{\lambda=1}^{\infty}\left[\frac{1}{\lambda}-\log\left(1+\frac{1}{\lambda}\right)\right]$$

gleichmäßig in jedem besprochenen Bereich gelten; γ heißt bekanntlich *Euler—Mascheronische Konstante*. Es ist offenbar

$$(3.6)\qquad \Psi(1) = -\gamma,$$

während auf Grund von (2.4), (3.1), (3.5)

$$(3.7)\qquad \Psi\left(\frac{1}{2}\right) = -(\gamma+2\log 2).$$

(3.3) und (3.4) liefern die *Partialbruchzerlegung von* $\Psi(z)$; daraus erhellt, daß $\Psi(z)$ eine *meromorphe* Funktion ist, deren sämtliche Pole (erster Ordnung) die Stellen $0, -1, -2, \ldots$ mit dem gemeinsamen *Residuum* -1 sind.

Die folgenden Eigenschaften von $\Psi(z)$ ergeben sich unmittelbar aus § 1, I bzw. IV für $n \to \infty$ unter Beachtung der Definition (3.1); sie gelten für jedes z, wobei sie einen Sinn haben.

I. Differenzengleichung:

$$(3.8)\qquad \Psi(z+1)-\Psi(z) = \frac{1}{z}.$$ [30]

[30] Aus (3.6) und (3.8) kann man auf $\Psi(\nu) = -\gamma+\beta_{-1}(1, \nu-1)$ $(\nu = 2, 3\ldots)$ schließen.

II. Multiplikationstheorem:

$$\sum_{k=0}^{m-1} \Psi\left(z+\frac{k}{m}\right) = m\,[\Psi(mz) - \log m]. \tag{3.9}$$

Was den Fall von § 1, III betrifft, so braucht man $\lim_{n\to\infty} \beta_{-1}(z-n, 2n)$. Die Existenz folgt für $z \neq 0, 1, 2, \ldots$ aus den vorigen und zwar ist

$$\lim_{n\to\infty} \beta_{-1}(z-n, 2n) = \frac{1}{z} + \sum_{\lambda=1}^{\infty}\left(\frac{1}{z+\lambda} + \frac{1}{z-\lambda}\right), \tag{3.10}$$

gleichmäßig in jedem abgeschlossenen, endlichen Bereich von nicht-ganzen z. Andererseits gewinnt man durch (2.40) bzw. (2.44)

$$\beta_{-1}(z-n, 2n) = \frac{(-1)^n \pi}{\sin \pi z} \int_0^1 \frac{\sin n\pi u}{\cos \frac{\pi u}{2}} \sin \pi u \left(z - \frac{1}{2}\right) du =$$

$$= \frac{(-1)^n \pi}{\sin \pi z} \int_0^1 \cos \pi z u \, \frac{\cos (2n+1)\frac{\pi u}{2}}{\cos \frac{\pi u}{2}} \, du - \frac{1}{z+n} \tag{3.11}$$

und somit wegen des Riemannschen Lemmas[31]

$$\lim_{n\to\infty} \beta_{-1}(z-n, 2n) = \tag{3.12}$$

$$= \lim_{n\to\infty} \left[\frac{(-1)^n \pi}{\sin \pi z} \int_0^1 \left(\cos \pi z u - \cos \pi z \, \frac{\cos (2n+1)\frac{\pi u}{2}}{\cos \frac{\pi u}{2}} \right) du + \pi \operatorname{ctg} \pi z \right] =$$

$$= \pi \operatorname{ctg} \pi z \qquad (z \neq 0, \pm 1, \pm 2, \ldots).$$

Wir registrieren gleich auf Grund von (3.10) und (3.12)

$$\pi \operatorname{ctg} \pi z = \frac{1}{z} + \sum_{\lambda=1}^{\infty}\left(\frac{1}{z+\lambda} + \frac{1}{z-\lambda}\right), \tag{3.13}$$

d. h. die *Partialbruchdarstellung des Kotangens.*

Setzt man nun in (2.10) $p = -1$ und läßt $n \to \infty$ streben, so ergibt sich (vgl. (3.1)) der

[31] Der Real- und Imaginärteil von $\left(\cos \frac{\pi u}{2}\right)^{-1} (\cos \pi z u - \cos \pi z)$ wird offenbar in $0 \leqq u \leqq 1$ stetig, wenn man die Grenzwerte für $u \to 1-0$ als Funktionswerte auffaßt.

Auf Grund von (3.4), (3.5) kann der Verlauf von $\Psi(z)$ (samt ihren Ableitungen) leicht untersucht worden; es läßt sich z. B. zeigen, daß alle Nullstellen von $\Psi(z)$ einfach, reell sind und je eine von ihnen in den Intervallen $(0, 2)$, $(-1, 0)$, $(-2, -1)$, ... liegt.

III. Ergänzungssatz:

$$\Psi(1-z)-\Psi(z)=\pi\operatorname{ctg}\pi z \qquad (z\neq 0,\pm 1,\pm 2,\ldots). \tag{3.14}$$

Aus (3.8), (3.9) und (3.14) durch *Differentiation*, oder aus (2.7), (2.10), (2.11) für $p=-\nu$ und $n\to\infty$ erhält man einfach die entsprechenden Eigenschaften von $(\nu=2,3,\ldots)$[32]

$$\Psi^{(\nu-1)}(z)=(-1)^{\nu}(\nu-1)!\lim_{n\to\infty}\beta_{-\nu}(z,n)=(-1)^{\nu}(\nu-1)!\,\zeta(\nu,z) \tag{3.15}$$
$$(z\neq 0,\pm 1,\pm 2,\ldots);$$

es seien bemerkt

$$\Psi^{(\nu-1)}(1)=(-1)^{\nu}(\nu-1)!\lim_{n\to\infty}\beta_{-\nu}(1,n)=(-1)^{\nu}(\nu-1)!\,\zeta(\nu) \tag{3.16}$$

und (vgl. (2.4))

$$\Psi^{(\nu-1)}\left(\frac{1}{2}\right)=(-1)^{\nu}(\nu-1)!\,(2^{\nu}-1)\,\zeta(\nu). \tag{3.17}$$

Die *Integralrelationen* (2.17), (2.18) ergeben durch Benutzung der gleichmäßigen Gültigkeit von (3.1)

$$\int_{z}^{z+1}\Psi(t)\,dt=\log z \qquad (|\arg z|<\pi) \tag{3.18}$$

mit einem die Halbachse $t\leqq 0$ vermeidenden Integrationsweg, ferner

$$\int_{\frac{1}{2}}^{1}\Psi(u)du=\lim_{n\to\infty}\left[\frac{1}{2}\log n-\log\frac{(2n)!!}{(2n-1)!!}\right]=-\frac{1}{2}\log\pi;^{33} \tag{3.19}$$

aus (3.18) folgt

$$\int_{1}^{\nu}\Psi(u)du=\sum_{k=1}^{\nu-1}\log k=\log(\nu-1)! \qquad (\nu=1,2,\ldots). \tag{3.20}$$

IV. Potenzreihenentwicklung. *Es ist*

$$\Psi(z+h)=\Psi(h)+\zeta(2,h)z-\zeta(3,h)z^2+-\cdots \tag{3.21}$$
$$\cdots+(-1)^{\nu+1}\zeta(\nu+1,h)\,z^{\nu}+R_{\nu+1}(z,h)\,z^{\nu+1}$$

[32] In den folgenden benutzen wir die übliche Bezeichnung der Hurwitzschen Zetafunktion:

$$\zeta(s,z)=\sum_{\lambda=0}^{\infty}{}'\frac{1}{(z+\lambda)^{s}} \qquad (\Re(s)>1;\ z\neq 0,-1,\ldots),$$

welche die Riemansche $\zeta(s)=\zeta(s,1)$ als Spezialfall umfaßt. — (3.3), (3.4) und (3.15) entsprechen dem bekannten Mittag-Lefflerschen Satz über meromorphe Funktionen.

[33] Das Wallissche Produkt kann bekanntlich auch als $\frac{(2n)!!}{(2n-1)!!}\sim\sqrt{\pi n}$ geschrieben werden.

mit

$$R_{\nu+1}(z,h)=\frac{1}{\nu!}\int_0^1 \psi^{(\nu+1)}(zu+h)(1-u)^\nu\,du \qquad (\nu=0,1,2,\ldots),$$

falls die Strecke $zu+h$ $(0\leqq u\leqq 1)$ *keine der Stellen* $0, -1, -2,\ldots$ *enthält; das Restglied strebt bei fixem* z *und* h *mit* $|z|<\min_{\lambda=0,1,2,\ldots}|\lambda+h|$ *für* $\nu\to\infty$ *gegen* 0 *und zwar gleichmäßig in jedem inneren, konzentrischen Kreis. Insbesondere gilt* $(h=1,\ \nu\to\infty)$

$$\psi(z+1)=-\gamma+\zeta(2)z-\zeta(3)z^2+-\cdots+(-1)^{\nu+1}\zeta(\nu+1)z^\nu+-\cdots \qquad (|z|<1). \tag{3.22}$$

Den BEWEIS (vgl. (3.15), (3.16)) liefert ohne weiteres die Taylorsche Formel bzw. der Entwicklungssatz der Potenzreihen.

V. Asymptotische Formeln. *Man hat für jedes positiv-ganze* N

$$\psi(z)=\log z+\sum_{\lambda=1}^{N}(-1)^{\lambda-1}\frac{B_\lambda}{\lambda z^\lambda}+N!\int_0^\infty\frac{\bar B_N(x)}{(z+x)^{N+1}}\,dx \tag{3.23}$$

und $(\nu=2,3,\ldots)$

$$\psi^{(\nu-1)}(z)=(-1)^{\nu-1}(\nu-2)!\sum_{\lambda=0}^{N}(-1)^\lambda B_\lambda\cdot\binom{\lambda+\nu-2}{\lambda}\frac{1}{z^{\lambda+\nu-1}}+$$
$$+(-1)^{\nu-1}(N+\nu-1)!\int_0^\infty\frac{\bar B_N(x)}{(z+x)^{N+\nu}}\,dx \tag{3.24}$$

mit $|\arg z|\leqq\pi-\varepsilon$ $(0<\varepsilon<\pi)$;[34] *die Restglieder sind (bei fixem* N, ν*)* $O\left(\frac{1}{|z|^N}\right)$ *bzw.* $O\left(\frac{1}{|z|^{N+\nu-1}}\right)$. *Es gelten speziell in jedem der genannten Sektoren*

$$\lim_{|z|\to\infty}[\psi(z)-\log z]=0, \tag{3.25}$$

$$\lim_{|z|\to\infty}\psi^{(\nu-1)}(z)=0 \qquad (\nu=2,3,\ldots). \tag{3.26}$$

BEWEIS. Die Behauptungen folgen unmittelbar aus (2.48) und (2.49) für $n\to\infty$, jedoch unter Beachtung der Tatsache, daß $\bar B_N(x)$ $(N\geqq 2)$ nach der Definition (vgl. (2.36)) für jedes x unter einer festen oberen Schranke bleibt.

[34] Die Restintegrale mit $\nu=1$ in (3.23) bzw. (3.24) können leicht auf andere übliche Formen gebracht, ferner in schwach oder stark konvergente Reihen entwickelt worden. (STIRLING, POISSON, BINET, HERMITE, JENSEN usw.) Man hat speziell Darstellungen der Form $\sum_n n!\,c_n/z(z+1)\cdots(z+n)$ (sog. *Fakultätenreihen*) zu erwähnen.

(3.25) und (3.26) implizieren speziell die *Limesrelationen*

$$\lim_{n\to\infty}[\Psi(z+n)-\log n]=0 \qquad (\Re(z)>0) \tag{3.27}$$

bzw. für jedes $\nu\geqq 2$

$$\lim_{n\to\infty}\Psi^{(\nu-1)}(z+n)=0 \qquad (\Re(z)>0). \tag{3.28}$$

Andererseits konstatiert man auf Grund der Definitionen sofort die *Beschränktheit* von $\Psi'(z)$, $\Psi''(z),\ldots$ in einer Halbebene $\Re(z)\geqq\delta$ $(\delta>0)$; insbesondere für jedes $\nu\geqq 2$

$$|\Psi^{(\nu-1)}(z)|\leqq(\nu-1)!\sum_{\lambda=0}^{\infty}\frac{1}{(\lambda+\Re(z))^{\nu}}\leqq(\nu-1)!\zeta(\nu) \qquad (\Re(z)\geqq 1). \tag{3.29}$$

Wir zeigen nun, daß die Eigenschaften (3.27), (3.28) oder (3.29), verknüpft mit der geeigneten Differenzengleichung, $\Psi(z)$ und ihre Ableitungen im wesentlichen schon determinieren. Schärfer und genauer:

VI. Charakteristische Eigenschaften.

1. *$\Psi(z)$ ist die einzige Lösung der Differenzengleichung*

$$F(z+1)-F(z)=\frac{1}{z} \tag{3.30}$$

$(z\neq 0,-1,-2,\ldots$ sonst beliebig), welche der Bedingung

$$\lim_{n\to\infty}[F(z+n)-\log n]=0 \tag{3.31}$$

genügt; das entsprechende gilt über $\dfrac{(-1)^{\nu-1}}{(\nu-1)!}\Psi^{(\nu-1)}(z)=-\sum_{\lambda=0}^{\infty}\dfrac{1}{(z+\lambda)^{\nu}}=$ $=-\zeta(\nu,z)$ *$(\nu\geqq 2$, fest), falls man* (3.30) *durch*

$$F(z+1)-F(z)=\frac{1}{z^{\nu}}, \tag{3.32}$$

während (3.31) *durch*

$$\lim_{n\to\infty}F(z+n)=0 \tag{3.33}$$

ersetzt.

2. *Wenn $F(z)$ für alle $z\neq 0,-1,-2,\ldots$ der Gleichung* (3.30) *genügt, ferner $F'(z)$ im Streifen $1\leqq\Re(z)<2$ vorhanden und beschränkt ist, so hat man $F(z)\equiv\Psi(z)+const.$*

Gleicherweise: eine in $1\leqq\Re(z)<2$ analytische und beschränkte Lösung $F(z)$ von (3.32) *$(z\neq 0,-1,-2,\ldots;\ \nu\geqq 2)$ muß von der Form* $\dfrac{(-1)^{\nu-1}}{(\nu-1)!}\Psi^{(\nu-1)}(z)+$ $+const.$ *sein.*

BEWEIS. 1. 1° Befriedigt $F(z)$ (3.30) und (3.31), so ergibt sich

$$(3.34)\quad F(z+n)-F(z)=\sum_{k=0}^{n-1}[F(z+k-1)-F(z+k)]=\beta_{-1}(z,n)\quad (n=1,2,\ldots)$$

und

$$F(z)=\lim_{n\to\infty}[F(z+n)-\log n]+\lim_{n\to\infty}[\log n-\beta_{-1}(z,n)]=\Psi(z).$$

2° Für eine Lösung von (3.32) mit (3.33) erhält man ähnlich

$$F(z+n)-F(z)=\beta_{-\nu}(z,n)\qquad (n=1,2,\ldots)$$

und

$$F(z)=\lim_{n\to\infty}F(z+n)-\lim_{n\to\infty}\beta_{-\nu}(z,n)=-\sum_{\lambda=0}^{\infty}{}'\frac{1}{(z+\lambda)^{\nu}}.$$

2. Tiefer ist die andere Hälfte des Satzes.

1° (3.30) impliziert (3.34), woraus man alle Werte von $F(z)$ mit $\Re(z)\geqq 2$ mittels derjenigen im Streifen $1\leqq\Re(z)<2$ ausdrücken kann; man erkennt zugleich die Regularität von $F(z)$ in der Halbebene $\Re(z)\geqq 1$.

Schreibt man aber (3.34) in der Form

$$F(z)=F(z+n)-\beta_{-1}(z,n),$$

so determiert sie die Funktionswerte $F(z)$ eines beliebigen Streifens $-n+1\leqq\Re(z)<-n+2$ $(n=1,2,\ldots)$ und zeigt: $F(z)$ ist auch für $\Re(z)<1$ bis auf die Punkte $-n$ $(n=0,1,2,\ldots)$ analytisch und an den letzteren besitzt sie, der Reihe nach, mit den Teilbrüchen $-\frac{1}{z+n}$ $(n=0,1,2,\ldots)$ verknüpfte Pole.

Da die meromorphe Funktion $\Psi(z)$ dieselbe Pole und zugehörigen Hauptteile hat, lehrt der Mittag-Lefflersche Satz, daß $D(z)=F(z)-\Psi(z)$ eine ganze Funktion sein soll; wegen (3.8) und (3.30) ist ferner $D(z+1)=D(z)$ $(z\neq 0,-1,\ldots)$, d. h. $D(z)$ mit 1 periodisch. Man sieht sofort, daß $D'(z)$ gleichfalls *ganz* ist und die *Periode* 1 hat; sie bleibt aber sogar, der Voraussetzung nach, im Streifen $1\leqq\Re(z)<2$ und somit in der ganzen Ebene beschränkt.

Nun wendet man den klassischen Satz von LIOUVILLE über ganze Funktionen an und findet $D'(z)=c_0$, $D(z)=c_0z+c_1$ mit konstanten c_0, c_1. Erinnert man sich auf die Periodizität von $D(z)$, so entspringt $D(z)=c_1$, d. h.

$$F(z)=\Psi(z)+c_1,$$

w. z. b. w.

2° Im Fall von (3.32) schließt man analog, daß $F(z)$ sich in die ganze z-Ebene hinein analytisch fortsetzen läßt und dieselben Singularitäten bzw. Partialbrüche besitzt, wie $\frac{(-1)^{\nu-1}}{(\nu-1)!}\Psi^{(\nu-1)}(z)=-\sum_{\lambda=0}^{\infty}{}'\frac{1}{(z+\lambda)^{\nu}}$ $(\nu\geqq 2)$; daraus erhellt nach MITTAG-LEFFLER, daß die beiden sich nur um eine ganze Funk-

tion unterscheiden können, welche zugleich offenbar mit 1 periodisch ist und somit, gemäß der Hypothese und (3.29), überall beschränkt ist.

Der Liouvillesche Satz ergibt also: $F(z)+\zeta(r,z)$ $(r \geqq 2)$ soll auf eine Konstante degenerieren, w. z. b. w.

VII. Geschlossene Darstellungen für $\Psi(z)$.

1. *Es gelten die Integralformeln:*

$$\Psi(z)+\gamma=\int_0^1 \frac{1-u^{z-1}}{1-u}\,du \qquad (\Re(z)>0),^{35} \tag{3.35}$$

$$\Psi(z)+\gamma+\frac{\pi}{2}\operatorname{ctg}\pi z=\begin{cases}\dfrac{\pi}{2}\displaystyle\int_0^1 \operatorname{tg}\frac{\pi u}{2}\left(u-\frac{\sin\pi z u}{\sin\pi z}\right)du-\frac{1}{2z}, \\ \qquad\qquad (z\neq 0,\pm 1,\pm 2,\ldots),^{36} \\ \dfrac{\pi}{2}\displaystyle\int_0^1 \sec\frac{\pi u}{2}\left[1-\frac{\cos\pi u\left(z-\frac{1}{2}\right)}{\sin\pi z}\right]du-2\log 2\end{cases} \tag{3.36}$$

und allgemeiner, für jedes $n=0,1,2,\ldots$

$$\Psi(z)+\gamma=\sum_{\lambda=0}^{n-1}\left(\frac{1}{\lambda+1}-\frac{1}{z+\lambda}\right)+\int_0^1 u^n\,\frac{1-u^{z-1}}{1-u}\,du \qquad (\Re(z)>0) \tag{3.37}$$

bzw.

$$\Psi(z)+\gamma=\sum_{\lambda=0}^{n-1}\left(\frac{1}{\lambda+1}-\frac{1}{z+\lambda}\right)-\frac{1}{z+n}+(-1)^n\frac{\pi}{2}\left(r_n(z)+\varrho_n(z)\right) \tag{3.38}$$
$$(z\neq 0,\pm 1,\pm 2,\ldots)$$

mit

$$r_n(z)=\int_0^1\left(\frac{\cos\pi z u}{\sin\pi z}-\operatorname{ctg}\pi z\right)\frac{\cos(2u+1)\frac{\pi u}{2}}{\cos\frac{\pi u}{2}}\,du, \tag{3.39}$$

[35] Das ist die sog. *Legendresche Formel,* welche indessen erst durch CAUCHY und LOBATSCHEWSKY streng bewiesen worden ist. — (3.35) oder (3.36) setzt übrigens in Evidenz einen Satz von GAUSS, daß $\Psi(z)$ für *rationale* z mittels Kreisfunktionen und Logarithmen ausdrückbar ist.

[36] Für eine andere Herleitung und Anwendungen sei auf meine Arbeit [29] verwiesen. — Etwa ein Halbjahr nach dem Erscheinen dieser Abhandlung habe ich durch Professor A. RÉNYI eine gefällige Information von Professor O. BORŮVKA (Brno) bekommen; demgemäß benutzt M. LERCH in einer böhmisch publizierten Arbeit ([22]) bei der Auswertung gewisser Fourierkoeffizienten eine Formel, welche dem ersten Teil von (3.36) (für reelle z) entspricht.

$$\varrho_n(z)=\int_0^1\left(u-\frac{\sin \pi z u}{\sin \pi z}\right)\frac{\sin(2u+1)\frac{\pi u}{2}}{\sin\frac{\pi u}{2}}\,du. \tag{3.40}$$

Die Restglieder von (3. 37), (3. 38) *streben für* $n\to\infty$ *in jedem abgeschlossenen, endlichen Teilgebiet des Gültigkeitsbereiches gleichmäßig gegen Null.*[37]

2. *Die Euler—Mascheronische Konstante läßt sich folgendermaßen darstellen:*[38]

$$\gamma=\int_0^1\left(\frac{1}{1-u}+\frac{1}{\log u}\right)du \tag{3.41}$$

oder

$$\gamma=\pi\int_0^1 \operatorname{tg}\frac{\pi u}{2}\left(\frac{1}{2}-\int_0^{\frac{1}{2}}\frac{\sin \pi x u}{\sin \pi x}\,dx\right)du. \tag{3.42}$$

BEWEIS. Es ist wegen (3. 1), (3. 5)

$$\Psi(z)+\gamma=\lim_{n\to\infty}[\beta_{-1}(1,n)-\beta_{-1}(z,n)]\quad(z\neq 0,-1,-2,\ldots) \tag{3.43}$$

und auf Grund von (2. 39) für $n\geqq 1$

$$\beta_{-1}(1,n)-\beta_{-1}(z,n)=\int_0^1\frac{1-u^{z-1}}{1-u}\,du-\int_0^1 u^n\frac{1-u^{z-1}}{1-u}\,du\qquad(\Re(z)>0)^{39} \tag{3.44}$$

mit Rücksicht auf die leicht nachweisbare Tatsache, daß $\lim_{u\to 1-0}\left(u^{n-1}\frac{u-u^z}{1-u}\right)$ existiert. Die Abschätzung

$$\left|\int_0^1 u^n\frac{1-u^{z-1}}{1-u}\,du\right|\leqq\frac{1}{n}\max_{0\leqq u\leqq 1}\left|\frac{u-u^z}{1-u}\right| \tag{3.45}$$

liefert unmittelbar (3. 35).

[37] Die Summe in (3. 37) und (3. 38) soll für $n=0$ Null bedeuten; dann gehen diese Formeln in (3. 35) bzw. (3. 36) über. Da aber für $n\to\infty$ (3. 4) entsteht, so handelt es sich zugleich um *Verschärfungen der Teilbruchdarstellung von* $\Psi(z)$. — Ist z ein reelles Parameter und bildet man die gewöhnliche Fourierentwicklung von $(\sin \pi z)^{-1}\pi\cos zx(-\pi\leqq x\leqq\pi)$, ferner die konjugierte Fourierreihe von $(\sin \pi z)^{-1}(x\sin \pi z-\pi\sin zx)$ $(-\pi\leqq x\leqq\pi)$ an der Stelle $x=\pi$, so sind die Restglieder eben durch $(-1)^{n+1}\pi r_n(z)$ bzw. $(-1)^{n+1}\pi\varrho_n(z)$ angegeben.

[38] (3. 41) rührt von EULER her; für (3. 42) vgl. MIKOLÁS [29].

[39] Das erste Integral rechts ist für $0<\Re(z)<1$ offenbar uneigentlich.

Was (3. 36) betrifft, so wollen wir mit der Formel (vgl. (2. 40))

$$\beta_{-1}(z,n)=\frac{\pi}{2\sin\pi z}\int_0^1 \sec\frac{\pi u}{2}\left[\cos\pi u\left(z-\frac{1}{2}\right)+\right.$$

$$\left.+(-1)^{n+1}\cos\pi u\left(z+n-\frac{1}{2}\right)\right]du=$$

(3. 46)

$$=\frac{1}{2z}-\frac{1}{z+n}+\frac{\pi}{2\sin\pi z}\left\{(-1)^n\int_0^1\cos\pi z u\,\frac{\cos(2n+1)\frac{\pi u}{2}}{\cos\frac{\pi u}{2}}\,du+\right.$$

$$\left.+\int_0^1\sin\pi z u\left[\operatorname{tg}\frac{\pi u}{2}+(-1)^{n+1}\frac{\sin(2n+1)\frac{\pi u}{2}}{\cos\frac{\pi u}{2}}\right]du\right\}.$$

ausgehen.[40] Andererseits ergeben (2. 35), (2. 42) und die Substitution $u=\frac{1}{2}(1-v)$

$$\beta_{-1}(z,n)=2\pi\int_0^{\frac{1}{2}}\left(\frac{1}{2}-u\right)\left[\operatorname{ctg}\pi u-\frac{\cos(2n+1)\pi u}{\sin\pi u}\right]du=$$

(3. 47)

$$=\frac{\pi}{2}\int_0^1 v\left[\operatorname{tg}\frac{\pi v}{2}+(-1)^{n+1}\frac{\sin(2n+1)\frac{\pi v}{2}}{\cos\frac{\pi v}{2}}\right]dv.$$

Benutzt man noch

(3. 48)

$$(-1)^n\int_0^1\frac{\cos(2n+1)\frac{\pi u}{2}}{\cos\frac{\pi u}{2}}\,du=\int_0^1\frac{\sin(2n+1)\frac{\pi x}{2}}{\sin\frac{\pi x}{2}}\,dx=1,$$

so folgt

$$\beta_{-1}(1,n)-\beta_{-1}(z,n)=-\frac{1}{2z}-\frac{\pi}{2}\operatorname{ctg}\pi z+$$

(3. 49)

$$+\frac{\pi}{2}\int_0^1\operatorname{tg}\frac{\pi u}{2}\left(u-\frac{\sin\pi z u}{\sin\pi z}\right)du+\frac{1}{z+n}+(-1)^{n+1}\frac{\pi}{2}\left[r_n(z)+\varrho_n(z)\right]$$

$$(z\neq 0,\pm 1,\pm 2,\ldots);$$

[40] Einen anderen, etwa ebenso kurzen Beweis ergibt die Benutzung des Satzes § 3, VI. 2.

da $\left(\cos\frac{\pi u}{2}\right)^{-1}\left(\frac{\cos\pi z u}{\sin\pi z}-\operatorname{ctg}\pi z\right)$ und $\left(\cos\frac{\pi u}{2}\right)^{-1}\left(u-\frac{\sin\pi z u}{\sin\pi z}\right)$ für $u\to 1-0$ endlich bleiben, strebt $r_n(z)$ und $\varrho_n(z)$ auf Grund des Riemannschen Lemmas für $n\to\infty$ gegen 0.

(3. 43) und (3. 48) führen also zu

$$\Psi(z)+\gamma=-\frac{1}{2z}-\frac{\pi}{2}\operatorname{ctg}\pi z+\frac{\pi}{2}\int_0^1\operatorname{tg}\frac{\pi u}{2}\left(u-\frac{\sin\pi z u}{\sin\pi z}\right)du. \tag{3.50}$$

Nun ist, wie eine leichte Umformung zeigt,

$$\frac{1}{z}+\pi\int_0^1\operatorname{tg}\frac{\pi u}{2}\left(\frac{\sin\pi z u}{\sin\pi z}-u\right)du= \tag{3.51}$$

$$=\pi\int_0^1\sec\frac{\pi u}{2}\left[\frac{\cos\pi u\left(z-\frac{1}{2}\right)}{\sin\pi z}-1\right]du+\pi\int_0^1\left(\sec\frac{\pi u}{2}-u\operatorname{tg}\frac{\pi u}{2}\right)du$$

und eine Teilintegration samt

$$\int_0^{\frac{1}{2}}\log(2\sin\pi x)\,dx=-\int_0^{\frac{1}{2}}\left(\sum_{\lambda=1}^{\infty}\frac{\cos 2\lambda\pi x}{\lambda}\right)dx=0 \tag{3.52}$$

ergibt

$$\pi\int_0^1\left(\sec\frac{\pi u}{2}-u\operatorname{tg}\frac{\pi u}{2}\right)du=\pi\lim_{\varepsilon\to+0}\int_\varepsilon^1\left[\operatorname{cosec}\frac{\pi v}{2}-(1-v)\operatorname{ctg}\frac{\pi v}{2}\right]dv=$$

$$=-2\lim_{\varepsilon\to+0}\left[\log\operatorname{tg}\frac{\pi\varepsilon}{4}-(1-\varepsilon)\log\sin\frac{\pi\varepsilon}{2}+\int_\varepsilon^1\log\sin\frac{\pi v}{2}\,dv\right]= \tag{3.53}$$

$$=2\lim_{\varepsilon\to+0}\log\left(2\cos^2\frac{\pi\varepsilon}{4}\right)-4\int_0^{\frac{1}{2}}\log\sin\pi x\,dx=4\log 2;$$

(3. 50), (3. 51), (3. 53) implizieren aber die zweite Gleichung unter (3. 36).

Zur Erledigung von (3. 37) und (3. 38) ist keine weitere Rechnung nötig; man erhält sie durch Substraktion aus (3. 35) und (3. 44) bzw. aus (3. 50) und (3. 49). Die Behauptung über die Reste ist offenbar der gleichmäßigen Gültigkeit von (3. 4) äquivalent, die wir schon bewiesen haben.

2. Auf Grund von (3. 18) und (3. 35) findet man

$$\gamma=\int_0^1[\Psi(x+1)+\gamma]\,dx=\int_0^1\int_0^1\frac{1-u^x}{1-u}\,du\,dx. \tag{3.54}$$

Es ist nun mühelos einzusehen, daß der Integrand im Rechteck $0 \leqq x \leqq 1$, $0 \leqq u \leqq 1$ *stetig* wird, wenn man jedem Punkte $(1, x)$ $(0 \leqq x \leqq 1)$ den entsprechenden Limes x als Funktionswert zuordnet; somit folgt durch Vertauschung der Reihenfolge der Integrationen unmittelbar (3. 41).

Ähnlich resultiert (3. 42) aus (3. 36).

Man braucht vor allem der Umformung (vgl. (3. 54), (3. 8)):

$$\gamma = \int_0^1 [\Psi(x+1) + \gamma]\, dx = \int_0^{\frac{1}{2}} [\Psi(x+1) + \Psi(2-x) + 2\gamma]\, dx =$$

$$= \int_0^{\frac{1}{2}} [\Psi(x) + \Psi(-x) + 2\gamma]\, dx + \log 2,$$

woher man durch (3. 50) und die Vertauschung der Integrationen nach den zwei Veränderlichen auf

$$\gamma = \pi \int_0^1 \operatorname{tg} \frac{\pi u}{2} \int_0^{\frac{1}{2}} \left(u - \frac{\sin \pi x u}{\sin \pi x} \right) dx\, du + \log 2 \tag{3. 55}$$

schließen kann.

