

MIKLÓS MIKOLÁS

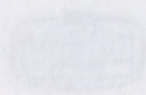
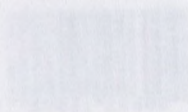
REAL FUNCTIONS,
ABSTRACT SPACES AND
ORTHOGONAL SERIES

AKADÉMIAI KIADÓ, BUDAPEST

**REAL FUNCTIONS,
ABSTRACT SPACES AND ORTHOGONAL SERIES**

MIKLÓS MIFOLAS

*Professor of Mathematics
Eötvös University, Budapest*



AKADÉMIAI KIADÓ, BUDAPEST

81238

REAL FUNCTIONS, ABSTRACT SPACES AND ORTHOGONAL SERIES

MIKLÓS MIKOLÁS

Professor of Mathematics
Technical University, Budapest

PART ONE

INTRODUCTION TO THE THEORY OF REAL FUNCTIONS

CHAPTER I: ELEMENTS OF SET THEORY

1.1 Set relations, operations with sets and characteristic functions	37
1.11 Fundamental notions - 1.12 Algebra of sets - 1.13 Sequences of sets - 1.14 Characteristic functions of sets	
1.2 Countable and non-countable sets	38
1.21 The concept of countability - 1.22 Examples - 1.23 The non-countability of continua - 1.24 Additional observations	
1.3 Comparison of cardinals and theorem of equivalence	34
1.31 Cardinality and equivalence of sets - 1.32 The Cantor-Schroeder theorem	
1.4 Sets of higher cardinality, transfinite cardinals	36
1.41 Cardinality of Euclidean spaces - 1.42 Sets of higher cardinality than the continuum power - 1.43 The notion of power set - 1.44 Transfinite induction and cardinals	
1.5 Important set structures: Boolean algebras and rings	40
1.51 Fields of sets; Boolean rings (algebra of sets) - 1.52 Quotient rings of sets, Boolean algebras	
1.6 Simple applications in elementary set theory	42
1.61 Countability of the set of all functions from a countable set to a countable set - 1.62 Cardinality of the set of all functions from a countable set to a continuum	

MTAK



AKADÉMIAI KIADÓ, BUDAPEST

0516

This book is the enlarged and revised version of the Hungarian "Valós függvénytan és ortogonális sorok" published by Tankönyvkiadó, Budapest

ABSTRACT SPACES AND
ORTHOGONAL SERIES

MIKOLÁS MIKLÓS

Professor of Mathematics
Technical University, Budapest

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

ISBN 963 05 6652 4

© 1994 M. Mikolás

Published by Akadémiai Kiadó
1117 Budapest
Prielle Kornélia u. 19-35.

All rights reserved. No part of this publication may be reproduced or stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the written permission of the publisher.

Printed in Hungary



M. TUD. AKADÉMIAI KÖNYVTÁRA

Könyvteltár 8471.../19 25... sz.

CONTENTS

From the preface to the Hungarian edition	17
Foreword to the English edition	21
Historical overview	23

PART ONE

INTRODUCTION TO THE THEORY OF REAL FUNCTIONS

CHAPTER 1: ELEMENTS OF SET THEORY

1.1 Set relations, operations with sets and characteristic functions	27
1.11 Fundamental notions – 1.12 Algebra of sets – 1.13 Sequences of sets – 1.14 Characteristic functions of sets	
1.2 Countable and non-countable sets	30
1.21 The concept of countability – 1.22 Examples – 1.23 The non-countability of continua – 1.24 Additional observations	
1.3 Comparison of cardinals and theorem of equivalence	34
1.31 Cardinality and equivalence of sets – 1.32 The Cantor–Bernstein theorem	
1.4 Sets of higher cardinality, transfinite cardinals	36
1.41 Cardinality of Euclidean spaces – 1.42 Sets of higher cardinality than the continuum power – 1.43 The notion of power-set – 1.44 Transfinite induction and cardinals	
1.5 Important set structures: Boolean and sigma-rings	40
1.51 Fields of sets: Boolean ring (algebra) and σ -ring (algebra) – 1.52 Generating fields of sets, Borel sets	
1.6 Simple applications in elementary analysis	42
1.61 Countability of discontinuities of the first kind – 1.62 Countability of break points – 1.63 Cardinal of the set of functions continuous in an interval	

CHAPTER 2: POINT SETS IN EUCLIDEAN SPACES

2.1	Distance and neighbourhood in space E_n	47
	2.11 Distance of points in E_n , fundamental inequalities – 2.12 Neighbourhood of a point and some ideas connected with that – 2.13 Topological mapping, foundations of the general theory of curves and surfaces, C. Jordan's curve theorem	
2.2	Convergence and clustering point of a sequence of points	50
	2.21 Fundamentals on convergent sequences of points, Cauchy's criterion – 2.22 Clustering point, upper and lower limit – 2.23 Connection between the notions of convergence and clustering point	
2.3	Notions and theorems concerning derivate sets	53
	2.31 Accumulation point of a point set, the Bolzano-Weierstrass theorem, compactness – 2.32 Condensation point, Lindelöf's theorem – 2.33 Derivate of a point set and related notions – 2.34 Closure of a set – 2.35 Perfectness of the set of condensation points – 2.36 On the cardinality of sets consisting of accumulation and condensation points	
2.4	Basic properties of closed and open sets	57
	2.41 Closure and boundary points – 2.42 Openness and related concepts – 2.43 Principle of duality; union and meet of closed and open sets – 2.44 The normal representation of a linear open set – 2.45 Cantor's ternary set	
2.5	The Cantor-Bendixson theorem, Cantor's meet theorem	62
	2.51 Distance of point sets – 2.52 The theorem of Cantor and Bendixson – 2.53 Diameter of a point set and the meet theorem of Cantor	
2.6	Covering theorems, outer measures, null sets	65
	2.61 Borel's covering theorem – 2.62 Jordan's and Lebesgue's outer measure – 2.63 Null sets – 2.64 Elementary properties of outer measures – 2.65 Sierpinski's lemma and Vitali's covering theorem	

CHAPTER 3: EXTENSION OF THE CONCEPT OF FUNCTION.

LIMIT, CONTINUITY AND DIFFERENTIABILITY OF A POINT FUNCTION

3.1	Functions on abstract sets (operators) and their classification	73
	3.11 The general concept of function, Cantor's theorem on the cardinality of operator sets – 3.12 Classification of operators; linearity, boundedness and norm of a functional (functional operation) – 3.13 The notion of linear operator	
3.2	Limit and cluster value of a point function, relating to a set	76
	3.21 The limit of a point function at an accumulation point of the basic set – 3.22 Cluster value (sublimit) of a point function, upper and lower limit – 3.23 Upper and lower fluctuation value – 3.24 The analogue of Cauchy's criterion, operations with generalized limits	
3.3	Properties of continuous and semi-continuous point functions	80
	3.31 Local continuity and semi-continuity relating to a basic set – 3.32 Characterization of point functions continuous on a set, closure with respect to the fundamental operations – 3.33 Main properties: extension of theorems of Weierstrass, Bolzano and Heine	

3.4	The approximation theorems of Weierstrass; Stone's theorem	83
3.41	The two approximation theorems of Weierstrass – 3.42 Uniform approximation on a prescribed set by elements of some class of functions, Stone's theorem – 3.43 Conclusions and remarks	
3.5	Sequences and series of continuous point functions	88
3.51	Uniform convergence on some set of sequence or series of point functions, the Weierstrass criterion – 3.52 Use of uniform convergence in case of continuous functions, quasi-uniform convergence – 3.53 Baire's and Young's classes of functions	
3.6	Generalized differentiability, derivate numbers	91
3.61	Total and partial differential, differential coefficients of a point function – 3.62 Generalized derivate numbers – 3.63 Geometric interpretation, main properties of differentiable functions	
3.7	Constructing everywhere continuous, nowhere differentiable functions	94
3.71	Historical remarks, example of Weierstrass – 3.72 An elementary example of a function continuous everywhere whose derivative does not exist at any point – 3.73 Extension of the latter construction, further developments	

CHAPTER 4: MONOTONIC, ABSOLUTELY CONTINUOUS FUNCTIONS AND FUNCTIONS OF BOUNDED VARIATION

4.1	Basic properties, continuous and pure jumping part of a monotonic function	99
4.11	Monotony and the existence of one-sided limits – 4.12 Constructing a pure saltus function with prescribed jumps – 4.13 Decomposition of an arbitrary bounded monotonic function into continuous and pure jumping part – 4.14 Continuity almost everywhere of a monotonic function	
4.2	Differentiability almost everywhere of monotonic functions	103
4.21	Lebesgue's theorem on monotonic functions (Proof after Rajchman and Saks) – 4.22 Historical comments	
4.3	Applications; Fubini's theorem for series of functions	106
4.31	The Denjoy–Saks–Young theorem on derivate numbers – 4.32 Theorem of Fubini on term-by-term differentiation of a series of monotonic functions	
4.4	Functions of bounded variation, the notion of total variation	108
4.41	Rectifiable curves and functions of bounded variation; the total variation – 4.42 Some elementary observations, additivity of total variation – 4.43 Closure of the class of functions of bounded variation with respect to the fundamental operations	
4.5	Jordan's and Lebesgue's theorems on functions of bounded variation ..	111
4.51	Jordan's theorem: representation of a function of bounded variation by means of monotonic functions – 4.52 Consequences, the theorem of Lebesgue – 4.53 Example of a monotone increasing and continuous, non-constant function with derivative vanishing almost everywhere	

4.6	Absolutely continuous functions	113
4.61	The idea of absolutely continuous function as a tightening of the notion of function of bounded variation – 4.62 Connection with uniform continuity and differentiability; ordinary Lipschitz condition – 4.63 Jordan's theorem for absolutely continuous functions – 4.64 The constancy of an absolutely continuous function with derivative vanishing almost everywhere – 4.65 Extension of the results of this chapter for point functions of several variables; utilization of interval functions	

CHAPTER 5: JORDAN MEASURE AND RIEMANN INTEGRAL. RIEMANN-STIELTJES INTEGRAL

5.1	Foundations of Jordan's measure theory	119
5.11	Fundamental concepts of the general measure theory – 5.12 Jordan measurability and Jordan measure in space E_n – 5.13 Basic properties of $\mu_J(S)$	
5.2	Riemann integral of a function of one variable as Jordan measure with sign; upper and lower Darboux integral	122
5.21	Connection between Riemann's integral concept and Jordan's theory of measure – 5.22 Darboux integrals and limit definition of the Riemann integral	
5.3	Riemann's and Lebesgue's criteria for (R) -integrability	125
5.31	Oscillation sum and Riemann integral – 5.32 Lebesgue's criterion: continuity almost everywhere of a Riemann integrable function	
5.4	Deeper study of (R) -integrable functions; limiting process under the (R) -integral sign	128
5.41	The role of the class of (R) -integrable functions in elementary analysis – 5.42 Operations with (R) -integrable functions; problem of passage to the limit term-by-term – 5.43 Example of a sequence (or series) of (R) -integrable functions whose limit (sum) is not (R) -integrable	
5.5	Use of a primitive; properties of the indefinite (R) -integral	131
5.51	The Leibniz-Newton formula for an arbitrary (R) -integrable function – 5.52 Investigation of the existence of a primitive – 5.53 Properties of (R) -integral functions – 5.54 Connection between the operations of (R) -integration and differentiation – 5.55 Volterra's example of a function differentiable in some closed interval whose derivative is bounded but not (R) -integrable	
5.6	Improper (R) -integral and Riemann integral of several dimensions	136
5.61	Kinds of improper (R) -integrals – 5.62 (R) -integral on a region in the space E_n – 5.63 Iterated (R) -integration, extended integral mean	
5.7	The Stieltjesian generalization of Riemann integral	138
5.71	Mathematical and physical background of the extension due to Stieltjes	
5.72	Definition of the (RS) -integral; calculation of $\int_a^b f(x) dg(x)$ if $f(x)$ is continuous and $g(x)$ is a step function – 5.73 Additivity, integration by parts for (RS) -integrals – 5.74 Conditions for existence of $\int_a^b f dg$; the cases when the integrator function g is monotonic, of bounded variation or absolutely continuous – 5.75 Main properties of the (RS) -integral, mean value theorems	
5.8	Connection of (RS) -integrals with (R) -integrals; Radon's and Burkill's integral	145

5.81 Transforming $\int_a^b f(x) dg(x)$ into an (R) -integral if $g'(x)$ exist and is (R) -integrable – 5.82 (RS) -integral of several dimensions (Radon's integral), derivative with respect to a region and the notion of density – 5.83 Burkil integral of an interval function as a generalized (RS) -integral

CHAPTER 6: LEBESGUE'S MEASURE AND INTEGRAL. MEASURABLE FUNCTIONS

6.1	Elements of Lebesgue's measure theory	149
	6.11 Defects of Riemann's concept of integral, the need of its developing – 6.12 Lebesgue measurable sets. Lebesgue measure in the space E_n – 6.13 Basic properties of $\mu_L(S)$ – 6.14 Total additivity of the (L) -measure	
6.2	Geometric introduction of the Lebesgue integral of a function of one variable; upper and lower Young integrals	154
	6.21 (L) -integral of a non-negative function as the (L) -measure of an ordinate set, Young integrals – 6.22 Extension to bounded function of arbitrary sign – 6.23 Some elementary features of the (L) -integral	
6.3	Lebesgue's original definition for the (L) -integral of a bounded function; (L) -integrability and measurability	157
	6.31 Definition of the (L) -integral by means of Lebesgue sums, the connection of the latter with Young sums – 6.32 The notion of measurable function, level sets – 6.33 Equivalence of the (L) -integrability and measurability of a bounded function	
6.4	Other definition forms; fundamental idea of the treatment due to F. Riesz	163
	6.41 Limit definition of the (L) -integral – 6.42 Use of (L) -integrals formed from characteristic functions – 6.43 The Riesz theory of (L) -integral	
6.5	Comparison of Lebesgue's concept of integral with that of Riemann; properties of measurable functions	165
	6.51 (L) -integrability and (R) -integrability – 6.52 Basic properties of measurable functions – 6.53 Closure with respect to the limit operation	
6.6	Almost uniform approximation of a measurable function by measurable functions; relationship between measurability and continuity	168
	6.61 Almost uniform convergence, Egorov's theorem – 6.62 Lusin's theorem – 6.63 "Correction" of a measurable function to a continuous one	

CHAPTER 7: PROPERTIES OF THE LEBESGUE INTEGRAL. GENERALIZED LEBESGUE INTEGRALS

7.1	Fundamental properties of the (L) -integral of a bounded function over a measurable set	173
	7.11 (L) -integral taken on a set – 7.12 (L) -equivalence, basic properties, mean value theorems – 7.13 Absolute (total) continuity of the (L) -integral – 7.14 Total additivity of the (L) -integral	

7.2	Integration of sequences and series of functions	176
7.21	The limit theorem of Lebesgue – 7.22 Convergence with respect to (L) -measure – 7.23 Term-by-term integration of a series consisting of measurable functions and Arzelà's theorem	
7.3	Indefinite (L) -integral with bounded integrand and the notion of primitive	179
7.31	Calculation of an (L) -integral by means of a primitive – 7.32 Determination of primitives for a bounded function by (L) -integration – 7.33 Properties of (L) -integral functions, notion of Lebesgue point – 7.34 The role of Lischitz condition in the theory, relation between (L) -integration and differentiation	
7.4	(L) -integral with unbounded basic set or integrand	183
7.41	Extension of Lebesgue's integral definition, generalized Lebesgue sum – 7.42 Reduction of the (L) -integral of an unbounded function to the case of bounded integrand – 7.43 Remarks on (L) -integrability in wider sense, the class $L(S)$	
7.5	Passage to the limit under the sign of a generalized (L) -integral	187
7.51	Lebesgue's main theorem on integration of a sequence of functions – 7.52 Fatou's lemma and the theorem of B. Levi – 7.53 A pendant of uniform continuity in the class $L(S)$	
7.6	Theorems on arbitrary (L) -integral functions	194
7.61	Extension of the Leibniz-Newton formula – 7.62 (L) -integrating the derivative of a function of bounded variation – 7.63 Absolute continuity of an arbitrary (L) -integral function – 7.64 Derivation of the indefinite (L) -integral and Lebesgue points – 7.65 Absolute continuity as a characteristic property, the total variation of an (L) -integral function – 7.66 (L) -integral function of an unbounded derivative, extensions of the (L) -integral due to Denjoy and Perron – 7.67 Integration by parts and by substitution for (L) -integrals	
7.7	(L) -integral of several dimensions; Fubini's theorem on successive integration	203
7.71	(L) -integral of a point function of several variables, Fubini's theorem for double (L) -integrals – 7.72 Remarks on the generalization of Lebesgue's theory.	

CHAPTER 8: MEASURES DEFINED ON BOOLEAN RINGS AND THE ABSTRACT LEBESGUE INTEGRAL. LEBESGUE-STIELTJES INTEGRAL

8.1	Extension of measures, Carathéodory's theorem	205
8.11	Introduction – 8.12 Measure and measure space defined on a Boolean ring – 8.13 The notion of completeness and Borel's measure; closure of a measure space – 8.14 Generating an abstract measure by means of an outer measure induced by premeasure – 8.15 Carathéodory's extension theorem, Lebesgue's continuation of premeasure	
8.2	Lebesgue integral on abstract measure space	208
8.21	Measurability of a function on abstract space, analogues of the abstract Lebesgue integral – 8.22 Definition of the abstract Lebesgue integral – 8.23 Properties of abstract (L) -integrals, Daniell integral	

8.3	Lebesgue–Stieltjes measure and integral	211
8.31	The Lebesgue–Stieltjes measure as a particular case of the notion of abstract measure – 8.32 (<i>LS</i>)-integral and its connection with the (<i>RS</i>)-integral – 8.33 Relation between (<i>LS</i>)-integral and (<i>L</i>)-integral – 8.34 Integration by parts and transformation of (<i>LS</i>)-integrals – 8.35 Development of the theory of (<i>LS</i>)-integral, applications in probability theory	
8.4	Measures and integrals on product spaces	217
8.41	(<i>L</i>)-measures of higher dimension generated from linear (<i>L</i>)-measures – 8.42 Cartesian product of semi-rings – 8.43 Product space and product measure originating from measure spaces; (<i>L</i>)-integral on a product space	
8.5	Fubini's theorem for product spaces; connections with probability theory	221
8.51	Preliminaries on μ -measures – 8.52 The theorem of Fubini for abstract Lebesgue integrals – 8.53 Use of the notions of this chapter in Kolmogorov's construction of probability theory	

PART TWO

FUNCTION SPACES AND ORTHOGONAL EXPANSIONS

CHAPTER 9: FOUNDATIONS OF FUNCTIONAL ANALYSIS. THE SPACES l^2 AND L^2

9.1	Orthogonal series and functional analysis	229
9.11	The notions of orthogonal system and series – 9.12 Geometrical background and role of Lebesgue's theory – 9.13 Bearing on the formation of functional analysis	
9.2	Vectorial interpretation of the Euclidean space E_n	231
9.21	Vectorial basic operations, norm in the space E_n – 9.22 Angle of inclination, orthogonality and projection – 9.23 Orthogonal system and its completeness – 9.24 Use of an orthogonal system as a base	
9.3	Transfer of fundamental ideas to the Hilbert space l^2	235
9.31	Definition of Hilbert's sequence space; vectorial basic operations, norm and notions connected with orthogonality – 9.32 Concept of distance, triangle inequality, theory of point sets in the space l^2 – 9.33 The role of l^2 -convergence in connection with the "decomposition problem"	
9.4	Further generalization: the function space L^2 as a vector field	238
9.41	The Hilbert space of square integrable functions – 9.42 Vectorial basic operations, norm, fundamental inequalities – 9.43 Notions issuing from norm and orthogonality – 9.44 Example of an incomplete orthonormal system of functions	
9.5	Introduction of the notions of distance and convergence in space L^2	243
9.51	Distance and L^2 -convergence – 9.52 Comparison of the L^2 -convergence with uniform and pointwise convergence – 9.53 The Riesz–Fischer convergence criterion – 9.54 Strong and weak convergence – 9.55 The space L^2 as a set of abstract points	
9.6	The problem of decomposition and the concept of general Fourier series	250
9.61	Decomposition of an element of the space L^2 in terms of a given orthogonal system – 9.62 The formal determination of coefficients – 9.63 General	

Fourier series. – 9.64 Parallelism between fundamental notions of the theory of orthogonal series and their geometric analogues

CHAPTER 10: SPECIAL ORTHOGONAL SYSTEMS AND SERIES

10.1	The trigonometric base system	255
	10.11 Orthogonality of the trigonometric base system – 10.12 Lagrange's lemma – 10.13 Completeness of the trigonometric system, the cosine and sine system	
10.2	The ordinary (trigonometric) Fourier series and its complex form	259
	10.21 Ordinary Fourier coefficients and Fourier series – 10.22 Complex form of the ordinary Fourier series – 10.23 Connection with complex power series and Laurent series	
10.3	Legendre polynomials	262
	10.31 Existence of a sequence of polynomials orthogonal in a finite interval – 10.32 Representation of Legendre polynomials by means of a Rodrigues formula – 10.33 Some important properties of the Legendre polynomials – 10.34 Generating function – 10.35 Recurrence formula and differential equation – 10.36 Norm and minimum property, Legendre series	
10.4	Properties of polynomial systems orthogonal with respect to weight functions	270
	10.41 Definition of the weighted orthogonality of polynomial sequences – 10.42 Recursive representation of orthogonal polynomials belonging to a given interval and to a prescribed weight function – 10.43 Basic features, minimum property – 10.44 Some pertinent abbreviations and notations, the class $L^2_{\rho(x)}(a, b)$	
10.5	Jacobi, Laguerre and Hermite polynomials	276
	10.51 Tchebyshev polynomials of the first and second kind, Tchebyshev's approximation – 10.52 Generalization: Gegenbauer (ultraspherical) and Jacobi (hypergeometric) polynomials – 10.53 Infinite basic intervals: Laguerre and Hermite polynomials – 10.54 Outline of the classical systems of orthogonal polynomials – 10.55 Generating functions and norms – 10.56 Fourier series relating to the classical orthogonal polynomial systems	
10.6	Orthogonal systems arising from differential equations of Sturm–Liouville type	282
	10.61 Main types of Sturm–Liouville boundary value problems – 10.62 Orthogonality of eigenfunctions – 10.63 Example: discussion of the differential equation of harmonic oscillation – 10.64 The orthogonal system formed from Bessel functions of the first kind, Fourier–Bessel series – 10.65 Main properties of orthogonal systems of Sturm–Liouville type, Sturm–Liouville series	
10.7	The Gram–Schmidt orthogonalization procedure	287
	10.71 Orthogonalization in the space $L^2(a, b)$ – 10.72 Expansion of an “arbitrary” function in terms of a system of functions prescribed “optionally”	
10.8	Applications: systems of Haar's type, Rademacher and Walsh functions	290
	10.81 Closure of an orthonormal system; complete orthogonal polynomial sequences obtained by orthogonalization – 10.82 Generating the Haar system from characteristic functions belonging to the dyadic partition of $[0, 1]$ – 10.83 Completeness of Haar's system – 10.84 The related $\{\mu^{(r)}_{k_1 \dots k_\nu}\}$ family of orthogonal functions – 10.85 Rademacher's orthonormal system and its importance in	

probability theory – 10.86 Extension of the Rademacher system due to Walsh, generalized Walsh systems and series

CHAPTER 11: PROPERTIES OF ARBITRARY ORTHOGONAL SYSTEMS AND SERIES IN HILBERT'S FUNCTION SPACE

11.1	Minimum property of the partial sums of a Fourier series	299
	11.11 Preliminary comments – 11.12 Fundamental identity for partial sums of orthogonal series – 11.13 Minimum property of the partial sums of general Fourier series and Bessel's identity – 11.14 Idea of approximation with respect to square integral (Bessel's approximation)	
11.2	Bessel's inequality and the Parseval–Hurwitz formula	301
	11.21 The Bessel inequality and its geometrical interpretation – 11.22 The Parseval–Hurwitz formula and its extended form – 11.23 Historical remarks	
11.3	L^2 -convergence of general Fourier series	303
	11.31 The general expansion theorem – 11.32 Consequences	
11.4	The Riesz–Fischer theorem: isomorphy of Hilbert's function space and sequence space	304
	11.41 Characterization of the elements of L^2 by a countable number of data – 11.42 The fundamental theorem of F. Riesz and E. Fischer – 11.43 The isomorphic connection of L^2 and l^2	
11.5	Convergence almost everywhere of arbitrary orthogonal series	307
	11.51 The need to use the classical concept of convergence – 11.52 The principal problems of convergence theory of orthogonal series – 11.53 The Rademacher–Menshov theorem	
11.6	The convergence and summation problem of general Fourier series	310
	11.61 The integral form of a partial sum of the general Fourier series and its kernel function – 11.62 The notions of Lebesgue function and singular integral – 11.63 Pertinent observations	

CHAPTER 12: CONVERGENCE OF SPECIAL FOURIER SERIES

12.1	Elementary convergence theorems on ordinary Fourier series; the Riemann–Lebesgue lemma	313
	12.11 Elementary remarks on trigonometric Fourier series – 12.12 The order of magnitude of Fourier constants in case of a k -times differentiable function; a simple expansion theorem – 12.13 Sharpenings; case of a function of bounded variation – 12.14 The Riemann–Lebesgue lemma – 12.15 Conclusions and additions	
12.2	Dirichlet's formula and Riemann's localization theorem	319
	12.21 Dirichlet's integral – 12.22 The localization theorem of Riemann – 12.23 Riemann's condition for the Fourier series representation of a function	
12.3	Convergence tests of Dini, Dirichlet, Jordan and Lipschitz	323
	12.31 Dini's and Jordan's condition – 12.32 Regular points and Lipschitz conditions – 12.33 Dirichlet's criterion – 12.34 Other local or global tests of convergence	

12.4	Example of a continuous function with ordinary Fourier series divergent at some point	329
12.41	Representation by Fourier series and continuity – 12.42 Uniform continuity of partial sums of the series $\sum (\sin nx)/n$ – 12.43 The Weierstrass example – 12.44 Du Bois–Reymond singularity and Schwarz’s example – 12.45 On points of divergence of the ordinary Fourier series of a continuous function	
12.5	The problem of uniqueness; term-by-term integrability of ordinary Fourier series	335
12.51	Connection between sum and coefficients of a convergent trigonometric series – 12.51 Outline of some pertinent results, set of uniqueness – 12.53 Term-by-term integration of ordinary Fourier series – 12.54 Theorem of term-by-term differentiation	
12.6	The conjugate series	339
12.61	The notion of conjugate series and integral form of its partial sums – 12.62 Convergence criterion of Pringsheim – 12.63 Consequences, theorem of F. Lukács	
12.7	Ordinary Fourier series of some special functions and harmonic analysis	342
12.71	The function $\{x\}$ – 12.72 Bernoulli polynomials and their applications – 12.73 The Hurwitz zeta function and Hurwitz’s formula – 12.74 Development of $\cos(vx)$ into cosine series and expansion of cotangent in terms of partial fractions – 12.75 Fourier series of $ \sin x $ – 12.76 Lerch’s zeta function and its functional equation – 12.77 Remarks on harmonic analysis; generation of Bessel functions by complex Fourier series	
12.8	Series of Haar’s type; equiconvergence theorems for Sturm–Liouville series	355
12.81	Representation of continuous functions by Haar’s series – 12.82 Proof of Haar’s expansion theorem on the basis of the connection between Haar’s system and some system of characteristic functions – 12.83 Convergence theorem for Fourier series of type $\mu_{k_1 \dots k_\nu}^{(r)}$ – 12.84 Haar’s equiconvergence theorem and its analogues for series of Sturm–Liouville type	

CHAPTER 13: SUMMATION OF ORDINARY FOURIER SERIES

13.1	Linear summation methods and problems of Tauber’s type	361
13.11	Introductory observations – 13.12 Foundations of the general theory of summation processes – 13.13 Classical summation methods: Cesàro’s, Hölder’s and Abel–Poisson summation – 13.14 Historical comments on the development of linear methods of summation – 13.15 Tauberian problems, theorem of Hardy and Landau	
13.2	$(C, 1)$ -summation of ordinary Fourier series; Fejér’s fundamental and approximation theorem	369
13.21	Closed form for the arithmetic means of the partial sums of ordinary Fourier series: Fejér’s integral – 13.22 The summation theorems of L. Fejér	
13.3	Some consequences of Fejér’s theorem	372
13.31	Applications relating to the limit of partial sums and to the codomain of their arithmetic means for ordinary Fourier series – 13.32 The Gibbs phenomenon – 13.33 Concluding from the $(C, 1)$ -summability of Fourier series to its convergence and to the approximation theorem of Weierstrass	

13.4	Lebesgue's summation theorem and other generalizations	376
13.41	Local absolute and ordinary integral mean – 13.42 The summation theorem of Lebesgue and its proof found by Fejér – 13.43 M. Riesz's extension for the Lebesgue theorem and "strong summation" of Fourier series in the sense of Hardy and Littlewood	
13.5	Summation of ordinary Fourier series by the Abel-Poisson method	380
13.51	Poisson's integral formula and its simplest consequences – 13.52 A sufficient condition for (A) -summability of Fourier series – 13.53 Connection with Cesàro's summation and with Dirichlet's problem in the plane	
13.6	Application of the recent (D) -method	384
13.61	On certain insufficiencies of the classical summation processes – 13.62 Definition of the (D) -summation and its connection with zeta-functions – 13.63 The pertinent singular integral and its kernel function – 13.64 (D) -summability of ordinary Fourier series – 13.65 Remarks on effectiveness of the (D) -method; related summation processes and generalized integro-differentiation	

CHAPTER 14: FOURIER INTEGRALS

14.1	Foundations of the theory of integral transforms of Fourier and related type	395
14.11	Trigonometric integrals and Fourier transforms, connection with the Laplace transformation – 14.12 Formal background of Fourier integrals as well as of Fourier's integral theorem and inversion formulae	
14.2	Fourier transformation in Hilbert's function space	398
14.21	Complex L -space and L^2 -space – 14.22 The Parseval formula for Fourier integrals – 14.23 Plancherel's theorem and Paley-Wiener theorem	
14.3	Properties of absolutely convergent Fourier transforms	408
14.31	The complex Fourier transforms as linear operators – 14.32 The notion of convolution and its simplest properties – 14.33 Proof of the main theorem on convolutions	
14.4	Fourier's integral theorem on classical convergence conditions	414
14.41	Transfer of the convergence theory of ordinary Fourier series – 14.42 Cesàro summation of integrals – 14.43 Analogue of the Fejér-Lebesgue summation theorem	
14.5	Non-absolutely convergent Fourier integrals	423
14.51	The fundamental problem – 14.52 A general equiconvergence theorem – 14.53 Invertibility of non-absolutely convergent Fourier transforms – 14.54 Complementary results; generalization of Poisson's summation formula	
14.6	Fourier-Stieltjes transforms	432
14.61	Fourier-Stieltjes transformation as a common extension of the Fourier transformation and ordinary Fourier series – 14.62 Probabilistic aspects, relation between distribution and characteristic function of a random variable	

CHAPTER 15: FURTHER ABSTRACT SPACES, APPLICATIONS

15.1 Complex L^p -spaces of F. Riesz	437
15.11 Grounds for a common extension of $L(a, b)$ and $L^2(a, b)$ - 15.12 Main points of the theory of complex L^p -spaces - 15.13 Riesz's theorems on bounded linear functionals - 15.14 The sequence spaces ℓ^p of Riesz	
15.2 Linearity, metrics and completeness; Banach's fixed point theorem	441
15.21 The most important structural properties of abstract spaces - 15.22 General concept of linear (vector) space as well as of metric and complete space - 15.23 The Banach fixed point theorem for complete metric spaces	
15.3 Abstract Hilbert space and Neumann's foundation of quantum mechanics	444
15.31 The abstract Hilbert space H as a common generalization of L^2 and ℓ^2 - 15.32 The Fréchet-Riesz theorem on linear operators in H and spectral decomposition theorem - 15.33 Linear algebraical interpretation of the theory - 15.34 Applications: equivalence of matrix and wave mechanics, the foundations of quantum mechanics by J. Neumann - 15.35 On the use of Dirac's impulse δ and related operators	
15.4 Banach space; concept of general topological space	450
15.41 The Banach space B as a further extension of H and the spaces L^p ; theorem of J. Neumann and P. Jordan - 15.42 Linear operators in B , dual Banach spaces - 15.43 Operator pendant of the geometric series (Neumann's series) - 15.44 General topological space and problem of metrization	
15.5 Extended applications concerning ordinary differential equations	454
15.51 Introductory comments - 15.52 General form of self-adjoint eigenvalue problems - 15.53 Extended orthogonality relations for eigenfunctions - 15.54 Transformation of self-adjoint eigenvalue problems into integral equations via Green functions	
15.6 Partial differential equations and orthogonal series	457
15.61 Solution of the problem of homogeneous vibrating string by Fourier's method - 15.62 Case of inhomogeneous string, use of Sturm-Liouville series - 15.63 Discussion of Dirichlet's problem in the plane by harmonic series of functions - 15.64 The Laplace (potential) equation in space; spherical harmonics and Laplace series - 15.65 Fourier's solution of the problem of linear heat conduction and of similar problems - 15.66 Membrane vibration and Fourier-Bessel series; boundary problems of further partial differential equations of second order in physics	
15.7 Solving integral equations by operator methods	465
15.71 Linear integral equation of Fredholm type - 15.72 C.G. Neumann's existence and expansion theorem; Fredholm's alternative - 15.73 Case of symmetric kernel function, the Hilbert-Schmidt expansion theorem	

BIBLIOGRAPHY	471
--------------------	-----

INDEX	483
-------------	-----

FROM THE PREFACE TO THE HUNGARIAN EDITION

The subject-matter and concept of this book have principally been matured in the course of my lectures given at the Faculty of Sciences of the L. Eötvös University in Budapest during the 1960's. According to both my own teaching experience and the findings of my colleagues in Hungary and abroad, as well as the remarks of young mathematicians, a certain gap has recently come into being between the lectures on analysis delivered at "graduate" and "postgraduate" levels (always striving after the most up-to-date treatment) and the introductory analysis courses for the first two classes; and this gap has become wider, following the acceleration of development. Due to the considerably abstract character and conciseness of those "up-to-date" lectures on analysis, the substance of knowledge presented by them seems to be strange compared to the "classical" material — for the majority of students; and thus, having no roots in the mind, it is soon forgotten.

Consequently, demands have lately arisen for a monograph which could moderate to a certain extent the above difficulties and so to say "builds a bridge" between the elementary and higher parts of analysis, with the aim of presenting new concepts and methods in accordance with *historical development* — even at the expense of some increase in size. Besides, a further requirement has been added by some Mathematics Departments of the Technical University of Budapest; at the Faculties of Mechanical and Electrical Engineering new programmes for engineer-mathematicians and other specialists have been started, which include much more knowledge of the modern theory of real functions together with elements of functional analysis; and these can only be understood with text-books of the type described.

The structure and content of my book are guided by the two objectives just explained; i.e. the work is intended primarily for graduate students of mathematics and physics, mechanical and electrical engineering, but generally for all who want a comprehensive view of the ramifications of modern analysis and its main applications. The readers are expected to have a preliminary knowledge of analysis at undergraduate level.

The first part of the book (consisting of eight chapters) discusses the foundations of the *theory of real functions*. After an introduction to set theory, it provides higher knowledge on the main function classes of elementary

analysis, then leads the reader to the contemporary theory of differentiation and integration. Within this scope — bearing in mind the worldwide development in the last quarter of the century — is applied consistently the new abstract algebraic terminology such as the notions of ring and semi-ring, Boolean algebra and sigma-algebra. Just after the preliminaries on sets, the basic concepts of the general theory of linear operators are exposed; further the Stieltjes integral as well as the Lebesgue integrals on abstract measure spaces are of greater moment than is usual. Vitali's covering theorem, this important expedient of the theory, is substituted throughout by the so-called Sierpinski lemma, the latter being simpler and easier to handle.

H. LEBESGUE's measure and integral theory is established in this book essentially in the classical way, i.e. on the one hand purely geometrically (based on Lebesgue's extension of the Jordan measure), on the other hand by using W. H. YOUNG's definition which means the direct generalization of the Riemann integral. In addition, the pertinent theory of F. RIESZ¹ is briefly elucidated. The Riesz method undoubtedly presents the most elegant and useful way to the Lebesgue integral, provided that we introduce the up-to-date integral concept immediately and independently, and the idea of this integral as a functional with advantageous properties is needed rather than Lebesgue's measure. There are, however, significant application fields — such as probability theory and statistics — where the concept of measure is preferred, so that a certain kind of "composed" treatment has become the most frequent in the latest literature.

Thus the authors, besides discussing the measure concept in the classical manner, also incorporate in the study of Lebesgue integral, at least in part, the Riesz theory starting from monotonic sequences of certain "simple" functions. (See the bibliography.) The conception in my book differs from this "model" only in the considerable attention given to the geometrical background and to the relation between Jordan's and Lebesgue's measure, as well as between Riemann's and Lebesgue's notion of integral.

The second part which is nearly of the same length deals with the historically primary application fields, first of all with the theory of *orthogonal series*, Fourier analysis and — in close connection with them — with foundations of *functional analysis*. The introduction of D. HILBERT's sequence space and function space, item of general Fourier series is followed by presentation of the most important special orthogonal systems and series, including the relevant fundamental facts on the independent theory of eigenvalue problems of Sturm-Liouville type. Then the convergence and summation theory of general and special Fourier series are examined in full, with particular regard to the ordinary (trigonometric) Fourier transformation. In the last chapter, a detailed survey on further abstract spaces and their wide

¹ See e.g. B. SZŐKEFALVI-NAGY, Introduction to real functions and orthogonal expansions (Oxford, 1964).

applications are dealt with, where the main consideration is given to the theory of linear operators as well as to differential and integral equations.

It is well-known that, following the fundamental results achieved by L. FEJÉR, F. and M. RIESZ, A. HAAR, the activity of mathematicians belonging to the Hungarian analytic schools has largely contributed to the establishment of modern real analysis, and this feature of development still exist.

For this reason, I felt justified to insert or at least to mention in the text some results found by Hungarian researchers (among them the author) which are in accordance with the didactic aims of the work and have not yet been published in any book of this kind. A few examples: a new general elementary procedure to construct everywhere continuous non-differentiable functions; the "natural" generating Haar's orthogonal system from characteristic functions belonging to the dyadic partition of $[0,1]$; simple discussion of the Hurwitz and Lerch zeta-functions (used in the chapter on summation theory) by means of Fourier analysis; Haar's equiconvergence theorem for Sturm-Liouville series; a new, short proof of Haar's expansion theorem; extension of Fejér's summation process; discussion of the so-called (D) -method for ordinary Fourier series which yields a very simple necessary and sufficient summability criterion and is based on fractional integration; some new results on non-absolutely convergent Fourier integrals with applications; J. NEUMANN's foundation of quantum mechanics by means of abstract Hilbert space which depends on the Riesz-Fischer theorem; etc.

My intention was, of course, to give the reader an overview of the disciplines discussed here which includes also some information on the latest results. Several foot-notes and sections of expository character are meant for this purpose, mainly in the closing chapters of the two separate parts of the book. I think that the publications referred to in the text and listed in the bibliography may present a basis for further, more profound study of the subject by the readers.

Miklós Mikolás

FOREWORD TO THE ENGLISH EDITION

The Hungarian book had a favourable reception in Hungary and abroad, being awarded a prize in 1979. The first edition published in 1978 and the second one issued in 1980 went out of print within a short time.

The main changes in this rewritten English version are a comprehensive extension of the discussion of Fourier integrals in a separate chapter, and the expansion of the whole text to comply with newer developments.

The references have been completely revised and supplemented, furthermore, a detailed bibliography and subject index have been added.

All these changes do not affect the principal objectives of the work, though the enlargements and various revisions have changed it from a textbook to a scientific monograph of special character.

I am very much indebted to all my colleagues in Hungary and abroad who helped me with their valuable comments on the first and the present version of my book. Among them, I am particularly grateful to Professors D. Králik in Budapest, K. Tandori in Szeged, E. R. Love in Melbourne, as well as to B. Ross in New Haven, Conn. My special thanks are due to Professor Králik for the recent tireless review of the manuscript, and to Dr. N. Thurston (London) for his searching linguistic observations.

Budapest, June 1992

Miklós Mikolás

HISTORICAL OVERVIEW

INTRODUCTION TO THE THEORY OF REAL FUNCTIONS

The origins of the contemporary *theory of real functions* lie with the critical attacks — culminating at the end of the 19th century — upon the foundations of analysis, and with the exact formulation and far-reaching generalization of the notions “function” and “limit”. This theory has played an important role in the history of science ever since the turn of the century, mainly after the emergence of the set theory of G. CANTOR and the measure and integral theory of H. LEBESGUE. Not only did it fertilize practically every branch of classical analysis from complex function theory and harmonic analysis to the theory of differential and integral equations, but it has proved indispensable in the strict foundation of probability theory by A. N. KOLMOGOROV (1933) and thus has helped in establishing this theory and the related fields as intertwined with functional analysis which was born at the early 1900’s and bears a relation to calculus similar to that between the latter and the algebra of real numbers.

Since *functional analysis* deals with functions in an extended sense (called operators) for which the role of the independent variable is played by an ordinary function, certain sets of function (called function spaces) are as indispensable in developing functional analysis, as sets of number tuples (i.e. Euclidean spaces) are in developing calculus. Therefore, the study of the simplest function spaces and other “*abstract spaces*,” in particular the investigation of the spaces named after D. HILBERT and of the series built from elements of the latter, the so-called *orthogonal series*, ordinary and generalized Fourier series — all this is one of the most natural and most significant fields where the theory of real functions can be applied.

As a result of the advance sketched above, the mathematical disciplines in question, which were known only to a handful of mathematicians in the first decades of this century, have become more and more important in the education of mathematicians and physicists at the world’s leading universities. After World War II, however, these disciplines, concurrently with the development of modern theories of generalized functions (distribution theory, operational calculus) have gradually penetrated into other fields of natural science and engineering — together with the rapid growth of approximation theory and numerical analysis, with the spreading use of integral transforms as well as the recent methods of probability theory and statistics.

PART ONE

INTRODUCTION TO THE THEORY OF REAL FUNCTIONS

ELEMENTS OF SET THEORY

1.1 Set relations, operations with sets and characteristic functions

1.11 In the modern theory of real functions, the set theory founded by G. CANTOR at the end of the last century is used at every turn, with special regard to point sets.¹ Thus we should have first of all a comprehensive overview on the well-known standard facts of the general theory of sets.

Let S be a *set* consisting of certain things (not necessarily numbers or geometrical objects). We denote by $\kappa \in S$ that κ is an element of S (" κ belongs to S ") and by $\kappa \notin S$ the contrary. If every element of a set T belongs also to S , we say that T is a *subset* (part) of S or that S *contains* T . These relations are expressed by the notations

$$T \subseteq S \quad \text{and} \quad S \supseteq T, \quad (1.1-1)$$

respectively, which are analogues of $\leq$ and $\geq$. Two sets are called *equal*, if their elements are identical, i.e., each of them contains the other; this is denoted by the sign of equality. If T is a part of S but $T \neq S$, we write $T \subset S$ or $S \supset T$.

According to the above definition, whatever the set S , it contains the set with no element, i.e., the so-called *empty set* (denoted by $\emptyset$) and itself. These are called the "improper" subsets of S . Any other subset of S is called a *proper subset*.

Sets with special elements are often written by using braces $\{\dots\}$. For instance, $\{1, 2\}$ or $\{2, 1\}$ means the set consisting of 1 and 2 (the order of elements is irrelevant!); or $\{x : 1 < x < 2\}$ denotes the aggregate of real numbers x for which $1 < x < 2$, in other words: the open interval $(1, 2)$. In general, $\{x : \Phi_x\}$ is the totality of elements x satisfying the condition Φ_x .

A set S is said to be *finite* if it consists of a finite number of elements; otherwise S is infinite. When denoting a finite set, its elements are mostly indicated separately; in the case of sequences² it suffices to write the so-called generic element.

¹ See e.g. Cantor [2].

² The word "sequence" without attribute always refers to an infinite sequence.

1.12 Let us deal with the fundamental notions of the algebra of sets.

By the *union* (join) of two sets S and T we mean the set of elements occurring in *at least one* of them. The *meet* or *intersection* (common part) of S and T is the set of their common elements, i.e., those which occur in both of them. The corresponding notations are¹

$$S \cup T \quad \text{and} \quad S \cap T \quad (1.1-2)$$

respectively. For more than two sets we define the concepts of union and meet (intersection) in the same manner; then the symbols $\cup, \cap$ can be applied in a similar sense to the signs $\sum$ and $\prod$. For example, in the case of $s_n = \{1, 2, \dots, n\}$ the symbol $\bigcup_{n=1}^{\infty} s_n$ means the union of the sets $s_1, s_2, \dots$, and $\bigcap_{n=1}^{\infty} s_n$ the meet of them; the former consists plainly of all natural numbers, the latter is 1.

The definition of the *difference*

$$S - T \quad (1.1-3)$$

is formed from S by omitting such elements of S which also occur in T . (The notation $S \setminus T$ is used as well.) If $S \supseteq T$, then an alternative name of $S - T$ is the *complement* of T with respect to S which is denoted by $(CT)_S$.² So we have $(CT)_S \cup T = S$ and $(CT)_S \cap T = \emptyset$.

For the sake of brevity, instead of the expressions "forming the union, meet (intersection), difference or complement" we shall speak about *union*, *intersection*, *subtraction* or *complementation* of sets. Note that subtraction may not be considered simply as the inverse of the union operation, since a set equation $A \cup B = C$ implies $A = C - B$ only if $A \cap B = \emptyset$. On the other hand, the subtraction is closely connected with the intersection operation by the obvious identity

$$S - T = S - (S \cap T). \quad (1.1-4)$$

As is easy to see, the union operation and the intersection operation are both commutative and associative separately, and distributive with each other. Moreover, there hold the so-called *De Morgan's identities*

$$\begin{cases} S - \bigcap_n S_n = \bigcup_n (S - S_n) \\ S - \bigcup_n S_n = \bigcap_n (S - S_n) \end{cases} \quad (1.1-5)$$

where $\{S_n\}$ means an arbitrary system of subsets of S .

¹ An analogy can be drawn between these set operations and ordinary addition and multiplication. Thus, the denominations *logical sum*, and *logical product*, as well as the notations $S + T$, and $S \cdot T$ occur in the literature, too. Both operations are, however, *idempotent* in the sense that $S \cup S = S \cap S = S$.

² If it is irrelevant to indicate S as a basic set, we can write simply CT .

Hence we obtain at once a duality principle which can be also stated in a slightly more general form as follows: if we have an identity on certain subsets of a basic set X in which the union and intersection as well as subtraction occur, this identity remains true if we replace the signs $\cup, \cap, -$, by the signs $\cap, \cup, -$, respectively. (In this replacement, the sign $\subseteq$ changes to $\supseteq$, X turns into $\bar{X}$, and vice versa.)

We mention that it is expedient for some applications to introduce the so-called *symmetric difference*, whose usual sign and definition is

$$S \Delta T = (S - T) \cup (T - S). \quad (1.1-6)$$

Clearly, we also have

$$S \Delta T = (S \cup T) - (S \cap T). \quad (1.1-7)$$

The *Cartesian product*¹ of certain sets $S_1, S_2, \dots, S_n$ is defined as the set of all ordered n -tuples $(\kappa_1, \kappa_2, \dots, \kappa_n)$ with $\kappa_1 \in S_1, \kappa_2 \in S_2, \dots, \kappa_n \in S_n$ and it is denoted by $S_1 \times S_2 \times \dots \times S_n$. (Instead of writing several $\times$, we sometimes use the sign Π .) For instance, the Cartesian product of two number lines is the number plane; that of three number lines is Cartesian 3-space; in general the Cartesian product of ν number lines represents the Cartesian (or Euclidean) ν -space.

Two or more sets are said to be *disjoint*, if they have no common element, i.e. their intersection is $\emptyset$. Naturally, the empty set is disjoint from every set, but there also exist, trivially, systems consisting of arbitrarily many non-empty sets which are pairwise disjoint. For example, the sequence of intervals $(0, 1), (1, 2), (2, 3), \dots$

1.13 A sequence of sets $S_1, S_2, \dots$ is called *increasing* if $S_n \subseteq S_{n+1}$, and *decreasing* when $S_n \supseteq S_{n+1}$ for all n . If these conditions hold with $\subset$ instead of $\subseteq$ and with $\supset$ instead of $\supseteq$, respectively, then we say that $\{S_n\}$ is *strictly increasing* and *strictly decreasing*, respectively. For example, the interval sequence $[-n, n]$ ($n = 1, 2, \dots$) is strictly increasing, but $[-\frac{1}{n}, \frac{1}{n}]$ ($n = 1, 2, \dots$) is strictly decreasing. More generally, any sequence of "nested" intervals is decreasing.

It is clear that in the case of $S_1 \neq \emptyset$, an increasing sequence of sets S_n ($n = 1, 2, \dots$) cannot be disjoint; we shall see later (cf. 2.5.2) that a similar assertion holds on an important class of decreasing set systems, too. We shall need the following easily verifiable identity:

$$S_1 = (S_1 - S_2) \cup (S_2 - S_3) \cup \dots \cup M \quad (1.1-8)$$

¹ The denomination *direct product* is also usual.

where $S_1, S_2, \dots$ is any decreasing sequence of sets for which the differences $S_1 - S_2, S_2 - S_3, \dots$ are disjoint pairwise and $M = \bigcap_{n=1}^{\infty} S_n$.

1.14 Finally let us include the notion of *characteristic function* (or *set indicator*).¹ If a basic set X is given and S is a subset of it, the notation and definition of the characteristic function of S are:

$$\chi_s = \chi_s(x) = \begin{cases} 1 & \text{for } x \in S, \\ 0 & \text{for } x \in X - S. \end{cases} \quad (1.1-9.)$$

Hence we get for any pairs $S, T \in X$ the relation

$$\chi_{S \cup T} = \chi_S + \chi_T - \chi_S \cdot \chi_T; \quad (1.1-10)$$

and in particular, if S and T are disjoint, then

$$\chi_{S \cup T} = \chi_S + \chi_T. \quad (1.1-11)$$

Note also that

$$\chi_{S \cap T} = \chi_S \cdot \chi_T \quad (1.1-12)$$

and

$$\chi_{X-S} = 1 - \chi_S. \quad (1.1-13)$$

Also $S \subseteq T$ holds if and only if $\chi_S(x) \leq \chi_T(x)$ for all $x \in X$.

Properties (1.1-10)—(1.1-13) can easily be extended for more than two sets. Their significance stems from the fact that operations with sets may be plainly “translated” by means of these formulae into ordinary addition, multiplication and subtraction.

1.2 Countable and non-countable sets

1.21 As the most important property of a finite set may be considered to be the number of its elements. This is also usually called the *cardinality* (*cardinal* or *power*) of the set. We are now going to study one of the most fundamental and fruitful problems of set theory: how might it be possible to extend the concept of cardinality to infinite sets as well, i.e., to classify these sets according to the “lot” of the elements contained by them.

Since in the case of a finite set the cardinal number is obtained simply by *counting* the elements, it stands to reason that, for an infinite set, we should try first of all to establish a one-to-one correspondence (*bijective mapping*) between the elements of the set and the sequence of positive

¹ The latter name is used mostly in the recent literature on probability theory, where the expression “characteristic function” has another meaning.

integers 1, 2, 3, If the elements of the set in question can be arranged in a sequence, then we say that the set is *countable*.¹

1.22 Some classical examples:

1.2.1 *The set of rational numbers is countable.*²

P r o o f. Obviously we may restrict ourselves to positive rationals, i.e., to numbers of the form α/β where α, β are positive integers. Among these rationals, let us collect those for which $\alpha + \beta$ is equal to 2, 3 and so on. Writing these groups successively yields the sequence

$$1; \quad \frac{1}{2}, \frac{2}{1}; \quad \frac{1}{3}, \frac{2}{2}, \frac{3}{1}; \quad \frac{1}{4}, \frac{2}{3}, \frac{3}{2}, \frac{4}{1}; \quad \dots, \quad (1.2-1)$$

in which each positive rational number occurs at least once. We then leave out repetitions and find that the set in consideration can indeed be written as a sequence.

* * *

The above argument may also be illustrated in such a way that we form the infinite matrix

$$\begin{bmatrix} \langle 1, 1 \rangle & \langle 1, 2 \rangle & \langle 1, 3 \rangle & \langle 1, 4 \rangle & \dots \\ \langle 2, 1 \rangle & \langle 2, 2 \rangle & \langle 2, 3 \rangle & \dots & \\ \langle 3, 1 \rangle & \langle 3, 2 \rangle & \dots & & \\ \langle 4, 1 \rangle & \dots & & & \\ \dots & & & & \end{bmatrix} \quad (1.2-2)$$

and collect, in succession after $\langle 1, 1 \rangle$, the pairs of numbers situated along suitable diagonals. This is the so-called *first diagonal process of Cantor*.

More generally we have:

1.2.2. *For any fixed $\nu \geq 2$, the totality of ν -tuples of natural numbers constitutes a countable set.*

This is because the method applied to (1.2-2) can be adapted immediately in case of ν -tuples of the form $(\alpha_1, \alpha_2, \dots, \alpha_\nu)$, inasmuch as we write one after another the tuples for which $\alpha_1 + \alpha_2 + \dots + \alpha_\nu$ is equal to $\nu, \nu + 1$, and so on. (Naturally, other ordering procedures are also possible.)

* * *

¹ Words formerly used for this concept: *denumerable, enumerable*.

² This is remarkable because the rational numbers cover the real axis "densely everywhere" (i.e., between any two points of the line there are points representing rational numbers). We emphasize, however, that sequence (1.2-1) is not monotonic.

As is well-known, any (real or complex) number ζ is said to be *algebraic*, if there exists an integral polynomial

$$p_n(x) = a_0 + a_1x + \dots + a_nx^n \quad (n \geq 1, a_n > 0) \quad (1.2-3)$$

such that ζ is one of its roots. (For $n = 1$ we get the rational numbers.) The non-algebraic numbers are called *transcendental*.

1.2.3 The set of algebraic numbers is countable.

P r o o f. As an auxiliary notion, we introduce the so-called height of $p_n(x)$ which is defined by

$$A = |a_0| + \dots + |a_{n-1}| + a_n + n. \quad (1.2-4)$$

There are plainly only two integral polynomials $p_n(x)$ with $A = 2$; these are x and 2 . The admissible polynomials with $A = 3$ are x^2 , $2x$, $x + 1$, $x - 1$ and 3 ; and so on. In general, because all the coefficients must be integers, the number of polynomials $p_n(x)$ with a prescribed (positive integral) height A is necessarily finite and, for any concrete value of A , the corresponding polynomials $p_n(x)$ can be given easily, ordered by their degrees. Writing successively all the roots of the non-constant polynomials just mentioned, we get a sequence consisting of all algebraic numbers. This completes the proof.

* * *

1.23 Since the set of algebraic numbers is much wider than that of rational numbers, our last result suggests that maybe all the real numbers can be arranged in a single sequence, too. Unfortunately, this is not true!

1.2.4 The points of any interval of number line (finite or infinite) and thus, also all real numbers do not form a countable set.

P r o o f. It suffices to deal with the stretch $0 \leq x \leq 1$, because any open interval of the number line can be mapped in one-to-one way onto the interval $(0,1)$.¹

Each real number belonging to $(0,1]$ can be uniquely represented by an infinite decimal fraction.² Suppose that, in contradiction to the assertion,

¹ Clearly any finite stretch $a < x < b$ may be transferred bijectively onto $0 < x' < 1$ by the function $x' = (x-a)/(b-a)$ and its inverse. Similar connections can be established between the interval $0 < x' < 1$ and, in turn, the intervals $-\infty < x < \infty$, $a < x < \infty$ and $-\infty < x < a$ ($|a| < \infty$) by means of the transformations $x' = 1/(1 + 2^{-x})$, $x' = (x-a)(|x|+1)$ and $x' = (a-x)/(|x|+1)$, respectively.

² In the case of a finite decimal fraction, we reduce the last nonzero digit by 1 and replace all subsequent 0 digits by 9.

the set of numbers $0 < x \leq 1$ is countable. Then there must exist a sequence of decimals, say

$$\begin{cases} c_1 = 0 \cdot \gamma_{11}\gamma_{12}\gamma_{13} \dots \\ c_2 = 0 \cdot \gamma_{21}\gamma_{22}\gamma_{23} \dots \\ c_3 = 0 \cdot \gamma_{31}\gamma_{32}\gamma_{33} \dots \\ \dots\dots\dots \end{cases} \quad (\gamma_{ij} = 0, 1, \dots \text{ or } 9) \quad (1.2-5)$$

which contains all the numbers in $(0,1]$.

Consider now the digits along the main diagonal on the right and form from them the decimal fraction $0 \cdot \gamma_{11}\gamma_{22}\gamma_{33} \dots$. If each digit of the latter is replaced by another integer between 1 and 9, then the number $c = 0 \cdot \gamma_1\gamma_2 \dots$, thus obtained, belongs again to $(0,1]$, and differs from each element of sequence (1.2-5), because $\gamma_i \neq \gamma_{ii}$ ($i = 1, 2, \dots$).

This contradiction proves the validity of our statement.

* * *

Note that the fundamental idea of the last argument is also very significant, it is called the *second diagonal process of Cantor*.

It can be shown, similarly, that "intervals" and "spheres" of several dimensions and, in general, arbitrary so-called *continua* (cf. 2.42) are *not countable*. Furthermore, by comparing 1.2.3 and 1.2.4 we see that all the real numbers cannot be algebraic, i.e. *there exist transcendental numbers*, too. In this connection we mention that J. LIOUVILLE was the first mathematician to present (in 1844) concrete examples of transcendental numbers. The famous "transcendence proofs" of CH. HERMITE for e and of F. LINDEMANN for π were published a few decades later (in 1873 and 1882, respectively).

1.24 Many more simple examples (e.g. the set of points with rational coordinates in the plane or space is countable; the set of irrational or transcendental numbers is not countable, etc.) can be deduced from the following observation:

1.2.5. *Any subset of a countable set and the union of countably many countable sets is also countable.*

Any set selected from the elements of a sequence can be arranged in sequence, too. Furthermore, by writing the union of certain sequences in tabular form like (1.2-2), the latter can always be rewritten as a single sequence by means of the first diagonal process.

* * *

As justified above, we split up the totality of sets into two large classes: finite sets and those infinite sets which can be arranged in sequence are

called *countable sets*, while all the others are the *non-countable sets*. The empty set $\emptyset$ as a “degenerate” finite set belongs to the former class. For infinite, countable sets the expression *countably infinite* is also in use.

In spite of the fact that countable sets often occur in fields of application, it has to be stated that, in a certain sense, most sets are non-countable. Though our original problem cannot be solved by directly “counting” the elements, it will be seen that a reasonable extension of the method might yet be successful.

1.3 Comparison of cardinals and theorem of equivalence

1.31 Bearing in mind the previous discussion, we want to build our further investigations on cardinalities upon two conventions: (I) the *cardinal of any countably infinite set* is identical with the cardinal of $\{1, 2, 3, \dots\}$ which is usually denoted by $\aleph_0$ (aleph null); (II) the cardinal of all points on the number line — called briefly the *power of the continuum* and denoted by $\aleph$ (without any index) — is higher than $\aleph_0$.

It is sufficient to refer, on the one hand to the definition of a countable set, on the other hand, to the fact that the set of all real numbers has countably infinite proper subsets (e.g. the natural or the rational numbers), but itself is non-countable according to 1.2.4. The question now arises as to whether there exists any set the cardinal of which exceeds even $\aleph$, and at all, may we speak about sets of arbitrarily high cardinality?

For an answer, we have to define the notion of cardinality precisely and to give the rule for comparing cardinalities in the case of arbitrary sets. Therefore, the following definitions are introduced.

To each set S , we attach a “number” characterizing the quantity of its elements; it is called the *cardinality* (*cardinal number*, *power*) of S and bears the sign $\overline{S}$. To two sets S_1 and S_2 is given the same cardinality (in sign: $\overline{S_1} = \overline{S_2}$) if and only if there is a one-to-one correspondence between the elements of S_1 and S_2 . In this case S_1 and S_2 are said to be *equivalent* or *equipotent*, and this is expressed by writing $S_1 \sim S_2$.

If $S_1 \not\sim S_2$, but S_1 has a subset equivalent to S_2 , then we say that S_1 is of *higher cardinality* than S_2 or S_2 is of *lower cardinality* than S_1 . Notations: $\overline{S_1} > \overline{S_2}$ and $\overline{S_1} < \overline{S_2}$, respectively.

The definition of equivalence in set theory obviously furnishes the properties expected from any equivalence concept in mathematics, i.e. it is *reflexive* ($S \sim S$), *symmetrical* ($S_1 \sim S_2$ implies $S_2 \sim S_1$) and *transitive* (if $S_1 \sim S_2$ and $S_2 \sim S_3$, then $S_1 \sim S_3$).

At this point, however, we have to mention a new type of phenomenon: *two infinite sets may be equivalent to each other even if one of them is a proper subset of the other*. Thus, according to the definition, any two

countably infinite sets are equivalent (their common cardinality being $\aleph_0$), though it is easy to present two such countable sets where one properly contains the other; for example: the set of all integers and that of the even numbers. We can remember also the fact (cf. 1.2.4) that the real line can be mapped, by a linear transformation, onto any half-line or finite interval, so that these point sets are all equivalent to each other (with the common cardinality $\aleph$). We may add another very demonstrative geometric example: the equivalence of a perpendicular side of a right-angled triangle (as a point set) to the hypotenuse. In this case, the required one-to-one mapping can be immediately implemented by drawing perpendiculars onto the smaller side in question.