Beachten wir noch (vgl. (3. 52))

$$\frac{\pi}{2} \int_0^1 (1-u) \operatorname{tg} \frac{\pi u}{2}\, du = 2\pi \int_0^{\frac{1}{2}} x \operatorname{ctg} \pi x\, dx = -2 \int_0^{\frac{1}{2}} \log \sin \pi x\, dx = \log 2, \tag{3. 56}$$

so entspringt, in der Tat, die Darstellung (3. 42).

Es sei darauf hingewiesen, daß man aus den eben bewiesenen Grundformeln mittels Modifizierung des Integrationsvariablen usw. zahlreiche andere herleiten kann.

Der Vergleich von (3. 41) und (3. 35) liefert z. B.

$$\Psi(z) = -\int_0^1 \left(\frac{u^{z-1}}{1-u} + \frac{1}{\log u} \right) du \qquad (\Re(z) > 0), \tag{3. 57}$$

die bei der Substitution $u = e^{-v}$ in eine *Formel von* GAUSS

$$\Psi(z) = \int_0^\infty \left(\frac{e^{-zv}}{e^{-v} - 1} + \frac{e^{-v}}{v} \right) dv \qquad (\Re(z) > 0) \tag{3. 58}$$

übergeht.

(3. 36) läßt sich, ähnlich wie (2. 40) mit $r=1$, in eine *komplexe Kreisintegralformel* umschreiben. Denn es ist nebst $0<\varepsilon<1$

$$\int_0^{1-\varepsilon}\frac{du}{\cos\frac{\pi u}{2}}=\int_{-1+\varepsilon}^{1-\varepsilon}\frac{du}{e^{i\frac{\pi u}{2}}+e^{-i\frac{\pi u}{2}}}=\frac{1}{i\pi}\overset{(\mathfrak{f}_1)}{\int_{-e^{i\pi\varepsilon}}^{-e^{-i\pi\varepsilon}}}\frac{dt}{t^{\frac{1}{2}}(t+1)},$$

$$\int_0^{1-\varepsilon}\frac{\cos\pi u\left(z-\frac{1}{2}\right)}{\cos\frac{\pi u}{2}}du=\int_0^{1-\varepsilon}\left(\frac{e^{i\pi zu}}{e^{i\pi u}+1}+\frac{e^{-i\pi zu}}{e^{-i\pi u}+1}\right)du=\frac{1}{i\pi}\overset{(\mathfrak{f}_1)}{\int_{-e^{i\pi\varepsilon}}^{-e^{-i\pi\varepsilon}}}\frac{t^{z-1}}{t+1}dt,$$

wobei $\mathfrak{f}_1$ den Kreisbogen $t=e^{i\pi u}$ $(-1+\varepsilon\leqq u\leqq 1-\varepsilon)$ bedeutet; daraus fließt aber

$$\psi(z)+\gamma+2\log 2+\frac{\pi}{2}\operatorname{ctg}\pi z=\frac{1}{2i}\oint\frac{1}{1+t}\left(t^{-\frac{1}{2}}-\frac{t^{z-1}}{\sin\pi z}\right)dt \tag{3. 59}$$
$$(z=0,\pm 1,\pm 2,\ldots)$$

mit $t^z=e^{z\log t}$, $t^{-\frac{1}{2}}=e^{-\frac{1}{2}\log t}$, erstreckt über den ganzen, im positiven Sinn durchlaufenen Einheitskreis.

VIII. Folgerungen über $\pi\operatorname{ctg}\pi z$ und $\psi^{(\nu)}(z)$.

1. *Man hat für jedes* $n=0,1,2,\ldots$

$$\pi\operatorname{ctg}\pi z=\begin{cases}\sum_{\lambda=0}^{n-1}{}'\left(\frac{1}{z+\lambda}+\frac{1}{z-\lambda-1}\right)+\int_0^1 u^n\frac{u^{z-1}-u^{-z}}{1-u}du & (0<\Re(z)<1),\\ \frac{1}{z}+\sum_{\lambda=1}^{n}{}'\left(\frac{1}{z+\lambda}+\frac{1}{z-\lambda}\right)+(-1)^{n+1}\pi r_n(z) & (z\neq 0,\pm 1,\pm 2,\ldots),\end{cases} \tag{3. 60}$$

ferner $(\nu\geqq 2)$

$$\psi^{(\nu-1)}(z)=\begin{cases}(-1)^\nu(\nu-1)!\sum_{\lambda=0}^{n-1}{}'\frac{1}{(z+\lambda)^\nu}+\int_0^1 u^{z-1}(\log u)^{\nu-1}\frac{u^n}{u-1}du & (\Re(z)>0),\\ (-1)^\nu(\nu-1)!\sum_{\lambda=0}^{n}{}'\frac{1}{(z+\lambda)^\nu}+(-1)^n\frac{\pi}{2}\left[r_n^{(\nu-1)}(z)+\varrho_n^{(\nu-1)}(z)\right] & (z\neq 0,\pm 1,\pm 2,\ldots)\end{cases} \tag{3. 61}$$

mit (3. 39) *und* (3. 40).

Die Reste konvergieren in jedem abgeschlossenen, endlichen Teil des entsprechenden Gültigkeitbereiches gleichmäßig gegen Null.[41]

2. *Es bestehen insbesondere* $(n=0)$ *für* $\nu=2,3,\ldots$

$$\zeta(\nu,z)=\frac{1}{(\nu-1)!}\int_0^1 u^{z-1}\left(\log\frac{1}{u}\right)^{\nu-1}\frac{1}{1-u}\,du \qquad (\Re(z)>0), \tag{3.62}$$

$$\zeta(\nu,z+1)= \tag{3.63}$$

$$=\frac{(-1)^\nu}{(\nu-1)!}\left[\frac{d^{\nu-1}}{dz^{\nu-1}}\left(\frac{1}{2z}-\frac{\pi}{2}\operatorname{ctg}\pi z\right)-\frac{\pi}{2}\int_0^1\operatorname{tg}\frac{\pi u}{2}\frac{\partial^{\nu-1}}{\partial z^{\nu-1}}\left(\frac{\sin\pi z u}{\sin\pi z}\right)du\right]$$

$$(z\neq 0,\pm1,\pm2,\ldots);$$

außerdem $(\mu=1,2,\ldots)$

$$\zeta(\nu)=\begin{cases}(-1)^{\mu-1}\dfrac{B_{2\mu}}{2(2\mu)!}(2\pi)^{2\mu}, & \text{wenn}\quad \nu=2\mu,\\[2ex] (-1)^{\mu+1}(2\pi)^{2\mu+1}\displaystyle\int_0^{\frac12}B_{2\mu+1}(u)\operatorname{ctg}\pi u\,du, & \text{falls}\ \nu=2\mu+1.\end{cases} \tag{3.64}$$

[42]

Beweis. (3.60) fließt unmittelbar aus den Formeln (3.37), (3.38) durch den Ergänzungssatz (3.14); (3.61), (3.62) und (3.63) resultieren einfach aus (3.35)—(3.38) durch Differentiation nach z (können aber natürlich auch aus (2.39), (2.40) für $n\to\infty$ erhalten werden). Daß die Ableitungen der (z als Parameter enthaltenden) Integrale formal, unter dem Integralzeichen, auszuführen sind, kann mühelos nachgewiesen werden. Dann die Integranden von (3.36), (3.39), (3.40) sind für $0<u<1$ beschränkte, analytische Funktionen von z ($z\neq 0,\pm1,\pm2,\ldots$) und für die in Betracht kommenden Wertepaare stetig in (z,u); in (3.35), (3.37) aber strebt der Faktor von u^{z-1} ja nach Multiplizieren mit einer positiven Potenz von u für $u\to+0$ gegen Null (vgl. die Fußnote[28]).

Schließlich entnimmt man (3.64) am leichtesten aus der Formel (2.42). Wir finden (vgl. (2.25), (3.48))

$$\beta_{-2\mu}(1,n)=(-1)^{\mu-1}(2\pi)\int_0^{\frac12}[B_{2\mu}(u)-B_{2\mu}(0)]\frac{\sin(2n+1)\pi u}{\sin\pi u}\,du+$$

$$+(-1)^{\mu-1}(2\pi)^{2\mu}\frac{B_{2\mu}}{2(2\mu)!},$$

[41] Für $n=0$ sind die in (3.60) und (3.61) rechts auftretenden Summen gleich Null zu setzen; unsere Formeln liefern dann einfache Integraldarstellungen, während für $n\to\infty$ in die Teilbruchzerlegungen (3.13) bzw. (3.15) übergehen.

[42] (3.64) entsteht auch aus (3.63) beim Grenzübergang $z\to 0$ (vgl. Mikolás [29]), so daß sie als ein Spezialfall derselben anzusehen mag; die Auswertung von $\zeta(\nu)$ $(\nu=2,3,\ldots)$ mit Hilfe der Bernoullischen Zahlen und Polynome geht auf Euler bzw. Sonin ([37], [38]) zurück. Über die Natur der Summen $\zeta(2\mu+1)$, gleichwie von γ, wissen wir auch heute nicht viel; vgl. hierzu z. B. Knopp [15], S. 246; Nielsen [30], S. 46; ferner Truesdell [40].

ferner

$$\beta_{-(2\mu+1)}(1,n) = (-1)^{\mu+1}(2\pi)^{2\mu+1}\int_0^{\frac{1}{2}} B_{2\mu+1}(u)\operatorname{ctg}\pi u\,du +$$

$$+(-1)^{\mu}(2\pi)^{2\mu+1}\int_0^{\frac{1}{2}}\frac{B_{2\mu+1}(u)}{\sin\pi u}\cos(2n+1)\pi u\,du,$$

weil die beiden letzten Integrale wegen $B_{2\mu+1}(0) = B_{2\mu+1}\left(\frac{1}{2}\right) = B_{2\mu+1}(1) = 0$ $(\mu = 1, 2, \ldots)$ (vgl. (2. 27), (2. 28), (2. 29)) für sich vorhanden sind. Beachtet man nun, daß die Faktoren von $\sin(2n+1)\pi u$ bzw. $\cos(2n+1)\pi u$ offenbar in $\left[0, \frac{1}{2}\right]$ beschränkt bleiben, so liefert das Riemannsche Lemma für $n\to\infty$ die Behauptung.

Was (3. 62) betrifft, so kann sie mit $u = e^{-v}$ in der Form

$$(3.65)\quad \zeta(\nu, z) = \sum_{\lambda=0}^{\infty}\frac{1}{(z+\lambda)^{\nu}} = \frac{1}{(\nu-1)!}\int_0^{\infty}\frac{e^{v(1-z)}}{e^{v}-1}v^{\nu-1}\,dv \qquad (\Re(z)>0;\ \nu = 2, 3, \ldots)$$

geschrieben werden, woher man für $z = 1$ die *Riemannsche Integraldarstellung von* $\zeta(\nu)$ erhält.[43]

IX. Trigonometrische Entwicklungen.

1. *Es gilt im Intervall* $0 < x < 1$

$$(3.66)\quad \Psi(x) = -\left[\gamma + \log(2\pi) + \frac{\pi}{2}\operatorname{ctg}\pi x + \sum_{\mu=1}^{\infty}\log\left(1+\frac{1}{\mu}\right)\frac{\sin(2\mu+1)\pi x}{\sin\pi x}\right]$$

und zwar in jedem abgeschlossenen Teilintervall desselben gleichmäßig.[44]

2. *Die Fourierentwicklung von* $\Psi^{(\nu-1)}(x)$ $(\nu = 1, 2, \ldots)$ *für ein Intervall* $\omega < x < \omega + 1$ $(\omega > 0)$ *lautet*

$$(3.67)\quad \Psi^{(\nu-1)}(x) = A_{\nu} + 2(\nu-1)!\sum_{n=1}^{\infty}(-2n\pi)^{\nu-1}[Ci_{\nu}(2n\pi\omega)\cos 2n\pi x + Si_{\nu}(2n\pi\omega)\sin 2n\pi x]$$

mit

$$(3.68)\quad A_{\nu} = \begin{cases} \log\omega, & \text{falls} \quad \nu = 1, \\ \dfrac{(\nu-2)!}{(-\omega)^{\nu-1}}, & \text{wenn} \quad \nu \geqq 2 \end{cases}$$

[43] Jedoch kommen diesmal nur ganzzahlige Werte von ν in Betracht.

[44] (3. 66) ist die sog. *Lerchsche Reihe* für $\Psi(x)\sin\pi x$ (vgl. [21]). Von einer trigonometrischen Entwicklung für $\Psi(x)$ selbst kann es offenbar keine Rede sein, da sie an $x = 0$ einen Pol besitzt.

und

$$(3.69)\qquad Ci_\nu(x)=-\int_x^\infty \frac{\cos u}{u^\nu}\,du, \qquad Si_\nu(x)=-\int_x^\infty \frac{\sin u}{u^\nu}\,du.^{45}$$

BEWEIS. 1. Die Funktion $\Psi(x)\sin \pi x$ ist für $0<x\leqq 1$ analytisch und hat für $x\to+0$ den Grenzwert $-\pi$ (vgl. (3. 3)); daraus erhellt, daß sie oder $\Phi(x)=\Psi(x)\sin\pi x+\frac{\pi}{2}\cos\pi x$ in $(0,1)$ konvergente Fourierentwicklungen zuläßt, also speziell die Sinusentwicklung

$$(3.70)\qquad \Phi(x)=\sum_{\lambda=1}^{\infty} b_\lambda \sin \lambda\pi x$$

mit

$$(3.71)\qquad b_\lambda=2\int_0^1 \Phi(x)\sin\lambda\pi x\,dx$$

bestehen muß.[46] Es bleibt nur die Berechnung der Koeffizienten übrig.

Dazu bemerken wir zunächst, daß man durch (3. 14) mühelos $\Phi(1-x)=$ $=\Phi(x)$ findet, d. h. das Bild von $\Phi(x)$ zur Geraden $x=\frac{1}{2}$ symmetrisch ist; somit folgt

$$(3.72)\qquad b_{2\mu}=2\left[\int_0^{\frac{1}{2}}\Phi(x)\sin 2\mu\pi x\,dx-\int_{\frac{1}{2}}^{0}\Phi(1-x)\sin 2\mu\pi(1-x)\,dx\right]=0$$

$$(\mu=1,2,\ldots).$$

Andererseits entnimmt man aus (3. 36) die Darstellung

$$(3.73)\qquad \Phi(x)=-(\gamma+2\log 2)\sin\pi x+$$

$$+\frac{\pi}{2}\int_0^1 \sec\frac{\pi u}{2}\left[\sin\pi x-\cos\pi u\left(x-\frac{1}{2}\right)\right]du \qquad (0<x<1)$$

und durch Benutzung derselben

$$(3.74)\qquad b_1=-2(\gamma+2\log 2)\int_0^1\sin^2\pi x\,dx+$$

$$+\pi\int_0^1\sec\frac{\pi u}{2}\int_0^1\left[\sin^2\pi x-\sin\pi x\cos\pi u\left(x-\frac{1}{2}\right)\right]dx\,du=$$

[45] Für den Fall $\nu=1$ vgl. NÖRLUND [37], S. 107; Ci_ν und Si_ν mögen Integralkosinus bzw. Integralsinus (im allgemeineren Sinn) heißen. — Nur der Kürze halber verzichten wir auf die Restglieder von (3. 66) und (3. 67) (die entsprechenden Dirichletschen Integrale).

[46] M. a. W.: wir bilden die Fourierreihe über $(-1,1)$ der aus $\Phi(x)$ gebildeten *ungeraden* Funktion.

$$= -(\gamma + 2\log 2) + \int_0^1 \left(\frac{\pi}{2} \sec \frac{\pi u}{2} - \frac{1}{1+u} - \frac{1}{1-u} \right) du =$$

$$= -(\gamma + 2\log 2) - \lim_{\varepsilon \to +0} \log \left[\left(\frac{2}{\varepsilon} - 1 \right) \operatorname{tg} \frac{\pi\varepsilon}{4} \right] = -\gamma - \log(2\pi),$$

(3. 75) $$b_{2\mu+1} = \pi \int_0^1 \sec \frac{\pi u}{2} \int_0^1 \sin(2\mu+1)\pi x \left[\sin \pi x - \cos \pi u \left(x - \frac{1}{2} \right) \right] dx\, du =$$

$$= \int_0^1 \left(\frac{1}{u-2\mu-1} - \frac{1}{u+2\mu+1} \right) du = \log \frac{\mu}{\mu+1} \qquad (\mu = 1, 2, \ldots).$$

Schreibt man die Werte (3. 72), (3. 74), (3. 75) in (3. 70) ein, so entsteht gerade (3. 66), multipliziert mit $\sin \pi x$.

Die Behauptung über gleichmäßige Konvergenz ist eine Folge der Tatsache, daß $b_{2\mu+1}$ für $\mu \to \infty$ *monoton* gegen 0 strebt.[47]

2. Es sei ω eine beliebige positive Zahl. Die zu $(\omega, \omega+1)$ gehörigen Fourierkoeffizienten (a_n, b_n) von $\Psi^{(\nu-1)}(x)$ können mittels (3. 4), (3. 8), (3. 18) und geeigneter Umformungen folgendermaßen bestimmt werden:

(3. 76) $$a_0 = \int_\omega^{\omega+1} \Psi^{(\nu-1)}(x)\, dx = \begin{cases} \log \omega & \text{für } \nu = 1, \\ \Psi^{(\nu-2)}(\omega+1) - \Psi^{(\nu-2)}(\omega) = (\nu-2)(-\omega)^{1-\nu} & \\ & \text{für } \nu = 2, 3, \ldots, \end{cases}$$

(3. 77) $$a_n = 2 \int_\omega^{\omega+1} \Psi^{(\nu-1)}(x) \cos 2n\pi x\, dx = 2(-1)^\nu (\nu-1)! \sum_{\lambda=0}^{\infty} \int_\omega^{\omega+1} \frac{\cos 2n\pi x}{(x+\lambda)^\nu} dx =$$[48]

$$= 2(-1)^\nu (\nu-1)! \int_\omega^\infty \frac{\cos 2n\pi v}{v^\nu} dv = 2(\nu-1)!\, (-2n\pi i)^{\nu-1} Ci_\nu(2n\pi\omega)$$

$$(n = 1, 2, \ldots)$$

und ähnlicherweise

(3. 78) $$b_n = 2 \int_\omega^{\omega+1} \Psi^{(\nu-1)}(x) \sin 2n\pi x\, dx = 2(\nu-1)!\, (-2n\pi i)^{\nu-1} Si_\nu(2n\pi\omega)$$

$$(n = 1, 2, \ldots).$$

Wegen der Regularität von $\Psi(x)$ $(x > 0)$ ist nun gewiß

(3. 79) $$\Psi^{(\nu-1)}(x) = a_0 + \sum_{n=1}^{\infty}{}' (a_n \cos 2n\pi x + b_n \sin 2n\pi x) \qquad (\omega < x < \omega+1),$$

w. z. b. w.

[47] Vgl. das bekannte Kriterium von DIRICHLET (z. B. ROGOSINSKI [35], S. 18).

[48] Wegen gleichmäßiger Konvergenz ist die gliedweise Integration der Teilbruchzerlegung von $\Psi^{(\nu-1)}(x)$ bekanntlich gestattet.

§ 4. Die Gammafunktion

Wie die Partialbruchzerlegungen (3. 3), (3. 4) und die Beziehungen (3. 18), (3. 19), (3. 20) zeigen, gelangt man aus $\Psi(z)$ durch Integration zu einer Funktion logarithmischen Charakters; damit ist man zur Idee geführt, $\exp \int \Psi(z)dz$ zu bilden. Wir betrachten genauer

$$\Gamma(z)=e^{\int_1^z \Psi(t)dt} \qquad (z \neq 0, -1, -2, \ldots) \tag{4. 1}$$

(*Gammafunktion*) mit einem rektifizierbaren Jordanschen Kurvenstück als Integrationsweg, welches zwischen 1 und z für $0 < \arg z \leqq \pi$ ganz oberhalb, für $-\pi < \arg z < 0$ ganz unterhalb der reellen Achse verläuft.[49] Somit ist nämlich wegen (3. 20)

$$\Gamma(\nu)=(\nu-1)! \qquad (\nu=1, 2, \ldots), \tag{4. 2}$$

d. h. (4. 1) bietet *eine äußerst einfache Fortsetzung der Fakultät* (als einer zahlentheoretischen Funktion) *ins Komplexe*. Wir schreiben zunächst die aus (3. 18) bzw. (3. 8) unmittelbar fließende Grundeigenschaft hinzu:

I. Differenzengleichung:

$$\Gamma(z+1)=z\Gamma(z) \qquad (z \neq 0, -1, -2, \ldots). \tag{4. 3}$$

Es verdient ferner den üblichen speziellen Wert (vgl. (3. 19))

$$\Gamma\left(\frac{1}{2}\right)=\pi^{\frac{1}{2}} \tag{4. 4}$$

zu erwähnen.

Benutzt man die (gleichmäßig gültige) Limesdefinition (3. 1) von $\Psi(z)$, so ergibt sich durch (2. 15)

$$\Gamma(z)=\lim_{n\to\infty}\left[n^{z-1}e^{-\int_1^z \beta_{-1}(t,\,n)\,dt}\right]=\lim_{n\to\infty}\left[n^{z-1}\frac{n!}{z(z+1)\cdots(z+n-1)}\right]= \tag{4. 5}$$

$$=\lim_{n\to\infty}\frac{n!\,n^z}{z(z+1)\cdots(z+n)} \qquad (z \neq 0, -1, -2, \ldots).$$

Das ist die *Gaußsche Form von* $\Gamma(z)$.

Integriert man, nach zweckmäßigen Umformungen mittels (3. 8), beide Seiten von (3. 3), $(3. 4)^{\nu 0}$ und berücksichtigt nebst (4. 1) $\Gamma(1)=1$, so folgt

[49] Das Integral hängt mithin vom Weg nicht ab und ist bekanntlich eine eindeutige, reguläre Funktion von z. — Wir erwähnen zugleich die Umkehrung von (4. 1): $\Psi(z)=\frac{d}{dz}\log\Gamma(z)=\frac{\Gamma'(z)}{\Gamma(z)}$.

nach Übergang zu den Exponenten

$$\Gamma(z+1)=\prod_{\lambda=1}^{\infty}\frac{\left(1+\frac{1}{\lambda}\right)^{z}}{1+\frac{z}{\lambda}}=e^{-\gamma z}\prod_{\lambda=1}^{\infty}\frac{e^{\frac{z}{\lambda}}}{1+\frac{z}{\lambda}}; \tag{4.6}$$

aus der letzten Beziehung entnimmt man durch (4.3) die *Weierstraßsche Produktdarstellung* von $\Gamma(z)^{-1}$:

$$\frac{1}{\Gamma(z)}=z\,e^{\gamma z}\prod_{\lambda=1}^{\infty}\left(1+\frac{z}{\lambda}\right)e^{-\frac{z}{\lambda}}. \tag{4.7}$$

(4.6) und (4.7) gelten offenbar für beliebige $z\neq 0,-1,-2,\ldots$ und zwar in jedem, die Ausnahmepunkte nicht enthaltenden, abgeschlossenen, endlichen Bereich gleichmäßig. (4.7) setzt in Evidenz, daß $\Gamma(z)^{-1}$ eine ganze (transzendente) Funktion ist, deren alle Nullstellen einfach und in $0,-1,-2,\ldots$ gelegen sind. Daraus erhellt aber, daß $\Gamma(z)$ *bis auf die genannten Stellen überall regulär und in* $z=-\lambda$ $(\lambda=0,1,2,\ldots)$ *mit einfachen Polen versehen, also meromorph ist.* Überdies ersieht man $\Gamma(z)\neq 0$ für jedes z und $\Gamma(z)>0$ für $z>0$.

Das Residuum an $0,-1,-2,\ldots$ errechnet sich bequem durch die aus (4.3) entspringende Relation

$$\Gamma(z+\lambda+1)=z(z+1)\cdots(z+\lambda)\Gamma(z) \qquad (\lambda=0,1,2,\ldots); \tag{4.8}$$

man erhält wegen $\Gamma(1)=1$

$$\operatorname*{res}_{z=-\lambda}\Gamma(z)=\lim_{z\to-\lambda}(z+\lambda)\Gamma(z)=$$
$$=\lim_{z\to-\lambda}\frac{\Gamma(z+\lambda+1)}{z(z+1)\cdots(z+\lambda-1)}=\frac{(-1)^{\lambda}}{\lambda!} \qquad (\lambda=0,1,2,\ldots). \tag{4.9}$$

Mithin liefert der Mittag-Lefflersche Satz die *Teilbruchzerlegung der Gammafunktion:*

$$\Gamma(z)=g(z)+\sum_{\lambda=0}^{\infty}{}'\frac{(-1)^{\lambda}}{\lambda!}\frac{1}{z+\lambda} \qquad (z\neq 0,-1,-2,\ldots), \tag{4.10}$$

wo $g(z)$ eine geeignete ganze Funktion bedeutet und die Reihe rechts offenbar in jedem endlichen, abgeschlossenen, die Ausnahmepunkte nicht enthaltenden Bereich gleichmäßig konvergiert.

Weitere Eigenschaften von $\Gamma(z)$ findet man einfach durch Integration aus den entsprechenden Beziehungen des vorangehenden Abschnittes; die dabei auftretenden Integrationskonstanten sind aber stets zu beachten.[52]

[50] Die gliedweise Integration ist durch die gleichmäßige Konvergenz der betreffenden Reihen ermöglicht. — Die Darstellungen (4.6) verdankt man Euler und Gauss bzw. Schlömilch.

[51] (4.10) rührt von Prym her.

[52] Die folgenden Sätze gelten überall, wo sie einen Sinn haben. — Bekanntlich hat (4.11) samt (4.3) schon Euler gefunden, während (4.13) für $z=0$ von Euler, für $m=2$ von Legendre, allgemein von Gauss festgestellt worden ist.

II. Ergänzungssatz:

$$\Gamma(z)\Gamma(1-z)=\frac{\pi}{\sin \pi z} \qquad (z \neq 0, \pm 1, \pm 2, \ldots). \tag{4.11}$$

BEWEIS. Es folgt aus (3.8) samt (4.1) und (4.4).

Die Formeln (4.11) und (4.7) oder (3.13) selbst implizieren die wichtige *Produktdarstellung der Sinusfunktion:*

$$\sin \pi z = \pi z \prod_{\lambda=1}^{\infty}\left(1-\frac{z^2}{\lambda^2}\right). \tag{4.12}$$

III. Multiplikationstheorem:

$$\prod_{k=0}^{m-1}\Gamma\left(z+\frac{k}{m}\right)=(2\pi)^{\frac{m-1}{2}} m^{\frac{1}{2}-mz}\Gamma(mz) \qquad (z \neq 0, -1, -2, \ldots). \tag{4.13}$$

BEWEIS. Der Fall $m=1$ ist trivial; es sei in den folgenden $m>1$ angenommen. Führt man in (3.9) die Gammafunktion hinein, so ergibt sich

$$\prod_{k=0}^{m-1}\Gamma\left(z+\frac{k}{m}\right)=C_m m^{-mz}\Gamma(mz) \tag{4.14}$$

mit einer vorläufig unbekannten, offenbar positiven Konstante C_m; es erübrigt nur noch die Berechnung derselben.

Man verwendet zu diesem Zweck (4.14) für $z=\frac{1}{m}$ und gewinnt wegen $\Gamma(1)=1$

$$\prod_{\lambda=1}^{m-1}\Gamma\left(\frac{\lambda}{m}\right)=\frac{1}{m}C_m, \tag{4.15}$$

oder anders geschrieben

$$\prod_{\lambda=1}^{m-1}\Gamma\left(1-\frac{\lambda}{m}\right)=\frac{1}{m}C_m. \tag{4.16}$$

Hieraus erschließt man durch (4.11)

$$\prod_{\lambda=1}^{m-1}\Gamma\left(\frac{\lambda}{m}\right)\Gamma\left(1-\frac{\lambda}{m}\right)=\prod_{\lambda=1}^{m-1}\frac{\pi}{\sin\frac{\lambda\pi}{m}}=\frac{1}{m^2}C_m^2. \tag{4.17}$$

Der Zerlegung

$$z^m-1=\prod_{\lambda=1}^{m-1}\left(z-e^{\frac{2\lambda\pi i}{m}}\right) \tag{4.18}$$

entnehmen wir aber speziell

$$\prod_{\lambda=1}^{m-1}\left|1-e^{\frac{2\lambda\pi i}{m}}\right|=\prod_{\lambda=1}^{m-1}\left(2\sin\frac{\lambda\pi}{m}\right)=\lim_{z\to 1}\left|\frac{z^m-1}{z-1}\right|=m,$$

also

$$\prod_{\lambda=1}^{m-1}\sin\frac{\lambda\pi}{m}=\frac{m}{2^{m-1}}; \tag{4.19}$$

setzt man dies in (4. 17) ein, so resultiert

$$C_m = m^{\frac{1}{2}} (2\pi)^{\frac{m-1}{2}},$$

w. z. b. w.

Nach (4. 1) kommt das Integral im Exponenten unter den (unendlich vielen) Logarithmus von $\Gamma(z)$ vor; wir schreiben weiterhin eindeutigerweise

$$\text{(4. 20)} \qquad \operatorname{Log}\Gamma(z) = \int_1^z \Psi(t)\,dt$$

indem nur die schon erwähnten Integrationswege zugelassen sind.[53]

Dann gilt die (3. 18) analoge *Raabesche Integralformel:*

$$\text{(4. 21)} \qquad \int_z^{z+1} \operatorname{Log}\Gamma(t)\,dt = z(\log z - 1) + \log\sqrt{2\pi} \qquad (|\arg z| < \pi),$$

welche man bekanntlich am schnellsten mittels (4. 13) herleiten kann.

Da nämlich (vgl. (3. 18))

$$\frac{d}{dz}\int_z^{z+1} \operatorname{Log}\Gamma(t)\,dt = \int_1^{z+1} \Psi(t)\,dt - \int_1^z \Psi(t)\,dt = \log z \qquad (|\arg z| \neq \pi),$$

also gewiß

$$\text{(4. 22)} \qquad \int_z^{z+1} \operatorname{Log}\Gamma(t)\,dt = z\log z - z + c$$

mit konstantem c besteht, ist es nur $c = \log\sqrt{2\pi}$ nachzuweisen.

Nun ergibt (4. 22) für $z > 0$, $z \to +0$

$$\text{(4. 23)} \qquad c = \int_0^1 \log\Gamma(u)\,du$$

(damit ist die Existenz des rechtsstehenden uneigentlichen Integrals gesichert); andererseits aus (4. 13) mit $z = \frac{1}{m}$ bekommt man, in der Tat,

$$\text{(4. 24)} \qquad \int_0^1 \log\Gamma(u)\,du = \lim_{m\to\infty} \frac{1}{m} \sum_{\lambda=1}^{m-1} \log\Gamma\left(\frac{\lambda}{m}\right) = \frac{1}{2}\log(2\pi).$$

Wie § 3, IV, erhellt ohne weiteres auch die folgende

IV. Potenzreihenentwicklung für $\operatorname{Log}\Gamma(z)$. *Man hat*

$$\text{(4. 25)} \qquad \begin{aligned} \operatorname{Log}\Gamma(z+h) = \operatorname{Log}\Gamma(h) + \Psi(h)z + \frac{1}{2}\zeta(2,h)z^2 + \cdots + \\ + \frac{(-1)^\nu}{\nu}\zeta(\nu,h)z^\nu + R^*_{\nu+1}(z,h)z^{\nu+1} \end{aligned}$$

[53] Es ist also jedenfalls $\operatorname{Log}\Gamma(1) = 0$, ferner $\operatorname{Log}\Gamma(z) = \log\Gamma(z)$ (reell) für $z > 0$.

mit

$$R_{\nu+1}^{*}(z,h)=\frac{1}{\nu!}\int_0^1 \psi^{(\nu)}(zu+h)(1-u)^{\nu}du \qquad (\nu=0,1,2,\ldots),$$

falls die Strecke $zu+h$ $(0\leqq u\leqq 1)$ *keinen Punkt der Folge* $0,-1,-2,\ldots$ *enthält; das Restglied strebt bei fixen* z *und* h *mit* $|z|<\min\limits_{\lambda=0,1,2,\ldots}|\lambda+h|$ *für* $\nu\to\infty$ *gegen* 0 *und zwar gleichmäßig in jedem inneren, konzentrischen Kreis. Speziell* $(h=1,\ \nu\to\infty)$

$$\operatorname{Log}\Gamma(z+1)=-\gamma z+$$

$$+\frac{\zeta(2)}{2}z^2-\frac{\zeta(3)}{3}z^3+-\cdots+(-1)^{\nu}\frac{\zeta(\nu)}{\nu}z^{\nu}+-\cdots \qquad (|z|<1).^{54} \tag{4.26}$$

Wir bemerken zunächst, daß man aus (4. 26) durch leichte Umformungen offenbar noch besser konvergente Entwicklungen gewinnen kann.

V. Asymptotische Entwicklung für $\operatorname{Log}\Gamma(z)$. *Es gilt für jedes positiv-ganze* N

$$\operatorname{Log}\Gamma(z)=\left(z-\frac{1}{2}\right)\log z-z+\log\sqrt{2\pi}+$$

$$+\sum_{\lambda=2}^{N}\frac{B_{\lambda}}{\lambda(\lambda-1)}\frac{1}{z^{\lambda-1}}-(N-1)!\int_0^{\infty}\frac{\bar{B}_N(x)}{(z+x)^N}dx \tag{4.27}$$

mit $|\arg z|\leqq\pi-\varepsilon$ $(0<\varepsilon<\pi)$; *das Restglied ist bei einem festen* $N\geqq 2$ $O\left(\frac{1}{|z|^{N-1}}\right)$. *Insbesondere besteht in jedem der genannten Sektoren*

$$\lim_{|z|\to\infty}\left[\operatorname{Log}\Gamma(z)-\left(z-\frac{1}{2}\right)\log z+z-\log\sqrt{2\pi}\right]=0, \tag{4.28}$$

während für $z=n+1$ $(n=0,1,2,\ldots)$

$$n!\sim\sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}.^{55} \tag{4.29}$$

[54] Natürlich kann auch $\Gamma(z)$ selbst nach dem Taylorschen Satz entwickelt werden, wobei aber die Koeffizienten sich nicht so einfach beherrschen lassen.

[55] (4. 27), (4. 28), (4. 29) pflegt man auf gleiche Weise als *Stirlingsche Formel* zu bezeichnen; die Allgemeingültigkeit von (4. 27) wurde erst von STIELTJES erkannt. Durch Umformung und zweckmäßige Entwicklung des Restes erhält man andere übliche Integraldarstellungen bzw. Reihendarstellungen für $\operatorname{Log}\Gamma(z)$ (BINET, BURNSIDE, GUDERMANN usw.). Für die Abschätzung des Restgliedes vgl. z. B. LÖSCH—SCHOBLIK [26].

BEWEIS. Die Integration von (3.23) liefert (vgl. (2.27))

$$\int_1^z \Psi(t)\,dt = \left(z-\frac{1}{2}\right)\log z - z + K + \tag{4.30}$$
$$+\sum_{\lambda=2}^{N} \frac{B_\lambda}{\lambda(\lambda-1)}\,\frac{1}{z^{\lambda-1}} - (N-1)!\int_0^\infty \frac{\overline{B}_N(x)}{(z+x)^N}\,dx$$

für jedes $N=1,2,\ldots$, falls der Integrationsweg ganz in einem Winkelraum $|\arg z| \leqq \pi-\varepsilon$ $(0<\varepsilon<\pi)$ verläuft; K bezeichnet hierbei eine (für alle in Betracht kommenden z eindeutig bestimmte) Konstante. Man hat nur ihren genauen Wert zu ermitteln, da die Behauptung über das Restglied und somit (4.28) und (4.29) wegen $|\overline{B}_N(x)| \leqq \frac{2}{(2\pi)^N}\zeta(N)$ $(N \geqq 2)$ (vgl. (2.36)) ja evident sind.

Verwenden wir (4.30) mit $N=2$ für $z=n+1$ und lassen $n\to\infty$ streben, so entspringt die Formel

$$K = \lim_{n\to\infty}\left[\log n! - \left(n+\frac{1}{2}\right)\log n + n\right];$$

setzt man noch in der Klammer $2n$ statt n, multipliziert beiderseits mit 2, so folgt nach Substraktion (vgl. (2.18), (3.19))

$$K = \lim_{n\to\infty}\left[2\log n! - \log(2n)! + 2n\log 2 - \frac{1}{2}\log n + \frac{1}{2}\log 2\right] =$$

$$= \lim_{n\to\infty}\left[\int_{\frac{1}{2}}^{1} \beta_{-1}(u,n)\,du - \frac{1}{2}\log n\right] + \frac{1}{2}\log 2 = \tag{4.31}$$

$$= \int_1^{\frac{1}{2}} \Psi(u)\,du + \frac{1}{2}\log 2 = \frac{1}{2}\log(2\pi),$$

w. z. b. w.

Es sei $h>0$, z beliebig, aber *fest*. Schreibt man in (4.28) h bzw. $z+h$ statt z, so ergibt eine Substraktion

$$\lim_{h\to\infty}\left[\operatorname{Log}\Gamma(z+h) - \operatorname{Log}\Gamma(h) - z\log h - \left(z+h-\frac{1}{2}\right)\log\left(1+\frac{z}{h}\right)+z\right] =$$

$$= \lim_{h\to\infty}\left[\operatorname{Log}\Gamma(z+h) - \operatorname{Log}\Gamma(h) - z\log h\right] = 0$$

oder

$$\Gamma(z+h) \sim \Gamma(h)\,h^z \qquad (h\to\infty); \tag{4.32}$$

eine *asymptotische Relation,* welche sich samt der speziellen ($h = n+1$, positiv-ganz)

$$\Gamma(z+n+1) \sim n!\, n^z \qquad (n \to \infty) \tag{4.33}$$

oft als nützlich erweist.[56]

Was das Verhalten im Unendlichen betrifft, so verdient es noch eine elementare Tatsache zu erwähnen, nämlich die *Beschränktheit* von $\Gamma(z)$ in jedem fixen Streifen $0 < \delta \leqq \Re(z) \leqq M$; dies leuchtet sofort ein, falls man die aus (4.6) wegen $|e^z| = e^{\Re(z)}$ und $|z| \geqq |\Re(z)|$ folgende Abschätzung

$$|\Gamma(z)| \leqq |\Gamma(\Re(z))| \qquad (z \neq 0, -1, -2, \ldots), \tag{4.34}$$

gleichwie die Stetigkeit von $\Gamma(z)$ auf der positiv-reellen Achse beachtet.

Wir zeigen nun, daß die Fundamentalgleichung (4.3) zusammen mit (4.33) oder der Regularität und Beschränktheit z. B. in $1 \leqq \Re(z) < 2$[57] schon ausreicht, die Gammafunktion zu charakterisieren. Genauer:

VI. Charakteristische Eigenschaften.

1. *$\Gamma(z)$ ist die einzige Lösung der Differenzengleichung*

$$F(z+1) = zF(z) \tag{4.35}$$

($z \neq 0, -1, -2, \ldots$, sonst beliebig), welche der Bedingung

$$\lim_{n\to\infty} \frac{F(z+n+1)}{n!\, n^z} = 1 \tag{4.36}$$

genügt.

2. *Befriedigt eine Funktion $F(z)$ für jedes $z \neq 0, -1, -2, \ldots$ die Gleichung* (4.35), *ferner ist sie im Streifen $1 \leqq \Re(z) < 2$ regulär und beschränkt, so muß $F(z) \equiv c\Gamma(z)$ (mit konstantem c) sein.*

Beweis. 1. Aus (4.35) folgt

$$F(z+n+1) = z(z+1)\ldots(z+n)\Gamma(z) \qquad (n = 0, 1, 2, \ldots),$$

woher man durch (4.36) und (4.5)

$$F(z) = \lim_{n\to\infty} \frac{n!\, n^z}{z(z+1)\ldots(z+n)} \cdot \lim_{n\to\infty} \frac{F(z+n+1)}{n!\, n^z} = \Gamma(z) \quad (z \neq 0, -1, -2, \ldots)$$

schließen kann.

2. Für eine einfache Überlegung bezüglich der anderen Behauptung, welche, ähnlich § 3, VI, den Liouvilleschen Satz benutzt, wollen wir auf die einschlägige neuere Literatur verweisen.[58]

[56] (4.33) kann übrigens auch durch (4.5) und (4.8) leicht erhalten werden.

[57] Ein beliebiger anderer Streifen $\omega \leqq \Re(z) < \omega + 1$ ($\omega > 0$) ist aber ebenso gut geeignet.

[58] Vgl. Knopp [16], S. 47; der dortige Beweis rührt von H. Wielandt her. — Die erste Hälfte dieses Satzes VI scheint mir vom Prym, die zweite aber von Scheeffer [36] gefunden worden zu sein.

VII. Integraldarstellungen für $\operatorname{Log}\Gamma(z)$, $\operatorname{Log}\sin\pi z$.

1. *Es gilt*

$$\operatorname{Log}\Gamma(z)=\int_0^1\frac{1}{\log u}\left[\frac{1-u^{z-1}}{1-u}-(z-1)\right]du \qquad (\Re(z)>0). \tag{4.37}$$

Ferner für $0<z<1$ *und beliebige nicht-reelle* z

$$\operatorname{Log}\Gamma(z)=\frac{1}{2}\left(\log\frac{\pi}{z}-\operatorname{Log}\sin\pi z\right)-z\log 2+ \tag{4.38}$$

$$+\pi z\int_0^1\int_0^{\frac{1}{2}}\operatorname{tg}\frac{\pi u}{2}\left(\frac{\sin\pi x u}{\sin\pi x}-\frac{\sin 2\pi x z u}{\sin 2\pi x z}\right)dx\,du$$

mit

$$\operatorname{Log}\sin\pi z=\pi\int_{\frac{1}{2}}^{z}\operatorname{ctg}\pi t\,dt;^{59} \tag{4.39}$$

der letzte Integrationsweg soll ein rektifizierbares, Jordansches Kurvenstück sein, dessen alle von $\frac{1}{2}$ *und* z *verschiedenen Punkte oberhalb, auf oder unterhalb der reellen Achse liegen, je nachdem* $0<\arg z<\pi$, $\arg z=0$ *bzw.* $-\pi<\arg z<0$ *ist. Mithin besteht die Beziehung*

$$\Gamma(z)=2^{-z}\sqrt{\frac{\pi}{z\sin\pi z}}\exp\left[\pi z\int_0^1\int_0^{\frac{1}{2}}\operatorname{tg}\frac{\pi u}{2}\left(\frac{\sin\pi x u}{\sin\pi x}-\frac{\sin 2\pi x z u}{\sin 2\pi x z}\right)dx\,du\right] \tag{4.40}$$

$$(0<z<1),$$

falls die Wurzel positiv gewählt wird.

Allgemein hat man $(n=0,1,2,\ldots)$

$$\operatorname{Log}\Gamma(z+1)=-\gamma z+\sum_{\lambda=1}^{n}\left[\frac{z}{\lambda}-\log\left(1+\frac{z}{\lambda}\right)\right]+P_n(z) \tag{4.41}$$

[59] Also speziell $\operatorname{Log}\sin\pi z=\log\sin\pi z$ (reell) für $0<z<1$ und $\operatorname{Log}\sin\pi z=\log\sin\pi z+2k\pi i$ (mit einem geeigneten ganzen k) allgemein. — Beachtet man, daß die Funktion $\sin\pi z$ mit 1 periodisch ist, ferner auf den Geraden $\Re(z)=2\nu-\frac{1}{2}$ $(\nu=0,\pm1,\pm2,\ldots)$ und nur auf diesen negativ wird, so kann man den Weg, welchen sie bei der Integration beschreibt, mühelos auf der Riemannschen log-Fläche verfolgen und somit den Wert von k genau bestimmen.

mit

$$(4.42)\quad P_n(z)=\begin{cases}\displaystyle\int_0^1 \frac{u^n}{1-u}\left(z-\frac{u^z-1}{\log u}\right)du & \text{für } \Re(z)>-1,\\[2ex] \displaystyle(-1)^n\frac{\pi z}{2}\int_0^1 [r_n(zu)+\varrho_n(zu)]\,du\ ^{60} & \text{für } -1<z<1 \text{ und jedes nicht-reelle } z.\end{cases}$$

2. *Es ist*

$$(4.43)\quad \operatorname{Log}\sin\pi z=\log\pi+\int_0^1\frac{1}{\log u}\left(\frac{u^{z-1}+u^{-z}-2}{1-u}+1\right)du \qquad (0<\Re(z)<1)$$

und allgemein $(n=0,1,2,\ldots)$

$$(4.44)\quad \log\sin\pi z=\log(\pi z)+\sum_{\lambda=1}^{n}\left[\log\left(1+\frac{z}{\lambda}\right)+\log\left(1-\frac{z}{\lambda}\right)\right]+Q_n(z)$$

mit

$$(4.45)\quad Q_n(z)=\begin{cases}\displaystyle\int_0^1 u^n\frac{u^z+u^{-z}-2}{(1-u)\log u}\,du & \text{für } 0<z<1 \text{ und nicht-reelle } z \text{ mit } -1<\Re(z)<1,\\[2ex] \displaystyle(-1)^{n+1}\pi z\int_0^1 r_n(zu)\,du & \text{für } 0<z<1 \text{ und jedes nicht-reelle } z.\end{cases}$$

3. *Die Restglieder* $P_n(z)$, $Q_n(z)$ *streben in jedem abgeschlossenen, endlichen Bereich der zugelassenen* z*-Werte für* $n\to\infty$ *gleichmäßig gegen Null.*

BEWEIS. 1. (4. 37) folgt unmittelbar aus (3. 57), wenn man unter dem Integralzeichen von 1 bis z integriert. Daß dies erlaubt ist, kann leicht nachgewiesen werden (vgl. Fußnote [28]); denn 1° der Integrand in (4. 37) für $u\to 1-0$ einen Grenzwert besitzt, was die Existenz von

$$J(z)=\int_0^1\frac{1}{\log u}\left[\frac{1-u^{z-1}}{1-u}-(z-1)\right]du$$

für $\Re(z)>0$ sichert; 2° es besteht $\lim_{u\to+0} u^{\vartheta}/(1-u)\log u=0$ $(\vartheta>0)$; 3° somit gilt $J'(z)=\Psi(z)$ $(\Re(z)>0)$; 4° $J(1)=\operatorname{Log}\Gamma(1)=0$.

Schreiben wir (3. 36) durch (3. 42) und (3. 56) in der Form:

$$(4.46)\quad \Psi(z)=-\frac{\pi}{2}\operatorname{ctg}\pi z-\frac{1}{2z}-\log 2+$$

$$+\frac{\pi}{2}\int_0^1\operatorname{tg}\frac{\pi u}{2}\left(2\int_0^{\frac{1}{2}}\frac{\sin\pi x u}{\sin\pi x}\,dx-\frac{\sin\pi z u}{\sin\pi z}\right)du \qquad (z\neq 0,\pm1,\pm2,\ldots),$$

[60] Vgl. (3. 39), (3. 40).

so liefert eine Integration von $\frac{1}{2}$ bis z [61]

$$(4.47)\quad \operatorname{Log}\Gamma(z)+\int_{\frac{1}{2}}^{z}\Psi(t)\,dt=-\frac{1}{2}(\operatorname{Log}\sin\pi z+\log z)-z\log 2+$$
$$+\frac{\pi}{2}\int_0^1 \operatorname{tg}\frac{\pi u}{2}\left(2z\int_0^{\frac{1}{2}}\frac{\sin\pi x u}{\sin\pi x}\,dx-\int_0^z\frac{\sin\pi t u}{\sin\pi t}\,dt\right)du;$$

$\Xi(z,u)=\int_0^z \frac{\sin\pi t u}{\sin\pi t}\,dt$ existiert gewiß, wenn der Integrationsweg geradlinig und $0<z<1$ oder z nicht-reell ist.[62] In diesem Fall ergibt eine leichte Integraltransformation und die Benutzung von (3.19) genau (4.38), welche beim Übergang zur Exponentialfunktion für $0<z<1$ offenbar in (4.40) übergeht.

(4.41) mit (4.42) erhält man einerseits aus (3.37) durch Integration von 1 bis z, indem z schließlich durch $z+1$ ersetzt wird, andererseits aus (3.38), falls man sie (vgl. (3.8)) in der Form

$$(4.48)\quad \Psi(z+1)=-\gamma+\sum_{\lambda=1}^{n}\left(\frac{1}{\lambda}-\frac{1}{z+\lambda}\right)+(-1)^n\frac{\pi}{2}[r_n(z)+\varrho_n(z)]$$
$$(z\neq 0,\pm 1,\pm 2,\ldots)$$

schreibt und dann von 0 bis z (geradlinig) integriert. Das Restintegral

$$(4.49)\quad \int_0^z[r_n(t)+\varrho_n(t)]\,dt=z\int_0^1[r_n(zu)+\varrho_n(zu)]\,du$$

ist wegen $\lim_{t\to 0}[r_n(t)+\varrho_n(t)]=0$ für $-1<z<1$ und jedes nicht-reelle z vorhanden.

2. Aus (3.18) und (3.14) gelangt man durch Integration von 1 bis z bzw. von $\frac{1}{2}$ bis z und durch Benutzung von (3.18) bzw. (3.19) zu den Relationen

$$(4.50)\quad \operatorname{Log}\Gamma(z+1)=\operatorname{Log}\Gamma(z)+\log z\quad (z\neq 0,-1,-2,\ldots),$$

$$(4.51)\quad \operatorname{Log}\Gamma(z)+\operatorname{Log}\Gamma(1-z)=\log\pi-\operatorname{Log}\sin\pi z$$
$$(0<z<1 \text{ oder } z \text{ nicht-reell});$$

sie implizieren

$$(4.52)\quad \operatorname{Log}\sin\pi z=\operatorname{Log}(\pi z)-[\operatorname{Log}\Gamma(1+z)+\operatorname{Log}\Gamma(1-z)].$$

[61] Für die hiesige Integration unter dem Integralzeichen vgl. Osgood [32], S. 307.

[62] Macht man auch von anderen Integrationswegen (z. B. Kreisbogen) Gebrauch, so läßt sich eine für jedes nicht-ganze z gültige (aber zugleich kompliziertere) Formel aus (4.47) gewinnen.

Verknüpft man nun (4. 51) mit (4. 37) bzw. (4. 32) mit (4. 41)—(4. 42), so entstehen gerade (4. 43) und (4. 44)—(4. 45).

Die Behauptung, daß $\lim_{n\to\infty} P_n(z) = 0$ und $\lim_{n\to\infty} Q_n(z) = 0$ in den besprochenen Bereichen gleichmäßig bestehen, ist ein Korollar der gleichmäßigen Gültigkeit von (4. 6) bzw. (4. 12).

Somit ist der Beweis vollendet.

Wir bemerken zunächst, daß (4. 41)—(4. 42) und (4. 44)—(4. 45) als *Verschärfungen der Produktdarstellungen* (4. 6), (4. 7) *bzw.* (4. 12) aufgefaßt werden können, indem man die entsprechenden logarithmusfreien Formeln

$$\Gamma(z+1) = e^{-\gamma z + P_n(z)} \prod_{\lambda=1}^{n} \frac{e^{\frac{z}{\lambda}}}{1+\frac{z}{\lambda}}, \tag{4. 53}$$

$$\sin \pi z = \pi z\, e^{Q_n(z)} \prod_{\lambda=1}^{n} \left(1-\frac{z^2}{\lambda^2}\right) \tag{4. 54}$$

betrachtet.

Was (4. 38) betrifft, so läßt sie z. B. eine Reihenentwicklung für $\log \Gamma(z)$ zu, welche für numerische Zwecke vielmehr brauchbar ist, als (4. 26).

Es ist nämlich, wie man sich davon mittels (2. 22) leicht überzeugt, für jedes u

$$\frac{\sin \pi t u}{\sin \pi t} = u + 2\sum_{\mu=1}^{\infty} (-1)^{\mu} B_{2\mu+1}\left(\frac{u+1}{2}\right)(2\pi t)^{2\mu} \qquad (|t|<1) \tag{4. 55}$$

und hiermit

$$\mathfrak{S}(z,u) = \int_0^z \frac{\sin \pi t u}{\sin \pi t}\, dt = uz + 2\sum_{\mu=1}^{\infty} (-1)^{\mu} \frac{(2\pi)^{2\mu}}{2\mu+1} B_{2\mu+1}\left(\frac{u+1}{2}\right) z^{2\mu+1}$$

nebst $|z|<1$ und geradlinigem Integrationsweg.

Setzt man dies in (4. 47) ein und beachtet die Darstellungen (3. 63) in der Form

$$\zeta(2\mu+1) = (-1)^{\mu} 2^{2\mu} \pi^{2\mu+1} \int_0^1 \operatorname{tg}\frac{\pi u}{2} B_{2\mu+1}\left(\frac{u+1}{2}\right) du \quad (\mu = 0,1,2,\ldots), \tag{4. 56}$$

so ergibt sich[63]

$$\begin{aligned} \log \Gamma(z) = {} & \frac{1}{2}(\log \pi - \log z - \log \sin \pi z) - \\ & - z \log 2 + z \sum_{\mu=1}^{\infty} \frac{\zeta(2\mu+1)}{2\mu+1}\left[\left(\frac{1}{2}\right)^{2\mu} - z^{2\mu}\right] \qquad (0<z<1), \end{aligned} \tag{4. 57}$$

[63] Die ausgeführte gliedweise Integration ist (gemäß leichten Abschätzungen) gestattet. (Vgl. MIKOLÁS [29], S. 156.)

oder, durch Verwendung von (4.50) und

$$\log(1+z)-\log(1-z)=2\sum_{\mu=0}^{\infty}\frac{1}{2\mu+1}z^{2\mu+1}\qquad(|z|<1),$$

die Entwicklung

$$\log\Gamma(z+1)=\frac{1}{2}\left(\log\frac{\pi z}{\sin\pi z}-\log\frac{1+z}{1-z}\right)+$$
$$+(1-\gamma)z-\sum_{\mu=1}^{\infty}\frac{\zeta(2\mu+1)-1}{2\mu+1}z^{2\mu+1}\qquad(0<z<1). \tag{4.58}$$

Das ist die sog. *Legendresche Reihe* für $\log\Gamma(z+1)$.

VIII. Trigonometrische Entwicklungen für $\log\Gamma(x)$. *Es gelten die Formeln*

$$\log\Gamma(x)=\frac{1}{2}\log\pi+(\gamma+\log\overline{2\pi})\left(\frac{1}{2}-x\right)-\frac{1}{2}\log\sin\pi x+$$
$$+\frac{1}{2}\sum_{\mu=2}^{\infty}\frac{\log\mu}{\mu}\sin 2\mu\pi x\qquad(0<x<1),^{64} \tag{4.59}$$

$$\log\Gamma(x)=\frac{1}{2}\log\frac{2\pi}{\omega}-\omega+x\log\omega+$$
$$+\frac{1}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}[Ci_1(2n\pi\omega)\sin 2n\pi x-Si_1(2n\pi\omega)\cos 2n\pi x]^{65} \tag{4.60}$$

und zwar in jedem abgeschlossenen Teilintervall von $(0,1)$ *bzw. im ganzen* $[\omega,\omega+1]$ *gleichmäßig.*

BEWEIS. Die fraglichen Beziehungen lassen sich fast unmittelbar aus (3.66), (3.67) entnehmen.

1° Integriert man in (3.66) beiderseits zwischen den Grenzen $\frac{1}{2}$ und x längs $(0,1)$ und beachtet die gleichmäßige Gültigkeit dieser Formel, so folgt (vgl. (3.19))

$$\log\Gamma(x)-\frac{1}{2}\log\pi=$$
$$=(\gamma+\log\overline{2\pi})\left(\frac{1}{2}-x\right)-\frac{1}{2}\log\sin\pi x-\sum_{\mu=1}^{\infty}\log\left(1+\frac{1}{\mu}\right)\int_{\frac{1}{2}}^{x}\frac{\sin(2\mu+1)\pi u}{\sin\pi u}\,du. \tag{4.61}$$

Nun liegt es nahe, die Identität

$$\sum_{\mu=1}^{m}[\log\mu-\log(\mu+1)]\frac{\sin(2\mu+1)\pi x}{\sin\pi x}=$$
$$=2\sum_{\mu=2}^{m}(\log\mu)\cos 2\mu\pi x-\log(m+1)\frac{\sin(2m+1)\pi x}{\sin\pi x} \tag{4.62}$$

[64] Das ist die berühmte sog. *Kummersche Reihe.*

[65] Für die Bezeichnungen vgl. § 3, IX. 2.

und die Abschätzung

$$\left|\int_{\frac{1}{2}}^{x} \frac{1}{\sin \pi u} \sin (2m+1)\pi u \, du\right| =$$

$$(4.63) \quad = \frac{1}{2m+1}\left|\frac{\cos (2m+1)\pi x}{\pi \sin \pi x} + \int_{\frac{1}{2}}^{x} \frac{\cos \pi u}{\sin^2 \pi u} \cos (2m+1)\pi u \, du\right| <$$

$$< \frac{1}{2m+1}\left(\frac{1}{\pi|\sin \pi x|} + \left|1 - \frac{1}{\sin \pi x}\right|\right) \qquad (0<x<1)$$

zu verwenden; man erhält

$$(4.64) \quad -\sum_{\mu=1}^{\infty} \log\left(1+\frac{1}{\mu}\right) \int_{\frac{1}{2}}^{x} \frac{\sin (2\mu+1)\pi u}{\sin \pi u} \, du = \sum_{\mu=2}^{\infty} \frac{\log \mu}{\pi \mu} \sin 2\mu\pi x$$

und mithin (4.59).

Da $\frac{\log \mu}{\mu}$ für $\mu = 3, 4, \ldots$ *monoton abnehmend* gegen Null strebt, ist auch die gleichmäßige Konvergenz von (4.64) in einem beliebigen Intervall $\varepsilon < x < 1-\varepsilon$ $(\varepsilon > 0)$ gesichert.[66]

2° Durch Integration bekommt man aus (3.67) mit $\nu = 1$[67]

$$\log \Gamma(x) = c_0 + x \log \omega +$$

$$(4.65) \quad + \frac{1}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} [Ci_1(2n\pi\omega) \sin 2n\pi x - Si_1(2n\pi\omega) \cos 2n\pi x]$$

$$(0 < \omega < x < \omega + 1),$$

wobei c_0 eine geeignete Konstante bezeichnet. Ihre Bestimmung gelingt ganz leicht, indem man von ω bis $\omega+1$ integriert und sich auf die Raabesche Formel (4.21) stützt; es resultiert

$$(4.66) \quad c_0 = \frac{1}{2} \log \frac{2\pi}{\omega} - \omega.$$

Die rechtsstehende Reihe in (4.65) ist ihrer Entstehung nach (als eine integrierte Fourierreihe) für jedes reelle x *gleichmäßig konvergent*[68] und stellt eine überall stetige, mit 1 periodische Funktion dar; letztere hat also für $x = \omega$ und $x = \omega + 1$ den Wert $\log \Gamma(\omega) - (c_0 + \omega \log \omega) = \log \Gamma(\omega+1) -$

[66] Vgl. z. B. Rogosinski [35], S. 18.

[67] Die gliedweise Integration der betreffenden Fourierreihe ist bekanntlich gestattet.

[68] Vgl. z. B. Rogosinski [35], S. 27; dies kann man auch durch das Weierstraßsche Kriterium mühelos nachprüfen, weil es für die Koeffizienten mittels Teilintegration $O\left(\frac{1}{n^2}\right)$ herauskommt.

$-[c_0+(\omega+1)\log\omega]$ und stimmt auf solche Art im ganzen Intervall $[\omega, \omega+1]$ mit $\log\Gamma(x)-(c_0+x\log\omega)$ überein, w. z. b. w.

Nun wollen wir uns mit der Teilbruchreihe in (4. 10) beschäftigen.

IX. Geschlossene Darstellungen der „unvollständigen Gammafunktion“ [69]

$$G(z)=\sum\frac{(-1)^\lambda}{\lambda!}\frac{1}{z+\lambda}.$$

Man hat

$$G(z)=\int_0^1 u^{z-1}e^{-u}\,du \qquad (\Re(z)>0) \tag{4. 67}$$

und

$$G(z)=\frac{\pi}{\sin\pi z}\int_0^1 e^{\cos\pi u}\cos(\pi u z+\sin\pi u)\,du \qquad (z\neq 0,\pm 1,\pm 2,\ldots). \tag{4. 68}$$

Allgemeiner gilt $(n=0, 1, 2, \ldots)$

$$G(z)=\sum_{\lambda=0}^{n}\frac{(-1)^\lambda}{\lambda!}\frac{1}{z+\lambda}+q_n(z) \tag{4. 69}$$

mit

$$q_n(z)=\begin{cases}\dfrac{(-1)^{n+1}}{n!}\displaystyle\int_0^1\int_0^1(1-x)^n u^{z+n}e^{-ux}\,dx\,du & \text{für } \Re(z)>0,\\[2ex] \dfrac{\pi}{n!\sin\pi z}\displaystyle\int_0^1\int_0^1(1-x)^n e^{x\cos\pi u}\cos(\pi u(z+n+1)+ & \\ \qquad +x\sin\pi u)\,dx\,du & \text{für jedes nicht-ganze } z.\end{cases} \tag{4. 70}$$

Die Relation $\lim\limits_{n\to\infty} q_n(z)=0$ *besteht gleichmäßig in jedem abgeschlossenen, endlichen Bereich der zugelassenen z-Werte.*

Beweis. Man benutzt aufs neue unmittelbar die Identitäten (2. 43), (2. 44), ferner die für jedes $n=0, 1, 2, \ldots$ und für alle ξ geltende Beziehung

$$e^\xi=\sum_{\lambda=0}^{n}\frac{\xi^\lambda}{\lambda!}+e^\xi\frac{\xi^{n+1}}{n!}\int_0^1 x^n e^{-\xi x}\,dx, \tag{4. 71}$$

welche mittels partieller Integration oder aus der Taylorschen Formel leicht gewonnen werden kann.

[69] Allgemeiner benutzt man dieselbe Bezeichnung für die (bis auf $z=0, -1, -2, \ldots$ überall existierende) Summe $\sum\limits_{\lambda=0}^{\infty}\frac{(-1)^\lambda}{\lambda!}\frac{\omega^{z+\lambda}}{z+\lambda}$ $(\omega>0)$, welche für $\Re(z)>0$ auch durch die Integraldarstellung $\int_0^\omega x^{z-1}e^{-x}\,dx$ definiert werden kann.

1° Aus (4. 10) entspringt einerseits durch (2. 43)

$$(4.72)\quad G(z)=\sum_{\lambda=0}^{\infty}\frac{(-1)^{\lambda}}{\lambda!}\frac{1}{z+\lambda}=\int_0^1 u^{z-1}\left(\sum_{\lambda=0}^{\infty}\frac{(-u)^{\lambda}}{\lambda!}\right)du=\int_0^1 u^{z-1}e^{-u}\,du$$
$$(\Re(z)>0),$$

da die gleichmäßige Konvergenz der hierauftretenden Potenzreihe die Vertauschung der Summation und Integration zuläßt; zugleich ist die Existenz des letzten (für $0<\Re(z)<1$ uneigentlichen) Integrals gesichert.