1.32 In the circumstances mentioned above, the following theorem is important:

1.3.1 (CANTOR-BERNSTEIN EQUIVALENCE THEOREM)¹ *Two sets, each of which is equivalent to a subset of the other, are also equivalent to each other.*

This theorem can be considered as a corollary of the following proposition which is also relevant in itself:

1.3.2 *Let S_0 be a given set and S_1, S_2 its subsets such that $S_1 \supseteq S_2$. Then the relation $S_2 \sim S_0$ implies $S_1 \sim S_0$.*

P r o o f. Suppose that S_2 is equivalent to S_0 , and so a bijective correspondence between the elements of S_2 and S_0 can be given.

Since $S_0 \supseteq S_1 \supseteq S_2$; this mapping attaches, at the same time, any element of S_1 to a single element of S_2 . Let us denote the set of the latter elements (which is a subset of S_2) by S_3 .

Therefore we have $S_1 \supseteq S_2 \supseteq S_3$ and also a bijective connection between S_1 and S_3 ; hence we see that the mapping just mentioned transfers S_2 onto a certain subset S_4 of S_3 , and so on.

Clearly, this procedure can be continued indefinitely and yields a *decreasing* sequence of sets for which $S_n \sim S_{n+2}$ ($n = 0, 1, 2, \dots$) and

$$S_n - S_{n+1} \sim S_{n+2} - S_{n+3} \quad (n = 0, 1, 2, \dots). \quad (1.3-1)$$

If we apply representation (1.1-8) to S_0 and to S_1 , we get

$$S_0 = T_0 \cup R, \quad S_1 = T_1 \cup R, \quad (1.3-2)$$

where

$$T_k = \bigcup_{n=k}^{\infty} (S_n - S_{n+1}) \quad \text{and} \quad R = \bigcap_{n=0}^{\infty} S_n. \quad (1.3-3)$$

¹ The theorem was discovered by G. CANTOR, but its first complete proof was given by F. BERNSTEIN in 1897. (See Borel [2], Fraenkel [1].)

Thus from (1.3-1) and (1.3-2), it follows that $T_0 \sim T_1$ and $S_0 \sim S_1$, respectively.

* * *

Using 1.3.2, 1.3.1, it can easily be verified:

Let us denote the sets in question by s, t and let $s \sim t_1, t \sim s_1$ where $s_1 \subseteq s$ and $t_1 \subseteq t$. Then there is a bijective connection between the elements of t and s_1 which transfers $t_1 \subseteq t$ onto a subset s_2 of s_1 . Thus $t_1 \sim s_2$ and since, by assumption, $s \sim t_1$, we have $s \sim s_2$ with $s \supseteq s_1 \supseteq s_2$. Now, by applying proposition 1.3.2 to s, s_1, s_2 , we obtain $s \sim s_1$ and by assumption also $s_1 \sim t$. Therefore $s \sim t$ and the theorem is proved.

The equivalence theorem makes it possible to show, that the comparison of cardinal numbers can be realized in all respects on the model of real numbers.

1.4. Sets of higher cardinality, transfinite cardinals

1.41 Looking for some kind of set the cardinality of which would be higher than $\aleph$, we might think first that higher dimensional numbers are connected with higher cardinalities, i.e. that the plane has a greater cardinality than the line, the space is of greater cardinality than the plane, etc. The following proposition shows that this conjecture is false.

1.4.1 *The set of points of an Euclidean space of arbitrary dimension has the cardinal number $\aleph$.*

P r o o f. It suffices obviously to show that those points $(x_1, x_2, \dots, x_\nu)$ of the ν -dimensional space for which $0 < x_n \leq 1$ ($n = 1, 2, \dots, \nu$), constitute a set equivalent to the interval $0 < x \leq 1$ of the real line. We have seen that the whole real line can be transferred, in one-to-one way, onto $(0, 1]$, and there is a completely similar bijective correspondence between the whole ν -dimensional space and its unit "ball" defined above. (Cf. 1.2.4.)

Now, the coordinates of any point of the latter can be written down in infinite decimal fractions of the form

$$x_n = 0.\xi_{1n} \xi_{2n} \dots \quad (n = 1, 2, \dots, \nu), \quad (1.4-1)$$

and this representation is unique if we transform the finite decimals to proper infinite ones in the well-known manner. (See the second foot-note to the proof of 1.2.4.)

The ν -dimensional point (1.4-1) is now mapped onto that point of $(0, 1]$ whose abscissa is defined by the decimal

$$x = 0.\xi_{11} \xi_{12} \dots \xi_{1\nu} \xi_{21} \xi_{22} \dots \xi_{2\nu} \dots \quad (1.4-2)$$

(the first ν digits are the *first* digits of $x_1, x_2, \dots, x_\nu$, respectively; the next ν digits are, in turn, the *second* digits of $x_1, x_2, \dots, x_\nu$; and so on). This is clearly a bijective mapping between the unit "ball"

$$\{(x_1, x_2, \dots, x_\nu) : 0 < x_n \leq 1 \quad (n = 1, 2, \dots, \nu)\} \quad (1.4-3)$$

of the ν -dimensional space and a subset of $(0,1]$.

Conversely, the interval $(0,1]$ can be mapped bijectively onto a subset of $(1.4-3)$ in a trivial way. For example:

$$\{(x_1, x_2, \dots, x_\nu) : 0 < x_1 \leq 1, \quad x_2 = x_3 = \dots = x_\nu = 0\} = (0,1].$$

Thus 1.3.1 yields at once the equivalence of (1.4-3) to $(0,1]$.¹

* * *

1.42 The simplest examples of sets whose cardinality really is higher than $\aleph$, are the following: the set ρ consisting of all subsets of the interval $[0,1]$, and the set U of all real functions defined in $[0,1]$. The second example partly reduces to the first, since the set of characteristic functions of the subsets of $[0,1]$ is equivalent both to ρ and the set of functions defined in $[0,1]$ whose range only contains the values 0 and 1. These functions, however, form a proper subset of U .

Going still further, we have the following proposition:

1.4.2 *All the subsets of a non-empty set S constitute a set, whose cardinality exceeds that of S .*

P r o o f. Denote by P the set consisting of all subsets of $S \neq \emptyset$ (including also S itself and $\emptyset$). From the definitions relating to the comparison of cardinalities, we have to prove, on the one hand, that $P \not\sim S$, and on the other hand, that a proper subset of P is equivalent to S . The latter condition is clearly satisfied by the collection of single elements of S (each of them being considered as a subset containing one element).

Assume now — contrary to the assertion — that the elements of S and P are connected bijectively and assign the elements of S to two classes. S_I consists of those elements each of which is contained in the corresponding subset of P , while S_{II} is the remaining part of S . (For instance, if E denotes the element of S which is mapped onto itself, then $E \in S_I$, but the element $\emptyset$ corresponding to $\emptyset$ belongs to S_{II} .) Since S_{II} is a subset of S , so that $S_{II} \in P$, there exists an element $\kappa \in S$ which is mapped onto S in P . We ask: in which class is this κ ?

¹ We have to emphasize that sets of different dimension cannot be mapped bijectively into one another, if the map is also required to be continuous. This follows from the "principle of the invariance of dimension" (L.E.J. BROUWER, 1911) which is one of the most important facts of classical topology. See also 2.13.

In the case $\kappa \in S_I$, the definition of S_I would imply $\kappa \in S_{II}$, which is a contradiction. However, if we suppose that $\kappa \in S_{II}$, then we would get, by the definition of S_{II} , the result $\kappa \notin S_{II}$ which contradicts this last assumption. Therefore, κ belongs neither to S_I nor to S_{II} ; but this is impossible, because $S_I \cup S_{II} = S$.

We see that our initial supposition is untrue, i.e. $P \not\sim S$.

* * *

1.43 We note that the set consisting of all subsets of a set S is called the *power-set* of S , and it is usually denoted by $P(S)$. For the corresponding cardinality we often write symbolically:

$$\overline{P(S)} = 2^{\overline{S}}. \quad (1.4-4)$$

This convention is based upon the elementary fact that for a finite set S with n elements, the number of the elements of $P(S)$ is just 2^n . Namely, $P(S)$ contains beside $\emptyset$ and S , the subsets consisting of $1, 2, \dots, n-1$ elements; the number of these subsets is, in turn, $\binom{n}{1}, \binom{n}{2}, \dots, \binom{n}{n-1}$, and by addition we obtain:

$$\binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n-1} = 2^n - 2.$$

Using the notation (1.4-4), the content of theorem 1.4.2 can briefly be expressed as follows:

$$2^{\overline{S}} > \overline{S}. \quad (1.4-5)$$

If we consider the case where S consists of all positive integers, so that $\overline{S} = \aleph_0$, then (1.4-5) can be sharpened to an equality which establishes — in symbolic form — a connection between $\aleph_0$ and $\aleph$. More precisely:

1.4.3 *The cardinality of all subsets of natural numbers is $\aleph$, i.e.*

$$\aleph = 2^{\aleph_0}. \quad (1.4-6)$$

P r o o f. Let $\{n_k\}$ be an arbitrary (finite or infinite) subset of the numbers $1, 2, 3, \dots$. To every n_k attach the value 1, and to all other elements of the sequence of positive integers the value 0. Then we get a sequence $j_1, j_2, j_3, \dots$ where each element is 0 or 1; and the dyadic fraction

$$0 \cdot j_1 j_2 j_3 \dots \quad (j_q = 0 \text{ or } 1) \quad (1.4-7)$$

represents a real number between 0 and 1 which is uniquely determined apart from the fractions containing only 0-s or only 1-s from a certain digit onwards.

Hence we see that the set of all subsets in question, after omitting an exceptional countable set, can be mapped bijectively onto the interval $(0,1)$. Since the union of a countable set and a set of continuum power has the cardinal number $\aleph$, (1.4-6) follows.

* * *

1.44 The above Theorem 1.4.2 enables us to present sets whose cardinality is as large as we like. Thus, the next step in building up the theory is: arranging *all* the possible infinite cardinals in an increasing sequence, to form, so to say, the "transfinite" continuation of the sequence of natural numbers and hence, establish a considerable extension of the principle of mathematical induction: the *transfinite induction*.

The realization of this program and, in connection with it, the further development of the notion of ordinal number (introduction of the concepts "order type" and "ordinal"), and the *well-ordering of sets* stand apart from the scope of our treatment. We must be content with mentioning that, amongst the infinite cardinals written in increasing order, the so-called *transfinite cardinals*¹, the first (least) one is $\aleph_0$ and the second one is $\aleph_1 \leq \aleph$. The so-called *generalized continuum hypothesis* asserts the validity of the relations

$$\aleph_{\alpha+1} = 2^{\aleph_{\alpha}} \quad (\alpha = 0, 1, 2, \dots). \quad (1.4-8)$$

As we see by (1.4-6), this famous conjecture of Cantor is equivalent in the case $\alpha = 0$ to the connection $\aleph_1 = \aleph$, i.e. to the assertion that in the sequence of transfinite cardinals the cardinal of natural numbers is immediately followed by the continuum power.²

¹ In this regard, the cardinals of finite sets, i.e. the non-negative integers, are called *finite cardinals*, too.

² It was shown by P.J. COHEN in 1962 that (1.4-8) cannot be proved on the basis of the usual axiomatic foundation of set theory, which is due to ZERMELO and FRAENKEL, completed by the so-called axiom of choice. (Cf.e.g. Fraenkel [1], Cohen [2].) This result is comparable, to a certain extent, with the fact that Euclid's parallel postulate is independent of the axioms of absolute geometry, i.e. that non-Euclidean geometries are also possible. The latter were discovered — as it is well-known — by J. BOLYAI and N.I. LOBATCHEVSKI (after some unpublished results of C.F. GAUSS) in the first decades of the 19th century.

1.5 Important set structures: Boolean and sigma-rings

1.51 A system $\mathcal{Z}$ of sets is said to be *closed with respect to a set operation*, if, by carrying out the latter on arbitrary sets in $\mathcal{Z}$, we always obtain elements of $\mathcal{Z}$. (In other words: the set operation in question "does not go out" of $\mathcal{Z}$.) Therefore, every such system of sets is at the same time a *structure* in the modern algebraic sense.

From the point of view of the applications, the following set structures, associated with two operations, are of primary significance: (I) *Boolean ring* or *simple field of sets*; (II) σ -ring or *Borel's field of sets*. (The terminology in the literature has not been unified up to now.)

A non-empty family $\mathcal{R}$ of subsets of a basic set X is called a *Boolean ring*, if it is closed with respect to union and subtraction operations, i.e. for any $S_1 \in \mathcal{R}$, $S_2 \in \mathcal{R}$ we have $S_1 \cup S_2 \in \mathcal{R}$ and $S_1 - S_2 \in \mathcal{R}$. The set $\emptyset$ obviously belongs to every Boolean ring because $S \in \mathcal{R}$ implies $\emptyset = S - S \in \mathcal{R}$. Besides, $\mathcal{R}$ contains by (1.1-6) the symmetric difference and, by (1.1-7), the intersection of any two elements. Finally, it follows by induction that both the union and the intersection operations can be carried out finitely many times, with results in $\mathcal{R}$.

It is useful to give a particular name to those Boolean rings which are also closed under the union operation applied a *countably infinite number of times*. These are the so-called σ -rings, which will be denoted by $\mathcal{S}$. The definition implies by (1.1-5) that any σ -ring also contains the intersection of a countably infinite number of its elements.

Remarks: 1. According to the above, the Boolean ring and sigma-ring can also be defined in such a way that, instead of the operations of union and subtraction, closure with respect to the intersection operation and the forming of symmetric differences is required.¹ 2. In that frequent special case when the Boolean or σ -ring in question contains the basic set X itself, too, the denominations *Boolean algebra* and σ -algebra are used. (Alternative definitions: we mean by Boolean algebra and sigma-algebra any non-empty system of sets which is closed with respect to the complementation and the union operations. The latter is applied finitely many times in the Boolean case and a countably infinite number of times in the sigma case.) The usual notations for the last-mentioned set structures are: $\tilde{\mathcal{R}}$ and $\tilde{\mathcal{S}}$.

Apart from the set consisting of $\emptyset$ only, the simplest *example* of a Boolean ring is the system of all subsets of a set X , i.e., the *power-set* of X [see (1.4-3)]. This is plainly a σ -ring and a σ -algebra at the same time.

¹ If the operation of forming symmetric differences is considered as an abstract "addition" and the intersection operation as an abstract "multiplication", then any Boolean ring of sets becomes a ring in the sense used in abstract algebra. All this explains the preference for the above terminology.

But if X is infinite and we consider only the system of all finite subsets of X , then we get such a Boolean ring which is neither a σ -ring, nor a Boolean algebra.

1.52 In case of the last-mentioned example, the following mode of expression presents itself: "Boolean ring generated by the finite subsets of X ". This suggests that a general theorem can be formulated:

1.5.1 For any non-empty set system $\mathcal{U}$ consisting of certain subsets of a basic set X , there exists a smallest one among the Boolean rings (or σ -rings) containing $\mathcal{U}$. It is called the Boolean ring (resp. σ -ring) generated by $\mathcal{U}$, and is denoted by $\mathcal{R}(\mathcal{U})$ and $\mathcal{S}(\mathcal{U})$, respectively.

P r o o f. Let us apply to the sets contained in $\mathcal{U}$, the union operation and subtraction in all possible ways finitely many times. The sets thus obtained plainly constitute such a Boolean ring $\mathcal{R}$ which contains $\mathcal{U}$; and in any other Boolean ring $\mathcal{R}^*$ to which $\mathcal{U}$ also belongs, all the elements of $\mathcal{R}$ must occur, so that $\mathcal{R}^* \supseteq \mathcal{R}$. Accordingly, $\mathcal{R}$ is the smallest Boolean ring containing $\mathcal{U}$, i.e. $\mathcal{R} = \mathcal{R}(\mathcal{U})$.

$\mathcal{S}(\mathcal{U})$ can be generated similarly, if we apply to the element of $\mathcal{U}$ the operation of subtraction and, a countably infinite number of times, the union operation.

* * *

It is clear that $\mathcal{R}(\mathcal{U})$ and $\mathcal{S}(\mathcal{U})$ can also be regarded as the common part of all Boolean rings and σ -rings, respectively, which contain the system $\mathcal{U}$. Furthermore we have:

$$\mathcal{U} \subseteq \mathcal{R}(\mathcal{U}) \subseteq \mathcal{S}(\mathcal{U}). \quad (1.5-1)$$

We may add that 1.5.1 still remains valid if the words Boolean ring and σ -ring are replaced by Boolean and σ -algebra, respectively. To verify this we have simply to use complements instead of differences in the above proof.

Two elementary examples connected with the concepts just introduced:

I. If $\mathcal{U}$ consists of all finite intervals of the line, including $\emptyset$ and the sets containing a single point, then $\mathcal{R}(\mathcal{U})$ contains all interval systems obtained as unions of finitely many elements of $\mathcal{U}$. (Therefore, these interval systems certainly constitute a Boolean ring.) — II. If $\mathcal{U}$ has the same meaning, the σ -ring generated by it contains the whole line, too, i.e., it is a σ -algebra. The elements of this structure are called the *Borel sets* of the line.

The notions of this section play an important role in measure theory and probability theory. (See 5.13, 6.14, 8.11–8.15, 8.31, 8.41–8.43, 8.53.) Let

us also stress that a significant branch of modern algebra, namely *lattice theory*, has developed in close connection with Boolean algebras during the last decades.¹

1.6 Simple applications in elementary analysis

1.61 We want to discuss some additional material showing that the fundamental notions of set theory can also be useful in the elementary theory of real functions.

Though most discontinuous functions occurring in applications have a finite number of discontinuities, it is easy to give examples of functions which are bounded, but *not continuous on a set of continuum power*. We mention e.g. the characteristic function of the set of rational numbers, the so-called *function of Dirichlet*:

$$\chi(x) = \begin{cases} 1 & \text{for rational } x, \\ 0 & \text{for irrational } x \end{cases} \quad (1.6-1)$$

which is plainly nowhere continuous.² (Cf. 1.2.4.)

In general, points of discontinuities are usually limited into two classes: we speak about a *discontinuity of the first kind* at a point where both of the unilateral limits of the function exist³; any other discontinuity is said to be *of the second kind*. Thus all discontinuities of the function $[x]$ ("integral part of x ") belong to the first class, and all of $\chi(x)$ to the second class.

Our next result concerns the cardinality of the discontinuities of the first kind on an interval.

1.6.1 *If the one-sided limits of a function $f(x)$ exist at every point of an interval, then the totality of its points of discontinuity there is countable.*

P r o o f. Denote by S the set of all points of discontinuities of $f(x)$ in a finite interval $[a, b]$. We shall show first: the subset S_δ of those points $\xi \in S$, for which both $|f(\xi) - f(\xi + 0)|$ and $|f(\xi) - f(\xi - 0)|$ exceed any fixed positive number δ , is necessarily finite. For brevity we refer below to this assertion by the symbol $(*)$.

¹ We mean by *lattice* an algebraic structure in which two operations are defined with properties corresponding to those of the union and intersection operation for sets. (Cf. 1.12.)

² There is also an interesting analytic representation of this function:

$$\chi(x) = \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} [\cos(m! \pi x)]^{2n}.$$

³ If the one-sided limits at a point are equal, but differ from the corresponding function value, then we have a so-called *removable discontinuity*.

Suppose the contrary, i.e., that the subset S_δ is infinite. Then the boundedness of S_δ implies that the corresponding set of discontinuities has in $[a, b]$ at least one accumulation point ξ_0 (by the Bolzano-Weierstrass theorem). Therefore, there exists a sequence of points $\xi_k \in S_\delta$ ($k = 1, 2, \dots$) tending from one (e.g. left-hand) side to ξ_0 , such that $|f(\xi_m) - f(\xi_n)| > \delta/2$ for m, n sufficiently large. On the other hand, we know that $f(\xi_0 \pm 0)$ exists, and hence we conclude that $|f(\xi_m) - f(\xi_n)| \rightarrow 0$ as $m, n \rightarrow \infty$. This contradiction verifies our assertion (*).

Let now $v(\xi)$ denote the smaller of the deviations $|f(\xi) - f(\xi \pm 0)|$, and let us consider the following subsets of S :

$$S_1 = \{\xi : v(\xi) > 1\}, \quad S_n = \left\{ \xi : \frac{1}{n+1} < v(\xi) \leq \frac{1}{n} \right\} \quad (n = 2, 3, \dots).$$

Obviously $S = \bigcup_{n=1}^{\infty} S_n$ and, since all the S_n are finite according to (*), we deduce from 1.2.5, that S is countable.

In the case when the domain of $f(x)$ is infinite, we express it as a countable union of finite intervals and, hence, reduce it to the previous case on the basis of 1.2.5.

* * *

1.62 Our second observation relates to the concept of differentiability.

Let us call the point a a *break point* of a function $f(x)$, if the unilateral derivatives

$$f'_+(a) = \lim_{x \rightarrow a+0} \frac{f(x) - f(a)}{x - a}, \quad f'_-(a) = \lim_{x \rightarrow a-0} \frac{f(x) - f(a)}{x - a} \quad (1.6-2)$$

exist simultaneously, but $f'_+(a) \neq f'_-(a)$. [Geometrically: we draw on the graph of $f(x)$ a right half-tangent at the point $(a, f(a+0))$, and a left half-tangent at the point $(a, f(a-0))$, having different directions.] Note that any break point of f must either be a point of continuity or a point of discontinuity of the first kind.

1.6.2 *The break points of any function $f(x)$ constitute a countable set.*

P r o o f. By virtue of 1.2.1, all the rational numbers can be arranged in a sequence; let ρ_n denote the n -th element of such a sequence. We shall use repeatedly the fact that the number line is densely covered everywhere by the rationals.

Consider the set $T = \{x : f'_-(x) < f'_+(x)\}$, and let a be a fixed element of it. By the definition of $\{\rho_n\}$ there exists a lowest index λ such that

$f'_-(a) < \rho_\lambda < f'_+(a)$. Furthermore, by (1.6-2) we can find a largest element ρ_μ in $\{\rho_n\}$, for which

$$\frac{f(x) - f(a)}{x - a} < \rho_\lambda \quad (\rho_\mu < x < a),$$

and another smallest element ρ_ν , for which

$$\frac{f(x) - f(a)}{x - a} > \rho_\lambda \quad (a < x < \rho_\nu).$$

Connecting these two inequalities, we get with $a \in (\rho_\mu, \rho_\nu)$ the condition:

$$f(x) - f(a) > \rho_\lambda(x - a) \quad (\rho_\mu < x < \rho_\nu, x \neq a). \quad (1.6-3)$$

Thus we have shown, that to any $a \in T$ there corresponds a triplet (λ, μ, ν) of indices such that (1.6-3) holds. It is easy to see that this correspondence is unique, i.e., that the same triplet cannot belong to any other point a_1 of T . In fact, assuming the contrary, we can substitute in (1.6-3) first $x = a_1$, and then $x = a$ with a replaced by a_1 , and obtain the impossible result:

$$f(a_1) - f(a) > \rho_\lambda(a_1 - a), \quad f(a) - f(a_1) > \rho_\lambda(a - a_1) \quad (a \neq a_1).$$

Because the set of all index triplets (λ, μ, ν) is countable by 1.2.2, the above set T and (on the basis of a similar argument) the set $\{x : f'_+(x) < f'_-(x)\}$ as well as their union, i.e. $\{x : f'_+(x) \neq f'_-(x)\}$, have the same property.

* * *

1.63 We give another elementary example of a non-countable set:

1.6.3 *The cardinal of the set of functions continuous in an interval is $\aleph$.*¹

P r o o f. Denote the basic interval in question by (a, b) and the set of functions continuous in (a, b) by F . Since F contains all the functions $y = C$ ($a < x < b$) where C is an arbitrary real number, it is plain that

$$\overline{F} \geq \aleph. \quad (1.6-4)$$

¹ We call the reader's attention to the fact that the *totality* of real functions defined on a given interval has a cardinality higher than $\aleph$. (See 1.4.2.)

We want also to show that

$$\overline{F} \leq \aleph. \quad (1.6-5)$$

To that end, we note, first of all, that F is equivalent to the set F_r of all sequences

$$f(r_1), f(r_2), \dots$$

where $\{r_n\}$ runs over the totality of rational numbers in (a, b) (cf. 1.2.1). Namely, the range of any continuous function is evidently determined by its values at the rational points [lying densely everywhere in (a, b)].

Consider now the set Y consisting of sequences of the form

$$y_1, y_2, \dots,$$

the elements of which run over all real numbers, independently of one another. If we transfer y_1 by a linear transformation into the interval $(0, 1)$, y_2 into $(1, 2)$ and so on, then it becomes clear that the cardinal of Y cannot exceed that of $(0, \infty)$, i.e., the continuum power.

Thus $\overline{F_r} \leq \overline{Y}$ implies $\overline{F_r} \leq \aleph$, which by $F \sim F_r$ is equivalent to the inequality (1.6-5).

Hence and from (1.6-4) we get finally

$$\overline{F} = \aleph, \quad (1.6-6)$$

as asserted above.

* * *

The set theoretic attitude — as is well-known — has been adopted not only in various branches of *analysis*, but also in *abstract algebra* and in other parts of mathematics. Since the turn of the century it has had a considerable and accelerating effect on the development of *mathematical logic* (i.e., the theory of judgement, logical functional calculus), and during the last decades it has been much entwined with the *theory of probability and statistics* as well as with the mathematical foundations of modern *computing technics*.

POINT SETS IN EUCLIDEAN SPACES

2.1 Distance and neighbourhood in space E_ν

2.11 In the elementary theory of functions, not only point sets on the line ("linear sets"), but those on the plane and in three-space are needed. Besides, it is indispensable to the uniformity and simplicity of discussing functions of several variables that we also consider *spaces of arbitrary dimensions*. As a matter of course, the foundations of modern analysis presuppose the further development of the latter theory by means of the general set theoretical expedients reviewed in the previous chapter. Therefore, we shall deal in the sequel with the systematic treatment of abstract point sets of arbitrary dimension, for the sake of a self-contained treatment premising some well-known basic knowledge.

Let ν be a fixed natural number. We denote by E_ν the *Euclidean space of ν dimensions*, i.e., the totality of ordered ν -tuples of real numbers, satisfying certain conditions below. Any such ν -tuple $(x_1, x_2, \dots, x_\nu)$ is called a *point* of E_ν , usually denoted by Roman capital letters, and x_r is the r -th *co-ordinate* of the point ($r = 1, 2, \dots, \nu$). The point whose all coordinates are 0, is the so-called *zero-point* (origin) and is denoted by O . Two points: $P(x_1, x_2, \dots, x_\nu)$ and $Q(y_1, y_2, \dots, y_\nu)$ are identical (equal) if $x_r = y_r$ for every r . The *distance* between P and Q is defined by

$$\overline{PQ} = [(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_\nu - y_\nu)^2]^{1/2}. \quad (2.1-1)$$

The motivation for (2.1-1) is that, when $\nu = 1, 2, 3$ (i.e., in case of the line, plane and three-space) it reduces to the well-known distance formula of analytic geometry. Furthermore, the requirement that any definition of distance possesses the following properties [obvious for $\nu \leq 3$, and referred later by the symbol (*):

- (*) 1. $\overline{PQ} = 0$ holds if and only if $P = Q$; 2. $\overline{PQ} = \overline{QP}$ for all P and Q ;
3. for any triplet P, Q, R the so-called *triangle inequality*

$$\overline{PQ} + \overline{QR} \geq \overline{PR} \quad (2.1-2)$$

is fulfilled.

We note that (2.1-2) is a corollary of *Cauchy's inequality*

$$\left(\sum_{r=1}^{\nu} u_r v_r \right)^2 \leq \left(\sum_{r=1}^{\nu} u_r^2 \right) \left(\sum_{r=1}^{\nu} v_r^2 \right), \quad (2.1-3)$$

valid for arbitrary real numbers u_r, v_r ; and (2.1-3) follows immediately from the so-called *Lagrange identity*

$$\frac{1}{2} \sum_{r,s=1}^{\nu} (u_r v_s - u_s v_r)^2 = \left(\sum_{r=1}^{\nu} u_r^2 \right) \left(\sum_{r=1}^{\nu} v_r^2 \right) - \left(\sum_{r=1}^{\nu} u_r v_r \right)^2 \quad (2.1-4)$$

by the non-negativity of the sum of squares on the left.

We mean by an *interval* in E_ν the set of points P with coordinates x_r in the linear intervals i_r ($r = 1, 2, \dots, \nu$).¹ A ν -dimensional interval I is closed or open when all the corresponding intervals $i_1, i_2, \dots, i_\nu$ are closed or open, respectively. I is said to be finite if all of its i_r -s are finite. (For instance, the square $|x_1| \leq 1, |x_2| \leq 1$ is a 2-dimensional closed interval, and the cube $0 < x_r < 1$ [$r = 1, 2, 3$] is a 3-dimensional open interval.) Any interval in E_ν may plainly also be considered as a Cartesian product of ν factors. (See 1.12.)

2.12 The ϵ -neighbourhood of a point $P_0 \in E_\nu$ is the set of those points of E_ν for which $\overline{P_0 P} < \epsilon$. In the case of E_1 it is a linear interval symmetric about P_0 , while for E_2 we have a circular disk and for E_3 a sphere, both with centre P_0 and radius ϵ . Thus the denomination "open sphere" is also in use for any ν . (The condition $\overline{P_0 P} \leq \epsilon$ determines a so-called closed sphere or ball.) The size of $\epsilon > 0$ is often irrelevant and in these instances we only mention the centre P_0 . In particular if a point P_0 of a set $S \subset E_\nu$ is considered, then the ϵ -neighbourhood of P_0 relating to S is the set of points of S whose distance from P_0 is less than ϵ .

Using these, the points of E_ν can now be classified with respect to any set $S \subset E_\nu$ in the usual manner. A point P is an *interior* (inner) *point* of S , if P has a neighbourhood lying entirely in S ; P is an *exterior point* of S , if it has a neighbourhood which does not contain any point of S ; and, lastly, P is a *boundary point* of S , if neither of these two cases hold (i.e. every neighbourhood of P contains a point which belongs to and another one which does not belong to S). According to these definitions, interior points always belong to S and exterior points never belong to S . In case of a boundary point, both possibilities may occur. The totality of boundary points is often briefly called the *boundary* of S .

A set $S \subset E_\nu$ is said to be *bounded*, if it can be included in a ν -dimensional finite interval or in any finite neighbourhood of 0. Among sets consisting of real numbers, it is useful to distinguish *boundedness from above* and *boundedness from below* in the usual sense, and hence, to introduce the notions of *unilateral bounds* as well as the concepts of *least upper bound* (*supremum*) and *greatest lower bound* (*infimum*). The latter will be denoted, in the case of a linear set S , by $\sup S$ and $\inf S$, respectively. Their existence

¹ Here i_r denotes a finite or infinite segment of the number line, i.e., the corresponding set of real numbers.

is assured when S is bounded from above and from below, respectively; otherwise we have $\sup S = +\infty$ or $\inf S = -\infty$.

2.13 By means of the notions of “distance” and “neighbourhood”, the concepts of *curve* and *surface* can exactly be defined. Furthermore, the foundations of theory of curves and surfaces can be extended to spaces of arbitrary dimensions.

Let us call a mapping of a set $S \subseteq E_\nu$ into a set $T \subseteq E_\nu$ *neighbourhood-preserving*¹ if, for any pair of corresponding points $P \in S$, $Q \in T$, we can attach to every ϵ -neighbourhood of Q in T a $\delta [= \delta(\epsilon)]$ -neighbourhood of P in S such that the latter neighbourhood is mapped into the former one. As we know, *topology* in a narrower sense (“analysis situs”) deals with bijective and neighbourhood-preserving mappings between geometric configurations. Hence, such a correspondence is also called a *topological map*, and the sets transferable by such a map into one another are said to be topologically equivalent or simply *homeomorphic*. For example, circular disk is homeomorphic to the interior of a rectangle or to any polygonal domain, but not homeomorphic to the interior of an annulus. The interval $(0,1)$ is homeomorphic to any circular or parabolic arc, but not homeomorphic to the square or cube. (Cf. 1.4.1.)

Now, it is easy to transfer the notions of *simple continuous arc* (*Jordan curve*) and *simple closed curve* (closed Jordan curve) to any space E_ν ($\nu \geq 2$). We have to consider the ν -dimensional point sets which are homeomorphic to a line segment and to a circle, respectively. These conditions are realizable analytically by such a *parametric representation* $x_r = x_r(t)$ ($\alpha \leq t \leq \beta$; $r = 1, 2, \dots, \nu$), where every $x_r(t)$ is a continuous function of the parameter t , and different t -values are always mapped into different points $(x_1, x_2, \dots, x_\nu)$, except for the values $t = \alpha$, $t = \beta$ in case of a closed curve. The representation in question yields, simultaneously, an *orientation* on the curve according to increasing values of t . In the case $\nu = 2$ we have a fundamental theorem which is seemingly evident but in reality is deep and cannot be discussed here in detail:

2.1.1 (JORDAN CURVE THEOREM) *Any simple closed curve $\mathcal{C}$ of two dimensions divides the plane into two disjoint point sets of which it is the common boundary. One of these sets is bounded and is called the interior of $\mathcal{C}$ and the other set is the exterior of the curve.*²

¹ In what follows, this notion will be applied exclusively to sets having the property that any two points are contained in a continuum. (Cf. 2.42.) In such cases, the term “*continuous mapping*” is more usual, because it is a further development of the concept of a continuous function.

² Each of these sets is also a *region* in the sense known from elementary analysis (see 2.42). Cf. C. Jordan [2].

We define in E_ν ($\nu \geq 3$), in a similar manner the notions of *simple continuous surface* and *simple closed surface* as the “topological images” of a square and the surface of a sphere, respectively. These surfaces are realized analytically by *biparametric representations* of Gauss type; and, if we replace the word “curve” with the word “surface”, the above theorem 2.1.1 also applies to the case $\nu = 3$. A considerable generalization of the theory of curves and surfaces yields the so-called Riemann–Poincaré theory of higher dimensional manifolds.

In the sequel everywhere — if no special restrictions are made — our statements concerning point sets are to be understood to apply for *arbitrary dimension number* ν .

2.2 Convergence and clustering point of a sequence of points

2.21 The transfer of the concepts of convergence and limit to Euclidean spaces is of decisive significance.

The most elementary notion about this topic is the following: a sequence of ν -dimensional points $P_1, P_2, \dots$ is said to be *convergent*, if there is a point $P \in E_\nu$ for which $\overline{P_n P} \rightarrow 0$ as $n \rightarrow \infty$. (This condition is plainly equivalent to the following: 1. any arbitrarily small neighbourhood of P should contain all but a finite number of elements of the sequence; or 2. each coordinate of P_n should tend to the corresponding coordinate of P as $n \rightarrow \infty$.) P is the *limit point* or, in short, the *limit* of the sequence. Non-convergent sequences are called *divergent*. Note that these statements: “the limit of $\{P_n\}$ is P ” and “ $\{P_n\}$ tends (converges) to P ” have the same meaning. The corresponding notations are:

$$\lim_{n \rightarrow \infty} P_n = P \quad \text{and} \quad P_n = P \quad (n \rightarrow \infty), \quad (2.2-1)$$

respectively.

The classical convergence principle for number sequences can be extended directly — by reduction to coordinate sequences — to higher dimensional cases:

2.2.1 (CAUCHY’S INNER CONVERGENCE CRITERION) A sequence $P_1, P_2, \dots$ of points is convergent if and only if to any given positive number ϵ , however small, we can find an index $n_0 = n_0(\epsilon)$ such that for $m \geq n_0$, $n \geq n_0$ the inequality $\overline{P_m P_n} < \epsilon$ holds.

Hence we see that the convergence of $\{P_n\}$ depends only on elements the indices of which are *sufficiently large*. Thus, if we omit a finite number of elements in a sequence of points, or rearrange it, no matter how, its convergence behaviour and limit remain unchanged. Furthermore, it is clear that any infinite subsequence of a convergent sequence of points is also convergent and tends to the same limit. We add to this last observation a fundamental lemma:

2.2.2 In every bounded infinite sequence of points, it is possible to find a convergent subsequence.

P r o o f.

Let $P_n = (x_{1n}, x_{2n}, \dots, x_{\nu n})$ be the n -th element of the given sequence and suppose that $a_1 \leq x_{1n} \leq b_1, \dots, a_\nu \leq x_{\nu n} \leq b_\nu$ ($n = 1, 2, \dots$), i.e., that $\{P_n\}$ is contained in the ν -dimensional interval I_0 corresponding to $[a_1, b_1], \dots, [a_\nu, b_\nu]$ where all linear intervals are finite.

If there is in $\{P_n\}$ an element which occurs infinitely many times, then we have nothing to prove, because we have got a trivially convergent subsequence. In the opposite case, we can find among the elements of $\{P_n\}$ an infinite number of distinct points. Let us denote the selected points, after arranging them in increasing order of the indices, by $Q_1, Q_2, \dots$. It is clear that $\{Q_n\} \subset I_0$. Now consider any subset of I_0 consisting of points whose r -th coordinate belongs to one of the two subintervals obtained by bisecting $[a_r, b_r]$ ($r = 1, 2, \dots, \nu$). The total number of these subsets is 2^ν , and among them there is certainly one, say I_1 , in which infinitely many elements of $\{Q_n\}$ are contained. Let us repeat, in case of I_1 , the above procedure of "halving". This produces a subset $I_2 \subset I_1$, which again contains infinitely many points of $\{Q_n\}$, and so on.

In this manner we obtain a strictly decreasing sequence of intervals: $I_0, I_1, I_2, \dots$. Starting with a point Q_{m_0} of I_0 , the construction allows a choice from I_1 , of a point $Q_{m_1} \neq Q_{m_0}$. Thereafter we can choose from I_2 a point Q_{m_2} which differs both from Q_{m_0} and Q_{m_1} , etc. Since the corresponding coordinates of the points just mentioned lie in a set of nested (linear) intervals, which have one single common point, the sequence $Q_{m_0}, Q_{m_1}, \dots$ is convergent. But this is, apart from the order of elements, a subsequence of $\{P_n\}$.

* * *

2.22 The preceding considerations suggest the introduction of a new notion. We say that a sequence of points $\{P_n\} \subset E_\nu$ *clusters* at a point $P \in E_\nu$, or P is a *clustering point*¹ of $\{P_n\}$, if it has a subsequence converging to P . According to 2.2.2, each bounded sequence has at least one clustering point, but it is easy to give instances of (bounded or unbounded) sequences clustering at several points. If $\nu = 1$, i.e., in case of a sequence of real numbers, we say *cluster value* instead of clustering point. Then we can find, among all the cluster values, maximal and minimal values in wider sense², i.e., $\pm\infty$ being also permitted, which are called *upper* and *lower limits* (limit

¹ Sometimes the word *sublimit* is used.

² It is clear that the supremum and infimum of the set of cluster values are themselves also cluster values of the sequence.

superior and limit inferior) of the sequence, respectively. The most usual symbols of the latter are: $\overline{\lim}$ and $\underline{\lim}$, respectively.

For example, the number sequence $1, 0, 1, 0, \dots$ has two cluster values, namely 1 and 0; the former is an upper and the latter a lower limit. As for the sequence (1.2-1) of positive rational numbers, this clusters at any point of the interval $[0, \infty)$ so that 0 is its lower limit and $+\infty$ its upper limit. Finally, in the case of $1, \frac{1}{2}, \frac{1}{3}, \dots$ we have $\overline{\lim}_{n \rightarrow \infty} \frac{1}{n} = \underline{\lim}_{n \rightarrow \infty} \frac{1}{n} = 0$ which is also the limit of our sequence.

2.23 The connection of the notions of convergence and clustering point is simple:

2.2.3 *A sequence of points is convergent if and only if it is bounded and has a unique clustering point.*

P r o o f. 1° Suppose that a sequence $\{P_n\}$ converges to a point P . Then there are only finitely many elements of $\{P_n\}$ outside any given neighbourhood of P , so that a suitably large interval contains the entire sequence, i.e., it is bounded. Furthermore we know that every subsequence of $\{P_n\}$ also tends to P , and thus P is the only clustering point of $\{P_n\}$.

2° Conversely, let $\{P_n\}$ be a sequence, all of whose elements lie in a finite interval I with a single clustering point P in I ; let $\tau_\epsilon \subset I$ denote an ϵ -neighbourhood of P . Then it follows from the definition of P , that there is a subsequence of $\{P_n\}$ tending to this point; and therefore, τ_ϵ must contain infinitely many elements of $\{P_n\}$. The complement $I - \tau_\epsilon$, however, contains only a finite number of points of the sequence, because otherwise the above lemma 2.2.2 would imply the existence of a subsequence of $\{P_n\}$ converging to a point different from P , and, thereby, also the existence of a second clustering point of $\{P_n\}$. With number $\epsilon > 0$ being arbitrarily small, we have shown that $\{P_n\}$ is convergent with the limit P .

* * *

According to the second part of the reasoning: if a bounded sequence of points has only one clustering point, so this is also a limit point; especially in the case $\nu = 1$, the existence of one single clustering point is equivalent to the statement that the upper and lower limit of the sequence are equal. If we allow the possible limits of a sequence $\{a_n\}$ of real numbers to include $\pm\infty$ (in the well-known sense), then a necessary and sufficient condition of the existence of $\lim_{n \rightarrow \infty} a_n$ — whether for a bounded or unbounded sequence — is the following:

$$\overline{\lim}_{n \rightarrow \infty} a_n = \underline{\lim}_{n \rightarrow \infty} a_n. \quad (2.2-2)$$

2.3 Notions and theorems concerning derivate sets

2.31 Let us now deal with arbitrary subsets of Euclidean spaces, which will be called briefly *point sets*. Considering the elements of a point set, any point may be counted only once. Thus point sets always consist of distinct elements, in contrast to sequences of points. However, each countably infinite point set may also be represented as a point sequence the elements of which are different from one another.

Like the definition of a clustering point, any point $P \in E_n$ is called an *accumulation point*¹ of a set $S \subseteq E_n$, if S contains a countably infinite subset whose limit is P . It is easy to see that this condition is equivalent to the following. 1. Each ϵ -neighbourhood of P contains infinitely many points of S . 2. In any neighbourhood of P , however small, we can find at least one point of S which differs from P . It must be emphasized that the accumulation point of a set can also be a point which does not belong to the set in question. (For example, 0 is the only accumulation point of the set $\{1, \frac{1}{2}, \frac{1}{3}, \dots\}$ and one of the accumulation points of the interval $0 < x < \infty$.) If a point $Q \in S$ is not an accumulation point of S , then it is called an *isolated point*; in other words, this means that there is a (sufficiently small) neighbourhood of Q which only contains the point Q .

The extension of 2.2.2 to point sets is the following:

2.3.1 (BOLZANO-WEIERSTRASS THEOREM) *Every bounded, infinite point set has at least one accumulation point.*

In fact, if a bounded set S consists of infinitely many points, we can apply to it the "halving" procedure presented in the proof of 2.2.2. Thus we can select from S a convergent sequence of distinct points, the limit of which is, according to the above definition, also an accumulation point of S .

* * *

Let us mention here, that a set is said to be *compact*, if every infinite subset of it contains at least one accumulation point. In E_n , by 2.3.1, all the bounded point sets — and, as it is easy to see, only these — possess the property of compactness.

2.32 We also need to sharpen the concept of an accumulation point. — A point $P \in E_n$ is called a *condensation point* of a point set $S \in E_n$, if every neighbourhood of P contains a non-countable infinity of points of S .

¹ In spite of the similarity of this notion to the "clustering point", the use of the same word might be confusing. For example, the sequence $\{1, 1, \dots\}$ clusters at 1 (even converges to this value), but $\{1\}$ as a set consisting of one point has no accumulation point in the above sense.

Such a point clearly cannot exist in the case of a countable set, so that we have:

2.3.2 (LINDELÖF'S THEOREM) *In any non-countable point set there is a condensation point.*¹

P r o o f. If the relevant set S is not bounded, then it may always be considered as the union of a countably infinite number of bounded sets and, among these, at least one, say S_0 , must also be non-countable by 1.2.5. In the case of S_0 , however, we may argue in the same manner as we did for 2.3.1, and it is enough to mention that if a non-countable infinity of points of S_0 is contained in a finite interval I , and the latter is divided into finitely many subintervals, then at least one part will "inherit" the property in question.

2.33 The totality of accumulation points of a set S — which characterizes the set in certain respects — will be called the *derived set* (briefly *derivate*) of S and denoted by S' . Examples:

In the case of $S = \{1, 2, \dots, n\}$ we have $S' = \emptyset$ (there is no accumulation point, the elements are isolated points). If S is the union of the sequences $\{\frac{1}{n}\}$ and $\{\frac{n+1}{n}\}$, then S' consists of the two accumulation points, namely of 0 and 1. For $S = (0, 1)$ we have $S' = [0, 1]$; and for the circle, $S = \{(x, y) : x^2 + y^2 = 1\}$, it is $S' = S$. (In the last two examples, every accumulation point is a condensation point, too.)

According to the set relations between S and S' , we distinguish three types of point sets: 1. S is said to be *closed*, if it contains all its accumulation points, i.e. $S \supseteq S'$; 2. S is *dense in itself*, if each point in it is an accumulation point of other elements of the set², i.e., $S \subseteq S'$; 3. S is *perfect*, if it has both of the properties just mentioned, i.e., $S = S'$.

Every finite point set and any infinite point set with no accumulation points are obviously closed, but not dense in themselves. The set of rational points of $(0, 1)$ or any open interval and sphere are dense in themselves, but not closed. All closed intervals and spheres are perfect. The empty set $\emptyset$ is also closed, but it is not considered to be dense in itself and so is not perfect. Note that a lot of point sets do not belong to any of the above three classes. This holds, for example, in the case of an arbitrary convergent sequence of distinct points.

We mention a consequence of the above property 2. Let I be a ν -dimensional interval and S a subset of it. We say that the set S is *dense in I* (everywhere), if each point of I is an accumulation point of S , i.e., $I \subseteq S'$. (For instance, the set of points with rational coordinates is everywhere dense

¹ Lindelöf [1].

² In other terms: if S has no isolated points.

in E_ν .) On the other hand, S is said to be *nowhere dense* in I , if no subinterval of I exists in which S is dense (in the sense just mentioned).

2.34 Turning to discussion of derived sets, we show next:

2.3.3 *The derivate of any point set is closed.*

P r o o f. If the set S in question is such that S' has no accumulation point, then the assertion is trivial. Otherwise, let us consider an accumulation point P of S' and a neighbourhood τ of P ; this τ must contain a point $P_1 \neq P$ of S' , and, because P_1 is an accumulation point of S , also a sequence consisting of distinct elements of S and converging to P . Thus P must also be an accumulation point of S , i.e., $P \in S'$. Therefore S' contains all its accumulation points.

* * *

If the formation of a derivate set is regarded as an operation, then it is easy to see that this operation may be carried out term by term in the case of a finite number of sets, i.e.,

$$(S_1 \cup S_2 \cup \dots \cup S_n)' = S'_1 \cup S'_2 \cup \dots \cup S'_n. \quad (2.3-1)$$

Hence, using 2.3.3, we can conclude that for any point set S , the set consisting of the accumulation points of

$$\bar{S} = S \cup S' \quad (2.3-2)$$

is equal to S' , so that $\bar{S} \supseteq (\bar{S})'$. Accordingly, the set (2.3-2) is always closed and it is called the *closure* (adherence) of S . This $\bar{S}$ yields a natural expedient in such cases when we need the enlargement of a non-closed set S to a closed set.

2.35 A further relevant problem is: what can be said about the *set of condensation points* of a point set S which will be denoted by $\dot{S}$. The set $\dot{S}$ is clearly a subset of S' ; we shall show that $\dot{S}$ is not only closed but also dense in itself when $\dot{S} \neq \emptyset$. That is to say:

2.3.4 *The set of condensation points of a non-countable point set is perfect.*

P r o o f. Let S be a non-countable set of points. It is assured by 2.3.2 that $\dot{S}$ is non-empty.

1° To verify that $\dot{S}$ is closed, consider any accumulation point P of the elements of $\dot{S}$. In an arbitrary neighbourhood τ of P , we can find a point $P_1 \neq P$ of $\dot{S}$. Furthermore, any $\tau_1 \subseteq \tau$ neighbourhood of P_1 (as a condensation point) contains a non-countable infinity of points of S . Since these points also belong to τ , P must be a condensation point of S , i.e., $P \in \dot{S}$. Therefore, it follows that $\dot{S}$ contains all its accumulation points.

2° As regards the assertion that $\dot{S}$ is dense in itself, let us choose any point Q of $\dot{S}$ and take a δ -neighbourhood T_0 of this point. If, for an arbitrary positive integer n , T_n denotes the $(\delta/2^n)$ -neighbourhood of Q and $\Delta T_n = T_{n-1} - T_n$, then we can write:

$$T_0 = \{Q\} + \bigcup_{n=1}^{\infty} \Delta T_n.$$

The non-countability of S implies that a non-countable infinity of points of this set are contained in T_0 . Thus, at least one of the sets $\Delta T_1, \Delta T_2, \dots$, say ΔT_m , has the same property by 1.2.5. Now applying theorem 2.3.2 to those points of S which belong to ΔT_m , we get that the latter set has a condensation point $Q_1 \neq Q$. Because $Q_1 \in \dot{S}$ and δ can be as small as we please, it follows that Q is an accumulation point of $\dot{S}$. This means that every point of $\dot{S}$ also belongs to $(\dot{S})'$.

* * *

2.36 On the basis of what has gone before, we can investigate, how far we can transmit the cardinality of a point set S to S' or $\dot{S}$; and, conversely, what can be said about the cardinality of a point set, if that of its derived set is known. These very natural questions are plainly connected with the cardinality problem of point sets having the properties 1., 2. or 3. above, since S' is always closed and $\dot{S}$ is besides dense in itself according to theorems 2.3.3 and 2.3.4, respectively.

It is clear that the derivate of a countable set can be finite, countably infinite or non-countable. (Examples: $\{1, \frac{1}{2}, \frac{1}{3}, \dots\}$; the union of the sequences $\{\frac{1}{n}\}, \{1 + \frac{1}{n}\}, \{2 + \frac{1}{n}\}, \dots$; the rational numbers in $[0,1]$.) Furthermore, a closed set or a set which is dense in itself can be countable or non-countable, too. (For instance: $\{1, 2, 3, \dots\}$ is closed and countable; the rationals of $(0,1)$ constitute an in itself dense and countable set; and the interval $[0,1]$ is closed, dense in itself and non-countable.)

Nevertheless, we have two fundamental statements on non-countable or perfect point sets:

2.3.5 *The derivate and the set of condensation points of a non-countable point set are also non-countable.*

2.3.6 Every perfect point set is non-countable.

P r o o f. The first assertion is a simple consequence of the second one. If S is a non-countable point set, then $\dot{S}$ is perfect by 2.3.4. Thus, assuming the validity of 2.3.6, it follows that $\dot{S}$ and (since $\dot{S} \subseteq S'$) S' are non-countable.

Now we restrict ourselves to the verification of 2.3.6.

Let T denote a perfect point set which is necessarily infinite by definition. Arguing indirectly, suppose that T is countable, i.e., it can be written as a point sequence $\{P_1, P_2, \dots\}$. Since T is dense in itself, every point of it is an accumulation point as well. Thus consider a point $Q_1 \neq P_1$ of T , and an ϵ -neighbourhood τ_1 of Q_1 which does not contain P_1 . There are infinitely many points of T in τ_1 and so we can pick out among them a point Q_2 differing both from P_2 and Q_1 . Thereafter take an $(\epsilon/2)$ -neighbourhood τ_2 of Q_2 which is a part of τ_1 and does not contain P_2 and Q_1 ; etc.

Continuing in this manner, we get a decreasing sequence of "spheres" $\tau_1, \tau_2, \dots$ such that the radius of τ_n tends to 0 as $n \rightarrow \infty$. Therefore, the sequence of points $Q_1, Q_2, \dots$ has a limit P_0 which is, since $Q_m \neq Q_n$, also an accumulation point of T . But T is closed as well and so $P_0 \in T$. Thus P_0 must be one of the points $P_1, P_2, \dots$, say P_N . But this is impossible, because P_0 lies in each of the neighbourhoods $\{\tau_n\}$ whereas, by construction, P_N is not contained in any of the neighbourhoods $\tau_{N+1}, \tau_{N+2}, \dots$.

This leads to a contradiction and so the original assertion is right.

* * *

We note that the first part of 2.3.5 is equivalent to the following lemma:

2.3.7 *If the derivate of a point set is countable, then the same holds for the set itself.*

2.4 Basic properties of closed and open sets

2.41 Let us compare the elementary notions: "interior point", "exterior point", "boundary point" (see 2.12) with the concept of accumulation point, introduced in the previous section.

On the basis of the relevant definitions, it is clear that every interior point of a set $S \subseteq E_\nu$ is also an accumulation (as well as a condensation) point of S , but no exterior point can be a point of accumulation. In the case of boundary points, any possibility may occur! 1. E.g. on the line, the end points of a finite interval or, in the plane, the boundary points of a circular disk are accumulation (and condensation) points, too. 2. The points of a ν -dimensional interval the coordinates of which are rational numbers or

irrational numbers constitute sets which consist of boundary points only. Furthermore, all the points of both these sets are also accumulation points and, in the case of irrational coordinates, condensation points as well. 3. Every point of an arbitrary set consisting of isolated points only (e.g. any finite point set) is necessarily a boundary point and such a set has no accumulation points at all. 4. In the plane the points of one of the coordinate axes or, more generally, the points $(x_1, x_2, \dots, x_\nu)$ of the space E_ν ($\nu \geq 2$) for which $x_i = 0$ (i fixed), also constitute sets consisting exclusively of boundary points which are also accumulation points of these sets.¹

After this, it is easy to verify a general theorem on the connection between closure and boundary points:

2.4.1 (I) *A point set S is closed if and only if it contains all of its boundary points.*

(II) *The boundary of a point set is always closed.*

Proof. (I) From the discussion above, it is clear that S consists of its interior points and of certain boundary points; the isolated points of S being among the latter. On the other hand, S' contains all the interior points of S , as well as the non-isolated boundary points (which are not necessarily elements of S). Thus $S \supseteq S'$ holds exactly on the condition that every non-isolated boundary point belongs to S , and this means that all the boundary points are contained by S .

(II) is an immediate consequence of the first part of (I), applied to a point set without interior points.

* * *

2.42 We now have to deal in detail with that important class of sets which is in a sense complementary to the class of closed sets. These are the *open sets*. A point set is called open if it does not contain boundary points at all, i.e., if it *contains interior points only*. Hence it is evident that any non-empty open set is dense in itself and has the cardinality $\aleph$. The empty set $\emptyset$ and its complement, the whole space E_ν , are considered to be closed and open at the same time.

Beside open intervals or balls of arbitrary dimension and systems of these, the most elementary examples of open sets in the plane and space are: any *convex set* (i.e., such a set which contains the line segment linking each pair of its points) without its boundary points and, more generally, the interior of any simple closed curve or surface (cf. 2.13). Every set of the

¹ These examples remind us of the importance of *clarifying in advance the dimension of a set in any discussion about concepts connected with the notion of neighbourhood*. Thus $(0, 1)$ as a part of the line has interior points only, but the same interval consists exclusively of boundary points if it is considered as a part of the plane.

type just mentioned, obviously becomes closed after including its boundary points. Furthermore, each of them possesses the property of *connectedness*, i.e., they cannot be represented as the union of two non-empty, disjoint, open point sets.¹ We also mention two other relevant notions. Any point set $S \subseteq E_n$ is called a *region*, if it is non-empty, open and connected; and a set S consisting of more than one point is called a *continuum*, if it is compact and connected (cf. 2.31).²

2.43 There is a kind of *duality* between open and closed sets in E_n , expressed by the following theorem:

2.4.2 *Omitting from any closed/open point set an open/closed subset, leaves a closed/open set.*

Epecially: the complement of a closed/open set with respect to the whole space is open/closed.

P r o o f. Let S be, e.g., a closed point set and T an open subset of it. We show that $S - T$ contains all its boundary points. In fact, let P be an accumulation point of $S - T$ and assume that $P \notin S - T$. Then P must be a boundary point both of $S - T$ and T . On the other hand, P is an accumulation point and, since S is closed, also an element of S . Thus the assumption that P does not belong to $S - T$ implies $P \in T$. But T is open and cannot contain a boundary point. Hence we must have $P \in S - T$.

The other part of the assertion can be discussed similarly.

* * *

On the application of the main set operations we have:

2.4.3 (I) *The union of a finite number of closed sets and the intersection of an arbitrary number of closed sets is closed.*

(II) *The union of an arbitrary number of open sets and the intersection of a finite number of open sets is open.*

P r o o f. For instance, the part of (I) about the union operation is a direct consequence of the definition of closure and of formula (2.3-1). The

¹ The last definition is equivalent to the characterization known from elementary analysis: a point set S is connected, if any two of its points can be linked by a "continuous path" lying in S .

² Certain special types of region in the plane are also often useful, namely the *simply* or *multiply connected* ones. A region $S \subset E_2$ is said to be m -fold connected, if it is homeomorphic to a point set obtained from the interior of a circle by omitting $(m - 1)$ closed and disjoint circular disks. (Cf. 2.13.) Note that any continuum has the cardinal $\aleph$ in accordance with 1.4.1.

remaining part of (I) can be obtained by observing that any accumulation point of an intersection of closed sets must also be an accumulation point of each of the intersecting sets, and (by closure) an element of each of them. It therefore belongs to their intersection as well.

The assertion (II) follows at once from (I) by using the “duality principle” 2.4.2.

* * *

Note that in 2.4.3 “finite number” may not be replaced by “arbitrary number”. For example, the union of the closed sets $\{1\}, \{1/2\}, \{1/2^2\}, \dots$ is obviously not closed, since it does not contain the origin; and the intersection of the open intervals $(-2^{-n}, 2^{-n})$ ($n = 1, 2, \dots$) (i.e., $\{0\}$) is not open.

2.44 The above results suggest a thorough investigation of the structure of open sets. First of all, we show that, *on the line*, no open sets exist other than systems of open intervals, mentioned among the preliminary examples in this section. More precisely:

2.4.4 *Any linear open set can be expressed as the union of countably many disjoint open intervals, which apart from ordering, is unique. This decomposition is called the normal representation of the set.*

P r o o f. Let ξ be an arbitrary point of the number line, belonging to the set S in question. Since S is open, there are real numbers a and b , such that $a < \xi < b$ with $(a, b) \subseteq S$. If $A = \inf a$ ($\geq -\infty$) and $B = \sup b$ ($\leq \infty$), then (A, B) is the largest open interval which contains ξ and consists of points of S only. In the case when S also has a point which does not belong to (A, B) , we repeat the procedure until all the points of S are exhausted.

In the resulting representation S is split up into a (uniquely determined) system of non-overlapping open intervals which can be ordered so that both the left and the right end points are strictly monotone increasing.

* * *

Theorem 2.4.4 cannot be extended to the open sets of $E_2, E_3, \dots$, etc. Although it is possible to find analogous decompositions into “balls” or intervals, but these can overlap and the representation is no longer unique. The situation is similar about the following statement. Any open point set $S \subseteq E_n$ can be written as the union of countably many *closed* n -dimensional intervals, if common boundary points are admitted. For example, we have:

$$(0, 1) = \bigcup_{n=1}^{\infty} [2^{-(n+1)}, 2^{-n}] + \bigcup_{n=1}^{\infty} [1 - 2^{-n}, 1 - 2^{-(n+1)}].$$

We also remark that the open and closed sets of the line can be interpreted as *Borel sets* by 2.4.4. (Cf. 1.52)

2.45 In the knowledge of the normal representation of open sets on the line, the first part of theorem 2.4.3 suggests that perhaps any linear *closed* point set would be generated in the form of an interval system. (That is as the union of a finite or countably infinite number of closed intervals.) That this is not possible, will be illustrated by the famous example of *Cantor's ternary set* ($\mathfrak{T}$)¹ which, in spite of being closed, does not contain a single closed interval.

The definition of $\mathfrak{T}$ is the following. Let us leave out from $[0, 1]$ the open interval $(1/3, 2/3)$. From the remaining subintervals $[0, 1/3]$ and $[2/3, 1]$ omit again the open middle thirds $(1/9, 2/9)$, $(7/9, 8/9)$, respectively. Repeat the procedure with the last obtained four closed subintervals and continue in a similar manner. The set $\mathfrak{T}$ consists of the points which remain from $[0, 1]$ after continuing the procedure ad infinitum. Now, it is easy to prove the following theorem.

2.4.5 The point set $\mathfrak{T}$ is (1) closed and dense in itself i.e. perfect; (2) nowhere dense in the interval $[0, 1]$; (3) non-countable.

Preliminary remarks: I. the comparison of the properties (1)–(2) shows that a point set can be dense in itself but nowhere dense in an interval at the same time (cf. 2.33); II. property (2) naturally precludes the existence of a subinterval $j \subset [0, 1]$, however small, such that every point of j belongs to $\mathfrak{T}$.

Proof. By 2.4.2 it follows that $\mathfrak{T}$ is closed, since its defining procedure starts from a closed set, from which only open sets are excluded. Property (2) is obvious, because any subinterval of $[0, 1]$ contains an open interval (obtained by successive division into three parts) which does not belong to $\mathfrak{T}$.

To verify that our set is dense in itself, consider a point $P \in (0, 1)$ of $\mathfrak{T}$. If this is not an accumulation point of $\mathfrak{T}$, then each point of some neighbourhood of P apart from P itself would belong to $C\mathfrak{T} = [0, 1] - \mathfrak{T}$. Thus $C\mathfrak{T}$ should contain two subintervals (of equal length) of $(0, 1)$ with a boundary point in common. But this would contradict the construction of $C\mathfrak{T}$. Thus in $(0, 1)$ every point of $\mathfrak{T}$ is an accumulation point. Both of the points 0 and 1 (belonging to $\mathfrak{T}$) are also accumulation points, since all the end points of the open intervals "left out" (among them $1/3, 1/9, \dots$ as well as $2/3, 8/9, \dots$) are clearly elements of $\mathfrak{T}$.

Finally, the non-countability of $\mathfrak{T}$ follows from property (1) and theorem 2.3.6.