Andererseits folgt durch (2. 44) analog für jedes nicht-ganze z

$$(4.73)\quad \begin{aligned} G(z)&=\frac{\pi}{\sin\pi z}\sum_{\lambda=0}^{\infty}\frac{1}{\lambda!}\int_0^1\cos\pi u(z+\lambda)\,du= \\ &=\frac{\pi}{\sin\pi z}\int_0^1\left(\sum_{\lambda=0}^{\infty}\frac{\cos\pi u(z+\lambda)}{\lambda!}\right)du; \end{aligned}$$

so gilt aber für beliebige u, z

$$(4.74)\quad \sum_{\lambda=0}^{\infty}\frac{1}{\lambda!}e^{i\pi u(z+\lambda)}e^{i\pi uz+e^{i\pi u}},$$

also

$$(4.75)\quad \sum_{\lambda=0}^{\infty}\frac{\cos\pi u(z+\lambda)}{\lambda!}=\frac{1}{2}\left(e^{i\pi uz+e^{i\pi u}}+e^{-i\pi uz+e^{-i\pi u}}\right)=e^{\cos\pi u}\cos(\pi uz+\sin\pi u),$$

womit (4. 73) in (4. 68) übergeht.

2° Für die entsprechende n-te Partialsumme erhalten wir durch (2. 43) und (4. 71) ($\xi=-u$) ohne weiteres

$$(4.76)\quad \begin{aligned} \sum_{\lambda=0}^{n}\frac{(-1)^{\lambda}}{\lambda!}\frac{1}{z+\lambda}&=\int_0^1 u^{z-1}\left(\sum_{\lambda=0}^{n}\frac{(-u)^{\lambda}}{\lambda!}\right)du= \\ &=\int_0^1 u^{z-1}e^{-u}\left[1-\frac{(-u)^{n+1}}{n!}\int_0^1 x^n e^{ux}\,dx\right]du. \end{aligned}$$

Die Verwendung von (2. 44) liefert dagegen

$$(4.77)\quad \sum_{\lambda=0}^{n}\frac{(-1)^{\lambda}}{\lambda!}\frac{1}{z+\lambda}=\frac{\pi}{\sin\pi z}\int_0^1\left(\sum_{\lambda=0}^{n}\frac{\cos\pi u(z+\lambda)}{\lambda!}\right)du$$

und betreffs der Summe in der Klammer gewinnt man wegen (vgl. (4. 71))

$$(4.78)\quad \sum_{\lambda=0}^{n}\frac{1}{\lambda}e^{i\pi u(z+\lambda)}=e^{i\pi uz+e^{i\pi u}}\left[1-\frac{1}{n!}e^{i(n+1)\pi u}\int_0^1 x^n e^{-xe^{i\pi u}}\,du\right]$$

mit Rücksicht auf die Eulerschen Relationen einfach die Darstellung

$$\sum_{\lambda=0}^{n} \frac{1}{\lambda!} \cos \pi u(z+\lambda) = e^{\cos \pi u} \cos(\pi u z + \sin \pi u) - \tag{4.79}$$

$$- \frac{1}{n!} \int_0^1 (1-x)^n e^{x \cos \pi u} \cos((z+n+1)\pi u + x \sin \pi u) du.$$

Vergleichen wir (4.67) mit (4.76), ferner (4.68) mit (4.77) samt (4.79), so kann durch Substraktion gerade (4.69) mit den Restgliedern (4.70) gefolgert werden.

3° Die Behauptung, daß $q_n(z)$ für $n \to \infty$ gleichmäßig gegen Null strebt, ist ein Korollar der gleichmäßigen Gültigkeit von (4.10).

Es liegt übrigens auf der Hand, das Integral (4.68) in *ein über den Einheitskreis erstrecktes Kurvenintegral* umzuformen; wir schreiben (vgl. (4.75))

$$\int_0^1 e^{\cos \pi u} \cos(\pi u z + \sin \pi u) du = \frac{1}{2} \int_0^1 (e^{i\pi u z + e^{i\pi u}} + e^{-i\pi u z + e^{-i\pi u}}) du = \tag{4.80}$$

$$= \frac{1}{2} \int_{-1}^{1} e^{i\pi u z + e^{i\pi u}} du,$$

also, durch Benutzung der Substitution $t = e^{i\pi u}$,

$$G(z) = \frac{1}{2i \sin \pi z} \oint t^{z-1} e^t dt, \tag{4.81}$$

wobei $t^z = e^{z \log t}$ zu beachten ist.

Was die Bestimmung der bisher näher nicht geprüften ganzen Funktion $g(z)$ in (4.10) betrifft, so kommt man am natürlichsten Weise zum Ziel, indem man die *Differenzengleichung von* $G(z)$ betrachtet.[70] Man erhält nämlich durch (vgl. (4.67))

$$\int_0^1 (z u^{z-1} - u^z) e^{-u} du = [u^z e^{-u}]_{u=0}^{u=1} = \frac{1}{e} \qquad (\Re(z) > 0) \tag{4.82}$$

die Relation

$$z G(z) - G(z+1) = \frac{1}{e}, \tag{4.83}$$

welche, wie die Darstellung $G(z) = \sum_{\lambda=0}^{\infty} \frac{(-1)^\lambda}{\lambda!} \frac{1}{z+\lambda}$ zeigt, zugleich für jedes $z \neq 0, -1, -2, \ldots$ gilt.

[70] Es ist von vornherein zu erwarten, daß $G(z)$ einer (4.3) ähnlichen Beziehung genügt. — (4.83) rührt von PRYM her.

Wird nun eine Lösung der Differenzengleichung von $\Gamma(z)$ gesucht, so hat man offenbar nur die obere Grenze in $\int_0^1 u^{z-1}e^{-u}du$ auf solche Art zu modifizieren, daß $u^z e^{-u}$ auch an dieser verschwindet. Hiermit gelangen wir zu den folgenden

X. Integraldarstellungen von $\Gamma(z)$. *Es ist in* (4.10)

$$g(z)=\int_1^\infty u^{z-1}e^{-u}du \qquad (z \text{ beliebig}) \tag{4.84}$$

und so

$$\Gamma(z)=\int_0^\infty u^{z-1}e^{-u}du \qquad (\Re(z)>0),^{71} \tag{4.85}$$

$$\Gamma(z)=\frac{\pi}{\sin \pi z}\int_0^1 e^{\cos \pi u}\cos(\pi u z+\sin \pi u)du+\int_1^\infty u^{z-1}e^{-u}du \tag{4.86}$$

$$(z \neq 0, \pm 1, \pm 2, \ldots).$$

Beweis. Das Integral in (4.84) ist für jedes feste z wegen $|u^{z-1}e^{-u}| = u^{\Re(z)-1}e^{-u} = O\left(\frac{1}{u^2}\right)$ konvergent und stellt offenbar eine überall analytische (ganze) transzendente Funktion von z dar.[72]

Mithin ist

$$F(z)=G(z)+\int_1^\infty u^{z-1}e^{-u}du \tag{4.87}$$

bis auf $z=0, -1, -2, \ldots$ regulär; man hat die Relation (vgl. (4.84))

$$zF(z)-F(z+1)=\frac{1}{e}+[u^z e^{-u}]_{u=1}^{u=\infty}=0 \qquad (z \neq 0, -1, -2, \ldots) \tag{4.88}$$

und im Streifen $1 \leqq \Re(z) < 2$ die (übrigens grobe) Abschätzung

$$|F(z)| \leqq \int_0^\infty u^{\Re(z)-1}e^{-u}du < \int_0^1 e^{-u}du+\int_1^\infty ue^{-u}du = 1+\frac{1}{e}, \tag{4.89}$$

so daß die Anwendung des Satzes § 4, VI. 2

$$F(z)=c\Gamma(z) \qquad (c \text{ konstant}) \tag{4.90}$$

[71] Das ist das sog. *Eulersche Integral zweiter Gattung*. Man erhält daher mühelos eine Cauchysche Darstellung, welche für einen Streifen $-\nu-1<\Re(z)<-\nu$ $(\nu=0,1,2,\ldots)$ gilt, oder für $\Re(z)>0$ das sog. Laplacesche Integral.

[72] Vgl. z. B. Osgood [32], S. 307.

ergibt. Beachtet man noch, daß

$$F(1)=\int_0^\infty e^{-u}du=1,$$

also $c=1$ sein muß, so folgt

$$F(z)=G(z)+\int_1^\infty u^{z-1}e^{-u}du \qquad (z\neq 0,-1,-2,\ldots). \tag{4.91}$$

Daher ist die Gültigkeit von (4.84), (4.85), (4.86) wegen $g(z)=\Gamma(z)-G(z)$ und durch (4.67) bzw. (4.68) sofort ablesbar.

Es sei vor allem bemerkt, daß die Transformationen $u=-\log v$ und $u=e^v$ (4.85) in die gleichfalls üblichen Darstellungen

$$\Gamma(z)=\int_0^1\left(\log\frac{1}{v}\right)^{z-1}dv \tag{4.92}$$

bzw.

$$\Gamma(z)=\int_{-\infty}^{\infty}e^{vz-e^v}dv \tag{4.93}$$

$(\Re(z)>0)$

übertragen.[73]

Ist $\varkappa$ eine beliebig angegebene komplexe Zahl mit $\Re(\varkappa)>0$ und schreibt man in (4.85) $u=\varkappa t$, so folgt

$$\Gamma(z)=\varkappa^z\int_0^\infty t^{z-1}e^{-\varkappa t}dt \qquad (\Re(z)>0), \tag{4.94}$$

erstreckt über die Gerade $\arg t=\arg\frac{1}{\varkappa}$; mit Hilfe des Cauchyschen Integralsatzes zeigt man leicht, daß der Integrationsweg wieder die positiv-reelle Achse sein kann.

Inbetreff (4.86) ergibt die Benutzung von (4.68) und (4.81)

$$\Gamma(z)=\frac{1}{2i\sin\pi z}\oint t^{z-1}e^t dt+\int_1^\infty u^{z-1}e^{-u}du \quad (z\neq 0,\pm 1,\pm 2,\ldots). \tag{4.95}$$

Schreiben wir zweckmäßig $u=-t$ im letzten Glied, so entsteht wegen $u^z=e^{z\log u}$ $(\log u\gtrless 0)$

[73] Bei EULER und LEGENDRE tritt übrigens lediglich die Form (4.92) (für rationale z bzw. $z>0$) auf. — Wir erwähnen einen Gedanken, der die Betrachtung dieses Integrals, als einer Verallgemeinerung der Fakultät, genügend zu motivieren scheint: setzt man in (2.39) $z=1$ und $n=1$, so folgt die Darstellung $(\nu-1)!=\int_0^1\left(\log\frac{1}{u}\right)^{\nu-1}du$ $(\nu=1,2,\ldots)$.

$$\int_1^{\infty} u^{z-1}e^{-u}\,du = e^{-i\pi z}\int_{-1}^{-\infty} t^{z-1}e^{t}\,dt = \frac{1-e^{-2\pi iz}}{2i\sin\pi z}\int_{-1}^{-\infty} t^{z-1}e^{t}\,dt =$$

(4.96)

$$= \frac{1}{2i\sin\pi z}\left(\int_{-\infty}^{-1} t^{z-1}e^{t-2\pi iz}\,dt + \int_{-1}^{-\infty} t^{z-1}e^{t}\,dt\right);$$

hiermit gilt für jedes nicht-ganze z

$$\text{(4.97)}\qquad \Gamma(z) = \frac{1}{2i\sin\pi z}\left[\int_{-\infty}^{-1}(t^{z-1})_{-}e^{t}\,dt + \oint t^{z-1}e^{t}\,dt + \int_{-1}^{-\infty}(t^{z-1})_{+}e^{t}\,dt\right],$$

wobei die Integration in der Mitte auf den (im positiven Sinn durchlaufenen) Einheitskreis, sonst auf die negativ-reelle Achse zu beziehen, ferner (vgl. (4.81)) $t^{z-1} = e^{(z-1)\log t}$, $(t^{z-1})_{\pm} = e^{(z-1)(\log|t|\pm i\pi)}$ zu verstehen ist.[74]

Nun erkennt man gleich, daß es *allgemein nebst* $R>0$, $\frac{1}{2}<\delta\leqq 1$

$$\text{(4.98)}\qquad \Gamma(z) = \frac{1}{2i\sin\pi z}\left[\int_{\mathfrak{h}_{-\delta}} t^{z-1}e^{t}\,dt + \int_{\mathfrak{k}_R} t^{z-1}e^{t}\,dt + \int_{\mathfrak{h}_{\delta}} t^{z-1}e^{t}\,dt\right]$$
$$(z\neq 0, \pm 1, \pm 2, \ldots)$$

besteht, wo $\mathfrak{h}_{\pm\delta}$ die beiden durch $t = ue^{\pm i\delta\pi}$ $(R\leqq u\leqq\infty)$ gekennzeichneten, gemäß wachsendem bzw. abnehmendem u zu durchlaufenden Halbgeraden, $\mathfrak{k}_R$ aber den Kreisbogen $t = Re^{i\vartheta}$ $(-\delta\leqq\vartheta\leqq\delta)$ bedeutet.

In der Tat: 1. das erste und dritte Integral in (4.98) konvergiert für jedes in Betracht kommende R, δ und z wegen

$$\text{(4.99)}\qquad \int_{\mathfrak{h}_{\pm\delta}} t^{z-1}e^{t}\,dt = \pm e^{\pm i\delta\pi z}\int_{R}^{\infty} u^{z-1}e^{ue^{\pm i\delta\pi}}\,du$$

und zwar, wie die Abschätzung

$$\text{(4.100)}\qquad \left|u^{z-1}e^{ue^{\pm i\delta\pi}}\right| = u^{\Re(z)-1}e^{u\cos\delta\pi} = O\left(\frac{1}{u^2}\right)$$

zeigt, bei festgelegtem R und z für $\frac{1}{2}+\varepsilon\leqq|\delta|\leqq 1$ $(\varepsilon>0)$ gleichmäßig; 2. der Integrand ist bis auf $t\leqq 0$ überall eindeutig-regulär, so daß der Cauchysche Hauptsatz, ferner die nebst $\frac{1}{2}<\delta_1<\delta_2\leqq 1$, $\xi = \Re(z)$, $\eta = \Im(z)$ und für alle

[74] M. a. W.: denkt man die t-Ebene länge $t\leqq 0$ aufgeschlitzt, so beziehen sich die geradlinigen Integrale auf das obere bzw. untere Ufer des Schnittes.

$\mathfrak{k}_\varrho$: $t=\varrho e^{i\pi\vartheta}$ $(\varrho>0,\ \delta_1<|\vartheta|<\delta_2)$ geltende Abschätzung

$$(4.101)\qquad \left|(\mathfrak{k}_\varrho)\int_{\varrho e^{\pm i\delta_1\pi}}^{\varrho e^{\pm i\delta_2\pi}} t^{z-1}e^t\,dt\right| = \left|\varrho^z i\pi\int_{\pm\delta_1}^{\pm\delta_2} e^{i\pi z\vartheta+\varrho e^{i\pi\vartheta}}\,d\vartheta\right| \leqq \pi e^{\pi\eta}\varrho^\xi e^{\varrho\cos\delta_1\pi}{}^{75}$$

die *Unabhängigkeit* der Klammer in (4.98) von der Wahl der Parameter R und δ liefert; 3. da $(\mathfrak{h}_{\pm\delta})\int t^{z-1}e^t\,dt = J_\pm(\delta)$ für festes R und z auf Grund von (4.100) an $\delta=1$ stetig sind, geht die rechte Seite von (4.98) mit $R=1$ und $\delta\to 1$ in diejenige von (4.97) über.

Es verdient *eine andere Form von* (4.98) zu bemerken, in welcher nur über reelle Wege erstreckte Integrale vorkommen: wegen (vgl. (4.99))

$$(4.102)\qquad \begin{aligned}(\mathfrak{h}_{-\delta})\int t^{z-1}e^t\,dt+(\mathfrak{h}_\delta)\int t^{z-1}e^t\,dt &= \int_R^\infty u^{z-1}(e^{i\delta\pi z+ue^{i\delta\pi}}-e^{-i\delta\pi z+ue^{-i\delta\pi}})\,du = \\ &= 2i\int_R^\infty u^{z-1}e^{u\cos\delta\pi}\sin(\delta\pi z+u\sin\delta\pi)\,du\end{aligned}$$

und (vgl. (4.75))

$$(4.103)\qquad (\mathfrak{k}_R)\int t^{z-1}e^t\,dt = R^z\pi i\int_{-\delta}^{\delta} e^{i\pi\vartheta z+Re^{i\pi\vartheta}}\,d\vartheta = 2\pi i\cdot R^z\int_0^\delta e^{R\cos\pi\vartheta}\cos(\pi\vartheta z+R\sin\pi\vartheta)\,d\vartheta$$

erhält man für $R>0$, $\frac{1}{2}<\delta\leqq 1$ und *beliebiges* z

$$(4.104)\qquad \begin{aligned}\Gamma(z)\sin\pi z = \pi R^z\int_0^\delta e^{R\cos\pi u}\cos(\pi uz+R\sin\pi u)\,du+ \\ +\int_R^\infty u^{z-1}e^{u\cos\delta\pi}\sin(\delta\pi z+u\sin\delta\pi)\,du,\end{aligned}$$

woraus (4.86) sich speziell für $R=\delta=1$ ergibt.

Schließlich sei erwähnt, daß (4.86), (4.97), (4.98), (4.104) mittels $\Gamma(z)\sin\pi z=\dfrac{\pi}{\Gamma(1-z)}$ (vgl. (4.11)) zugleich in (für alle z gültige) *Integraldarstellungen der reziproken Gammafunktion* umgeschrieben werden können, ferner, daß der (aus $\mathfrak{h}_{-\delta}$, $\mathfrak{k}_R$, $\mathfrak{h}_\delta$ zusammengesetzte) Integrationsweg in (4.98) sich auf Grund des Cauchyschen Hauptsatzes und von (4.101) durch einen beliebigen Weg $\mathfrak{L}$ ersetzen läßt, welcher 1° vom Punkte ∞ ausgeht und bis

[75] Die letzte obere Schranke strebt bei festem z und $\frac{1}{2}<\delta_1\leqq 1$ für $\varrho\to\infty$ offenbar gegen Null.

zu einem Punkt $\omega > 0$ in der unteren Halbebene verläuft, 2° dann durch die obere Halbebene wieder in den Punkt ∞ zurückgeht und 3° indessen ganz innerhalb einer Kreisscheibe $|t| \leqq \varrho$, gleichwie eines Winkelraums $|\arg t| \geqq$ $\geqq \delta\pi > \frac{\pi}{2}$ bleibt.

Mithin hat man die Darstellung (vgl. (4.98))

$$\frac{1}{\Gamma(z)} = \frac{1}{2\pi i}\,(\mathfrak{L})\int t^{-z}e^{t}\,dt \qquad (z \text{ beliebig}) \tag{4.105}$$

als eine *Erweiterung des Hankelschen Schleifenintegrals* für $\Gamma(z)^{-1}$.[76]

(Eingegangen am 21. Dezember 1954.)

Literaturverzeichnis

[1] W. P. Alexejewsky, Über die Funktionen, welche den Gammafunktionen ähnlich sind, *Sammlung Mitteil. math. Ges. Charkow* (2), **1** (1890), S. 169—238 (russisch).
[2] E. Artin, Einführung in die Theorie der Gammafunktion, *Hamburger math. Einzelschriften*, **11** (1931).
[3] E. W. Barnes, The theory of the gamma-function, *Messenger* (2), **29** (1899), S. 64—128.
[4] E. W. Barnes, On the theory of the multiple gamma-function, *Cambr. Trans.*, **19** (1904), S. 374—425.
[5] L. Bendersky, Sur la fonction gamma généralisée, *Acta Math.*, **61** (1933), S. 263—322.
[6] O. Borůvka, Sur les racines imaginaires de l'équation $\Gamma(z) = a$, *Publ. Fac. Sci. Univ. Masaryk*, **26** (1923), (böhmisch).
[7] G. Brunel, *Bestimmte Integrale*, Enzykl. math. Wiss. II, 1 (1), 3 (Leipzig, 1916), S. 135—188.
[8] A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, *Higher transcendental functions*. I—II (New York—Toronto—London, 1953).
[9] L. Ginzel, Die konforme Abbildung durch die Gammafunktion, *Acta Math.*, **56** (1931), S. 273—353.
[10] *Correspondance d'*Hermite *et de* Stieltjes. I—II (Paris, 1905).

[76] Bei jener Formel ist der Integrationsweg wesentlich spezieller, indem nur eine solche „Schleife" in Betracht kommt, welche durch zwei mit der negativ-reellen Achse parallele Geraden ins Unendliche (also in $-\infty$) reicht. Man pflegt übrigens die erwähnte Darstellung auch mit den Namen von Weierstrass, Pochhammer und Heine in Verbindung zu bringen. — Wie bekannt, gibt es eine große Anzahl bestimmter Integrale, welche mit Hilfe der soeben angegebenen Darstellungen durch die Gammafunktion ausdrückbar sind; in dieser Hinsicht ist vor allem das sog. Eulersche Integral erster Gattung („Betafunktion")

$$\int_0^1 u^{x-1}(1-u)^{y-1}\,du = \Gamma(x)\,\Gamma(y)/\Gamma(x+y)$$ zu erwähnen.

[11] T. H. GRONWALL, The gamma-function in the integral calculus, *Annals of Math.* (2), **20** (1918), S. 35—124.

[12] J. L. W. V. JENSEN, An elementary exposition of the theory of the gamma-function (transl. from the Danish by T. H. GRONWALL), *Annals of Math.* (2), **17** (1916), S. 124—166.

[13] N. S. KOSCHLIAKOFF, Sur quelques applications du calcul des résidus, *Proc. Math. Labor. Crimean (Tauric) Univ.*, **1** (1920), S. 189—254; **2** (1921), S. 1—79 (russisch).

[14] N. S. KOSCHLIAKOFF, Über eine Summenformel, *Math. Ann.*, **90** (1923), S. 26—29.

[15] K. KNOPP, *Theorie und Anwendung der unendlichen Reihen* (Berlin—Heidelberg, 1947), IV. Auflage.

[16] K. KNOPP, *Funktionentheorie.* II (Berlin, 1949), VII. Auflage, S. Göschen.

[17] M. LERCH, Über die Fundamental-Eigenschaften der Eulerschen Integrale, *Prag. Ber.*, (1889), S. 188—222 (böhmisch).

[18] M. LERCH, Sur une point concernant la théorie de la fonction gamma, *Prag. Ber.*, (1893), S. 1—8.

[19] M. LERCH, Theorie der Gammafunktion, *Vestnik Čes. Akad.*, **2** (1893), S. 238—247, 305—317, 462—472; **8** (1899), S. 308—324 (böhmisch).

[20] M. LERCH, Weitere Studien auf dem Gebiete der Malmsténschen Reihen, *Rožpr. Čes. Akad.*, **3** (1894), 28, S. 1—63 (böhmisch).

[21] M. LERCH, Sur la différentiation d'une classe de séries trigonométriques, *Ann. Éc. Norm.* (3), **12** (1895), S. 351—361.

[22] M. LERCH, Verschiedenes über die Gammafunktion, *Rožpr. Čes. Akad.*, **5** (1896), 14. S. 1—37 (böhmisch).

[23] M. LERCH, Über einige Entwicklungen auf dem Gebiete der unvollständigen Eulerschen Integrale zweiter Art, *J. für Math.*, **128** (1905), S. 211—221.

[24] M. LERCH, Einige Reihenentwicklungen der unvollständigen Gammafunktion, *J. für Math.*, **130** (1905), S. 47—65.

[25] M. LERCH, Über einige Punkte der Theorie der Eulerschen Integrale, *Monatsh.*, **19** (1908), S. 119—147.

[26] F. LÖSCH und F. SCHOBLIK, *Die Fakultät (Gammafunktion) und verwandte Funktionen* (Leipzig, 1951).

[27] HJ. MELLIN, Abriß einer einheitlicher Theorie der Gamma- und der hypergeometrischen Funktionen, *Math. Ann.*, **68** (1910), S. 305—377.

[28] M. MIKOLÁS, Sur une extension de la formule d'Euler—Maclaurin, se rapportant à des intégrales curvilignes complexes, *Comptes Rendus du Premier Congrès des Math. Hongrois* (Budapest, 1952), S. 541—550.

[29] M. MIKOLÁS, Über die Beziehung zwischen der Gammafunktion und den trigonometrischen Funktionen, *Acta Math. Acad. Sci. Hung.*, **4** (1953), S. 145—159.

[30] N. NIELSEN, *Handbuch der Theorie der Gammafunktion* (Leipzig, 1906).

[31] N. E. NÖRLUND, *Differenzenrechnung* (Berlin, 1924).

[32] W. F. OSGOOD, *Lehrbuch der Funktionentheorie* (Leipzig—Berlin, 1912).

[33] B. VAN DER POL, Note on the gamma function, *Canadian J. Math.*, **6** (1954), S. 18—22.

[34] A. PRINGSHEIM, Zur Theorie der Gammafunktion, *Math. Ann.*, **31** (1888), S. 455—481.

[35] W. ROGOSINSKI, *Fouriersche Reihen* (Berlin—Leipzig, 1930), S. Göschen.

[36] L. SCHEEFFER, Zur Theorie der Funktionen $\Gamma(z), P(z), Q(z)$, *J. für Math.*, **97** (1884), S. 230—242.

[37] N. J. SONIN, Bernoullische Polynome und ihre Anwendungen, *Warsch. Univ. Nachr.*, (1888) (russisch).

[38] N. J. SONIN et CH. HERMITE, Sur les polynomes de Bernoulli. (Extrait d'une correspondance entre Sonin et Hermite), *J. für Math.*, **116** (1896), S. 133—156.

[39] E. C. TITCHMARSH, *The theory of the Riemann zeta-function* (Oxford, 1951).

[40] C. TRUESDELL, Generalization of Euler's summation of the series $\sum_{n=1}^{\infty} n^{-2n}$, $\sum_{n=1}^{\infty}(-1)^n \cdot (2n+1)^{-2n-1}$ etc., *Annals of Math.* (2), **46** (1945), S. 194—195.

[41] W. WIRTINGER, Einige Anwendungen der Euler—Maclaurinschen Summenformel, insbesondere auf eine Aufgabe von Abel, *Acta Math.*, **26** (1902), S. 255—272.

[42] E. T. WHITTAKER and G. N. WATSON, *A course of modern analysis* (Cambridge, 1927), IV. Auflage.

К ТЕОРИИ ГАММА-ФУНКЦИИ, ζ-ФУНКЦИИ РИМАНА И ДРУГИХ, РОДСТВЕННЫХ ИМ ФУНКЦИЙ. I

М. Миколаш (Будапешт)

(Резюме)

Ввиду важности упомянутых функций понятно, что их обширная литература содержит не только работы, излагающие вполне новые результаты, но и такие, которые лишь с новой точки зрения трактуют классическую теорию. Ограничиваясь лишь рассмотрением литературы последних десятилетий, можно указать прежде всего на работы Алексеевского, Артина, Бернса, Бендерского, Гронвалла, Йенсена, Лерха, Меллина, Прингсхейма, а также на монографии [42], [26], [8], [39].

В основе настоящей работы лежат новые результаты работы [29] о гамма-функции и о других родственных ей функциях. Автору удалось дальше развить эти результаты в разных направлениях и вставить их в сжатое новое построение всей теории.

Исходным пунктом служит исследование рациональной функции $\beta_p(z,n)=\sum_{k=0}^{n-1}(z+k)^p$ при любом целом показателе p—обобщающей сумму степеней $1^p+2^p+\cdots+n^p$ (§ 2). После этого в § 3 исследуются мероморфная функция $\psi(z)=\lim_{n\to\infty}[\log n-\beta_{-1}(z,n)]$ и производные $\psi^{(\nu-1)}(z)=(-1)^\nu(\nu-1)!\lim_{n\to\infty}\beta_{-\nu}(z,n)$, а в § 4 функция $\log\Gamma(z)=\int_1^z\psi(t)\,dt$ и гамма-функция

$$\Gamma(z)=\exp\int_1^z\psi(t)\,dt=\lim_{n\to\infty}\left[n^{z-1}\,e^{-\int_1^z\beta_{-1}(t,n)\,dt}\right]=\lim_{n\to\infty}\frac{n!\,n^z}{z(z+1)\cdots(z+n)}.$$

Каждая глава начинается с соответствующим разностным уравнением и вытекающими из него другими функциональными уравнениями. После этого следуют различные разложения в ряд, интегральные представления, простейшие теоремы охарактеризования и асимптотические разложения. В § 1 легко получаемая производящая функция (1.1) делает удобным введение полиномов Бернулли в качестве вспомогательных функций на основе (1.2); тогда имеем $\beta_p(z,n)=p![B_{p+1}(z+n)-B_{p+1}(z)]$ $(p=0,1,2,\ldots)$, и одновременно легко получаем соотношения для $B_\lambda(z)$, и формулу суммирования Эйлера—Маклорена, которая нам часто понадобится в дальнейшем (в особенности в связи с

асимптотическими и интегральными представлениями). Интегральные формулы для упомянутых нецелых функций выводятся непосредственно из соответствующего определения суммы, на основе тождеств

$$\frac{1}{z+k}=\int_0^1 u^{z+k-1}\,du \qquad (\Re(z)>0)$$

и

$$\frac{1}{z+k}=\frac{(-1)^k\pi}{\sin\pi z}\int_0^1 \cos(z+k)\pi u\,du \qquad (z\neq 0,\pm 1,\pm 2\ldots);$$

таким образом для каждой из этих формул получаются два представления „в конечном виде", одно из которых справедливо в полуплоскости $\Re(z)>0$, а другое для всех нецелых значений z. (Как известно, все интегральные представления классической теории — за исключением одного, т. н. интеграла Ганкеля — справедливы только в одной полуплоскости.)

Из преимуществ указанного метода изложения можно отмечать следующие: 1. Все встречающиеся представления в виде ряда или произведения можно снабжать остаточным членом; некоторый класс получаемых таким образом представлений — разложения в частичные дроби — переходит в простейшем случае в классические интегральные формулы (Эйлера, Лежандра, и т. д.), являясь таким образом обобщением этих последних. 2. Асимптотические разложения (в их числе и формула Стирлинга) получаются нетрудным вычислением из соответствующих формул для функций $\beta_{-\nu}(z, n)$, эти же формулы можно непосредственно записать, следуя Эйлеру—Маклорену. 3. В соответствии со сказанным получается ряд новых интегральных представлений для вещественных путей интегрирования (с тригонометрическими функциями в интегранде); эти представления — в отличие от известных до сих пор — остаются в силе для всей плоскости z. 4. Упомянутые интегралы простыми преобразованиями переводятся в комплексные криволинейные интегралы по замкнутым или незамкнутым контурам, так что становится излишним применение относящейся сюда идеи Римана из теории функций. 5. Всегда легко установить, какое именно значение (бесконечнозначного) комплексного логарифма или какую ветвь функции $\operatorname{Log}\Gamma(z)$ следует взять, между тем как в существующей до сих пор литературе этот вопрос часто не получал удовлетворительное освещение, что иногда даже приводил к недоразумениям. 6. Вдобавок удобно получаются: разложение котангенса и синуса на простейшие дроби и в бесконечное произведение соответственно (с остаточными членами); выражения в конечном виде (с помощью чисел и полиномов Бернулли) для сумм рядов $\zeta(\nu)=\sum_{\lambda=1}^{\infty}\frac{1}{\lambda^{\nu}}$ $(\nu=2, 3, \ldots)$ принадлежащие Эйлеру и Сонину; новое представление для константы γ Эйлера—Маскерони; новое, очень простое доказательство для т. н. тригонометрического разложения Лерха и вместе с тем для ряда Куммера, и т. д.

Так, например, для $\Psi(z)=\frac{d\operatorname{Log}\Gamma(z)}{dz}=\frac{\Gamma'(z)}{\Gamma(z)}$ получим представления (1. 3), где значение $Q_n(z)$ то же, что и в (1. 4). (См. § 3, VII.) — Для значений ζ-функции Гурвица $\zeta(\nu+1, z+1)=\sum_{k=1}^{\infty}\frac{1}{(z+k)^{\nu+1}}$ отсюда вытекает (1. 5). Само соотношение (1. 3) переходит для $n=0$ в формулу Лежандра (1. 6) или соответственно в новое представление (1. 7), между тем как для $n\to\infty$ из него получается разложение на простейшие дроби функции $\Psi(z)$. Из следствий (1. 6) и (1. 7) следует указать на (1. 9), (1, 10) и (1. 11). (1. 7) легко преобразуется в комплексный контурный интеграл: в (1. 12) следует взять

за контур интегрирования единичную окружность $t = e^{i\pi u}$ $(-1 \leqq u \leqq 1)$, описанную в положительном направлении.

Другим примером служит представление (1. 13), где $g(z)$ означает целую функцию $\int_0^\infty u^{z-1} e^{-u} du$, а $q_n(z)$—один из интегралов (1. 14). Положив здесь $n = 0$, получим эйлеровый интеграл (1. 15) или соответственно новую интегральную формулу (1. 16); если же $n \to \infty$, то получается т. н. разложение Прым'а $\Gamma(z)$ на простейшие дроби. Из (1. 16) для $t = e^{i\pi u}$ непосредственно вытекает (1. 18), где пути интегрирования суть: для среднего члена единичная окружность (от -1 до $+1$ в положительном направлении), а для первого и третьего члена верхний и нижний край соответственно сечения вдоль $t \leqq 0$ римановой поверхности для логарифма. Применяя основную теорему Коши, теперь уже нетрудно получить представление (1. 20), где путь интегрирования $\mathfrak{L}$—произвольная „петля", находящаяся от ∞ до точки $\omega > 0$ в нижней полуплоскости плоскости t, а отсюда обратно к ∞ в верхней полуплоскости, оставаясь при этом в некоторой области $|t| \leqq \varrho$, $|\arg t| \geqq \delta\pi > \frac{\pi}{2}$ (ϱ, δ—постоянные). (1. 20) является обобщением т. н. интегральной формулы Ганкеля.

Из других новых результатов отметим еще представление (1. 21) справедливое для $0 < z < 1$ (§ 4, VII), а также теоремы охарактеризования из § 3, VI. 2, доказанные с помощью теорем Лиувилля и Миттаг-Леффлера. По этим теоремам функция $\Psi(z)$ и любая из ее производных характеризуются соответствующим разностным уравнением, дополненным условием регулярности и ограниченности для некоторого полоса.

ON APPROXIMATELY DIFFERENTIABLE FUNCTIONS OF TWO VARIABLES

By
LÁSZLÓ GEHÉR (Miskolc)
(*Presented by* G. ALEXITS)

A. KHINTCHINE [1][1] proved the following theorem: In order that a finite measurable function $f(x)$ defined on a measurable set[2] $M \subset R_1$ may be approximately derivable at almost all the points of this set, it is necessary and sufficient that the set M shall be representable as union of a set of measure zero and of a finite or enumerable monotone ascending sequence of measurable sets on which the function $f(x)$ is absolutely continuous.

This paper deals with the following generalization of the above-mentioned theorem of KHINTCHINE, concerning functions of two variables:

THEOREM. *In order that a finite measurable function $F(x, y)$ defined on a measurable set $M \subset R_2$ may be approximately differentiable at almost all the points of M, it is necessary and sufficient that the set M shall be representable as union of a set of measure zero and of a finite or enumerable sequence of measurable (closed) sets on which the function $F(x, y)$ fulfils the Lipschitz condition.*

The proof of our Theorem is based on four lemmas.

LEMMA 1. *If the finite measurable function of two variables $F(x, y)$ defined on a set $M \subset R_2$ is approximately differentiable at almost all the points of this set, then the set M is representable as union of a set of measure zero and of a finite or enumerable sequence of measurable (closed) sets on which the function $F(x, y)$ fulfils the Lipschitz condition.*

PROOF. In accordance with the theorem of LUSIN we may suppose that the set M is closed and $F(x, y)$ is continuous on M. We denote by A_n the set of the points (x, y) of M which fulfil the following condition for every two dimensional interval Q:[3]

$$(1) \qquad L_2\left(E_{(\xi,\eta)}\left[\left|\frac{F(\xi,\eta)-F(x,y)}{\sqrt{(x-\xi)^2+(y-\eta)^2}}\right| \leqq n;\ (\xi,\eta) \in M \cap Q\right]\right) \geqq \frac{3}{4} L_2(Q)$$

whenever[4] $\delta(Q) \leqq \frac{2}{n}$ and $(x, y) \in Q$.

[1] See Bibliography.
[2] R_n denotes the n-dimensional Euclidean space.
[3] L_n denotes the n-dimensional Lebesgue measure.
[4] $\delta(Q)$ denotes the diameter of Q.

As the function $F(x, y)$ is approximately differentiable at almost all the points of M and from the density theorem it follows

$$L_2\left(\bigcup_{n=1}^{\infty} A_n\right) = L_2(M).$$

Let $A_n^{(i,k)}$ be the intersection of the set A_n and of the closed square $Q_{i,k} = \left[\frac{i}{n}, \frac{i+1}{n}; \frac{k}{n}, \frac{k+1}{n}\right]$ $(i, k = 0, \pm 1, \pm 2, \ldots)$.

Let $(x_1, y_1) \in A_n^{(i,k)}$ and $(x_2, y_2) \in A_n^{(i,k)}$, then putting $Q = [x_1, x_2; y_1, y_2]$ the sets

$$\underset{(\xi,\eta)}{E}\left[\left|\frac{F(\xi,\eta)-F(x_1,y_1)}{\sqrt{(\xi-x_1)^2+(\eta-y_1)^2}}\right| \leqq n; (\xi,\eta) \in M \cap Q\right]$$

and

$$\underset{(\xi,\eta)}{E}\left[\left|\frac{F(\xi,\eta)-F(x_2,y_2)}{\sqrt{(\xi-x_2)^2+(\eta-y_2)^2}}\right| \leqq n; (\xi,\eta) \in M \cap Q\right]$$

have according to (1) a point (x_3, y_3) in common whenever $x_1 \neq x_2$ and $y_1 \neq y_2$. Hence

$$|F(x_1,y_1)-F(x_2,y_2)| \leqq |F(x_1,y_1)-F(x_3,y_3)| + |F(x_3,y_3)-F(x_2,y_2)| \leqq$$
$$\leqq n\left(\sqrt{(x_1-x_3)^2+(y_1-y_3)^2} + \sqrt{(x_3-x_2)^2+(y_3-y_2)^2}\right) \leqq 2n\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}.$$

If $x_1 = x_2$ or $y_1 = y_2$, then let

$$Q = \left[x_1, x_1 + \frac{1}{2}(y_2-y_1); y_1, y_2\right]$$

or

$$Q = \left[x_1, x_2; y_1, y_1 + \frac{1}{2}(x_2-x_1)\right],$$

respectively.

The set M is closed and the function $F(x, y)$ is continuous on M, consequently $F(x, y)$ fulfils the Lipschitz condition on the closure $\overline{A_n^{(i,k)}} \subset M$ of $A_n^{(i,k)}$. We see at once from the construction that

$$L_2\left(\bigcup_{n=1}^{\infty}\bigcup_{i=-\infty}^{\infty}\bigcup_{k=-\infty}^{\infty}\overline{A_n^{(i,k)}}\right) = L_2(M)$$

and this completes the proof.

LEMMA 2. *Let $F(x, y)$ be a finite measurable function defined on a measurable set $M \subset R_2$. If the set M is representable as union of a set of measure zero and of a finite or enumerable sequence of measurable sets on which the function $F(x, y)$ fulfils the Lipschitz condition, then there exist the finite approximate partial derivatives $F_{ap_x}(x, y)$ and $F_{ap_y}(x, y)$ at almost all the points of this set.*

PROOF. It is sufficient to prove that if the function $F(x, y)$ fulfils the Lipschitz condition on a measurable plane set L, then the approximate partial derivatives $F_{ap_x}(x, y)$ and $F_{ap_y}(x, y)$ exist at almost all the points of L. Let us now denote for any number η by $L^{[\eta]}$ the linear set of the points x such that $(x, \eta) \in L$. The linear set $L^{[\eta]}$ is measurable for almost every η. Hence for almost all η the function $F(x, \eta)$, as function of x, is approximately derivable with respect to x at almost all the points of $L^{[\eta]}$. Since further, the set of the points of L at which the function $F(x, y)$ is approximately partially derivable with respect to one variable is measurable, it follows at once that the function $F(x, y)$ is approximately partially derivable with respect to x at almost all the points of L. Similarly, we establish the corresponding result concerning approximate partial derivability of $F(x, y)$ with respect to y.

LEMMA 3. *If the finite measurable function of two variables $F(x, y)$, defined on a measurable set $M \subset R_2$, is approximately partially derivable at almost all the points of this set, then M is representable as union of a set of measure zero and of a finite or enumerable sequence of measurable (closed) sets on which the function $F(x, y)$ fulfils the Lipschitz condition.*

PROOF. We may clearly assume that M is bounded. We may furthermore assume, in accordance with the theorem of LUSIN, that M is closed, further $F(x, y)$, $F_{ap_x}(x, y)$, $F_{ap_y}(x, y)$ exist and are continuous on M.

Suppose that

$$|F_{ap_x}(x, y)| + |F_{ap_y}(x, y)| + 1 < k \quad \text{for} \quad (x, y) \in M.$$

We denote by B_n the set of the points (x, y) of M for which

$$(2) \qquad L_1\left(E_{\xi}\left[\left|\frac{F(\xi, y) - F(x, y)}{\xi - x}\right| \leqq K; (\xi, y) \in M; \alpha \leqq \xi \leqq \beta\right]\right) \geqq \frac{3}{4}(\beta - \alpha)$$

whenever $\beta - \alpha \leqq \frac{1}{n}$ and $x \in [\alpha, \beta]$.

Each set B_n is closed. To see this, put $(x_k, y_k) \in B_n$ and $\lim_k (x_k, y_k) = (x_0, y_0)$. Then, the set M being closed, $(x_0, y_0) \in M$. Let $\alpha_0 \leqq x_0 \leqq \beta_0$ and put $\alpha_k = \alpha_0 + x_k - x_0$, $\beta_k = \beta_0 + x_k - x_0$ so that we have $\alpha_k \leqq x_k \leqq \beta_k$, $\beta_k - \alpha_k = \beta_0 - \alpha_0 \leqq \frac{1}{n}$ and according to (2) $L_1(E_k) \geqq \frac{3}{4}(\beta_k - \alpha_k)$ $(k = 1, 2, \ldots)$ for

$$E_k = E_{\xi}\left[\left|\frac{F(\xi, y_k) - F(x_k, y_k)}{\xi - x_k}\right| \leqq K; (\xi, y_k) \in M; \alpha_k \leqq \xi \leqq \beta_k\right]$$
$$(k = 0, 1, 2, \ldots).$$

As the sets E_k are closed and, in consequence of the continuity of $F(x, y)$ on M, $E_0 \supset \limsup_k E_k$, we have $L_1(E_0) \geqq \limsup_k L_1(E_k) \geqq \frac{3}{4}(\beta_0 - \alpha_0)$ so that finally $(x_0, y_0) \in B_n$.

As the approximate partial derivatives exist everywhere (at the points of M) and almost all points of M are points of density in the direction of the x axis, it follows

$$L_2\left(M-\bigcup_{n=1}^{\infty} B_n\right)=0.$$

Let us divide the plane with the straight lines $x=\frac{m}{n}$ into stripes parallel to the y axis, we denote by $B_{n,m}$ $(m=0, \pm 1, \pm 2, \ldots)$ the intersection of the set B_n and of the stripe $\frac{m}{n} \leqq x \leqq \frac{m+1}{n}$. On the set $B_{n,m}$ the function $F(x, y)$ fulfils the Lipschitz condition in the direction of the x axis.

Namely, let $(x_1, y) \in B_{n,m}$ and $(x_2, y) \in B_{n,m}$; the sets

$$E_{\xi}\left[\left|\frac{F(\xi, y)-F(x_1, y)}{\xi-x_1}\right| \leqq K;\ (\xi, y) \in M;\ \xi \in [x_1, x_2]\right]$$

and

$$E_{\xi}\left[\left|\frac{F(\xi, y)-F(x_2, y)}{\xi-x_2}\right| \leqq K;\ (\xi, y) \in M;\ \xi \in [x_1, x_2]\right]$$

have then by (2) a point x_3 in common. For the points x_3 we have

$$|F(x_1, y)-F(x_2, y)| \leqq |F(x_1, y)-F(x_3, y)|+|F(x_3, y)-F(x_2, y)| \leqq$$
$$\leqq K(|x_1-x_3|+|x_2-x_3|)=K|x_1-x_2|.$$

Let us resolve the analogue construction in the direction of the other axis, i. e. let $B_{n,m}^{(i)}$ be the set of the points (x, y) of $B_{n,m}$ for which

$$(3)\quad L_1\left(E_{\eta}\left[\left|\frac{F(x, \eta)-F(x, y)}{\eta-y}\right| \leqq K;\ (x, \eta) \in B_{n,m};\ \alpha \leqq \eta \leqq \beta\right]\right) \geqq \frac{3}{4}(\beta-\alpha)$$

whenever $\beta-\alpha \leqq \frac{1}{i}$ and $y \in [\alpha, \beta]$. Let us divide the plane with straight lines $y=\frac{k}{i}$ $(k=0, \pm 1, \pm 2, \ldots)$ into stripes parallel to the x axis and denote by $B_{n,m}^{(i,k)}$ the intersection of the set $B_{n,m}^{(i)}$ and of the stripe $\frac{k}{i} \leqq y \leqq \frac{k+1}{i}$.

Let $(x_1, y_1) \in B_{n,m}^{(i,k)}$ and $(x_2, y_2) \in B_{n,m}^{(i,k)}$. Then according to (3) the closed sets

$$E_{\eta}\left[\left|\frac{F(x_1, \eta)-F(x_1, y_1)}{\eta-y}\right| \leqq K;\ (x_1, \eta) \in B_{n,m};\ \eta \in [y_1, y_2]\right]$$

and

$$E_{\eta}\left[\left|\frac{F(x_2, \eta)-F(x_2, y_2)}{\eta-y_2}\right| \leqq K;\ (x_2, \eta) \in B_{n,m};\ \eta \in [y_1, y_2]\right]$$

have a point y_3 incommon.

Accordingly

$$|F(x_1, y_1)-F(x_2, y_2)| \leqq |F(x_1, y_1)-F(x_1, y_3)|+|F(x_1, y_3)-F(x_2, y_3)|+$$
$$+|F(x_2, y_3)-F(x_2, y_2)| \leqq K(|y_1-y_3|+|x_1-x_2|+|y_3-y_2|) \leqq$$
$$\leqq 2K\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}.$$

The set M is closed and the function $F(x, y)$ is continuous on M, consequently $F(x, y)$ fulfils the Lipschitz condition on the closure $\overline{B_{n,m}^{(i,k)}} \subset M$ of $B_{n,m}^{i,k}$. We see at once from the construction that

$$L_2\left(\bigcup_{n=1}^{\infty}\bigcup_{m=-\infty}^{+\infty}\bigcup_{i=1}^{\infty}\bigcup_{k=-\infty}^{+\infty}\overline{B_{n,m}^{(i,k)}}\right)=L_2(M)$$

and this completes the proof.

LEMMA 4. *Under the hypotheses of Lemma 2 it follows that $F(x, y)$ is approximately differentiable at almost all the points of M.*

PROOF. This follows at once from the extension theorem of functions fulfilling the Lipschitz condition [2] and from the theorem of RADEMACHER [3].

Now we are able to prove the Theorem stated at the beginning of the paper. We shall prove it in two ways:

1° It follows from Lemmas 1 and 2 and from the theorem of STEPANOFF [4].

2° It follows from Lemmas 1 and 4 (without using the theorem of STEPANOFF).

COROLLARY (Theorem of STEPANOFF). *In order that a finite function $F(x, y)$ which is measurable on a measurable plane set M may be approximately differentiable at almost all the points of M, it is necessary and sufficient that $F(x, y)$ shall be almost everywhere in M approximately partially derivable with respect to each variable.*

When this is the case, the approximate partial derivatives $F_{ap_x}(x, y)$ and $F_{ap_y}(x, y)$ are at almost all the points (x, y) of M the coefficients of the approximate differential of $F(x, y)$.

PROOF. *Necessity*. From Lemmas 1 and 2.

Sufficiency. From Lemmas 3 and 4.

(Received 8 February 1955)

Bibliography

[1] A. KHINTCHINE, Sur la dérivation asymptotique, *Comptes Rendus Acad. Sci. Paris*, **164** (1917), pp. 142—144.

[2] J. CZIPSZER and L. GEHÉR, Extension of functions satisfying a Lipschitz condition, *Acta Math. Acad. Sci. Hung.*, **6** (1955), pp. 213—220.

[3] H. RADEMACHER, Über partielle und totale Differenzierbarkeit. I, *Math. Ann.*, **79** (1919), pp. 340—359.

[4] W. STEPANOFF, Sur les conditions de l'existence de la différentielle totale, *Rec. Math. Soc. Math. Moscou*, **32** (1925), pp. 511—526.

ОБ АППРОКСИМАТИВНОЙ ДИФФЕРЕНЦИРУЕМОСТИ ФУНКЦИЙ ДВУХ ПЕРЕМЕННЫХ

Л. Гехер (Мишкольц)

(Резюме)

В настоящей работе дается обобщение на функции многих переменных теоремы Хинчина [1]. Именно справедлива следующая

Теорема. Конечная измеримая функция двух переменных $F(x, y)$, определенная на измеримом множестве M, является аппроксимативно дифференцируемой почти всюду на M, тогда и только тогда, если M разложимо на сумму множества меры нуль и конечного или счетного числа измеримых множеств, на каждом из которых $F(x, y)$ удовлетворяет условию Липшица. (Измеримые множества можно считать и замкнутыми.)

Из лемм настоящей работы вытекает как следствие новое доказательство следующей теоремы Степанова:

Конечная измеримая функция $F(x, y)$, определенная на измеримом множестве M, аппроксимативно дифференцируема почти всюду на M тогда и только тогда, если почти всюду на M существуют аппроксимативные частные производные $F_{ap_x}(x, y)$ и $F_{ap_y}(x, y)$.

Тогда почти всюду на M $F_{ap_x}(x, y)$ и $F_{ap_y}(x, y)$ являются коэффициентами аппроксимативного дифференциала $F(x, y)$.

A SPECIAL SYMMETRIC EQUATION IN A FINITE FIELD

By
L. CARLITZ (Durham, USA)
(*Presented by* L. RÉDEI)

§ 1. Introduction

MORDELL [2] has discussed the congruence

$$a_1\sigma_1 + \cdots + a_4\sigma_4 \equiv a \pmod{p}, \tag{1.1}$$

where σ_r is the r-th elementary symmetric function of $x_1, \ldots, x_4$; he proved that the number of solutions of (1.1) is equal to

$$p^3 + O(p^{\frac{3}{2}}) \qquad (p \to \infty), \tag{1.2}$$

except in certain excluded cases. In the present note we consider the special case

$$\sigma_3 \equiv a \pmod{p} \tag{1.3}$$

and determine explicitly the number of solutions of (1.3). Since it is no more difficult we take $a, x_r \in GF(q)$, where $q = p^n$ and p is an arbitrary prime; in particular the case $p = 2$ is not excluded. For the final results, see Theorems 1 and 2 below.

It may be remarked that the equation

$$a_1x_2x_3x_4 + \cdots + a_4x_1x_2x_3 = a \qquad (a_1a_2a_3a_4 \neq 0) \tag{1.4}$$

is really no more general than (1.3) as is clear on making the substitution $x_r = a_r y_r$ $(r = 1, \ldots, 4)$ in (1.4).

§ 2. The case q odd

Let $M = M(a)$ denote the number of solutions of $\sigma_3 = a$. Also put

$$A = N\{x_1x_2 + x_2x_3 + x_3x_1 \neq 0\}, \tag{2.1}$$

$$B(a) = N\{x_1x_2 + x_2x_3 + x_3x_1 = 0,\ x_1x_2x_3 = a\}, \tag{2.2}$$

where the symbol $N\{\cdots\}$ denotes the number of solutions of the indicated system. It is clear that

$$M(a) = A + qB(a). \tag{2.3}$$

We shall first consider the case $a \neq 0$. Then clearly $B(a)$ is equal to the number of solutions of the system

$$x_1^2x_2^2 + a(x_1 + x_2) = 0 \qquad (x_1x_2 \neq 0). \tag{2.4}$$

If we define $\Psi(a) = +1, -1, 0$ according as a is a square, a non-square or 0 of $GF(q)$, then it is seen that

$$B(a) = \sum_{x \neq 0} \{1 + \Psi(a^2 - 4ax^3)\}. \tag{2.5}$$

If we put

$$J(a) = \sum_{x} \Psi(a - 4x^3) \qquad (J(0) = 0), \tag{2.6}$$

then (2.5) becomes

$$B(a) = q - 2 + \Psi(a) J(a). \tag{2.7}$$

In the next place returning to (2.1), it is easy to show that

$$A = q^3 - q^2. \tag{2.8}$$

Hence, by (2.3) and (2.7), we get

$$M(a) = q^3 - 2q + q\Psi(a) J(a). \tag{2.9}$$

We next evaluate $J(a)$. In the first place it is clear from (2.6) that

$$J(ab^3) = \Psi(b) J(a) \qquad (b \neq 0), \tag{2.10}$$

which implies

$$J^2(ab^3) = J^2(a) \qquad (b \neq 0). \tag{2.11}$$

It is also evident that

$$J(a) = 0 \qquad (3 \nmid q - 1), \tag{2.12}$$

so that we may now assume that $q \equiv 1 \pmod 3$.

Now consider the sum

$$S = \sum_{a \neq 0} J^2(a) = \sum_{x, y} \sum_{a} \Psi((a - 4x^3)(a - 4y^3)).$$

In view of the known formula

$$\sum_{a} \Psi(a^2 - d) = -1 \qquad (d \neq 0), \tag{2.13}$$

we get

$$S = \sum_{x^3 \neq y^3} (-1) + \sum_{x^3 = y^3} (q - 1) = -(q^2 - 3q + 2) + (q - 1)(1 + 3(q - 1)),$$

which reduces to

$$S = 2q(q - 1). \tag{2.14}$$

In the next place if w denotes a number of $GF(q)$ such that $w^2 + w + 1 = 0$, then it is clear from (2.11) that

$$S = \frac{1}{3}(q - 1)\{J^2(1) + J^2(w) + J^2(w^2)\}.$$

Hence (2.14) yields

$$J^2(1) + J^2(w) + J^2(w^2) = 6q. \tag{2.15}$$

By (2.6) we have also (since $\Psi(w) = 1$)

$$J(1) + J(w) + J(w^2) = 0; \tag{2.15}'$$

thus (2.25) becomes

$$J^2(1)+J(1)J(w)+J^2(w)=3q. \tag{2.16}$$

We also note that (2.16) implies

$$J(1)\equiv J(w)\equiv 1 \pmod 3; \tag{2.17}$$

moreover

$$J(1)=\sum_x (1-x^3)^{\frac{1}{2}(q-1)}-\binom{\frac{1}{2}(q-1)}{\frac{1}{3}(q-1)}\sum_x x^{q-1}$$

and consequently when $p\equiv 1 \pmod 3$

$$J(1)\equiv -\binom{\frac{1}{2}(p-1)}{\frac{1}{3}(p-1)}^n \pmod p. \tag{2.18}$$

The last step is a consequence of the familiar congruence

$$\binom{a_0+a_1p+\cdots+a_rp^r}{b_0+b_1p+\cdots+b_rp^r}\equiv\binom{a_0}{b_0}\binom{a_1}{b_1}\cdots\binom{a_r}{b_r} \pmod p$$

for $0\leqq b_i\leqq a_i\leqq p-1$.

When $p\equiv 2 \pmod 3$, J can be explicitly evaluated. Indeed we have

$$J(w^2)=J(w^p)=\sum_x \Psi(w^p-4^px^{3p})=\sum_x \Psi(w-4x^3)=J(w),$$

so that (2.15)′ and (2.16) imply $J^2(w)=q$, $J(w)=\pm q^{\frac{1}{2}}$. Hence, using (2.17), we get

$$J(1)=-2J(w),\quad J(w)=(-1)^{\frac{1}{2}n}q^{\frac{1}{2}} \qquad (p\equiv 2 \pmod 3). \tag{2.19}$$

We have now evaluated $J(a)$ in all cases and hence determined $M(a)$ for $a\neq 0$. (For a detailed discussion of Jacobsthal sums see WHITEMAN [3].)

To determine $M(0)$ we note that

$$M(0)=q^4-\sum_{a\neq 0}M(a)=q^4-(q-1)(q^3-2q)-q\sum_a \Psi(a)J(a).$$

Now by (2.13)

$$\sum_a \Psi(a)J(a)=\sum_x\sum_a \Psi(a^2-4ax^3)=0,$$

and therefore

$$M(0)=q^3+2q^2-2q. \tag{2.20}$$

It is not difficult to give a direct proof of (2.20).

We may now state

THEOREM 1 (q odd). *The number of solutions of* $\sigma_3 = a$ *is determined by* (2.20) *when* $a = 0$ *and by* (2.9), (2.12), (2.18) *and* (2.19) *when* $a \neq 0$. *In particular*

$$M(a) = q^3 - 2q \qquad (a \neq 0, q \equiv 1 \pmod 3),$$

$$M(a) = q^3 - 2q + (-1)^{\frac{1}{2}n} t q^{\frac{3}{2}} \qquad (a \neq 0, q \equiv 1, p \equiv 2 \pmod 3),$$

where $t = -2$ *or* 1 *according as* a *is or is not a cube of* $GF(q)$.

§ 3. The case q even

We again use the notation (2.1), (2.2). Then both (2.3) and (2.4) hold. Now define $l(a) = +1$ or -1 according as the polynomial $u^2 + u + a$ is reducible or irreducible in $GF[u, q]$. It follows that

$$l(a+b) = l(a)l(b). \tag{3.1}$$

We now put

$$L(a) = \sum_x l(ax^3), \tag{3.2}$$

so that in particular

$$L(a) = L(a^2) = L\left(\frac{1}{a}\right) \qquad (a \neq 0) \tag{3.3}$$

and

$$L(ab^3) = L(a) \qquad (b \neq 0). \tag{3.4}$$

It is convenient to define the function

$$k(a) = \begin{cases} -1 & (a \neq 0), \\ q-1 & (a = 0). \end{cases} \tag{3.5}$$

It follows from the definition of $l(a)$ that

$$\sum_{x \neq 0} l(ax) = k(a). \tag{3.6}$$

In the next place, using (3.1) and (3.6), we get

$$\sum_{a \neq 0} L^2(a) = \sum_{x,y} \sum_{a \neq 0} l(a(x^3 + y^3)) = \sum_{x,y} k(x^3 + y^3) =$$
$$= (q-1) + 3(q-1)^2 - 2(q-1) - \{(q-1)^2 - 3(q-1)\} = 2q(q-1). \tag{3.7}$$

Now clearly

$$L(a) = 0 \qquad (3 \nmid q-1), \tag{3.8}$$

so that we may assume $q \equiv 1 \pmod 3$. If then $w^2 + w + 1 = 0$, $w \in GF(q)$, we get, using (3.4) and (3.7),

$$L^2(1) + L^2(w) + L^2(w^2) = 6q. \tag{3.9}$$

But we have also

$$L(1) + L(w) + L(w^2) = 0 \tag{3.10}$$

and by (3.3), $L(w)=L(w^2)$ so that (3.9) reduces to

$$L^2(w)=q, \quad L(w)=\pm q^{\frac{1}{2}}. \tag{3.11}$$

To remove the ambiguity we note that (3.2) implies $L(1)\equiv L(0)\equiv 1 \pmod 3$, so that $L(w)\equiv 1 \pmod 3$; consequently (3.11) yields

$$L(w)=(-1)^{\frac{1}{2}n}q^{\frac{1}{2}}. \tag{3.12}$$

Returning to (2.1) we apply some results of [1], § 9 to get A; we find that

$$A=q^3-q^2. \tag{3.13}$$

As for $B(a)$ examination of (2.4) yields

$$B(a)=\sum_{x\neq 0}\left\{1+l\left(\frac{x^3}{a}\right)\right\}=q-2+L(a) \qquad (a\neq 0). \tag{3.14}$$

Hence by (2.3)

$$M(a)=q^3-2q+qL(a) \qquad (a\neq 0). \tag{3.15}$$

As for $a=0$, we have

$$M(0)=q^4-\sum_{a\neq 0}(q^3-2q+qL(a))=q^4-(q-1)(q^3-2q)-q\sum_{a\neq 0}L(a);$$

using (3.4) and (3.10) this reduces to

$$M(0)=q^3+2q^2-2q. \tag{3.16}$$

We now state

THEOREM 2 ($q=2^n$). *The number of solutions of $\sigma_3=a$ is determined by* (3.16) *when $a=0$. When $a\neq 0$ we have*

$$M(a)=q^3-2q \qquad (n \text{ odd}), \tag{3.17}$$

$$M(a)=q^3-2q+(-1)^{\frac{1}{2}n}tq^{\frac{1}{2}} \qquad (n \text{ even}), \tag{3.18}$$

where $t=-2$ or 1 according as a is or is not a cube of $GF(q)$.

DUKE UNIVERSITY,
DURHAM, USA

(*Received 4 March 1955*)

References

[1] L. CARLITZ, Pairs of quadratic equations in a finite field, *American Journal of Mathematics*, **76** (1954), pp. 137—154.

[2] L. J. MORDELL, Note on the linear symmetric congruence in n variables, *Canadian Journal of Mathematics*, **5** (1953), pp. 433—438.

[3] A. L. WHITEMAN, Cyclotomy and Jacobsthal sums, *American Journal of Mathematics*, **74** (1952), pp. 89—99.

ОДНО ОСОБЕННОЕ СИММЕТРИЧЕСКОЕ УРАВНЕНИЕ В КОНЕЧНОМ ПОЛЕ

Л. Карлиц (Даргем, США)

(Резюме)

Пусть σ_r $(r = 1, \ldots, 4)$ означает симметрической многочлен степени r от $x_1, \ldots, x_4$. Морделл [2] доказал, что число решений сравнения

$$a_1\sigma_1 + \cdots + a_4\sigma_4 \equiv a \pmod{p},$$

несмотря на некоторые исключения, равно

$$p^3 + O\left(p^{\frac{3}{2}}\right) \quad \text{для} \quad p \to \infty.$$

В настоящей работе определяется элементарным путем точное число $M(a)$ решений уравнения $\sigma_3 = a$ в любом конечном поле $GF(q)$; (отмечается, что более общее уравнение

$$a_1x_2x_3x_4 + \cdots + a_4x_1x_2x_3 = a \qquad (a_1a_2a_3a_4 \neq 0)$$

с помощью подстановкой $x_r = a_r y_r$ $(r = 1, \ldots, 4)$ легко сводится к уравнению $\sigma_3 = a$).