* * *

¹ $\mathfrak{T}$ and the more general point sets constructed in a similar manner also have the name "*Cantor's discontinua*" which stresses the fact that they are *not connected*.

So, the structure of the linear closed set $\mathfrak{T}$ turns out to be rather complicated. This can be said all the more about some closed sets of $E_2, E_3, \dots$. Hence we see, that the closed continua and some finite systems of them (mentioned at the beginning of the present section) may be ranked amongst the most elementary examples of closed point sets.

Nevertheless, it is possible to point out something definite about the structure of an arbitrary non-countable closed set in E_n . An interesting *decomposition theorem* allows us (essentially) to reduce the discussion of such sets to the case of perfect sets.

2.5 The Cantor–Bendixson theorem, Cantor’s meet theorem

2.51 Before the treatment of the result in question, we have to deal briefly with the concept of *distance between two non-empty point sets* S and T . If P denotes an arbitrary point of S and Q that of T , then we mean by the distance between S and T the greatest lower bound $\rho \geq 0$ of the set of all possible values PQ . (For instance, the distance of the set $\{0, 1, 2\}$ from the interval $(2, \infty]$ is 0, and from the interval $[1/2, 3/4]$ it is $1/4$.) If we assume that S and T are *closed* and at least one of them is *bounded*, then $\rho = \rho(S, T)$ is simultaneously the *minimum* of PQ relating to the totality of point pairs $P \in S, Q \in T$; i.e., we can find a point P_0 in S and another point Q_0 in T such that $\overline{P_0Q_0} = \rho$.¹

In particular, the distance of a point Q (as a point set having a single element) from a point set $S \neq \emptyset$ is equal to 0 only if this point belongs to the closure of S [cf. (2.3-2)]. Thus $\rho(\{Q\}, S) > 0$, if S is closed and $Q \notin S$.

2.52 Let S be a non-empty closed point set which is countable. Then we have $S = S' \cup A$ where A , the so-called *isolated part* of S , is disjoint from S' and non-empty (by 2.3.6). Furthermore, both S' and A are countable. (Of course, S' may be empty, e.g. if S is finite, or infinite and consists of isolated points.)

If S is a non-countable closed set, then the corresponding decomposition can be sharpened in the following way:

2.5.1 (CANTOR’S AND BENDIXSON’S THEOREM)² (I) *Any non-countable, closed point set S can be written as the union of a perfect set S_1 and a countable set M which is disjoint from S_1 .*

(II) *This representation is unique since S_1 is $\dot{S}$, the set of condensation points of S , and $M = S - \dot{S}$.*

¹ The existence of such a pair of points P_0, Q_0 can be easily verified by using 2.2.2.

² Cantor [3], Bendixson [1].

P r o o f. The set $\dot{S}$ is non-empty by 2.3.2 and is a subset of S since the latter is closed. Moreover S is perfect by 2.3.4. Thus we have to show: 1° the difference $M = S - \dot{S}$ is countable; 2° if S is any perfect subset of S for which $S - S_1$ is countable, then $S_1 = \dot{S}$.

1° If $M = \emptyset$, i.e., $S = \dot{S}$, the assertion is trivial; so let us suppose that M is non-empty. Since $\dot{S}$ is closed, the distance Δ of any point of M from $\dot{S}$ must be positive. Let M_1 denote the subset of M whose points satisfy the condition $\Delta \geq 1$; let $M_2, M_3, \dots$ be the subsets of M characterized by $1/2 \leq \Delta < 1$, by $1/3 \leq \Delta < 1/2, \dots$, respectively. These subsets are pairwise disjoint, and we have:

$$M = \bigcup_{n=1}^{\infty} M_n.$$

Assume that one of these subsets, say M_k , is non-countable. Then it has at least one condensation point Q (cf. 2.3.2) which, since $M_k \subseteq S$, is a condensation point of S as well. On the other hand, the definition of M_k implies that the distance of Q from $\dot{S}$ is at least $1/k$. Thus Q would be such a condensation point of S which does not belong to $\dot{S}$: a contradiction!

Consequently, all the subsets M_n are countable and, by 1.2.5, so is M .

2° Let S_1 be a point set satisfying the stated conditions. Then, since $S - S_1$ has no condensation point, it follows that $\dot{S} = \dot{S}_1$. We have to show that $S_1 = \dot{S}_1$.

The relation $S_1 \supseteq \dot{S}_1$ is a direct consequence of the fact that S_1 is closed. So we have to verify that $S_1 \subseteq \dot{S}_1$, i.e., that every point $P_1 \in S_1$ is also a condensation point of S_1 . (*)

To that end, consider a closed "ball" $\tau \subset S$ with centre P_1 and that part S_τ of S_1 which is contained in τ . Just like S_1 , the set S_τ is also perfect. In fact, every accumulation point of S_τ belongs to S_1 and is contained in τ , therefore S_τ is closed. Furthermore, for any point $P \in S_\tau$ we can find a sequence of points in S_1 which converges to P (because S_1 is dense in itself); and only a finite number of elements of this sequence can lie outside τ . Thus each element of S_τ is also an accumulation point of this set, i.e. S_τ is dense in itself.

Summing up, the above assertion (*) follows at once, since the radius of the closed neighbourhood τ can be as small as we like, and S_τ (as a perfect set) is not countable by 2.3.6.

* * *

Observe that the second part of our argument yields a result which is interesting in itself: *any perfect set S is identical to the set of its condensation points*. (Therefore, in our case, not only $S = S'$, but also $S = \dot{S}$

holds.) This statement can also be considered as a corollary of the Cantor-Bendixson theorem: if S is perfect, then it satisfies the conditions of 2.5.1 (cf. 2.3.6), and clearly $S = \dot{S} \cup \bigcirc = \dot{S}$.

Some usual notations: in 2.5.1, the subset $S_1 (= S)$ is called the *perfect kernel* of S , while $M (= S - \dot{S})$ is called the *separate part* of S . The former is the “largest” perfect part of S , the latter is countable.

2.53 With the notion of distance a further fundamental concept is connected: the *diameter of a point set*. Let S be a bounded point set. Let us consider the distance between two points of S , and let $d(S)$ denote the least upper bound for all pairs of points of S of the set of non-negative numbers thus obtained. We call $d(S)$ the *diameter* of S . It is zero if S consists of a single point, otherwise $d(S) > 0$. (It is clearly a generalization of the notion “diameter of a circle or a sphere”.) If S is closed, then — as we know — there exist at least two points of S whose distance is just $d(S)$. In this case, we may write it in the form:

$$d(S) = \max_{P_1, P_2 \in S} \overline{P_1, P_2}. \quad ^1 \quad (2.5-1)$$

The concept of diameter makes possible the extension of certain theorems, relating to special geometric configurations, to arbitrary bounded, closed point sets. For instance, the elementary lemma on “nested intervals” — which is often applied in the foundations of analysis — takes the following form in the general theory of point sets:

2.5.2 (CANTOR’S MEET THEOREM)² *For any decreasing sequence $S_1, S_2, \dots$ of non-empty, bounded and closed point sets, the intersection*

$$\Pi = \bigcap_{n=1}^{\infty} S_n \quad (2.5-2)$$

cannot be empty. Furthermore, if we also assume that $d(S_n) \rightarrow 0$ as $n \rightarrow \infty$, then Π must consist of a single point.

P r o o f. Choose from each set S_n a point P_n , and consider the sequence of points $P_1, P_2, \dots$. Since $P_n \in S_1$, for all n , this sequence $\{P_n\}$ is bounded. If any element Q of $\{P_n\}$ occurs infinitely many times, then there exist sets S_n with arbitrary large indices n , containing Q . Since $\{S_n\}$ is decreasing,

¹We see that for both $\rho(S, T)$ and $d(S)$, the definition in terms of an extreme value is related to the notion of *closure*. This property is always useful when the solvability of *extreme value problems* is investigated.

² Cf. Cantor [2].

this is possible if and only if Q belongs to every S_n , i.e., $Q \in \Pi \neq \emptyset$. Otherwise, i.e., if each element of $\{P_n\}$ occurs only finitely many times, then, according to 2.2.2, we can select from $\{P_n\}$ a convergent subsequence of distinct points. Since the limit P of this subsequence is an accumulation point of every set S_n and, since all the S_n are closed, this implies that $P \in S_n$ for all n . Thus $P \in \Pi$ and Π is again non-empty, as asserted above.

The added condition $d(S_n) \rightarrow 0$ (*) implies that Π cannot contain two distinct points. Otherwise, we would have a pair of points $Q_1 \neq Q_2$ such that $Q_1 \in S_n$, $Q_2 \in S_n$ ($n = 1, 2, \dots$), and $d(S_n) \geq Q_1 Q_2$ for all n , which contradicts the condition (*).

2.6 Covering theorems, outer measures, null sets

2.61 In the theory of real functions, it is often necessary to have for a given point set S a system of (possibly simple) sets whose union contains all the points of S . If there such a set system R exists, then we say that R covers S .¹

It is an important fact than if we consider bounded closed sets covered by open sets, we can always restrict ourselves to the case of a finite number of covering sets. Formulating this in detail, we have

2.6.1 (BOREL'S COVERING THEOREM)² *If a bounded and closed point set S can be covered by an infinite system of open sets, then there exists a finite subsystem of the latter which also covers S .*

P r o o f. Our conditions include the existence of a finite interval I for which $S \subseteq I$. Let $\{M_\lambda\}$ be such an infinite system of open sets which covers S . Then $Z = \bigcup_\lambda M_\lambda \supseteq S$, and Z is also open by 2.4.3(II).

1° If $\{M_\lambda\}$ is countable, then its elements can be arranged in a sequence $M_1, M_2, \dots$. Let us denote the union of $M_1, M_2, \dots, M_n$ by Z_n , the complement of the latter, with respect to the whole space, by CZ_n , and write $S_n = S \cap CZ_n$ ($n = 1, 2, \dots$).

Since $\{Z_n\}$ is increasing, so $\{CZ_n\}$ and $\{S_n\}$ are decreasing sequences of sets. Also each S_n is closed by our assumptions [cf. 2.4.3(I)]. Furthermore, $\{S_n\}$ is bounded by $S_n \subseteq S \subseteq I$. Now, if $S_n \neq \emptyset$ for every n , we can apply

¹ The system R may consist of a finite number of sets, of countably or non-countably many sets. Moreover, not only disjoint covering sets, but also sets which overlap, or have common boundary points are permitted.

² Borel [2].

theorem 2.5.2 to get that $\bigcap_{n=1}^{\infty} S_n = S \cap CZ \neq \emptyset$. Hence it follows that CZ contains at least one point of S , which contradicts the condition $Z \supseteq S$.

Thus there must exist an index m such that $S_m = CZ_m \cap S = \emptyset$, i.e., $Z_m \supseteq S$. In other words, for some finite m , the system $\{M_1, M_2, \dots, M_m\}$ covers S .

2° If $\{M_\lambda\}$ is non-countable, then the proof can be reduced to the case discussed above.

We have only to utilize some results and comments of section 2.44 on the general structure as well as on certain characteristic properties of open sets.

* * *

2.62 As a rule, Borel's theorem is applied so, that the covering systems consist of intervals or "spheres". Moreover we want the sum of diameters of the covering sets to be as small as possible. This last requirement leads naturally to the very significant concept of outer (exterior) measure.

Let us define the *measure of a ν -dimensional interval i* as the product of the lengths of linear intervals determining i .¹ Let S be an arbitrary point set in E_ν , and consider the set of all *finite* systems of ν -dimensional intervals which cover S . Taking the total measure² of each system in question, we obtain a set consisting of positive numbers. The greatest lower bound of this set of numbers is called the (ν -dimensional) *outer (J)-measure* of S and will be denoted by $\bar{\mu}_J(S)$.

If the definition given above is extended from "finite" to "finite or countably infinite" systems, then we get the ν -dimensional *outer (L)-measure* of S , denoted by $\bar{\mu}_L(S)$. (The indices in the symbols $\bar{\mu}_J, \bar{\mu}_L$ refer to the names of Jordan and Lebesgue, respectively.)

For example, if S is a set of finite, disjoint linear intervals, then $\bar{\mu}_J$ and $\bar{\mu}_L$ are both equal to the total length of the system.³ On the other hand, if S is the set of rational numbers belonging to $(0,1]$, we have $\bar{\mu}_J = 1$ and $\bar{\mu}_L = 0$, since every finite covering interval system must contain all the points of $(0,1]$. Whereas, if the covering system is countably infinite, we can cover the rationals in $(0,1]$ by covering the first one by an interval of length $\epsilon/2$, the second by an interval of length $\epsilon/2^2$, and so on. The interval system so obtained contains, for arbitrary values of ϵ , every rational number in $(0,1]$

¹ Thus the measure of a line segment is its length, that of a rectangle its area, that of a rectangular parallelepiped its volume, etc. (The fact that i is closed or open is irrelevant.)

² By this we mean the sum of measures of the intervals figuring in the system.

³ It is evident that $\bar{\mu}_J$ and $\bar{\mu}_L$ cannot exceed this value. The other part of the assertion is also easy to see.

and has the total length:

$$0 < l(\epsilon) = \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \frac{\epsilon}{2} \frac{1}{1-0.5} = \epsilon \quad (2.6-1)$$

which can be made as small as we like.

2.63 In what follows we shall frequently need point sets whose outer (L)-measure is 0, as in case of the last example. These will be called *null sets*. In other words, $S \subset E_\nu$ is said to be a ν -dimensional null set if, given any positive number ϵ , however small, we can find a finite or countably infinite system of ν -dimensional intervals covering S , whose total measure is less than ϵ .

It must be emphasized that *not only all finite point sets, but also all countably infinite ones are null sets*. This follows from the typical argument above, using the geometric series (2.6-1). Applying the same idea to a sequence of null sets, we obtain, more generally:

2.6.2 *The union of countably many null sets is also a null set.*

However, the totality of null sets is wider than that of the countable sets, i.e. there are non-countable null sets, too. For instance, the set $\mathfrak{T}$ introduced in 2.45 is non-countable by 2.4.5; but, if we consider the subintervals of $[0,1]$ remaining after the r -th step of the construction of $\mathfrak{T}$, then this system obviously covers $\mathfrak{T}$ and has the total length:

$$1 - \left(\frac{1}{3} + 2 \frac{1}{3^2} + \dots + 2^{r-1} \frac{1}{3^r} \right) = 1 - \left[1 - \left(\frac{2}{3} \right)^r \right] = \left(\frac{2}{3} \right)^r,$$

which becomes arbitrarily small for r sufficiently large. Therefore, *Cantor's ternary set is a null set*.

2.64 It turns out from the definitions above, that every bounded point set has finite outer (J)- and (L)-measures, while both $\bar{\mu}_J$ and $\bar{\mu}_L$ are infinite in the case of an unbounded set.

Furthermore, since a wider class of coverings is permitted in the definition of $\bar{\mu}_L(S)$ than in that of $\bar{\mu}_J(S)$, it follows that, for any bounded point set S :

$$0 \leq \bar{\mu}_L(S) \leq \bar{\mu}_J(S) < \infty. \quad (2.6-2)$$

It is easy to see that $\bar{\mu}_J$ and $\bar{\mu}_L$ possess the following simple properties:

1. If $s \subseteq S$, then

$$\bar{\mu}_J(s) \leq \bar{\mu}_J(S) \quad \text{and} \quad \bar{\mu}_L(s) \leq \bar{\mu}_L(S), \quad (2.6-3)$$

2. The outer measures of the union of the bounded sets $S_1, S_2, \dots, S_n$, cannot exceed the corresponding total measures of this system of sets.

3. Let $S_1, S_2, \dots$ be bounded sets whose union is bounded as well. Then

$$\bar{\mu}_L \left(\bigcup_{k=1}^{\infty} S_k \right) = \lim_{n \rightarrow \infty} \bar{\mu}_L \left(\bigcup_{k=1}^n S_k \right) \leq \sum_{k=1}^{\infty} \bar{\mu}_L(S_k). \quad ^1 \quad (2.6-4)$$

We remark that the outer (L)-measure of a bounded set $S(\subset E_\nu)$ can also be defined in the following equivalent way:

$$\bar{\mu}_L(S) = \inf_{Z \supseteq S} \{ \bar{\mu}_L(Z) \}, \quad (2.6-5)$$

where Z denotes any ν -dimensional, bounded and *open* set which covers S .

In fact, by 2.44, Z can always be considered as the union of countably many, disjoint open or closed intervals. Thus $\bar{\mu}_L(Z)$ may be interpreted as the total measure of such a system of intervals.

2.65 Finally, by means of the concept of outer (L)-measure, we shall deal with a useful one-dimensional covering theorem for arbitrary linear sets.

Next we have to say that Borel's theorem 2.6.1 is not valid, if the covered set S is not closed. To verify this, let us consider the countable set $s = \{1/2, 1/3, 1/4, \dots\}$ which is not closed because it does not contain its only accumulation point 0, and about each point $1/n$ ($n = 2, 3, \dots$) take an open interval of length $[n(n+1)]^{-1}$, centred on $1/n$. This gives us an infinite system $\{j\}$ which covers s , but clearly, no finite subsystem of $\{j\}$ will cover all of s . The total length of this $\{j\}$ is:

$$\sum_{n=2}^{\infty} \frac{1}{n(n+1)} = \left(\frac{1}{2} - \frac{1}{3} \right) + \left(\frac{1}{3} - \frac{1}{4} \right) + \dots = \frac{1}{2}.$$

Now, our example shows that the part of s which is not contained in the union of (disjoint) intervals covering the first N elements of s , can be squeezed into any given small interval $(0, \epsilon)$ for N large enough. We want to show that a similar result is true for coverings of arbitrary bounded linear sets. We first mention a useful abbreviating term.

Suppose that a linear set S is covered by an infinite system of intervals $\{I\}$. We say that S can be *essentially covered* by finite subsystems of $\{I\}$ if for any small $\epsilon > 0$, we can find a decomposition of S into two disjoint sets, T_1 and $T_2 = S - T_1$, with the following properties: T_1 is covered even by finitely many (suitably selected) intervals of the system $\{I\}$ and T_2 has an outer (L)-measure less than ϵ .

¹ Here and in the previous property, equality holds if the pairwise distances between the sets S_k are positive, and for (2.6-4) we have also $\sum_{k=1}^{\infty} \bar{\mu}_L(S_k) < \infty$.

2.6.3 (SIERPINSKI'S LEMMA)¹ Let $\{I\}$ be an infinite system of closed intervals on the line which covers a bounded infinite linear set S in such manner that each point P of S is the left-hand side² end point of any interval covering P . Then S can be essentially covered by finite subsystems of $\{I\}$ consisting of non-overlapping intervals.

P r o o f. 1° Let us denote by s_n the totality of such points $x \in S$ to which at least one interval $[x, x + l]$ longer than $1/n$ ($n = 1, 2, \dots$) can be attached. As a subset of S , s_n is bounded; furthermore we have $s_1 \subseteq s_2 \subseteq \dots$ and $S = s_1 \cup s_2 \cup \dots$.

Therefore, by (2.6-3) and (2.6-4), $\{\bar{\mu}_L(s_n)\}$ is a monotone increasing number sequence and, since $s_1 \cup s_2 \cup \dots \cup s_n = s_n$ we have

$$\lim_{n \rightarrow \infty} \bar{\mu}_L(s_n) = \bar{\mu}_L(S). \quad (2.6-6)$$

Hence we see that, given $\epsilon > 0$, there exists an index $N = N_\epsilon$ such that

$$\bar{\mu}_L(S) - \frac{\epsilon}{4} < \bar{\mu}_L(s_n) \leq \bar{\mu}_L(S). \quad (2.6-7)$$

2° Now consider the set $s_N (\subseteq S)$ on the number line (Fig. 1).

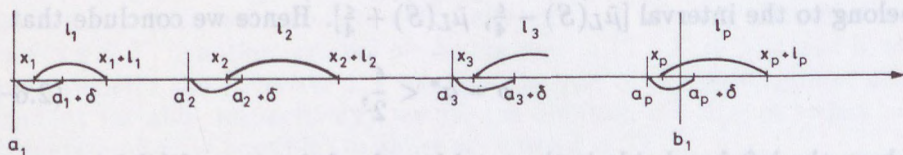


Fig. 1

Let the greatest lower and the least upper bound of s_N be a_1 and b_1 , respectively (both are finite); and let δ denote a positive number to be fixed later. The definition of a_1 implies that s_N contains a point x_1 for which $a_1 \leq x_1 < a_1 + \delta$. Let $I_1 = [x_1, x_1 + l_1]$ be an interval of the system $\{I\}$ covering x_1 and of length $l_1 > 1/N$. (By the definition of s_N , the existence of such an interval is assured.) If there are points $x \in s_N$ for which $x \geq x_1 + l_1$, take the greatest lower bound of them. Then we can find an $x_2 \in s_N$ for which $(x_1 + l_1) \leq a_2 \leq x_2 \leq a_2 + \delta$ and an interval $I_2 = [x_2, x_2 + l_2] \in \{I\}$ whose length satisfies $l_2 > 1/N$. If there are points $x \in s_N$ with $x \geq x_2 + l_2$, we can select from $\{I\}$ an interval $[x_3, x_3 + l_3]$, lying on the right-hand side of $x_2 + l_2$ and satisfying $l_3 > 1/N$; etc.

The lengths of the closed intervals thus obtained exceed a common lower bound, they are non-overlapping, and the left-hand side end point of each

¹ Sierpinski [1].

² The assertion obviously remains valid, if the term "left-hand side" is replaced by "right-hand side".

of them belongs to the finite interval $[a_1, b_1]$. As is seen, the number of the pertinent intervals must be finite, say p . Moreover, for $p > 1$ we can write $l_1 + l_2 + \dots + l_{p-1} < b_1 - a_1$, which by $l_k > 1/N$ ($k = 1, 2, \dots, p$) yields the inequalities $(p-1)/N < b_1 - a_1$ and

$$p < N(b_1 - a_1) + 1. \quad (2.6-8)$$

3° On the basis of the definition of $\bar{\mu}_L(S)$, there exists an (in general infinite) interval system $\{i\}$, whose total length σ approximates $\bar{\mu}_L(S)$ from above with an error less than $\epsilon/4$, i.e., for which

$$\bar{\mu}_L(S) \leq \sigma < \bar{\mu}_L(S) + \frac{\epsilon}{4}. \quad (2.6-9)$$

Let us denote the union of the intervals I_k ($k = 1, 2, \dots, p$) by J and that of the intervals $[x_k - \delta, x_k]$ ($k = 1, 2, \dots, p$) by S_δ . Then, that part of the system $\{i\}$ which is contained in $J + S_\delta$ obviously covers s_N .¹ Denoting by σ^* the total length of this covering set, we get

$$\bar{\mu}_L(s_N) \leq \sigma^* \leq \sigma. \quad (2.6-10)$$

The comparison of (2.6-7), (2.6-9) and (2.6-10) shows that both σ and σ^* belong to the interval $[\bar{\mu}_L(S) - \frac{\epsilon}{4}, \bar{\mu}_L(S) + \frac{\epsilon}{4}]$. Hence we conclude that

$$\sigma - \sigma^* < \frac{\epsilon}{2}, \quad (2.6-11)$$

where the left-hand side is the total length of the part of $\{i\}$ lying outside $J + S_\delta$.

As far as the interval system S_δ itself is concerned, its total length cannot exceed the value $p\delta$. But according to 2°, p is less than a fixed upper bound [cf. (2.6-8)]. It therefore follows that

$$p\delta < \frac{\epsilon}{2}, \quad (2.6-12)$$

provided that δ is chosen to be sufficiently small.

Collecting the results obtained above, we have established: $\epsilon > 0$ being given, N and δ can always be determined so that: 1. the disjoint intervals $I_1, I_2, \dots, I_p$, selected from $\{I\}$ in the manner described, contain a part of the set S ; 2. the remaining part of S can be covered by a system of intervals whose total length is less than $(\epsilon/2) + (\epsilon/2) = \epsilon$. Since it follows from the last statement that the outer (L)-measure of this remainder is also $< \epsilon$, the assertion is proved.

¹ Namely, our conditions imply that no point of s_N lies between $x_k + l_k$ and a_k , furthermore we have $[x_k - \delta, x_k] \supset [a_k, x_k]$ ($k = 1, 2, \dots, p$).

We mention still another classical result of similar character, which is, however, deeper and much more difficult to verify.

2.6.4 (VITALI'S COVERING THEOREM)¹ Assume that an infinite system $\{j\}$ of closed intervals covers a bounded infinite linear set S , so that each point of S can be included into a subsystem j_n ($n = 1, 2, \dots$) of $\{j\}$ for which $|j_n| \rightarrow 0$ ($n \rightarrow \infty$). Then we can select from $\{j\}$ countably many, pairwise disjoint intervals which together cover S as well, apart possibly from a null set.

their classification

3.1.1 As is well known, Dirichlet's elementary concept of function means that all elements of a number set X ("domain of definition") are associated with some elements of another number set Y ("range of values" or "codomain" of the function). Briefly, the function yields a mapping of X into Y (denoted simply by $X \rightarrow Y$). On the basis of the set theoretical foundation, it is natural to replace both X and Y by any kind of abstract sets; hereby we obtain the following general notion of function:

Let X and Y be two arbitrary sets. We call *abstract function* or *operator* and denote by T a map of X into Y ; i.e., T is a rule of correspondence according to which we associate every element $x \in X$ with one or more elements $y \in Y$. Another symbol of the concept is $\{T : X \rightarrow Y\}$; and if we want to register the elements x, y (playing the role of the independent and dependent variable, respectively), we use the notation $y = T[x]$ or (when no misunderstanding is possible) in short we write $y = Tx$.

If the number of image elements y attached to any element $x \in X$ is at most p ($1 \leq p \leq \infty$), then we say that the operator T is *p-valued*¹ in X . The widest basic set X considered in the definition of T is called its *domain*, the totality of the corresponding elements $y \in Y$ is the so-called *codomain* of T . By the *inverse* of T we mean the mapping which transfers each element y of the codomain of T into the one or more elements $x \in X$ whose image is y ; this operator is denoted by T^{-1} . Thus we have $T[T^{-1}[y]] = y$ for any y belonging to the codomain, and it holds also for each $x \in X$ the relation $T^{-1}[T[x]] = x$ provided that T is one-valued. If both of T and T^{-1} are one-valued, then there is a one-to-one correspondence between the elements of domain and codomain.

Applying to the codomain of an operator $y = T_1[x]$ another operator $z = T_2[y]$, we obtain the *composed operator* $z = T_2[T_1[x]] = T_2T_1x$, which is called by the latter symbolic notation also the *product* of T_1 and T_2 .² This

¹ For $p = 1$ the term "single-valued", for $p > 1$ the denomination "multiple-valued" is also used.

² For T_1 and T_2 we use the symbol $T_1 \circ T_2$, too. In case of identical factors, we apply naturally $T \circ T = T^2$, $T \circ T \circ T = T^3$, etc.

¹ Cf. Vitali [3], Vitali-Sansone [1].

EXTENSION OF THE CONCEPT OF FUNCTION. LIMIT, CONTINUITY AND DIFFERENTIABILITY OF A POINT FUNCTION

3.1 Functions on abstract sets (operators) and their classification

3.11 As is well-known, Dirichlet's elementary concept of function means that all elements of a number set X ("domain of definition") are associated with some elements of another number set Y ("range of values" or "codomain" of the function). Briefly, the function yields a *mapping* of X into Y (denoted simply by $X \Rightarrow Y$). On the basis of the set theoretical foundation, it is natural to replace both X and Y by any kind of abstract sets; hereby we obtain the following *general notion of function*:

Let X and Y be two arbitrary sets. We call *abstract function* or *operator* and denote by $\mathcal{T}$ a map of X into Y ; i.e., $\mathcal{T}$ is a rule of correspondence according to which we associate every element $x \in X$ with one or more elements $y \in Y$. Another symbol of the concept is $\{\mathcal{T} : X \Rightarrow Y\}$; and if we want to register the elements x, y (playing the role of the independent and dependent variable, respectively), we use the notation $y = \mathcal{T}[x]$ or (when no misunderstanding is possible) in short we write $y = \mathcal{T}x$.

If the number of image elements y attached to any element $x \in X$ is at most p ($1 \leq p \leq \infty$), then we say that the operator $\mathcal{T}$ is *p-valued*¹ in X . The widest basic set X considered in the definition of $\mathcal{T}$ is called its *domain*, the totality of the corresponding elements $y \in Y$ is the so-called *codomain* of $\mathcal{T}$. By the *inverse* of $\mathcal{T}$ we mean the mapping which transfers each element y of the codomain of $\mathcal{T}$ into the one or more elements $x \in X$ whose image is y ; this operator is denoted by $\mathcal{T}^{-1}$. Thus we have $\mathcal{T}[\mathcal{T}^{-1}[y]] = y$ for any y belonging to the codomain, and it holds also for each $x \in X$ the relation $\mathcal{T}^{-1}[\mathcal{T}[x]] = x$ provided that $\mathcal{T}$ is one-valued. If both of $\mathcal{T}$ and $\mathcal{T}^{-1}$ are one-valued, then there is a one-to-one correspondance between the elements of domain and codomain.

Applying to the codomain of an operator $y = \mathcal{T}_1[x]$ another operator $z = \mathcal{T}_2[y]$, we obtain the *composed operator* $z = \mathcal{T}_2[\mathcal{T}_1[x]] = \mathcal{T}_2\mathcal{T}_1x$, which is called by the latter symbolic notation also the *product* of $\mathcal{T}_1$ and $\mathcal{T}_2$.² This

¹ For $p = 1$ the term "*single-valued*", for $p > 1$ the denomination "*multiple-valued*" is also used.

² It is possible to use the symbol $\mathcal{T}_2 \circ \mathcal{T}_1$, too. In case of identical factors, we apply naturally the power notation and write: $\mathcal{T} \circ \mathcal{T} = \mathcal{T}^2$, $\mathcal{T} \circ \mathcal{T} \circ \mathcal{T} = \mathcal{T}^3$, etc.

“operator multiplication” is obviously not commutative ($\mathcal{T}_1\mathcal{T}_2 \neq \mathcal{T}_2\mathcal{T}_1$), but it is an associative operation. If a one-valued $\mathcal{T}$ has a one-valued inverse, then $\mathcal{T}\mathcal{T}^{-1} = \mathcal{T}^{-1}\mathcal{T} = \mathcal{I}$, where $\mathcal{I}$ denotes the *identity (unit) operator*.

Like 1.4.2, we can easily verify the following result of Cantor:

3.1.1 *If X and Y are non-empty sets and Y consists of more than one elements, then the cardinality of operators of X into Y is higher than the power of X .*

3.12 *A few important classes of abstract functions (operators):*

I. The name of $y = \mathcal{T}[x]$ is ν -dimensional point function, if X consists of certain points of the Euclidean space E_ν , while the elements of Y are real numbers. (That is to say, the point functions are the scalar functions of one or several variables, known from elementary analysis.)

II. We get the various sorts of *vector functions*, if the elements of at least one of X and Y are (one or several dimensional) vectors. If both x and y are vectors and a homogeneous linear connection of these variables is considered, the simplest variant of the concept of *tensor* is obtained.

III. We speak about *set function*, if the elements of X are themselves point sets and those of Y are real numbers. In particular, the terms *interval function* and *line function* are used, if the role of x is played by a variable interval and curve arc, respectively. Examples: $\bar{\mu}_J(S)$, $\bar{\mu}_L(S)$, respectively, any definite integral or line integral in the case of a fixed integrand and varying path of integration.

IV. $y = \mathcal{T}[x]$ is called *functional* or *functional operator* if X consists of point or vector functions, or of systems of these, while Y is a number set. Thus any functional signifies the dependence of a real or complex variable upon a varying function or system of functions. Examples: $y = \sup_{a \leq t \leq b} \{x(t)\}$

where $x = x(t)$ runs over all the functions bounded, real-valued in $[a, b]$; or any integral taken on a fixed basic domain and with a varying integrand.

V. An integral $\mathcal{T}[x] = \int_a^b x(t)\rho(t) dt$ with fixed (a, b) and a given bounded function $\rho(t)$ considered as a functional of the varying bounded function $x(t)$, possesses the properties of *additivity* and *homogeneity*:

$$\mathcal{T}[x_1 + x_2] = \mathcal{T}[x_1] + \mathcal{T}[x_2] \quad (3.1-1)$$

$$\mathcal{T}[cx] = c\mathcal{T}[x] \quad (c \text{ constant}); \quad (3.1-2)$$

here x_1, x_2 and x have to vary in the class X of those functions for which the occurring integrals exist. Besides, $\mathcal{T}[x]$ is *bounded* in the sense that $|\mathcal{T}[x]|$ can be estimated from above by means of $|x(t)|$:

$$|\mathcal{T}[x]| \leq M \sup_{a \leq t \leq b} \{|x(t)|\}, \quad (3.1-3)$$

M denoting a suitable constant (which depends of course on the choice of $\rho(t)$).

Since (3.1-1)—(3.1-2) are characteristic of the homogeneous linear point or vector functions, generally we say a functional $\mathcal{T}[x] = \mathcal{T}x$ to be *linear*, if we have for each pair of elements $x_1 \in X$, $x_2 \in X$ and for arbitrary number pairs c_1, c_2 the relation

$$\mathcal{T}[c_1x_1 + c_2x_2] = c_1\mathcal{T}[x_1] + c_2\mathcal{T}[x_2]; \quad (3.1-4)$$

moreover the boundedness of $\mathcal{T}[x]$ is defined by

$$|\mathcal{T}[x]| \leq M \cdot \|x\| \quad (\forall x \in X) \quad (3.1-5)$$

where M means a suitable positive number, and $\|x\|$ a functional with befitting special properties.

More precisely, it is required from the latter — on the model of the right-hand term of (3.1-3) — that the conditions

$$\|x\| \geq 0, \|x_1 + x_2\| \leq \|x_1\| + \|x_2\|, \|cx\| = |c| \|x\| \quad (c \text{ const.}) \quad (3.1-6)$$

are always satisfied in X and that $\|x\| = 0$ holds if and only if x equals to the so-called *zero element* of X . So $\|x\|$ may be regarded as an extension of the concept of absolute value (modulus); it is called the *norm* of x .

If it is possible to introduce a norm for all elements of a set X , we say that X is *normed*.

VI. We use for $y = \mathcal{T}[x]$ mostly the term *functional transformation*, if the operator transfers a point or vector function into another point or vector function, so that both X and Y consist of functions in the ordinary sense. Classical examples of this type of operator are: the homogeneous linear *coordinate transformation* (written in vectorial form); furthermore, the $\mathcal{T} = d/dt$ *differential* or the $\mathcal{T} = \int \dots dt$ *integral operator*. All three of them are additive and homogeneous in the above sense, but boundedness can come up at the first only.

3.13 On this basis it is expedient to develop further the notion of linear functional: any additive and homogeneous abstract function, i.e., which satisfies the condition (3.1-4), is called a *linear operator*.²

In the special case when both the domain X and the co-domain of a linear operator $\mathcal{T}$ are normed and we have the inequality [as a pendant of (3.1-5)]

$$\|\mathcal{T}[x]\| \leq M \cdot \|x\| \quad (\forall x \in X) \quad (3.1-7)$$

¹ Here for brevity we include — according to recent conventions of the literature — the homogeneity in the linearity; further we assume that X also contains together with however many of its elements, the linear combinations of the latter, i.e., that X is a *linear manifold*.

² The denomination “*linear transformation*” is also used in the literature.

with a suitable constant $M > 0$, we say that $\mathcal{T}$ is *bounded*. The greatest lower bound of all the numbers M in question is called simply the *norm* of $\mathcal{T}$ and it is denoted by $\|\mathcal{T}\|$.

Two relevant kinds of bounded linear operators are given by the homogeneous linear coordinate transformations (mentioned before) and by their "continuous" analogues, the *integral transformations* of the form

$$y(s) = \int_a^b K(s, t) x(t) dt, \quad (3.1-8)$$

where $K(s, t)$ is the so-called *kernel function*.

Next we want to deal mainly with point functions, since the case of vector functions can be reduced to these; beside short surveys of the pertinent topics (on a higher level), it will also be possible to intensify the theory in many directions. Meanwhile we shall still investigate certain set functions — above all in connection with modern integral concepts.

As regards the general discussion of linear functionals and operators, respectively, these are parts of *functional analysis*, which will be discussed later.

3.2 Limit and cluster value of a point function, relating to a set

3.21 Let S be a given point set in E_ν and $P(x_1, x_2, \dots, x_\nu)$ an arbitrary point of S ; we shall study a *one-valued point function* $f(P)$, defined in S^1 which is at the same time a real-valued function of the real variables $x_1, x_2, \dots, x_\nu$. As is known, $f(P)$ is said to be *bounded* or especially *bounded from above* and *from below*, resp., on S , if the number set $\{f(P)\} \ (P \in S)$ possesses the corresponding property (cf. 2.12).

Consider an accumulation point P_0 of S (which must not belong itself to S); and denote by S_0 the set consisting of all the points of S different from P_0 .² We say that $f(P)$ has a *limit at P_0 relating to S_0* and it equals to a number A ; by symbols:

$$\lim_{P \rightarrow P_0} f(P) = A \quad (P \in S_0) \quad \text{or} \quad f(P) \rightarrow A \quad (P \rightarrow P_0; P \in S_0), \quad (3.2-1)$$

if for *every* point sequence $P_1, P_2, \dots$ with $P_n \in S_0$ and $P_n \rightarrow P_0 \ (n \rightarrow \infty)$, the sequence of the corresponding values $f(P_1), f(P_2), \dots$ tends to A . The latter condition is plainly equivalent to the following one: to any $\varepsilon > 0$ we

¹ This expression also always means the finiteness of the function values in question.

² Thus $P_0 \notin S_0$ but $P_0 \in S'_0 = S'$. (Cf. 2.34.)

can find a neighbourhood of P_0 such that for each point $P \in S_0$ lying in it we have

$$|f(P) - A| < \varepsilon. \quad (3.2-2)$$

The first version of the definition is valid in the case of $A = \pm\infty$, too; but in the second version, ε is replaced by an arbitrarily large number $G > 0$ for $A = \pm\infty$, and instead of (3.2-2) we have to write $f(P) \geq G$ or $f(P) \leq -G$ according to the cases $A = +\infty$ and $A = -\infty$, respectively.

We can say that the existence and value of $\lim_{P \rightarrow P_0} f(P)$ ($P \in S_0$) depends essentially upon the basic set S . For instance, the function $((x)) = x - [x]$ (the so-called "fractional part of x ")¹ has no limit at the origin if S is an ordinary (two-sided) linear neighbourhood of the point 0; but $((x)) \rightarrow 0$ and $((x)) \rightarrow 1$, if x tends to 0 over the points of an interval $(0, \delta)$ and $(-\delta, 0)$, respectively (i.e., on the right or on the left), separately. More generally: if we consider the limit of a function f of one real variable relating to the (open) right-hand or left-hand neighbourhood of a point a , we obtain the concept of *right-hand or left-hand limit* of f at a , denoted by

$$\lim_{x \rightarrow a+0} f(x) = f(a+0), \quad \lim_{x \rightarrow a-0} f(x) = f(a-0) \quad (3.2-3)$$

in elementary analysis.

3.22 Like the sequences of points (see 2.22), we can also introduce for point functions such natural extensions of the limit as the concept of cluster value, lower or upper limit. Preserving the above meaning of S , P_0 and S_0 , let us start with the following definition: we say that $f(P)$ has at P_0 the *cluster value* A (or $\pm\infty$) *relating to* S_0 , if we can select from the set S_0 at least one point sequence $\{P_n\}$ tending to P_0 for which $\{f(P_n)\}$ tends to A (or to $\pm\infty$). The term *sublimit* instead of cluster value is also used.

Since every sequence of function values contains a subsequence tending to a finite value, to $+\infty$ or $-\infty$ (cf. 2.2.2), so at least one (finite or infinite) cluster value of $f(P)$ belongs to any point $P_0 \in S'$. In general, several (maybe infinitely many) sublimits can be attached to the same point P_0 , and it is easy to see that we can always find among them a maximal and a minimal one, in a wider sense.² The former is called the *upper limit* (or limit superior), the latter is the so-called *lower limit* (or limit inferior) of $f(P)$ at point P_0 relating to the set S_0 . The corresponding notations are:

$$\overline{\lim}_{P \rightarrow P_0} f(P) \quad \text{or} \quad \bar{f}(P_0) \quad (P \in S_0), \quad (3.2-4)$$

¹ In addition to $((x))$, the notation $\{x\}$ is used.

² Namely the least upper and the greatest lower bound of the set of sublimits at P_0 .

$$\lim_{P \rightarrow P_0} f(P) \quad \text{or} \quad \underline{f}(P_0) \quad (P \in S_0).$$

Note that the above definitions imply:

$$-\infty \leq \underline{f}(P_0) \leq \bar{f}(P_0) \leq \infty. \quad (3.2-5)$$

Some examples: the functions I. $\text{sgn } x$; II. $((x))$; III. $1/x$; IV. $\sin(1/x)$; V. $x \sin(1/x)$; VI. $\chi(x)$ [cf. (1.6-1)] have the following upper and lower limit at the origin relating to a "punctured" ε -neighbourhood of this point (i.e., to such one which is deprived of 0 itself). The values $\bar{f}(0)$, $\underline{f}(0)$ are for I. 1 and -1 ; for II. 1 and 0; for III. $+\infty$ and $-\infty$; for IV. 1 and -1 ; for V. 0 and 0; for VI. 1 and 0.

The number of cluster values at the origin is for function V. one, for I-III. and VI. two, but for IV. "infinite". For the function V., 0 is also limit in the classical sense; but for IV., each number $\xi \in [-1, 1]$ is sublimit at the same time.

Hence we see that $\bar{f}(P_0)$ and $\underline{f}(P_0)$ determine, so to say, the extreme bounds between which the values in question of $f(P)$ can "fluctuate" as $P \rightarrow P_0$. In order to interpret the situation more exactly, we need to premise with specifying terms.

3.23 Let s be any subset of the domain of $f(P)$ and let $P_0 \in \bar{s}$;¹ consider those points of s which belong to the ε -neighbourhood τ_ε of P_0 . Let $M_\varepsilon(P_0)$ and $m_\varepsilon(P_0)$ denote the least upper bound and greatest lower bound, respectively, of the function values $f(P)$ taken on the set $s_\varepsilon = s \cap \tau_\varepsilon$ just mentioned; then the limits

$$M(P_0) = \lim_{\varepsilon \rightarrow +0} M_\varepsilon(P_0), \quad m(P_0) = \lim_{\varepsilon \rightarrow +0} m_\varepsilon(P_0) \quad (3.2-6)$$

certainly exist in a wider sense (i.e., $\pm\infty$ being also permitted). In fact, both $M_\varepsilon(P_0)$ and $m_\varepsilon(P_0)$ behave for $\varepsilon \rightarrow +0$ monotonically, namely, the former is non-increasing, the latter non-decreasing, because s_ε gets ever "tighter" when ε becomes smaller. $M(P_0)$ and $m(P_0)$ are called the *upper and lower fluctuation value* of $f(P)$ at the point P_0 , for the basic set s .

Now we can formulate the following lemma:

3.2.1 *The basic set S_0 being fixed, $\bar{f}(P_0)$ and $\underline{f}(P_0)$ are identical with the upper and lower fluctuation value of $f(P)$ at P_0 , respectively.*

P r o o f. Let us consider the above $M_\varepsilon(P_0)$ with $s = S_0$; by definition, any left-hand side neighbourhood of this upper bound contains a function

¹ Here $\bar{s}$ denotes the closure (adherence) of the set s . [Cf. (2.3-2)].

value $f(P)$ taken on the set $S_{0,\varepsilon} = S_0 \cap \tau_\varepsilon$. Therefore, if $\varepsilon > 0$ runs over a sequence $\{\varepsilon_n\}$ tending to 0, we can attach to ε_n a point $P_n \in S_{0,\varepsilon_n}$ such that $f(P_n) \rightarrow \lim_{n \rightarrow \infty} M_{\varepsilon_n}(P_0) = M(P_0)$. Since $\lim_{n \rightarrow \infty} P_n = P_0$, we obtain that $M(P_0)$ is a cluster value at P_0 relating to S_0 . On the other hand, for $M(P_0) < \infty$ any number $M(P_0) + \eta$ ($\eta > 0$) is an upper bound of the function values taken in $S_{0,\varepsilon}$ provided that ε is chosen sufficiently small [cf. (3.2-6)]; this excludes that a given number $B > M(P_0)$ would be another cluster value of $f(P)$ ($P \in S_0$) at point P_0 .¹

Summing up, the result is that $M(P_0)$ with the basic set S_0 cannot differ from $\lim_{P \rightarrow P_0} f(P)$ ($P \in S_0$).

The equality $m(P_0) = \underline{f}(P_0)$ (relating also to S_0) can be verified similarly.

* * *

It is to be stressed that 3.2.1 does not hold generally, if in the case of $P_0 \in S$ the point P_0 itself is included in the basic set of $M(P_0)$ and $m(P_0)$ ($s = S$); then we have obviously

$$\begin{cases} M(P_0) = \max[f(P_0), \bar{f}(P_0)], \\ m(P_0) = \min[f(P_0), \underline{f}(P_0)] \end{cases} \quad (3.2-7)$$

and so

$$M(P_0) \geq \bar{f}(P_0), \quad m(P_0) \leq \underline{f}(P_0). \quad (3.2-8)$$

For instance, for the function

$$g(x) = \begin{cases} 0 & (x \neq 0) \\ 1 & (x = 0) \end{cases}$$

we get $M(0) = 1$ relating to a (complete) neighbourhood of the origin; hence $M(0) > \bar{g}(0) = 0$.

3.2.4 From the above considerations we can easily deduce the analogues of condition (2.2-2) as well as of Cauchy's criterion 2.2.1:

3.2.2 For the existence of a finite function limit $\lim_{P \rightarrow P_0} f(P)$ ($P \in S_0$) it is necessary and sufficient that

(I) $\bar{f}(P_0)$ and $\underline{f}(P_0)$ are finite and have the same value,² or

¹ Observe that, η being chosen in a suitable manner, we have $f(P) \leq M(P_0) + \eta < B$ in every point $P \in S_0$ provided that $\overline{PP_0}$ is small enough.

² This condition can also be expressed as $f(P)$ has one single and finite cluster value at P_0 with respect to S_0 .

(II) to any $\varepsilon > 0$ we can find a positive number $\delta = \delta(\varepsilon)$ such that

$$|f(P) - f(Q)| < \varepsilon \quad (3.2-9)$$

holds for each point pair $P \in S_0$, $Q \in S_0$ which belongs to the δ -neighbourhood of P_0 . The condition $\bar{f}(P_0) = \underline{f}(P_0)$ can be extended also to the case of infinite limits and it implies always that

$$\lim_{P \rightarrow P_0} f(P) = \bar{f}(P_0) = \underline{f}(P_0) \quad (P \in S_0). \quad (3.2-10)$$

P r o o f. We see by 3.2.1 e.g. for finite $\bar{f}(P_0)$, $\underline{f}(P_0)$:

1° Any interval $(\underline{f}(P_0) - \varepsilon, \bar{f}(P_0) + \varepsilon)$ contains all the values of $f(P)$ for which $P \in S_0$ and $\overline{PP_0}$ is sufficiently small; thus the existence of $\lim_{P \rightarrow P_0} f(P)$ ($P \in S_0$) is equivalent to the condition $\bar{f}(P_0) = \underline{f}(P_0)$.

2° If the (obviously necessary) condition (II) is satisfied, then to any given $\varepsilon > 0$, $P \in S_0$ and $Q \in S_0$ can be chosen in such a way that $\overline{PP_0}$ and $\overline{QP_0} < \delta(\varepsilon)$, furthermore all three of $|f(P) - \bar{f}(P_0)|$, $|f(Q) - \underline{f}(P_0)|$ and $|f(P) - f(Q)|$ are less than ε . But hence we get:

$$|\bar{f}(P_0) - \underline{f}(P_0)| < |\bar{f}(P_0) - f(P)| + |f(P) - f(Q)| + |f(Q) - \underline{f}(P_0)| < 3\varepsilon,$$

and — ε being as small as we like — it follows the equality $\bar{f}(P_0) = \underline{f}(P_0)$, i.e., the existence of $\lim_{P \rightarrow P_0} f(P)$ ($P \in S_0$).

* * *

We can state that: since the generalized limit (3.2-1) is defined by means of sequences, the elementary rules on the *limit of a sum, product, quotient* can be transferred at once.

In the sequel, each function limit at P_0 without giving explicitly a basic set is to be referred to a neighbourhood of P_0 deprived of P_0 itself. In considering the existence of such a limit, we include also its *finiteness*.

3.3 Properties of continuous and semi-continuous point functions

3.31 Closely related to the notion of limit is the well-known concept of continuity of a point function. According to the former, this will be developed further, too.

Let $f(P)$ be a point function defined on a set $S \subseteq E_n$, and let P_0 denote such an accumulation point of S which belongs also to S .¹ We say that $f(P)$

¹ As contrasted with the definition of limit, here we require that $f(P)$ should also be defined at the point P_0 .

is continuous at P_0 relating to S , if $\lim_{P \rightarrow P_0} f(P)$ ($P \in S$, $P \neq P_0$) exists and equals the (finite) function value $f(P_0)$. By the definition of limit, these conditions may also be expressed in the following forms:

1. for every sequence $P_n \in S$, $P_n \neq P_0$ ($n = 1, 2, \dots$) tending to P_0 we have $f(P_n) \rightarrow f(P_0)$ ($n \rightarrow \infty$); or
2. to any number $\varepsilon > 0$, however small, we can find a neighbourhood of P_0 such that for each point $P \in S$, $P \neq P_0$

$$|f(P) - f(P_0)| < \varepsilon \quad (3.3-1)$$

holds.

For the sake of simpler wording, the restriction that P_0 should be an accumulation point and, accordingly, also the condition $P \neq P_0$ in connection with (3.3-1), is sometimes omitted. In such a sense, we can say that $f(P)$ is continuous at any isolated point of S , because P_0 being the only point of S in a sufficiently small neighbourhood of the latter, (3.3-1) is fulfilled trivially. If f is a function of one variable and we take a one-sided neighbourhood of a point as a basic set, the well-known notions of *right-hand* and *left-hand* continuity arise.

At such a point P_0 , where $\overline{f}(P_0) \neq \underline{f}(P_0)$, the limit of $f(P)$ ($P \in S$) cannot exist and therefore $f(P)$ is certainly not continuous. Nevertheless it often occurs that at some point P_0 either the condition

$$\overline{\lim}_{P \rightarrow P_0} f(P) = f(P_0) \quad (P \in S, P \neq P_0) \quad (3.3-2)$$

or its pendant

$$\underline{\lim}_{P \rightarrow P_0} f(P) = f(P_0) \quad (P \in S, P \neq P_0) \quad (3.3-3)$$

is satisfied separately; then $f(P)$ is said to be *upper* or *lower semicontinuous* at P_0 relating to S .

Examples: among the functions mentioned after (3.2-5), I., III., IV., V. are neither upper nor lower semi-continuous, but II. is lower, VI. upper semicontinuous at the origin, with respect to a neighbourhood of this point. We note also: function V. becomes continuous at the origin, if there the value of the function is taken as 0; III. can be made semi-continuous at point 0, if $+\infty$ or $-\infty$ are also allowed as function values; finally, VI. is for any rational x upper, for any irrational x lower semi-continuous.

3.32 The above definitions and (3.2-7) yield immediately the following two criteria of continuity:

3.3.1 A function $f(P)$ defined on a set S is continuous at a point $P_0 \in S$ if and only if (I) $f(P)$ is at P_0 both upper and lower semi-

continuous; or (II) the upper and lower fluctuation value of $f(P)$ at P_0 (relating to S) have the same value.¹

As far as the fundamental properties of continuous functions are concerned, first of all we want to stress: by the formulae on the limit of a sum, product, quotient, the continuity of $f_1(P)$ and $f_2(P)$ at a common point P_0 relating to S implies also the continuity at P_0 of

$$f_1(P) + f_2(P), \quad f_1(P) \cdot f_2(P), \quad f_1(P)/f_2(P),$$

relating to the same basic set — with the restriction $f_2(P) \neq 0$ in the last case. Since the assertion can be extended at once to the cases of sums or products of more than two terms or factors, we may state that the application finitely many times of the four basic operations does not take out of the class of continuous functions, provided that the functions, figuring as divisors do not vanish. (Cf. 1.51.)

3.33 As any local property, the continuity or semi-continuity can also be attributed to a function at every point of some set S , relating to S as a common basic set; then $f(P)$ is said to be *continuous* and *semi-continuous*, respectively, on S . The pertinent generalizations of theorem due to Weierstrass, Bolzano and Heine, known from elementary analysis, can be easily verified:

3.3.2 Let S be a bounded closed set and $f(P)$ a function defined on S . If $f(P)$ is upper semi-continuous on S , then there exists a maximal value, and if $f(P)$ is lower semi-continuous on S , so there exists a minimal value in the co-domain of f .²

3.3.3 If $f(P)$ is continuous in a closed region (continuum) S , then it takes every value between its maximum and minimum on S .

3.3.4 Any function $f(P)$ continuous on a bounded closed set S is there also uniformly continuous.

The meaning of the term *uniform continuity* is: given any $\varepsilon > 0$, we can find a number $\delta = \delta(\varepsilon) > 0$ such that for each pair of points with $P \in S$, $Q \in S$, $\overline{PQ} < \delta$ we have $|f(P) - f(Q)| < \varepsilon$. As an illustration, we now present a simple proof of 3.3.4, which is based upon Borel's covering theorem.

Let $\varepsilon > 0$ be fixed. $f(P)$ being continuous on S , we can attach to any point $P_0 \in S$ such a δ -neighbourhood τ_δ (depending on P_0), that for any point $P \in S$ belonging to τ_δ the inequality $|f(P) - f(P_0)| < \varepsilon/2$ (*) holds.

¹ This number cannot be other than $f(P_0)$.

² The content of this theorem is expressed sometimes by saying that any function upper (lower) semi-continuous on a bounded closed set *attains* the least upper (greatest lower) bound of its co-domain.

Consider the $(\delta/2)$ -neighbourhood of an arbitrary point of S ; the totality of these sets covers S and therefore, by 2.6.1, we can find a *finite* system of them having the same property. Let us denote the centre of the k -th element of the last-mentioned system by O_k , its radius by $\delta_k/2$ ($k = 1, 2, \dots, r$), and write $\delta_0 = \min(\delta_k/2)$. If we choose in S two points P, Q for which $\overline{PQ} < \delta_0$, then it holds obviously $\overline{O_k P} < \delta_k/2$ for a suitable index k ; on the other hand, (2.1-2) yields $\overline{O_k Q} \leq \overline{O_k P} + \overline{PQ} < (\delta_k/2) + \delta_0 \leq \delta_k$.

We see that the above inequality (*) is applicable to both of the point pairs $\{O_k, P\}$ and $\{O_k, Q\}$; thus we obtain:

$$|f(P) - f(Q)| \leq |f(P) - f(O_k)| + |f(O_k) - f(Q)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

which implies the assertion.

3.4 The approximation theorems of Weierstrass; Stone's theorem

3.4.1 We know that Heine's theorem 3.3.4 forms the basis of a fact of great practical importance: the graph of a function of one variable can be *approximated uniformly* on every closed interval of continuity by a suitable broken line, i.e., by such a continuous function which is piecewise linear.

Next we have to remember some deeper problems of classical analysis, namely the approximability of a continuous function $f(x)$ by an *entire rational function* (polynomial)

$$R_N(x) = C_0 + C_1x + \dots + c_Nx^N \quad (C_N \neq 0) \quad (3.4-1)$$

or by a so-called *trigonometric polynomial* of the form

$$S_N(x) = A_0 + \sum_{n=1}^N (A_n \cos nx + B_n \sin nx) \quad (A_N^2 + B_N^2 > 0) \quad (3.4-2)$$

with real coefficients. On this topic, two well-known theorems are valid:

3.4.1 (APPROXIMATION THEOREMS OF WEIERSTRASS)

(I) If $f(x)$ is continuous on a finite line segment $[a, b]$, then to every prescribed number $\varepsilon > 0$ we can find an entire rational function $R_N(x)$ for which

$$|f(x) - R_N(x)| < \varepsilon \quad (a \leq x \leq b). \quad (3.4-3)$$

(II) If $f(x)$ is an everywhere continuous function with period 2π , then there exists to any positive ε a trigonometric polynomial $S_N(x)$ such that

$$|f(x) - S_N(x)| < \varepsilon \quad (-\infty < x < \infty). \quad ^1 \quad (3.4-4)$$

We omit here direct proofs of 3.4.1 (I)–(II). On the one hand, it is expedient to treat now, instead of these classical results, a much more general theorem of M.H. STONE which has been discovered in 1947 and concerns the case of however many variables. On the other hand, an investigation of trigonometric series will later yield not only a very simple verification but a kind of sharpening, too. (Cf. 13.33.)

3.42 In order to shorten the formulation of Stone's extended result, we introduce the following mode of expression: we say that a given point function f is *uniformly approximable* on a certain point set $S \subset E_\nu$ by the elements of a function class $\mathcal{Z}$, if to each number $\varepsilon > 0$ we can find a point function $\varphi \in \mathcal{Z}$ for which

$$|f - \varphi| < \varepsilon \quad (3.4-5)$$

at every point of S .

We want to discuss:

3.4.2 (STONE'S THEOREM)² Let S be a bounded closed set in E_ν , and let $\mathcal{Z}_\varphi$ denote a class of real-valued, on S continuous functions φ which satisfies the following conditions:

(I) $\mathcal{Z}_\varphi$ contains together with any of its elements also the product of the latter by an arbitrary constant, further, together with any two of its elements also the sum and product of these.

(II) To any two distinct points $P_1 \in S$, $P_2 \in S$, we can find such an element $\varphi(P) \in \mathcal{Z}_\varphi$ for which $\varphi(P_1) \neq \varphi(P_2)$.

(III) To any point $P_0 \in S$ there exists an element of $\mathcal{Z}_\varphi$ which does not vanish at P_0 .

Then every real-valued, on S continuous function is uniformly approximable on this set by the elements of $\mathcal{Z}_\varphi$.

P r o o f. Let us denote by $\overline{\mathcal{Z}}_\psi$ the totality of continuous, real-valued functions $\psi(P)$ ($P \in S$) which are uniformly approximable on S by the elements of $\mathcal{Z}_\varphi$. By (I) we can see that $\psi \in \overline{\mathcal{Z}}_\psi$ implies $\alpha\psi \in \overline{\mathcal{Z}}_\psi$, where α is an arbitrary real number, furthermore in case of $\psi_1 \in \overline{\mathcal{Z}}_\psi$, $\psi_2 \in \overline{\mathcal{Z}}_\psi$ we have also $\psi_1 + \psi_2 \in \overline{\mathcal{Z}}_\psi$ and $\psi_1 \psi_2 \in \overline{\mathcal{Z}}_\psi$. (*)

We wish to show that $\overline{\mathcal{Z}}_\psi$ contains *all* the functions continuous and real-valued on the set S . This will be carried out in five steps.

¹ In other words: on our conditions, the function $f(x)$ can be approximated uniformly by $R_n(x)$ in $[a, b]$ or by $S_N(x)$ over $(-\infty, \infty)$ with an arbitrarily small error. Of course, in the case of a lower ε the increase of the index N (i.e., of the degree of $R_N(x)$, or of the order of $S_N(x)$) is needed.

² See Stone [1]

1° We begin by verifying that $\bar{\mathcal{Z}}_\psi$ contains together with any of its elements the modulus of the latter, too.

Assume that $\psi(P) \in \bar{\mathcal{Z}}_\psi$, and consider $M = \max |\psi(P)|$ relating to S whose existence is assured by 3.3.2. If $M = 0$, then also $\psi(P) = 0$ ($P \in S$), i.e., $|\psi(P)| = \psi(P) \in \bar{\mathcal{Z}}_\psi$. In turn, if $M > 0$, we have to form the auxiliary function $\Psi(P) = M^{-1} \psi(P)$ for which $\max_{P \in S} |\Psi(P)| = 1$ and which belongs to $\bar{\mathcal{Z}}_\psi$ by (*).

Studying $|\Psi|$, let us start from the trivial inequality

$$(\Psi^2 + \eta^2)^{1/2} - |\Psi| \leq \eta \quad (\eta > 0 \text{ however small}),$$

and use the expansion of $(x + \eta^2)^{1/2}$ in terms of the powers of $x - 1$, which converges uniformly in the interval $0 \leq x \leq 1$. Thus we can write for a suitable partial sum $r_\eta(x)$ of the Taylor series in question:

$|r_\eta(x) - (x + \eta^2)^{1/2}| < \eta$ ($0 \leq x \leq 1$); and especially $|r_\eta(0) - \eta| < \eta$, $|r_\eta(0)| < 2\eta$. Consequently we obtain:

$$\begin{aligned} |r_\eta(t^2) - r_\eta(0) - |t|| &\leq |r_\eta(t^2) - r_\eta(0) - (t^2 + \eta^2)^{1/2}| + |(t^2 + \eta^2)^{1/2} - |t|| \leq \\ &\leq |r_\eta(t^2) - (t^2 + \eta^2)^{1/2}| + |r_\eta(0)| + \eta < \eta + 2\eta + \eta = 4\eta \quad (|t| \leq 1). \end{aligned}$$

If we put here $t = \Psi(P)$, it follows that $|\Psi(P)|$ can be approximated uniformly on the set S by a polynomial without the constant term of $\Psi(P)$, in such a way that the approximation error is less than 4η . Since a polynomial of the mentioned type belongs clearly to $\bar{\mathcal{Z}}_\psi$, the same holds about $|\Psi(P)|$.

But $|\psi(P)| = M |\Psi(P)|$ so that we get $|\psi(P)| \in \bar{\mathcal{Z}}_\psi$ ($P \in S$), as asserted before.