Имеет место соотношение

$$M(0) = q^3 + 2q^2 - 2q.$$

В случае $a \neq 0$ значение $M(a)$ зависит от четности или нечетности числа q, и во втором случае также от того, в каком классе вычетов mod 3 находятся число q и характеристика от $GF(q)$, но всегда имеем

$$M(a) = q^3 + O\left(q^{\frac{3}{2}}\right) \quad \text{для} \quad a \neq 0.$$

SUR LES FONCTIONS LOCALEMENT MONOTONES AU SENS GÉNÉRALISÉ

Par
ÁKOS CSÁSZÁR (Budapest)
(*Présenté par* G. ALEXITS)

1. Dans un récent ouvrage [1], j'ai introduit une notion de fonction localement croissante ou décroissante au sens généralisé, cette généralisation consistant à négliger les ensembles qui appartiennent à une famille $\mathfrak{N}$ d'ensembles donnée et soumise à quelques conditions générales. Dans la note présente, je vais examiner les propriétés des fonctions qui sont localement monotones au sens généralisé en question, et cela en tous les points d'un ensemble donné. Les résultats principaux qui vont être acquis nous montrent qu'une fonction de ce type appartient au plus à la deuxième classe de Baire, qu'elle coïncide, excepté les points d'un ensemble de mesure aussi petite que l'on veut, avec une fonction monotone par segments, enfin qu'elle est continue, excepté les points d'un ensemble dénombrable, lorsqu'on néglige les ensembles appartenant à la famille $\mathfrak{N}$.[1]

2. Pour donner aux propositions précédentes une formulation précise, convenons des notations suivantes.

Soit $\mathfrak{N}$ une famille héréditaire et σ-additive d'ensembles linéaires, c'est-à-dire une famille de sous-ensembles de la droite R satisfaisant aux conditions suivantes:

(2. 1) Si $A \in \mathfrak{N}$ et si $B \subset A$, on a $B \in \mathfrak{N}$.

(2. 2) Si $A_i \in \mathfrak{N}$ $(i = 1, 2, \ldots)$, on a $\sum_{i=1}^{\infty} A_i \in \mathfrak{N}$.

Comme famille $\mathfrak{N}$, on peut choisir par exemple celle des ensembles dénombrables, de mesure nulle, de première catégorie, ou celle des sous-ensembles d'un ensemble donné (vide ou non), etc. Toutes les notions qui vont être introduites dépendent du choix de la famille $\mathfrak{N}$, mais, pour simplifier les notations, nous n'indiquerons pas cette dépendance, ce qui ne causera aucune confusion, la famille $\mathfrak{N}$ étant fixée une fois pour toutes.

Nous désignerons par $\overline{\mathfrak{N}}$ la famille composée des ensembles de la forme $N + D$, où $N \in \mathfrak{N}$ et D est dénombrale. Cette famille satisfait évidemment à deux conditions analogues à (2. 1) et (2. 2).

[1] Qu'il me soit permis d'exprimer ma reconnaissance à M. G. ALEXITS qui m'a posé une partie des questions résolues dans cette note.

Soit $f(x)$ une fonction réelle et finie, définie pour $x \in R$. $\delta > 0$ étant donné, nous dirons que $f(x)$ jouit de la propriété $C_+(\delta)$ au point x_0, si

$$\underset{x}{E}\left[\frac{f(x)-f(x_0)}{x-x_0} \leqq 0, \quad 0 < x-x_0 < \delta\right] \in \mathfrak{N}, \tag{2. 3}$$

et qu'elle jouit de la propiété $C_+^*(\delta)$ au même point, si

$$\underset{x}{E}\left[\frac{f(x)-f(x_0)}{x-x_0} < 0, \quad 0 < x-x_0 < \delta\right] \in \mathfrak{N}.$$

En remplaçant en (2. 3) le signe $\leqq$ par $\geqq$ et par $>$, on arrive à la définition de la propriété $D_+(\delta)$ et $D_+^*(\delta)$ respectivement. Nous définissons les propriétés $C_-(\delta)$, $C_-^*(\delta)$, $D_-(\delta)$ et $D_-^*(\delta)$ en substituant dans les définitions précédentes $-\delta < x-x_0 < 0$ à la formule $0 < x-x_0 < \delta$.

Nous dirons que $f(x)$ possède la propriété C_+ au point x_0, si elle y possède la propriété $C_+(\delta)$ pour au moins un nombre $\delta > 0$ (et alors pour tout nombre $\delta > 0$ suffisamment petit, d'après (2. 1)). On définit de façon analogue les propriétés C_+^*, D_+, D_+^*, C_-, C_-^*, D_- et D_-^*.

En supprimant le signe $+$ ou $-$ d'une des notations précédentes, on veut dire que $f(x)$ possède toutes les deux propriétés, celle marquée par le signe $+$ et celle marquée par le signe $-$. Par exemple $f(x)$ jouit au point x_0 de la propriété C, si elle y possède les propriétés C_+ et C_- en même temps; cela veut dire que $f(x)$ est localement croissante au sens strict au point x_0, en négligeant les ensembles $\mathfrak{N}$. De façon analogue, les symboles C^*, D et D^* désignent que $f(x)$ est respectivement localement croissante, décroissante au sens strict et décroissante, lorsqu'on néglige les ensembles $\mathfrak{N}$.

Nous dirons que $f(x)$ possède au point x_0 la propriété $M(\delta)$, si elle y possède au moins l'une des propriétés $C(\delta)$ et $D(\delta)$. On définit pareillement les propriétés $M^*(\delta)$, M et M^*. Ces dernières expriment que $f(x)$ est localement monotone, au sens strict ou au sens large respectivement, au point considéré, en négligeant les ensembles $\mathfrak{N}$.

Nous dirons enfin que $f(x)$ jouit de la propriété $\bar{E}$ ou $\underline{E}$ au point x_0, si elle y possède les propriétés C_- et D_+, ou celles D_- et C_+, selon le cas, c'est-à-dire si elle a en ce point un maximum ou un minimum au sens strict, en négligeant les ensembles $\mathfrak{N}$, et qu'elle possède la propriété E, si elle possède au moins une de celles $\bar{E}$ et $\underline{E}$.

Comme d'habitude, nous dirons que la fonction $f(x)$ est croissante [décroissante] sur l'ensemble H, si l'on a pour $\alpha, \beta \in H$, $\alpha < \beta$, l'inégalité $f(\alpha) \leqq f(\beta)$ [$f(\alpha) \geqq f(\beta)$]. On dira qu'elle est strictement croissante [strictement décroissante] sur H, si les mêmes conditions impliquent $f(\alpha) < f(\beta)$ [$f(\alpha) > f(\beta)$]. On dit que $f(x)$ est [strictement] monotone sur H, si elle est ou bien [strictement] croissante, ou bien [strictement] décroissante sur cet ensemble.

On dit que la fonction $f(x)$ est [strictement] monotone par segments sur l'ensemble H, s'il existe une suite $\{x_i\}$ $(i = 0, \pm 1, \pm 2, \ldots)$ de nombres réels tels que $x_i < x_{i+1}$, $x_i \to +\infty$ pour $i \to +\infty$, $x_i \to -\infty$ pour $i \to -\infty$ et que $f(x)$ est [strictement] monotone sur chacun des ensembles $[x_i, x_{i+1}] \cdot H$.

Désignons par U l'ensemble composé par les points $x_0 \in R$ tels qu'on a $(x_0, x_0 + \delta) \notin \mathfrak{N}$ et $(x_0 - \delta, x_0) \notin \mathfrak{N}$ pour tout $\delta > 0$.

Dans ce qui suit, pour plus de brièveté, nous résumerons parfois les énoncés de plusieurs propositions de structure et de démonstration analogues en un seul énoncé; les changements nécessaires seront indiqués entre crochets et, au cas où le nombre des propositions résumées dépasse deux, ils seront séparés les uns des autres par des points-virgules; quant à la démonstration, elle ne sera détaillée en général que pour la proposition mentionnée en premier lieu.

3. Nous sommes capables à présent de formuler les résultats principaux de ce paragraphe.

THÉORÈME I. *Si la fonction $f(x)$ ne possède la propriété E en aucun point d'un ensemble S, mais elle y jouit de la propriété M^* [M], on a $S = \sum_{n=1}^{\infty} S_n$, où $S_n \subset S_{n+1}$ et $f(x)$ est monotone [strictement monotone] par segments sur chacun des ensembles S_n. La première hypothèse peut être remplacée par celle $S \subset U$.*

THÉORÈME II. *Les conditions du théorème précédent étant admises, $f(x)$ appartient sur S à la deuxième classe de Baire au plus.*

THÉORÈME III. *Si les conditions du théorème I sont remplies et si, en outre, S est mesurable, à tout nombre $\varepsilon > 0$ correspond un ensemble mesurable $S_\varepsilon \subset S$ tel que $|S - S_\varepsilon| < \varepsilon$ et que $f(x)$ est monotone [strictement monotone] par segments sur S_ε.*

La démonstration de ces théorèmes se compose de plusieurs propositions auxiliaires.

(3. 1) *On a $R - U \in \overline{\mathfrak{N}}$.*

DÉMONSTRATION. En désignant par H_1 l'ensemble des points $x_0 \in R - U$ tels qu'il existe un $\delta > 0$ pour lequel on a $(x_0 - \delta, x_0) \in \mathfrak{N}$ et $(x_0, x_0 + \delta) \in \mathfrak{N}$, par H_2 l'ensemble des points $x_0 \in R - U$ tels qu'il existe un $\delta > 0$ pour lequel on a $(x_0, x_0 + \delta) \in \mathfrak{N}$, mais qu'on a pour tout $\varepsilon > 0$ $(x_0 - \varepsilon, x_0) \notin \mathfrak{N}$, enfin par H_3 celui des $x_0 \in R - U$ tels qu'il existe un $\delta > 0$ pour lequel on a $(x_0 - \delta, x_0) \in \mathfrak{N}$, mais qu'on a pour tout $\varepsilon > 0$ $(x_0, x_0 + \varepsilon) \notin \mathfrak{N}$, on a évidemment $R - U = H_1 + H_2 + H_3$. Or, pour $x_0 \in H_1$, il existe un $\delta > 0$ tel qu'on a $(x_0 - \delta, x_0 + \delta) \in \overline{\mathfrak{N}}$, et ceci entraîne, en vertu du théorème de LINDELÖF, que $H_1 \in \overline{\mathfrak{N}}$. Pour $x_0 \in H_2$, il existe un $\delta > 0$ tel que $(x_0, x_0 + \delta) \in \mathfrak{N}$, donc aucun point de l'intervalle $(x_0, x_0 + \delta)$ n'appartient à H_2, ce qui entraîne que l'en-

semble H_2 est dénombrable, et la même conclusion s'obtient pour H_3; tout cela démontre la proposition.

(3. 2) *Si $f(x)$ ne possède la propriété E en aucun point d'un ensemble S, l'ensemble $S-U$ est dénombrable.*

DÉMONSTRATION. Avec les notations de la démonstration précédente, c'est évident que $f(x)$ possède les propriétés $\overline{E}$ et $\underline{E}$ en tout point de l'ensemble H_1. Les ensembles H_2 et H_3 étant dénombrables, la proposition est la conséquence immédiate de l'égalité $S-U=S(R-U)=S(H_1+H_2+H_3)$.

La proposition suivante est une conséquence évidente des définitions correspondantes :

(3. 3) *x_0 étant un point de l'ensemble U, aucune fonction $f(x)$ ne peut avoir au point x_0 en même temps les propriétés C_+ et D_+^* [C_+^* et D_+; C_- et D_-^*; C_-^* et D_-].*

La proposition (3. 4) qui suit coïncide avec le théorème (3. 23) de mon ouvrage [1] cité plus haut.

(3. 4) *Si $f(x)$ jouit de la propriété E en tout point d'un ensemble S, on a $S\in\overline{\mathfrak{N}}$.*

(3. 5) *Si $f(x)$ possède au point x_0 la propriété $C_+(\delta)$ [$C_+^*(\delta)$; $D_+(\delta)$; $D_+^*(\delta)$] et si l'on a pour un point x_1 $0<x_1-x_0<\delta$ et $f(x_1)-f(x_0)\leqq 0$ [$<$; $\geqq$; $>$], $f(x)$ possède au point x_1 la propriété $\underline{E}$ [$\underline{E}$; $\overline{E}$; $\overline{E}$]. Si elle possède en x_0 la propriété $C_-(\delta)$ [$C_-^*(\delta)$; $D_-(\delta)$; $D_-^*(\delta)$] et si l'on a pour un point x_1 $-\delta<x_1-x_0<0$ et $f(x_1)-f(x_0)\geqq 0$ [$>$; $\leqq$; $<$], elle possède en x_1 la propriété $\overline{E}$ [$\overline{E}$; $\underline{E}$; $\underline{E}$].*

DÉMONSTRATION. Les hypothèses $f(x_1)\leqq f(x_0)$ et $0<x_1-x_0<\delta$ entraînent

$$\underset{x}{E}\left[\frac{f(x)-f(x_1)}{x-x_1}\geqq 0,\ x_0<x<x_1\right]\subset \underset{x}{E}[f(x)\leqq f(x_0),\ x_0<x<x_1]\subset$$

$$\subset \underset{x}{E}\left[\frac{f(x)-f(x_0)}{x-x_0}\leqq 0,\ x_0<x<x_0+\delta\right]\in\mathfrak{N},$$

puisque $f(x)$ jouit de la propriété $C_+(\delta)$ au point x_0. On a par un raisonnement analogue

$$\underset{x}{E}\left[\frac{f(x)-f(x_1)}{x-x_1}\leqq 0,\ x_1<x<x_0+\delta\right]\subset \underset{x}{E}[f(x)\leqq f(x_0),\ x_1<x<x_0+\delta]\subset$$

$$\subset \underset{x}{E}\left[\frac{f(x)-f(x_0)}{x-x_0}\leqq 0,\ x_0<x<x_0+\delta\right]\in\mathfrak{N},$$

et ces deux relations donnent ensemble la proposition à démontrer.

Les propositions (3. 6) découlent immédiatement de (3. 5), et celles (3. 7) de (3. 6).

(3. 6) *Si $f(x)$ ne possède la propriété E ni en α, ni en β, et si l'on a $\alpha<\beta<\alpha+\delta$, l'hypothèse que $f(x)$ possède en α la propriété $C_+(\delta)$ [$C_+^*(\delta)$;*

$D_+(\delta)$; $D^*_+(\delta)$], ou qu'elle possède en β la propriété $C_-(\delta)$ [$C^*_-(\delta)$; $D_-(\delta)$; $D^*_-(\delta)$], entraîne l'inégalité $f(\alpha) < f(\beta)$ [$\leqq$; $>$; $\geqq$].

(3. 7) *Si $f(x)$ ne possède la propriété E en aucun point d'un ensemble S de diamètre moindre que δ et si elle possède celle $C_+(\delta)$ [$C_-(\delta)$; $C^*_+(\delta)$; $C^*_-(\delta)$; $D_+(\delta)$; $D_-(\delta)$; $D^*_+(\delta)$; $D^*_-(\delta)$] en tout point de S, elle est strictement croissante [strictement croissante; croissante; croissante; strictement décroissante; strictement décroissante; décroissante; décroissante] sur S.*

(3. 8) *Si $f(x)$ ne possède la propriété E ni en α, ni en β, si elle jouit de celle $C^*(\delta)$ en α et de celle $D^*(\delta)$ en β et si $|\alpha-\beta| < \delta$, on a $f(\alpha) = f(\beta)$. La première hypothèse peut être remplacée par celle qu'on ait $\alpha, \beta \in U$.*

DÉMONSTRATION. D'après (3. 6), les hypothèses entraînent tant l'inégalité $f(\alpha) \leqq f(\beta)$, que celle $f(\alpha) \geqq f(\beta)$, donc la proposition à démontrer. Si $\alpha, \beta \in U$, il découle, en vertu de (3. 3), de ce que $f(x)$ possède en α et en β ou bien la propriété C^*, ou bien celle D^*, qu'elle n'y possède pas la propriété E.

(3. 9) *Si $f(x)$ ne possède la propriété E ni en α, ni en β, et si elle jouit de celle $C(\delta)$ en α et de celle $D(\delta)$ en β, on a $|\alpha-\beta| \geqq \delta$. La première hypothèse peut être remplacée par celle qu'on ait $\alpha, \beta \in U$.*

DÉMONSTRATION. Si l'on avait $|\alpha-\beta| < \delta$, ceci impliquerait d'après (3. 6) les inégalités $f(\alpha) < f(\beta)$ et $f(\alpha) > f(\beta)$, ce qui est impossible. Quant à la dernière assertion de l'énoncé, voir la démonstration précédente.

Les deux théorèmes suivants présentent quelque intérêt de ce point de vue que, dans leurs hypothèses, ne figurent que des conditions unilatérales concernant la fonction $f(x)$.

(3. 10) *Si $f(x)$ ne possède la propriété E en aucun point de l'ensemble S et si elle jouit en tout point de S d'au moins l'une des propriétés C^*_+, C^*_-, D^*_+ ou D^*_- [C_+, C_-, D_+ ou D_-], on a une décomposition $S = \sum_{k=1}^{\infty} S_k$, où $f(x)$ est monotone [strictement monotone] sur chacun des ensembles S_k.*[2]

DÉMONSTRATION. Désignons par $T_n^{(1)}$, $T_n^{(2)}$, $T_n^{(3)}$ et $T_n^{(4)}$ respectivement les ensembles des points $x_0 \in S$ tels que $f(x)$ possède en x_0 la propriété $C^*_+\left(\frac{1}{n}\right)$, $C^*_-\left(\frac{1}{n}\right)$, $D^*_+\left(\frac{1}{n}\right)$ ou $D^*_-\left(\frac{1}{n}\right)$. En employant la notation

$$T_{nm}^{(i)} = T_n^{(i)} \cdot \left[\frac{m}{n+1}, \frac{m+1}{n+1}\right]$$

pour $m = 0, \pm 1, \pm 2, \ldots, n = 1, 2, \ldots$ et $i = 1, \ldots, 4$, on a évidemment

$$S = \sum_{i,\, m,\, n} T_{nm}^{(i)},$$

et d'après (3. 7), $f(x)$ est monotone sur chacun des ensembles $T_{nm}^{(i)}$.

[2] Elle est donc (VBG) sur l'ensemble S, cf. [2], p. 221.

(3. 11) *Si $f(x)$ possède en tout point d'un ensemble S au moins l'une des propriétés C_+^*, C_-^*, D_+^* ou D_-^* [C_+, C_-, D_+ ou D_-], on a une décomposition $S = N + \sum_{k=1}^{\infty} S_k$, où $N \in \overline{\mathfrak{N}}$ et $f(x)$ est monotone [strictement monotone] sur chacun des ensembles S_k.*

DÉMONSTRATION. On n'a qu'à appliquer (3. 4) et (3. 10).

La proposition suivante coïncide avec théorème I.

(3. 12) *Si $f(x)$ ne possède la propriété E en aucun point d'un ensemble S, et si elle jouit en tout point de S de celle M^* [M], on a $S = \sum_{n=1}^{\infty} S_n$, où $S_n \subset S_{n+1}$ et $f(x)$ est monotone [strictement monotone] par segments sur chacun des ensembles S_n. La première hypothèse peut être remplacée par celle $S \subset U$.*

DÉMONSTRATION. Soit S_n l'ensemble des points $x_0 \in S$ tels que $f(x)$ possède en x_0 au moins une des propriétés $C^*\left(\frac{1}{n}\right)$ ou $D^*\left(\frac{1}{n}\right)$ $\left[C\left(\frac{1}{n}\right)\right.$ ou $\left.D\left(\frac{1}{n}\right)\right]$ $(n = 1, 2, \ldots)$. On a évidemment $S_n \subset S_{n+1}$. En considérant pour un entier k quelconque l'ensemble $S_{nk} = S_n \cdot \left[\frac{k}{n+1}, \frac{k+1}{n+1}\right]$, ou bien il existe deux points $\alpha \in S_{nk}$ et $\beta \in S_{nk}$ tels que $f(x)$ possède la propriété $C^*\left(\frac{1}{n}\right)\left[C\left(\frac{1}{n}\right)\right]$ en α et celle $D^*\left(\frac{1}{n}\right)\left[D\left(\frac{1}{n}\right)\right]$ en β, et en ce cas $f(x)$ est constante, donc monotone sur S_{nk}, en vertu de (3. 8) [cette possibilité est exclue d'après (3. 9)], ou bien $f(x)$ possède en tout point de S_{nk} la même propriété $C^*\left(\frac{1}{n}\right)$ ou $D^*\left(\frac{1}{n}\right)\left[C\left(\frac{1}{n}\right)\right.$ ou $\left.D\left(\frac{1}{n}\right)\right]$, et alors elle est également monotone [strictement monotone] sur S_{nk}, d'après (3. 7). En résumant nos conclusions, on arrive à ce que $f(x)$ est monotone [strictement monotone] par segments sur l'ensemble S_n.

La proposition suivante est identique au théorème II.

(3. 13) *Les hypothèses de* (3. 12) *impliquent que $f(x)$ appartient sur S à la deuxième classe de Baire au plus.*

DÉMONSTRATION. En appliquant les notations de la démonstration précédente, définissons la fonction $f_n(x)$ de façon qu'elle soit égale à $f(x)$ sur l'ensemble S_n, qu'elle soit monotone sur chacun des plus petits intervalles I_{nk} (ouverts, fermés ou à moitié fermés selon le cas) qui contiennent les ensembles S_{nk}, et qu'elle soit constante sur chacun des ensembles $[k, k+1] - (S_{n,k-1} + S_{n,k+1} + I_{nk})$. Puisque $f(x)$ est monotone sur chacun des ensembles

S_{nk}, on peut satisfaire aux conditions que nous venons de poser. On voit immédiatement que $f_n(x) \to f(x)$ pour $x \in S$. D'ailleurs chacune des fonctions $f_n(x)$ n'a qu'une infinité dénombrable de discontinuités au plus, elle est donc limite d'une suite de fonctions continues (pour le voir, rangeons en une suite $\{x_1, x_2, \ldots\}$ les nombres rationnels et les points de discontinuité de $f_n(x)$; en désignant par $f_{nk}(x)$ la fonction qui est égale à $f_n(x)$ aux points $x_1, \ldots, x_k$, à $f_n(\alpha)$ pour $x \leqq \alpha$ et à $f_n(\beta)$ pour $x \geqq \beta$, où $\alpha = \min(x_1, \ldots, x_k)$ et $\beta = \max(x_1, \ldots, x_k)$, et qui est linéaire entre deux points voisins quelconques des points $x_1, \ldots, x_k$, on a $\lim_{k\to\infty} f_{nk}(x) = f_n(x)$ pour un x quelconque). La proposition en découle immédiatement.

Nous pouvons enfin démontrer la proposition suivante, identique au théorème III.

(3. 14) *Si $f(x)$ ne possède la propriété E en aucun point d'un ensemble S, si elle jouit de celle M*[M] en tout point de S, et si S est mesurable, il correspond, à tout $\varepsilon > 0$, un ensemble mesurable $S_\varepsilon \subset S$ tel que $|S - S_\varepsilon| < \varepsilon$ et que $f(x)$ est monotone [strictement monotone] par segments sur S. La première hypothèse peut être remplacée par celle $S \subset U$.*

DÉMONSTRATION. Soit $S_k = S \cdot [k, k+1]$; $f(x)$ est mesurable sur S_k, d'après (3. 13). Un $\varepsilon > 0$ étant donné, on peut trouver donc en vertu du théorème de LUSIN un ensemble fermé $F_k \subset S_k$ sur lequel $f(x)$ est continue et tel que $|S_k - F_k| < \varepsilon \cdot 2^{-|k|-3}$, on peut encore supposer que $F_k \subset (k, k+1)$. Désignons par F_{kn} l'ensemble des points $x_0 \in F_k$ tels que $f(x)$ possède en x_0 au moins une des propriétés $C^*\left(\frac{1}{n}\right)$ ou $D^*\left(\frac{1}{n}\right)$. En appliquant le raisonnement de la démonstration de (3. 12), on constate que $f(x)$ est monotone sur l'ensemble F_{kn} et qu'elle conserve cette propriété sur la fermeture $\bar{F}_{kn} \subset F_k$ de F_{kn}, puisqu'elle est continue sur F_k. Puisqu'on a $F_{kn} \subset F_{k,n+1}$ et $F_k = \sum_{n=1}^{\infty} F_{kn}$, on a aussi $\bar{F}_{kn} \subset \bar{F}_{k.n+1}$ et $F_k = \sum_{n=1}^{\infty} \bar{F}_{kn}$, donc, $\bar{F}_{kn}$ étant mesurable, on a pour un indice n_k suffisamment grand

$$|F_k - \bar{F}_{kn_k}| < \varepsilon \cdot 2^{-|k|-3}.$$

Posons $S_\varepsilon = \sum_{k=-\infty}^{\infty} \bar{F}_{kn_k}$; S_ε est mesurable et, en vertu de la relation

$$S - S_\varepsilon \subset \sum_{k=-\infty}^{\infty} [(S_k - F_k) + (F_k - \bar{F}_{kn_k})],$$

on a

$$|S - S_\varepsilon| \leqq \sum_{k=-\infty}^{\infty} (\varepsilon \cdot 2^{-|k|-3} + \varepsilon \cdot 2^{-|k|-3}) = \varepsilon \cdot \frac{3}{4} < \varepsilon.$$

D'après la relation $\bar{F}_{in_i} \subset F_i \subset (i, i+1)$, on a $S \cdot [k, k+1] = \bar{F}_{kn_k}$ et $f(x)$ étant

monotone sur $\overline{F}_{kn_k}$, elle est monotone par segments sur S_ε, ce qu'il fallait démontrer.

[Pour démontrer la variante entre crochets de la proposition, on définit les ensembles S_k et F_k comme avant, puis on désigne par F_{kn} l'ensemble des points $x_0 \in F_k$ tels que $f(x)$ possède en x_0 au moins une des propriétés $C\left(\frac{1}{n}\right)$ ou $D\left(\frac{1}{n}\right)$. On trouve que $f(x)$ est strictement monotone sur F_{kn} et qu'elle conserve cette propriété sur l'ensemble $F_{kn}^* \subset \overline{F}_{kn} \subset F_k$, composé par les points d'accumulation bilatéraux de F_{kn}. L'ensemble $\overline{F}_{kn} - F_{kn}^*$ est évidemment dénombrable, donc F_{kn}^* est mesurable et $|F_{kn}^*| = |\overline{F}_{kn}|$. En choisissant l'indice n_k de la même manière que plus haut, on a

$$|F_k - F_{kn_k}^*| = |F_k - \overline{F}_{kn_k}| < \varepsilon \cdot 2^{-|k|-3},$$

et en posant $S_\varepsilon = \sum_{k=-\infty}^{\infty} F_{kn_k}^*$, S_ε sera mesurable, on aura $|S - S_\varepsilon| < \varepsilon$ et $f(x)$ sera strictement monotone par segments sur S_ε.]

Le théorème suivant montre qu'une fonction ne peut posséder une même propriété C, ou C^*, ou D, ou D^*, en tout point d'un intervalle que si elle y est monotone.

(3. 15) *Si $f(x)$ possède la propriété C [C^*; D; D^*] en tout point de l'intervalle $(a, b) = I \subset U$, elle est strictement croissante [croissante; strictement décroissante; décroissante] sur I. La même conclusion s'applique aux intervalles $[a, b]$, ou $[a, b)$, ou $(a, b]$, si, en dehors des conditions précédentes, $f(x)$ possède la propriété C_+^* [C_+^*; D_+^*; D_+^*] en a, ou celle C_-^* [C_-^*; D_-^*; D_-^*] en b, ou toutes les deux, selon le cas.*

DÉMONSTRATION. Les hypothèses impliquent, d'après (3. 3), que $f(x)$ ne possède la propriété E en aucun point de I, donc, en vertu de (3. 6) qu'elle est localement strictement croissante (au sens classique) en tout point de I, et, au cas où les conditions supplémentaires sont remplies, c'est valable même pour le point a, ou b, ou pour tous les deux. D'où, par un raisonnement classique, on parvient à la proposition à démontrer.

4. Dans ce dernier paragraphe, nous nous occuperons des questions de continuité généralisée, c'est-à-dire de continuité sous négligence des ensembles $\mathfrak{N}$, des fonctions qui sont localement monotones lorsqu'on néglige les ensembles en question. Avant de les formuler, convenons d'abord des définitions suivantes.

$f(x)$ étant, comme toujours, une fonction réelle et finie, et $\mathfrak{N}$ une famille d'ensembles qui satisfait aux conditions (2. 1) et (2. 2), désignons par

$$\overline{\lim}_{\substack{x \to x_0 \\ \mathfrak{N}}} f(x)$$

la borne inférieure (finie ou infinie) des valeurs y telles qu'il existe un inter-

valle ouvert I pour lequel on a $x_0 \in I$ et

$$\underset{x}{E}[f(x) > y] \in \mathfrak{N}.$$

De façon analogue, soit

$$\underline{\lim}_{x \to x_0 \mathfrak{N}} f(x)$$

la borne supérieure des valeurs y telles qu'il existe un intervalle ouvert I pour lequel on a $x_0 \in I$ et

$$\underset{x}{E}[f(x) < y] \in \mathfrak{N}.$$

On voit immédiatement que $x_0 \in U$ implique

$$\underline{\lim}_{x \to x_0 \mathfrak{N}} f(x) \leqq \overline{\lim}_{x \to x_0 \mathfrak{N}} f(x); \tag{4.1}$$

si le signe d'égalité est valable dans cette relation, la valeur commune des deux membres sera désignée par

$$\lim_{x \to x_0 \mathfrak{N}} f(x).$$

La fonction $f(x)$ est dite continue au point x_0 en négligeant les ensembles $\mathfrak{N}$, si

$$f(x_0) = \lim_{x \to x_0 \mathfrak{N}} f(x).$$

On constate aisément que cette définition de la continuité sous négligence des ensembles $\mathfrak{N}$ équivaut à celle due à R. BAIRE (cf. [3], p. 202).

Nous démontrerons le théorème suivant qui fournit une généralisation d'un théorème classique sur les discontinuités des fonctions monotones:

THÉORÈME IV. *Si $f(x)$ possède la propriété M^* en tout point d'un ensemble $S \subset U$, elle est continue en négligeant les ensembles $\mathfrak{N}$ en tout point de S, excepté les points d'un ensemble dénombrable.*

DÉMONSTRATION. En désignant par r un nombre rationnel quelconque, considérons l'ensemble S_r composé par les points $x_0 \in S$ tels que

$$\underline{\lim}_{x \to x_0 \mathfrak{N}} f(x) < r < \overline{\lim}_{x \to x_0 \mathfrak{N}} f(x),$$

puis l'ensemble S_r' des points $x_0 \in S$ tels que

$$\overline{\lim}_{x \to x_0 \mathfrak{N}} f(x) < r < f(x_0),$$

enfin celui S_r'' des points $x_0 \in S$ tels que

$$f(x_0) < r < \underline{\lim}_{x \to x_0 \mathfrak{N}} f(x).$$

C'est évident que l'ensemble $\sum_r (S_r + S_r' + S_r'')$ contient tous les points de S où $f(x)$ n'est pas continue en négligeant les ensembles $\mathfrak{N}$. Il suffit donc de démontrer que chacun des ensembles S_r, S_r' et S_r'' est dénombrable.

Soit $x_0 \in S_r$ et supposons que $f(x)$ possède en x_0, disons, la propriété C^*. Si l'on a $f(x_0) \geqq r$, on aura pour un nombre $\delta > 0$ convenable

$$\underset{x}{E}[f(x) < r, x_0 < x < x_0 + \delta] \subset \underset{x}{E}[f(x) < f(x_0), x_0 < x < x_0 + \delta] \in \mathfrak{N},$$

l'inégalité $x_0 < x_1 < x_0 + \delta$ entraîne donc $\overline{\lim}_{x \to x_1}{}_{\mathfrak{R}} f(x) \geqq r$, c'est-à-dire que x_1 ne peut pas appartenir à S_r. D'autre part, si $f(x_0) < r$, on aura pour un $\delta > 0$ convenable

$$\underset{x}{E}[f(x) > r, \quad x_0 - \delta < x < x_0] \subset \underset{x}{E}[f(x) > f(x_0), \quad x_0 - \delta < x < x_0] \in \mathfrak{N},$$

l'inégalité $x_0 - \delta < x < x_0$ implique donc $\overline{\lim}_{x \to x_1}{}_{\mathfrak{R}} f(x) \leqq r$, et par conséquent $x_1 \notin S_r$. En tout cas, x_0 n'est pas point d'accumulation bilatéral de S_r, et on arrive à la même conclusion en supposant que $f(x)$ jouit de la propriété D^* en x_0. La dénombrabilité de S_r en découle immédiatement.