2° On this basis, it is easy to see that in the case of $\psi_1(P) \in \bar{\mathcal{Z}}_\psi$, $\psi_2(P) \in \bar{\mathcal{Z}}_\psi, \dots, \psi_n(P) \in \bar{\mathcal{Z}}_\psi$ the functions

$$M_n(P) = \max[\psi_1(P), \dots, \psi_n(P)], \quad m_n(P) = \min[\psi_1(P), \dots, \psi_n(P)]$$

also belong to the class $\bar{\mathcal{Z}}_\psi$.

In fact: if we apply again the above statement (*) as well as the previous result 1°, then the relationships (3.4-6)

$$\max(\psi_1, \psi_2) = \frac{1}{2}(\psi_1 + \psi_2 + |\psi_1 - \psi_2|), \quad \min(\psi_1, \psi_2) = \frac{1}{2}(\psi_1 + \psi_2 - |\psi_1 - \psi_2|) \quad (3.4-6)$$

imply at once our assertion for $n = 2$; and hence the general result arises simply by induction.

3° We shall still need the following property of $\mathcal{Z}_\varphi$: to any two distinct points P_1, P_2 of S and to any prescribed pair of real numbers α_1, α_2 we can find a function $\Phi(P) \in \mathcal{Z}_\varphi$ such that $\Phi(P_1) = \alpha_1$, $\Phi(P_2) = \alpha_2$.

Namely, condition (II) of the theorem assures the existence of such a $\varphi(P) \in \mathcal{Z}_\varphi$ for which $\varphi(P_1) \neq \varphi(P_2)$. If the latter function value, e.g. $\varphi(P_1)$ is zero [so that $\varphi(P_2) \neq 0$], then let us select from $\mathcal{Z}_\varphi$ another element, say φ^* such that $\varphi^*(P_1) \neq 0$ [this is made possible by condition (III)]. Thereafter consider the function $\varphi^{**}(P) = \varphi(P) + c\varphi^*(P)$ where c is a real parameter. It is clear that for a suitable value of c we have simultaneously $\varphi^{**}(P_1) \neq 0$, $\varphi^{**}(P_2) \neq 0$ and $\varphi^{**}(P_1) \neq \varphi^{**}(P_2)$. Therefore, there certainly exists a function $\varphi_0(P) \in \mathcal{Z}_\varphi$ which vanishes neither at P_1 nor at P_2 and takes different values at these points.

Put now $\Phi(P) = \lambda\varphi_0(P) + \mu[\varphi_0(P)]^2$ where λ, μ are real parameters. Since the determinant of the system of equations

$$\begin{cases} \Phi(P_1) = \lambda\varphi_0(P_1) + \mu[\varphi_0(P_1)]^2 = \alpha_1, \\ \Phi(P_2) = \lambda\varphi_0(P_2) + \mu[\varphi_0(P_2)]^2 = \alpha_2 \end{cases}$$

is $\neq 0$ by the above, λ and μ can be determined in such a way that $\Phi(P)$ should have the property in question.

4° After these preparations, let us consider an arbitrary function $f(P)$ which is continuous and real-valued on S , and let a number $\varepsilon > 0$ be given. By step 3°, we can attach to any point pair $p \neq q$ belonging to S such a function $\Phi_{pq}(P) \in \mathcal{Z}_\varphi$ for which $\Phi_{pq}(p) = f(p)$, $\Phi_{pq}(q) = f(q)$. Let S_{pq} denote that subset of S on which the value of f differs from the corresponding value of Φ_{pq} by less than $\varepsilon/2$, i.e.,

$$S_{pq} = \left\{ P : P \in S, f(P) - \frac{\varepsilon}{2} < \Phi_{pq}(P) < f(P) + \frac{\varepsilon}{2} \right\}. \quad (3.4-7)$$

Putting here $P = p$ resp. $P = q$ we see that also the points p, q belong to S_{pq} . Our nearest object is to present an inequality of the form:

$$\Phi_q(P) - f(P) < \frac{\varepsilon}{2} \quad (P \in S) \quad (3.4-8)$$

where $\Phi_q(P)$ means a certain "minimum function", to be constructed below by means of the auxiliary function $\Phi_{pq}(P)$ with fixed q .

For this, we remark first: since $\Phi_{pq}(P) - f(P)$ is a continuous point function, so each point $P \in S_{pq}$ has an (open) neighbourhood $\tau(P)$ such that $\tau(P) \cap S \subseteq S_{pq} (*)$.

Fixing a point $q \in S$, to any other point p of S let us attach a neighbourhood $\tau(p)$ of the just mentioned type. Because the totality of these ν -dimensional spheres covers the set S which has been supposed to be bounded and closed, by 2.6.1 there is a finite subset of $\{\tau(p)\}$, say $\tau(p_1), \tau(p_2), \dots, \tau(p_{n_q})$, which also covers S . Therefore we have

$$\bigcup_{1 \leq i \leq n_q} \tau(P_i) \supseteq S \text{ and } \bigcup_{1 \leq j \leq n_q} S_{p_j, q}(P) = S \text{ by } (*).$$

Let now introduce the notation

$$\Phi_q(P) = \min [\Phi_{p_1,q}(P), \dots, \Phi_{p_{n_q},q}(P)] \quad (P \in S). \quad (3.4-9)$$

Since any point p of S belongs to at least one neighbourhood $\tau(p_i)$, the corresponding function value $\Phi_{p_i,q}(P)$ is less than $f(P) + (\varepsilon/2)$ according to (3.4-7); and this implies by (3.4-9) the inequality (3.4-8) at once. It should be stressed that by 2° $\Phi_q(P) \in \bar{\mathcal{Z}}_\psi$, namely any function $\Phi_{p_i,q}$ is an element of $\mathcal{Z}_\varphi$ and so of $\bar{\mathcal{Z}}_\psi$, too.

5° Finally, let us carry out the above construction to each point $q \in S$ and consider the intersection

$$s_q = S_{p_1,q} \cap S_{p_2,q} \cap \dots \cap S_{p_{n_q},q}$$

which contains the point q . Thus there is an (open) neighbourhood $\tau^*(q)$ of q for which $\tau^*(q) \cap S \subseteq s_q$. The totality of all neighbourhoods $\tau^*(q)$ in question covers S , and Borel's covering theorem can again be applied. In such a way, we get a system $\tau^*(q_1), \tau^*(q_2), \dots, \tau^*(q_l)$ of ν -dimensional open spheres for which $\bigcup_{1 \leq i \leq l} \tau^*(q_i) \supseteq S$ and at the same time, $\bigcup_{1 \leq j \leq l} s_{q_j} = S$ holds.

Like the preceding argument, consider next the "maximum function"

$$\Phi(P) = \max [\Phi_{q_1}(P), \dots, \Phi_{q_l}(P)] \quad (P \in S). \quad (3.4-10)$$

The inequality (3.4-8) being valid also with indices $q_1, q_2, \dots, q_l$, we can write

$$\Phi(P) < f(P) + \frac{\varepsilon}{2} \quad (P \in S). \quad (3.4-11)$$

On the other hand, any point $P \in S$ is contained by a set s_{q_j} and thus by the corresponding sets S_{p_i,q_j} ($i = 1, 2, \dots, n_{q_j}$). But at points of the latter, we have the inequalities (3.4-7); therefore we obtain $\Phi_{p_i,q_j}(P) > f(P) - \frac{\varepsilon}{2}$ ($i = 1, 2, \dots, n_{q_j}$) and hence by (3.4-9) — (3.4-10) also $\Phi_{q_j}(P) > f(P) - \frac{\varepsilon}{2}$ as well as

$$\Phi(P) > f(P) - \frac{\varepsilon}{2} \quad (P \in S). \quad (3.4-12)$$

Comparing (3.4-11) and (3.4-12), we find that the function (3.4-10) approximates $f(P)$ on the set S with an error less than $\varepsilon/2$. Since $\Phi(P)$, together with all the functions $\Phi_q(P)$ belongs to the class $\bar{\mathcal{Z}}_\psi$ (cf. 2°) and so it can be approximated on S by some function $\varphi(P) \in \mathcal{Z}_\varphi$ with an error $< \varepsilon/2$, too, it follows

$$|\varphi(P) - f(P)| < \varepsilon \quad (P \in S), \quad (3.4-13)$$

as required.

* * *

3.43 The most important special cases of Stone's theorem are:

1. Let S be a ν -dimensional finite closed interval and $\mathcal{Z}_\varphi$ the aggregate of the entire rational functions (polynomials) of ν variables, with real coefficients. Such a polynomial of degree N can be written in the form:

$$R_N(P) = \sum C_{n_1, n_2, \dots, n_\nu} x_1^{n_1} x_2^{n_2} \dots x_\nu^{n_\nu} \quad (3.4-14)$$

$[N = \max(n_1 + n_2 + \dots + n_\nu)]$ and obviously fulfils the conditions 3.4.2 (I)-(III).

As can be seen, we have verified by the above not only the *I. approximation theorem of Weierstrass*, but also its pendant for spaces of several dimensions.

2. Let $\nu = 2$, choose for S the unit circle $x^2 + y^2 = 1$, and let us characterize an arbitrary point P of the latter by the amplitude (polar angle) $\vartheta = \arctan(y/x)$. If we take as elements of $\mathcal{Z}_\varphi$ the trigonometric polynomials of order N , i.e., the expressions¹

$$T_N(\vartheta) = \sum_{n=0}^N (A_n \cos n\vartheta + B_n \sin n\vartheta) \quad (A_N^2 + B_N^2 > 0), \quad (3.4-15)$$

then conditions (II) and (III) of theorem 3.4.2 are plainly satisfied, while (I) follows easily from the identities (valid for arbitrary integers m, n):

$$\begin{cases} \cos m\vartheta \cdot \cos n\vartheta = \frac{1}{2} [\cos(m+n)\vartheta + \cos(m-n)\vartheta] \\ \sin m\vartheta \cdot \sin n\vartheta = \frac{1}{2} [\cos(m-n)\vartheta - \cos(m+n)\vartheta] \\ \sin m\vartheta \cdot \cos n\vartheta = \frac{1}{2} [\sin(m+n)\vartheta + \sin(m-n)\vartheta] \end{cases} \quad (3.4-16)$$

as well as by induction.

Considering that a continuous function f of ϑ has necessarily the period 2π and its codomain is determined completely by the function values belonging to $0 \leq \vartheta \leq 2\pi$, 3.4.2 yields in this case the *II. approximation theorem of Weierstrass*.

3.5 Sequences and series of continuous point functions

3.51 As is well-known, we mean by the (local) *convergence* at a given point P_0 of some *sequence of functions* $\{f_n(P)\}$ or *series of functions* $\sum f_n(P)$ the convergence of the number sequence $\{f_n(P)\}$ and that of the numerical series $\sum f_n(P_0)$, respectively. The (ordinary) convergence on a point set S means the convergence at all the points of S , simultaneously.

¹ Note that $T_N(\vartheta)$ may also be interpreted as a polynomial of the variables $x_1 = \cos \vartheta$, $x_2 = \sin \vartheta$ in the (3.4-14) sense.

The concept of *uniform convergence* is cognate in a certain sense to the uniform continuity. We say that a sequence of functions $f_1(P), f_2(P), \dots$ tends (converges) uniformly on a set S to the *limit* (limit function) $f(P)$, if there exists to every given $\varepsilon > 0$ a number N (depending upon ε , but independent of P) such that for each $P \in S$ and $n \geq N$ the inequality

$$|f_n(P) - f(P)| < \varepsilon \quad (3.5-1)$$

holds. The uniform convergence of a series of functions $f_1(P) + f_2(P) + \dots$ on S means the uniform convergence of the sequence of partial sums

$$s_n(P) = f_1(P) + \dots + f_n(P) \quad (n = 1, 2, \dots) \quad (3.5-2)$$

in the just mentioned sense; here the role of the limit function is played by the *sum function* $s(P) = \lim_{n \rightarrow \infty} s_n(P)$.

By Cauchy's criterion 2.2.1, $\{f_n(P)\}$ is convergent on a set S if and only if to any given $\varepsilon > 0$ we can find such a number n_0 (common for all points of S) that

$$|f_m(P) - f_n(P)| < \varepsilon \quad (3.5-3)$$

for any pair of indices $m, n \geq n_0$ and at every point $P \in S$. On the other hand, we have the classical sufficient *criterion (test) of Weierstrass*: $\sum f_n(P)$ is uniformly convergent on a set S , if it can be majorized there by a convergent numerical series.¹

3.52 We reiterate that the notion of uniform convergence — which means of course a stronger condition than convergence at all the points of the basic set — is a useful expedient if term by term limiting processes are investigated, and in connection with the problem how a common property of the terms is transmissible to the limit or sum function, respectively. The simplest theorem of this kind concerns continuous functions:

3.5.1 If $\{f_n(P)\}$ converges uniformly on S to the limit function $f(P)$, then at every common point of continuity $P_0 \in S$ of the functions $f_n(P)$ ($n = 1, 2, \dots$), $f(P)$ is also continuous relating to S .

P r o o f. Let $\varepsilon > 0$ be fixed arbitrarily; by the concept of uniform convergence, there exists a number N such that at every point $P \in S$

$$|f_n(P) - f(P)| < \frac{\varepsilon}{3} \quad (n = N, N + 1, \dots). \quad (3.5-4)$$

If each of the terms $f_n(P)$ is continuous at P_0 (with respect to S), then we get especially for $n = N$ the inequality

$$|f_N(P) - f_N(P_0)| < \frac{\varepsilon}{3} \quad (3.5-5)$$

¹ In other words: if we can find such a series of positive constant terms, say $\sum \gamma_n$ that $|f_n(P)| \leq \gamma_n$ for n sufficiently large, at all points $P \in S$.

at any such point P of S which belongs to a sufficiently small neighbourhood τ of P_0 .

Thus we get by (3.5-4) — (3.5-5) on the set $S \cap \tau$

$$|f(P) - f(P_0)| \leq |f(P) - f_N(P)| + |f_N(P) - f_N(P_0)| + |f_N(P_0) - f(P_0)| < \varepsilon, \quad (3.5-6)$$

which means the continuity of $f(P)$ at P_0 (relating to S).

* * *

From this consideration we can see that the uniform convergence is not necessary in its classical form to the validity of 3.5.1; discussing the conditions of validity deeper, we obtain a weaker variant of this concept, the so-called *quasi-uniform convergence*. In turn, if we assume on $\{f_n(P)\}$ (beside continuity of the terms) monotony on S only, then the semi-continuity of $f(P)$ can be concluded.¹ Moreover, the pertinent theorems are invertible in a certain sense (BAIRE).

3.53 We add still that the last-mentioned statements can easily be rewritten for the case of series of functions, as well as that it is convenient in the so-called descriptive theory of functions to classify according to the representation possibility by iterated sequence forming and limiting processes (*Baire's and Young's classes of functions*).² Since every continuous function can be represented as a limit of a sequence of entire rational functions on the basis of the I. approximation theorem of Weierstrass (cf. 3.4.1), the Baire's classes contain all functions which are formed from the independent variables by applying finitely many times the basic arithmetical operations and the limiting process (the so-called "analytic operations" altogether).

Speaking about point functions, their domain are mostly open or closed regions (especially intervals), so that continuity relates also to these types

¹ More precisely: if the terms are continuous at P_0 , then $f(P)$ is at this point *upper* or *lower semi-continuous* according to the case of a monotone decreasing or increasing $\{f_n(P)\}$.

² Baire's classification uses arbitrary convergent function sequences, Young's procedure operates on monotonic sequences of functions. So — $f_n(x)$ denoting a continuous function — the first class of Baire consists of all the non-continuous functions which can be written in the form $\lim_{n \rightarrow \infty} f_n(x)$; the two kinds of Young's first class are those subsets of the former family of functions in which the monotony of $\{f_n(x)\}$ is also required. The functions constituting Baire's or Young's second class are obtained from the corresponding first classes by iterating the limiting process, etc. Examples: $\lim_{n \rightarrow \infty} x^n$ ($0 \leq x \leq 1$) belongs to the first class according to Baire or Young at the same time, but $\chi(x)$ [cf. (1.6-1)] is a function of Baire's second class.

See Baire [1], W.H. Young [3], L.C. Young [1].

of sets. In the sequel, if no basic sets are mentioned explicitly, pointwise continuity always concerns the complete neighbourhood of any point in consideration, too. (This is *continuity in narrower sense*.) In such a case, any point at which the function is not continuous will be called a *point of discontinuity*; for functions of one variable also the well-known terms "discontinuity of the first or second kind" are used (cf. 1.61).

3.6 Generalized differentiability, derivate numbers

3.61 The other most fundamental concept connected with the notion of limit is the differentiability. The version relating to an arbitrary basic set is the following:

Let $f(P) = f(x_1, x_2, \dots, x_\nu)$ be a function defined on a set $S \subseteq E_\nu$, $P_0(a_1, a_2, \dots, a_\nu)$ an accumulation point of S and S_0 the set obtained from S by removing P_0 . We say that $f(P)$ is (totally) *differentiable* at P_0 with respect to S_0 , if there are such constants $D_1, D_2, \dots, D_\nu$ for which

$$\lim_{P \rightarrow P_0} \frac{f(P) - f(P_0) - [D_1(x_1 - a_1) + D_2(x_2 - a_2) + \dots + D_\nu(x_\nu - a_\nu)]}{PP_0} = 0 \quad (3.6-1)$$

$(P \in S).$

this condition — as we know — can also be written in the form:

$$\begin{cases} f(P) - f(P_0) = [D_1(x_1 - a_1) + D_2(x_2 - a_2) + \dots + D_\nu(x_\nu - a_\nu)] + \\ + [\varepsilon_1(x_1 - a_1) + \varepsilon_2(x_2 - a_2) + \dots + \varepsilon_\nu(x_\nu - a_\nu)], \end{cases} \quad (3.6-2)$$

where any of $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_\nu$ tends to 0 as $P \rightarrow P_0$ over the points of S_0 . The "main term" on the right, i.e.,

$$(df)_{P_0} = D_1(x_1 - a_1) + \dots + D_\nu(x_\nu - a_\nu) \quad (3.6-3)$$

is called the *total (complete) differential* of $f(P)$ at P_0 ; its k -th term: $D_k(x_k - a_k)$ is the so-called *partial differential* with respect to x_k of the function at point P_0 . From (3.6-2) we get by putting $x_i = a_i$ ($i \neq k$), $x_k \neq a_k$ and by passage to the limit $x_k \rightarrow a_k$ a direct representation of the constants $D_1, D_2, \dots, D_\nu$:

$$D_k = \lim_{x_k \rightarrow a_k} \frac{f(a_1, \dots, a_{k-1}, x_k, a_{k+1}, \dots, a_\nu) - f(a_1, \dots, a_\nu)}{x_k - a_k} \quad (3.6-4)$$

$(k = 1, 2, \dots, \nu).$

Therefore, D_k is nothing else but

$$\left(\frac{\partial f}{\partial x_k} \right)_{P_0} = f'_{x_k}(P_0), \quad (3.6-5)$$

i.e., the *partial differential coefficient* (derivate) of f with respect to x_k and S_0 at point P_0 . Using this notation, (3.6-3) may be written as

$$(df)_{P_0} = \sum_{k=1}^{\nu} f'_{x_k}(P_0)(x_k - a_k), \quad (3.6-6)$$

i.e., the total differential is a homogeneous linear form of the increments $\Delta x_1 = x_1 - a_1, \dots, \Delta x_{\nu} = x_{\nu} - a_{\nu}$, with the partial derivatives as coefficients.

In the case $\nu = 1$, the condition (3.6-1) becomes, after omitting the indices,

$$\lim_{x \rightarrow a} \frac{f(x) - f(a) - D(x - a)}{|x - a|} = 0 \quad (x \in S_0), \quad (3.6-7)$$

which is equivalent to the existence of the finite *differential coefficient* (ordinary derivative)

$$D = f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \quad (x \in S_0), \quad (3.6-8)$$

and yields herewith also the most elementary form of the *differentiability* (derivability) condition for a function $f(x)$ at a point a with respect to the linear set S_0 .¹ The corresponding *differential* is simply

$$(df)_a = f'(a)(x - a) \quad (x \in S_0). \quad (3.6-9)$$

3.62 If we modify the above definitions of derivatives in such a way that we use instead of "limit" the words "cluster value" and "upper or lower limit", respectively, so we get further useful characteristic data of the function, the so-called *generalized derivate numbers*. For instance, in the case of a function of one variable we can consider the numbers

$$\begin{cases} \overline{D} = (\overline{D}f)_a = \overline{\lim}_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}, \\ \underline{D} = (\underline{D}f)_a = \underline{\lim}_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}, \end{cases} \quad (3.6-10)$$

called the *upper and lower derivative*, respectively, of f at the point a , with respect to S_0 . Taking S_0 especially for a unilateral neighbourhood of a , we obtain instead of $\overline{D}$ and $\underline{D}$ the so-called *Dini derivatives*, namely, the right-hand ones:

$$\begin{cases} D^+ = (D^+f)_a = \overline{\lim}_{x \rightarrow a+0} \frac{f(x) - f(a)}{x - a}, \\ D_+ = (D_+f)_a = \underline{\lim}_{x \rightarrow a+0} \frac{f(x) - f(a)}{x - a}, \end{cases} \quad (3.6-11)$$

¹ According to the former statement, S_0 does not contain point a .

and the left-hand ones:

$$\begin{cases} D^+ = (D^+ f)_a = \overline{\lim}_{x \rightarrow a-0} \frac{f(x) - f(a)}{x - a}, \\ D_- = (D_- f)_a = \underline{\lim}_{x \rightarrow a-0} \frac{f(x) - f(a)}{x - a}. \end{cases} \quad (3.6-12)$$

It is clear that D^+, D_+, D^-, D_- exist for an arbitrary function at every point of its domain of definition in wider sense, i.e., if the values $\pm\infty$ are also admitted.

In the remarks below, derivatives are mentioned mostly without indication of a basic set; then S_0 should be a two-sided or a right-hand or left-hand neighbourhood of the point in question. From the criterion 3.2.2 and the definitions given above we see that *the necessary and sufficient condition for the existence of a finite or infinite right-hand derivative $f'_+(a)$ is $D^+ = D_+$; and similar assertion holds on the existence of $f'_-(a)$* .¹ [Cf. (1.6-2).] The corresponding condition in the case of ordinary derivability is that D^+ and D_- or the four Dini derivatives should be finite and equal to one another. Applying the latter criterion to the function of one variable $\varphi(x_k) = f(a_1, \dots, a_{k-1}, x_k, a_{k+1}, \dots, a_n)$, we get the conditions for the existence of $f'_{x_k}(P_0)$. If, in turn, the above derivatives (derivates) are considered as point functions, the differentiation process leads to various kinds of second derivate numbers of f ; the derivatives of the latter are the so-called third derivate numbers of f ; etc. The *differentials of higher order $d^k f$* ($k = 2, 3, \dots$), as well as the concept of *repeated differentiability* are defined similarly by iteration.

Examples: I. we have for $y = |x|$ at the origin $D^+ = D_+ = 1$, $D^- = D_- = -1$, $\overline{D} = 1$, $\underline{D} = -1$; II. in case of the function

$$y = \begin{cases} \gamma x \sin \frac{1}{x} & (\gamma \neq 0 \text{ const.}) \text{ for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases} \quad (3.6-13)$$

the values of the Dini derivatives at point 0 are: $D^+ = |\gamma|$, $D_+ = -|\gamma|$, $D^- = |\gamma|$, $D_- = -|\gamma|$ so that $\overline{D} = |\gamma|$, $\underline{D} = -|\gamma|$; III. for the function

$$y = \begin{cases} x^2 \sin \frac{1}{x} & (x \neq 0) \\ 0 & (x = 0) \end{cases} \quad (3.6-14)$$

the corresponding derivate numbers at 0 are: $D^+ = D_+ = D^- = D_- = 0$ and at the same time $\overline{D} = \underline{D} = 0$. Accordingly, for example I. $y'_+(0) = 1$, $y'_-(0) = -1$, and the function has a break point at 0 [cf. (1.6-2)]; function II. is neither from the right, nor from the left differentiable at point 0; function

¹ If $f'_+(a)$ [or $f'_-(a)$] exists and is finite, then we say that $f(x)$ is *differentiable from the right* (or from the left) at point a .

III. is differentiable both from the right and from the left at the origin and also $y'(0)$ exists (it is 0).

3.63 From the foregoing it turns out that the derivate numbers have a simple *geometric meaning*: if we draw the secant line going through the points Q_0 with abscissa a and Q with abscissa x on the image curve of f , then D^+ and D_+ or D^- and D_- yield the side slopes of two angles with vertex Q_0 , and $\overline{Q_0Q}$ may "oscillate" in the two angle domains thus obtained as $x \rightarrow a + 0$ and $x \rightarrow a - 0$, respectively.

Note that example II) also illustrates that at certain isolated points the values of all four derivate numbers can be prescribed arbitrarily, if we consider such functions whose diagram oscillates in the right or left neighbourhood of these points inside suitable angle domains. We shall see later that "globally" there is another situation (cf. 4.31).

The *fundamental properties of differentiable functions* are expressed (apart from the so-called general differentiation rules) by the well-known *Darboux theorem* as well as the *mean value theorems* of *Lagrange (Rolle)* and *Cauchy*; on the interchangeability of the order of partial differentiations we have *Schwarz's* and *Young's theorem*. The transcription of all these results to point functions can be carried out immediately.

3.7 Constructing everywhere continuous, nowhere differentiable functions

3.71 Let us deal now with the *relationship of differentiability and continuity*!

We get from (3.6-2) at once that the first mentioned property locally implies the second one, i.e., *the continuity of $f(P)$ at P_0 with respect to S_0 is necessary to its differentiability at point P_0 with respect to the same set*. Our assertion is not invertible, i.e., from continuity at a point differentiability at the same point does not follow. This can be seen already from the well-known example of $|x|$ at the origin.

But considering that the actual continuous functions occurring in applications are also differentiable apart from certain isolated points which can easily be characterized, the opinion was widespread till the second half of the last century, that differentiability would follow from continuity maybe "globally". Speaking more precisely: it was thought that a function which is continuous in a region S , must also be differentiable at some points, possibly at "most" points of S .

This belief was first disproved by the famous *example of Weierstrass* in 1872. (B. BOLZANO could not publish an example found earlier, so his result remained completely unknown.) K. WEIERSTRASS has shown during

his university courses that the function:

$$W_{\lambda,q}(x) = \sum_{n=0}^{\infty} q^n \cos(\lambda^n \pi x) \quad (3.7-1)$$

where $\lambda > 1$ denotes an odd integer, $0 < q < 1$ and $\lambda q > 1 + (3\pi/2)$, is continuous for every real x , but not differentiable at any point.¹

Nowadays a lot of other pertinent examples are known, among them such functions which can be discussed in a quite easy way.² In what follows a particularly elementary, vividly descriptive example will be given.³

3.72 Our construction is based upon two obvious observation:

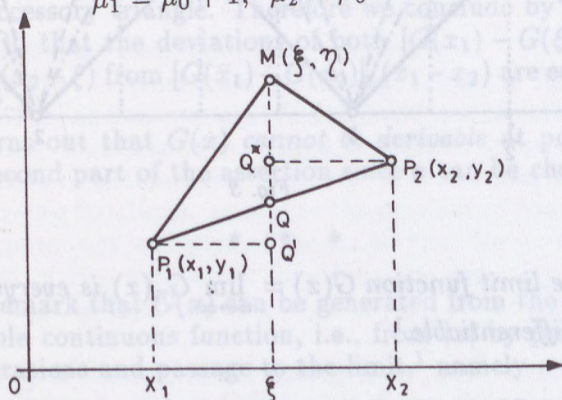
(I) If $f(x)$ is differentiable at a point a , and x_1, x_2 are two distinct points for which $a \in [x_1, x_2]$, then

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} \rightarrow f'(a) \quad (3.7-2)$$

as $x_1 \rightarrow a - 0, x_2 \rightarrow a + 0$.

(II) Let $P_1(x_1, y_1), P_2(x_2, y_2)$ be two points of the plane with $x_1 < x_2$ and Q the midpoint of $\overline{P_1P_2}$; let us denote the projection of P_1 on the line parallel to the ordinate axis through Q by Q' , and let $M(\xi, \eta)$ be that point of the mentioned parallel line which arises from Q by marking off the segment $\overline{P_1Q'}$ in the direction of the positive ordinate axis (Fig. 2). If the slopes of the segments $\overline{P_1P_2}, \overline{P_1M}, \overline{P_2M}$ are, in turn, μ_0, μ_1 and μ_2 , then we have

$$\mu_1 = \mu_0 + 1, \quad \mu_2 = \mu_0 - 1. \quad (3.7-3)$$



¹ Weierstrass [1].

² See e.g. Dieudonné [1], Faber [1], H. Hahn [1], Knopp [1], Lebesgue [10], Orlicz [1], Van der Waerden [1], etc. As far as the deeper theoretical background of the topic is concerned, we refer to Bruckner [1], Hardy [3], Mikolás [3], Zygmund [4].

³ Cf. Mikolás [3], 50–53. As is well-known, there are trivial examples of continuous, non-differentiable functions in complex analysis ($\bar{z}, |z|, \arg z, \operatorname{Re} z, \operatorname{Im} z$, etc.), because derivability of a function of one complex variable on a region means an essentially stronger condition than the existence of the derivative of a real function $f(x)$ in an interval.

In fact: (I) follows from the elementary property of the difference quotient that $[f(x_1) - f(x_2)] / (x_1 - x_2)$ is always between $[f(x_1) - f(a)] / (x_1 - a)$ and $[f(x_2) - f(a)] / (x_2 - a)$ ($x_1 < a < x_2$). As regards (II), we can see in Fig. 2 where $y_1 \leq y_2$ is assumed, and $\overline{MQ} = \overline{P_1Q'}$, $\overline{QQ'} = \overline{QQ^*}$:

$$\mu_1 = \frac{\overline{MQ}}{\overline{P_1Q'}} + \frac{\overline{QQ'}}{\overline{P_1Q}} = 1 + \mu_0, \quad \mu_2 = -\frac{\overline{MQ} - \overline{QQ^*}}{\overline{P_2Q^*}} = -1 + \mu_0.$$

The case $y_1 > y_2$ can be arranged alike.

For the sake of brevity, let us call now the segments $\overline{MP_1}$ and $\overline{MP_2}$ occurring in construction (II) the "accessories" of $\overline{P_1P_2}$, the triangle with the vertices P_1, P_2, M the "accessory triangle". Consider the following sequence of functions: let $G_0(x) \equiv 0$; let $G_1(x)$ denote the function represented by the broken line ("zigzag line") the "sides" of which are the accessories of the intervals $[k, k+1]$ ($k = 0, \pm 1, \pm 2, \dots$); when drawing the accessories of the last-mentioned ones, we obtain the diagram of $G_2(x)$; applying again the operation of "forming accessories" to $G_2(x)$, it results $G_3(x)$; and so on (Fig. 3).

We assert that

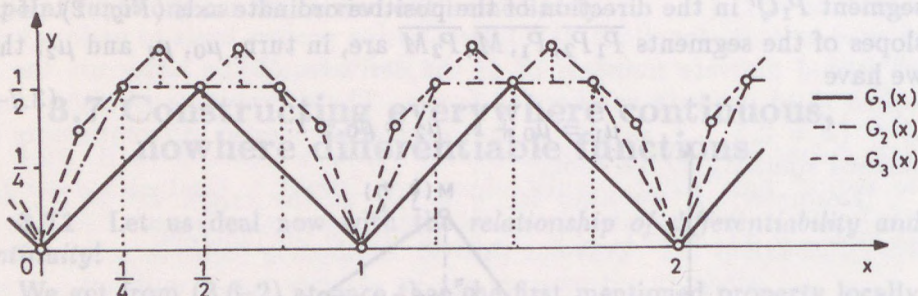


Fig. 3

3.7.1 The limit function $G(x) = \lim_{n \rightarrow \infty} G_n(x)$ is everywhere continuous but nowhere differentiable.¹

Proof. 1° Obviously $G_1(x) - G_0(x) \leq 1/2$, $G_2(x) - G_1(x) \leq 1/4$, etc. In general, $G_k(x) - G_{k-1}(x) \leq 2^{-k}$ ($k = 1, 2, \dots$) for every x . Thus the series $\sum_{k=1}^{\infty} [G_k(x) - G_{k-1}(x)]$ is uniformly convergent on $(-\infty, \infty)$ by the criterion of Weierstrass (cf. 3.51), and the same holds for the corresponding sequence

¹ It is evident that the above procedure yields whichever point of the image of $G(x)$, but all of its points i.e., the whole graph of $G(x)$ cannot be represented.

of functions $\{G_n(x)\}$. Thus $G(x) = \lim_{n \rightarrow \infty} G_n(x)$ exists and is *continuous* at any point x .

2° Suppose that $g(x)$ is differentiable at a given point a .

Then the above observation (I) yields for any two pairs of values $t_1 < t_2$, $u_1 < u_2$ that the deviation of $[G(t_1) - G(t_2)] / (t_1 - t_2)$ from $[G(u_1) - G(u_2)] / (u_1 - u_2)$ is as small as we want, provided that point a is sufficiently approached by t_1 and u_1 from the left, by t_2 and u_2 from the right. [Cf. 3.2.2 (II).] Accordingly we should have

$$\left| \frac{G(t_1) - G(t_2)}{t_1 - t_2} - \frac{G(u_1) - G(u_2)}{u_1 - u_2} \right| < \frac{1}{2} \quad (3.7-4)$$

for each pertinent pair $\{t_1, t_2\}, \{u_1, u_2\}$ which belongs to a sufficiently small interval $[\delta_1, \delta_2]$ containing a in its interior.

On the other hand, we see that the end points of any pair of accessories in our construction belong to the graph of the forthcoming functions $G_n(x)$ and so to the diagram of $G(x)$, too. Moreover, since the projection on the x axis of any side of the broken line representing $G_n(x)$ has the length 2^{-n} , thus each subset of the image of $G(x)$, whose projection on the x axis is of a prescribed length, contains three points which can be considered as the vertices of an accessory triangle. Hence it follows that we can also find on the above interval $[\delta_1, \delta_2]$ a triad of points, say $\{x_1, \xi, x_2\}$ such that $a \in [x_1, x_2]$ and the end points of the corresponding ordinates $G(x_1), G(\xi), G(x_2)$ determine an accessory triangle. Therefore we conclude by the observation (II) [cf. (3.7-3)], that the deviations of both $[G(x_1) - G(\xi)] / (x_1 - \xi)$ and $[G(x_2) - G(\xi)] / (x_2 - \xi)$ from $[G(x_1) - G(x_2)] / (x_1 - x_2)$ are equal to 1, which contradicts (3.7-4).

Thus it turns out that $G(x)$ *cannot be derivable* at point a ; and this furnishes the second part of the assertion since a can be chosen arbitrarily.

* * *

3.73 We remark that $G(x)$ can be generated from the simplest locally non-differentiable continuous function, i.e., from $|x|$ by periodic continuation, linear operations and passage to the limit,¹ namely

$$G(x) = \sum_{k=0}^{\infty} 2^{-k} G_1(2^k x). \quad (3.7-5)$$

More generally: we can extend the fundamental idea of the above construction in such a way that it results a quite *comprehensive method* to

¹ The occurring break points of the zigzag lines so to say "accumulate" as $n \rightarrow \infty$.

obtain geometrically or analytically defined families of continuous, non-differentiable real functions. One of these function families can be represented in the form:

$$\sum_{k=0}^{\infty} q^k \Phi(\lambda^k x) \quad (0 < q < 1; \lambda = 3, 5, \dots) \quad (3.7-6)$$

where $\Phi(u)$ is a continuous function satisfying certain additional conditions; the type (3.7-6) contains also the Weierstrass function (3.7-1) as a special case (cf. Mikolás [3], Th. 3-7).

The discovery of the first examples of everywhere continuous, nowhere derivable real functions created a great stir in the literature at the end of the nineteenth century, and — since it drew attention to the subtle investigations relating to the foundations of analysis — it gave a strong impulse to the development of the modern theory of real functions. Thus the impression was given for some time that a function may be nowhere differentiable even if its graphic representation is relatively simple, i.e., that other simple geometric features do not indicate the existence of tangents at all.

Surprisingly, H. LEBESGUE showed just after the turn of the century, that the situation changes if the diagram of the function is supposed to be monotonic or (more generally) *rectifiable* — instead of the assumption of continuity. These deep results will be presented in the next chapter.

MONOTONIC, ABSOLUTELY CONTINUOUS FUNCTIONS AND FUNCTIONS OF BOUNDED VARIATION

4.1 Basic properties, continuous and pure jumping part of a monotonic function

4.11 The concept of a function monotonic in a given interval is well-known from elementary analysis. The notion can immediately be extended to an arbitrary basic set E as follows. Any function $f(x)$ defined on a linear set E is said to be *monotone increasing*, or *monotone decreasing*¹ on E , if the inequality $f(x_1) \leq f(x_2)$ or $f(x_1) \geq f(x_2)$ holds whenever $x_1, x_2 \in E$ with $x_1 < x_2$. We use the terms *strictly* monotone increasing, or decreasing, if the sign of equality is excluded in the conditions just mentioned. A function is called simply *monotonic on a set* if it is monotone increasing or monotone decreasing thereon. A function which is constant on a set may be regarded as monotone increasing or monotone decreasing on that set.

Examples: $\log x$ is strictly monotone increasing in $(0, \infty)$, whereas the function

$$y = \begin{cases} x \sin \frac{1}{x} & (x \neq 0) \\ 0 & (x = 0) \end{cases} \quad (4.1-1)$$

is not monotonic at any neighbourhood of the origin.

Note that if $f(x)$ is monotone decreasing, then the function $-f(x)$ is monotone increasing and conversely. Consequently, when studying the properties of monotonic functions it is enough to restrict ourselves to monotone increasing functions. Furthermore, it is clear that the sum of any number of monotone increasing functions, as well as the product of non-negative monotone increasing functions is always of the same type. However, the difference and the quotient of two monotonic functions are generally not monotonic. For instance, $x - x^2$ in $[0, 1]$ and $(\log x)/x$ in $[1, 3]$ are not monotonic.

4.1.1 If $f(x)$ is monotone increasing in a finite interval $[a, b]$ then it has both right-hand and left-hand limit at each interior point $x_0 \in (a, b)$ with

$$-\infty < f(x_0 - 0) \leq f(x_0) \leq f(x_0 + 0) < \infty. \quad (4.1-2)$$

At the end points these are replaced by the inequalities $f(a + 0) \geq f(a)$ and $f(b - 0) \leq f(b)$.

¹ The word "monotone" is (for the sake of brevity) sometimes omitted.

P r o o f. Let $x_0 \in (a, b)$. Consider a right-hand neighbourhood $(x_0, x_0 + \delta) \subset [a, b]$ of the point x_0 , further the infimum m_0 of f in this interval (cf. 2.12). Then, on the one hand, $f(x) \geq m_0$ for all $x \in (x_0, x_0 + \delta)$. On the other hand, being given an arbitrarily small $\varepsilon > 0$, a point ξ of $(x_0, x_0 + \delta)$ can be found such that $f(\xi) < m_0 + \varepsilon$. Since f is increasing, we have $m_0 \leq f(x) \leq f(\xi) < m_0 + \varepsilon$ whenever $x_0 < x < \xi$. Thus $f(x) \rightarrow m_0$ as $x \rightarrow x_0 + 0$. Similarly it follows that $f(x_0 - 0)$ is the supremum M_0 of f taken in a left-hand neighbourhood of x_0 , and (4.1-2) clearly follows from the monotony of f as well as from the inequalities $M_0 \leq f(x_0), m_0 \geq f(x_0)$.

For the end points a, b we have only to repeat the reasoning either for the right-hand or the left-hand side alone.

* * *

The above statement is valid also for monotone decreasing functions if right-hand limits are replaced by left-hand ones, and conversely. Thus, if a function f is monotonic in an interval $[a, b]$, then every interior point of $[a, b]$ is either a point of continuity or a point of discontinuity of the first kind (cf. 1.61). In addition, if $f(a)$ and $f(b)$ are defined as equal to $f(a + 0)$ and $f(b - 0)$, respectively, then $f(x)$ is continuous at a from the right and at b from the left. The quantities $|f(x_0 + 0) - f(x_0 - 0)|$, $|f(x_0 + 0) - f(x_0)|$, $|f(x_0) - f(x_0 - 0)|$ are called, in turn, the *total jump*, *right jump* and *left jump* of the function at point x_0 . The points of continuity are characterized by the fact that both of the corresponding unilateral jumps are zero.

Any monotonic function may have not only a finite, but also an infinite number of discontinuities of the first kind. In the finite case, any of the right or left jumps may take however large values.

4.12 It is easy to construct a function which is monotone increasing in a finite interval (a, b) and has a finite number of discontinuities at some interior points $\xi_1 < \xi_2 < \dots < \xi_{n-1}$ of (a, b) , such that the right and left jumps at any ξ_k have some prescribed values $u_k^+ \geq 0$ or $u_k^- \geq 0$ with $u_k^+ + u_k^- > 0$, ($k = 1, 2, \dots, n - 1$). The simplest example is the following step function:¹

$$J(x) = \begin{cases} \sum_{r=1}^k (u_r^- + u_r^+) & \text{for } \xi_k < x < \xi_{k+1} \quad (k = 0, 1, \dots, n - 1), \\ \sum_{r=1}^k (u_r^- + u_r^+) + u_{k+1}^- & \text{for } x = \xi_{k+1} \quad (k = 0, 1, \dots, n - 2) \end{cases} \quad (4.1-3)$$

¹ So we have $J(x) \equiv 0$ for $x \in (a, \xi_1)$; furthermore, $J(\xi_1) = u_1^-$; $J(x) = u_1^- + u_1^+$ ($\xi_1 < x < \xi_2$); $J(\xi_2) = (u_1^- + u_1^+) + u_2^-$; $J(x) = (u_1^- + u_1^+) + (u_2^- + u_2^+)$ ($\xi_2 < x < \xi_3$); etc.

where $\xi_0 = a$, $\xi_n = b$ and $\sum_{r=1}^k$ means 0 for $k = 0$.

Observe that (4.1-3) can also be written in the form:

$$J(x) = \sum_{\xi_r \leq x} u_r^- + \sum_{\xi_r < x} u_r^+ \quad (x \in (a, b)) \quad (4.1-4)$$

for any partition $\{\xi_r\} \subset (a, b)$. This representation has the advantage that it can also be used for infinite (a, b) and an arbitrary (finite or infinite) countable set of discontinuities $\{\xi_k\}$ in (a, b) , provided that

$$\sum_r (u_r^- + u_r^+) < \infty. \quad (4.1-5)$$

The monotone increasing function (4.1-4) will be called the *pure jumping function* associated with the discontinuities $\{\xi_k\}$ and the corresponding jumps $\{u_r^-, u_r^+\}$.²

4.13 The sum of $J(x)$ and of any function which is continuous and monotone increasing in (a, b) is obviously again monotone increasing with exactly the same discontinuities and jumps as $J(x)$. We shall now show that all bounded, monotone increasing functions can be obtained in this way.

4.1.2 Let $f(x)$ [$x \in (a, b)$] be a bounded, monotone increasing function. Then

(I) the set of the discontinuities of f is countable and the sum of its total jumps is finite;

(II) there exists a monotone increasing and continuous function $\varphi(x)$ such that

$$f(x) = \varphi(x) + J(x), \quad x \in (a, b) \quad (4.1-6)$$

where $J(x)$ denotes the pure jumping function associated with the discontinuities and the corresponding jumps of $f(x)$.

P r o o f. 1° The countability of the discontinuities of f , i.e., that they can be written as a (finite or infinite) sequence $\{\xi_r\} \subset (a, b)$, is an immediate consequence of 4.1.1, if 1.6.1 is also taken into consideration.

Denoting by u_r^+ and u_r^- the one-sided jumps associated with ξ_r , we see that all the intervals $(f(\xi_r) - u_r^-, f(\xi_r) + u_r^+) = (f(\xi_r - 0), f(\xi_r + 0))$ are

¹ Both sums in (4.1-4) are extended to all terms whose indices r fulfill the corresponding inequality indicated under the summation sign.

² It is worth noting that $\{\xi_r\}$ can be a set *dense everywhere* in (a, b) (e.g. all rational points of this interval). In such a case, we get an example of a monotonic function which is discontinuous at points lying densely on a given interval.

disjoint to one another and contained in the interval $(f(a+0), f(b-0))$ which is finite by the boundedness of f . Furthermore, 4.1.1 implies that $f(x) \geq f(\xi_r + 0)$ for $x > \xi_r$ and $f(x) \leq f(\xi_r - 0)$ for $x < \xi_r$ so that $f(\xi_i + 0) \leq f(\xi_k - 0)$ as $\xi_i < \xi_k$. Hence it follows that the total length of the intervals in question is:

$$\sum_r (u_r^- + u_r^+) \leq f(b-0) - f(a+0) < \infty.$$

2° Consider now the pure jumping function (4.1-4) corresponding to the discontinuities and jumps of $f(x)$ and the difference function:

$$\varphi(x) = f(x) - J(x). \quad (4.1-7)$$

This φ is continuous not only at the points of continuity of f but also at every point ξ_r . We have e.g. for the right jump:

$$\varphi(\xi_r + 0) - \varphi(\xi_r) = [f(\xi_r + 0) - f(\xi_r)] - [J(\xi_r + 0) - J(\xi_r)] = u_r^+ - u_r^- = 0;$$

similarly, $\varphi(\xi_r) - \varphi(\xi_r - 0) = 0$. On the other hand, $\varphi(x)$ is monotone increasing in (a, b) since, for $[x_1, x_2] \subset (a, b)$,

$$\varphi(x_2) - \varphi(x_1) = [f(x_2) - f(x_1)] - [J(x_2) - J(x_1)]$$

and $f(x_2) - f(x_1)$ is clearly not less than the sum of all jumps of f associated with $[x_1, x_2]$ which is identical with $J(x_2) - J(x_1)$.

Thus the function $\varphi(x)$ possesses the properties required in the proposition and (4.1-7) implies (4.1-6).

* * *

4.14 Because of one of our previous comment, 4.1.2 has an immediate analogue for monotone decreasing functions. In (4.1-6), $\varphi(x)$ is called the *continuous part* and $J(x)$ the *pure jumping part* of $f(x)$. Since every countable set is a null set (cf. 2.6.2), our last result implies that *the discontinuities of any monotonic function constitute a null set*. This assertion is usually expressed by saying that *every monotonic function is continuous almost everywhere*.

In general, whenever a statement holds for all points of a set E except certain points of it constituting a null set, we say that the statement in question is valid on E *almost everywhere* or *almost always* or *for almost all* $x \in E$. (The usual abbreviations are *a.e.*, *a.a.* or sometimes in French *p.p.* [= "presque partout"], respectively.)

4.2 Differentiability almost everywhere of monotonic functions

4.21 Having discussed the connection between monotony and continuity, let us now turn our attention to the problem of the relationship between monotony and derivability. In 1904, H. LEBESGUE made a rather unexpected discovery about this. For any monotonic function, he proved that the points at which the function has no finite derivate form a null set. In other terms:

4.2.1 (LEBESGUE'S THEOREM ON MONOTONIC FUNCTIONS) *If $f(x)$ is monotonic in an interval (a, b) , then it is differentiable almost everywhere there.*¹

P r o o f. Since the union of an arbitrary sequence of null sets is again a null set and the negative of a monotone decreasing function increases, it is sufficient to prove the theorem for the case where (a, b) is finite and f is a monotone increasing function.

1°. The points x_0 where $f(x)$ is not differentiable can be divided into three classes: 1. points having a neighbourhood in which $[f(x) - f(x_0)] / (x - x_0)$ is unbounded; 2. those having a neighbourhood with $|f(x) - f(x_0)| \leq \text{const. } |x - x_0|$ and such that at least one of the one-sided derivatives $f'_+(x_0)$ and $f'_-(x_0)$ does not exist; 3. those where both $f'_+(x_0)$ and $f'_-(x_0)$ exist but have different values. The points of the last class are known to constitute a null set by 1.6.2, and in order to show the same for the first two classes, we need the Dini derivate numbers, introduced in 3.62. [Cf. (3.6-11) — (3.6-12).]

The function f being monotone increasing, all difference quotients are non-negative, so that we have at every point $x_0 \in (a, b)$ the inequalities $D^+ \geq D_+ \geq 0$, $D^- \geq D_- \geq 0$. Since any result for the right-hand derivate numbers gives a similar statement for the left-hand ones (e.g. by applying the former to $-f(-x)$), we see that our theorem is true if (I) the points at which $D^+ = +\infty$ and (II) the points characterized by $D_+ < D^+$ both form null sets.

2° In order to verify assertion (I), let us denote by $E_\infty \subset (a, b)$ the set of points where $D^+ = +\infty$ and let $G > 0$ be an arbitrary positive number. If E_∞ is finite (and thus a null set), then we have nothing more to prove. Otherwise, for each $x \in E_\infty$, choose $l_x > 0$ and construct an interval

¹ Note that monotony (and also continuity) is a much weaker condition than differentiability. It is therefore surprising that one may yet infer (at least up to a null set) derivability from monotony. The theorem 4.2.1 was published first in Lebesgue [4].

$[x, x + l_k]$ such that $x + l_x < b$ and

$$\frac{f(x + l_x) - f(x)}{l_x} > G. \quad (4.2-1)$$

(This is possible by the definition of E_∞ .) To the resulting system of intervals $\{I\}$, which covers E_∞ , Sierpinski's lemma (2.6.3) is applicable. Accordingly, we can select from $\{I\}$ a finite number of non-overlapping intervals $I_k = [x_k, x_k + l_{x_k}]$ ($k = 1, 2, \dots, p$), in such a way that the outer Lebesgue measure of the set $E_\infty - \bigcup_{k=1}^p I_k$ is less than $\mu/2$ where $\mu = \bar{\mu}_L(E_\infty)$.

Now the sum of the outer (L)-measures of $\bigcup_{k=1}^p I_k$ and $E_\infty - \bigcup_{k=1}^p I_k$ is at least μ , which implies that

$$l_{x_1} + l_{x_2} + \dots + l_{x_p} \geq \mu - \bar{\mu}_L(E_\infty - \bigcup_{k=1}^p I_k) > \mu - \frac{\mu}{2} = \frac{\mu}{2}. \quad (4.2-2)$$

Since the intervals $I_1, I_2, \dots, I_p$ are non-overlapping and all of them are contained in (a, b) , the monotone increasing function f satisfies

$$\sum_{k=1}^p [f(x_k + l_{x_k}) - f(x_k)] \leq f(b) - f(a). \quad (4.2-3)$$

Hence, by (4.2-1) — (4.2-2), we get the inequality $f(b) - f(a) > G(l_{x_1} + l_{x_2} + \dots + l_{x_p}) > G\mu/2$ and therefore

$$\mu < \frac{2[f(b) - f(a)]}{G}. \quad (4.2-4)$$

Considering that G is arbitrarily large and μ non-negative, it follows that $\mu = \bar{\mu}_L(E_\infty) = 0$, as required.

3° As far as statement (II) is concerned, let us examine the set E_D of points $x \in (a, b)$ at which $D_+ < D^+$. If E_D consists of a finite number of points, the proof is finished; so we suppose in the sequel E_D to be infinite. We shall split up our set into certain subsets defined by using rational numbers: denoting by u and v ($u > v$) two prescribed *positive rational numbers* and considering the totality $E(u, v)$ of points $x \in E_D$ with $D^+ > u > v > D_+$, we have

$$E_D = \bigcup_{u, v} E(u, v), \quad (4.2-5)$$

where the union is over all pairs of rationals satisfying the given conditions. Since (4.2-5) contains a countable family of subsets (cf. 1.2.1–1.2.2), by 2.6.2 it will be sufficient to show that every $E(u, v)$ is a null set.

Using indirect reasoning, let us assume that there exists a pair $u > v$ of rationals for which $E(u, v)$ is infinite and $\mu^* = \bar{\mu}_L[E(u, v)] > 0$. Then (by the definition (2.6-5) of outer Lebesgue measure) we can find for any $\varepsilon > 0$, however small, a linear open set S such that $E(u, v) \subseteq S \subseteq (a, b)$ and $\bar{\mu}_L(S) < \mu^* + \varepsilon$.

Let x be a point of $E(u, v)$. Since the definition of $E(u, v)$ implies $D_+ < v$ and the set S [covering $E(u, v)$] may be regarded as the union of disjoint open intervals (cf. 2.4.3), an interval $[x, x + \lambda_x] \subset S$ ($\lambda > 0$) can be found for which

$$\frac{f(x + \lambda_x) - f(x)}{\lambda_x} < v. \quad (4.2-6)$$

From the intervals associated with the individual points of $E(u, v)$ in the manner described above, we can select — again because of Lemma 2.6.3 — a finite, non-overlapping system $i_k = [x_k, x_k + \lambda_{x_k}]$ ($k = 1, 2, \dots, q$) such that the part $E_1(u, v)$ of $E(u, v)$ contained in $\bigcup_{k=1}^q i_k$ has outer (L)-measure larger than $\mu^* - \varepsilon$. Plainly $\bigcup_{k=1}^q i_k \subset S$, and the outer (L)-measure of S is less than $\mu^* + \varepsilon$; hence we obtain

$$\lambda_{x_1} + \lambda_{x_2} + \dots + \lambda_{x_q} < \mu^* + \varepsilon, \quad (4.2-7)$$

and it results by (4.2-6) for the sum of the corresponding increments of f the upper bound:

$$\sum_{k=1}^q [f(x_k + \lambda_{x_k}) - f(x_k)] < v(\mu^* + \varepsilon). \quad (4.2-8)$$

A lower bound can be deduced by applying the inequality $D^+ > u$ which holds at each point ξ of the above subset $E_1(u, v)$. Thus, for every $\xi \in E_1(u, v)$ we can find a $\rho_\xi > 0$ and an interval $[\xi, \xi + \rho_\xi] \subset \bigcup_{k=1}^q i_k$ such that

$$\frac{f(\xi + \rho_\xi) - f(\xi)}{\rho_\xi} > u. \quad (4.2-9)$$

By Sierpinski's lemma 2.6.3, there is a finite number of non-overlapping intervals, say $j_k = [\xi_k, \xi_k + \rho_{\xi_k}]$ ($k = 1, 2, \dots, r$), in the family just introduced, such that $\bigcup_{k=1}^r j_k$ covers a subset of $E_1(u, v)$ whose outer (L)-measure is greater than $\bar{\mu}_L[E_1(u, v)] - \varepsilon > \mu^* - 2\varepsilon$, and any j_k is contained in one of the intervals i_k ($k = 1, 2, \dots, q$). Hence we have

$$\sum_{k=1}^r [f(\xi_k + \rho_{\xi_k}) - f(\xi_k)] > u(\mu^* - 2\varepsilon) \quad (4.2-10)$$

and, f being monotone increasing,

$$\sum_{k=1}^r [f(\xi_k + \rho_{\xi_k}) - f(\xi_k)] \leq \sum_{k=1}^q [f(x_k + \lambda_{x_k}) - f(x_k)], \quad (4.2-11)$$

so that the combination of (4.2-10)—(4.2-11) with (4.2-8) yields

$$u(\mu^* - 2\varepsilon) < v(\mu^* + \varepsilon).$$

But $\varepsilon > 0$ can be chosen as small as we please, and therefore we get $u \leq v$, whereas we initially assumed that $u > v$. Thus $\mu^* > 0$ turns out to be impossible, i.e., $E(u, v)$ must be a null set for every pair $\{u, v\}$ with $u > v$ of positive rational numbers. This implies by (4.2-5) that E_D is a null set, too.

* * *

4.22 H. LEBESGUE himself originally proved the deep theorem 4.2.1 only for *continuous* monotonic functions, working within the framework of his new theory of integration which we shall discuss in Chapters 6–7. Seven years later, G.C. and W.H. YOUNG as well as G. FABER showed that the condition of continuity is unnecessary, and that Lebesgue's argument can be simplified at certain points.¹ The proof presented above is essentially that published in an article by the Polish mathematicians A. RAJCHMAN and S. SAKS in 1923. Another proof, which is easier for the case of continuous monotonic functions, was found in 1931 by F. RIESZ (1880–1956) the world-famous professor in Szeged and (after World War II) in Budapest. Riesz's proof is also contained in his joint monograph with B. SZŐKEFALVI-NAGY.²

4.3 Applications; Fubini's theorem for series of functions

4.31 On the basis of the above, it can be shown that the Dini derivate numbers [see (3.6-11)—(3.6-12)] obey “almost everywhere” certain restrictions by pairs. More precisely, the following result holds (found by three authors between 1915 and 1924), which is given here only for orientation:

4.3.1 (THEOREM OF DENJOY-SAKS-YOUNG)³ *Apart from the points of a null set, the Dini derivatives of a function satisfy everywhere the conditions:*

¹ Cf. G.C. and W.H. Young [1], Faber [2].

² See Rajchman-Saks [1]; F. Riesz [23] and F. Riesz-B. Szőkefalvi-Nagy [1].

³ Cf. Denjoy [2], G.C. Young [1] and Saks [1].

(I) Two derivatives relating to the same side are either equal and finite, or unequal with at least one infinite of the corresponding sign.

(II) As regards D^+ and D_- , either both of them are finite and equal to each other, or both are infinite, namely $D^+ = +\infty$, $D_- = -\infty$.

(III) The last statement remains valid, if D^+ is replaced by D^- and D_- by D_+ .

Hence we get specially for the upper and lower derivative [cf. (3.6–10)]: except for a null set, at any point either $\overline{D} = +\infty$ and $\underline{D} = -\infty$ holds, or both $\overline{D}$ and $\underline{D}$ are finite and equal (case of differentiability). If it concerns a monotonic function, then only the second possibility can occur at any interior point, so that the assertion reduces to Lebesgue's theorem 4.2.1.

4.32 A further interesting application possibility is afforded by the problem of termwise differentiation of a series of monotonic functions. We have:

4.3.2 (FUBINI'S THEOREM ON FUNCTION SERIES)¹ If a series of functions $f_1(x) + f_2(x) + \dots = s(x)$ is convergent in a finite interval $[a, b]$ and every term $f_n(x)$ is there monotonic of the same type, then the series can be differentiated term-by-term almost everywhere in $[a, b]$, i.e.

$$f'_1(x) + f'_2(x) + \dots = s'(x) \quad (4.3-1)$$

at a.e. point $x \in [a, b]$.

P r o o f. Plainly we may restrict ourselves to the case of increasing functions and we may suppose that $f_n(a) = 0$ ($n = 1, 2, \dots$), since additive constants do not affect the result of differentiation.

According to the assumptions, $s_n(x) = f_1(x) + \dots + f_n(x)$ ($n = 1, 2, \dots$) and also the sum $s(x)$ are monotone increasing and so, by 4.2.1, a.e. derivable in $[a, b]$. Furthermore, the inequalities $s_n(x+l) - s_n(x) \leq s_{n+1}(x+l) - s_{n+1}(x)$ and $s_n(x+l) - s_n(x) \leq s(x+l) - s(x)$ ($l > 0$; $n = 1, 2, \dots$) imply after dividing by l and passage to the limit $l \rightarrow 0$:

$$s'_n(x) \leq s'_{n+1}(x), \quad s'_n(x) \leq s'(x) \quad (n = 1, 2, \dots) \quad (4.3-2)$$

for almost all $x \in E[a, b]$. Because of (4.3-2), the sequence of functions $\{s'_n(x)\}$ is monotone increasing and bounded above, and therefore it is convergent at the points x in question. We have still to show that

$$\lim_{n \rightarrow \infty} s'_n(x) = s'(x) \quad (4.3-3)$$

¹ Fubini [2].

a.e. in $[a, b]$; by monotony of the partial sums $s'_n(x) = f'_1(x) + \dots + f'_n(x)$ it is sufficient to see that there are special indices $n_1 < n_2 < \dots$ for which $s'_{n_k}(x) \rightarrow s'(x)$ ($k \rightarrow \infty$).

To that end, consider the number sequence $\{s_n(b)\}$ tending increasingly to $s(b)$; and select such a subsequence $s_{n_1}(b), s_{n_2}(b), \dots$ of it which satisfies the inequalities $s(b) - s_{n_k}(b) < 1/k^2$ ($k = 1, 2, \dots$). Then convergence of $\sum_{k=1}^{\infty} [s(b) - s_{n_k}(b)]$ is assured, and this numerical series majorizes in $[a, b]$ the series of functions

$$\sum_{k=1}^{\infty} [s(x) - s_{n_k}(x)] \quad (4.3-4)$$

by $0 \leq s(x) - s_{n_k}(x) \leq s(b) - s_{n_k}(b)$ ($a \leq x \leq b$; $k = 1, 2, \dots$). Thus the Weierstrass criterion (cf. 3.51) yields the uniform convergence of (4.3-4) on the segment $[a, b]$. But every term of the series (4.3-4) is a monotone increasing function of x , and herewith this is of the same type as the original series $\sum f_n(x)$. By the above argument, such a series has the property that its derivative formed term-by-term converges almost everywhere. Consequently, $\sum [s'(x) - s'_{n_k}(x)]$ is also convergent a.e. in $[a, b]$, which implies $s'(x) - s'_{n_k}(x) \rightarrow 0$ ($k \rightarrow \infty$), i.e., $\lim_{k \rightarrow \infty} s'_{n_k}(x) = s'(x)$ for almost all $x \in [a, b]$.

Hereby, the verification of the assertion is completed.

4.4 Functions of bounded variation, the notion of total variation

4.41 We have to deal also with another important class of functions, closely related to monotonic functions. This is the class of the so-called functions of bounded variation which arises in the most natural way through the discussion of the concepts of *arc length* and *rectifiability* of a curve.

Let $f(x)$ be a single-valued function, defined in a finite interval $a \leq x \leq b$. As is well-known, we mean by the arc length of the image of f in $[a, b]$, i.e., of the point set $\{(x, y) : a \leq x \leq b, y = f(x)\}$ the least upper bound of the set of total lengths of the corresponding inscribed polygons. Such a polygon length has the form $\sum_{k=1}^n \overline{Q_{k-1}Q_k}$ where Q_k denotes that point of the image curve which belongs to point x_k , a partition $a = x_0 < x_1 < \dots < x_n = b$ of $[a, b]$ being given. Since $\overline{Q_{k-1}Q_k} = \sqrt{(x_k - x_{k-1})^2 + [f(x_k) - f(x_{k-1})]^2}$, so that we have

$$|f(x_k) - f(x_{k-1})| \leq \overline{Q_{k-1}Q_k} \leq (x_k - x_{k-1}) + |f(x_k) - f(x_{k-1})|$$

and

$$\sum_{k=1}^n |f(x_k) - f(x_{k-1})| \leq \sum_{k=1}^n \overline{Q_{k-1}Q_k} \leq (b - a) + \sum_{k=1}^n |f(x_k) - f(x_{k-1})|,$$

it is evident: the graph of $f(x)$ ($a \leq x \leq b$) has a finite arc length, i.e., it is rectifiable if and only if the sum $\sum |f(x_k) - f(x_{k-1})|$ (occurring on both sides of the last inequality) remains under a common upper bound for all partitions of $[a, b]$.

We know that the functions in question are defined exactly by the condition just mentioned. In other terms: $f(x)$ is said to be of *bounded variation* in a finite interval $[a, b]$, if the set of sums of the form

$$V_f(a, x_1, \dots, x_{n-1}, b) = \sum_{k=1}^n |f(x_k) - f(x_{k-1})|, \quad (4.4-1)$$

relating to arbitrary partitions of $[a, b]$, has a finite least upper bound; the latter (which depends already on the totality of the values $f(x)$ ($a \leq x \leq b$)) is called the *total variation* of f with respect to $[a, b]$ and denoted by $V_f(a, b)$.¹

4.42 Observe that $V_f(a, x_1, \dots, x_{n-1}, b)$ becomes $|f(b) - f(a)|$ for $n = 1$, and thus it can be regarded as a "generalized increment" of the function f . Moreover there always hold the inequalities

$$|f(b) - f(a)| \leq V_f(a, x_1, \dots, x_{n-1}, b) \leq V_f(a, b), \quad (4.4-2)$$

which imply the *boundedness of any function of bounded variation*.

It is easy to find concrete examples of functions with finite total variation: *all functions monotonic in a closed interval are of bounded variation*, therefore the former class of functions is a subset of the latter one. In fact, if $f(x)$ increases in the whole $[a, b]$ or decreases throughout this interval, then obviously

$$V_f(a, x_1, \dots, x_{n-1}, b) = |f(b) - f(a)|, \quad (4.4-3)$$

since all increments $f(x_k) - f(x_{k-1})$ are of the same sign. Thus now the sum (4.4-1) is independent of the partition and we have $V_f(a, b) = |f(b) - f(a)|$.

More generally: any function which is *piecewise monotonic* in a segment $[a, b]$, is also of bounded variation there. We mean by piecewise monotony that we can find a partition of $[a, b]$ into a finite number of subintervals such that the function is monotonic in each of them (whether increasing or decreasing). Namely, the definition implies immediately the following *additive property*: if $f(x)$ is of bounded variation in certain intervals which form a partition of some interval $[a, b]$, then f is of bounded variation also in the latter. Conversely: if a function is of bounded variation in $[a, b]$, the

¹ Let us remark that $V_f(a, b)$ has the character of an interval function in case of variable a and b (cf. 3.12); furthermore that it can also be written as the Riemann integral of $|f'(x)|$ on $[a, b]$, provided that $f'(x)$ exists and is (R) -integrable in this interval. (Cf. 5.22.)

same holds for any subinterval of $[a, b]$. The corresponding property of total variation holds as well:

$$V_f(a, b) = V_f(a, c) + V_f(c, b) \quad (a < c < b) \quad (4.4-4)$$

The additivity, however, cannot be extended to the case of infinitely many subintervals. A simple counterexample is the function (4.1-1)¹: its diagram in $[-\delta, \delta]$ consists of infinitely many monotonic (alternately increasing or decreasing) pieces, and the function is yet not of bounded variation there, namely its total variation is $+\infty$.

4.43 Numerous further examples can be obtained on the basis of the following proposition:

4.4.1 *If both of $\varphi(x)$ and $\psi(x)$ are of bounded variation in a finite interval $[a, b]$, then the same holds for $\varphi(x) + \psi(x)$, $\varphi(x) - \psi(x)$, $\varphi(x) \cdot \psi(x)$ and in the case of the boundedness of $\psi(x)^{-1}$ also for $\varphi(x)/\psi(x)$.*

P r o o f. Let $a = x_0 < x_1 < \dots < x_n = b$ be a partition of $[a, b]$. Then by

$$\begin{aligned} |[\varphi(x_k) + \psi(x_k)] - [\varphi(x_{k-1}) + \psi(x_{k-1})]| &\leq |\varphi(x_k) - \varphi(x_{k-1})| + \\ &+ |\psi(x_k) - \psi(x_{k-1})| \quad (k = 1, 2, \dots, n) \end{aligned}$$

we get at once:

$$V_{\varphi+\psi}(a, b) \leq V_{\varphi}(a, b) + V_{\psi}(a, b), \quad (4.4-5)$$

and this implies the assertion on $\varphi(x) + \psi(x)$. The case of $\varphi(x) - \psi(x) = \varphi(x) + (-\psi(x))$ needs no separate discussion.

As far as $\varphi(x)\psi(x)$ is concerned, we can write for every k :

$$\begin{aligned} |\varphi(x_k)\psi(x_k) - \varphi(x_{k-1})\psi(x_{k-1})| &\leq |\varphi(x_k)\psi(x_k) - \varphi(x_{k-1})\psi(x_k)| + \\ &+ |\varphi(x_{k-1})\psi(x_k) - \varphi(x_{k-1})\psi(x_{k-1})| \leq M_{\psi}|\varphi(x_k) - \varphi(x_{k-1})| + \\ &+ M_{\varphi}|\psi(x_k) - \psi(x_{k-1})|, \end{aligned}$$

where $M_{\varphi} = \sup_{[a, b]} \{|\varphi(x)|\}$, $M_{\psi} = \sup_{[a, b]} \{|\psi(x)|\}$ (both finite).

Hence follows the inequality:

$$V_{\varphi\psi}(a, b) \leq M_{\psi}V_{\varphi}(a, b) + M_{\varphi}V_{\psi}(a, b) \quad (4.4-6)$$

and herewith the assertion on the product $\varphi\psi$.

¹ Any part of this graph containing the origin yields a simple example of a *continuous but non-rectifiable curve*.

Finally, the case of $\varphi(x)/\psi(x)$ reduces to the former, if we put $\Omega_\psi = \sup_{[a,b]} \{|\psi(x)|^{-1}\}$ and use the estimation:

$$|\psi(x_k)^{-1} - \psi(x_{k-1})^{-1}| \leq \Omega_\psi^2 |\psi(x_k) - \psi(x_{k-1})| \quad (k = 1, 2, \dots, n).$$

* * *

By 4.4.1 we can establish that *applying finitely many times the four basic operations to functions of bounded variation we get again a function of the same type, provided that the above restriction on division is fulfilled.* (We saw before that a similar statement can be made about continuous functions, but the class of monotonic functions does not possess the corresponding property. Cf. 3.32 and 4.11.)

4.5 Jordan's and Lebesgue's theorems on functions of bounded variation

4.51 According to our last result, the difference of two functions which are monotone increasing in an interval $[a, b]$ (i.e., the sum of a monotone increasing and of a monotone decreasing function) is surely of bounded variation in $[a, b]$. C. JORDAN who introduced the notion of functions of bounded variation about hundred years ago, also showed that this statement is invertible. Thus it has turned out that the functions of bounded variation, though their class is wider than that of monotonic functions, can still be generated by means of the latter.

4.5.1 (JORDAN'S REPRESENTATION THEOREM)¹ *Any function of bounded variation can be written as the difference of two monotone increasing functions.*

P r o o f. Let $f(x)$ be of bounded variation in $[a, b]$. Consider its total variation in a subinterval $[a, b]$ ($a < x < b$) as a function of x , and write briefly $V_f(a, x) = V(x)$.

By (4.4-4) and (4.4-2) we have:

$$V(x+l) - V(x) = V_f(x, x+l) \leq |f(x+l) - f(x)| \quad (0 < l < b-x). \quad (4.5-1)$$

Hence we can conclude on the one hand

$$V(x+l) - V(x) \geq 0$$

¹ Cf. C. Jordan [2].

i.e.

$$V(x+l) \geq V(x) \quad (a < x < x+l < b); \quad (4.5-2)$$

on the other hand, also $V(x+l) - V(x) \geq f(x+l) - f(x)$ and therefore

$$V(x+l) - f(x+l) \geq V(x) - f(x) \quad (a < x < x+l < b). \quad (4.5-3)$$

Since both $V(x)$ and $V(x) - f(x)$ are monotone increasing by (4.5-2) — (4.5-3), our assertion follows immediately from the decomposition

$$f(x) = V(x) - [V(x) - f(x)]. \quad (4.5-4)$$

* * *

Note that the representation

$$f(x) = \frac{1}{2} [V(x) + f(x)] - \frac{1}{2} [V(x) - f(x)] \quad (4.5-5)$$

satisfies the requirements of 4.5.1, too. (4.5-1) implies at the same time $V(x+l) - V(x) \geq -[f(x+l) - f(x)]$ i.e.,

$$V(x+l) + f(x+l) \geq V(x) + f(x)$$

which means that $V(x) + f(x)$ is monotone increasing.

4.5.2 On the basis of the above proposition, many properties of monotonic functions can be transferred at once to functions of bounded variation. Thus by 4.1.1 and 4.1.2 we see that

4.5.2 *If $f(x)$ is of bounded variation in $[a, b]$, then it is continuous in that interval, except at most countably infinitely many values of x ; furthermore, $f(x)$ has at every inner point of discontinuity finite one-sided limits and at each end point a limit taken from the inner side.*

Lebesgue's theorem 4.2.1 yields the important corollary:

4.5.3 *Any function of bounded variation is differentiable almost everywhere.*

Let us observe that 4.5.2 and 4.5.3 disclose the "global" connection of functions of bounded variation with continuity and differentiability. It is remarkable that, accordingly, though the diagram of a function of bounded variation may contain some discontinuities or break points, the image curve must have a tangent at "most" points.

In spite of the "intertwinedness" between monotony and functions of bounded variation, the wide application possibilities of the latter give sufficient reasons to regard them as a separate class of functions: beside the

relationship with the theory of curves (concept of arc length), they are often used e.g. in modern integral theory and in the theory of Fourier series.

4.53 Such being the case, we have also to touch upon the natural question suggested by Lebesgue's theorem 4.5.3: does it hold for functions of bounded variation (admitting an exceptional null set) one of the most fundamental theorems of calculus, according to which the derivative of a function differentiable on a line segment can vanish only if the function is constant?

Unfortunately, the answer is no; we can easily also find in the narrower class of monotonic functions such a non-constant function whose derivative is 0 everywhere apart from a null set. One of the simplest pertinent examples can be constructed by means of Cantor's ternary set $\mathfrak{T}$ (cf. 2.4.5). Let us denote by $\mathfrak{T}(x)$ the function defined for $0 \leq x \leq 1$ in the following way: (I) $\mathfrak{T}(0)=0$, $\mathfrak{T}(1)=1$; (II) on the 2^{n-1} open intervals left out from $[0,1]$ at the n -th step of the construction of $\mathfrak{T}$, $\mathfrak{T}(x)$ should be equal, in turn, to the constants $1/2^n, 3/2^n, \dots, (2^n - 1)/2^n$ ($n = 1, 2, \dots$); (III) for any point $x_0 \in (0,1)$ of the set $\mathfrak{T}$, we put $\mathfrak{T}(x_0) = \sup \{\mathfrak{T}(x)\}$, where $x \in [0,1] - \mathfrak{T}$ and $0 < x < x_0$.

4.5.4 $\mathfrak{T}(x)$ ($0 \leq x \leq 1$) is a non-constant, monotone increasing and continuous function for which $\mathfrak{T}'(x)$ exists and equals zero almost everywhere.

P r o o f. On the basis of the above definition, $\mathfrak{T}(x)$ is obviously non-constant and monotone increasing. Its *continuity* follows partly from the monotony, partly from the fact that on the (pairwise disjoint) subintervals of $[0,1]$ left out by the first n steps according to the definition of $\mathfrak{T}$, the function $\mathfrak{T}(x)$ takes the values $k/2^n$ ($k = 1, 2, \dots, 2^n - 1$). The monotony implies that $\mathfrak{T}(x)$ has unilateral limits at all the inner points and a right-hand limit at 0, a left-hand limit at 1 (cf. 4.1.1). And any of these must be equal to the corresponding function value, because our second observation shows: the whole codomain of $\mathfrak{T}(x)$ is *everywhere dense* in $[0,1]$ [cf. (4.1-2)].

Now, the existence a.e. of $\mathfrak{T}'(x)$ is a consequence of 4.2.1 (or of 4.5.3). The value of the derivative is 0 at almost every point $x \in [0,1]$, since on the one hand $\mathfrak{T}$ is a null set (cf. 2.6.2); on the other hand, $\mathfrak{T}(x)$ is defined to be constant on all the open sets constituting the set $[0,1] - \mathfrak{T}$.

4.6 Absolutely continuous functions

4.61 Considering the above, it seems desirable to develop further the concept of a function of bounded variation in order that the "anomaly" indicated by examples of the type 4.5.4 should be eliminated. This is indeed

possible and leads to such a class of functions which takes a prominent part in modern integral theory.

Let $f(x)$ be a function defined in a finite interval $[a, b]$. Instead of an ordinary partition of $[a, b]$, let us start with an arbitrary system $[\alpha_k, \beta_k]$ ($k = 1, 2, \dots, n$) of non-overlapping segments $\subseteq [a, b]$ and form the associated "deviation sum"

$$\sum_{k=1}^n |f(\beta_k) - f(\alpha_k)|. \quad (4.6-1)$$

If not only the boundedness of the totality of sums (4.6-1) for all admitted systems $\{[\alpha_k, \beta_k]\}$ is required, but we demand also that this sum tends to 0 whenever the total length of the system $\{[\alpha_k, \beta_k]\}$ itself tends to 0 no matter how, then the function $f(x)$ is said to be *absolutely* (completely) *continuous* in $[a, b]$.

It is easy to see that the definition may be formulated in a slightly modified form: the condition just mentioned is equivalent to the assumption that any $\varepsilon > 0$ being given, we should have

$$\sum_k |f(\beta_k) - f(\alpha_k)| < \varepsilon \quad (4.6-2)$$

for whatever *finite or countably infinite* system of non-overlapping intervals $[\alpha_k, \beta_k] \subseteq [a, b]$, provided that the total length of the latter is small enough, i.e.,

$$\sum_k (\beta_k - \alpha_k) < \delta. \quad (4.6-3)$$

In fact, in the case of infinitely many intervals $[\alpha_k, \beta_k]$, the sums (4.6-2)—(4.6-3) are infinite series of non-negative terms, the convergence of which can be replaced by the boundedness of the corresponding partial sums.¹

4.62 The absolute continuity clearly implies *continuity*, moreover *uniform continuity* in the same interval, but not conversely. Further, of course, every absolutely continuous function is at the same time of *bounded variation*, but there are trivially such functions of bounded variation at which absolute continuity fails. Any piecewise monotonic function with some discontinuities belong to this latter class; and we shall see below that the function $\mathfrak{T}(x)$ in 4.5.4 is not absolutely continuous either, though it is not only of bounded variation, but also continuous everywhere. Consequently, the class of absolutely continuous functions is narrower than that of continuous functions or the class of functions of bounded variation.

Anyway, the most usual examples of absolutely continuous functions are those, the difference quotient of which is uniformly bounded in an interval $[a, b]$, i.e., which satisfy a condition of the form:

$$|f(x_1) - f(x_2)| \leq K|x_1 - x_2| \quad (x_1, x_2 \in [a, b]; \quad K > 0 \text{ constant}). \quad (4.6-4)$$

¹ It can also be verified simply that we may write $|\sum_k [f(\beta_k) - f(\alpha_k)]|$ instead of the sum (4.6-2) in this latter definition form.

This is the so-called *ordinary Lipschitz condition*.¹ Thus all the functions *differentiable* on a closed segment are absolutely continuous.

In order to fix shortly the *relations between the function classes* discussed hitherto, let us denote the set of functions which are differentiable, satisfying an ordinary Lipschitz condition, absolutely continuous, continuous, of bounded variation, monotonic in some interval $[a, b]$, in turn by D, Lip, AC, C, V, M . Then we have:

$$D \subset Lip \subset AC \subset \begin{cases} V \\ C \end{cases}; \quad \text{and} \quad M \subset V.$$

Like the class of functions of bounded variation, the totality of absolutely continuous functions is also essentially *closed with respect to the four basic operations*. Namely, the following pendant of proposition 4.4.1 holds, whose proof can be carried out entirely analogously:

4.6.1 *If two functions are absolutely continuous on a segment $[a, b]$, then their sum, difference, product and in case of a denominator $\neq 0$, also their quotient have the same property.*

4.63 As regards the deeper properties of absolutely continuous functions, Jordan's theorem 4.5.1 can be transferred to this function class in the form:

4.6.2 *Every absolutely continuous function can be represented as the difference of two monotone increasing and absolutely continuous functions.*

P r o o f. Let $f(x)$ be absolutely continuous on a segment $[a, b]$, and denote its total variation $V_f(a, x)$ relating to a subinterval $[a, x]$ ($a < x \leq b$) briefly by $V(x)$.

Since $f(x)$ is *a fortiori* of bounded variation, $V(x)$ is finite at any point $x \in [a, b]$, and both $V(x)$ and $V(x) - f(x)$ are monotone increasing by (4.5-2)—(4.5-3). Considering the decomposition

$$f(x) = V(x) - [V(x) - f(x)] \quad (4.6-5)$$

as it occurs in the verification of 4.5.1 too, as well as in theorem 4.6.1, it remains only to show that $V(x)$ *should be absolutely continuous in our case*.

To that end, let a number $\varepsilon > 0$ be given, and set out from the following statement: because of the absolute continuity of f , there exists such a number $\delta > 0$ that for *every* system of non-overlapping intervals $\subseteq [a, b]$ whose total length is less than δ , the corresponding "deviation sum" of the

¹ Cf. 12.32.

type (4.6-1) is smaller than $\varepsilon = \varepsilon/(b-a+1)$. Let $[\alpha_k, \beta_k]$ ($k = 1, 2, \dots, n$) be such an individual interval system for which $\sum_{k=1}^n (\beta_k - \alpha_k) < \delta$.

On the basis of the definition of total variation, there is a partition $\alpha_k = t_{0k} < t_{1k} \dots < t_{p_k, k} = \beta_k$ of any $[\alpha_k, \beta_k]$, such that the associated deviation sum approaches $V_f(\alpha_k, \beta_k)$ from below by an approximation error less than $\eta_k = \varepsilon^*(\beta_k - \alpha_k)$, and so

$$V(\beta_k) - V(\alpha_k) = V_f(\alpha_k, \beta_k) < \sum_{j=1}^{p_k} |f(t_{jk}) - f(t_{j-1, k})| + \eta_k. \quad (4.6-6)$$

In view of the fact that the total length of the unified system of intervals $[t_{j-1, k}, t_{jk}]$ ($j = 1, 2, \dots, p_k; k = 1, 2, \dots, n$) is $\sum_{k=1}^n (\beta_k - \alpha_k)$, i.e., also $< \delta$, we can write according to the above:

$$\sum_{k=1}^n \sum_{j=1}^{p_k} |f(t_{jk}) - f(t_{j-1, k})| < \varepsilon^*. \quad (4.6-7)$$

But by (4.6-6)—(4.6-7), the deviation sum of $V(x)$ relating to the interval system $[\alpha_k, \beta_k]$ ($k = 1, 2, \dots, n$) satisfies:

$$\sum_{k=1}^n |V(\beta_k) - V(\alpha_k)| < \sum_{k=1}^n \sum_{j=1}^{p_k} |f(t_{jk}) - f(t_{j-1, k})| + \sum_{k=1}^n \eta_k <$$

$$< \varepsilon^* + \sum_{k=1}^n \varepsilon^*(\beta_k - \alpha_k) < \varepsilon^*(1 + b - a) = \varepsilon;$$

and because this holds for any interval system of the above mentioned type (with total length $< \delta$), we get the absolute continuity of $V(x)$ in $[a, b]$.

* * *

4.64 Any absolutely continuous function being also of bounded variation, Lebesgue's theorem 4.5.3 implies immediately that *every absolutely continuous function is differentiable almost everywhere*. Therefore, absolute continuity on a line segment $[a, b]$ is on the one hand — as seen before — a property which is stronger than continuity everywhere and much weaker than differentiability everywhere; on the other hand, it still assures the existence of a derivative apart from the points of a null set.

Thus it is understandable that in numerous classical theorems of analysis which concern the connection between differentiation and integration, assumptions on the derivability of certain occurring functions can be weakened to conditions of absolute continuity. The basis of our later pertinent

results is the following theorem, which shows simultaneously that the insufficiency of the concept of a function of bounded variation, as was discussed at the end of the previous section, holds off in case of an absolutely continuous function:

4.6.4 *If the derivative of an absolutely continuous function vanishes on a segment $[a, b]$ almost everywhere, then the function is constant in $[a, b]$.*

P r o o f. Let $f(x)$ be a function which is absolutely continuous in $[a, b]$, and for which $f'(x)$ vanishes on this segment except for a null set. Let E_x denote that subset of an interval $[a, x]$ ($a < x < b$) on which $f' = 0$.

After prescribing a number $\varepsilon > 0$, we can find to any point $t \in E_x$ a right-hand neighbourhood $I = [t, t + l]$ in which

$$\frac{|f(u) - f(t)|}{u - t} < \varepsilon \quad (t < u \leq t + l). \quad (4.6-8)$$

If we assign to each point of E_x such an interval I , these cover together the set E_x and satisfy the conditions of Sierpinski's lemma 2.6.3. Thus in the case of any $\delta > 0$, however small, we can select from $\{I\}$ a finite system of intervals $I_1 = [t_1, t_1 + l_1]$, $I_2 = [t_2, t_2 + l_2]$, ..., $I_p = [t_p, t_p + l_p]$ ($a \leq t_1 \leq \dots \leq t_p < t_p + l_p \leq x$) which also covers E_x , apart from certain points $t \in E_x$ constituting a subset of outer (L) -measure $< \delta$.

Consider the complementary set $J_x = [a, x] - \bigcup_{k=1}^p I_k$ which is the union of the intervals $[a, t_1]$, $[t_1 + l_1, t_2]$, ..., $[t_{p-1} + l_{p-1}, t_p]$, $[t_p + l_p, x]$. (Admitting that some of the latter should "degenerate" to single points.) The total length of the system J_k is less than δ ; namely the complement of E_x is a null set, and so it can be written:

$$x - a \leq \bar{\mu}_L(E_x) + \bar{\mu}_L([a, x] - E_x) < \sum_{k=1}^p l_k + \delta$$

whence $(x - a) - (l_1 + l_2 + \dots + l_p) < \delta$.

We supposed f to be absolutely continuous; therefore δ can be chosen so small that the deviation sum of f of the type (4.6-1), belonging to J_x , should be $< \varepsilon$. Then using the notation

$$\begin{aligned} \sum_{J_x} = & [f(t_1) - f(a)] + [f(t_2) - f(t_1 + l_1)] + \dots + [f(t_p) - f(t_{p-1} + l_{p-1})] + \\ & + [f(x) - f(t_p + l_p)], \end{aligned}$$

we shall have

$$|\sum_{J_x}| < \varepsilon. \quad (4.6-9)$$

On the other hand, the application of (4.6-8) yields:

$$|\sum_{J_x} - [f(x) - f(a)]| \leq \sum_{k=1}^p |f(t_k) - f(t_k + l_k)| < \sum_{k=1}^p \varepsilon l_k < \varepsilon(x - a). \quad (4.6-10)$$

After all, by (4.6-9)—(4.6-10) we conclude

$$|f(x) - f(a)| \leq |\sum_{J_x}| + |f(x) - f(a) - \sum_{J_x}| < \varepsilon(1 + x - a)$$

and herewith, ε being arbitrarily small,

$$f(x) = f(a).$$

Since x may be any point of $[a, b]$, we see that the premises of our theorem involve the constancy of $f(x)$ ($a \leq x \leq b$), as asserted above.

* * *

We have to mention also an immediate consequence of 4.6.4: *the function $\mathfrak{T}(x)$, presented in 4.53, is plainly not absolutely continuous*, since otherwise it would be a constant by our last result.

4.65 It is clear that each of the function classes treated in this chapter *can be extended to Euclidean spaces of higher dimension*.

In fact, a function $f(x)$ monotone increasing on a line segment (a, b) can be characterized in such a way that the corresponding interval function $\Delta_f(x_1, x_2) = f(x_1) - f(x_2)$ ($x < x_2$) takes only non-negative values on every subinterval (x_1, x_2) of (a, b) ; and $\Delta_f(x_1, x_2)$ is additive in the sense that $\Delta_f(x_1, x_3) = \Delta_f(x_1, x_2) + \Delta_f(x_2, x_3)$ ($x_1 < x_2 < x_3$). Furthermore, the sum (4.4-1) can be written also as

$$\sum_{k=1}^n |\Delta_f(x_{k-1}, x_k)|. \quad (4.6-11)$$

Thus in the ν -dimensional Euclidean space E_ν , the role of monotone increasing functions is taken over by non-negative, additive interval functions defined for ν -dimensional intervals. (Writing here “non-positive” instead of “non-negative”, we get the corresponding concept of monotone decreasing function.) A ν -dimensional interval function is said to be of bounded variation, if the ν -dimensional pendant of the sum (4.6-11) is bounded. We define analogously the absolutely continuous functions of higher dimension, using a suitable modification of the sum (4.6-2) involving interval function.

Though the main results found for the one-dimensional case remain valid also for interval functions of several dimensions, we cannot give a detailed account of this kind of generalization within this framework.

JORDAN MEASURE AND RIEMANN INTEGRAL. RIEMANN-STIELTJES INTEGRAL

5.1 Foundations of Jordan's measure theory

5.11 The most general function class of elementary analysis — that of functions integrable in classical sense — arises from RIEMANN'S concept of integral,¹ used mostly in the case of continuous or piecewise continuous functions (CAUCHY). It is also well-known that the notion of Riemann integral for a function of one variable or two variables is in close connection with the elementary (JORDAN) theory of contents. Thus, before we summarize briefly the most important knowledge of the classical notion of integral and develop it further by means of our recent findings, it is advisable first to deal with the geometric background contemplating point sets of however many dimensions.

The *measure theory* intends to extend the concept of content (length, area, volume) to various set structures, keeping in view certain natural requirements. The characteristic quantity attached to a set E in this manner is called the *measure of E* and is usually denoted by $\mu(E)$; every measure yields an example of one-valued *set function* (cf. 3.12). The pertinent natural requirements for μ are: I. $\mu(\emptyset) = 0$; II. $0 \leq \mu(E) \leq \infty$ (*definiteness*); III. the measure of the union of pairwise disjoint sets should be equal to the sum of measures of the individual sets (*additivity*). We shall see that the third postulate leads to essentially different concepts of measure if the number of joined sets must be finite, and if we admit also countably infinitely many sets to join. Note still that in all concrete cases to be discussed, $\mu(E)$ is also *movement invariant* in the sense that congruent sets have the same measure.

As far as the point sets of E_ν are concerned, the first step has been taken already in 2.62: by the *measure of a ν -dimensional interval I* we mean the product $l_1 \cdot l_2 \cdot \dots \cdot l_\nu$ where l_k is the length of the k -th line segment determining I ($k = 1, 2, \dots, \nu$); we denote this measure briefly by $|I|$. To prepare further steps, we recall a well-known classical method of elementary geometry: the use of inscribed and circumscribed polygons in planimetry. For instance, the area of a circular disc is arbitrarily approached by the contents of circumscribed polygons from above, by those of inscribed polygons from below; and these polygons may plainly be replaced by certain

¹ The main part of this theory of integral is to be found in a famous work of B. RIEMANN (1854). Cf. Riemann [1].

finite systems of non-overlapping two-dimensional intervals covering the disk or belonging to its interior, respectively.

5.12 If this model is followed in case of an arbitrary bounded ν -dimensional point set E , then we have to consider on the one hand the infimum of the measure sums of all those *finite* systems consisting of ν -dimensional intervals, which cover E ; on the other hand, the supremum of the measure sums of all such *finite* systems of non-overlapping ν -dimensional intervals which consist of points of E only. It is clear that the former number is nothing but the *outer* (exterior) (J)-measure of E , introduced already in 2.62 and denoted by $\bar{\mu}_J(E)$; correspondingly, the last mentioned number is called the *inner* (interior) (J)-measure of E and denoted by $\underline{\mu}_J(E)$.

There is also another definition of the inner (J)-measure, equivalent to the previous one:

$$\underline{\mu}_J(E) = |I| - \bar{\mu}_J(I - E), \quad (5.1-1)$$

where I means a (ν -dimensional) closed interval containing E . (5.1-1) makes it possible to reduce the problem of determining the inner (J)-measure of E to that of calculating the outer (J)-measure of a complement of E ; it is obvious that $e \subseteq E$ implies $\underline{\mu}_J(e) \leq \underline{\mu}_J(E)$.

Now, a point set $E \subset E_\nu$ is said to be *Jordan measurable*, briefly (J)-measurable, if its outer and inner (J)-measures are equal. In this case, we mean by the *Jordan measure* [in short (J)-measure] of E the non-negative number $\bar{\mu}_J(E) = \underline{\mu}_J(E)$; it is simply denoted by $\mu_J(E)$.¹

On the basis of the above definitions we see that

$$0 \leq \underline{\mu}_J(E) \leq \bar{\mu}_J(E). \quad (5.1-2)$$

Thus $\bar{\mu}_J(E) = 0$ implies at the same time $\underline{\mu}_J(E) = 0$; hence it follows also that $\mu_J(E)$ exists and equals 0.

Examples: any finite point set or convergent sequence of points has the (J)-measure 0. The same holds for every bounded, closed null set E , since by 2.6.1, we can select from any system of open intervals covering E a finite subsystem which also contains E . Furthermore, also (J)-measurable is any finite system of finite, non-overlapping intervals in arbitrary space E_ν (the (J)-measure of the system equalling the sum of measures of the individual intervals); in the plane, the interior of any rectifiable closed curve is again a (J)-measurable set, etc. On the other hand, the set consisting of the rational points of $[0,1]$ (and so its complement, i.e., the set of irrational points) is not (J)-measurable. In this case we have namely $\bar{\mu}_J = 1$, but $\underline{\mu}_J = 0$ because the set does not contain a single line segment.

We can easily verify the following useful criterion of measurability:

¹ C. Jordan [2].

5.1.1 A bounded point set is (J) -measurable if and only if the set of its boundary points has the (J) -measure 0.

5.1.3 The notion of (J) -measure evidently satisfies the general measure theoretical requirements mentioned in 5.11. In particular, the definition of μ_J and the properties of $\bar{\mu}_J$ yield:

5.1.2 (I) If E is a (J) -measurable set and e is a (J) -measurable subset of it, then $\mu_J(E - e) = \mu_J(E) - \mu_J(e)$ and $e \subseteq E$ implies $\mu_J(e) \leq \mu_J(E)$. (Monotony.)

(II) The union of finitely many, pairwise disjoint, (J) -measurable sets is also (J) -measurable, and the (J) -measure of the latter is equal to the sum of (J) -measures of the individual sets. (Finite additivity.)

Regarding the second property we add: 1. The union of infinitely many (J) -measurable sets is not necessarily (J) -measurable, e.g. the set of rational numbers in $[0,1]$ — as mentioned before — has no (J) -measure, though each individual point of the set is trivially (J) -measurable with (J) -measure 0. 2. The outer (J) -measure is in general not additive; for $\bar{\mu}_J$, the pendant of 5.1.2 (II) is the inequality mentioned in 2.64 (so-called subadditivity).¹ 3. In case of two sets, there holds a more general statement:

5.1.3 If E_1 and E_2 are (J) -measurable point sets, then their union and intersection also have this property, moreover

$$\mu_J(E_1 \cup E_2) = \mu_J(E_1) + \mu_J(E_2) - \mu_J(E_1 \cap E_2). \quad (5.1-3)$$

Therefore we have

$$\mu_J(E_1 \cup E_2) \leq \mu_J(E_1) + \mu_J(E_2), \quad (5.1-4)$$

with equality if and only if the (J) -measure of $E_1 \cap E_2$ is 0.

P r o o f. (5.1-3) is obvious, if E_1 and E_2 are intervals. Otherwise, we can proceed in the following manner:

¹⁰ We show that arbitrary bounded point sets e_1, e_2 fulfil the condition

$$\bar{\mu}_J(e_1 \cup e_2) + \bar{\mu}_J(e_1 \cap e_2) \leq \bar{\mu}_J(e_1) + \bar{\mu}_J(e_2). \quad (5.1-5)$$

Let us denote by $\{i'\}$ a finite system of intervals covering e_1 , by $\{i''\}$ another such system covering e_2 , and let the total measures of these be μ_1 and μ_2 , respectively. Forming $I' = \cup i'$ and $I'' = \cup i''$, the set $I' \cup I''$

¹ Cf. e.g. the decomposition of $[0,1]$ into rational and irrational points; both subsets and also the interval $[0,1]$ have an outer (J) -measure of value 1. As can be seen, a usable concept of measure can be obtained only by applying the outer and inner measures simultaneously.

will contain $e_1 \cup e_2$. But $I' \cup I''$ may also be regarded as the union of three such systems of non-overlapping intervals, say I_1, I_2, I_3 , among which $I_1 \supseteq e_1 - m$, $I_2 \supseteq m$, $I_3 \supseteq e_2 - m$ with $m = e_1 \cap e_2$. If the total measure of I_r is denoted by σ_r ($r = 1, 2, 3$), then we can write according to the definition of outer (J)-measure: $\bar{\mu}_J(e_1 \cup e_2) + \bar{\mu}_J(e_1 \cap e_2) \leq (\sigma_1 + \sigma_2 + \sigma_3) + \sigma_2 = (\sigma_1 + \sigma_2) + (\sigma_2 + \sigma_3) \leq \mu_1 + \mu_2$. Hence (5.1-5) follows immediately, because the deviation of μ_1 and $\bar{\mu}_J(e_1)$, further that of μ_2 and $\bar{\mu}_J(e_2)$ can be taken as small as we please.

2⁰ If I is an interval containing $e_1 \cup e_2$, then replacing e_1 and e_2 by their complements with respect to I in (5.1-5) and using the formulae (1.1-5), it results:

$$\bar{\mu}_J[I - (e_1 \cap e_2)] + \bar{\mu}_J[I - (e_1 \cup e_2)] \leq \bar{\mu}_J(I - e_1) + \bar{\mu}_J(I - e_2).$$

Considering (5.1-1), this is equivalent to the inequality

$$\underline{\mu}_J(e_1 \cup e_2) + \underline{\mu}_J(e_1 \cap e_2) \geq \underline{\mu}_J(e_1) + \underline{\mu}_J(e_2). \quad (5.1-6)$$

3⁰ Assume E_1, E_2 to be (J)-measurable and apply to them (5.1-5)—(5.1-6). We get

$$\bar{\mu}_J(E_1 \cup E_2) + \bar{\mu}_J(E_1 \cap E_2) \leq \mu_J(E_1) + \mu_J(E_2) \leq \underline{\mu}_J(E_1 \cup E_2) + \underline{\mu}_J(E_1 \cap E_2);$$

and this is possible by (5.1-2) only if both $E_1 \cup E_2$ and $E_1 \cap E_2$ are (J)-measurable. This being the case, the inequalities (5.1-5) and (5.1-6) yield together (5.1-3).

* * *

It turns out from propositions 5.1.2 and 5.1.3: the (J)-measurable sets of the space E_ν constitute a *Boolean ring* (Cf. 1.51.).

5.2 Riemann integral of a function of one variable as Jordan measure with sign; upper and lower Darboux integral

5.21 Let us apply the foregoing in case of certain planar sets!

Consider a function $y = f(x)$, defined in a finite interval $a \leq x \leq b$, single-valued and bounded there; let us call the totality of the ordinates $f(x)$ belonging to $[a, b]$ (as a point set) the *set of ordinates* of f on $[a, b]$ and denote it by $Y(f; [a, b])$.

For $f(x) \geq 0$, this set can be written in the form:

$$Y(f; [a, b]) = \{(x, y) : a \leq x \leq b, 0 \leq y \leq f(x)\};$$

its outer and inner (J)-measure are the so-called upper and lower Darboux integrals of f in $[a, b]$, denoted briefly by the symbols

$$\int_a^b f(x) dx \quad \text{and} \quad \int_{\bar{a}}^{\bar{b}} f(x) dx,$$

respectively. If the last-mentioned two numbers are equal, i.e., if $Y(f; [a, b])$ is (J)-measurable, then $f(x)$ is said to be RIEMANN integrable, or simply (R)-integrable in $[a, b]$. The Jordan measure (content) of the planar set $Y(f; [a, b])$ is the well-known Riemann integral [briefly (R)-integral] of $f(x)$ relating to $[a, b]$, denoted by

$$\int_a^b f(x) dx.$$

Take now a partition of $[a, b] : a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$, and let us denote by M_k and m_k the supremum and infimum of the function values $f(x)$ belonging to $[x_{k-1}, x_k]$ (the " k -th subinterval"), respectively ($k = 1, 2, \dots, n$). The so-called *upper Darboux sum*

$$S = \sum_{k=1}^n (x_k - x_{k-1}) M_k \quad (5.2-1)$$

and *lower Darboux sum*

$$s = \sum_{k=1}^n (x_k - x_{k-1}) m_k \quad (5.2-2)$$

mean geometrically the total area of a finite system of planar intervals, covering $Y(f; [a, b])$ in the first case, and consisting only of points of $Y(f; [a, b])$ in the second case, respectively.

Since to any given finite system R of planar intervals which covers $Y(f; [a, b])$, we can find a sum S (corresponding to a suitable partition of $[a, b]$) which does not exceed the total area of R , thus the infimum of the set of sums S associated with all possible partitions of $[a, b]$ cannot be but the

outer (J)-measure of $Y(f; [a, b])$, i.e., $\int_a^b f(x) dx$. We can conclude similarly

that the supremum of the totality of sums s is necessarily $\int_{\bar{a}}^{\bar{b}} f(x) dx$. All this can be expressed briefly by the formulae:

$$\int_a^b f(x) dx = \inf \{S\}, \quad (5.2-3)$$

$$\int_a^b f(x) dx = \sup\{s\}. \quad (5.2-4)$$

From the well-known elementary fact that s increases, and S decreases when new points of subdivision are inserted, furthermore from the boundedness of s and S from above and from below, respectively, it follows that s tends to the lower and S to the upper Darboux integral of f over $[a, b]$ as $\max(x_k - x_{k-1}) \rightarrow 0$.

Therefore, if the values (5.2-3), (5.2-4) are the same, we have

$$\int_a^b f(x) dx = \inf\{S\} = \sup\{s\}; \quad (5.2-5)$$

and here both of the symbols “inf” and “sup” may be replaced by “lim”, if we introduce this latter notation for the limiting process just mentioned, too.

5.22 The sum S and s can also be formed for functions taking values of *arbitrary sign*; only their geometric interpretation changes slightly, if the codomain of f contains both positive and negative numbers. Consequently, the above notions can be transferred mechanically to any *bounded*, one-valued function. Thus the *upper and lower Darboux integral* are defined by (5.2-3)—(5.2-4); we say that $f(x)$ is *Riemann integrable* [(R)-integrable] in $[a, b]$, if both of the latter have the same value i.e. $\inf\{S\} = \sup\{s\}$. In this case the *Riemann integral* [(R)-integral] of f over $[a, b]$ is given by (5.2-5).¹

Instead of S and s , we may consider naturally any so-called *Riemann sum* (“rectangle sum”):

$$\sigma = \sum_{k=1}^n (x_k - x_{k-1}) f(\xi_k), \quad (5.2-6)$$

where ξ_k is an arbitrary point of $[x_{k-1}, x_k]$ ($k = 1, 2, \dots, n$). Since $m_k \leq f(\xi_k) \leq M_k$ for each k , this implying

$$s \leq \sigma \leq S, \quad (5.2-7)$$

we get the following alternative representation of the (R)-integral:

$$\lim_{\max \Delta x_k \rightarrow 0} \sigma = \int_a^b f(x) dx \quad (\Delta x_k = x_k - x_{k-1}). \quad (5.2-8)$$

¹ The terms “ $f(x)$ is (R)-integrable in $[a, b]$ ” and “the (R)-integral of $f(x)$ on $[a, b]$ exists” have the same meaning.

Let us recall at the same time: according to (5.2-8), any given $\varepsilon > 0$ can be associated with a number $\delta = \delta(\varepsilon)$ such that

$$\left| \sigma - \int_a^b f(x) dx \right| < \varepsilon$$

whenever $\max_k \Delta x_k < \delta$, independently of the choice of the "middle" points ξ_k ($k = 1, 2, \dots, n$). This is usually expressed by saying that " σ tends to the (R) -integral $\int_a^b f(x) dx$, as the partition of $[a, b]$ is refined unboundedly".

The formal properties of the classical concept of integral can simply be deduced, as is well-known, from the corresponding features of Jordan measure. Note that this statement can be reversed in some degree, because the outer and inner (J)-measure and the Jordan measure of a bounded linear set can also be written as an upper and lower Darboux integral and a Riemann integral, respectively. More precisely, we have:

$$\bar{\mu}_J(E) = \int_a^b \chi_E dx, \quad \underline{\mu}_J(E) = \int_a^b \chi_E dx, \quad \mu_J(E) = \int_a^b \chi_E dx, \quad (5.2-9)$$

where χ_E denotes the characteristic function of a linear set E , and $[a, b]$ is a finite basic interval containing E .

(5.2-9) obviously makes easier the discussion of some properties of the Jordan measure; for example, (5.1-3) can be obtained from (1.1-10) by integration immediately.

5.3 Riemann's and Lebesgue's criteria for (R) -integrability

5.31 Next we have to deal with the *conditions of Riemann integrability*. First of all, let us introduce the notion of oscillation with respect to an interval and to a point.

Let $f(x)$ be defined, one-valued and bounded in $\alpha \leq x \leq \beta$, and let $M(\alpha, \beta)$, $m(\alpha, \beta)$ denote the supremum and infimum of the values of f in $[\alpha, \beta]$, respectively. The (non-negative) difference

$$\Omega(\alpha, \beta) = M(\alpha, \beta) - m(\alpha, \beta) \quad (5.3-1)$$

is called the *oscillation* of $f(x)$ in $[\alpha, \beta]$. If $\xi \in (\alpha, \beta)$, then the oscillation of $f(x)$ at point ξ is defined as follows:

$$\Omega(\xi) = \lim_{\varepsilon \rightarrow 0} \Omega(\xi - \varepsilon, \xi + \varepsilon). \quad (5.3-2)$$

The existence of the last limit is assured by the simultaneous existence of

$$M(\xi) = \lim_{\varepsilon \rightarrow +0} M(\xi - \varepsilon, \xi + \varepsilon), \quad m(\xi) = \lim_{\varepsilon \rightarrow +0} m(\xi - \varepsilon, \xi + \varepsilon), \quad (5.3-3)$$

i.e., of the so-called *upper and lower fluctuation value* of f at ξ . [Cf. (3.2-6).] Since $\Omega(\xi - \varepsilon, \xi + \varepsilon)$ decreases as $\varepsilon \rightarrow +0$, the inequality holds:

$$\Omega(\xi) = M(\xi) - m(\xi) \leq \Omega(\alpha, \beta). \quad (5.3-4)$$

Hence and from the 3.3.1 (II) criterion of continuity we see that *the points of continuity of $f(x)$ are characterized by the condition $\Omega(\xi) = 0$ and its points of discontinuity by $\Omega(\xi) > 0$.*

Let us form now

$$S - s = \sum_{k=1}^n (x_k - x_{k-1}) \Omega(x_{k-1}, x_k), \quad (5.3-5)$$

the so-called *oscillation sum* [cf.(5.2-1)—(5.2-2)]. Since the (R) -integrability means — as said before — that S and s tend to the same number as $\max(x_k - x_{k-1}) \rightarrow 0$, it follows:

5.3.1 (RIEMANN'S CRITERION OF INTEGRABILITY) *The Riemann integral of a function $f(x)$, bounded in a finite interval $[a, b]$, exists there if and only if the oscillation sum (5.3-5) tends to 0 as the partition of $[a, b]$ is refined unboundedly.*

5.3.2 This condition, though significant theoretically, is often not easy to apply in case of concrete integrands; besides, it yields no really descriptive characterization of (R) -integrable functions. Both requirements are satisfied, however, by another elegant condition, found by LEBESGUE at the beginning of our century.

5.3.2 (LEBESGUE'S CRITERION)¹ *A function $f(x)$ is Riemann integrable in a finite interval $[a, b]$ if and only if it is bounded there and continuous almost everywhere.*

P r o o f. ¹⁰ (*Necessity.*) Let us suppose that the (R) -integral $\int_a^b f(x) dx$ exists. Then $f(x)$ must be bounded in $[a, b]$; or else there would exist such a subinterval of $[a, b]$, where the infimum and supremum of the function values are not finite so that the sums (5.2-1), (5.2-2) would become meaningless. We have still to show that the points of discontinuity of $f(x)$ in $[a, b]$ constitute a null set.

¹ Lebesgue [4].

To that end consider first the set E_η of those points $\xi \in [a, b]$, at which $\Omega(\xi)$ is larger than a number $\eta > 0$ prescribed arbitrarily. If E_η is empty, then it is a null set. Otherwise, according to 5.3.1, we can find to any $\varepsilon > 0$ such a partition of $[a, b]$ for which

$$\sum_{k=1}^n (x_k - x_{k-1}) \Omega(x_{k-1}, x_k) < \frac{1}{2} \eta \varepsilon. \quad (5.3-6)$$

From the oscillation sum on the left, select only those terms, which pertain to subintervals containing at least one point of E_η inside.

In the sum thus obtained, each factor $\Omega(x_{k-1}, x_k)$ is $> \eta$ by the definition of E_η and (5.3-4), so that it follows

$$\sum_{k=1}^n (x_k - x_{k-1}) \Omega(x_{k-1}, x_k) > \eta \sum^* (x_k - x_{k-1}), \quad (5.3-7)$$

where the sum $\sum^*$ is to extend over the subintervals just mentioned.

The comparison of (5.3-6)—(5.3-7) yields the inequality

$$\sum^* (x_k - x_{k-1}) < \frac{1}{2} \varepsilon.$$

Hence we see that the totality of those points of E_η which belong to the interior of one of the subintervals, can be covered by a finite system of intervals of total length $< \varepsilon/2$. On the remaining points of E_η (which are themselves also points of subdivision), it holds trivially the same assertion, since the number of these points is finite. Therefore E_η can be covered in its entirety by an interval system of total length smaller than $\varepsilon/2 + \varepsilon/2 = \varepsilon$, i.e., E_η is a null set.

Putting here $\eta = 1/n$, we conclude by 2.6.2 that $E = \bigcup_{n=1}^{\infty} E_{1/n}$ must be a null set, too; but E plainly consists of those points ξ of $[a, b]$ at which $\Omega(\xi) > 0$ i.e., of all the points of discontinuity of $f(x)$ in $[a, b]$.

2^0 (*Sufficiency*.) Assuming now that $f(x)$ is bounded and continuous in $[a, b]$ apart from a null set E , we have to verify that the sum (5.3-5) is less than a positive number ε , however small, if only the partition is chosen suitably — this implying by 5.3.1 the existence of $\int_a^b f(x) dx$.

If E is empty, i.e., $f(x)$ is continuous in the whole interval $[a, b]$, then (by the Weierstrass theorem) M_k and m_k mean in the sums (5.2-1)—(5.2-2) simply the maximal or minimal value of f on the segment $[x_{k-1}, x_k]$ ($k = 1, 2, \dots, n$), and so we can write:

$$S - s \leq \max_{1 \leq k \leq n} (M_k - m_k) \sum_{k=1}^n (x_k - x_{k-1}) = (b - a) \max_{1 \leq k \leq n} \Omega(x_{k-1}, x_k). \quad (5.3-8)$$

Since by the uniform continuity of $f(x)$ (Heine's theorem), the last factor on the right becomes arbitrarily small for a partition of $[a, b]$ refined enough, the sum (5.3-5) tends indeed to 0 as $\max(x_k - x_{k-1}) \rightarrow 0$.

If $E \neq \emptyset$, include E in such a system $\{I\}$ consisting of countably infinitely many open intervals which has a total length $< \varepsilon [2\Omega(a, b)]^{-1}$ [cf. (5.3-1)].¹ The oscillation being 0 at any point of continuity, every such point can be covered by an open interval on which the oscillation of f is less than $\varepsilon [2(b-a)]^{-1}$. Let us denote by $\{i\}$ the system of line segments having this latter property.

The interval systems $\{I\}$ and $\{i\}$ together cover $[a, b]$; by 2.6.1 the same can be asserted on a suitable system $\{j\}$ of *finitely many* open intervals, consisting of certain elements of $\{I\}$ and $\{i\}$, respectively. Let $a = x_0 < x_1 < \dots < x_n = b$ be the partition of $[a, b]$ generated by the end points of intervals belonging to $\{j\}$ (a and b also being included). Then any subinterval $[x_{k-1}, x_k]$ is obviously contained by a stretch $\subset \{i\}$; correspondingly, split up the oscillation sum relating to $\{j\}$ into two parts, denoted by $\sum'$ and $\sum''$, respectively.

Considering the definitions of $\{I\}$, $\{i\}$, these latter sums can be estimated in the following way:

$$\sum'(x_k - x_{k-1}) \Omega(x_{k-1}, x_k) \leq \Omega(a, b) \cdot \sum'(x_k - x_{k-1}) <$$

$$< \Omega(a, b) \cdot \varepsilon [2\Omega(a, b)]^{-1} = \frac{\varepsilon}{2},$$

$$\sum''(x_k - x_{k-1}) \Omega(x_{k-1}, x_k) < \varepsilon [2(b-a)]^{-1} \cdot \sum''(x_k - x_{k-1}) \leq$$

$$\leq \varepsilon [2(b-a)]^{-1} \cdot (b-a) = \frac{\varepsilon}{2},$$

so that we get by addition:

$$\sum_{k=1}^n (x_k - x_{k-1}) \Omega(x_{k-1}, x_k) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Qu.e.d.

5.4 Deeper study of (R) -integrable functions; limiting process under the (R) -integral sign

5.41 After these, it is easy to clear the bearings of the *Riemann integrability* upon other elementary properties of functions.

¹ This is possible, because E is supposed to be a null set.

First we want to note: *any function continuous in a finite closed interval is there (R)-integrable*; of course, the assertion does hold a fortiori for differentiable functions.¹ In fact, by the Weierstrass theorem, any function continuous in a finite closed interval is also bounded, and therefore the assertion is a simple corollary of 5.3.2. (Otherwise, a little part of the proof of the theorem just mentioned — cf. (5.3-8) — yields a direct verification, too.)

We also establish: *if $f(x)$ is monotonic or of bounded variation in a finite interval $[a, b]$, then it is there (R)-integrable*. In fact, such a function $f(x)$ is in $[a, b]$ bounded and differentiable almost everywhere, therefore it is also p.p. continuous. (Cf. 4.2.1 and 4.5.3.) Hence by 5.3.2 we conclude at once the (R)-integrability of f . (We can argue a simpler way, if we use in the case of a monotonic $f(x)$ the inequality [cf. (4.4-3)]

$$S - s \leq \max(x_k - x_{k-1}) \cdot |f(b) - f(a)|$$

and 5.3.1, otherwise we refer to 4.5.1.)

As it can be seen, *the class of functions (R)-integrable on a stretch $[a, b]$ contains simultaneously the totality of continuous, differentiable, monotonic functions and those of bounded variation in the same interval* — but, naturally, a lot of other functions, too. For instance, the function

$$y = \begin{cases} x \sin(1/x) & (x \neq 0) \\ 1 & (x = 0) \end{cases}$$

(which differs from (4.1-1) at one single point) is continuous everywhere with the exception of $x = 0$, and thus it is (R)-integrable in any finite neighbourhood of the origin; nevertheless, it does not belong to any of the four function classes specified above. (Cf. 4.42.)

On the other hand, there are trivial examples of functions *non-(R)-integrable* in a finite interval. According to 5.3.2, all unbounded functions, e.g. x^α with $\alpha < 0$ or $\log x$ in $[0, 1]$ as well as $\tan x$ and $\cot x$ in $[0, \pi/2]$, are of this type. Moreover, the *Dirichlet function* $\chi(x)$ [cf. (1.6-1)] is bounded, but not (R)-integrable in every interval $[a, b]$, because it is not continuous at any point.

Thus the ordinate set of $\chi(x)$ belonging to an arbitrary interval yields an example of a planar point set which is not (J)-measurable.

5.42 As far as the *basic operations with (R)-integrable functions* are concerned, the criterion 5.3.2 proves again to be very useful. On this basis, it is clear (cf. 3.32):

¹ Here and in the sequel, *continuity or differentiability on a closed interval* is always meant in that way that the function has the property in question *at the end points from the inner side only*.

5.4.1 The application finitely many times of the module-forming, as well as of the four basic operations does not take out from the class of Riemann integrable functions, if in the case of a quotient its boundedness is also required.

It can be verified similarly: any composite function of the form $\varphi(x) = g[f_1(x), \dots, f_\nu(x)]$ is (R) -integrable on such a segment $a \leq x \leq b$, where each of the "inner" functions $t_1 = t_1(x), \dots, t_\nu = t_\nu(x)$ is (R) -integrable, provided that the "outer" function $g(t_1, t_2, \dots, t_\nu)$ is bounded and continuous on the ν -dimensional set of the corresponding tuples $(t_1, t_2, \dots, t_\nu)$.

We have another situation when applying infinitely many steps, e.g. certain limiting processes to (R) -integrable functions. As we know from elementary analysis, any uniformly convergent sequence or series of continuous functions may be integrated term-by-term, and the proof of this statement can be transferred also to the case of arbitrary (R) -integrable functions. Assuming, however, only ordinary convergence instead of the uniform one, the (R) -integrability on a common interval $[a, b]$ of the terms of $\{f_n(x)\}$ or of $\sum_{n=1}^{\infty} f_n(x)$ does not imply the (R) -integrability in $[a, b]$ of $\lim_{n \rightarrow \infty} f_n(x)$ and $\lim_{n \rightarrow \infty} [f_1(x) + \dots + f_n(x)]$; not even in the case if the terms and partial sums, respectively, are majorized (in absolute value) by a common constant.

Thus it can be stated:

5.4.2 The integral relations

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \left[\lim_{n \rightarrow \infty} f_n(x) \right] dx, \quad (5.4-1)$$

$$\int_a^b f_1(x) dx + \int_a^b f_2(x) dx + \dots = \int_a^b [f_1(x) + f_2(x) + \dots] dx \quad (5.4-2)$$

are generally not valid for (R) -integrals.

5.4.3 A concrete example: let $\varrho_1, \varrho_2, \dots$ be a sequence containing all the rational numbers and each of them exactly once (cf. 1.2.1), further consider the function sequence

$$f_n(x) = \begin{cases} 1 & \text{for } x = \varrho_k \quad (k = 1, 2, \dots, n), \\ 0 & \text{for any other } x. \end{cases} \quad (5.4-3)$$

Since $f_n(x)$ differ from the function identically 0 only at finitely many points, it is trivially (R) -integrable and its (R) -integral has the value 0. So we have

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \lim_{n \rightarrow \infty} 0 = 0.$$

On other hand, $\lim_{n \rightarrow \infty} f_n(x)$ is nothing else but the function $\chi(x)$ mentioned above, which is not (R) -integrable on any $[a, b]$. Herewith we showed

that the integral on the right of (5.4-1) does not exist in our case, and therefore (5.4-1) cannot hold.

Obviously $f_1(x) + [f_2(x) - f_1(x)] + [f_3(x) - f_2(x)] + \dots$ [where $f_n(x)$ has the meaning (5.4-3)] is a series of functions for which (5.4-2) is not true.

5.5 Use of a primitive; properties of the indefinite (R)-integral

5.51 In the classical analysis, the so-called *Leibniz-Newton formula* on the connection of integration and derivation :

$$\int_a^b f(x) dx = F(b) - F(a) \quad (5.5-1)$$

is of fundamental importance. Here $f(x)$ is supposed to be continuous in $[a, b]$, while $F(x)$ is a *primitive* of it, i.e., a function whose derivative is $f(x)$ in the whole interval $a \leq x \leq b$.¹

It is easy to see that (5.5-1) holds not only for a continuous integrand but also for the most general natural condition:

5.5.1 If $f(x)$ is (R)-integrable in a finite interval $[a, b]$ and has there a primitive $F(x)$, then we have (5.5-1).

P r o o f. For any partition $a = x_0 < x_1 < \dots < x_n = b$, the mean-value theorem of Lagrange assures the existence of a point $\xi_k \in (x_{k-1}, x_k)$ satisfying $F(x_k) - F(x_{k-1}) = (x_k - x_{k-1})F'(\xi_k)$ ($k = 1, 2, \dots, n$); and by $F'(x) = f(x)$ ($a \leq x \leq b$), the last factor equals $f(\xi_k)$. Thus we can write:

$$F(b) - F(a) = \sum_{k=1}^n [F(x_k) - F(x_{k-1})] = \sum_{k=1}^n (x_k - x_{k-1})f(\xi_k),$$

which becomes (5.5-1) as $\max (x_k - x_{k-1}) \rightarrow 0$. [Cf. (5.2-8).]

* * *

5.52 As regards the existence and discussion of primitives, in the first place we recall the elementary fact that a function f continuous in $[a, b]$ always has primitives and their totality can be written in the form:

$$F(x) = \int_a^x f(t) dt + C \quad (a \leq x \leq b), \quad (5.5-2)$$

C denoting an arbitrary constant.

¹ At points a and b , $F'(x)$ is to be replaced by $F'_+(a)$ and $F'_-(b)$, respectively. [Cf. (1.6-2).]

Now, it is an immediate consequence of the above:

5.5.2 (I) Any primitive of an (R) -integrable function f , if it exists at all, is necessarily of the form (5.5-2).

(II) There are (R) -integrable functions having no primitive.¹

P r o o f. (I) If f is (R) -integrable and has a primitive F on $[a, b]$, then putting x instead of b , and t instead of x in (5.5-1), we obtain:

$$\int_a^x f(t) dt = F(x) - F(a) \quad (x \in [a, b]), \quad (5.5-3)$$

i.e.,

$$F(x) = \int_a^x f(t) dt + F(a) \quad (a \leq x \leq b).$$

(II) Let $f^*(x)$ be a function which takes values different from 0 at certain finitely many points $x \in [a, b]$, and vanishes otherwise. Since f^* is (R) -integrable in $[a, b]$, any possible primitive of it should have the form $\int_a^x f^*(t) dt + C = C$ (constant), whose derivative is, however, identically zero and thus is not equal to $f^*(x)$ at every point $x \in [a, b]$. Therefore f^* has no primitive in $[a, b]$.

* * *

5.53 By what has gone before, it is reasonable to deal separately with the function:

$$\Phi(x) = \int_a^x f(t) dt \quad (x \in [a, b]), \quad (5.5-4)$$

in case of an arbitrary (R) -integrable f ; $\Phi(x) + C$ (C constant) will be called the (R) -integral function or indefinite (R) -integral of $f(x)$. We shall need below the *Lipschitz condition* (4.6-4) which means — as we have seen — more than absolute continuity but less than differentiability everywhere.

5.5.3 Let $f(x)$ be (R) -integrable in $[a, b]$. Then (I) the function $\Phi(x)$ satisfies a Lipschitz condition for $a \leq x \leq b$; (II) at every inner point of continuity of f and so almost everywhere in $[a, b]$ we have the relation

$$\Phi'(x) = f(x). \quad (5.5-5)$$

¹ Remark that — by a theorem of Darboux — the same holds for any function which does not take all the real numbers between two distinct values of its codomain.

P r o o f. Our assumption on f implies the boundedness, i.e., the existence of a constant $K > 0$ such that $|f(t)| \leq K$ ($a \leq t \leq b$). Hence for any pair of points $x_1 \in [a, b]$, $x_2 \in [a, b]$ the inequality holds

$$|\Phi(x_1) - \Phi(x_2)| = \left| \int_{x_1}^{x_2} f(t) dt \right| \leq K|x_1 - x_2|, \quad (5.5-6)$$

by which the first part of the assertion (and simultaneously the continuity of $\Phi(x)$) is verified.

Furthermore, we can write for any point $x \in (a, b)$, if $|l|$ is taken sufficiently small:

$$\left| \frac{\Phi(x+l) - \Phi(x)}{l} - f(x) \right| = \frac{1}{|l|} \left| \int_x^{x+l} [f(t) - f(x)] dt \right| \leq M_l, \quad (5.5-7)$$

where $l \neq 0$ and $M_l = \sup \{|f(t) - f(x)|\}$, the variable t running between x and $x+l$. If x is a point of continuity of f , then $M_l \rightarrow 0$ as $l \rightarrow 0$; therefore (5.5-7) yields

$$\Phi'(x) = \lim_{l \rightarrow 0} \frac{\Phi(x+l) - \Phi(x)}{l} = f(x).$$

Since the points of discontinuity of f constitute a null set by 5.3.2, the validity of (5.5-5) follows for almost all $x \in [a, b]$.

* * *

5.54 Let us call shortly (R) -integration the application of Riemann's indefinite integral (5.5-4) as an operator (cf. 3.12, VI.); we want to elucidate the connection of (R) -integration with differentiation (derivation). More exactly, we are going to study the following problems:

(1) If the (R) -integration of $f(x)$ on a segment can be carried out, whether it is true that the differentiation of the (R) -integral function of f yields the function itself?

(2) If $f'(x)$ exists on a segment, whether it is true that we always can regain function $f(x)$ from it by (R) -integration?