Soit $x_0 \in S_r'$ et supposons que $f(x)$ possède la propriété C^* en x_0. On a pour un $\delta > 0$ convenable

$$\underset{x}{E}[f(x) < r, \quad x_0 < x < x_0 + \delta] \subset \underset{x}{E}[f(x) < f(x_0), \quad x_0 < x < x_0 + \delta] \in \mathfrak{N},$$

donc, pour $x_0 < x_1 < x_0 + \delta$, on a $\overline{\lim}_{x \to x_1}{}_{\mathfrak{R}} f(x) \geqq r$, ce qui entraîne la relation $x_1 \notin S_r'$ en vertu de (4. 1), valable puisque $x_0 \in S - U$. Par conséquent, x_0 n'est pas point d'accumulation bilatéral de S_r'; la même conclusion s'ensuit de l'hypothèse que $f(x)$ possède la propriété D^* en x_0. S_r' est donc, lui aussi, dénombrable. Un raisonnement analogue s'applique pour S_r'', ce qui termine la démonstration.

(Reçu le 18 mars 1955.)

Ouvrages cités

[1] Á. CSÁSZÁR, Sur les ensembles de niveau des fonctions à une variable, à paraître en *Colloquium Mathematicum*.

[2] S. SAKS, *Theory of the integral* (Warszawa—Lwów, 1937).

[3] H. HAHN, *Reelle Funktionen* (Leipzig, 1932).

О ФУНКЦИЯХ, ЛОКАЛЬНО МОНОТОННЫХ В ОБЩЕМ СМЫСЛЕ

А. Часар (Будапешт)

(Резюме)

Пусть $\mathfrak{N}$ наследственная, σ-аддитивная система линейных точечных множеств. Говорим, что функция $f(x)$ локально $\mathfrak{N}$-монотонна в точке x_0, если существует такое положительное число δ, что или

$$\underset{x}{E}\left[\frac{f(x)-f(x_0)}{x-x_0}<0,\quad 0<|x-x_0|<\delta\right]\in\mathfrak{N},$$

или же

$$\underset{x}{E}\left[\frac{f(x)-f(x_0)}{x-x_0}>0,\quad 0<|x-x_0|<\delta\right]\in\mathfrak{N}.$$

Пусть U множество точек x_0 числовой прямой, для которых при всяком $\delta>0$ имеет место $(x_0, x_0+\delta)\notin\mathfrak{N}$ и $(x_0, x_0-\delta)\notin\mathfrak{N}$.

Главные результаты работы состоят в следующем:

Если функция $f(x)$ локально $\mathfrak{N}$-монотонна во всех точках множества $S\subset U$, то $S=\sum_{n=1}^{\infty}S_n$, где $S_n\subset S_{n+1}$ и $f(x)$ кусочно монотонна на всяком из множеств S_n. При тех же условиях $f(x)$ на множестве S не может принадлежать классу Бэра выше второго. Если при указанных условиях множество S является и измеримым, то для любого положительного ε можно найти такое измеримое множество $S_\varepsilon\subset S$, что $|S-S_\varepsilon|<\varepsilon$ и функция $f(x)$ кусочно монотонна на S_ε. Наконец, при прежних условиях $f(x)$ непрерывна во всех точках S, хотя бы за исключением счетного множества точек, если пренебрегать множествами $\mathfrak{N}$.

A NOTE ON SUBDIRECT PRODUCTS

By
ISIDORE FLEISCHER (Paris)
(*Presented by* G. HAJÓS)

In a previous paper [3] L. FUCHS has proved that for groups, rings, or Boolean algebras, every subdirect product of two systems can be realized as the set of all couples whose components have equal images with respect to a pair of homomorphisms defined on the systems; and he has suggested that this property is related to the fact that homomorphisms of such structures are completely determined by the set of elements congruent to a zero element. In this note the validity of the theorem for more general algebraic systems is discussed; it is shown in particular that the result depends on a completely different property of these structures: namely, the permutability of their congruence relations.

Without being overly specific, let us agree to call *algebra* a set A in which various algebraic laws operate some of which may even involve operators from some set outside A. It is known that a quotient algebra is completely determined by an equivalence relation ϑ in A which has the substitution property for all the algebraic operations concerned; this property enables one to transfer the operations from A to the set of equivalence classes A/ϑ.

We shall need the following concepts from the theory of equivalence relations: One writes $\vartheta \subset \vartheta'$ and one says that ϑ is contained in (or implies) ϑ' if every class of ϑ is contained in a class of ϑ'. Thus ϑ' induces a unique equivalence relation in A/ϑ and every equivalence relation in A/ϑ is induced in this way (First Isomorphism Theorem for equivalence relations). The classes of the conjunction of ϑ and ϑ', $\vartheta \cap \vartheta'$, are by definition the non-void intersections of a class of ϑ and a class of ϑ'. The product $\vartheta\vartheta'$ is defined as the binary relation: $x\vartheta\vartheta' x'$ if there exists a y with $x\vartheta y$ and $y\vartheta' x'$. In case $\vartheta\vartheta'$ equals $\vartheta'\vartheta$, ϑ and ϑ' are said to be permutable: this is the necessary and sufficient condition that $\vartheta\vartheta'$ be an equivalence relation (and hence the smallest equivalence relation containing ϑ and ϑ').

The following Lemma is the essential step in the proof of a Second Isomorphism Theorem for equivalence relations.

LEMMA. *The class of ϑ containing x intersects the class of ϑ' containing x' if and only if $x\vartheta\vartheta' x'$.*

We turn now to subdirect products.

Corresponding to a subdirect product representation of A there exist two projections onto the factors which in turn give rise to congruence relations ϑ and ϑ'; the fact that distinct elements of A have at least one unequal component implies that $\vartheta \cap \vartheta' = \varepsilon$, the identity relation. Conversely, two such congruence relations yield an embedding of A as a subdirect product of its quotient algebras A/ϑ and A/ϑ'.[1]

When will A, represented in this way, be the collection of all couples having the same image by a pair of homomorphisms sending A/ϑ and A/ϑ' into a fourth algebra? Observe that every function on A/ϑ induces in A an equivalence relation ϑ^* containing ϑ; the same relation must be induced by the function on A/ϑ', because the components of a couple coming from an element of A are to be sent on the same element. Translated into an assertion on equivalence classes in A, the requirement that *every* couple whose components have the same image come from an element of A reads: A class of ϑ and a class of ϑ' contained in the same class of ϑ^* always intersect. According to our Lemma, this can only happen if $\vartheta^* = \vartheta\vartheta'$. Conversely, if ϑ and ϑ' are permutable, so that $\vartheta^* = \vartheta\vartheta'$ is an equivalence relation, then A is represented in the desired manner with respect to the mappings of A/ϑ and A/ϑ' onto A/ϑ^*.

THEOREM. *A necessary and sufficient condition that a representation of A as a subdirect product of A/ϑ and A/ϑ' consist of all pairs whose components have the same image by a pair of homomorphisms, is that ϑ and ϑ' be permutable.*

COROLLARY. *If for a class of algebras every pair of congruence relations is permutable, then every subdirect product is of the type described in the Theorem.*

All the algebraic structures discussed in [3] are examples of groups with operators; for these the congruence relations are given by normal subgroups and hence are all permutable.

On the other hand, let G be a non-commutative group and take for the algebras sets A admitting G as operators on the left. Of course G is itself such an algebra, and by following through the usual proof of the homomorphism theorem one easily verifies that the congruence relations on G are just its decompositions into the left cosets of a subgroup.[2] Now if H and H' are subgroups of G which do not commute, the congruence relations they define will not be permutable; therefore the representation of $G/H \cap H'$ as a subdirect product of G/H and G/H' will not be of the desired type.

[1] For the most general possible algebras it may happen that the operations of A/ϑ and A/ϑ' do not induce exactly the operations of A; this in no way affects the discussion.

[2] This is therefore a type of structure in which such a relation is determined by the class of elements equivalent to a neutral element.

In conclusion, we might remark that the remaining theorems 3—7 of [3] are easily proved in the general framework of our discussion by a consistent use of the correspondence between homomorphisms and congruence relations.

(Received 4 May 1955)

References

[1] G. Birkhoff, *Lattice theory* (New York, 1948).
[2] P. Dubreil, *Algèbre*, Tome 1 (Paris, 1954).
[3] L. Fuchs, On subdirect unions. I, *Acta Math. Acad. Sci. Hung.*, **3** (1952), pp. 103—120.

ЗАМЕЧАНИЕ О ПОДПРЯМЫХ ПРОИЗВЕДЕНИЯХ

И. Флейшер (Париж)

(Резюме)

Л. Фукс доказал, что в случае групп, колец или булевых алгебр всякое подпрямое произведение двух систем представляется как совокупность всех таких пар элементов, что образы двух компонентов пары при некоторых гомоморфизмах совпадают. В настоящей работе автором доказывается, что в случае любой алгебраической системы A и любых соотношений конгруентности ϑ и ϑ' представление A в виде подпрямого произведения A/ϑ и A/ϑ' осуществляется всеми парами элементов со совпадающими для некоторых гомоморфизмов образами тогда и только тогда, если ϑ и ϑ' перестановочны.

ÜBER DIFFERENZIERTE FOLGEN DER LAGRANGESCHEN INTERPOLATION

Von

G. FREUD (Budapest)

(*Vorgelegt von* G. ALEXITS)

1. In vorliegender Arbeit wird folgendes Problem behandelt: Es sei $f(x)$ eine m-mal stetig differenzierbare Funktion; wir betrachten die Lagrangeschen Interpolationspolynome dieser Funktion bezüglich des Grundpunktsystems $\{x_{kn}\}$:

$$L_n(f;x)=\sum_{k=1}^{n} f(x_{kn})\,l_{kn}(x).$$

Bilden wir die m-ten Derivierten $L_n^{(m)}(f;x)$ der Interpolationspolynome und suchen hinreichende Bedingungen für das Bestehen von

$$\lim_{n=\infty} L_n^{(m)}(f;x)=f^{(m)}(x).$$

Es werden die Fälle behandelt, wo die $\{x_{kn}\}$ Nullstellen einer orthogonalen Polynomenfolge sind. Wir werden zwei Sätze beweisen. In Satz I wird bezüglich des Orthogonalpolynomsystems nur die recht allgemeine Bedingung (2) vorausgesetzt, dann muß aber die Funktionenklasse, zur welchen $f(x)$ gehört, stärker eingeschränkt werden. Für $m=0$ geht Satz I in einen entsprechenden Satz der Theorie des Lagrangeschen Interpolationsverfahrens über, welcher schon eingehend behandelt wurde (vgl. J. SHOHAT [8]; G. GRÜNWALD und P. TURÁN [6]; P. ERDŐS und P. TURÁN [2], S. 527; G. ALEXITS [1]).

Im Satz II werden bezüglich des Orthogonalpolynomsystems schärfere Voraussetzungen getroffen (vgl. (12) und (13)), aber die Funktionenklasse, auf welche unser Satz anwendbar ist, wird wesentlich ausgedehnt. Setzt man in Satz II $m=0$, dann erhält man ein früheres Resultat des Verfassers [5]. Die Voraussetzungen (12) und (13) sind erfüllt, wenn $w(x)=(1-x)^{\alpha}(1+x)^{\beta}$ mit $\alpha, \beta > -1$ ist, in welchem Fall die $p_n(x)$ mit den Jakobischen Polynomen identisch sind.

Die Sätze I und II werden aus der Abschätzung der Summe

$$\sum_{k=1}^{n} |\,l_{kn}^{(m)}(x)\,|$$

gefolgert, welche die Rolle der Lebesgueschen Funktion unseres Grenzprozesses spielt. Der Beweis ergibt sich aus einer Kombination zweier Verfah-

ren, die in zwei vorigen Arbeiten des Verfassers ([4], [5]) auseinandergesetzt wurden.

Am Ende der Arbeit wird gezeigt, daß die Ungleichung für kein Grundpunktsystem durch eine der Größenordnung nach schärfere ersetzt werden kann.

2. In der Arbeit [5] des Verfassers wurde folgendes Problem behandelt: $\tau_n(x)$ durchlaufe alle Polynome höchstens n-ten Grades mit

$$\int_{-1}^{+1} [\tau_{n-1}(t)]^2 w(t)\,dt \leqq 1 \qquad (0 \leqq w(t) \in L) \tag{1}$$

und man suche das Maximum $M_{n,m}(x)$ von $\tau_{n-1}^{(m)}(x)$. Es wurde gezeigt, daß

$$M_{n,m}(x) = \left\{\sum_{k=0}^{n-1} [p_\nu^{(m)}(x)]^2\right\}^{\frac{1}{2}}$$

ist, wobei $p_\nu(x)$ jenes normierte Orthogonalpolynom ν-ten Grades bedeutet, welches zur Gewichtsfunktion $w(x)$ und zum Orthogonalitätsintervall $(-1, +1)$ gehört. Es wurde ferner gezeigt, daß unter der Voraussetzung

$$w(x) \geqq \mu > 0 \quad \text{für} \quad a \leqq x \leqq b \qquad ((a, b) \subset (-1, +1)), \tag{2}$$

in jedem inneren Teilintervall von (a, b) die Abschätzung

$$M_{n,m}(x) = O\left(n^{m+\frac{1}{2}}\right) \tag{3}$$

gleichmäßig besteht. Wir wollen jetzt $M_{n,k}(x)$ mit den Derivierten der Lagrangeschen Grundpolynome ausdrücken. Es seien $x_{1n}, x_{2n}, \ldots, x_{nn}$ die Nullstellen von $p_n(x)$ und die Lagrangeschen interpolatorischen Grundfunktionen, die zu diesen Grundpunkten gehören, seien $l_{kn}(x)$ $(k = 1, 2, \ldots, n)$. Infolge der Gauß—Jacobischen mechanischen Quadraturformel gilt

$$\int_{-1}^{+1} [\tau_{n-1}(x)]^2 w(x)\,dx = \sum_{k=1}^{n} \lambda_{kn} [\tau_{n-1}(x_{kn})]^2, \tag{4}$$

und aus der Lagrangeschen Interpolationsformel erhalten wir

$$\tau_{n-1}^{(m)}(x) = \sum_{k=1}^{n} \tau_{n-1}(x_{kn}) l_{kn}^{(m)}(x). \tag{5}$$

Hieraus folgt unter Beachtung von (1) und (4)

$$|\tau_{n-1}^{(m)}(x)| \leqq \sum_{k=1}^{n} \lambda_{kn}^{\frac{1}{2}} |\tau_{n-1}(x_{kn})| \lambda_{kn}^{-\frac{1}{2}} |l_{kn}^{(m)}(x)| \leqq$$

$$\leqq \left\{\sum_{k=1}^{n} \lambda_{kn} [\tau_{n-1}(x_{kn})]^2\right\}^{\frac{1}{2}} \left\{\sum_{k=1}^{n} \frac{[l_{kn}^{(m)}(x)]^2}{\lambda_{kn}}\right\}^{\frac{1}{2}} \leqq \left\{\sum_{k=1}^{n} \frac{[l_{kn}^{(m)}(x)]^2}{\lambda_{kn}}\right\}^{\frac{1}{2}}$$

und die Gleichheit wird tatsächlich erreicht, falls

$$\tau_{n-1}(t) = \frac{\sum_{k=1}^{n} \frac{l_{kn}^{(m)}(x)\, l_{kn}(t)}{\lambda_{kn}}}{\left\{\sum_{k=1}^{n} \frac{[l_{kn}^{(m)}(x)]^2}{\lambda_{kn}}\right\}^{\frac{1}{2}}}$$

ist. Mit Hilfe von (4) zeigt man leicht, daß mit dieser Wahl von τ_{n-1} auch (1) befriedigt ist. Somit wurde gezeigt, daß

$$M_{n,m}(x) = \left\{\sum_{k=1}^{n} \frac{[l_{kn}^{(m)}(x)]^2}{\lambda_{kn}}\right\}^{\frac{1}{2}} \tag{6}$$

gilt. Dieser Zusammenhang (sowie die Ungleichung (7), die wir aus ihr folgern) wurde andeutungsweise und ohne Beweis schon in einer Arbeit von P. ERDŐS und P. TURÁN ([2], S. 525) ausgesprochen. Aus (3) und (6) folgt

$$\sum_{k=1}^{n} \frac{[l_{kn}^{(m)}(x)]^2}{\lambda_{kn}} = O(n^{2m+1}) \tag{7}$$

gleichmäßig in jedem inneren Teilintervall von (a, b). Somit gilt wegen $\sum_{k=1}^{n} \lambda_{kn} = \int_{-1}^{+1} w(x)dx$ auch die Abschätzung

$$\sum_{k=1}^{n} |l_{kn}^{(m)}(x)| \leqq \left(\sum_{k=1}^{n} \lambda_{kn}\right)^{\frac{1}{2}} \left\{\sum_{k=1}^{n} \frac{[l_{kn}^{(m)}(x)]^2}{\lambda_{kn}}\right\}^{\frac{1}{2}} = O\left(n^{m+\frac{1}{2}}\right) \tag{8}$$

gleichmäßig in jedem inneren Teilintervall von (a, b).

SATZ I. *Als Grundpunkte der Lagrangeschen Interpolationsfolge $L_n(f; x)$ wählen wir die Nullstellen von Orthogonalpolynomen, welche durch eine die Bedingung* (2) *befriedigende Gewichtsfunktion $w(x)$ bestimmt sind. Die Funktion $f(x)$ sei in $[-1, +1]$ m-mal stetig differenzierbar und es sei $f^{(m)}(x) \in \operatorname{Lip} \alpha$ $\left(\alpha > \frac{1}{2}\right)$; dann gilt gleichmäßig in jedem inneren Teilintervall von (a, b)*

$$\lim_{n=\infty} L_n^{(m)}(f; x) = f^{(m)}(x). \tag{9}$$

Die Forderung $f^{(m)}(x) \in \operatorname{Lip} \alpha$ mit $\alpha > \frac{1}{2}$ kann durch folgende schwächere ersetzt werden: es gilt gleichmäßig in x_1 und x_2 $(-1 \leqq x_1,\ x_2 \leqq +1)$

$$f^{(m)}(x_1) - f^{(m)}(x_2) = o\left(|x_1 - x_2|^{\frac{1}{2}}\right). \tag{10}$$

BEWEIS. Laut (10) gibt es Polynome höchstens $n-1$-ten Grades $\psi_n(x)$ für welche gleichmäßig in $x \in [-1, +1]$

$$f(x) - \psi_n(x) = o\left(n^{-m-\frac{1}{2}}\right) \tag{11}$$

und

$$f^{(m)}(x)-\psi_n^{(m)}(x)=o\left(n^{-\frac{1}{2}}\right)$$

besteht. Aus der Identität

$$\psi_n(x)=\sum_{k=1}^{n}\psi_n(x_{kn})l_{kn}(x)$$

folgt durch m-maliges Differenzieren

$$\psi_n^{(m)}(x)=\sum_{k=1}^{n}\psi_n(x_{kn})l_{kn}^{(m)}(x). \tag{12}$$

Endlich ist wegen (12) und (8)

$$|f^{(m)}(x)-L_n^{(m)}(f;x)|=\left|\sum_{k=1}^{n}[f(x_{kn})-\psi_n(x_{kn})]\,l_{kn}^{(m)}(x)+o\left(n^{-\frac{1}{2}}\right)\right|\leqq$$

$$\leqq \operatorname{Max}|f(x_{kn})-\psi_n(x_{kn})|\sum_{k=1}^{n}|l_{kn}^{(m)}(x)|+o\left(n^{-\frac{1}{2}}\right)=o(1),$$

w. z. b. w.

3. Im weiteren soll die nichtnegative Gewichtsfunktion $w(x)\in L$ und die dazu gehörende Folge der normierten Orthogonalpolynome $\{p_n(x)\}$ den folgenden stärkeren Voraussetzungen genügen:

Sei

$$\mu_1\leqq w(x)\leqq\mu_2>0 \quad \text{für} \quad a\leqq x\leqq b \tag{13}$$

und

$$|p_n(x)|\leqq K \quad \text{für} \quad a\leqq x\leqq b;\quad n=0,1,2,\ldots. \tag{14}$$

SATZ II. *Sind die Bedingungen* (13) *und* (14) *erfüllt, ist ferner* $f(x)$ *in* $[-1,+1]$ *m-mal stetig differenzierbar und*

$$f^{(m)}(x_1)-f^{(m)}(x_2)=o\left(\log^{-1}\frac{1}{|x_1-x_2|}\right) \quad \text{für} \quad a\leqq x_1,x_2\leqq b; \tag{15}$$

so besteht (9) *gleichmäßig in jedem inneren Teilintervall von* (a,b).

BEWEIS. Sei (a_1,b_1) ein festes inneres Teilintervall von (a,b) und $x\in(a_1,b_1)$. Nach den Ergebnissen der Arbeit [4] des Verfassers gilt unter den Bedingungen (13) und (14)

$$\sum_{x_{kn}\in(a_1,\,b_1)}|l_{kn}(x)|=O(\log n) \tag{16}$$

und

$$\sum_{x_{kn}\in(a_1,\,b_1)}|l_{kn}(x)|=O(1), \tag{17}$$

wobei diese Abschätzungen gleichmäßig in x bestehen, falls x im inneren Teilintervall (a_1,b_1) von (a,b) variiert. Es bestehen also dieselben Abschätzungen auch für die Summen

$$\sum_{x_{kn}\in(a_1,\,b_1)}\varepsilon_{kn}l_{kn}(x) \quad \text{und} \quad \sum_{x_{kn}\in(a_1,\,b_1)}\varepsilon_{kn}l_{kn}(x),$$

wo die Zahlen ε_{kn} die Werte $+1, -1$ und 0 in beliebiger Anordnung annähmen. Nach der Ungleichung von BERNSTEIN ist somit

$$\sum_{x_{kn} \in (a_1, b_1)} \varepsilon_{kn} l_{kn}^{(m)}(x) = O(n^m \log n)$$

und

$$\sum_{x_{kn} \in (a_1, b_1)} \varepsilon_{kn} l_{kn}^{(m)}(x) = O(n^m).$$

Diese Abschätzungen gelten gleichmäßig in x für jedes innere Teilintervall von (a_1, b_1) und die numerischen Konstanten sind sogar von der Wahl der Zahlen ε_{kn} unabhängig.

Wählen wir

$$\varepsilon_{kn} = \operatorname{sg} l_{kn}^{(m)}(x),$$

dann erhalten wir

$$\sum_{x_{kn} \in (a_1, b_1)} |l_{kn}^{(m)}(x)| = O(n^m \log n) \tag{18}$$

und

$$\sum_{x_{kn} \in (a_1, b_1)} |l_{kn}^{(m)}(x)| = O(n^m). \tag{19}$$

Wir betrachten jetzt die Funktion $f(x)$. Wegen der Annahme (15) gibt es nach bekannten Sätzen eine Polynomfolge $\{\psi_n(x)\}$, so daß $\psi_n(x)$ von höchstens n-tem Grade ist und die Beziehung

$$f(x) - \psi_n(x) = \begin{cases} o(n^{-m} \log^{-1} n) & (x \in (a_1, b_1)), \\ o(n^{-m}) & (x \notin (a_1, b_1)) \end{cases} \tag{20}$$

besteht. Hieraus ergibt sich infolge der Bernsteinschen Ungleichung

$$f^{(m)}(x) - \psi_n^{(m)}(x) = o(1) \qquad (x \in (a_1, b_1)). \tag{21}$$

Der letzte Teil des Beweises verläuft nun auf altbekanntem Weg:

$$f^{(m)}(x) - L_n^{(m)}(f; x) = f^{(m)}(x) - \psi_{n-1}(x) - L_n^{(m)}(f - \psi_{n-1}; x), \tag{22}$$

wegen (18), (19) und (20) ergibt sich somit

$$\begin{aligned} |L_n^{(m)}(f - \psi_{n-1}; x)| &\leqq o(n^{-m} \log^{-1} n) \sum_{x_{kn} \in (a_1, b_1)} |l_{kn}^{(m)}(x)| + \\ &+ o(n^{-m}) \sum_{x_{kn} \in (a_1, b_1)} |l_{kn}^{(m)}(x)| = o(1). \end{aligned} \tag{23}$$

Aus (21), (22) und (23) folgt unmittelbar die Behauptung unseres Satzes.

Es bleibt noch übrig zu zeigen, daß die Abschätzung (18) für kein Grundpunktsystem verbessert werden kann, nicht einmal dann, wenn man auf die Forderung, daß die $\{x_{kn}\}$ Nullstellen eines Orthogonalpolynomsystems seien, verzichtet.

Wir setzen voraus, es gebe eine Grundpunktmatrix $\{x_{kn}\}$ für welche

$$\Sigma |l_{kn}^{(m)}(x)| = \varepsilon_n n^m \log n, \qquad \varepsilon_n \to o \tag{24}$$

gilt und leiten daraus einen Widerspruch her. Betrachten wir dazu das lineare Funktional

$$A_n(\varphi) = L_n^{(m)}\left\{\int_{-1}^{x}(x-t)^{m-1}\varphi(t)dt\,;\,x\right\}. \tag{25}$$

Es sei $\varphi(t)$ eine stetige Funktion und M das Maximum des Betrages von $\varphi(t)$ im abgeschlossenen Intervall $[-1, +1]$. Die Funktion

$$\Phi(x) = \int_{-1}^{x}(x-t)^{m-1}\varphi(t)\,dt,$$

deren m-te Ableitung gleich $\varphi(x)$ ist, kann infolge des Satzes von D. JACKSON mit einem Polynom höchstens $n-1$-ten Grades $\psi_{n-1}(x)$ approximiert werden, so daß

$$|\varphi(x)-\psi_{n-1}(x)| \leqq C_m M n^{-m}$$

besteht, wobei C_m eine nur von m abhängige positive Zahl bedeutet. Man erhält hieraus mit derselben Schlußweise, wie am Ende des Teiles 2 dieser Arbeit

$$|\varphi(x)-A_n(\varphi)| \leqq C_m^* \varepsilon_n \log n \cdot M, \tag{26}$$

wo auch C_m^* nur von m abhängt. Das lineare Funktional (25) hat die Eigenschaft, daß es jedes Polynom höchstens $n-m-1$-ten Grades in sich überführt. Aus einem Lemma von LOSINSKY, CHARSCHILADZE und NIKOLAJEW (vgl. I. P. NATANSON [7], Anhang 3, Lemma 3 und L. FEJÉR [3]) folgt aber, daß (26) falsch ist, die Voraussetzung (24) kann also nicht bestehen, womit der angekündigte Widerspruch hergestellt ist.

(Eingegangen am 1. Juli 1955.)

Literaturverzeichnis

[1] G. ALEXITS, Eine Bemerkung zur Konvergenzfrage des Lagrangeschen Interpolationsverfahren, *Acta Math. Acad. Sci. Hung.*, **4** (1953), S. 233—236.

[2] P. ERDŐS and P. TURÁN, On interpolation. III, *Annals of Math.*, **41** (1940), S. 510-553 (insbesondere S. 527).

[3] L. FEJÉR, Die Abschätzung eines Polynoms in einem Intervalle wenn Schranken für seine Werte und ersten Ableitungswerte in einzelnen Punkten des Intervalles gegeben sind, und ihre Anwendung auf die Konvergenzfrage Hermitescher Interpolationsreihen, *Mat. Zeitschr.*, **32** (1930), S. 426—457.

[4] G. FREUD, Über die Lebesguesche Funktionen der Lagrangeschen Interpolation, *Acta Math. Acad. Sci. Hung.*, **4** (1953), S. 137—142.

[5] G. FREUD, Über das gliedweise Differenzieren einer orthogonalen Polynomreihe, *Acta Math. Acad. Sci. Hung.*, **6** (1955), S. 221—226.

[6] G. Grünwald und P. Turán, Über Interpolation, *Annali della Scuola Norm. Sup. di Pisa*, **7** (1938), S. 137—146.
[7] И. П. Натансон, Конструктивная теория функций (Москва—Ленинград, 1949).
[8] J. Shohat, On interpolation, *Annals of Math.*, **34** (1933), S. 130—146.

О ДИФФЕРЕНЦИРОВАНИИ ИНТЕРПОЛЯЦИОННЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ЛАГРАНЖА

Г. Фрайд (Будапешт)

(Резюме)

Пусть $\{p_n(x)\}$ система многочленов ортогональных относительно весовой функции $0 \leqq w(x) \in L$ $(-1 \leqq x \leqq 1)$, $x_{1n}, x_{2n}, \ldots, x_{nn}$ корни многочлена n-ой степени $p^n(x)$, а $l_{1n}(x), l_{2n}(x), \ldots, l_{nn}(x)$— основные интерполяционные функции Лагранжа, соответствующие этим основным точкам.

Пусть $f(x)$ функция, m-кратно непрерывно дифференцируема на сегменте $[-1, +1]$, и пусть

$$L_n^{(m)}(f; x) = \sum_{k=1}^{n} f(x_{kn})\, l_{kn}^{(m)}(x).$$

Если для $x \in (a, b)$ имеет место $w(x) \geqq \mu > 0$, то во всяком внутреннем подинтервале интервала (a, b) имеем

$$\sum_{k=1}^{n} |l_{kn}^{(m)}(x)| = O\left(n^{m+\frac{1}{2}}\right)$$

равномерно по x, а следовательно, если $f(x)$ на $[-1, +1]$ равномерно удовлетворяет условию

$$f(x_1) - f(x_2) = o\left(|x_1 - x_2|^{1/2}\right),$$

то во всяком внутреннем подинтервале интервала (a, b) имеет место $\lim_{n=\infty} L_n^{(m)}(f; x) = f^{(m)}(x)$, равномерно по x.

Если же для $x \in (a, b)$ имеет место $\mu_1 \geqq w(x) \geqq \mu_2 > 0$, и в том же интервале нормированные ортогональные многочлены $p_n(x)$ ограничены в совокупности, то взяв произвольный внутренний подинтервал (a_1, b_1) интервала (a, b), во всяком внутреннем подинтервале из (a_1, b_1) имеют место, равномерно по x, соотношения

(I) $$\sum_{x_{kn} \in (a_1, b_1)} |l_{kn}^{(m)}(x)| = O(n^m \log n)$$

и

(II) $$\sum_{x_{kn} \notin (a_1, b_1)} |l_{kn}^{(m)}(x)| = O(n^m).$$

Из (I) и из (II) вытекает, что если

$$f(x_1) - f(x_2) = o\left(\log^{-1} \frac{1}{|x_1 - x_2|}\right) \qquad (a \leqq x_1, x_2 \leqq b),$$

то во всяком внутреннем подинтервале из (a, b) равномерно выполняется

$$\lim L_n^{(m)}(f; x) = f^{(m)}(x).$$

Показывается также, что оценка (I) не может быть улучшена никаким выбором (даже независимым от ортогональных многочленов) системы основных точек.

ON GROUPS EVERY CYCLIC SUBGROUP OF WHICH IS A POWER OF THE GROUP

By
F. SZÁSZ (Debrecen)

(*Presented by* L. RÉDEI)

To the memory of my beloved master T. SZELE

§ 1. Introduction

In recent years many group-theoretical investigations originated with the purpose of giving an explicit characterization of some class of groups defined by some fixed property. The present paper is devoted to a similar problem.

Let k be a fixed natural number not necessarily different from zero and let G be an arbitrary not necessarily abelian group written multiplicatively. *The k-th power of the group* G is a subgroup denoted by G^k, generated by the set of the k-th powers of all elements of G.

We shall call an arbitrary group G *a group with property* P, if to every cyclic subgroup H of G there exists a natural number k for which $H = G_k$ holds. Every cyclic group is a group with property P. The aim of this paper is to give an elementary proof of the converse statement, namely we are going to prove that a group with property P is necessarily cyclic, i. e. property P is characteristic for cyclic groups.