We might say that differentiation is completely invertible by means of (R) -integration or that (R) -integration and derivation are wholly inverse operations to one another, only if both questions (1)-(2) can be answered affirmatively.

Anyway, our answer to the first question is "essentially yes"! In fact, by virtue of 5.5.2, any function f which is (R) -integrable in $[a, b]$ satisfies almost everywhere the relation

$$\frac{d}{dx} \int_a^x f(t) dt = f(x). \quad (5.5-8)$$

Unfortunately, we have a less definite answer regarding the second question. If we suppose not only the existence but also the (R) -integrability of $f'(x)$ in $[a, b]$, then 5.5.2 yields indeed

$$\int_a^x f'(t) dt = f(x) - f(a) \quad (a \leq x \leq b); \quad (5.5-9)$$

but the mere existence and boundedness of f' do not involve even the existence of the (R) -integral in (5.5-9).

In other terms:

5.5.4 *It is possible that a function $f(x)$ is differentiable for $a \leq x \leq b$, and its derivative is bounded but not (R) -integrable in $[a, b]$.*

5.55 The first example of such a function was published by the Italian mathematician V. VOLTERRA in 1881.¹ His fundamental idea is: starting from the function (3.6-14), which is differentiable everywhere and oscillates infinitely many times in any neighbourhood of the origin, we construct by suitable transformations another function $Y(x)$, for which $Y'(x)$ exists and is bounded in $[0, 1]$, but has there infinitely many points of discontinuity *not* constituting a null set.

We still need a simple preliminary observation:

5.5.5 *There exists such a modification $\mathfrak{T}_\vartheta$ ($0 < \vartheta < 1$) of Cantor's ternary set $\mathfrak{T}$ whose outer (L) -measure is at least ϑ .*

P r o o f. Consider the definition of $\mathfrak{T}$ in 2.45, and instead of any open subinterval left out from $[0, 1]$, let us remove another open subinterval which is $(1 - \vartheta)$ -fold shorter. Then the total length of the removed interval system J_ϑ will be

$$(1 - \vartheta) \left(\frac{1}{3} + 2 \frac{1}{9} + 4 \frac{1}{27} + \dots \right) = \frac{1 - \vartheta}{3} \sum_{n=0}^{\infty} \left(\frac{2}{3} \right)^n = 1 - \vartheta;$$

and the remaining part of $[0, 1]$, i.e., the complement $\mathfrak{T}_\vartheta = [0, 1] - J_\vartheta$ has obviously an outer (L) -measure $\geq \vartheta$.²

* * *

Volterra's function can now be represented as follows:

¹ Volterra [1].

² Otherwise the union of $\mathfrak{T}_\vartheta$ and J_ϑ , that is, the interval $[0, 1]$ should be covered by an interval system of total length < 1 which is impossible.

At the points of any interval $I = (x_1, x_2)$ belonging to the above system J_ϑ we define:

$$Y(x) = \begin{cases} v(x - x_1) & \text{for } x_1 < x \leq x_1 + \xi; \\ v(\xi) & \text{for } x_1 + \xi < x < x_2 - \xi; \\ v(x_2 - x) & \text{for } x_2 - \xi \leq x < x_2, \end{cases} \quad (5.5-10)$$

where

$$v(x) = x^2 \sin \frac{1}{x} \quad (x \neq 0) \quad (5.5-11)$$

and ξ means the greatest root of $v'(x) = 2x \sin(1/x) - \cos(1/x)$ in the interval $0 < x < \frac{1}{2}(x_2 - x_1)$.

We see at once that for $0 < |x| \leq 1$ the inequality $|v'(x)| \leq 2|x| + 1 \leq 3$ holds; furthermore, by

$$\overline{\lim}_{x \rightarrow 0} v'(x) = 1, \quad \underline{\lim}_{x \rightarrow 0} v'(x) = -1 \quad (5.5-12)$$

$v'(x)$ is not continuous at $x = 0$. From the definition (5.5-10) we can conclude that $Y(x)$ is differentiable not only on I but also at every point of the interval system J_ϑ ; moreover by $|v'(x)| \leq 3$ ($|x| \leq 1$) we have $|Y'(x)| \leq 3$ ($x \in J_\vartheta$).

As regards the complementary set $\mathfrak{T}_\vartheta$, let us put at its points $Y(x) = 0$; we show that at the same time $Y'(x) = 0$ ($x \in \mathfrak{T}_\vartheta$). In fact, let x_0 be any fixed point of $\mathfrak{T}_\vartheta$ and consider the difference quotient $q(x, x_0) = [Y(x) - Y(x_0)] / (x - x_0)$ ($x \neq x_0$). If also $x \in \mathfrak{T}_\vartheta$, then $q(x, x_0)$ and its limit for $x \rightarrow x_0$ vanish trivially. In the case where x belongs to any interval I , we can write by (5.5-10) and (5.5-11): $|Y(x) - Y(x_0)| = |Y(x)| \leq (x - x_1)^2 \leq (x - x_0)^2$, therefore $|q(x, x_0)| \leq |x - x_0| \rightarrow 0$ as $x \rightarrow x_0$, i.e., $Y'(x_0) = 0$.

Summing up, we obtain that $Y'(x)$ exists and is bounded in the whole interval $[0, 1]$. On the other hand, $Y'(x)$ is clearly not continuous for $x \in \mathfrak{T}_\vartheta$; this set consists namely of the end points of every open interval I (left out from $[0, 1]$) as well as of the accumulation points of the latter, and at such points the limit of $Y'(x)$ cannot exist on the basis of (5.5-10) and (5.5-12). Since $\mathfrak{T}_\vartheta$ is a set of positive (L)-measure by 5.5.5 (i.e., it is not a null set), it follows from Lebesgue's criterion 5.3.2: $Y'(x)$ is not (R)-integrable in $[0, 1]$, as asserted above.

* * *

$Y(x)$ being a primitive of $Y'(x)$, we see from Volterra's example at the same time: *even a not (R)-integrable function can have a primitive*. So we may establish: 1. the existence of a primitive and the (R)-integrability are independent features, neither of them involving the other; 2. even if a primitive exists, possibly it cannot be obtained by (R)-integration of the original function at all.

Finally we remark: $Y'(x)$ furnishes also a further example of a function bounded but not (R) -integrable in a finite interval. (Cf. 5.41.)

5.6 Improper (R) -integral and Riemann integral of several dimensions

5.61 It should be emphasized that the classical ("ordinary") concept of Riemann integral always concerns a *finite interval* and a *bounded integrand*. If we want to extend the notion of (R) -integral to the case of an infinite basic interval or of an unbounded integrand (*improper (R) -integrals*), then we use mostly the well-known Cauchy-Dirichlet method. In the first case, the integral is defined as limit of the (R) -integral taken on a suitable finite subinterval; in the latter case, the so-called points of infinity of the integrand are temporarily excluded from the basic interval together with their proper neighbourhoods, and then we form the limit of the sum of (R) -integrals relating to the remaining subintervals for the case as the end points of these subintervals tend to the corresponding "infinity points" (always on the assumption that the result obtained so is not meaningless).

Some examples:

$$\int_1^{\infty} x^{-\lambda} dx = \lim_{\xi \rightarrow \infty} \int_1^{\xi} x^{-\lambda} dx = \frac{1}{\lambda - 1} \quad (\lambda > 1),$$

$$\int_0^1 x^{\alpha} dx = \lim_{\varepsilon \rightarrow +0} \int_{\varepsilon}^1 x^{\alpha} dx = \frac{1}{\alpha + 1} \quad (-1 < \alpha < 0),$$

$$\int_0^1 \log x dx = \lim_{\varepsilon \rightarrow +0} \int_{\varepsilon}^1 \log x dx = -1;$$

and the first two improper (R) -integrals do not exist (i.e., are divergent) if $\lambda \leq 1$ and $\alpha \leq -1$, respectively.

5.62 The main ideas on the extension to spaces of higher dimension are also well-known:

Let I be a finite ν -dimensional interval, consisting of points $(x_1, x_2, \dots, x_{\nu})$ with $a_k \leq x_k \leq b_k$ ($k = 1, 2, \dots, \nu$), a_k and b_k being given; consider a one-valued, bounded point function $f(P)$, defined on I . Any system of partitions $a_k = x_0^{(k)} < x_1^{(k)} < \dots < x_{m_k}^{(k)} = b_k$ of the linear intervals $[a_k, b_k]$ ($k = 1, 2, \dots, \nu$) generates a decomposition of I into ν -dimensional subintervals. Thus we can form the pendants of the upper and lower sum (5.2-1)–(5.2-2) or of the sum (5.2-6). The so-called ν -dimensional (R) -integral of $f(P)$ on I , denoted by

$$\int_I f(P) dP, \quad (5.6-1)$$

arises from these expedients in the same manner as for the one-dimensional case.

Going further, it is natural to introduce the slightly more general concept of (R) -integral relating to a point set (or region),¹ where the role of an interval I is transferred to an arbitrary (J) -measurable ν -dimensional basic set E . Let $e_1, e_2, \dots, e_n$ be pairwise disjoint, (J) -measurable point sets in the space E_ν for which $\bigcup_{k=1}^n e_k = E$, and let us denote by P_k an arbitrary point of e_k ($k = 1, 2, \dots, n$). Then the above-mentioned integral $\int_{(E)} f(P) dP$

can directly be defined as the limit of a sum $\sum_{k=1}^n f(P_k) \mu_J(e_k)$ for the case when the maximal diameter of the sets e_k tends to 0 [cf. (2.5-1)]. We can proceed, however, in another way: if the basic set E is embedded in a finite ν -dimensional interval I , then the (R) -integral over E can be reduced to an integral over I by putting

$$\int_{(E)} f(P) dP = \int_{(I)} f_0(P) dP, \quad (5.6-2)$$

where

$$f_0(P) = \begin{cases} f(P) & (P \in E), \\ 0 & (P \notin E). \end{cases}$$

The existence criteria 5.3.1-5.3.2 are plainly transferable to higher-dimensional (R) -integrals,² and the connection between the notions of (R) -integral and (J) -measure is retained in a certain sense, too. For instance, the (R) -integral of a non-negative function $f(x_1, x_2)$ relating to a planar set E means geometrically the Jordan volume of the ordinate set of f over E .

5.63 It is well-known from elementary analysis that a ν -dimensional (R) -integral can be written under suitable conditions also as a *multiple* (ν -fold) integral. [This is why we often write instead of the notation (5.6-1)—(5.6-2) several integral signs and instead of dP the symbol $dx_1 dx_2 \dots dx_\nu$.] So the (R) -integral of a function $f(x_1, x_2)$ relating to a rectangle $T = [a_1, b_1] \times [a_2, b_2]$ can be expressed as a double integral in the following form:

$$\int_{(T)} f(x_1, x_2) dx_1 dx_2 = \int_{a_2}^{b_2} \left[\int_{a_1}^{b_1} f(x_1, x_2) dx_1 \right] dx_2, \quad (5.6-3)$$

¹ In the special cases $\nu = 2$ and $\nu = 3$ the terms "area integral" and "volume integral" are also usual.

² In the analogue of 5.3.2, of course, we have the notion of ν -dimensional null set. (Cf. 2.63.)

provided that $f(x_1, x_2)$ is (R) -integrable in T and the inner (R) -integral on the right exists for any fixed $x_2 \in [a_2, b_2]$. (Rule of *successive integration*.)

We mention still that the first *mean-value theorem of integral calculus*, which belongs to the foundation of analysis, can be transferred together with its proof to (R) -integrals of higher dimension. Accordingly we can state: if $f(P)$ is (R) -integrable in a ν -dimensional interval I , further $\varrho(P)$ is another function which is (R) -integrable and of constant sign in I , its (R) -integral on I being different from zero, then the quotient

$$m_\varrho[f] = \frac{\int\limits_{(I)} f(P)\varrho(P) dP}{\int\limits_{(I)} \varrho(P) dP}, \quad (5.6-4)$$

called the *integral mean of $f(P)$ on I with respect to the weight function $\varrho(P)$* , lies between $\inf_{P \in I} \{f(P)\}$ and $\sup_{P \in I} \{f(P)\}$. Especially, if $f(P)$ is *continuous* in the closed interval I , then there exists such a point ξ in the interior of I for which $m_\varrho[f] = f(\xi)$.

In case of $\nu \leq 3$, the integral mean (5.6-4) has important applications in mechanics (centre of gravity, calculation of moments); hence the term "*weight function*".

5.7 The Stieltjesian generalization of Riemann integral

5.71 We know that the extension of Riemann's integral to functions of several variables is possible not only in the form presented in the previous section: besides the so-called ν -dimensional (R) -integrals (i.e., double integral for $\nu = 2$, triple integral for $\nu = 3$), we can speak about *line integrals* of diverse kinds of functions of two or three variables, about *surface integrals* of functions of three variables, about the *vectorial variants* of the latter, and so on.

A characteristic example is the so-called *line integral with respect to arc length*, whose definition is:

$$\int\limits_{(\mathcal{L})} f(P) ds = \lim \sum_{k=1}^n f(P_k) (s_k - s_{k-1}). \quad (5.7-1)$$

Here $\mathcal{L}$ denotes a continuous and rectifiable planar or spatial curve arc, $f(P)$ is a single-valued point function defined along $\mathcal{L}$ (i.e., f may be regarded also as a function of two or three variables, respectively). The sum on the right relates to a decomposition of $\mathcal{L}$ into n subarcs, P_k being an arbitrary point of the k -th subarc, and s_k denoting the length of the curve arc

between the initial point of $\mathcal{L}$ and the end point of the k -th subarc. Finally the passage to the limit is $\max (s_k - s_{k-1}) \rightarrow 0$, i.e., the decomposition (partition) of $\mathcal{L}$ is to be refined unboundedly. It should be pointed out that the "approximating sum" on the right-hand side in (5.7-1) differs from the corresponding Riemann sum (5.2-6) formally not more than in the former the increment $\Delta s_k = s_k - s_{k-1}$ of the arc length s is written instead of the increment $\Delta x_k = x_k - x_{k-1}$ of the variable x . Hence we see that, properly speaking, *two functions* are necessary to the definition (5.7-1): $f(P)$ and the "arc length function" $s = s(P)$, i.e., the length of the curve arc from the initial point of $\mathcal{L}$ to the variable point P of $\mathcal{L}$.

Limits of similar type occur at every turn in numerous chapters of mathematical analysis and of geometry, in probability theory (expected value, variance moments), as well as in several branches of physics. Particular emphasis is to be laid on calculations of moments connected with inhomogeneous mass distributions in mechanics and on the study of charge distributions and of potentials in electrostatics. Thus it is justified to discuss below at length the further development of Riemann's integral concept in the direction just mentioned, attached in the first place to the name of a Dutch mathematician: T.J. STIELTJES (1894).¹

5.72 Taking the simplest case for our basis, let us consider a pair $f(x)$, $g(x)$ of single-valued functions, both being defined in a finite interval $[a, b]$; and on the model of (5.2-6), let us form the so-called *Stieltjesian approximating sum*:

$$\sum[f; g] = \sum_{k=1}^n f(\xi_k) [g(x_k) - g(x_{k-1})], \quad (5.7-2)$$

where $a = x_0 < x_1 < \dots < x_n = b$ is a partition of $[a, b]$, while ξ_k denotes an optional point of the subinterval $[x_{k-1}, x_k]$ ($k = 1, 2, \dots, n$). If $\sum[f; g]$ approaches arbitrarily a real number as $\max (x_k - x_{k-1}) \rightarrow 0$, and this is independent of the manner of refinement of the partition as well as of the choice of the point ξ_k , then the number in question is called the *Stieltjes* (or *Riemann-Stieltjes*) *integral of f with respect to g on the segment $[a, b]$* . Its symbol is

$$\int_a^b f(x) dg(x) \quad \text{or in short} \quad \int_a^b f dg. \quad (5.7-3)$$

It is usual to give both for f and for g separate names: $f(x)$ is the *integrand*, $g(x)$ the *integrator* of (5.7-3).

In the sequel, we shall use instead of the term Riemann-Stieltjes integral mostly the abbreviation *(RS)-integral*; and the existence of (5.7-3) will often

¹ Stieltjes [1]. The concept of Stieltjes integral became known, however, first after a famous result of F. RIESZ on linear operators (1909). (Cf. F. Riesz [7] and 15.1.2.)

be expressed by saying: $f(x)$ is (RS) -integrable with respect to $g(x)$ in $[a, b]$. The Riemann integral $\int_a^b f(x) dx$ is obviously a special case of (5.7-3), namely it corresponds to $g(x) = x$; note also that for all $g(x)$ we have

$$\int_a^b c dg = c[g(b) - g(a)] \quad (c \text{ constant}). \quad (5.7-4)$$

As a non-trivial example, let us calculate the (RS) -integral of an arbitrary continuous function $f(x)$ with respect to a so-called *step function* (i.e., piecewise constant function).

Accordingly, suppose that there is a decomposition of the finite interval $[a, b]$, say $a = t_0 < t_1 < \dots < t_p = b$, such that $g(x) = \gamma_k$ (constant) for $t_{k-1} < x < t_k$ ($k = 1, 2, \dots, p$) and all γ_k are distinct. Furthermore, let us denote the jump (with sign) of $g(x)$ at point t_i by u_i ($k = 0, 1, \dots, p$); more precisely, we put

$$u_i = \begin{cases} g(a+0) - g(a) & \text{for } i = 0, \\ g(t_i+0) - g(t_i-0) & \text{for } i = 1, 2, \dots, p-1, \\ g(b) - g(b-0) & \text{for } i = p. \end{cases} \quad (5.7-5)$$

If $a = x_0 < x_1 < \dots < x_n = b$ is another partition for which $\max(x_k - x_{k-1})$ is less than the length of any subinterval (t_{i-1}, t_i) , then each of the segments $[x_{k-1}, x_k]$ may contain at most one single point of discontinuity of $g(x)$. Thus $g(x_k) - g(x_{k-1})$ takes the value 0, if $g(x)$ has no discontinuity along $[x_{k-1}, x_k]$, and otherwise this difference is equal to one of the numbers $u_0, u_1, \dots, u_p$. Therefore, the result of the limiting process $\max(x_k - x_{k-1}) \rightarrow 0$ is the following:

$$\int_a^b f(x) dg(x) = \sum_{i=1}^p f(t_i) u_i. \quad (5.7-6)$$

5.73 An immediate consequence of the above definition is the *finite additivity* of $\int_a^b f dg$ relating to the basic interval, to the integrand and to the integrator alike [cf. 5.1.2 (II)]. Moreover, it is always possible to factor out a constant c under the integral sign:

$$\int_a^b cf dg = \int_a^b f d(CG) = c \int_a^b f dg. \quad (5.7-7)$$

There is an important pendant of *partial integration* for (RS) -integrals which is however, not so evident:

5.7.1 If $f(x)$ is (RS) -integrable with respect to $g(x)$ in $[a, b]$, then $g(x)$ is also (RS) -integrable with respect to $f(x)$ in the same interval, and we have the formula:

$$\int_a^b f dg + \int_a^b g df = f(b)g(b) - f(a)g(a). \quad (5.7-8)$$

P r o o f. A rearrangement of the terms in $\sum[f; g]$ yields the following identity:

$$\begin{aligned} \sum_{k=1}^n f(\xi_k)[g(x_k) - g(x_{k-1})] &= f(b)g(b) - f(a)g(a) - \\ &- \sum_{k=0}^n g(x_k)[f(\xi_{k+1}) - f(\xi_k)], \end{aligned} \quad (5.7-9)$$

where $\xi_0 = a$, $\xi_{n+1} = b$.

Since the last sum is clearly of the form $\sum[g; f]$ where, however, the role of the points $\{x_k\}$ and $\{\xi_k\}$ is also inverted, further the conditions $\max(x_k - x_{k-1}) \rightarrow 0$ and $\max(\xi_{k+1} - \xi_k) \rightarrow 0$ are equivalent, (5.7-9) becomes (5.7-8) as the partition $a = x_0 < x_1 < \dots < x_n = b$ of $[a, b]$ is refined unboundedly.

* * *

(5.7-8) shows that there is a close connection between the (RS) -integrals $\int_a^b f dg$ and $\int_a^b g df$: any statement on the former has a counterpart about the latter. Thus e.g. the above result that the (RS) -integral of a continuous function with respect to a step function does always exist [cf.(5.7-6)], implies at once the (RS) -integrability of any step function with respect to a continuous function.

5.74 Herewith we have arrived at the natural question: which are the most general conditions on f and g to assure the *existence of an* (RS) -integral $\int_a^b f dg$?

Mindful of the fact that the (R) -integral corresponds to the choice $g(x) = x$, moreover of the requirements of the most important applications, we may assume that the case of *monotone increasing integrator functions* should take a prominent part concerning the subject. Another pertinent essential aspect is: if $g(x)$ increases, then its increments occurring in the Stieltjesian approximating sum are all *non-negative*, which must be considered if we want to transfer all the relevant properties of the classical concept of integral also to the (RS) -integral.

Therefore let $g(x)$ now be monotone increasing in $[a, b]$. Then a consideration quite similar to that given in section 5.3 verifies the counterparts of the criteria 5.3.1—5.3.2 for (RS) -integrals:

5.7.2 (I) *The (RS) -integral of a bounded function $f(x)$ with respect to a monotone increasing function $g(x)$ in a finite interval $[a, b]$ exists if and only if $f(x)$ is continuous at every point of discontinuity of $g(x)$ and the condition*

$$\lim_{\max \Delta x_k \rightarrow 0} \sum_{k=1}^n \Omega_f(x_{k-1}, x_k) [g(x_k) - g(x_{k-1})] = 0 \quad (5.7-10)$$

is fulfilled. Here x_k ($k = 0, 1, \dots, n$) are arbitrary points of subdivisions in a partition of $[a, b]$, $\Delta x_k = x_k - x_{k-1}$; and $\Omega_f(x_{k-1}, x_k)$ means the oscillation of f belonging to the subinterval $[x_{k-1}, x_k]$ i.e.,

$$\Omega_f(x_{k-1}, x_k) = \sup_{x \in [x_{k-1}, x_k]} \{f(x)\} - \inf_{x \in [x_{k-1}, x_k]} \{f(x)\}.$$

(II) *A function $f(x)$ is (RS) -integrable with respect to a monotone increasing function $g(x)$ in a finite segment $[a, b]$ if and only if $f(x)$ is bounded in $[a, b]$, moreover it is there continuous except possibly some points constituting a null set with respect to $g(x)$.*

The last mode of expression means the following: a point set E is said to be *null set with respect to $g(x)$* , if to any $\varepsilon > 0$, however small, countably many such intervals covering together E can be attached, each of them ends at points of continuity of $g(x)$; furthermore, the sum of increments of $g(x)$ corresponding to these intervals is less than ε .¹ It is evident that for $g(x) = x$ we get the definition of the ordinary null set.

The generalization 5.7.2 (II) of Lebesgue's criterion furnishes valuable *sufficient conditions* for existence of Stieltjes integrals:

5.7.3 *An (RS) -integral $\int_a^b f dg$, taken on a finite segment, exists surely if in $[a, b]$ (I) $f(x)$ is continuous and $g(x)$ is of bounded variation or (II) $f(x)$ is (R) -integrable and $g(x)$ is absolutely continuous.*

P r o o f. (I) If $f(x)$ is continuous in $[a, b]$, then the set of its points of discontinuity is empty, i.e., it is trivially a null set with respect to any monotone increasing function. Hence by 4.5.1 and by the finite additivity

¹ Hereby we mean a sum of the form $\sum [g(\beta) - g(\alpha)]$, extended to all intervals (α, β) of the covering system in question. In the case $g(x) = x$, this reduces simply to the total length of the interval system.

of (RS) -integrals, the (RS) -integrability of $f(x)$ with respect to any $g(x)$ of bounded variation follows, too.

(II) Suppose that $f(x)$ is (R) -integrable in $[a, b]$ and let $g(x)$ be a function monotone increasing and absolutely continuous in the same interval. Then $f(x)$ is bounded and its points of continuity form a null set E_f by 5.3.2; it is easy to see that E_f is at the same time a null set with respect to $g(x)$.

Let us regard a finite or countably infinite system of not overlapping intervals $[\alpha, \beta] \subset [a, b]$, which covers E_f . By the definition of absolute continuity, the sum $\sum [g(\beta) - g(\alpha)]$ becomes arbitrarily small as $\sum (\beta - \alpha) \rightarrow 0$ [cf. (4.6-2), (4.6-3)]; but E_f being a null set, $\sum (\beta - \alpha)$ can be as small as we please.

Thus the criterion 5.7.2 (II) yields the existence of $\int_a^b f dg$. Hence follows, however, also the existence of $\int_a^b f dg$, where $g(x)$ is an arbitrary absolutely continuous function; namely by 4.6.2, any such function can be written as the difference of two monotone increasing absolutely continuous functions.

* * *

5.75 As mentioned before, in the case of an increasing integrator function, certain *fundamental properties* of the (R) -integral can be transferred to the (RS) -integral.

For instance, the definition of (RS) -integral and 5.7.3 yield:

5.7.4 For monotone increasing $g(x)$ we have

$$\int_a^b f_1(x) dg(x) \leq \int_a^b f_2(x) dg(x) \quad (5.7-11)$$

if only

$$f_1(x) \leq f_2(x) \quad (a \leq x \leq b) \quad (5.7-12)$$

and both integrals in (5.7-11) exist. Especially, any (RS) -integral with non-negative continuous integrand and increasing integrator is itself non-negative.

Another simple proposition of this kind:

5.7.5 (MEAN VALUE THEOREM FOR STIELTJES INTEGRALS) If $f(x)$ is bounded and $g(x)$ is monotone increasing in a finite interval $[a, b]$, then

$$(\inf_{[a,b]} f) [g(b) - g(a)] \leq \int_a^b f dg \leq (\sup_{[a,b]} f) [g(b) - g(a)] \quad (5.7-13)$$

whenever the (RS) -integral in the middle exists. Especially, in case of a continuous integrand and increasing integrator the equality

$$\int_a^b f dg = f(\xi) \cdot [g(b) - g(a)] \quad (5.7-14)$$

holds, where ξ is a suitable point of (a, b) .¹

P r o o f. (5.7-13) is a direct consequence of 5.7.4, (5.7-4) and (5.7-7) on the basis of the assumptions regarding the functions $f(x)$ and $g(x)$. The continuity of f and the monotony of g assures the existence of $\int_a^b f dg$ by 5.7.3 (I). As regards the formula (5.7-14), in the case $g(a) = g(b)$ it follows trivially from (5.7-13); if $g(a) \neq g(b)$, we have to take into consideration that the quotient

$$m[f; g] = \frac{\int_a^b f dg}{\int_a^b dg}, \quad (5.7-15)$$

the so-called *Stieltjes integral mean* of f with respect to g on $[a, b]$, is a number between $\max_{[a, b]} f$ and $\min_{[a, b]} f$ by (5.7-13), so that f must take the value (5.7-15) somewhere in (a, b) by the Weierstrass theorem.

* * *

We also insert a particular estimation formula for (RS) -integrals.

5.7.6 If $f(x)$ is continuous and $g(x)$ is of bounded variation on a finite segment $[a, b]$, we have the inequality:

$$\left| \int_a^b f(x) dg(x) \right| \leq (\max_{[a, b]} f) \cdot V_g(a, b) \quad (5.7-16)$$

where $V_g(a, b)$ denotes the total variation of $g(x)$ in $[a, b]$.²

¹ As we see, (5.7-14) generalizes the so-called *first mean value theorem* of the integral calculus. We note still that its counterpart, the *second mean value theorem* (which is applied much more rarely) can be proved on the basis of the theory of Stieltjes integral more easily, than in the framework of Riemann's theory.

² Cf. 4.41 and 5.7.3.

In fact, in case of such a pair of functions f, g the Stieltjes sum (5.7-2) can be estimated as follows:

$$\begin{aligned} |\Sigma(f; g)| &\leq (\max_{[a, b]} f) \sum_{k=1}^n |g(x_k) - g(x_{k-1})| \leq \\ &\leq (\max_{[a, b]} f) \cdot V_g(a, b). \end{aligned} \quad (5.7-17)$$

Since this is valid for an arbitrary partition of $[a, b]$, by the limiting process $\max (x_k - x_{k-1}) \rightarrow 0$ (5.7-17) becomes (5.7-16).

5.8 Connection of (RS) -integrals with (R) -integrals; Radon's and Burkill's integral

5.81 Let us deal now with the simplest general method of calculation of the Riemann-Stieltjes integral. It will be shown that if we restrict slightly one of the existence conditions presented in 5.7.3 — namely on the assumption that $g'(x)$ exist and is (R) -integrable — $\int_a^b f dg$ is transformable into a Riemann integral.

5.8.1 If $f(x)$ is (R) -integrable and $g(x)$ has (R) -integrable derivative on a finite segment $[a, b]$, then

$$\int_a^b f(x) dg(x) = \int_a^b f(x) g'(x) dx. \quad (5.8-1)$$

P r o o f. Let f and g be a pair of functions satisfying our conditions and let us consider a partition $a = x_0 < x_1 < \dots < x_n = b$ of $[a, b]$, together with certain points $\xi_k \in [x_{k-1}, x_k]$ ($k = 1, 2, \dots, n$).

A number $\varepsilon > 0$ being given, we can choose $\delta_1 = \delta_1(\varepsilon)$ in such a way that for $\max (x_k - x_{k-1}) < \delta_1$ the inequalities

$$\left| \frac{g(x_k) - g(x_{k-1})}{x_k - x_{k-1}} - g'(\xi_k) \right| < \frac{\varepsilon}{2M(b-a)} \quad (k = 1, 2, \dots, n)$$

are valid, where $M = \sum_{x \in [a, b]} \{|f(x)|\}$. This follows from the derivability of $g(x)$ in $[a, b]$ and from an elementary property of the difference quotient [cf. (3.7-2)].

Therefore we get for the deviation of the sums

$$\sum[f; g] = \sum_{k=1}^n f(\xi_k) [g(x_k) - g(x_{k-1})],$$

$$\sum^*[f; g] = \sum_{k=1}^n f(\xi_k) g'(\xi_k) (x_k - x_{k-1})$$

on the condition $\max(x_k - x_{k-1}) < \delta_1$:

$$|\sum[f; g] - \sum^*[f; g]| \leq \sum_{k=1}^n |f(\xi_k)| \cdot \frac{\varepsilon}{2M(b-a)} (x_k - x_{k-1}) \leq \frac{\varepsilon}{2}. \quad (5.8-2)$$

On the other hand, $\sum^*[f; g]$ is nothing else but the Riemann sum of the function $f(x)g'(x)$ associated with $[a, b]$ [cf. (5.2-6)], and fg' is (R) -integrable in $[a, b]$ by assumption. So there exists a number $\delta_2 = \delta_2(\varepsilon) > 0$ such that

$$\left| \sum^*[f; g] - \int_a^b f(x)g'(x) dx \right| < \frac{\varepsilon}{2} \quad \text{if only } \max(x_k - x_{k-1}) < \delta_2. \quad (5.8-3)$$

The combination of (5.8-2) with (5.8-3) gives:

$$\begin{aligned} \left| \sum[f; g] - \int_a^b f(x)g'(x) dx \right| &\leq |\sum[f; g] - \sum^*[f; g]| + \\ &+ \left| \sum^*[f; g] - \int_a^b f(x)g'(x) dx \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

provided that $\max(x_k - x_{k-1}) < \min(\delta_1, \delta_2)$.

Since $\varepsilon > 0$ can be however small, the result is:

$$\lim_{\max(x_k - x_{k-1}) \rightarrow 0} \sum[f; g] = \int_a^b f(x) dg(x) = \int_a^b f(x)g'(x) dx$$

i.e., the assertion.

* * *

Note that (5.8-1) can be written by the notation $g'(x) = \varrho(x)$ ($a \leq x \leq b$) in the form:

$$\int_a^b f(x)\varrho(x) dx = \int_a^b f(x) dg(x). \quad (5.8-4)$$

Accordingly, the (R) -integral of the product of two functions can always be interpreted as an (RS) -integral, if only one of the factors possesses a primitive; namely, this primitive function will be the integrator of the (RS) -integral in question. Let us recall the fact that the (R) -integrability of a function and the existence of a primitive of it are two independent features (cf. 5.55).

5.82 The extension of the concept of Riemann–Stieltjes integral to the case of an infinite basic interval or of an unbounded integrand can be realized in the same way as we have seen it in 5.61 in connection with the Riemann integral. (*Improper (RS) -integrals.*)

The generalization to functions of several variables starts with the observation that the difference $g(x_k) - g(x_{k-1})$ figuring in the approximation sum (5.7–2) can also be considered as the value belonging to $[x_{k-1}, x_k]$ of an additive interval function $g(\beta) - g(\alpha)$, which is non-negative in the distinguished case when $g(x)$ is monotone increasing.

After these, the Riemann–Stieltjes integral of a one-valued function $f(P)$, given in a ν -dimensional finite interval I and associated with an additive interval function $G(j)$ whose domain consists of some subintervals of I , is defined and denoted by

$$\int_{(I)} f(P)G(dj) = \lim \sum_{k=1}^n f(P_k)G(j_k). \quad (5.8-5)$$

(The alternative, shorter name of this notion is *Radon integral*.)¹ On the right-hand side j_k ($k = 1, 2, \dots, n$) is such a finite system of pairwise disjoint ν -dimensional intervals for which $\bigcup_{k=1}^n j_k = I$; P_k means an optional point of j_k , and the limiting process is that the measure of each j_k tends to 0.

For the integrals of type (5.8–5), we have results analogous to those found in the one-dimensional case, if first the notions: monotone increasing, absolute continuous functions and the concept: function of bounded variation are introduced in a suitable way. (Cf. 4.65.)

The higher-dimensional Stieltjes integrals can also be transformed into (R) -integrals on rather general conditions. Concerning this topic, there is a useful term, the so-called *derivative with respect to region*, a typical example of it being the definition of *local density*:

$$\varrho(P) = \lim_{\Delta\mu \rightarrow 0} \frac{\Delta M}{\Delta\mu}. \quad (5.8-6)$$

Here $\Delta\mu$ denotes the (J) -measure of a region containing the point P , and ΔM means the corresponding quantity of mass. [Cf. (5.6–4), (5.7–15), (5.8–1), (5.8–4).]

¹ Radon [1].

5.83 In any case, the introduction of interval functions into the theory of (RS) -integral involves a further generalization possibility. In particular, if we consider products like $[g(\beta) - g(\alpha)] f(\xi)$ and $G(j)f(P)$ [cf. (5.7-2), (5.8-5)] as such extended kinds of interval functions, whose values depend also on the choice of $\xi \in [\alpha, \beta]$ and $P \in j$, respectively, then we reach the limit expression of the following type:

$$\int_{(I)} \varphi(dj) = \lim \sum_{k=1}^n \varphi(j_k). \quad (5.8-7)$$

Here is a fixed, j a variable interval in E_ν , $\varphi(j)$ denotes an arbitrary (not necessarily additive) interval function defined for every subinterval of I ; furthermore, the right-hand side sum relates to a decomposition of I into finitely many disjoint subintervals j_k ($k = 1, 2, \dots, n$), and the passage to the limit is the same as in (5.8-5).

Though the so-called *Burkill integral* (5.8-7) permits interesting applications in analysis and geometry¹, in what follows we prefer instead that branch of integral theory which has attained the highest importance since the turn of the century. It can be attributed mainly to the work of H. LEBESGUE.

¹ Cf. Burkill [1].

LEBESGUE'S MEASURE AND INTEGRAL. MEASURABLE FUNCTIONS

6.1 Elements of Lebesgue's measure theory

6.11 Riemann's concept of integral is sufficiently general for classical analysis, geometry and physical applications, though it is not difficult to construct bounded functions which are not (R) -integrable. (Cf. e.g. 5.41 and 5.55.) In Chapter 5, however, *certain inadequacies of the (R) -integral* were also elucidated:

- I. The limiting process generally takes out of the class of (R) -integrable functions, and thus the termwise integration of sequences or series of functions can often be carried out on too strong conditions.
- II. The classical integral concept does not allow to invert fully the operation of differentiation and to give a primitive in every case, when such a function exists at all.

Both defects gain importance if we think over that the limiting process is a fundamental operation of the analysis, and the inversion of derivation, as well as the search for primitives are also all-important problems.

The development of mathematics in this century has shown that Riemann's integral theory is *quite modelled on the properties of continuous functions*¹. Consequently — as will be demonstrated later — the (R) -integral as a mathematical expedient is not fine enough to disclose some deeper and overall new connections which have arisen since the early 1900's in analysis and its application fields (first, in the theory of the so-called orthogonal series).

On the basis of all these, we are going to treat thoroughly the more general integral theory of Lebesgue which eliminates for the most part the difficulties above mentioned, and is founded on an appropriately extended theory of measure.²

6.12 Let E be a bounded set of the Euclidean ν -space. The concept of *outer (exterior) Lebesgue measure* $\bar{\mu}_L(E)$ [briefly: *outer (L) -measure*] has

¹ Cf. 5.3.2.

² The first systematic discussion of these topics was given by H. Lebesgue, later professor at the University of Paris, within the framework of his thesis entitled "Intégrale, longueur, aire" (1902). Independently of him, during the years at the turn of the century, also E. Borel [2], G. Vitali [1] and W.H. Young [1]–[2] considerably contributed to the formation of modern measure theory.

already been introduced in connection with covering theorems in 2.62. We know: this is more general than the outer (J)-measure $\bar{\mu}_J(E)$ only so far as in the definition of $\bar{\mu}_L(E)$ the number of intervals covering E may be finite, and countably infinite, too.

Now, like the relationship (5.1-1) between outer and inner (J)-measure, the *inner* (interior) *Lebesgue measure* $\underline{\mu}_L(E)$ (in short: inner (L)-measure) is defined by

$$\underline{\mu}_L(E) = |I| - \bar{\mu}_L(I - E). \quad (6.1-1)$$

where I denotes a ν -dimensional finite closed interval containing E . (The existence of such an I is assured by the boundedness of E .)

Consequently, the inner (L)-measure of E and the outer (L)-measure of its complement with respect to I are interlinked by the formula

$$\underline{\mu}_L(E) + \bar{\mu}_L(CE) = |I| \quad (E \subseteq I).$$

It is clear that $\underline{\mu}_L(E)$ does not depend on the choice of I and it also possesses the property (2.6-3) of $\bar{\mu}_L(E)$.

We verify at once that it holds for any bounded set E :

$$0 \leq \underline{\mu}_L(E) \leq \bar{\mu}_L(E). \quad (6.1-2)$$

Let us cover both of E and $CE = I - E$ by a system of ν -dimensional open intervals whose total measure is μ and μ' , respectively. Every point of I is plainly inside one of such an interval; this remains true if suitable subsystems of *countably many* intervals in the two covering interval systems are considered. Since the sum of total measures of the latter is at least $|I|$, we have the inequality (*) $\mu + \mu' \geq |I|$. But in (*), μ and μ' can approach arbitrarily the values $\bar{\mu}_L(E)$ and $\bar{\mu}_L(CE)$, respectively; therefore we obtain

$$\bar{\mu}_L(E) + \bar{\mu}_L(CE) \geq |I|.$$

Hence follows

$$\bar{\mu}_L(E) \geq |I| - \bar{\mu}_L(CE) \geq 0$$

which is by (6.1-1) equivalent to (6.1-2).

Especially any set E for which $\underline{\mu}_L(E) = \bar{\mu}_L(E)$ is said to be *Lebesgue measurable* (briefly: (L)-measurable). In this case, the common value of the outer and inner (L)-measure is called the *Lebesgue measure* [(L)-measure] of E , and it is denoted by $\mu_L(E)$ or simply by $|E|$.¹

¹ The latter notation refers to the fact that it is a strong extension of the concept of interval measure. According to the above, the notion of (L)-measure is defined for bounded sets only; so here and in the sequel, (L)-measurability always implies the boundedness of the set.

Each measure being non-negative, the condition $\bar{\mu}_L(E) = 0$ implies $\mu_L(E) = 0$ by (6.1-2); thus every null set is (L) -measurable with (L) -measure 0. (This is also why the term "set of measure 0" is used.)

Evidently, any (J) -measurable set E is at the same time (L) -measurable, and we have $\mu_L(E) = \mu_J(E)$; so every interval or an arbitrary system of finitely many intervals is (L) -measurable, and their (L) -measure is identical with the (ordinary) measure. On the other hand, it is easy to give examples of (L) -measurable but not (J) -measurable sets, such as on the number line the set of rational points of a finite stretch, or on the plane the ordinate set belonging to a finite $[a, b]$ of the function (1.6-1). (Both have the (L) -measure 0; cf. 2.6.2.) Therefore, the (L) -measure is a generalization of Jordan's measure; we have to examine, whether it possesses the *basic properties* mentioned in 5.11, which are natural requisites of each concept of measure.

6.13 First of all, we point out that 5.1.2 (I) plainly remains true for $\mu_L(E)$. As regards the additivity of measure concept, the proposition 5.1.3 can also be transferred together with its proof, so that we obtain:

6:1.1 Let E_1, E_2 denote arbitrary ν -dimensional point sets.

(I) If E_1 and E_2 are bounded, there hold the inequalities:

$$\begin{cases} \bar{\mu}_L(E_1 \cup E_2) + \bar{\mu}_L(E_1 \cap E_2) \leq \bar{\mu}_L(E_1) + \bar{\mu}_L(E_2), \\ \underline{\mu}_L(E_1 \cup E_2) + \underline{\mu}_L(E_1 \cap E_2) \geq \underline{\mu}_L(E_1) + \underline{\mu}_L(E_2). \end{cases} \quad (6.1-3)$$

(II) If E_1 and E_2 are (L) -measurable, then $E_1 \cup E_2$ and $E_1 \cap E_2$ have the same property, furthermore

$$|E_1 \cup E_2| + |E_1 \cap E_2| = |E_1| + |E_2|. \quad (6.1-4)$$

In particular, we have

$$|E_1 \cup E_2| = |E_1| + |E_2|, \quad (6.1-5)$$

when the intersection of E_1 and E_2 is a null set.

(6.1-3) implies e.g. in the case of $E_1 \cup E_2 = \bigcirc$ the inequality $\bar{\mu}_L(E_1 \cup E_2) \geq \bar{\mu}_L(E_1)$; moreover its successive application to the pairs of sets $(E_1 \cup E_2)$ and E_3 , after to $(E_1 \cup E_2 \cup E_3)$ and E_4 etc. yields:

$$\mu_L(E_1 \cup E_2 \cup \dots \cup E_n) \geq \mu_L(E_1) + \mu_L(E_2) + \dots + \mu_L(E_n), \quad (6.1-6)$$

where $E_1, E_2, \dots, E_n$ are all bounded and disjoint by pairs. [For outer (L) -measures, an opposite inequality is valid also in case of non-disjoint sets; cf. (2.6-4).]

From the second part of 6.1.1, it follows similarly by induction for optional (L) -measurable sets:

$$|E_1 \cup E_2 \cup \dots \cup E_n| \leq |E_1| + |E_2| + \dots + |E_n|; \quad (6.1-7)$$

and here the sign of equality holds, if the sets have no point in common with one another.

6.14 It is to be emphasized: the great significance of Lebesgue's concept of measure is due just to the fact, that the last statement — contrasted with the case of Jordan's measure — may be extended to countably infinitely many sets, too.¹

6.1.2 (TOTAL ADDITIVITY OF THE LEBESGUE MEASURE)² Let $E_1, E_2, \dots$ be pairwise disjoint, (L) -measurable sets whose union is bounded.

Then the set $\bigcup_{n=1}^{\infty} E_n$ is (L) -measurable as well, namely

$$\left| \bigcup_{n=1}^{\infty} E_n \right| = \sum_{n=1}^{\infty} |E_n|. \quad (6.1-8)$$

P r o o f. Since now $\mu_L(E_n) = \bar{\mu}_L(E_n) = |E_n|$ ($n = 1, 2, \dots$), we have by (6.1-6) for every n :

$$\mu_L \left(\bigcup_{n=1}^{\infty} E_n \right) \geq \mu_L \left(\bigcup_{k=1}^n E_k \right) \geq \sum_{k=1}^n |E_k|;$$

hence we can conclude the convergence of the right-hand series in (6.1-8) and the inequality

$$\sum_{n=1}^{\infty} |E_n| \leq \mu_L \left(\sum_{n=1}^{\infty} E_n \right). \quad (6.1-9)$$

On the other hand, using the properties (2.6-4) of the outer (L) -measure we get

$$\bar{\mu}_L \left(\bigcup_{n=1}^{\infty} E_n \right) \leq \sum_{n=1}^{\infty} \bar{\mu}_L(E_n) = \sum_{n=1}^{\infty} |E_n|; \quad (6.1-10)$$

¹ The outer (or inner) (L) -measure alone does not possess the property of additivity. It is also remarkable that — although the so-called choice or Zermelo's axiom enables us to give a not too complicated example of *non- (L) -measurable set* (cf. Vitali [2]) — Lebesgue's concept of measure is the only one which satisfies certain natural requirements of the general theory of measure. (For example, the condition that the measure of any bounded open linear set should be equal to the total length of the intervals in its normal representation; cf. 2.4.4.)

² Alternative terms for this property: *countably additivity* or *σ -additivity*.

and combining this result with (6.1-9), it follows:

$$\bar{\mu}_L \left(\bigcup_{n=1}^{\infty} E_n \right) \leq \sum_{n=1}^{\infty} |E_n| \leq \mu_L \left(\bigcup_{n=1}^{\infty} E_n \right). \quad (6.1-11)$$

By (6.1-2), this is possible only if the first and third terms are equal, i.e., $E_1 \cup E_2 \cup \dots$ is (L) -measurable; thus (6.1-11) becomes (6.1-8).

* * *

If $E_1, E_2, \dots$ are arbitrary (L) -measurable point sets in the Euclidean ν -space which are contained by a finite interval I , then denoting by H_n the meet of $E_1 \cup E_2 \cup \dots \cup E_{n-1}$ and E_n ($n \geq 2$), the sets $E_1, E_2 - H_2, E_3 - H_3, \dots$ will be pairwise disjoint. So application of 6.1.2 to the latter yields:

6.1.3 (I) The union of countably many (L) -measurable sets E_n is always (L) -measurable if it is bounded; (II) the (L) -measure of this union satisfies the inequality¹

$$\left| \bigcup_n E_n \right| \leq \sum_n |E_n|. \quad (6.1-12)$$

We may add that on the same assumptions on E_n , (2.6-4) also implies the limit relation:

$$\left| \bigcup_{n=1}^{\infty} E_n \right| = \lim_{n \rightarrow \infty} \left| \bigcup_{k=1}^n E_k \right|. \quad (6.1-13)$$

Furthermore, it is a simple corollary of 6.1.3 (I):

6.1.4 The intersection of countably many (L) -measurable sets E_n is always bounded and (L) -measurable.

In fact, we have by (1.1-4) for the complements with respect to the basic interval I :

$$I - \bigcap_n E_n = \bigcap_n (I - E_n);$$

and since every term on the right-hand side is (L) -measurable, the same holds for the difference on the left, i.e., for the bounded intersection $\bigcap_n E_n$, too.

At the same time, we obtain the pendant of (6.1-13):

$$\left| \bigcap_{n=1}^{\infty} E_n \right| = \lim_{n \rightarrow \infty} \left| \bigcap_{k=1}^n E_k \right|, \quad (6.1-14)$$

presuming that the right-hand side limit is finite.

¹ The series on the right may also diverge (i.e., its sum may be equal to $\pm\infty$), then (6.1-12) becomes trivial.

From what has been said we see that all the (L) -measurable sets of the Euclidean ν -space form such a Boolean ring, of which not even the application *countably infinitely many times* of the union and meet operations does take out, provided that the restriction about boundedness of the occurring sets is dropped, and $+\infty$ is also admitted as an (improper) (L) -measure. We know that a set system of the mentioned type is called σ -ring or *Borel's field of sets*. We may speak also about a σ -algebra, because the whole space E_ν can be represented as the union of countably infinitely many ν -dimensional intervals. (Cf. 1.51.)

6.2 Geometric introduction of the Lebesgue integral of a function of one variable; upper and lower Young integrals¹

6.21 As we have shown, the (L) -measure proves to be more flexible and more applicable concerning limiting processes than the Jordan measure. Therefore it is to be expected, that the utilization of Lebesgue's measure theory furnishes (through an analogous train of thought which has been conducted in Section 5.2 from the (J) -measure of planar sets to the ordinary (R) -integral) another concept of integral — not only more general than that of Riemann, but also more valuable in several respects mentioned at the beginning of this chapter.

Let us consider a function $y = f(x)$ defined on a finite interval $[a, b]$, single-valued and *bounded* there; assume first also that $f(x) \geq 0$ ($a \leq x \leq b$). Since the ordinate set $Y(f; [a, b])$ of $f(x)$ (cf. 5.21) is bounded, so the outer and inner (L) -measures of this planar set are finite. The former is called the upper, the latter the lower Young integral of $f(x)$ on $[a, b]$; they

are denoted by the symbols: $(L) \int_a^b f(x) dx$ and $(L) \int_a^b f(x) dx$, respectively. If

these two numbers are equal, i.e., if $Y(f, [a, b])$ is an (L) -measurable planar set, then we say that $f(x)$ is Lebesgue integrable [briefly (L) -integrable] in $[a, b]$. The (L) -measure $|Y(f; [a, b])|$ is called the Lebesgue integral [in short:

(L) -integral] of $f(x)$ on $[a, b]$, and it is denoted by $(L) \int_a^b f(x) dx$.

The question can be raised immediately: is it possible to express these three numbers by means of sums of similar type as those given by the formulae (5.2-1) and (5.2-2)?

The answer is affirmative: let us divide the line segment $[a, b]$ (instead of subintervals) into *finitely or countably infinitely many*, pairwise disjoint,

¹ Cf. Lebesgue [1],[3] and [4], further W.H. Young [3]. L.C. Young [1].

(L)-measurable subsets E_k in any possible way¹, and form in the case of each such decomposition the so-called *upper and lower Young sums*, respectively:

$$\bar{S} = \sum_k |E_k| M_k, \quad (6.2-1)$$

$$\underline{s} = \sum_k |E_k| m_k, \quad (6.2-2)$$

where M_k and m_k mean the supremum and infimum of the values of $f(x)$ taken on E_k . (In the case of infinitely many subsets E_k , the convergence of the right-hand side infinite series is assured, because we have by 6.1.2 $\sum_k |E_k| = b - a$, and the numbers $|M_k|, |m_k|$ — by the boundedness of $f(x)$ —

do not exceed a common bound for every k .)

It is easy to see that the infimum of all possible sums $\bar{S}$ is the upper Young integral, the supremum of all the sums $\underline{s}$ is the lower Young integral, and if both Young integrals are equal, their common value is $(L) \int_a^b f(x) dx$.

The product $|E_k| M_k$ obviously yields the (L)-measure of the ordinate set belonging to E_k of the function $y = M_k$ (constant), while $\bar{S}$ is the (L)-measure of the union U of all such ordinate sets (cf. 2.4.4, 6.1.2). Any set U associated with a decomposition of $[a, b]$ contains $Y(f; [a, b])$; on the other hand, every (finite or countably infinite) planar interval system covering $Y(f; [a, b])$ can be reduced to either of the sets U by omitting certain points. Thus the outer (L)-measure of $Y(f; [a, b])$ must be equal to the infimum of all the numbers $|U| = \bar{S}$, as asserted above. The case of the inner (L)-measure is reducible to the former by (6.1-1).

For instance, in the case of the Dirichlet function (1.6-1) we have on every interval $[a, b]$:

$$(L) \int_a^b \chi(x) dx = (L) \int_a^b \chi(x) dx = (L) \int_a^b \chi(x) dx = 0.$$

In fact, for $\chi(x)$ every sum $\underline{s}$ has the value 0. At the same time, there exists such a decomposition of $[a, b]$ for which $\bar{S} = 0$; namely it is enough to put $E_1 = [a, b] - \{\rho_n\}$, $E_2 = \rho_1, E_3 = \rho_2, \dots$, where $\{\rho_n\}$ is a sequence containing all the rational numbers in $[a, b]$. Furthermore, it holds clearly for an arbitrary decomposition: $\bar{S} \geq 0$.

6.22 Now, the above notions and notations can be transferred directly to bounded functions of *arbitrary sign*, considering that the Young sums

¹ E_k may consist of one single point, too. The reason that we also need here decompositions into (countably) infinitely many parts of $[a, b]$, further, we may not restrict ourselves to subintervals only, can be attributed to the concept of outer (L)-measure.

$\bar{S}, \underline{s}$ keep their meaning for all such $f(x)$. Accordingly, the definitions of upper and lower Young integral are the following:

$${}^{(L)}\int_a^b f(x) dx = \inf\{\bar{S}\}, \quad (6.2-3)$$

$${}^{(L)}\int_a^b f(x) dx = \sup\{\underline{s}\}, \quad (6.2-4)$$

where the infimum and supremum relate to the totality of decompositions of $[a, b]$ of the above mentioned type. Since — like the case of Darboux sums — for any pair of decompositions in question the inequality $\underline{s} \leq \bar{S}$ holds, we get:

$${}^{(L)}\int_a^b f(x) dx \leq {}^{(L)}\int_a^b f(x) dx. \quad (6.2-5)$$

If here the sign of equality is valid, then $f(x)$ is said to be (L) -integrable in $[a, b]$, and the common value of the Young integrals, i.e., the number

$${}^{(L)}\int_a^b f(x) dx = \inf\{\bar{S}\} = \sup\{\underline{s}\} \quad (6.2-6)$$

is called the *Lebesgue integral* [briefly (L) -integral] of $f(x)$ on $[a, b]$.¹

From the above we get simply an *existence criterion* which corresponds to 5.3.1: $f(x)$ is (L) -integrable in $[a, b]$ if and only if to any given $\varepsilon > 0$ we can find a decomposition $[a, b] = \bigcup_k E_k$ of the above type, for which the so-called *oscillation sum*

$$\bar{S} - \underline{s} = \sum_k |E_k| (M_k - m_k) \quad (6.2-7)$$

is less than ε ; the name of the difference $\Omega(E_k) = M_k - m_k$ is the *oscillation* of $f(x)$ on the set E_k .

6.23 Some trivial consequences of the definition of (L) -integral are:

$${}^{(L)}\int_a^b [f_1(x) + f_2(x)] dx = {}^{(L)}\int_a^b f_1(x) dx + {}^{(L)}\int_a^b f_2(x) dx \quad (6.2-8)$$

and

$${}^{(L)}\int_a^b c f(x) dx = c {}^{(L)}\int_a^b f(x) dx \quad (c \text{ constant}) \quad (6.2-9)$$

¹ Instead of the term " $f(x)$ is (L) -integrable", we can also say: "the (L) -integral of $f(x)$ exists".

whenever the (L) -integrals on the right exist. In particular: $(L) \int_a^b 0 dx = 0$,

$$(L) \int_a^b 1 dx = b - a.^1$$

We add still:

6.2.1 If $f(x)$ is (L) integrable in a finite interval $[a, b]$, then

(I) its (L) -integral exists also on every subinterval of $[a, b]$;

(II) $|f(x)|$ and the so-called positive and negative part of $f(x)$, i.e., the functions

$$f_{(+)}(x) = \frac{1}{2} [f(x) + |f(x)|], \quad f_{(-)}(x) = \frac{1}{2} [f(x) - |f(x)|], \quad (6.2-10)$$

are (L) -integrable on the same segment.²

These statements enable us to clarify the *geometric meaning* of the (L) -integral for any bounded function:

$$(L) \int_a^b f(x) dx = |Y_+(f; [a, b])| - |Y_-(f; [a, b])|, \quad (6.2-11)$$

where $Y_+(f; [a, b])$ and $Y_-(f; [a, b])$ denote the parts of the ordinate set $Y(f; [a, b])$ lying above and below the x -axis, respectively.

6.3 Lebesgue's original definition for the (L) -integral of a bounded function; (L) -integrability and measurability

6.31 Although the representation (6.2-6) of (L) -integral is the most natural further development of Riemann's concept of integral, the Young sums $\bar{S}$ and $\underline{s}$ (which are generally infinite series) are difficult to handle. An idea due to Lebesgue offers a more favourable procedure: considering that the existence of the integral depends on the oscillation sum (6.2-7), select the subsets E_k so that we know the corresponding oscillations Ω_k from the beginning, then try to substitute $\bar{S}$ and $\underline{s}$ for certain *finite* approximating sums relating to the subsets just mentioned. This can be realized plainly in such a manner that we *divide first a segment of the y -axis* (instead of

¹ It is advisable to write also for (L) -integrals: $\int_a^b f(x) dx = - \int_b^a f(x) dx$.

² Therefore $f_{(+)}(x)$ [or $f_{(-)}(x)$] is identical with $f(x)$ at points where the function takes positive (or negative) values, and vanishes at any other point.

the x -axis) into finitely many subintervals I_k , and choose as E_k the point set on the axis of abscissae which is associated with I_k , i.e., put $E_k = \{x : f(x) \in I_k\}$.¹ At the same time the process renders it possible to arrange the approximating sums according to increasing function values (ordinates), and a discussion of the "refinement" of the decomposition becomes easier, than in the case of the sums (6.2-1)—(6.2-2), modelled on Riemann's pattern. The example of the function (1.6-1) can illustrate as well, how useful the indicated transcription of $\bar{S}$ and s may be, especially in the case of very *discontinuous functions*: $\chi(x)$ "oscillates" on any (however small) segment of the x -axis between 0 and 1, but its codomain on the y -axis is extremely simple, consisting of two points only.

Let us now study *Lebesgue's method* in detail for some single-valued, bounded function $y = f(x)$ ($a \leq x \leq b$).

Let $[A, B]$ be a finite segment of the y -axis which contains all the values of $f(x)$ taken in $[a, b]$. Any subinterval $[\alpha, \beta] \subseteq [A, B]$ is associated with the totality of those points x of the axis of abscissae where the value of f belongs to $[\alpha, \beta]$, i.e., with the set $\{x : \alpha \leq f(x) \leq \beta\}$.² Next consider a partition $A = y_0 < y_1 < \dots < y_n = B$ of $[A, B]$, and attach to it the sums

$$S_L = \sum_{k=1}^n |e_k| y_k, \quad (6.3-1)$$

$$s_L = \sum_{k=1}^n |e_k| y_{k-1}, \quad (6.3-2)$$

where $e_k = \{x : y_{k-1} \leq f(x) < y_k\}$ ($k = 1, 2, \dots, n-1$) and $e_n = \{x : y_{n-1} \leq f(x) \leq y_n\}$. S_L and s_L are the so-called *upper* and *lower Lebesgue sums* of f belonging to $[A, B]$, respectively.³

In order to form the sums (6.3-1)—(6.3-2), all the (L) -measures $|e_k|$ ($k = 1, 2, \dots, n$) need to exist, i.e., each set e_k is (L) -measurable. (In Fig. 4, e_k is the intersection of the system of ordinate sets drawn obliquely with the x -axis.) Thus the application of Lebesgue's procedure presupposes the following restriction on the function f :

(*) For every subinterval $[\alpha, \beta] \subset [A, B]$, the corresponding point set $\{x : \alpha \leq f(x) \leq \beta\}$ should be (L) -measurable.

According to the definition of Lebesgue sums, the sets e_k are pairwise disjoint and their union is the whole interval $a \leq x \leq b$, further by 6.1.2 we

¹ It is worth mentioning that the same method fails for simple continuous functions, if we use the Jordan measure $\mu_J(E_k)$ instead of $|E_k|$. The concept of (L) -measure may be regarded as the main result of Lebesgue.

² If such a point x does not exist, then $\{x : \alpha \leq f(x) \leq \beta\} = \emptyset$ which is a null set by definition.

³ In case of a *monotonic* continuous function, S_L and s_L are Darboux sums of the function f as well. [Cf. (5.2-1)—(5.2-2).]

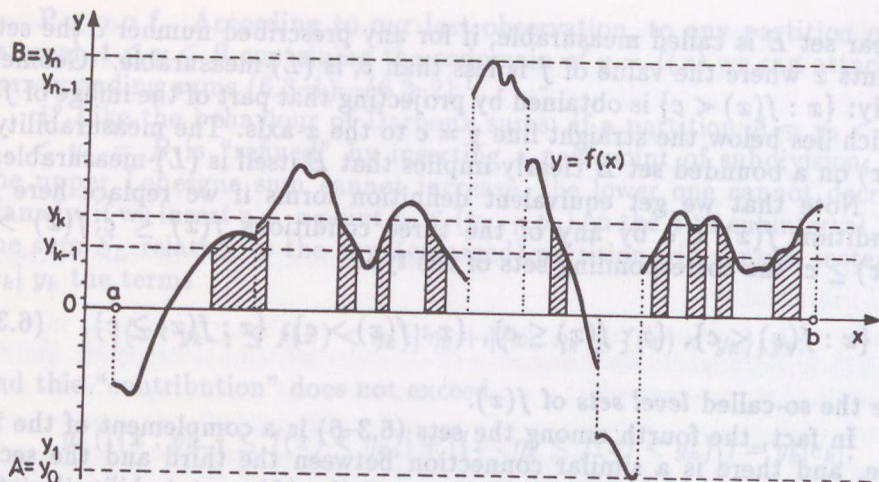


Fig. 4

have $|e_1| + |e_2| + \dots + |e_n| = b - a$; hence the sums (6.3-1)—(6.3-2) are, so to say, “almost” specialized Young sums. [Cf. (6.2-1)—(6.2-2).]

More precisely, on the one hand for the sums

$$\bar{S}_L = \sum_{k=1}^n |e_k| \bar{M}_k \quad \left(\bar{M}_k = \sup_{x \in e_k} \{f(x)\} \right), \quad (6.3-3)$$

$$\underline{s}_L = \sum_{k=1}^n |e_k| \underline{m}_k \quad \left(\underline{m}_k = \inf_{x \in e_k} \{f(x)\} \right); \quad (6.3-4)$$

on the other hand for the above Lebesgue sums we have by $y_{k-1} \leq \underline{m}_k \leq \bar{M}_k \leq y_k$ the inequalities

$$s_L \leq \underline{s}_L \leq \bar{S}_L \leq S_L; \quad (6.3-5)$$

and there holds $s_L = \underline{s}_L$ or $S_L = \bar{S}_L$ in the case if $f(x)$ takes on e_k also the values y_{k-1} and y_k , respectively, or at least it approaches these numbers arbitrarily.

6.32 After this, it is justified to look for a connection between the Lebesgue integral and the sums (6.3-1)—(6.3-2), provided that the latter are not meaningless, i.e., that condition (*) is fulfilled. Going further, we want to show that restriction (*) is equivalent to the existence of the integral

$(L) \int_a^b f(x) dx$, and that the expression of (L)-integral by means of Lebesgue sums is equivalent to the representation (6.2-6) as well.

First it is advisable to introduce the concept of *measurable function*, related to condition (*). A single-valued function $f(x)$, defined on a bounded

linear set E is called measurable, if for any prescribed number c the set of points x where the value of f is less than c , is (L) -measurable.¹ Geometrically: $\{x : f(x) < c\}$ is obtained by projecting that part of the image of $f(x)$ which lies below the straight line $y = c$ to the x -axis. The measurability of $f(x)$ on a bounded set E clearly implies that E itself is (L) -measurable.

Note that we get equivalent definition forms if we replace here the condition $f(x) < c$ by any of the three conditions $f(x) \leq c$, $f(x) > c$, $f(x) \geq c$; the corresponding sets of the type

$$\{x : f(x) < c\}, \{x : f(x) \leq c\}, \{x : f(x) > c\}, \{x : f(x) \geq c\}, \quad (6.3-6)$$

are the so-called *level sets* of $f(x)$.

In fact, the fourth among the sets (6.3-6) is a complement of the first one, and there is a similar connection between the third and the second set. Thus it is sufficient to show that e.g. the (L) -measurability for every c of $\{x : f(x) > c\}$ follows from the same property of $\{x : f(x) \geq c\}$, and conversely; but this is immediately seen by means of the representations (cf. 6.14):

$$\{x : f(x) > c\} = \bigcup_{n=1}^{\infty} \left\{ x : f(x) \geq c + \frac{1}{n} \right\},$$

$$\{x : f(x) \geq c\} = \bigcap_{n=1}^{\infty} \left\{ x : f(x) > c - \frac{1}{n} \right\}.$$

The measurability of a function $f(x)$ defined in $[a, b]$ is equivalent to the above condition (*). Any set $\{x : \alpha \leq f(x) \leq \beta\}$ can be written in the form $\{x : f(x) \leq \beta\} - \{x : f(x) < \alpha\}$, furthermore, we have for every c

$$\{x : f(x) \leq c\} = \bigcup_{n=1}^{\infty} \{x : c - n \leq f(x) \leq c\}.$$

Now, we can point out:

6.3.1 If $f(x)$ is bounded and measurable in $[a, b]$, then the infimum of its upper Lebesgue sums is equal to the supremum of its lower Lebesgue sums; moreover $f(x)$ is (L) -integrable in $[a, b]$ and

$${}^{(L)}\int_a^b f(x) dx = \inf\{S_L\} = \sup\{s_L\}. \quad (6.3-7)$$

¹ The notion "measurability of a function" relates always to the totality of points x in consideration; if no basic set is given, then to the domain of definition of the function.

P r o o f. According to our last observation, to any partition of the interval $A \leq y \leq B$ containing the codomain of $y = f(x)$ we can attach the corresponding sums (6.3-1)—(6.3-2).¹

1° Like the behaviour of Darboux sums: if a partition $A = y_0 < y_1 < \dots < y_n = B$ is “refined” by inserting a new point of subdivision, then the upper Lebesgue sum cannot increase, the lower one cannot decrease. Namely, if we insert e.g. a point $\eta_k \in (y_{k-1}, y_k)$ on the k -th subinterval, then the sum S_L relating to the new (enlarged) partition will contain instead of $|e_k| y_k$ the terms

$$|\{x : y_{k-1} \leq f(x) < \eta_k\}| \eta_k + |\{x : \eta_k \leq f(x) < y_k\}| y_k;$$

and this “contribution” does not exceed

$$y_k (|\{x : y_{k-1} \leq f(x) < \eta_k\}| + |\{x : \eta_k \leq f(x) < y_k\}|) = y_k |e_k|,$$

while the remaining terms of S_L are the same as before.

The assertion on s_L can be verified similarly.

2° We know (cf. (6.3-5)) that $s_L \leq S_L$, if both Lebesgue sums belong to the same partition.

Consider now a lower sum s_L relating to a decomposition $A = y_0 < y_1 < \dots < y_n = B$ and an upper sum S_L^* belonging to another partition $A = Y_0 < Y_1 < \dots < Y_n = B$. If these two decompositions are joined into one single partition $A = u_0 < u_1 < \dots < u_m = B$, the latter will contain at least one new point of subdivision in comparison with the original two partitions; so we obtain by 1° the inequalities $s'_L \geq s_L, S'_L \leq S_L^*$, where s'_L, S'_L denote the lower and upper Lebesgue sums, respectively, associated with the joined partition of $[A, B]$. Since $s'_L \leq S'_L$, it follows that $s_L \leq S_L^*$.

Accordingly, each element of the set of lower (L)-sums is bounded from above by any element of the set of upper (L)-sums, and hence also

$$\sup\{s_L\} \leq \inf\{S_L\}. \quad (6.3-8)$$

3° Finally, estimate the difference of S_L and s_L in the case of a common partition. Using the foregoing we get

$$0 \leq S_L - s_L = \sum_{k=1}^n |e_k| (y_k - y_{k-1}) \leq \max(y_k - y_{k-1}) (b - a), \quad (6.3-9)$$

and the last bound tends to 0 as $\max(y_k - y_{k-1}) \rightarrow 0$. This result implies that the sign of equality holds in (6.3-8), i.e.,

$$\inf\{S_L\} = \sup\{s_L\}. \quad (6.3-10)$$

¹ It is evident that $\int_a^b f(x) dx$ does not depend on the choice of $[A, B]$, because in the Lebesgue sums all the terms vanish which relate to subintervals $[y_{k-1}, y_k]$ containing no value of $f(x)$.

With regard to (6.3-5), we conclude from (6.3-10) that at the same time $\inf\{\bar{S}_L\} = \sup\{\underline{S}_L\}$, and this value must be identical with the value in (6.3-10). In other words [cf. (6.2-6)]: $f(x)$ is (L) -integrable on $[a, b]$, and we have (6.3-7) as asserted above.

* * *

6.33 Theorem 6.3.1 involves an important criterion on the existence of (L) -integral.

6.3.2 *A function bounded in a finite interval is (L) -integrable there if and only if it is measurable.*

P r o o f. Let us denote the interval in question by $[a, b]$, and the single-valued function defined and bounded on it by $y = f(x)$. We have seen before that the measurability of $f(x)$ implies its (L) -integrability in $[a, b]$.

Conversely: let us suppose that $f(x)$ is (L) -integrable on the segment $a \leq x \leq b$. Then the whole ordinate set $Y(f; [a, b])$ is (L) -measurable (cf. 6.2); for some given number c , consider that part of the ordinate set of the function $f(x) - c$ which lies in a strip $-\varepsilon < y \leq 0$ and denote it by $Y_{-\varepsilon}$. The negative part of $f(x) - c$ is (L) -integrable in $[a, b]$ and its ordinate set is (L) -measurable; thus $Y_{-\varepsilon}$ (as the intersection of the latter with the rectangle $\{(x, y) : a \leq x \leq b, -\varepsilon < y \leq 0\}$) is (L) -measurable, too.

If we lengthen each ordinate belonging to $Y_{-\varepsilon}$ up to the straight line $y = -\varepsilon$ and omit from $Y_{-\varepsilon}$ all the ordinates of length $< \varepsilon$, then we get plainly the ordinate sets of the function $y \equiv -\varepsilon$ attached to $\{x : f(x) - c < 0\}$ and $\{x : f(x) - c \leq -\varepsilon\}$, respectively. The outer and inner (planar) (L) -measures of these ordinate sets can be simply obtained as the product of ε and the corresponding linear (L) -measure of the "basic set" on the x -axis. Therefore, by the (L) -measurability of $Y_{-\varepsilon}$ we can write [cf. (2.6-3)]:

$$\varepsilon \cdot \bar{\mu}_L(\{x : f(x) - c \leq -\varepsilon\}) \leq |Y_{-\varepsilon}| \leq \varepsilon \cdot \underline{\mu}_L(\{x : f(x) - c < 0\}),$$

and hence

$$\bar{\mu}_L(\{x : f(x) \leq c - \varepsilon\}) \leq \underline{\mu}_L(\{x : f(x) < c\}) \quad (\varepsilon > 0). \quad (6.3-11)$$

Put here $\varepsilon = 1/n$ ($n = 1, 2, \dots$). Since $\{x : f(x) < c\}$ may be regarded as union of the sets $\{x : f(x) \leq c - (1/n)\}$ ($n = 1, 2, \dots$) which form an increasing sequence, on the basis of (2.6-4) we have:

$$\lim_{n \rightarrow \infty} \bar{\mu}_L\left(\left\{x : f(x) \leq c - \frac{1}{n}\right\}\right) = \bar{\mu}_L(\{x : f(x) < c\}),$$

and (6.3-11) becomes:

$$\bar{\mu}_L(\{x : f(x) < c\}) \leq \underline{\mu}_L(\{x : f(x) < c\}).$$

Because of (6.1-2), this is possible only with sign of equality, i.e., if the set $\{x : f(x) < c\}$ is (L) -measurable; and c being an arbitrary real number, our result means the measurability of the function $f(x)$ in $[a, b]$.

* * *

On the basis of 6.3.2 — as can be seen — the problem of (L) -measurability of planar sets is reducible to the study of (L) -measurability of linear sets. We may remark that the application of Lebesgue's method renders unnecessary the further discussion of the oscillation sum (6.2-7) for the purpose of characterizing (L) -integrable functions.

6.4 Other definition forms; fundamental idea of the treatment due to F. Riesz

6.41 From the verification of theorem 6.3-1, we can also easily get a *limit definition of the Lebesgue integral*.

In fact, consider a partition $A = y_0 < y_1 < \dots < y_n = B$ on the y -axis, and mark out an (optional) point η_k in each subinterval $[y_{k-1}, y_k]$. Then the so-called (ordinary) *Lebesgue sum*

$$\sigma_L = \sum_{k=1}^n |e_k| \eta_k, \quad (6.4-1)$$

where $e_k = \{x : y_{k-1} \leq f(x) < y_k\}$ ($k = 1, 2, \dots, n-1$), $e_n = \{x : y_{n-1} \leq f(x) \leq y_n\}$, can be regarded as a common extension of the sums (6.3-1) — (6.3-2) and by $y_{k-1} \leq \eta_k \leq y_k$ ($k = 1, 2, \dots, n$), the inequalities

$$s_L \leq \sigma_L \leq S_L \quad (6.4-2)$$

hold.

Since $S_L - s_L \rightarrow 0$ as $\Delta y_k = y_k - y_{k-1} \rightarrow 0$ [cf. (6.3-9)], it follows from (6.4-2) and (6.3-7):

$${}^{(L)}\int_a^b f(x) dx = \lim_{\max \Delta y_k \rightarrow 0} \sigma_k. \quad (6.4-3)$$

In a more detailed form: if $f(x)$ is bounded and measurable in $[a, b]$, then to any given $\varepsilon > 0$ we can find a number $\delta = \delta(\varepsilon)$ such that

$$|\sigma_L - {}^{(L)}\int_a^b f(x) dx| < \varepsilon \quad (6.4-4)$$

holds for every partition with $\max \Delta y_k < \delta$, independently of the choice of the "middle" points η_k ($k = 1, 2, \dots, n$). A further usual formulation of the content of (6.4-3) is: "the Lebesgue sum (6.4-1) approaches arbitrarily the integral $\int_a^b f(x) dx$ as the corresponding partition on the y -axis is refined unboundedly".

6.42 Another way to the concept of (L) -integral is furnished by the introduction of (L) -measures as integrals of characteristic functions. [Cf. (5.2-9)].

Let E be an (L) -measurable set contained by a finite interval $[a, b]$, and χ_E its characteristic function with respect to $[a, b]$; then the *integral form of the outer and inner (L) -measure as well as of the (ordinary) Lebesgue measure of E* follow from (6.2-3)—(6.2-4) and (6.2-6):

$$\bar{\mu}_L(E) = {}^{(L)}\int_a^b \chi_E dx, \quad \underline{\mu}_L(E) = {}^{(L)}\int_a^b \chi_E dx,$$

$$|E| = {}^{(L)}\int_a^b \chi_E dx. \quad (6.4-5)$$

Thus the Lebesgue sum (6.4-1) itself can also be presented as an (L) -integral:

$$\sigma_L = {}^{(L)}\int_a^b \varphi_n dx, \quad (6.4-6)$$

where

$$\varphi_n = \sum_{k=1}^n \eta_k \chi_{e_k}. \quad (6.4-7)$$

(As we see, function φ_n takes on the sets $e_1, e_2, \dots, e_n$ — disjoint by pairs — the values $\eta_1, \eta_2, \dots, \eta_n$, so that its whole codomain consists of n numbers.)

Let us regard such a sequence of partition of $[A, B]$ on the y -axis for which $n \rightarrow \infty$ and $\max \Delta y_k \rightarrow 0$. Then by (6.4-3) the (L) -integral of $f(x)$ — whenever it exists — can be written as the limit of the sequence consisting of integrals of the corresponding functions φ_n :

$${}^{(L)}\int_a^b f(x) dx = \lim_{n \rightarrow \infty} {}^{(L)}\int_a^b \varphi_n dx. \quad (6.4-8)$$

In particular, if we put $\eta_k = y_{k-1}$ or $\eta_k = y_k$ (case of the lower and upper Lebesgue sum, respectively), moreover, we take such a sequence of partitions in which each decomposition arises from the preceding one by inserting a

new point of subdivision, the integrals $(L)\int_a^b \varphi_n dx$ will even constitute a monotone increasing or decreasing sequence.

6.43 The representation (6.4–8) forms the basis of the *theory of (L)-integral due to F. RIESZ*.¹

The starting-point of this theory is the discussion of that simple type of function φ_n in which the role of the sets $e_1, e_2, \dots, e_n$ is played, in turn, by subintervals of $[a, b]$, namely by $[a, x_1], [x_1, x_2], \dots, [x_{n-1}, b]$. The integral of such a *step function* is defined (by the geometric interpretation) as follows:

$$(L)\int_a^b \varphi_n dx = \sum_{k=1}^n \gamma_k(x_k - x_{k-1}),$$

where γ_k means the value of φ_n taken in the k -th subinterval.

Next we consider those monotone increasing sequences of step functions whose termwise integration yields convergent number sequences. If $\{\varphi_n\}$ denotes such a sequence, then it is easy to see that $\{\varphi_n\}$ converges almost everywhere to some limit function $f(x)$; so it is natural to define the (L) -integral of $f(x)$ by the formula (6.4–8).

The limit functions of the sequences of step functions just introduced constitute the so-called C_1 class of Riesz; it turns out that forming the difference of any two functions $\in C_1$, we get the totality of functions (L) -integrable in the former sense.

In the treatment according to Riesz, the theory deals first with the concept of (L) -integral and after (as the integral of a characteristic function) with the notion of (L) -measure.

6.5 Comparison of Lebesgue's concept of integral with that of Riemann; properties of measurable functions

6.51 As regards the comparison of (L) -integral with the classical concept of integral and, further, the relation of (L) -integrability to (R) -integrability, we point out: since the Young sums (6.2–1)—(6.2–2) are strong extensions of the Darboux sums (5.2–1)—(5.2–2), thus the upper and lower Young integrals generalize the upper and lower Darboux integrals, respectively, furthermore the Lebesgue integral is a generalization of that of Riemann.

Going into details:

¹ See e.g. F. Riesz [10], [15], [27]; F. Riesz-Sz.-Nagy [1]; Sz.-Nagy [16].

(I) By the extension of any bounded number set its infimum does not increase, and its supremum does not decrease. *So we always have the inequalities*

$$\int_a^b f(x) dx \leq {}^{(L)}\int_a^b f(x) dx \leq {}^{(L)}\int_a^b f(x) dx \leq {}^{(L)}\int_a^b f(x) dx \quad (6.5-1)$$

Consequently:

(II) *If $f(x)$ is (R) -integrable in the interval $[a, b]$, then it is (L) -integrable there and*

$${}^{(L)}\int_a^b f(x) dx = \int_a^b f(x) dx.$$

(III) *The converse is not true, i.e., the totality of (L) -integrable functions is wider than that of (R) -integrable functions.* For example, the bounded function $\chi(x)$ [cf. (1.6-1)] is (L) -integrable on any finite segment $[a, b]$ and its (L) -integral is equal to 0 (cf. 6.21), but it is not (R) -integrable (cf. 5.41).¹

From the foregoing it is clear that the class of functions (L) -integrable in a finite interval $[a, b]$ contains all the functions which are continuous or of bounded variation (or differentiable, monotonic, absolutely continuous) on the same segment.

6.52 Let us now study the *properties of functions bounded, (L) -integrable in some interval $[a, b]$.* By 6.3.2 this class is identical with that of functions bounded and measurable on the same stretch; it is advisable to deal (a little more generally) with the *properties of arbitrary measurable functions.*

From the definitions and the properties of (L) -measure, it follows simply:

The absolute value of a measurable function, further the sum, product and quotient of two measurable functions — the latter with the restriction that the divisor does not vanish — are measurable. Moreover, the measurability of $f(x)$ implies that of $|f(x)|^\alpha$ ($\alpha > 0$) as well.

In addition to this, holds the proposition:

6.5.1 *If $t_1 = f_1(x)$, $t_2 = f_2(x), \dots, t_\nu = f_\nu(x)$ are measurable on a bounded set σ of the x -axis and $g(t_1, t_2, \dots, t_\nu)$ is continuous on the set E of the corresponding points of the Euclidean ν -space, then the composite function $\psi(x) = g(f_1(x), \dots, f_\nu(x))$ is measurable on the set σ .*

¹ More generally: any such function is (L) -integrable but not (R) -integrable in $[a, b]$, which takes a constant c_1 on an everywhere dense, countable subset E of $[a, b]$, and equals to another constant $c_2 (\neq c_1)$ on the complementary set $[a, b] - E$.

P r o o f. Those points $(t_1, t_2, \dots, t_\nu)$ of E , which for a given c satisfy the inequality $g(t_1, t_2, \dots, t_\nu) < c$, plainly form a $(\nu$ -dimensional) open set by the continuity of the function g . So (cf. 2.44) the set of the points in question is the union of countably many ν -dimensional closed intervals, say $I_1, I_2, \dots$. If e.g. the closed linear intervals determining I_k are: $a_1^{(k)} \leq t_1 \leq b_1^{(k)}, \dots, a_\nu^{(k)} \leq t_\nu \leq b_\nu^{(k)}$, then the corresponding linear sets $\{x : a_1^{(k)} \leq f_1(x) \leq b_1^{(k)}\}, \dots, \{x : a_\nu^{(k)} \leq f_\nu(x) \leq b_\nu^{(k)}\}$ in σ are all (L) -measurable, since the functions $f_1, f_2, \dots, f_\nu$ are assumed to be measurable on σ .

Consequently, the set $\{x : \psi(x) < c\} \subset \sigma$ is the union of countable many bounded, (L) -measurable linear sets, and thus it is measurable as well (cf. 6.14). This result means that $\psi(x)$ is measurable on the set σ , as asserted above.

* * *

6.53 As a corollary of 6.5.1 we get: the application finitely many times of the basic operations does not take out of the set of measurable functions. It is an important fact that — in contrast with the (R) -integrable functions — *the limiting process does take out of the totality of measurable functions.* (Cf. 5.42.) In order to verify this, we introduce the concepts of upper and lower envelope function.

Let $\{f_\lambda(x)\}$ be a system of functions defined on some bounded set σ . Each point $x \in \sigma$ can be associated with a set of real numbers, namely with the values of all the functions f_λ at x ; consider the supremum and infimum of this set (especially in the case of a finite number of function values the greatest and the least of them.) The numbers $\sup \{f_\lambda(x)\}$ and $\inf \{f_\lambda(x)\}$ thus obtained depend on x , and therefore they generate a pair of functions $\bar{f}(x)$, $\underline{f}(x)$. These are called the *upper and lower envelope of $\{f_\lambda(x)\}$, respectively.*¹

We see immediately: *the upper and lower envelope of a system $\{f_\lambda(x)\}$ consisting of countably many measurable functions is also measurable.*

Indeed we have the trivial representations:

$$\{x : \bar{f}(x) > c\} = \bigcup_n \{x : f_n(x) > c\},$$

$$\{x : \underline{f}(x) < c\} = \bigcup_n \{x : f_n(x) < c\},$$

and from the (L) -measurability of any set on the right-hand sides we can infer to (L) -measurability of the two sets on the left.

Now:

¹ The geometric background of these denominations is evident.

6.5.2 The upper and lower limit of a bounded sequence of functions measurable on a bounded set, as well as in case of convergence also its limit function are measurable.

Proof. Let $f_1(x), f_2(x), \dots$ denote functions measurable on a common bounded set σ . Forming the upper envelope $\bar{f}_n(x)$ and the lower envelope $\underline{f}_n(x)$ of the subsequence $f_n(x), f_{n+1}(x), \dots$, clearly $\{\bar{f}_n(x)\}$ is monotone decreasing and bounded from below, $\{\underline{f}_n(x)\}$ is monotone increasing and bounded from above at any point x ; moreover, the limit of the former sequence is plainly the greatest cluster value, the limit of the latter is the least cluster value of $\{f_n(x)\}$. (Cf. 2.22.) But the limit of a monotone increasing sequence is at the same time its supremum and the limit of a

monotone decreasing sequence is identical with its infimum; thus $\overline{\lim}_{n \rightarrow \infty} f_n(x)$ may also be regarded as the lower envelope of the sequence $\{\bar{f}_n(x)\}$, while $\underline{\lim}_{n \rightarrow \infty} f_n(x)$ is the upper envelope of the sequence $\{\underline{f}_n(x)\}$.

Applying twice the premised remark, it follows from the measurability of the functions $f_1(x), f_2(x), \dots$ the same property of $\overline{\lim}_{n \rightarrow \infty} f_n(x)$ and $\underline{\lim}_{n \rightarrow \infty} f_n(x)$; and if the latter limits are identical with each other, i.e., $\{f_n(x)\}$ converges at every point $x \in \sigma$, then we come to the conclusion that $\lim_{n \rightarrow \infty} f_n(x)$ is measurable as well.

6.6 Almost uniform approximation of a measurable function by measurable functions; relationship between measurability and continuity

6.61 Since the continuous functions are measurable, by 6.5.2 all the functions belonging to any BAIRE or YOUNG class are measurable, too. (See 3.53.) This means a quite strong generality, being e.g. the Dirichlet function from the second Baire class [cf. (1.6-1)] *nowhere* continuous.

Hence it is very surprising that D.F. EGOROV and N.N. LUSIN, the founders of a well-known school of modern real analysis in Moscow, had already elucidated in the early 1900's: the concept of measurable function does retain certain connections with that of continuity. The possibility of this is given by a relevant property of any a.e. convergent sequence of measurable functions, discovered by Egorov, according to which such sequences are "*almost uniformly convergent*" in a sense. (Cf. 3.51.)

Formulated more precisely, the following assertion holds:

6.6.1 (EGOROV'S THEOREM)¹ Let $f_1(x), f_2(x), \dots$ be measurable on a bounded set E , and assume that $\{f_n(x)\}$ is convergent at almost every point x of E . Then to any given $\eta > 0$, however small, we can find such an (L) -measurable subset E^* of E , for which $|E - E^*| < \eta$ and on which the convergence of $\{f_n(x)\}$ is uniform.

P r o o f. By hypothesis, the set E is (L) -measurable, and we obtain from it by omitting a suitable null set such an (L) -measurable subset E_0 , on which the sequence $\{f_n(x)\}$ consisting of measurable functions converges *everywhere* to a likewise measurable function $f(x)$. (Cf. 6.5.2.)

Let $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n, \dots$ be a monotone decreasing null sequence of positive numbers, and let us denote by $e_{p,r}$ the set of points $x \in E_0$ for which there exists an index p such that

$$|f_n(x) - f(x)| \geq \varepsilon_r \quad (n \geq p). \quad (6.6-1)$$

Since $e_{p,r}$ is nothing else but the union of the level sets $\{x \in E_0 : |f_n(x) - f(x)| \geq \varepsilon_r\}$ ($n = p, p+1, \dots$), so it is also (L) -measurable. On the other hand, the limit relation $|f_n(x) - f(x)| \rightarrow 0$ ($n \rightarrow \infty$) implies that any point $x \in E_0$ belongs only to finitely many $e_{p,r}$, for every fixed r ; thus by $e_{1,r} \cap e_{2,r} \cap \dots = \emptyset$ and $e_{1,r} \supseteq e_{2,r} \supseteq \dots$, on the basis of (6.1-3) we get

$$\lim_{p \rightarrow \infty} |e_{p,r}| = 0.$$

Accordingly, after fixing r and $\eta > 0$ we can find an index $p = p(r)$ such that $|e_{p(r),r}| < \eta/2^r$; therefore the (L) -measure of the set $e = \bigcup_{r=1}^{\infty} e_{p(r),r}$ satisfies the inequality:

$$|e| < \sum_{r=1}^{\infty} \frac{\eta}{2^r} = \eta. \quad (6.6-2)$$

Consider next the set $E^* = E - e$. Because of the definition of e , E^* does not contain any point x which belongs to one of the sets $e_{p,r}$; this holds in particular for all points x satisfying (6.6-1) with an optional index r and $n \geq p(r)$. In other words, we have at each point $x \in E^*$ $|f_n(x) - f(x)| < \varepsilon_r$, if only $n \geq p(r)$. The index r being however large, on a subset E^* of E with $|E - E^*| < \eta$ the convergence of $\{f_n(x)\}$ is uniform.

* * *

6.62 Given this, there is no difficulty in showing that any measurable function can, so to say, be "corrected to a continuous one". For that purpose we need merely a relatively slight modification of its domain of definition.

¹ Egorov [1]. Concerning the operator theoretic significance of this result we refer to F. Riesz [16] and F. Riesz [20].

6.6.2 (LUSIN'S THEOREM)¹ If $f(x)$ is finite and measurable on a bounded set E , then to an arbitrarily small number $\varepsilon > 0$ we can find an (L) -measurable subset E^* of E such that $|E - E^*| < \varepsilon$, and that after reducing the domain of definition of f to E^* , the function becomes continuous.

P r o o f. If the function $f(x)$ is decomposed into its positive and negative part (cf. (6.2-10)), then we see that it is sufficient to deal with the case $f \geq 0$.

1° Given a partition of the bounded, (L) -measurable set E into a finite number of pairwise disjoint (L) -measurable subsets, and if $f(x)$ is defined as constant on each of the latter (i.e., it is about a step function in wider sense), then the assertion is trivial: we have only to exclude from the domain of f all the points of discontinuity of the function, together with their fitting (small) neighbourhoods.

2° Let f now be an arbitrary non-negative, measurable function with finite values, and let us regard the auxiliary functions defined as follows:

$$f_n(x) = \begin{cases} \frac{k-1}{n} & \text{for } \frac{k-1}{n} \leq f(x) < \frac{k}{n} \quad (k = 1, 2, \dots, n) \\ n & \text{for } f(x) \geq n, \end{cases} \quad (6.6-3)$$

where n runs over all the natural numbers. The sequence (6.6-3) of "step functions" converges to $f(x)$ on the set E , so that Egorov's theorem is applicable to $\{f_n(x)\}$.

Hence to any prescribed $\varepsilon > 0$, we can select an (L) -measurable subset $e \subset E$ such that $|e| < \varepsilon/2$ and $\{f_n(x)\}$ tends uniformly to $f(x)$ on set $E - e$. On the other hand, following the instructions in 1°, we can choose for each f_n separately an (L) -measurable subset $e_n \subset E$ such that $|e_n| < \varepsilon/2^{n+1}$ and f_n is continuous on the set $E - e_n$. Consequently, every $f_n(x)$ is continuous on the set $E^* = E - e - \bigcup_{n=1}^{\infty} e_n$, further the sequence $\{f_n(x)\}$ converges uniformly to $f(x)$ on E^* as well; but then by 3.5.1 $f(x)$ is continuous at every point of E^* , too.

In view of the estimation

$$|E - E^*| = |e \cup e_1 \cup e_2 \cup \dots| < \frac{\varepsilon}{2} + \sum_{n=1}^{\infty} 2^{-(n+1)} = \varepsilon,$$

the verification of the theorem is accomplished.

* * *

6.6.3 It should be emphasized that the conclusion of 6.6.2 is naturally not valid at such an interpretation of the word "continuous" that the values

¹ Lusin [1].

of the function taken on E^* are understood together as a part of its codomain belonging to the whole set E . In other words: when we apply Lusin's theorem, we have to disregard completely the values of $f(x)$ on the set $E - E^*$. For instance, the Dirichlet function (1.6-1) becomes continuous on the set obtained by omitting from $[0,1]$ all the rational points (i.e., at points x of irrational abscissae between 0 and 1) only on the condition that the part of the definition of $\chi(x)$ about rational points is entirely ignored, or if the definition is changed in such a way so that the rational points are also associated with the function value 0.

As is suggested by this example, Lusin's theorem can also be stated in another, still more pregnant form:

Any measurable function can be transformed into a continuous function in such a manner that one modifies suitably its values on a subset of arbitrarily small (L) -measure in the domain of definition.

(b) *Introduced in the preceding chapter.*

Before coming to the thorough study of (L) -integrals, we want to make a formal extension. Let E be an (L) -measurable linear set, $f(x)$ a single-valued and bounded function, defined on E . We mean by the (L) -integral of $f(x)$ over E [like (6.4-3)] the limit:

$$\int_E f(x) dx = \lim_{\max \Delta y_k \rightarrow 0} \sum_{k=1}^n \eta_k e_k, \quad (7.1-1)$$

and in case of its existence we say that $f(x)$ is (L) -integrable over E .

The sum (7.1-1) relates to a partition $A = y_0 < y_1 < \dots < y_n = B$ of a finite segment $[A, B]$ on the y -axis which contains the codomain of f associated with E . We write $\Delta y_k = y_k - y_{k-1}$ and use the notations $\eta_k \in [y_{k-1}, y_k]$ ($k = 1, 2, \dots, n$), further $e_k = \{x : y_{k-1} \leq f(x) < y_k\}$ ($1 \leq k \leq n-1$), $e_n = \{x : y_{n-1} \leq f(x) \leq y_n\}$, so that the subsets e_k are disjoint by pairs and

$$\bigcup_{k=1}^n e_k = E. \quad (7.1-2)$$

(The condition $\max \Delta y_k \rightarrow 0$ naturally implies $n \rightarrow \infty$, but the converse assertion is not true.)

The (L) -integral extended over a set is easily reducible to the (L) -integral relating to an interval. Namely, if $[a, b]$ is a finite segment containing the set E , and $f^*(x)$ is such a function which is identical with $f(x)$ or 0 according to $x \in E$ or $x \in [a, b] - E$, then plainly

$$\int_E f(x) dx = \int_a^b f^*(x) dx. \quad (7.1-3)$$

Thus all the results of the preceding chapter can be transferred to the integral (7.1-1), and by 6.3.2 it can be stated: a bounded measurable function is always (L) -integrable.

PROPERTIES OF THE LEBESGUE INTEGRAL. GENERALIZED LEBESGUE INTEGRALS

7.1 Fundamental properties of the (L) -integral of a bounded function over a measurable set

7.1.1 In the sequel, Lebesgue's concept of integral and its generalizations solely occur, and there is no danger of mistaking it for the Riemann integral. Therefore — in accordance with the special literary convention — *(L) -integrals will always be denoted by mere integral signs*, instead of the symbol ${}^{(L)}\int$ introduced in the preceding chapter.

Before coming to the thorough study of (L) -integrals, we want to make a formal extension. Let E be an (L) -measurable linear set, $f(x)$ a single-valued and *bounded* function, defined on E . We mean by the *(L) -integral of $f(x)$ over E* [like (6.4-3)] the limit:

$$\int_{(E)} f(x) dx = \lim_{\max \Delta y_k \rightarrow 0} \sum_{k=1}^n \eta_k |e_k|, \quad (7.1-1)$$

and in case of its existence we say that $f(x)$ is *(L) -integrable over E* .

The sum (7.1-1) relates to a partition $A = y_0 < y_1 < \dots < y_n = B$ of a finite segment $[A, B]$ on the y -axis which contains the codomain of f associated with E . We write $\Delta y_k = y_k - y_{k-1}$ and use the notations $\eta_k \in [y_{k-1}, y_k]$ ($k = 1, 2, \dots, n$), further $e_k = \{x : y_{k-1} \leq f(x) < y_k\}$ ($1 \leq k \leq n-1$), $e_n = \{x : y_{n-1} \leq f(x) \leq y_n\}$, so that the subsets e_k are disjoint by pairs and

$$\bigcup_{k=1}^n e_k = E. \quad (7.1-2)$$

(The condition $\max \Delta y_k \rightarrow 0$ naturally implies $n \rightarrow \infty$, but the converse assertion is not true.)

The (L) -integral extended over a set is easily reducible to the (L) -integral relating to an interval. Namely, if $[a, b]$ is a finite segment containing the set E , and $f^*(x)$ is such a function which is identical with $f(x)$ or 0 according to $x \in E$ or $x \in [a, b] - E$, then plainly

$$\int_{(E)} f(x) dx = \int_a^b f^*(x) dx. \quad (7.1-3)$$

Thus all the results of the preceding chapter can be transferred to the integral (7.1-1), and by 6.3.2 it can be stated: *a bounded measurable function is always (L) -integrable*.

7.12 Let us survey the simplest properties of (L) -integrals taken on sets.

Since the alteration of the codomain of $f(x)$ on a null set does not modify the value of the sum $\sum \eta_k |e_k|$ and of the integral (7.1-1), so we can establish: *the value of an (L) -integral is the same for all (L) -integrable functions which differ from one another over a null set only.* In what follows, these functions will be called (L) -equivalent ones, and will be considered as identical with respect to (L) -integration. In particular: if $f(x)$ vanishes almost everywhere on an (L) -measurable set E , then it is (L) -equivalent there to the function identically zero, and therefore (*) $\int_{(E)} f(x) dx = 0$. If $f(x) \geq 0 (x \in E)$, then

the latter assertion is invertible: the validity of (*) implies that $f(x) = 0$ almost everywhere on E .¹

By the above, let us fix the convention that *in case of any (L) -integral whose integrand $f(x)$ is not defined on a null set E_0 , we put $f(x) = 0 (x \in E_0)$.*

We conclude from the definition (7.1-1) immediately

$$\int_{(E)} 1 dx = |E|, \quad (7.1-4)$$

furthermore (cf. 6.23)

$$\int_{(E)} c f(x) dx = c \int_{(E)} f(x) dx \quad (c \text{ constant}) \quad (7.1-5)$$

$$\int_{(E)} [f_1(x) + f_2(x)] dx = \int_{(E)} f_1(x) dx + \int_{(E)} f_2(x) dx, \quad (7.1-6)$$

provided that the integrals on the right-hand side exist. (By iteration this additive property can be extended to any infinite number of functions $f_k(x)$, too.)

Hence we get: if $\varphi_1(x) \leq \varphi_2(x)$ almost everywhere on an (L) -measurable set E , then supposing the (L) -integrability of both functions on E , it holds also

$$\int_{(E)} \varphi_1(x) dx \leq \int_{(E)} \varphi_2(x) dx. \quad (7.1-7)$$

Since $\varphi_2(x) - \varphi_1(x) \geq 0$ a.e. $x \in E$, the (L) -integral of this difference on E is non-negative as well, and the result follows.

Some trivial consequences of (7.1-7) are:

$$\left| \int_{(E)} f(x) dx \right| \leq \int_{(E)} |f(x)| dx, \quad (7.1-8)$$

¹ This can be verified in the easiest way by regarding $\int_{(E)} f(x) dx$ as the (L) -measure of the ordinate set $Y(f; E)$.

and

$$m \int_{(E)} g(x) dx \leq \int_{(E)} f(x)g(x) dx \leq M \int_{(E)} g(x) dx \quad (7.1-9)$$

with $g(x) \geq 0 (x \in E)$, $m = \inf_{(E)} \{f(x)\}$, $M = \sup_{(E)} \{f(x)\}$, whenever the occurring (L) -integrals exist.¹

For $g(x) = 1$, (7.1-9) becomes

$$m|E| \leq \int_{(E)} f(x) dx \leq M|E|; \quad (7.1-10)$$

and the application of the second part of (7.1-10) to $|f|$ as well as its combination with (7.1-8) leads to the useful estimation:

$$\left| \int_{(E)} f(x) dx \right| \leq |E| \sup_{(E)} \{|f(x)|\}. \quad (7.1-11)$$

The inequalities (7.1-9)—(7.1-11) can also be interpreted as expressing a kind of connection between the *integral means* $|E|^{-1} \int_{(E)} f(x) dx$ or

$\left[\int_{(E)} g(x) dx \right]^{-1} \int_{(E)} f(x)g(x) dx$ and the codomain of $f(x)$ on E . [Cf. (5.6-4), (5.7-15).]

7.13 The last inequality also forms a basis for a relevant property of $\int_{(e)} f(x) dx$ as a function of the set e .

7.1.1. (ABSOLUTE OR TOTAL CONTINUITY OF THE LEBESGUE INTEGRAL) Suppose that $\int_{(E)} f(x) dx$ exists. Then

(I) $\int_{(e)} f(x) dx$ exists for each (L) -measurable subset $e \subset E$, too;

(II) to any given $\varepsilon > 0$, however small, we can find a number $\delta = \delta(\varepsilon)$ such that

$$\left| \int_{(e)} f(x) dx \right| < \varepsilon \quad \text{whenever} \quad |e| < \delta. \quad (7.1-12)$$

Proof. We know that the (L) -integrability of the (bounded) function $f(x)$ on E is equivalent to its measurability, i.e., to the assumption that for all c the set $\{x : f(x) < c\} (x \in E)$ is (L) -measurable. Thus the intersection of any set $e \subseteq E$ with the last-mentioned set is also (L) -measurable by

¹ (7.1-9) yields plainly a pendant of the *first mean value theorem* of integral calculus. Note that the *second mean value theorem* (cf. 5.75) holds in a sense for (L) -integrals, too.

6.1.1. (II); and this means the measurability of $f(x)$ on e , in other words, the existence of the integral $\int_{(e)} f(x) dx$.

As regards (7.1-12), let us observe that the supremum of $\{|f|\}$ on a subset e of E cannot exceed that on E . Therefore we can write by (7.1-11):

$$\left| \int_{(e)} f(x) dx \right| \leq |e| \sup_{(E)} \{|f(x)|\},$$

and the right-hand term is arbitrarily small for $|e|$ small enough.

* * *

7.14 Another important property is a strong extension of the elementary additivity of (L) -integral with respect to intervals.

7.1.2 (TOTAL ADDITIVITY OF THE LEBESGUE INTEGRAL) *Let a bounded linear set E be the union of finitely or countably infinitely many, (L) -measurable disjoint sets e_n . Then we have*

$$\int_{(E)} f(x) dx = \sum_n \int_{(e_n)} f(x) dx, \quad (7.1-13)$$

if only the left-hand side integral exists.

P r o o f. The set E is (L) -measurable by 6.1.2 and any subintegral $\int_{(e_n)} f(x) dx$ exists by 7.1.1 (I).

If now the integral of f on E is considered as the (L) -measure "with sign" of the ordinate set $Y(f; E)$ [cf. (6.2-8)], so (7.1-13) is an obvious consequence of total additivity of the concept of planar (L) -measure.

7.2 Integration of sequences and series of functions

7.21 We have seen in 5.43 that Riemann's concept of integral does not even assure the termwise integration of such convergent sequences of functions whose every term has in absolute value a common upper bound. Now, we are going to show that the use of (L) -integrals enables us to eliminate the insufficiency in question; this may be considered as one of the main benefits of the Lebesgue integral. The verification relies essentially on the fact that the limit function of a convergent sequence of measurable functions is also measurable. (Cf. 6.53.)

7.2.1 (LEBESGUE'S LIMIT THEOREM)¹ Let $f_1(x), f_2(x), \dots$ be a sequence of functions which is convergent on an (L) -measurable set E and the elements of which are: 1. measurable on E as well; 2. bounded in their totality, i.e. at each point $x \in E$ and for every index n :

$$|f_n(x)| \leq K \quad (K \text{ const.})^2 \quad (7.2-1)$$

holds.

Then the limit function $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ is also (L) -integrable on E and

$$\lim_{n \rightarrow \infty} \int_{(E)} f_n(x) dx = \int_{(E)} f(x) dx. \quad (7.2-2)$$

P r o o f. The existence of $\int_{(E)} f(x) dx$ follows therefrom that by (7.2-1)

we have at the same time $|f(x)| \leq K$ ($x \in E$), and that by 6.5.2 the measurability of $f_n(x)$ ($n = 1, 2, \dots$) implies that of $f(x)$. For the sake of brevity write $f(x) - f_n(x) = \varphi_n(x)$ ($n = 1, 2, \dots$); then it suffices to show that

$$\lim_{n \rightarrow \infty} \int_{(E)} \varphi_n(x) dx = 0, \quad (7.2-3)$$

if only every $\varphi_n(x)$ is measurable on E , further (*) $\varphi_n(x) \rightarrow 0$ and $|\varphi_n(x)| \leq 2K$ ($x \in E$; $n = 1, 2, \dots$).

Let $\eta > 0$ be fixed arbitrarily and let us regard the set E_0 consisting of such points x of E at which $|\varphi_n(x)| < \eta$ ($n = 1, 2, \dots$); this E_0 is the intersection of the sets $\{x : |\varphi_1(x)| < \eta\}$, $\{x : |\varphi_2(x)| < \eta\}$, $\dots$, and so it is (L) -measurable by assumption (cf. 6.14). For similar reasons, we can establish the (L) -measurability of the set E_p of points x for which $|\varphi_p(x)| \geq \eta$ and simultaneously $|\varphi_n(x)| < \eta$ ($n = p + 1, p + 2, \dots$) — p denoting any fixed natural number. According to the first part of condition (*), each point of E belongs to one of the subsets E_p just mentioned; since these sets are obviously pairwise disjoint, we can write by 6.1.2:

$$|E| = \sum_{n=0}^{\infty} |E_p|. \quad (7.2-4)$$

Introducing the notations $e_n^{(1)} = \{x : |\varphi_n(x)| \geq \eta\}$, $e_n^{(2)} = \{x : |\varphi_n(x)| \geq \eta\}$ and using the result of the previous section, the integral in (7.2-3) can be estimated as follows:

$$\left| \int_{(E)} \varphi_n(x) dx \right| \leq \int_{(e_n^{(1)})} |\varphi_n(x)| dx + \int_{(e_n^{(2)})} |\varphi_n(x)| dx \leq \eta \cdot |e_n^{(1)}| + 2K |e_n^{(2)}|.$$

¹ Lebesgue [4].

² The conditions 1-2. imply by 6.3.2 that all the functions $f_n(x)$ are (L) -integrable on E .

Here we have by $e_n^{(1)} \subseteq E$ the inequality $|e_n^{(1)}| \leq |E|$, moreover by $e_n^{(2)} = E_n \cup E_{n+1} \cup \dots$ it holds $|e_n^{(2)}| \leq \sum_{r=n}^{\infty} |E_r|$; from the last inequality, owing to the convergence of the series (7.2-4), we can conclude: $|e_n^{(2)}| < \eta$ if only $n \geq N_\eta$.

Summing up, we obtain:

$$\left| \int_{(E)} \varphi_n(x) dx \right| < \eta(|E| + 2K) \quad (n \geq N_\eta),$$

which yields (7.2-3), since η is as small as we please.

* * *

7.22 Note that the above argument yields at the same time: if $\{f_n(x)\}$ converges to $f(x)$ almost everywhere on E , then for any $\varepsilon > 0$ the limit relation

$$\lim_{n \rightarrow \infty} |\{x : |f_n(x) - f(x)| \geq \varepsilon\}| = 0 \quad (7.2-5)$$

holds. This is mostly expressed by saying that any sequence of functions which converges a.e. is necessarily *convergent with respect to (L)-measure* to the same limit function as well.

The converse assertion is not valid, i.e., the convergence with respect to (L)-measure does not imply the pointwise convergence almost everywhere. Nevertheless, a theorem due to F. RIESZ assures that any sequence of functions converging with respect to (L)-measure contains an a.e. convergent subsequence. (Cf. 2.2.2.)¹

7.23 Applying 7.2.1 to the partial sums of a series of functions, Lebesgue's theorem can be formulated in the following manner (equivalent to the preceding form):

7.2.2 *If a function series $f_1(x) + f_2(x) + \dots$ is convergent on an (L)-measurable set E , moreover its terms are measurable and its partial sums are bounded in their totality for $x \in E$, then the sum $s(x)$ of $\sum f_n(x)$ is also (L)-integrable and the series may be integrated term-by-term:*

$$\int_{(E)} f_1(x) dx + \int_{(E)} f_2(x) dx + \dots = \int_{(E)} s(x) dx. \quad (7.2-6)$$

This theorem renders possible the termwise integration of a series of functions on a much more general condition than uniform convergence.

¹ F. Riesz [6].

In particular, if E is an interval $[a, b]$, further we assume (R) -integrability on the terms instead of measurability, and every (L) -integral is replaced by an (R) -integral, the assertion loses its validity (cf. 5.42). The situation changes, however, if we require also the (R) -integrability of $s(x)$; then 7.2.2 becomes a classical theorem found by C. ARZELÀ in 1885.¹

Note that (7.2-6) written in the form:

$$\int_{(E)} \left[\sum_{n=1}^{\infty} f_n(x) \right] dx = \sum_{n=1}^{\infty} \int_{(E)} f_n(x) dx \quad (7.2-7)$$

signifies the *total additivity of the (L) -integral as a functional with respect to the integrand*. (Cf. 3.12 and 7.1.2.)

7.3 Indefinite (L) -integral with bounded integrand and the notion of primitive

7.31 What can be said about the calculation of Lebesgue integrals by means of *primitive functions*? The answer is given by the following generalization of 5.5.1:

7.3.1 If $f(x)$ is bounded in a finite interval $[a, b]$ and there exists a primitive function $F(x)$ of it, then f is also (L) -integrable in $[a, b]$ and its (L) -integral is equal to the corresponding increment of $F(x)$. Briefly:

$$\int_a^b f(x) dx = [f(x)]_a^b = F(b) - F(a). \quad (7.3-1)$$

P r o o f. By our assumptions, there is a constant $K > 0$ such that $|f(x)| \leq K$ ($a \leq x \leq b$), furthermore $F'(x) = f(x)$ ($a \leq x \leq b$). It is advisable to extend the definition of $F(x)$ to a segment $(b, b + \delta]$: put $F(x) = F(b) + F'(b)(x - b)$ ($b < x \leq b + \delta$), i.e., let the graph of $F(x)$ beyond b be the lengthening of the left half-tangent belonging to the point $(b, F(b))$ of the curve $y = F(x)$.

Attach to any point x of $[a, b]$ a sequence of difference quotients:

$$\psi_n(x) = \frac{F(x + l_n) - F(x)}{l_n} \quad (n = 1, 2, \dots) \quad (7.3-2)$$

where $0 < l_n < \delta$ and $\lim_{n \rightarrow \infty} l_n = 0$. The sequence $\{\psi_n(x)\}$ fulfils the conditions of Lebesgue's theorem 7.2.1, because $\lim_{n \rightarrow \infty} \psi_n(x) = F'(x) =$

¹ See Arzelà [1]. It is relevant, that the premises of 7.2.1 or 7.2.2. involve the (L) -integrability of $f(x)$ and $s(x)$, respectively.

$f(x)$ ($a \leq x \leq b$), moreover each function $\psi_n(x)$ is continuous (besides differentiable together with F), and we have by the mean value theorem of Lagrange for every $x \in [a, b]$: $|\psi_n(x)| = |F'(x + \vartheta l_n)| = |f(x + \vartheta l_n)| \leq K$ ($0 < \vartheta < 1$).

Thus by 7.2.1, on the one hand it follows the existence of the (L) -integral of f in $[a, b]$; on the other hand we get the representation:

$$\begin{aligned} \int_a^b f(x) dx &= \lim_{n \rightarrow \infty} \int_a^b \psi_n(x) dx = \\ &= \lim_{n \rightarrow \infty} l_n^{-1} \left[\int_{a+l_n}^{b+l_n} F(x) dx - \int_a^b F(x) dx \right]. \end{aligned} \quad (7.3-3)$$

Since by (7.1-6) $\int_{a+l_n}^{b+l_n} F(x) dx - \int_a^b F(x) dx = \int_b^{b+l_n} F(x) dx - \int_a^{a+l_n} F(x) dx$, and the first mean value theorem of integral calculus yields ($0 < \vartheta_1 < 1$):

$$\lim_{n \rightarrow \infty} l_n^{-1} \int_x^{x+l_n} F(t) dt = \lim_{n \rightarrow \infty} F(x + \vartheta_1 l_n) = F(x) \quad (x \in [a, b]),$$

we obtain from (7.3-3) the formula (7.3-1).

* * *

7.3.2 $F(x)$ being a primitive of $F'(x)$, as an immediate corollary of 7.3.1 we can state:

7.3.2 Any derivative which exists and is bounded in a finite interval $[a, b]$, is necessarily (L) -integrable there.

Volterra's example shows that the corresponding assertion does not hold for (R) -integration. (Cf. 5.5.4 and 5.55.)

A further important consequence of theorem 7.3.1 is:

7.3.3 If a bounded function $f(x)$ ($a \leq x \leq b$) has a primitive at all, then the latter is necessarily of the form

$$\int_a^x f(t) dt + C \quad (x \in [a, b]; C \text{ constant}). \quad (7.3-4)$$

P r o o f. If we denote the primitive function in question by $F(x)$, i.e., $F'(x) = f(x)$, so the application of 7.3.1 furnishes the (L) -integrability of f on any subinterval $[a, x]$ of $[a, b]$, and the formula

$$\int_a^x f(t) dt = F(x) - F(a)$$

or

$$F(x) = \int_a^x f(t) dt + F(a) \quad (a \leq x \leq b). \quad (7.3-5)$$

* * *

Hence it can be seen that Lebesgue's concept of integral enables us to solve another important problem: *the representation of all primitives of a bounded function*. At this point, we have to recall the fact that the counterpart of proposition 7.3.3 is true for (R) -integrals only if we assume on $f(x)$ beside the boundedness also the Riemann integrability. On the other hand, a primitive can be attached even to a non- (R) -integrable function, as was pointed out at the end of Section 5.5.

7.33 Let us deal now with the properties of the function (7.3-4) which is called the (L) -integral function or *indefinite (L) -integral* of f . We shall need a notion due to Lebesgue: a point x is said to be *Lebesgue point* of f , if the condition

$$\lim_{l \rightarrow 0} \frac{1}{l} \int_x^{x+l} |f(t) - f(x)| dt = 0 \quad (l \neq 0) \quad (7.3-6)$$

is fulfilled. It is obvious that any *point of continuity* of f is at the same time a Lebesgue point of the function. [Cf. (7.1-11).]

Theorem 5.5.3 has the following pendant for (L) -integrals:

7.3.4 Let $f(x)$ be bounded and measurable in $[a, b]$. Then the (L) -integral function

$$\Phi(x) = \int_a^x f(t) dt \quad (7.3-7)$$

satisfies a Lipschitz condition for $a \leq x \leq b$ and is differentiable almost everywhere. Moreover at any Lebesgue point of f we have:

$$\Phi'(x) = f(x). \quad (7.3-8)$$

P r o o f. (7.1-11) implies at once that an inequality $|f(x)| \leq K$ ($a \leq x \leq b$) assures the condition

$$|\Phi(x_1) - \Phi(x_2)| \leq K|x_1 - x_2| \quad (7.3-9)$$

for any pair of points $x_1, x_2 \in [a, b]$. Hence we can conclude that $\Phi(x)$ is absolutely continuous (and so of bounded variation) on $[a, b]$, consequently $\Phi'(x)$ exists almost everywhere (cf. 4.62).

The remaining part of the assertion follows from the estimation [cf. (7.1-8)]:

$$\left| \frac{\Phi(x+l) - \Phi(x)}{l} - f(x) \right| \leq \frac{1}{|l|} \left| \int_x^{x+l} |f(t) - f(x)| dt \right| \quad (x, x+l \in [a, b])$$

* * *

It will be pointed out later (see 7.64), that (7.3-6) holds for any (L) -integrable function at every point apart from a null set, i.e., that *almost all points are Lebesgue points*. Thus the validity of (7.3-8) is also assured almost everywhere just as in the case of the (R) -integral — though the integrand of an (L) -integral is eventually nowhere continuous.

7.34 It is much easier to see, that the Lipschitz condition signifies a characteristic property of an indefinite (L) -integral with bounded integrand.

In other words:

7.3.5 A function $F(x)$ defined on a finite segment $[a, b]$ can be written as the (L) -integral function of a bounded function f , if and only if it satisfies in $[a, b]$ a Lipschitz condition.

Proof. The necessity of the condition in question is a direct consequence of 7.3.4.

To verify the sufficiency, let us suppose that $F(x)$ fulfils for $x_1, x_2 \in [a, b]$ the inequality $|F(x_1) - F(x_2)| \leq K|x_1 - x_2|$. Then it is absolutely continuous and a.e. differentiable in $[a, b]$ according to section 4.6; moreover $|F'(x)| \leq K$ apart from a null set, further the (L) -integral $\int_a^b F'(x)dx$ exists by 7.3.2.

We obtain on the basis of Lebesgue's limit theorem 7.2.1, like the proof of (7.3-1), the representation:

$$F(x) = \int_a^x F'(t) dt + F(a) \quad (a \leq x \leq b); \quad (7.3-10)$$

and this means that $F(x)$ is the indefinite (L) -integral of the bounded function $f(x) = F'(x)$ on $[a, b]$.

* * *

Comparing our results, we establish:

7.3.6 *If only bounded functions are considered, then formation of the indefinite (L)-integral and derivation are inverse operations.*¹

Indeed, (I) if $f(x)$ is bounded and the indefinite (L)-integral (7.3-7) exists in $[a, b]$, then by the above

$$\frac{d}{dx} \left[\int_a^x f(t) dt + C \right] = f(x) \quad \text{for a.a. } x \in [a, b];$$

(II) if $f'(x)$ exists and is bounded in $[a, b]$, so we have by 7.3.1:

$$\int_a^x f'(t) dt + f(a) = f(x) \quad (a \leq x \leq b).$$

7.4 (L)-integral with unbounded basic set or integrand

7.41 Up to now, we have studied exclusively (L)-integrals of bounded, measurable functions, taken on bounded, (L)-measurable sets. We have to observe, however, that *the concept of Lebesgue integral can be extended in natural manner to other cases, too.*²

Namely, let $y = f(x)$ be a single-valued function which is defined³ almost everywhere on a *bounded or unbounded* linear set E and is *measurable* there; in the case of an unbounded E , the latter property means that $f(x)$ is measurable on each bounded subset of E . Since $f(x)$ may be unbounded on E , its codomain on the y -axis possibly cannot be included in a finite interval. Having regard to this, let us take an infinite partition of the *whole* y -axis with some points of division y_k :

$$\dots < y_{-k} < y_{-k+1} < \dots < y_{-1} < y_0 < y_1 < \dots < y_{k-1} < y_k < \dots, \quad (7.4-1)$$

and assuming that all the subinterval lengths $\Delta y_k = y_k - y_{k-1}$ ($k = 0, \pm 1, \pm 2, \dots$) are less than a common bound δ . Such a decomposition (7.4-1) will be called briefly " δ -fine scale".

Modelled on the sum (7.1-1), let us form the so-called *general Lebesgue sum*:

$$\sum_L = \sum_{k=-\infty}^{\infty} \eta_k |e_k|, \quad (7.4-2)$$

¹ Lebesgue [4].

² Lebesgue [9].

³ According to our earlier specifications, $f(x)$ is said to be defined at a point only if its value is *finite* there.

where $e_k = \{x : y_{k-1} \leq f(x) < y_k\}$, $y_{k-1} \leq \eta_k \leq y_k$ ($k = 0, \pm 1, \pm 2, \dots$) and (*) the series on the right is supposed to be *absolutely convergent*. Then the series (7.4-2) is — by certain known facts in the theory of series — absolutely convergent for any other scale of the mentioned type as well; and as $\delta \rightarrow 0$ (i.e., when the partition will be refined unboundedly), $\sum_L$ approaches arbitrarily some number $\int_L$, independently of the individual choice of scales and “middle” points η_k . The proof of this statement can be carried out by means of the specialized sums

$$\sum_{k=-\infty}^{\infty} y_k |e_k| \quad \text{and} \quad \sum_{k=-\infty}^{\infty} y_{k-1} |e_k|,$$

like the verification of theorem 6.3.1.

Thus it is reasonable to introduce the following extended wording: a function $f(x)$ measurable on E is said to be (L) -integrable or *summable after Lebesgue* on this set, if the above condition (*) is fulfilled. For such a function $f(x)$, the number $\int_L$ just mentioned is called its (L) -integral relating to E and denoted by the symbol $\int_{(E)} f(x) dx$. In short, we can write:

$$\int_{(E)} f(x) dx = \lim_{\delta \rightarrow 0} \sum_{k=-\infty}^{\infty} \eta_k |e_k| \quad (\sup\{\Delta y_k\} < \delta). \quad (7.4-3)$$

For instance, in case of the function

$$f(x) = \begin{cases} \log x & (0 < x \leq 1), \\ 1/\sqrt{x-1} & (1 < x < 2), \\ 1/(x-1)^2 & (2 \leq x < \infty) \end{cases} \quad (7.4-4)$$

(see Fig. 5) a $(1/n)$ -fine scale — with a fixed positive integer n — on the y -axis can be associated with a decomposition infinite on both sides of the interval $0 < x < \infty$. Splitting up the latter into three parts (corresponding to the subintervals $0 < x \leq 1$, $1 < x < 2$, $2 \leq x < \infty$ separately), the limiting process $n \rightarrow \infty$ yields the existence of the (L) -integral of the function (7.4-4) relating to $(0, \infty)$, further the result

$$\int_0^{\infty} f(x) dx = -1 + 2 + 1 = 2.$$

We have to point out: for unbounded E it can occur that the sum (7.4-2) can be formed in case of any bounded, (L) -measurable subset e of E , but this is impossible in case of the whole set E [e.g. $f(x) = \text{constant}$, $E = (-\infty, \infty)$]. Under such circumstances, the (L) -integral relating to E is defined as the appropriate limit of the (L) -integral taken on e .

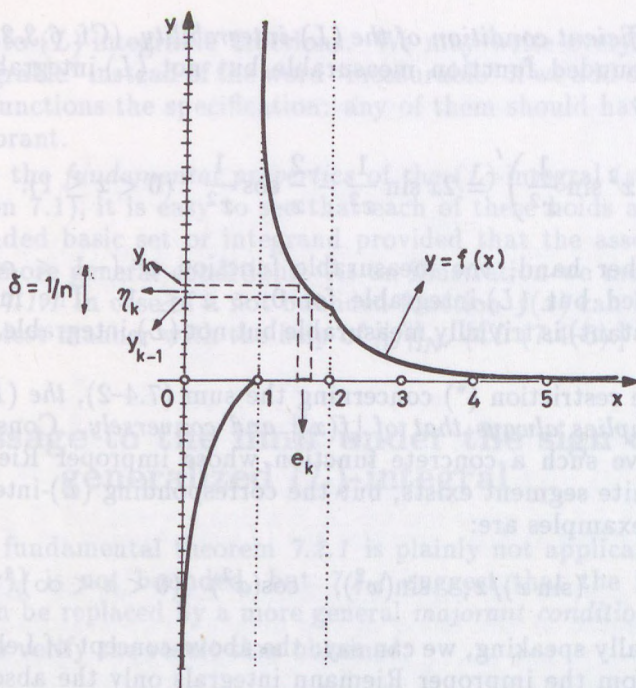


Fig. 5

7.42 The example (7.4-4) clearly shows that the discussion of the (L) -integral of a function $\varphi(x)$ unbounded on some set E is facilitated by the use of the following auxiliary function (bounded from above):

$$[\varphi(x)]_N = \begin{cases} \varphi(x) & \text{for } \varphi(x) \leq N \\ N & \text{for } \varphi(x) > N \end{cases} \quad (N > 0 \text{ integer}) \quad (7.4-5)$$

which is called (by its simple geometric interpretation) "function φ cut off at height N ".

In fact, after decomposing $f(x)$ into its positive and negative parts (cf. 6.23), the definition (7.4-3) can obviously be replaced by

$$\int_{(E)} f(x) dx = \lim_{N_1 \rightarrow \infty} \int_{(E)} [f_{(+)}(x)]_{N_1} dx - \lim_{N_2 \rightarrow \infty} \int_{(E)} [|f_{(-)}(x)|]_{N_2} dx, \quad (7.4-6)$$

provided that both integrals on the right hand side exist. Specifically, if the basic set E is bounded and (L) -measurable, then the (L) -integral of an unbounded function can be reduced by (7.4-6) to that of a bounded integrand.

7.43 Let us stress that in the general case — i.e., if unbounded basic set and integrand are also admitted — the measurability is only necessary

but not sufficient condition of the (L) -integrability. (Cf. 6.3.2.) An example of an unbounded function measurable but not (L) -integrable on a finite segment is:

$$\left(x^2 \sin \frac{1}{x^2}\right)' = 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2} \quad (0 < x \leq 1). \quad (7.4-7)$$

On the other hand, the measurable function x^α ($-1 < \alpha < 0$) is also not bounded but (L) -integrable for $0 < x \leq 1$. The function $y \equiv c$ ($c \neq 0$ constant) is trivially measurable but not (L) -integrable on any infinite interval.