§ 2. Preliminaries

The characterization to be given of cyclic groups will be more easy to survey, if we prove first of all three lemmas. Before this, let us make some terminological remarks.

We shall not assume the commutativity of the groups considered and therefore we use multiplicative notation throughout. We denote the order of the element g by the symbol $O(g)$. A group is called *torsion free* if it contains no element of finite order other than the unity. On the other hand, groups every element of which is of finite order are called *torsions groups*. If a torsion group G contains an element of maximal order m, then we say that G is an *m-bounded group*. A group is termed a *p-group* if all its elements have orders which are powers of the fixed prime number p. A finite p-group always has a non-trivial center [2].[1] In any finite group G of order $n = p_1^{\alpha_1} \dots p_i^{\alpha_i} \dots p_r^{\alpha_r}$ where $p_1, \dots, p_i, \dots, p_r$ are different prime numbers, there exists necessarily a p_i-subgroup of order $p_i^{\alpha_i}$, called a *Sylow* p_i*-group* of the group G [2]. A group is called *metabelian* if its center contains its derived group, i. e., the subgroup generated by the commutators $[a, b] = a^{-1}b^{-1}ab$ of

[1] The numbers in brackets refer to the Bibliography given at the end of this paper.

all group elements a and b. In metabelian groups the relation $[a, bc] = [a, b]\cdot[a, c]$ holds for arbitrary elements a, b, c [2].

Now we prove the lemmas.

LEMMA 1. *A group with property* P *is a torsion free or a finite group.*

PROOF. Let G be an arbitrary not torsion free group with property P, and let $g \in G$ where $O(g) = n > 1$. If $\{g\} = G^k$, then G is obviously an m-bounded group, where $m \leqq nk$. In this case the set of all different powers of G is finite. Therefore the number of all cyclic subgroups of G is also finite, hence the torsion group G contains but a finite number of elements.

LEMMA 2. *Every homomorphic image G' of a group G with property* P *is a group with property* P.

PROOF. Let G be an arbitrary group with property P and φ a homomorphic mapping of G onto the group G'. If H' is an arbitrary cyclic subgroup generated by the element h' of G', then there exists an element h of G for which $h\varphi = h'$ holds. Let $\{h\} = H = G^k$ and thus $h = g_1^k \cdots g_s^k$ with certain elements $g_1, \ldots, g_s$ of G. Obviously, we have $H\varphi = H'$ and $h' = (g_1\varphi)^k \cdots (g_s\varphi)^k$, whence $(G^k)\varphi = (G')^k$. On the other hand, if $g \in G$, then $(g^k)\varphi = (g\varphi)^k \in H\varphi$, therefore we can write $(G')^k \subseteq (G^k)\varphi$, and thus $H' = (G')^k$.

LEMMA 3. *A finite group with property* P *has always a non-trivial center.*

PROOF. Let ε be an arbitrary endomorphism of the group G with property P and H a cyclic subgroup of G. Then, by the proof of Lemma 2, $(G^n)\varepsilon = (G\varepsilon)^n$ is valid for any natural number n, and thus $H = G^k$ implies $H\varepsilon = (G^k)\varepsilon = (G\varepsilon)^k \subseteq G^k = H$. But then every subgroup S is also mapped into itself, because $g \in S$ implies $\{g\} \subseteq S$ and $g\varepsilon \in \{g\}\varepsilon \subseteq \{g\} \subseteq S$ for every endomorphism ε of the group G. Thus every *Sylow p-group* of a finite group G is normal in G, and so G is the direct product of its *Sylow p-groups*. Since finite *p-group* is never without center and the center of a direct product is the direct product of the centers of the direct factors, therefore G has a non-trivial center.

§ 3. A characterization of the cyclic groups

Now we prove the following

THEOREM. *An arbitrary group is cyclic if and only if it has property* P.[2]

[2] By our theorem it is clear that if a group G has property P then to every subgroup S of G there exists a natural number n, for which $S = G^n$ holds. In a previous paper written in Hungarian with German summary (Über zyklische Gruppen, *Acta Scientiarum Universitatis Debreceniensis de Ludovico Kossuth* nominatae) I have proved this theorem in the following weaker form: if for every subgroup S of a group G there exists a natural number n for which $S = G^n$ holds, then G is necessarily cyclic. The proof of this weaker theorem made use of the fundamental theorem of R. BAER on the structure of hamiltonian groups [1].

PROOF. Let first G be an arbitrary finite group with property P and let m be the order of G. In this case we prove our statement by induction on m. The theorem is trivially true for $m=1$, and we assume that it holds for all finite groups with property P of order s less than n. By Lemma 3 the center C of G differs from $\{1\}$. By our induction hypothesis the factor group G/C is cyclic, because it is by Lemma 2 a group with property P and on the other hand the order of G/C is $< m$. Considering that the factor group with respect to the center is never cyclic, we conclude that G is a finite abelian group with property P. By the fundamental theorem on finite abelian groups there exists in every *Sylow p-group* $S_p \neq \{1\}$ of G a cyclic direct factor $A_p \neq \{1\}$ of S_p, so that $S_p = A_p \times B_p$ with a certain subgroup B_p of S_p. But $A_p = G^k$ and $G = A_p \times B_p \times M$ imply $A_p = A_p^k \times B_p^k \times M^k$. Here we must have $B_p^k = M^k = \{1\}$ and $A_p = A_p^k$, so that $(k, p) = 1$. Then $B_p = \{1\}$, and thus each p-component of G being cyclic, G itself is a finite cyclic group.

If G is infinite, then, by Lemma 1, the group G with property P is necessarily torsion free. Let now a and b be arbitrary elements of G. The subgroup $T = \{a, b\}$ is metabelian, because by the proof of Lemma 3 every subgroup of G is normal, therefore the commutator $c = [a, b] = a^{-1}b^{-1}ab$ is contained in the intersection $\{a\} \cap \{b\}$, and so in the center of the subgroup T of G. If $c = a^s = b^t$, then $c^{st} = [a, b]^{st} = [a^s, b^t] = [c, c] = 1$ and thus $[a, b] = 1$. Therefore G is an abelian torsion free group. If $n \neq 0$, then the mapping φ of G into itself, where $g\varphi = g^n$, is an isomorphism. Let $g \in G$ and $g \neq 1$. If $\{g\} = G^k$, then G itself is necessarily cyclic.

On the other hand, every cyclic group has obviously the property P; this completes the proof of our Theorem.

(Received 4 July 1955)

Bibliography

[1] R. BAER, Situation der Untergruppen und Struktur der Gruppe, *S. B. Heidelberg. Akad.*, **2** (1933), pp. 12—17.

[2] А. Г. Курош, Теория групп (Москва, 1953).

О ГРУППАХ, КАЖДАЯ ЦИКЛИЧЕСКАЯ ПОДГРУППА КОТОРЫХ ЯВЛЯЕТСЯ НЕКОТОРОЙ СТЕПЕНЬЮ ГРУППЫ

Ф. Сас (Дебрецен)

(Резюме)

Произвольная группа G называется группой со свойством P, если каждая циклическая подгруппа H группы G является некоторой степенью G^k группы G, где k-й степень G^k группы G есть подгруппа, порожденная множеством k-ых степеней всех элементов группы G. Элементарным способом доказывается, что произвольная группа G является циклической тогда и только тогда, если она имеет свойство P.

BEMERKUNG ZU EINER ARBEIT VON T. SZELE

Von

OTTÓ STEINFELD (Szeged)

(*Vorgelegt von* L. RÉDEI)

Herrn Professor L. KALMÁR *zum 50. Geburtstag gewidmet*

§ 1

Wir nennen den Unterring (das Ideal, das einseitige Ideal usw.) $\mathfrak{a}$ des Ringes R *trivial*, wenn $\mathfrak{a} = (0)$ oder $\mathfrak{a} = R$ ist. T. SZELE hat in einer vor kurzem erschienenen Arbeit folgenden Satz bewiesen:

Ein nur triviale Linksideale enthaltender Ring ist entweder ein Schiefkörper oder ein Zeroring mit p Elementen, wobei p eine Primzahl bezeichnet.[1]

Natürlich gilt dasselbe auch für die Ringe, die nur triviale Rechtsideale enthalten, deshalb bekommen wir sogleich folgende

FOLGERUNG. *Enthält der Ring R mindestens ein nicht-triviales Linksideal, so besitzt er auch mindestens ein nicht-triviales Rechtsideal.*

Wir werden in unserer Arbeit *aus einem gegebenen nicht-trivialen Linksideal* des Ringes R ein nicht-triviales Rechtsideal von R bestimmen. In der Bestimmung des Rechtsideals spielen die linksseitigen und die rechtsseitigen Annullatoren eine wichtige Rolle. Es bezeichne H eine Teilmenge des Ringes R. Das Element ϱ von R, wofür die Bedingung

$$\varrho H = (0)$$

gilt, nennen wir *einen linksseitigen Annullator von H*. Es ist klar, daß alle linksseitigen Annullatoren von H ein Linksideal von R bilden. Dieses Linksideal heißt *das annullierende Linksideal von H*, und wird mit $\mathfrak{l}_H$ bezeichnet. Ebenso können wir *das annullierende Rechtsideal von H* definieren, welches mit $\mathfrak{r}_H$ bezeichnet wird.

Wir brauchen zu den folgenden den Begriff des Quasiideals.[2] Wir nennen den Untermodul $\mathfrak{a}$ des Ringes R ein *Quasiideal*, wenn für $\mathfrak{a}$ die Bedingung

$$R\mathfrak{a} \cap \mathfrak{a}R \subseteq \mathfrak{a} \tag{1}$$

[1] T. SZELE, Die Ringe ohne Linksideale, *Buletin Ştiinţific (Bucureşti)*, **1** (1950), S. 783—789. Im Zusammenhang mit diesem Satz machte mich Herr Professor L. RÉDEI darauf aufmerksam, daß dieser Satz ein Korollar des Wedderburn—Artinschen Satzes der halbeinfachen Ringen ist.

[2] Das Quasiideal wird zuerst in unserer Arbeit „On ideal-quotients and prime ideals" (*Acta Math. Acad. Sci. Hung.*, **4** (1953), S. 289—298) definiert. Es ist leicht zu sehen, daß man den Begriff des Quasiideals für den Doppeltoperatormodul verallgemeinern kann.

gilt. Es ist klar, daß die Ideale, die einseitigen Ideale und der Durchschnitt eines Links- und Rechtsideals alle Quasiideale vom Ringe R sind. Wir werden hier beweisen: Wenn der Ring R mindestens ein nicht-triviales Quasiideal enthält, dann enthält R auch nicht-triviale Links- und Rechtsideale. Weiterhin bestimmen wir aus einem gegebenen nicht-trivialen Quasiideal des Ringes R ein nicht-triviales Links- und Rechtsideal von R.

§ 2

Wir werden jetzt aus einem nicht-trivialen Linksideal des Ringes R ausgehend ein nicht-triviales Rechtsideal von R bestimmen. Es gilt der folgende

SATZ 1.[3] *Ist* $\mathfrak{a}$ *ein nicht-triviales Linksideal vom Ringe* R, *so gibt es unter den folgenden Moduln mindestens ein nicht-triviales Rechtsideal:*

1. αR *für ein geeignetes Element* $\alpha(\in \mathfrak{a})$;
2. *die Menge der in* $\mathfrak{a}$ *enthalteten linksseitigen Annullatoren von* R, *d. h.* $\mathfrak{a} \cap \mathfrak{l}_R$;
3. *das annullierende Rechtsideal* $\mathfrak{r}_\alpha$ *eines geeigneten Elementes* α *von* $\mathfrak{a}$;
4. *die Menge der (in* R *enthalteten) linksseitigen Annullatoren von* $\mathfrak{a}$, *d.h.* $\mathfrak{l}_\mathfrak{a}$.

BEWEIS. 1. Gibt es ein Element α in $\mathfrak{a}$ mit

$$\alpha R \neq \begin{cases} (0), \\ R, \end{cases}$$

so ist offenbar αR ein nicht-triviales Rechtsideal von R.

2. Wenn wir kein solches Element $\alpha(\in \mathfrak{a})$ finden, so muß für irgendein Element $\xi(\in \mathfrak{a})$ entweder

$$\xi R = (0) \qquad (\xi \in \mathfrak{a}) \tag{2}$$

oder

$$\xi R = R \qquad (\xi \in \mathfrak{a}) \tag{3}$$

bestehen. Sei $\xi(\neq (0))$ ein Element von $\mathfrak{a}$, wofür (2) richtig ist.[4] Da dieses Element $\xi(\in \mathfrak{a})$ ein linksseitiges Annullator von R ist, ist die Menge der in $\mathfrak{a}$ enthalteten linksseitigen Annullatoren von R, d. h. der Durchschnitt $\mathfrak{d}$ des Linksideals $\mathfrak{a}$ und des annullierenden Linksideals $\mathfrak{l}_R$ von R vom Nullideal verschieden:

$$\mathfrak{d} = \mathfrak{l}_R \cap \mathfrak{a} \neq (0).$$

Wegen $\mathfrak{d} \subseteq \mathfrak{a} \subset R$ ist $\mathfrak{d}$ auch von R verschieden. Nachher brauchen wir nur zu zeigen, daß $\mathfrak{d}$ ein Rechtsideal von R ist. Da es für $\mathfrak{d}$ nach der Definition

[3] Für die gefällige Hilfe in der Abfassung dieses Satzes und seines Beweises spreche ich meinen herzlichen Dank Herrn Professor L. FUCHS aus.

[4] In diesem Falle ist auch das durch das Element ξ erzeugte Rechtsideal nicht-trivial.

$\mathfrak{d} \subseteq \mathfrak{l}_R$ gilt, bekommen wir

$$\mathfrak{d} R = (0) \subset \mathfrak{d}.$$

Damit haben wir bewiesen, daß $\mathfrak{d}$ ein nicht-triviales Rechtsideal von R ist. (Sogar ist $\mathfrak{d}$ wegen seiner Definition auch ein nicht-triviales Ideal von R.)

3. Wenn (2) unter den Elementen von $\mathfrak{a}$ nur für $\xi = 0$ erfüllt ist, dann muß für jedes Element $\xi(\neq 0, \in \mathfrak{a})$ (3) bestehen. Gibt es ein Element $\xi(\neq 0)$ in $\mathfrak{a}$, wofür neben (3) auch

$$\xi\varrho = 0 \qquad (\xi \neq 0, \in \mathfrak{a}) \tag{4}$$

erfüllt ist, wo ϱ ein vom Nullelement verschiedenes Element in R bezeichnet, so ist wegen (3) und (4) das annullierende Rechtsideal $\mathfrak{r}_\xi$ des obigen Elementes $\xi(\neq 0, \in \mathfrak{a})$ vom Nullideal und von R verschieden, deshalb ist $\mathfrak{r}_\xi$ ein nicht-triviales Rechstideal von R.

4. Wenn zuletzt für jedes Element $\xi(\neq 0)$ von $\mathfrak{a}$ die Bedingung (3) gilt, und (4) für jedes $\xi(\neq 0, \in \mathfrak{a})$ nur mit dem Element $\varrho = 0$ besteht,[5] dann wird das annullierende Linksideal $\mathfrak{l}_\mathfrak{a}$ von $\mathfrak{a}$ ein nicht-triviales Rechtsideal von R sein. Es ist gleich einsehbar, daß $\mathfrak{l}_\mathfrak{a}$ ein Rechtsideal von R ist. Ferner gibt es wegen (3) ein Element ε in R, wofür

$$\xi\varepsilon = \xi \qquad (\xi \neq 0, \in \mathfrak{a})$$

gilt. Zu einem gegebenen Element $\xi(\neq 0, \in \mathfrak{a})$ finden wir nur ein einziges Element ε. Wenn nämlich auch $\xi\eta = \xi$ erfüllt ist, so bekommen wir

$$\xi(\varepsilon - \eta) = 0,$$

woraus wegen unserer Voraussetzung $\varepsilon - \eta = 0$, d. h.

$$\varepsilon = \eta$$

folgt. Wir werden jetzt nachweisen, daß das Element ε ein linksseitiges Einselement in R ist. Wegen $\xi\varepsilon = \xi$ ist für jedes Element $\varrho(\in R)$ die Gleichung $\xi\varepsilon\varrho = \xi\varrho$ erfüllt, deshalb gilt auch

$$\xi(\varepsilon\varrho - \varrho) = 0,$$

woraus wieder wegen unserer Voraussetzung

$$\varepsilon\varrho = \varrho \qquad (\varrho \in R)$$

folgt. Da das Element ε wegen $\varepsilon\mathfrak{a} = \mathfrak{a} \neq (0)$ in $\mathfrak{l}_\mathfrak{a}$ nicht enthalten ist, ist $\mathfrak{l}_\mathfrak{a}$ von R verschieden.

Bezeichne $\xi'(\in R)$ ein Element, wofür

$$\xi\xi' = \varepsilon \qquad (\xi' \in R, \ \xi \in \mathfrak{a})$$

gültig ist. Wir behaupten, daß auch

$$\xi'\xi = \varepsilon \qquad (\xi' \in R, \ \xi \in \mathfrak{a}) \tag{5}$$

[5] Es ist nicht schwer auch folgendes zu beweisen: Es sei $\mathfrak{a}$ ein nicht-triviales Linksideal vom Ringe R. Wenn für jedes Element $\xi(\neq 0)$ von $\mathfrak{a}$ die Bedingung (3) gilt, und für jedes $\xi(\neq 0, \in \mathfrak{a})$ (4) nur mit dem Element $\varrho = 0$ besteht, so ist $\mathfrak{a}$ ein Schiefkörper und das annullierende Linksideal $\mathfrak{l}_\mathfrak{a}$ von $\mathfrak{a}$ ist von dem Nullideal verschieden. Ferner ist $\mathfrak{a} = R\varepsilon$, wo ε das Einselement in $\mathfrak{a}$ und zugleich ein linksseitiges Einselement in R ist.

besteht. Da

$$\xi(\xi'\xi) = (\xi\xi')\xi = \varepsilon\xi = \xi$$

ist, deshalb ist auch (5) wegen $\xi\varepsilon = \xi$ und wegen der Eindeutigkeit von ε richtig. Aus $\xi \in \mathfrak{a}$ und (5) folgt, daß das linksseitige Einselement ε im Linksideal $\mathfrak{a}$ enthalten ist. Wegen $\varepsilon \in \mathfrak{a}$ und $\mathfrak{a} \neq R$ gibt es mindestens ein Element ϱ in R mit

$$\varrho\varepsilon \neq \varrho \qquad (\varrho \in R).$$

Da es wegen $\varepsilon^2 = \varepsilon$ für das Element $\varrho\varepsilon - \varrho (\neq 0, \in R)$

$$(\varrho\varepsilon - \varrho)\varepsilon = 0$$

erfüllt ist, ist das Element $\varrho\varepsilon - \varrho (\neq 0, \in R)$ ein linksseitiges Annullator von $\varepsilon\mathfrak{a} = \mathfrak{a}$, deshalb ist das annullierende Linksideal $\mathfrak{l}_\mathfrak{a}$ von $\mathfrak{a}$ auch vom Nullideal verschieden. Somit haben wir Satz 1 bewiesen.

BEISPIELE. 1. Im vollen Matrizenring K_2 des Ranges 4 über einem Schiefkörper K bildet die Menge $\mathfrak{a}$ aller Matrizen $\begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix}$ $(a, b \in K)$ ein nicht-triviales Linksideal. Bezeichne 1 das Einselement von K. Z. B. ist für das Element $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} (\in \mathfrak{a})$

$$\alpha K_2$$

ein nicht-triviales Rechtsideal von K_2. (Fall 1.)

2. Im Ringe R bestehend aus den Elementen $0, \alpha, \beta, \alpha + \beta$ mit den definierenden Relationen

$$\alpha + \alpha = \beta + \beta = 0, \quad \alpha^2 = \alpha, \quad \alpha\beta = \beta, \quad \beta\alpha = \beta^2 = 0$$

ist $\mathfrak{b} = (0, \beta)$ ein nicht-triviales Linksideal. Wegen $\beta R = (0)$ ist βR ein triviales Rechtsideal in R, aber der Durchschnitt des Linksideals $\mathfrak{b}$ und des annullierenden Linksideals $\mathfrak{l}_R$ von R ist ein nicht-triviales Rechtsideal von R. (Fall 2.)

Für das nicht-triviale Linksideal $\mathfrak{a} = (0, \alpha)$ ist das annullierende Linksideal $\mathfrak{l}_\mathfrak{a}$ von $\mathfrak{a}$ ein nicht-triviales Rechtsideal von R. (Fall 4.)

§ 3

Nachher werden wir die ähnlichen Ergebnisse für die Quasiideale bekanntmachen. Ähnlich zu der erwähnten Folgerung des obigen Satzes von SZELE gilt

SATZ 2. *Enthält der Ring R mindestens ein nicht-triviales Quasiideal, so besitzt er mindestens auch ein nicht-triviales Linksideal und nicht-triviales Rechtsideal.*

Zum Beweis des Satzes 2 benutzen wir den folgenden

SATZ 3. *Der Schiefkörper K hat nur triviale Quasiideale.*

BEWEIS. Es sei $\mathfrak{a}(\neq(0))$ ein Quasiideal von K. Betrachten wir ein beliebiges Element $a(\neq 0)$ in $\mathfrak{a}$. Da die Gleichungen

$$x = a\xi, \quad x = \eta a \qquad (a \in \mathfrak{a};\ x, \xi, \eta \in K)$$

für jedes Element x von K lösbar sind, deshalb ist x in $\mathfrak{a}K$ und $K\mathfrak{a}$ enthalten. Also ist $x(\in K)$ nach (1) ein Element von $\mathfrak{a}$, d. h. $\mathfrak{a} = K$.

BEWEIS DES SATZES 2. Es ist klar, daß es den Fall des Linksideals zu betrachten genügt. Im Gegensatz zu unserer Behauptung setzen wir voraus, daß R mindestens ein nicht-triviales Quasiideal aber kein nicht-triviales Linksideal enthält. Dann ist R nach dem Satz von SZELE entweder ein Schiefkörper oder ein Zeroring mit p Elementen, wobei p eine Primzahl bezeichnet. Nach Satz 3 hat ein Schiefkörper nur triviale Quasiideale. Da ein Zeroring mit p Elementen nur triviale Quasiideale besitzt, deshalb muß R wegen dieses Widerspruchs mindestens ein nicht-triviales Linksideal enthalten.

Jetzt werden wir aus einem gegebenen nicht-trivialen Quasiideal des Ringes R je ein nicht-triviales Links- und Rechtsideal von R genau bestimmen. Wegen Satz 1 können wir uns auf den Fall beschränken, in dem das Quasiideal von R kein einseitiges Ideal von R ist.

SATZ 4. *Wenn das nicht-triviale Quasiideal $\mathfrak{a}$ kein einseitiges Ideal vom Ringe R ist, so ist $R\mathfrak{a}$ ein nicht-triviales Linksideal und $\mathfrak{a}R$ ein nicht-triviales Rechtsideal von R.*[6]

Zum Beweis des Satzes 4 benutzen wir den folgenden

HILFSSATZ 2. *Wenn es für das Quasiideal $\mathfrak{a}$ des Ringes R entweder $\mathfrak{a} \subseteq R\mathfrak{a}$ oder $\mathfrak{a} \subseteq \mathfrak{a}R$ gilt, so ist $\mathfrak{a}$ der Durchschnitt eines Linksideals und eines Rechtsideals, ferner ist auch*

$$\mathfrak{a} = (\mathfrak{a}, R\mathfrak{a}) \cap (\mathfrak{a}, \mathfrak{a}R) \tag{4}$$

erfüllt.[7]

BEWEIS. Wir können uns auf den Fall $\mathfrak{a} \subseteq R\mathfrak{a}$ beschränken, weil der Fall $\mathfrak{a} \subseteq \mathfrak{a}R$ analog ist. Es bezeichne D den Durchschnitt von $(\mathfrak{a}, R\mathfrak{a})$ und $(\mathfrak{a}, \mathfrak{a}R)$, d. h.

$$D = (\mathfrak{a}, R\mathfrak{a}) \cap (\mathfrak{a}, \mathfrak{a}R). \tag{5}$$

Aus (5) folgt

$$\mathfrak{a} \subseteq D. \tag{6}$$

Wegen der Voraussetzung $\mathfrak{a} \subseteq R\mathfrak{a}$ ergibt sich statt (5)

$$D = R\mathfrak{a} \cap (\mathfrak{a}, \mathfrak{a}R). \tag{5'}$$

[6] Es ist klar, daß der Satz 2 ein Korollar des Satzes 4 und der Folgerung des Satzes von SZELE ist, es war jedoch zweckmäßig für Satz 2 einen selbständigen Beweis zu geben.

[7] Aus dem Hilfssatz 2 folgt sofort das folgende Korollar: Wenn der Ring R ein einseitiges Einselement enthält, so sind alle Quasiideale von R als Durchschnitt eines Linksideals und eines Rechtsideals darstellbar.

Betrachten wir jetzt ein beliebiges Element δ von D. Wegen (5′) können wir das Element δ in der Form

$$\delta = \sum_{i=1}^{m} \varrho_i a_i = a + \sum_{j=1}^{n} \beta_j \sigma_j \qquad (a, a_i, \beta_j \in \mathfrak{a}\,;\ \varrho_i, \sigma_j \in R) \tag{7}$$

schreiben. Das Element

$$\sum_{j=1}^{n} \beta_j \sigma_j = \sum_{i=1}^{m} \varrho_i a_i - a$$

gehört wegen $\sum_{i=1}^{m} \varrho_i a_i$, $a \in R\mathfrak{a}$ zum Modul $R\mathfrak{a}$ und wegen der Definition auch zu $\mathfrak{a}R$, weshalb $\sum_{j=1}^{n} \beta_j \sigma_j$ nach (1) ein Element von $\mathfrak{a}$ ist. Daraus folgt nach (7), daß $\delta (\in D)$ in $\mathfrak{a}$ enthalten, d. h.

$$D \subseteq \mathfrak{a} \tag{8}$$

ist. Die Bedingungen (6) und (8) beweisen gerade die Behauptung (4), womit der Beweis beendet ist.

BEWEIS DES SATZES 4. Selbstverständlich ist es genug die Behauptung für $R\mathfrak{a}$ nachzuweisen. Im Gegensatz zu unserer Behauptung setzen wir voraus, daß $R\mathfrak{a}$ ein triviales Linksideal von R ist, d. h. entweder $R\mathfrak{a} = (0)$ oder $R\mathfrak{a} = R$ gilt. Da $\mathfrak{a}$ wegen der Voraussetzung kein Linksideal ist, ist der Fall $R\mathfrak{a} = (0)$ unmöglich, also muß $R\mathfrak{a} = R$ bestehen. Folglich ist $\mathfrak{a} \subseteq R\mathfrak{a}$ und der Hilfssatz 2 ist anwendbar. Nach (4) bekommen wir jetzt

$$\mathfrak{a} = R \cap (\mathfrak{a}, \mathfrak{a}R) = (\mathfrak{a}, \mathfrak{a}R),$$

woraus $\mathfrak{a}R \subseteq \mathfrak{a}$ folgt, was bedeutet, daß $\mathfrak{a}$ im Gegensatz zu unserer Voraussetzung ein Rechtsideal von R ist. Damit haben wir Satz 4 bewiesen.

(Eingegangen am 28. Juli 1955.)

ЗАМЕЧАНИЕ К ОДНОЙ СТАТЬЕ Т. СЕЛЕ

О. Штеинфельд (Сегед)

(Резюме)

Развивая дальше результат Т. Селе, автор находит с помощью данного нетривиального левого идеала $\mathfrak{a}$ кольца R (т. е. такого, для которого $\mathfrak{a} \neq (0)$, $\mathfrak{a} \neq R$) некоторый нетривиальный правый идеал. Далее выводятся подобные теоремы для введенного автором квазиидеала. Именно доказывается, что если кольцо R содержит нетривиальный квазиидеал, то оно содержит также нетривиальные левый и правый идеалы, и из нетривиального квазиидеала конструируется нетривиальный левый и правый идеал.

Les Acta Mathematica paraissent en russe, français, anglais et allemand et publient des mémoires du domaine des sciences mathématiques.

Les Acta Mathematica sont publiés sous forme de fascicules qui seront réunis en un volume.

On est prié d'envoyer les manuscrits destinés à la rédaction et écrits à la machine à l'adresse suivante:

Acta Mathematica, Budapest 62, Postafiók 440.

Toute correspondance doit être envoyée à cette même adresse.

Le prix de l'abonnement est de 110 forints par volume.

On peut s'abonner à l'Entreprise pour le Commerce Extérieur de Livres et Journaux *„Kultúra"* (Budapest, VI., Sztálin út 21. Compte-courant No. 43-790-057-181) ou à l'étranger chez tous les représentants ou dépositaires.

The Acta Mathematica publish papers on mathematics in Russian, French, English and German.

The Acta Mathematica appear in parts of various size, making up one volume.

Manuscripts should be typed and addressed to:

Acta Mathematica, Budapest 62, Postafiók 440.

Correspondence with the editors and publishers should be sent to the same address.

The rate of subscription to the Acta Mathematica is 110 forints a volume. Orders may be placed with *„Kultúra"* Foreign Trade Company for Books and Newspapers (Budapest, VI., Sztálin út 21. Account No. 43-790-057-181) or with representatives abroad.

Die Acta Mathematica veröffentlichen Abhandlungen aus dem Bereiche der mathematischen Wissenschaften in russischer, französicher, englischer und deutscher Sprache.

Die Acta Mathematica erscheinen in Heften wechselnden Umfanges. Mehrere Hefte bilden einen Band.

Die zur Veröffentlichung bestimmten Manuskripte sind, mit Maschine geschrieben, an folgende Adresse zu senden:

Acta Mathematica, Budapest 62, Postafiók 440.

An die gleiche Anschrift ist auch jede für die Redaktion und den Verlag bestimmte Korrespondenz zu richten.

Abonnementspreis pro Band 110 Forints. Bestellbar bei dem Buch- und Zeitungs-Außenhandels-Unternehmen *„Kultúra"* (Budapest, VI., Sztálin út 21. Bankkonto Nr. 43-790-057-181) oder bei seinen Auslandsvertretungen und Kommissionären.

64,— forint

INDEX

Fejér, L., Verschiedene Bemerkungen elementarer Natur über die Grundpolynome, die bei den parabolischen Interpolationen auftreten 227
T. Sós, Vera and *Turán, P.*, On some new theorems in the theory of diophantine approximations 241
Turán, P., On the instability of systems of differential equations 257
Rédei, L., Die gruppentheoretischen Zetafunktionen und der Satz von Hajós 271
Hájek, J. and *Rényi, A.*, Generalization of an inequality of Kolmogorov 281
Rényi, A., On a new axiomatic theory of probability 285
Császár, Á., Sur la structure des espaces de probabilité conditionnelle 337
Takács L., On stochastic processes connected with certain physical recording apparatuses 363
Mikolás, M., Zur Theorie der Gammafunktion, der Riemannschen Zetafunktion und verwandter Funktionen. I. 381
Gehér, L., On approximately differentiable functions of two variables 439
Carlitz, L., A special symmetric equation in a finite field 445
Császár, Á., Sur les fonctions localement monotones au sens généralisé 451
Fleischer, I., A note on subdirect products 463
Freud, G., Über differenzierte Folgen der Lagrangeschen Interpolation 467
Szász, F., On groups every cyclic subgroup of which is a power of the group . . . 475
Steinfeld, O., Bemerkung zu einer Arbeit von T. Szele 479

Technikai szerkesztő: Molnár Ferenc

A kiadásért felelős: Az Akadémiai Kiadó igazgatója — Műszaki felelős: Farkas Sándor
A kézirat beérkezett: 1955. X. 14. — Terjedelem: 22,50 (A/5) ív, 4 ábra.

Csongrádmegyei Nyomdaipari Vállalat, Szeged 55-3605 — Felelős vezető: Vincze György