By the restriction (*) concerning the sum (7.4-2), the (L) -integrability of $f(x)$ implies always that of $|f(x)|$ and conversely. Consequently, it is easy to give such a concrete function whose improper Riemann integral on an infinite segment exists, but the corresponding (L) -integral does not; pertinent examples are:

$$(\sin x)/x, \quad \sin(x^2), \quad \cos(x^2) \quad (0 < x < \infty).^1$$

Generally speaking, we can say: the above concept of Lebesgue integral includes from the improper Riemann integrals only the absolutely convergent ones. Anyway, there is no objection to define "improper (L) -integrals" using the classical Cauchy-Dirichlet method (see 5.61) — and thus to get an extension fully spanning both the ordinary and the improper (R) -integrals. From the point of view of the applications, however, only Lebesgue integrals of type (7.4-3) are of importance.

The totality of functions whose (L) -integral relating to a set E exists in the (7.4-3) sense, is usually denoted by $L(E)$; therefore $f(x) \in L(E)$ means simply that $f(x)$ is (L) -integrable ("summable after Lebesgue") over E . Note that the definition implies: any function belonging to the class $L(E)$ takes finite values almost everywhere on E .

The representation (7.4-3) of (L) -integral and the inequality (7.1-8) yield immediately:

7.4.1 If a function $f(x)$ measurable on E has an (L) -integrable majorant function, i.e., there exists a function $g(x) \in L(E)$ for which $|f(x)| \leq g(x)$ ($x \in E$), then $f(x) \in L(E)$.²

By means of this criterion, all the statements of 6.52 and 6.53 about measurable functions (among them theorems 6.5.1 and 6.5.2) can directly

¹ As it is well-known, the improper (R) -integrals on $(0, \infty)$ of these functions have, in turn, the values: $\pi/2$, $(1/2)\sqrt{\pi/2}$, $(1/2)\sqrt{\pi/2}$.

² If $f(x)$ is bounded on an (L) -measurable set E , so it can be majorized there by a (positive) constant. Therefore 7.4.1 involves one of our previous results, namely that a bounded measurable function is at the same time (L) -integrable. (Cf. 7.11.)

be transferred to (L) -integrable functions. We may write everywhere the term " (L) -integrable" instead of the word "measurable" if we add concerning the occurring functions the specification: any of them should have an (L) -integrable majorant.

As regards the *fundamental properties* of the (L) -integral (studied already in Section 7.1), it is easy to see that each of these holds also in the case of unbounded basic set or integrand provided that the assertion has a sense of the more general conditions. As an illustration we mention e.g. that the proof 7.1.1 in case of a not bounded function $f(x)$ can be carried out in the simplest manner with the help of $[|f|]_N$. [Cf. (7.4-5).]

7.5 Passage to the limit under the sign of a generalized (L) -integral

7.51 The fundamental theorem 7.2.1 is plainly not applicable, if the sequence $\{f_n(x)\}$ is not bounded, but 7.4.1 suggest that the inequality $|f_n(x)| \leq K$ can be replaced by a more general *majorant condition*.

We want to verify the result thus obtained:

7.5.1 (LEBESGUE'S MAIN THEOREM ON INTEGRATION OF FUNCTION SEQUENCES)¹ Let $f_1(x), f_2(x), \dots$ be a sequence of measurable functions, converging on a set E almost everywhere, and suppose that its terms are majorizable on E by an (L) -integrable function $g(x)$, i.e.,

$$|f_n(x)| \leq g(x) \quad (n = 1, 2, \dots). \quad (7.5-1)$$

Then the limit function $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ is also (L) -integrable on the set E and $\int_{(E)} f_n(x) dx$ tends to $\int_{(E)} f(x) dx$ as $n \rightarrow \infty$.

P r o o f. Instead of a direct extension of the train of thought connected with 7.2.1 there will be another short line of argument based upon Egorov's theorem 6.6.1.

Assume that $\{f_n(x)\}$ satisfies the above conditions and let f and g vanish on the null set $\epsilon \in E$ where $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ does not exist. Then (7.5-1) holds on the whole set E , and $f(x)$ is there (L) -integrable by 7.4.1.

¹ See Lebesgue [9]. We omit the transcription of the theorem for the case of function series, which can be made mechanically following 7.2.2. Remember that the above type of convergence which is associated with the condition (7.5-1) is called *majorized* or *dominated convergence*.

Introducing the notations $e_k = \{x : k-1 \leq g(x) < k\}$ and $E_m = \bigcup_{k=m+1}^{\infty} e_k = \{x : g(x) \geq m\}$, we can write by 7.1.2:

$$\int_{(E)} g(x) dx = \sum_{k=1}^{\infty} \int_{(e_k)} g(x) dx;$$

and because of the absolute convergence of the series on the right, there exists an integer m such that

$$\int_{(E_m)} g(x) dx = \sum_{k=m}^{\infty} \int_{(e_k)} g(x) dx < \frac{\varepsilon}{5},$$

ε denoting an arbitrarily prescribed positive number.

After fixing this m , consider the (L) -measurable set $E - E_m = \bigcup_{k=1}^m e_k$ on which $g(x) < m$. Applying theorem 6.6.1, we see that $E - E_m$ can be represented as the union of two disjoint, (L) -measurable sets E^* , E^{**} , such that E^{**} has an (L) -measure as small as we please, while on E^* the convergence of $\{f_n(x)\}$ is uniform. Choose E^{**} in such a way that $|E^{**}| < \varepsilon(5m)^{-1}$, further let N be so large, that for all points of E^* we have $|f_n(x) - f(x)| < \varepsilon(5|E^*|)^{-1}$ ($n \geq N$).

Now it is easy to observe that for $n \geq N$ in the decomposition

$$\int_{(E)} (f_n - f) dx = \int_{(E_m)} f_n dx - \int_{(E_m)} f dx + \int_{(E^{**})} f_n dx - \int_{(E^{**})} f dx + \int_{(E^*)} (f_n - f) dx$$

the absolute value of each integral on the right-hand side is less than $\varepsilon/5$. Thus we obtain finally:

$$\left| \int_{(E)} [f_n(x) - f(x)] dx \right| < \varepsilon \quad (n \geq N = N(\varepsilon)),$$

and the result follows.

* * *

7.52 For the sake of later applications, we insert two further important theorems on termwise integration of function sequences. In these, the condition (7.5-1) is replaced by the boundedness of the integrals of terms, which means a much weaker restriction; on the other hand, we require also the non-negativity or monotony of all term functions, respectively.

7.5.2 (FATOU'S LEMMA)¹ If the terms of a sequence of functions $\{f_n(x)\}$ are (L) -integrable and non-negative on a set E , moreover the sequence is convergent a.e. on E and satisfies for all n the condition

$$\int_{(E)} f_n(x) dx \leq K \quad (K \text{ constant}), \quad (7.5-2)$$

then $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ is (L) -integrable on E , and we have also

$$\int_{(E)} f(x) dx \leq K. \quad (7.5-3)$$

P r o o f. It holds for almost all $x \in E$ by hypothesis $f_n(x) \rightarrow f(x)$ ($n \rightarrow \infty$) and therefore also $\lim_{n \rightarrow \infty} [f_n(x)]_N = [f(x)]_N$ [Cf. (7.4-5)]. Since $[f_n(x)]_N \leq N$ for every n and $x \in E$, we can write by 7.5.1:

$$\lim_{n \rightarrow \infty} \int_{(E)} [f_n(x)]_N dx = \int_{(E)} [f(x)]_N dx. \quad (7.5-4)$$

But the conditions $f_n(x) \geq 0$ ($n = 1, 2, \dots$) and (7.5-2) imply

$$\int_{(E)} [f_n(x)]_N dx \leq \int_{(E)} f_n(x) dx \leq K,$$

so that from (7.5-4) we get

$$\int_{(E)} [f(x)]_N dx \leq K.$$

This involves $f(x) \in L(E)$ and (7.5-3), as $N \rightarrow \infty$.

* * *

7.5.3 (THEOREM OF B. LEVI)² Let $\{f_n(x)\}$ be a monotone increasing sequence, consisting of functions (L) -integrable on a set E , which satisfies the condition (7.5-2).

Then $\{f_n(x)\}$ is convergent almost everywhere and its limit function $f(x)$ is (L) -integrable on E , furthermore

$$\lim_{n \rightarrow \infty} \int_{(E)} f_n(x) dx = \int_{(E)} f(x) dx. \quad (7.5-5)$$

¹ Fatou [1].

² Levi [1].

Proof. Consider the sequence of functions $\varphi_n(x) = f_n(x) - f_1(x)$ ($n = 1, 2, \dots$). According to our assumptions, each $\varphi_n(x)$ is non-negative and (L) -integrable on E ; moreover $\{\varphi_n(x)\}$ is monotone increasing on the same set and by (7.5-2)

$$\int_{(E)} \varphi_n(x) dx \leq K + \int_{(E)} |f_1(x)| dx = K^* \quad (n = 1, 2, \dots). \quad (7.5-6)$$

Since the integrals on the left-hand side in (7.5-6) constitute an increasing sequence which is bounded from above, so the (finite) limit

$$A = \lim_{n \rightarrow \infty} \int_{(E)} \varphi_n(x) dx \quad (7.5-7)$$

exists. We have to show that also

$$\lim_{n \rightarrow \infty} \varphi_n(x) = \varphi(x) \quad (7.5-8)$$

exists (is finite) for almost all $x \in E$, further $\varphi(x) \in L(E)$ and

$$\int_{(E)} \varphi(x) dx = A. \quad (7.5-9)$$

1° Let us denote by E_∞ the set of points $x \in E$ with $\lim_{n \rightarrow \infty} \varphi_n(x) = \infty$. This may be rewritten in the form:

$$E_\infty = \bigcap_{r=1}^{\infty} \bigcup_{n=1}^{\infty} E_{nr},$$

where $E_{nr} = \{x : x \in E, \varphi_n(x) > r\}$.

Because of $\varphi_n(x) \geq 0$ we have $\int_{(E)} \varphi_n(x) dx \geq \int_{(E_{nr})} \varphi_n(x) dx \geq r \cdot |E_{nr}|$,

and thus (7.5-6) yields for every n and r :

$$|E_{nr}| \leq \frac{1}{r} \int_{(E)} \varphi_n(x) dx \leq \frac{K^*}{r}. \quad (7.5-10)$$

By $E_{1r} \subseteq E_{2r} \subseteq \dots$ it is $\bigcup_{p=1}^n E_{pr} = E_{nr}$; hence we see that (7.5-10)

remains valid if E_{nr} is replaced by $\bigcup_{p=1}^{\infty} E_{pr}$. On the other hand, there holds

$E_\infty \subseteq \bigcup_{p=1}^{\infty} E_{pr}$ for all r . Combining the latter result with the former, we get:

$$|E_\infty| \leq \left| \bigcup_{p=1}^{\infty} E_{pr} \right| \leq \frac{K^*}{r} \quad (r = 1, 2, \dots).$$

But the last term tends to 0 as $r \rightarrow \infty$ and therefore $|E_\infty| = 0$, i.e., E_∞ is a null set. Accordingly, $\{\varphi_n(x)\}$ has indeed a finite limit a.e. on E .

2° Observing still that also $\int_E \varphi_n(x) dx \leq A$ for all n , since the integral

on the left-hand side tends increasingly to A as $n \rightarrow \infty$ [cf. (7.5-7)], we can apply lemma 7.5.2 to the sequence

$\{\varphi_n(x)\}$.

So we obtain the (L) -integrability of $\varphi(x)$ on E and the inequality

$$\int_E \varphi(x) dx \leq A. \quad (7.5-11)$$

At the same time, we have $\varphi(x) \geq \varphi_n(x)$ ($n = 1, 2, \dots$) a.e. on E , whence it follows after integration [cf. (7.1-7)] and for $n \rightarrow \infty$ the opposite estimation:

$$\int_E \varphi(x) dx \geq A. \quad (7.5-12)$$

Therefore it yields (7.5-9), i.e., the formula

$$\int_E \varphi(x) dx = \lim_{n \rightarrow \infty} \int_E \varphi_n(x) dx,$$

and this is equivalent to (7.5-5).

* * *

7.53 If one knows about a function $f(x)$ only that it is (L) -integrable in some finite interval $[a, b]$, then the validity of the limit relation

$$\lim_{l \rightarrow 0} |f(x+l) - f(x)| = 0$$

is not assured even at a single point $x \in [a, b]$, because it may happen that $f(x)$ is nowhere continuous. [See e.g. the Dirichlet function (1.6-1).]

So much the more surprising is that if we first integrate the deviation $|f(x+l) - f(x)|$ on the segment $a \leq x \leq b$ then carry out the limiting process $l \rightarrow 0$, we get *always* 0. This fact which was found by LEBESGUE in 1905 can be regarded with good reason as one of the most relevant properties of the (L) -integral, namely for the class $L(a, b)$ it is a counterpart of the uniform continuity of functions continuous in $[a, b]$. [Hence the term " $L(a, b)$ -continuity" is sometimes used.]¹

In addition, we want to stress that the passage to the limit cannot now be effected under the integral sign; i.e., the pertinent theorems just discussed are not applicable, and we need a particular method to make the verification.

¹ Cf. Lebesgue [5], [9].

7.5.4 If $f(x)$ is (L) -integrable on a finite segment $[a, b]$ and vanishes outside this interval, then

$$\lim_{l \rightarrow 0} \int_a^b |f(x+l) - f(x)| dx = 0. \quad (7.5-13)$$

P r o o f. For the sake of brevity write:

$$J_l[f] = \int_a^b |f(x+l) - f(x)| dx. \quad (7.5-14)$$

Then we have for all l the inequality $J_l[f] \leq 2 \int_a^b |f(x)| dx$, and hence for any pair f_1, f_2 of admitted functions:

$$J_l[f_1] \leq J_l[f_2] + J_l[f_1 - f_2] \leq J_l[f_2] + 2 \int_a^b |f_1(x) - f_2(x)| dx. \quad (7.5-15)$$

More generally, it is evident that

$$J_l[c_1 f_1 + c_2 f_2] \leq |c_1| J_l[f_1] + |c_2| J_l[f_2], \quad (7.5-16)$$

where c_1 and c_2 are arbitrary constants.

Now, the validity of (7.5-13) follows at once from theorem 3.3.4, if $f(x)$ is continuous in $[a, b]$. Thus by (7.5-15), the assertion can also be transferred to piecewise continuous functions, namely if the latter are approached by functions continuous in the whole interval. The case of other admitted functions will be discussed in three parts.

1° As far as the step functions (in a wider sense) are concerned, consider first of all the characteristic function χ_E of an (L) -measurable set $E \subset [a, b]$. According to the definition of (L) -measure, E can be covered by a system of countably many finite intervals whose total length M satisfies the inequality $M < |E| + \varepsilon$, denoting by ε an arbitrarily small positive number; moreover we can select from this interval system a *finite* subsystem I for which $|I| > |E|$. Thus we can write:

$$\int_a^b |\chi_E - \chi_I| dx < (M - |E|) + (M - |I|) < 2\varepsilon,$$

and by (7.5-15)

$$J_l[\chi_E] \leq J_l[\chi_I] + 2 \int_a^b |\chi_E - \chi_I| dx \leq J_l[\chi_I] + 4\varepsilon. \quad (7.5-17)$$

But χ_l being piecewise continuous, $J_l[\chi_l]$ becomes less than ε for $|l|$ small enough. Consequently, we obtain from (7.5-17):

$$J_l[\chi_E] < \varepsilon + 4\varepsilon = 5\varepsilon, \quad (7.5-18)$$

provided that $|l| < \lambda = \lambda(\varepsilon)$. By (7.5-16) follows the validity of (7.5-13) for every so-called generalized step function of the form $K_1\chi_{e_1} + K_2\chi_{e_2} + \dots + K_m\chi_{e_m}$, where $e_1, e_2, \dots, e_m$ are arbitrary pairwise disjoint, (L) -measurable subsets of $[a, b]$ and $K_1, K_2, \dots, K_m$ denote constants.

2° Let $\varphi(x)$ be a function bounded, (L) -integrable in $[a, b]$, which vanishes outside $[a, b]$. Then the definition of Lebesgue integral assures the existence of a sequence $\{\varphi_n(x)\}$ consisting of generalized step functions for which $\int_a^b \varphi_n(x) dx$ tends increasingly to $\int_a^b \varphi(x) dx$. (Cf. 6.42.)

So to any given $\eta \rightarrow 0$, n can be chosen in such a way that

$$\int_a^b [\varphi(x) - \varphi_n(x)] dx < \frac{\eta}{3}; \quad (7.5-19)$$

on the basis of 1°, we shall have the inequality $J_l[\varphi_n] < \eta/3$ for $|l|$ sufficiently small, n being fixed according to (7.5-19).

This implies again by using (7.5-15):

$$J_l[\varphi] \leq J_l[\varphi_n] + 2 \int_a^b [\varphi(x) - \varphi_n(x)] dx < \eta \quad [|l| < l_0(\eta)]. \quad (7.5-20)$$

3° If $f(x)$ is not bounded, but (L) -integrable in $[a, b]$ and takes the value 0 outside $[a, b]$, then its (L) -integral on $[a, b]$ can be approximated arbitrarily by suitable (L) -integrals of certain bounded functions. [Cf 7.4-5.]

Let $f_N(x)$ be a function which is identical with $f(x)$ for $f(x) \leq N$ and 0 otherwise. Then on the one hand, the (L) -integral of $|f(x) - f_N(x)|$ in $[a, b]$ is less than η , provided that N is large enough; on the other hand, we conclude by (7.5-20) that $J_l[f_N] < \eta$, if only $|l|$ is sufficiently small.

Using still (7.5-15), we obtain the inequality

$$J_l[f] \leq J_l[f_N] + 2 \int_a^b [f(x) - f_N(x)] dx < 3\eta \quad [|l| < l^*(\eta)]. \quad (7.5-21)$$

Since $\eta > 0$ is as small as we please, herewith our argumentation is finished.

7.6 Theorems on arbitrary (L) -integral functions

7.61 Let us study, how and in which degree can be extended the results of Section 7.3 to the widest class of (L) -integrable functions.

Since the proofs 7.3.1 and 7.3.2 are founded upon the limit theorem 7.2.1, it is clear that the theorems in question can be transferred to the case of unbounded integrands, too, if we apply the so-called "great" Lebesgue theorem 7.5.1 instead of 7.2.1. Thus we get:

7.6.1 Let $f(x)$ be (L) -integrable and of finite values on a finite segment $[a, b]$.

(I) If we can find to $f(x)$ a primitive function $F(x)$, then the pendant of the Newton-Leibniz formula holds:

$$\int_a^b f(x) dx = [F(x)]_a^b = F(b) - F(a). \quad (7.6-1)$$

(II) If there exists a primitive function of $f(x)$ in $[a, b]$ at all, this is necessarily an (L) -integral function, i.e., it must be of the form:

$$F(x) = \int_a^x f(t) dt + C \quad (C \text{ constant}). \quad (7.6-2)$$

We add that 7.6.1 (I) may be expressed also in the following form, equivalent to the former:

Is $F(x)$ differentiable everywhere on $[a, b]$ with an (L) -integrable derivative, so we have the connection:

$$\int_a^b F'(x) dx = F(b) - F(a). \quad (7.6-3)$$

7.62 Next we deal with the very natural question: if $F(x)$ belongs to one of the important special classes of functions treated in Chapter 4 (monotonic, absolutely continuous functions and functions of bounded variation), for which the existence of derivative is assured only *almost everywhere*, what can be said about the (L) -integrability of $F'(x)$ and the validity of (7.6-3)?

In connection with 7.6.1 (II), we have first to clarify whether the (L) -integral function

$$\Phi(x) = \int_a^x f(t) dt \quad (a \leq x \leq b) \quad (7.6-4)$$

is indeed a primitive of $f(x)$ or whether the relation

$$\Phi'(x) = f(x) \quad (7.6-5)$$

does at least hold *almost everywhere* in $[a, b]$; moreover, whether we can obtain from a primitive function of $f(x)$ the totality of primitives by addition of an arbitrary constant.

As regards the question concerning (7.6-3), the answer to this is — unfortunately — partly negative:

7.6.2 If $F(x)$ is monotone increasing in a finite interval $[a, b]$, so $F'(x)$ is (L) -integrable there, further we have the connection

$$\int_a^b F'(x) dx \leq F(b) - F(a). \quad (7.6-6)$$

which cannot always be valid with the sign of equality.

In case of a monotone decreasing $F(x)$, a similar statement holds but with $\geq$ instead of $\leq$ in (7.6-6).

P r o o f. Let $F(x)$ be monotone increasing in $[a, b]$, and let us extend its definition to a right-hand neighbourhood of b as follows: $F'(x) = F(b)$ ($b < x \leq b+1$). The function thus obtained is differentiable p.p. in $[a, b]$ by 4.2.1; where $F'(x)$ exists, it can be represented by means of a specialized difference quotient in the form:

$$F'(x) = \lim_{n \rightarrow \infty} n[F(x + \frac{1}{n}) - F(x)] \geq 0.$$

Considering the sequence $q_n(x) = n[F(x + \frac{1}{n}) - F(x)]$ ($n = 1, 2, \dots$), we see that all its terms are non-negative and (R) -integrable [a fortiori (L) -integrable] in $[a, b]$; moreover $\{q_n(x)\}$ is convergent for a.e. $x \in [a, b]$, and by the first mean value theorem of integral calculus:

$$\int_a^b q_n(x) dx = n \int_a^{b+(1/n)} F(t) dt - n \int_a^{a+(1/n)} F(t) dt \leq F(b) - F(a) \quad (n = 1, 2, \dots).$$

So the application of lemma 7.5.2 yields immediately the (L) -integrability of $F'(x)$ in $[a, b]$ and the inequality (7.6-6). That the sign of equality cannot hold generally, is shown already by the function $\mathfrak{T}(x)$ discussed in 4.53; since on the basis of 4.5.4 we can write

$$0 = \int_0^1 \mathfrak{T}'(x) dx < \mathfrak{T}(1) - \mathfrak{T}(0) = 1. \quad (7.6-7)$$

If $F(x)$ is monotone decreasing, the whole argumentation above remains valid, but the inequalities take an opposite sense.¹

* * *

¹ By 7.6.2 and 7.6.4 it is easy to see, that every monotone increasing or decreasing function can be written in the form $f(x) = \varphi_a(x) + \varphi_s(x) + \sum(x)$, where $\varphi_a(x)$ is absolutely continuous, $\varphi_s(x)$ is "singular" in the sense that [like $\mathfrak{T}(x)$] it is continuous and has a derivative vanishing a.e., finally $\sum(x)$ denotes the so-called pure jumping function associated with the discontinuities of $f(x)$ (cf. 4.1.2). There holds an analogous representation for any function of bounded variation; its name is *canonical decomposition*.

Because of the well-known connection between functions of bounded variation and monotonic functions (cf. 4.5.1), our result implies the following:

7.6.3 *On a finite segment, any function of bounded variation has (L)-integrable derivative, but the latter does not always satisfy the relation (7.6-3).*

7.6.3 Our next subject is the investigation of the function (7.6-4). The first pertinent problem can be solved essentially in a positive sense, as was indicated already in 7.33.

7.6.4 *If $f(x)$ is (L)-integrable on a finite segment $[a, b]$, then the (L)-integral function (7.6-4) is in $[a, b]$ absolutely continuous and its derivative is equal to $f(x)$ almost everywhere.*

P r o o f. 1° As regards the absolute continuity of $\Phi(x)$, this follows in the case of a bounded function $f(x)$ immediately from the fact that $\Phi(x)$ then fulfils a Lipschitz condition. (Cf. 4.62, 7.3.4.)

Assume now that $f(x)$ is not bounded in $[a, b]$. Then we write: $f(x) = f_N(x) + [f(x) - f_N(x)]$ with $f_N(x) = [f_{(+)}(X)]_N - [|f_{(-)}(x)|]_N$ [cf. (7.4-5) — (7.4-6)]. Since $|f(x) - f_N(x)| \leq |f(x)|$ and $|f(x) - f_N(x)| \rightarrow 0$ ($N \rightarrow \infty$), the application of 7.5.1 yields $\int_a^b |f(x) - f_N(x)| dx = 0$ ($N \rightarrow \infty$). Hence to any given $\eta > 0$, there exists an index $N_0(\eta)$ such that

$$\int_a^b |f(x) - f_N(x)| dx < \frac{\eta}{2} \quad \text{if only } N \geq N_0(\eta).$$

N being fixed in this manner, consider a system of non-overlapping intervals $[\alpha_k, \beta_k] \subseteq [a, b]$ with a total length $< \eta/(2N)$. Then we have:

$$\begin{aligned} \sum_k |\Phi(\beta_k) - \Phi(\alpha_k)| &= \left| \sum_k \int_{\alpha_k}^{\beta_k} f(x) dx \right| \leq \left| \sum_k \int_{\alpha_k}^{\beta_k} f_N(x) dx \right| + \\ &+ \left| \sum_k \int_{\alpha_k}^{\beta_k} [f(x) - f_N(x)] dx \right| \leq N \sum_k (\beta_k - \alpha_k) + \int_a^b |f(x) - f_N(x)| dx < \\ &< N \cdot \frac{\eta}{2N} + \frac{\eta}{2} = \eta, \end{aligned}$$

which means by (4.6-2) — (4.6-3) the absolute continuity of $\Phi(x)$ for $a \leq x \leq b$.

2° From the foregoing we know that $\Phi'(x)$ exists in $[a, b]$ apart from some null set E_0 . It is still to be shown that on the set $(a, b) - E_0$ we have (7.6-5) almost everywhere.

For that purpose, let us take a pair of rationals $u < v$ and denote by $E(u, v)$ that subset of $(a, b) - E_0$ on which $\Phi'(x) > v > u > f(x)$. Since by 7.6.3 $\Phi'(x)$ is (L) -integrable, thus $E(u, v)$ is necessarily measurable; we assert that

$$|E(u, v)| = 0. \quad (7.6-8)$$

First of all, being given some $\varepsilon > 0$, let us choose a number $\delta > 0$, less than ε , such that

$$\left| \int_{(e)} f(t) dt \right| < \varepsilon \quad \text{for} \quad |e| < \delta. \quad (7.6-9)$$

(Cf. 7.1.1.) Let us denote by S such an open subset of (a, b) , which contains $E(u, v)$ and satisfies the inequality $|S| < |E(u, v)| + \delta$.

Now let us attach to any point $x \in E(u, v)$ a segment $[x, x+l] \subseteq S$ such that

$$\frac{\Phi(x + \Delta x) - \Phi(x)}{\Delta x} > v \quad (0 < \Delta x \leq l), \quad (7.6-10)$$

which is clearly possible by the definition of $E(u, v)$. The system R of the stretches just mentioned wholly covers $E(u, v)$ and fulfils the conditions of the Sierpinski lemma 2.6.3. Consequently, to an arbitrarily prescribed number $\eta > 0$ we can select from R a finite subsystem of non-overlapping intervals $I_1 = [x_1, x_1 + l_1], I_2 = [x_2, x_2 + l_2], \dots, I_p = [x_p, x_p + l_p]$ ($a \leq x_1 \leq \dots \leq x_p < x_p + l_p$) which also covers $E(u, v)$, except a subset of $E(u, v)$ of outer (L) -measure less than η .

If we apply (7.6-10) to these intervals I_k , we get

$$\int_{x_k}^{x_k + l_k} f(t) dt > v \cdot l_k \quad (k = 1, 2, \dots, p),$$

and hence with the notation $J = \bigcup_{k=1}^p I_k$ the inequality $\int_{(J)} f(t) dt > v|J|$.

By 2.6.3 η may be however small, so that $|J|$ can approximate arbitrarily $|E(u, v)|$ from above; therefore we can assume that $|J| - |E(u, v)| < \delta < \varepsilon$, i.e., $|J| = |E(u, v)| + \vartheta \varepsilon$ ($0 \leq \vartheta < 1$). Then (7.6-10) furnishes the inequality

$$\int_{(J)} f(t) dt > v(|E(u, v)| + \vartheta \varepsilon) \quad (0 \leq \varepsilon < 1) \quad (7.6-11)$$

On the other hand, we have by the above $J \subseteq S$ and hence $J - E(u, v) \subseteq S - E(u, v)$. Since by $|J - E(u, v)| \leq |S - E(u, v)| < \delta$ and the (L) -integral of $f(t)$ on the set $J - E(u, v)$ is $< \varepsilon$ [cf. (7.6-9)], the estimation follows:

$$\int_{(J)} f(t) dt < \int_{(E(u, v))} f(t) dt + \varepsilon \quad (7.6-12)$$

But the definition of $E(u, v)$ implies also $f(t) < u$, therefore the right-hand integral in (7.6-12) is $\leq |E(u, v)|$. Thus (7.6-11) and (7.6-12) yield together:

$$v(|E(u, v)| + \vartheta\varepsilon) < u|E(u, v)| + \varepsilon,$$

whence for $\varepsilon \rightarrow +0$ we obtain

$$v|E(u, v)| \leq u|E(u, v)|.$$

This can be true by the assumption $u < v$ only if (7.6-8) holds.

3° Let E denote the totality of points $x \in (a, b) - E_0$ where $\Phi'(x) > f(x)$. Then

$$E = \bigcup_{u, v} E(u, v),$$

being extended over all the pairs $u < v$ of rational numbers. According to 2°, each subset $E(u, v)$ on the right is a null set, so by 2.6.2 we can infer also

$$|E| = 0. \quad (7.6-13)$$

In other terms: on the set $(a, b) - E_0$ we have almost everywhere $\Phi'(x) \leq f(x)$. Using the same argumentation for the function $-f(x)$, we come to the conclusion that $\Phi'(x) \geq f(x)$ at a.e. points $x \in (a, b) - E_0$. These two statements together imply the validity of the relation $\Phi'(x) = f(x)$ almost everywhere in $[a, b]$.

* * *

7.64 The theorem just proved has certain important consequences. In the first place, we shall sharpen in a sense our result on (7.6-5), verifying at the same time an earlier statement concerning Lebesgue points. (Cf. 7.3.)

7.6.5 *For every function (L) -integrable in a finite interval $[a, b]$, the relation (7.6-5) holds at any Lebesgue point of $f(x)$, and almost all points $x \in [a, b]$ are Lebesgue points.*

P r o o f. The first part of our assertion can be seen in the same way as in case of a bounded function $f(x)$. (Cf. 7.3.4.)

Let ϱ denote a rational number. Since $|f(t) - \varrho| \in L(a, b)$, in virtue of 7.6.4 we have at a.e. point $x \in [a, b]$

$$\frac{d}{dx} \int_a^x |f(t) - \varrho| dt = \lim_{l \rightarrow 0} \frac{1}{l} \int_x^{x+l} |f(t) - \varrho| dt = |f(x) - \varrho|; \quad (7.6-14)$$

i.e., the points of $[a, b]$ where (7.6-14) does not hold form a null set $E(\varrho)$. Let us denote by E_∞ the set of points where $|f(x)| = \infty$, and let $\varrho_1, \varrho_2, \dots$

be a sequence consisting of all rational numbers (cf. 1.2.1). Then by 2.6.2, $E_0 = E_\infty \cup E(\varrho_1) \cup E(\varrho_2) \cup \dots$ is a null set.

Next we show that any point $x \in [a, b] - E_0$ is a Lebesgue point of the function f .

Let $x_0 \in [a, b] - E_0$ and an arbitrarily small $\varepsilon > 0$ be given. If ϱ_n is chosen in such way that its deviation from $f(x_0)$ is less than $\varepsilon/3$, then ($l \geq 0$):

$$\begin{aligned} & \left| \frac{1}{l} \int_{x_0}^{x_0+l} |f(t) - \varrho_n| dt - \frac{1}{l} \int_{x_0}^{x_0+l} |f(t) - f(x_0)| dt \right| \leq \\ & \left| \frac{1}{l} \int_{x_0}^{x_0+l} |f(t) - \varrho_n - [f(t) - f(x_0)]| dt \right| < \\ & \left| \frac{1}{l} \int_{x_0}^{x_0+l} \frac{\varepsilon}{3} dt \right| = \frac{\varepsilon}{3}. \end{aligned} \quad (7.6-15)$$

On the other hand, (7.6-14) is valid at x_0 with $\varrho = \varrho_n$, since $x_0 \in E(\varrho_n)$. Thus for $|l|$ sufficiently small

$$\left| \frac{1}{l} \int_{x_0}^{x_0+l} |f(t) - \varrho_n| dt - |f(x_0) - \varrho_n| \right| < \frac{\varepsilon}{3},$$

i.e., by $|f(x_0) - \varrho_n| < \varepsilon/3$ we have

$$\left| \frac{1}{l} \int_{x_0}^{x_0+l} |f(t) - \varrho_n| dt \right| < \frac{2}{3} \varepsilon. \quad (7.6-16)$$

On the basis of (7.6-15) and (7.6-16) we obtain:

$$\left| \frac{1}{l} \int_{x_0}^{x_0+l} |f(t) - f(x_0)| dt \right| < \frac{\varepsilon}{3} + \frac{2}{3} \varepsilon = \varepsilon \quad \text{for } |l| < \delta(\varepsilon)$$

which means just that x_0 is a Lebesgue point.

* * *

7.65 A further relevant corollary of 7.6.4 is: *the absolutely continuous functions constitute such a distinguished subclass of the class of functions of bounded variation, for which (7.6-3) holds.* (Cf. theorem 7.6.3.) Established in a slightly more general form:

7.6.6 If $F(x)$ is absolutely continuous on a finite segment $[a, b]$, then it is equal to the (L) -integral function of its derivative, namely

$$F(x) = \int_a^x F'(t) dt + F(a) \quad (a \leq x \leq b).^1 \quad (7.6-17)$$

Proof. As we know, absolute continuity of $F(x)$ involves the existence p.p. of $F'(x)$ and by 7.6.3 its (L) -integrability in $[a, b]$, too.

Therefore we can form the (L) -integral function

$$F^*(x) = \int_a^x F'(t) dt + F(a) \quad (x \in [a, b]), \quad (7.6-18)$$

for which the application of 7.6.4 [with $f(t) = F'(t)$] yields: $F^*(x)$ is absolutely continuous as well and satisfies the relation $[F^*(x)]' = F'(x)$ almost everywhere in $[a, b]$.

But if $[F^*(x) - F(x)]' = 0$ for a.e. $x \in [a, b]$, then $F^*(x) - F(x)$ is necessarily a constant by 4.6.3; and the latter must be 0 because $F^*(a) - F(a) = 0$.

Thus we get $F^*(x) = F(x)$ ($a \leq x \leq b$), and (7.6-18) turns into (7.6-17).

* * *

Let us emphasize: 7.6.6 enables us to give a positive answer concerning the second question on 7.6.1 (II). For, if $f(x)$ possesses in $[a, b]$ a primitive function $F(x)$, i.e., we have $F'(x) = f(x)$ at all the points $x \in [a, b]$, then $F(x)$ must be absolutely continuous, so that 7.6.6 implies:

$$F(x) = \int_a^x f(t) dt + F(a) = \Phi(x) + F(a) \quad (a \leq x \leq b). \quad (7.6-19)$$

Hence we see that *the difference of any two primitive functions must be constant*.

The situation changes when $F'(x) = f(x)$ holds only *almost everywhere*. [Then $F(x)$ may not be called a primitive of $f(x)$.] In such a case the validity of (7.6-19) is not assured; for example, taking $F(x) = \mathfrak{T}(x)$, $f(x) = 0$, $[a, b] = [0, 1]$, we have $F'(x) = f(x)$ p.p., but $F(x)$ is not identical with $\int_a^x f(t) dt + F(a)$ ($\equiv 0$). (Cf. 4.5.4.)

¹ Observe that for $x = b$, (7.6-17) becomes (7.6-3).

Moreover, the connection of 7.6.4 with 7.6.6 shows that *absolute continuity is the characteristic property of (L)-integral functions for arbitrary (L)-integrable integrands.*

Thus the counterpart of 7.3.5 in the general case can be formulated as follows:

7.6.7 A function $F(x)$ defined in a finite interval $[a, b]$ can be represented as the (L)-integral function of a bounded or unbounded function $f(x)$, if and only if $F(x)$ is absolutely continuous in $[a, b]$.

Finally we mention another remarkable consequence of 7.6.6, namely that *the total variation of the (L)-integral function $\Phi(x)$ (as a function of bounded variation) is simply the (L)-integral of $|f(x)|$; in symbols:*

$$V_{\Phi}(a, b) = \int_a^b |f(x)| dx. \quad (7.6-20)$$

As a background to this, we merely refer to the elementary formula:

$$V_F(a, b) = \int_a^b |F'(x)| dx, \quad (7.6-21)$$

where F' is supposed to exist and be (R)-integrable in $[a, b]$. (Cf. 4.41.) The latter representation can be transferred to the case where F is absolutely continuous and we have an (L)-integral on the right.

7.66 Looking over the results of this section on the relation between differentiation and (L)-integration, we have to point out the regrettable fact that *the concept of Lebesgue integral renders possible the inversion of derivation* (i.e., to solve wholly the problem of primitives) *only in the case of a bounded integrand.* (Cf. 7.32.)

For, if we know about the derivative of $F(x)$ nothing more than it exists on a finite segment $[a, b]$, therefrom we may by no means conclude the (L)-integrability of $F'(x)$ and the validity of (7.6-17). On the contrary: in case of unbounded $F'(x)$ it may happen that $\int_a^x F'(t) dt$ as a Lebesgue integral does not exist. [A pertinent example is furnished by the function (7.4-7) on $0 < x \leq 1$.]

The circumstance at issue led to further extensions of the integral concept — by A. DENJOY and O. PERRON — in the first decades of this century.¹ Nevertheless, these generalizations have much less significance in respect of applications.

¹ Cf. Denjoy [1], Perron [1], further F. Riesz [25].

7.67 In order to terminate to a certain extent the study of (L) -integral functions, we touch upon the rules of *integration by parts and by substitution*. The former is mostly used in the theory of (L) -integral in the following form:

7.6.8 If $f(x)$, $g(x)$ are (L) -integrable on a finite segment $[a, b]$, and $F(x)$ is an (L) -integral function of f , $G(x)$ that of g , then the products $f(x)G(x)$, $F(x)g(x)$ are also (L) -integrable in $[a, b]$, and there holds the formula:

$$\int_a^b f(x)G(x) dx + \int_a^b F(x)g(x) dx = F(b)G(b) - F(a)G(a). \quad (7.6-22)$$

Proof. On our conditions, $F(x)$ and $G(x)$ are absolutely continuous in $[a, b]$ by 7.6.4, and we have almost everywhere $F'(x) = f(x)$, $G'(x) = g(x)$. Thus on the basis of 4.6.1, $F(x)G(x)$ is also absolutely continuous, moreover differentiable at a.e. point $x \in [a, b]$ and

$$[F(x)G(x)]' = f(x)G(x) + F(x)g(x). \quad (7.6-23)$$

Since the last two terms are obviously (L) -integrable in $[a, b]$ by the boundedness of $F(x)$ and $G(x)$, further by (7.6-17) we can write:

$$\int_a^b [F(x)G(x)]' dx = F(b)G(b) - F(a)G(a),$$

the (L) -integration in $[a, b]$ on both sides of (7.6-23) leads to (7.6-22).

* * *

Note that our result can be transformed immediately into the classical variant of the rule of integration by parts:

$$\int_a^b F'(x)G(x) dx + \int_a^b F(x)G'(x) dx = F(b)G(b) - F(a)G(a) \quad (7.6-24)$$

— where $F(x), G(x)$ are *absolutely continuous* functions — if we put $F'(x)$ instead of $f(x)$ and $G'(x)$ instead of $g(x)$ in (7.6-22).

The rule of *integration by substitution* is:

$$\int_a^b f(x) dx = \int_{\alpha}^{\beta} f(\gamma(t)) \cdot \gamma'(t) dt. \quad (7.6-25)$$

It holds on the conditions: let $\gamma(t)$ be strictly monotonic and absolutely continuous in the finite interval $[\alpha, \beta]$, further let $f(x)$ be (L) -integrable in

the corresponding interval $[a, b]$ with $a = \gamma(a)$, $b = \gamma(\beta)$. The significance of (7.6-25) turns out in the theory of the so-called Lebesgue-Stieltjes integral, to be presented later. For this reason, the formula (7.6-25) will be discussed within a more general framework in 8.34.

7.7 (L)-integral of several dimensions; Fubini's theorem on successive integration

7.71 Since the (L)-measure was defined in 6.12 for point sets of however many dimensions, it is natural to introduce the (L)-integral of a point function in the Euclidean ν -space by way of the associated concept of measure, patterned after the one-dimensional case.

For instance, let $\nu = 2$. We define the (L)-integral of a single-valued (bounded or unbounded) function $z = f(x, y)$ over some rectangle $T = [a, b] \times [c, d]$ as follows:

$$\int\limits_{(T)} f(x, y) dx dy = \lim_{\delta \rightarrow 0} \sum_{k=-\infty}^{\infty} \zeta_k |T_k| \quad (\sup\{\Delta z_k\} < \delta) \quad (7.7-1)$$

where the sum on the right relates to a δ -fine scale

$$\dots < z_{-k} < \dots \quad z_{-1} < z_0 < z_1 < \dots < z_k < \dots \quad (7.7-2)$$

on the z -axis, further $z_{k-1} \leq \zeta_k < z_k$, $\Delta z_k = z_k - z_{k-1}$ and T_k means the set of points $(x, y) \in T$ for which $z_{k-1} \leq f(x, y) < z_k$ ($k = 0, \pm 1, \pm 2, \dots$). We have to assume, of course, that the integrand f is *measurable* in the sense that the planar set consisting of points $(x, y) \in T$ with $f(x, y) < C$ is (L)-measurable for any C ; moreover we need the restriction that the sum (7.7-1) should be *absolutely convergent* in the case of at least one admitted decomposition of the z -axis.

Like the double (R)-integral, an (L)-integral of the type (7.7-1) can also be calculated by *successive (L)-integration* with respect to the variables x and y . More precisely:

7.7.1 (FUBINI'S THEOREM ON DOUBLE INTEGRALS)¹ Suppose that the (L)-integral (7.7-1) exists. Then

(I) Λ denoting the set of $x \in [a, b]$ for which $f(x, y)$ is (L)-integrable with respect to y in the interval $[c, d]$, the difference $[a, b] - \Lambda$ is a null set;

(II) the integral $\int_c^d f(x, y) dy$ is (L)-integrable with respect to x on the set Λ ;

¹ Fubini [1].

(III) the formula holds:

$$\int_{(T)} f(x, y) dx dy = \int_{(A)} \left[\int_c^d f(x, y) dy \right] dx. \quad (7.7-3)$$

If $f(x, y) \geq 0$, the integral (7.7-1) means geometrically the spatial (L)-measure of the ordinate set of the integrand on T .

7.72 A transformation rule analogous to that just given is valid for ν -dimensional (L)-integrals. It is advisable to carry out in detail the verification of 7.7.1 first after having accomplished the extension of Lebesgue's theory in the direction suggested by modern axiomatic theory of probability.

The same domain yields the most natural background for the pertinent pendant of *Stieltjes integral* which was originally defined by Lebesgue in the form:

$$\int_a^b f(x) dg(x) = \int_A^B f(g_{-1}(y)) dy. \quad (7.7-4)$$

Here $f(x)$ is assumed to be (L)-integrable, $y = g(x)$ denotes a function strictly monotone increasing in $[a, b]$; $x = g_{-1}(y)$ is the inverse of the latter, further $A = g(a)$, $B = g(b)$. (Cf. 5.74.)

The above-mentioned topics, i.e., the foundations of a "Lebesgue-like" measure and integral theory relating to *abstract sets* will be dealt with in the forthcoming chapter.

MEASURES DEFINED ON BOOLEAN RINGS AND THE ABSTRACT LEBESGUE INTEGRAL. LEBESGUE-STIELTJES INTEGRAL

8.1 Extension of measures, Carathéodory's theorem

8.11 The development of several branches of mathematics, first of all of probability theory in the first half of this century necessitated the transfer of the concepts of measure and integral to certain basic sets in which we cannot speak about "distance" or "neighbourhood" at all, and the geometric background usual in the theory of point sets is completely missing. In such a case it is advisable to take Lebesgue's theory as a model in all respects: postulating some elementary properties of the Lebesgue measure $\mu_L(E)$, we can guarantee the possibility of introducing an abstract concept of integral which possesses the well-proved, advantageous features of (L) -integral.

As regards Lebesgue's measure theory in the Euclidean ν -space, we have started from E_ν as a basic set and considered a σ -algebra $\tilde{S}_L$ consisting of certain subsets of E_ν (namely the totality of sets E which are (L) -measurable in a wider sense¹), in order to define $\mu_L(E)$ as a set function associated with $\tilde{S}_L$, which vanishes for $\bigcirc$ and is otherwise non-negative, as well as totally additive. (Cf. 1.51., 5.11 and 6.14.)

8.12 Let now be given an arbitrary basic set $X \neq \bigcirc$ and a σ -algebra $\tilde{S}$ consisting of certain subsets $E \in X$; suppose that we can define uniquely for the elements of $\tilde{S}$ some set function $\mu(E)$ such that I. $\mu(\bigcirc) = 0$; II. $0 \leq \mu(E) \leq \infty$; III. $\mu(\cup E_k) = \sum \mu(E_k)$ in the case of any countable system of pairwise disjoint sets $E_k \in \tilde{S}$. Then $\mu(E)$ is called an *abstract measure* (in short μ -measure) belonging to the pair $(X, \tilde{S})^2$; the sets $E \in \tilde{S}$ are said to be μ -measurable, while the name of the triplet $(X, \tilde{S}, \mu)$ is: *measure space*.

It is easy to present instructions to give explicitly an arbitrary number of concrete measure spaces. Such an instruction is e.g. the following: let X be a countable set $\{x_k\}$ and $\tilde{S} = P(X)$, i.e., the *power-set* of X [cf. (1.4-3), 1.51], furthermore let us take:

$$\mu(E) = \begin{cases} 0 & \text{if } E = \bigcirc, \\ \sum_k \varrho(\xi_k) & \text{if } E = \{\xi_k\}, \end{cases} \quad (8.1-1)$$

¹ I.e., which have a finite or infinite (L) -measure.

² Sometimes we also say that X and $\tilde{S}$ form together a *measurable space*.

where $\{\xi_k\} \subseteq \{x_k\}$ and $\varrho(x)$ is some single-valued, non-negative real function, defined for every $x \in X$.

8.13 Anyway, those measure spaces are mainly of significance for the applications whose basic set X is non-countable and at which $\tilde{\mathcal{S}}$ contains together with any element E of μ -measure zero also all the subsets of E ($\subset X$). If this latter condition is fulfilled, we say that both the measure space $(X, \tilde{\mathcal{S}}, \mu)$ and the associated concept of measure are *complete*.

In particular, the triplet $(E_\nu, \tilde{\mathcal{S}}_L, \mu_L)$ and the Lebesgue measure $\mu_L(E)$ obviously possess the property of completeness, but this cannot be asserted about every classical concept of measure. For instance, the so-called *Borel's measure* $\mu_B(E)$ is not complete with respect to the Euclidean ν -space as a basic set. The measure μ_B was introduced by E. BOREL in 1898 by extending the notion of measure from intervals of E_ν in a certain "natural" way to the totality of sets belonging to the σ -algebra $\tilde{\mathcal{S}}_B$ generated by systems of ν -dimensional intervals.¹ (By "natural" extension is meant: to the union of pairwise disjoint sets we always attach the sum of measures of the latter, and to the difference of a set E and of some subset of E we always attach the difference of the corresponding measures as the measure of the result.) It is easy to see that every set measurable in Borel's sense is (L) -measurable at the same time, but the converse is not true; thus there are sets E for which $\mu_L(E)$ exists and $\mu_B(E)$ does not, though in such a case it is always possible to find some Borel measurable set $E_1 \supseteq E$ for which $\mu_B(E_1) = \mu_L(E)$.

The last statement shows that the incomplete measure space $(E_\nu, \tilde{\mathcal{S}}_B, \mu_B)$ can be expanded to the complete measure space $(E_\nu, \tilde{\mathcal{S}}_L, \mu_L)$ thereby that the elements of the former are "widened" by all the subsets of null sets with respect to μ_B . We can proceed similarly in the case of any other incomplete measure space so that the following proposition holds:

8.1.1 *To each space $(X, \tilde{\mathcal{S}}, \mu)$ there exists a narrowest complete expansion which is called the closure of $(X, \tilde{\mathcal{S}}, \mu)$. The σ -algebra containing $\tilde{\mathcal{S}}$ which occurs in this closure, as well as the complete measure generalizing μ are determined uniquely.*

8.14 As regards the construction of measure spaces fitting the purposes of an integral theory, a further important method is furnished by studying the logical connection between Jordan's and Lebesgue's theory of measure.

¹ Hence the terms "Borel set" or "Borel's field of sets" have originated. (Cf. Borel [2].) By means of a strong widening of Borel's method, A. HAAR succeeded in introducing an abstract measure concept for certain algebraic structures — namely in the case of the so-called locally compact topological groups — too (1933). This is a notion of great importance; its name is "Haar's measure". (Cf. Haar [2].)

The (J)-measure is defined not on a σ -algebra, but only on a Boolean ring (cf. 5.13); and the idea of the outer (L)-measure which is based implicitly upon the Jordan measure, enables us to extend $\mu_J(E)$ to the Lebesgue measure $\mu_L(E)$ which is defined on the wider set structure $\tilde{\mathcal{S}}_L$. [Cf. (3.6-5), (6.1-1).]

In order to generalize this train of thought, we shall use a notion analogous to $\mu_J(E)$, for which it is advisable to introduce a new name and symbol. Let $\mathcal{R}$ be a Boolean ring consisting of certain subsets E of an arbitrary basic set $X \neq \emptyset$. We call any single-valued set function defined on $\mathcal{R}$ a *premeasure* with respect to $(X, \mathcal{R})$ and denote it by $\mu^*(E)$, when it has the following properties: I. $\mu^*(\emptyset) = 0$; II. $0 \leq \mu^*(E) \leq \infty$; III. $\mu^*(\cup E_k) = \sum \mu^*(E_k)$ for every countable system of pairwise disjoint sets $E_k \in \mathcal{R}$ satisfying the condition $\cup E_k \in \mathcal{R}$.¹

After these, it is natural to form the so-called *outer measure induced by* $\mu^*(E)$ [cf. (2.6-5)]:

$$\bar{\mu}(E) = \inf_{U \supseteq E} \{\mu^*(U)\} \quad (8.1-2)$$

where E denotes an arbitrary subset of X and U is a countable system of subsets belonging to $\mathcal{R}$. Since $\bar{\mu}(E)$ means an extension of $\mu^*(E)$ in a certain direction, it is to be expected and indeed can be verified without difficulty that these two set functions are identical with each other on $\mathcal{R}$ itself:

$$\bar{\mu}(E) = \mu^*(E) \quad \text{if } E \in \mathcal{R}. \quad (8.1-3)$$

On the other hand, for any (L)-measurable set E we can write $\bar{\mu}_L$ instead of μ_L in formula (6.1-1); further ibidem on the right I may be replaced by another (L)-measurable set H , if at the same time on the left $H \cap E$ is substituted for E . Thus the idea presents itself that the relation

$$\bar{\mu}(H \cap E) + \bar{\mu}(H - E) = \bar{\mu}(H) \quad [H \in P(X)] \quad (8.1-4)$$

should be utilized for the characterization of the abstract measure concept corresponding to the Lebesgue measure. More precisely, we see that the subsets $E \in X$ which satisfy (8.1-4) for any $H \in P(X)$, always constitute a σ -algebra $\tilde{\mathcal{S}}$, and that $\bar{\mu}(E)$ is simultaneously a measure in the sense given before, if it is restricted to $\tilde{\mathcal{S}}$. This measure concept $\mu(E)$ associated with $(X, \tilde{\mathcal{S}})$ is called *the measure induced by $\bar{\mu}(E)$* .

8.15 If now starting from a premeasure μ^* , we form first the outer measure $\bar{\mu}(E)$ by (8.1-2) and afterwards we define $\mu(E)$ on the basis of (8.1-4), then, in the way just sketched, we get one of the relevant results of general measure theory:

¹ The last requirement means, of course, a restriction only in the case of *infinitely many* sets, because otherwise it is a consequence of the fact that $\mathcal{R}$ is a *Boolean ring*.

8.1.2 (CARATHÉODORY'S EXTENSION THEOREM) *Let us consider a Boolean ring $\mathcal{R}$ belonging to an arbitrary basic set $X \neq \emptyset$ and a σ -algebra $\tilde{\mathcal{S}}(\mathcal{R})$ generated by $\mathcal{R}$. Every premeasure with respect to $(X, \mathcal{R})$ can be extended to a μ -measure associated with $(X, \tilde{\mathcal{S}}(\mathcal{R}))$.*

Note that 8.1.2 can even be sharpened, if we still assume that X can be represented as the union of its countably many, pairwise disjoint subsets belonging to $\mathcal{R}$ and being of finite premeasure. In such a case we say that the pertinent premeasure is σ -finite on $\mathcal{R}$. In the given case, the extension in question is unique, i.e., there exists one single measure defined on $\tilde{\mathcal{S}}(\mathcal{R})$ which is identical with the premeasure just mentioned on the Boolean ring $\mathcal{R}$.¹

The extension of a premeasure μ^* , together with expanding the induced measure $\mu(E)$ into a complete one, and furthermore, the resulting measure itself are called the *Lebesgue continuation* of μ^* .

8.2 Lebesgue integral on abstract measure space

8.21 Let us investigate, how the concept of (L) -integral can be transferred to functions given on abstract measure spaces, the measurability of these functions being naturally meant according to the notion of measure valid in the pertinent measure space.

Consider an arbitrary *non-countable* basic set X , a σ -algebra $\tilde{\mathcal{S}}$ consisting of certain subsets of X , further a measure $\mu(E)$ defined on $\tilde{\mathcal{S}}$, this triplet forming together a measure space $(X, \tilde{\mathcal{S}}, \mu)$. The latter, as well as $\mu(E)$ are always supposed in the sequel to be *complete* — without any particular specification.

A real-valued function $f(x)$, defined uniquely for the elements of X , is called μ -measurable on a set $E \subseteq X$, if both E and the level set $\{x : f(x) < c\}$ with any real c are μ -measurable, i.e., belong to $\tilde{\mathcal{S}}$ (cf. 8.12). Since this definition is formally the same as that given in 6.32, it is clear that all the *properties* of measurable functions, discussed at the end of Chapter 6, remain valid also for μ -measurable functions. In addition, we want to mention the following:

Let $\varphi(x)$ be called μ -step function, if its codomain consists of countably many real numbers and φ takes each of its values at the points of a μ -measurable set. (Such a function is obviously μ -measurable.) Having this notion, we easily obtain an elementary criterion of μ -measurability:

8.2.1 *A function $f(x)$ is μ -measurable on a set $E \in \tilde{\mathcal{S}}$ if and only if it can be represented as the limit function of a uniformly convergent sequence of μ -step functions.*

¹ Cf. Carathéodory [1].

P r o o f. The *sufficiency* of the condition follows at once from the fact that proposition 6.5.2 holds for any μ -measurable basic set and for arbitrary μ -measurable functions as well.

To show the *necessity*, suppose that $f(x)$ is μ -measurable on E , and form the μ -step functions

$$f_n(x) = \frac{k}{n} \quad \text{with} \quad \frac{k}{n} \leq f(x) < \frac{k+1}{n} \quad (k = 0, \pm 1, \pm 2, \dots), \quad (8.2-1)$$

n denoting an arbitrary natural number. We see that the sequence $\{f_n(x)\}$ converges on E uniformly to $f(x)$, since $|f(x) - f_n(x)| \leq 1/n$ at every point x .

* * *

The counterparts of ordinary null sets concerning μ -measurable functions are the sets e for which $\mu(e) = 0$. If some assertion holds for all values of a variable except a so-called μ -null set, then we say that it is valid μ -almost everywhere (in short: " μ -a.a."). When no misunderstanding is possible, the reference to μ -measure may be omitted. If two functions: $f(x)$ and $g(x)$ are equal μ -almost everywhere, i.e., the μ -measure of the set $\{x : f(x) \neq g(x)\}$ is equal to 0, then we say also that f and g are μ -equivalent to one another.

8.22 Now, having regard to the fact that the Lebesgue sum (7.4-2) can also be interpreted as the (L) -integral of a generalized step function [cf.(6.4-6)], it is advisable to carry out the extension of the integral concept in two steps, by use of the integral of some μ -step function.

Let $\varphi(x)$ be a μ -step function, defined on a set $E \in \tilde{\mathcal{S}}$, the values of which can be arranged in a finite or infinite sequence of distinct real numbers. If the series corresponding to (7.4-2), i.e.,

$$\sum_k c_k \mu(\{x : x \in E, \varphi(x) = c_k\}) \quad (8.2-2)$$

is absolutely convergent, then we say that $\varphi(x)$ is μ -integrable on the set E and the sum of (8.2-2) is called the μ -integral relating to E . In symbols:

$$\int_{(E)} \varphi(x) d\mu = \sum_k c_k \mu(\{x : x \in E, \varphi(x) = c_k\}). \quad (8.2-3)$$

An arbitrary function $f(x)$ ($x \in E$) is said to be μ -integrable on a set $E \in \tilde{\mathcal{S}}$, if there exists a sequence $f_n(x)$ ($n = 1, 2, \dots$) of μ -step functions, which are μ -integrable on E in the sense just mentioned and converge uniformly to $f(x)$. Then we call the limit

$$\int_{(E)} f(x) d\mu = \lim_{n \rightarrow \infty} \int_{(E)} f_n(x) d\mu \quad (8.2-4)$$

the μ -integral or abstract Lebesgue integral of $f(x)$ on set E .

This definition of integral is satisfactory in every respect, since it is easy to show: (I) the limit on the right exists surely on our conditions; (II) this limit does not depend upon the choice of the sequence of μ -step functions $f_n(x)$ ($n = 1, 2, \dots$); (III) if $f(x)$ is itself a μ -step function, so the definitions of μ -integrability and μ -integral are in accordance with those given by (8.2-2)—(8.2-3).

In fact: (I) for any pair of terms of $\{f_n(x)\}$ we have the evident estimation:

$$\left| \int_{(E)} f_m(x) d\mu - \int_{(E)} f_n(x) d\mu \right| \leq \mu(E) \sup_{x \in E} \{|f_m(x) - f_n(x)|\};$$

(II) if two μ -step functions: $f_n^{(1)}(x)$ and $f_n^{(2)}(x)$ are μ -integrable on E and they converge uniformly to $f(x)$ as $n \rightarrow \infty$, so the corresponding μ -integrals relating to E tend necessarily to the same value; (III) in case of a μ -step function $\varphi(x)$, the application of the sequence $f_n(x) = \varphi(x)$ ($n = 1, 2, \dots$) yields the above definition of $\int_{(E)} \varphi(x) d\mu$.

8.23 We remark that by the foregoing the μ -integrability always implies the μ -measurability but the converse is not true (cf. 8.2.1). On the other hand, since bounded μ -step functions are trivially μ -integrable according to (8.2-3), each bounded and μ -measurable function is also μ -integrable. Moreover, it follows from the absolute convergence of the series (8.2-2) that $f(x)$ and $|f(x)|$ are always μ -integrable at the same time.

In general, we can say: the fundamental properties of (L) -integral and all the results concerning limiting processes under the integral sign (cf. 7.1-7.2, 7.4-7.5) can be transferred without difficulty to μ -integrals, but the theorems about the connection of differentiation and integration keep their validity only partly and on strong restrictions, respectively.

By the comparison of definitions of μ -integral and higher dimensional (L) -integrals we see immediately that the latter are special cases of the abstract Lebesgue integral, relating to measure spaces of the form $(E_\nu, \tilde{S}_L, \mu_L)$.

It remains to be mentioned: another type of abstract (L) -integral can be obtained if we apply to arbitrary linear functionals a method like that by which F. Riesz constructed his theory of (L) -integral, starting from termwise integration of certain "simple" functions. (Cf. 6.43.) This way was realized by the concept of integral due to P.J. DANIELL.¹

¹ Daniell [1].

8.3 Lebesgue–Stieltjes measure and integral ¹

8.31 We want to make a more detailed study of that important particular case of μ -integral, which corresponds in the Lebesgue theory to the Riemann–Stieltjes integral discussed in Sections 5.7–5.8. In order to introduce the approximate notion of measure, we need a steady point function $g(x)$, the so-called *integrator*, on which we assume in the sequel that it is (I) defined uniquely for $-\infty < x < \infty$; (II) monotone increasing on the whole number line; (III) left-hand continuous everywhere.²

Let us take as a basic set X the interval $(-\infty, \infty)$, and consider those interval systems I on the number line which consist of finitely many non-overlapping finite closed segments — admitting that the latter may “degenerate” into single points. The interval systems in question together with the empty set constitute a Boolean ring $\mathcal{R}_B$ (cf. 1.52). Let us attach to any interval system $I = \bigcup_{k=1}^n [\alpha_k, \beta_k]$ belonging to $\mathcal{R}_B$ the pertaining sum of increments of the function $g(x)$, and to $\emptyset$ the zero; therefore, if the arising interval function is denoted by μ_g^* , we shall have:

$$\mu_g^*(\emptyset) = 0, \quad \mu_g^*(I) = \sum_{k=1}^n [g(\beta_k) - g(\alpha_k)]. \quad (8.3-1)$$

The conditions postulated in 8.14 for $\mu^*(E)$ are obviously satisfied by $\mu_g^*(I)$, so that the latter may be regarded as a *premeasure* (i.e., a generalized “total length”) of I with respect to g .

But by Carathéodory’s theorem 8.1.2 μ_g^* is “Lebesgue continuable” into a measure μ_g which is defined on a σ -algebra $\tilde{\mathcal{S}}_B$ generated by $\mathcal{R}_B$, i.e., on the so-called Borel sets of the line; and by 8.1.1, the definition of μ_g can immediately be expanded to the closure of $\tilde{\mathcal{S}}_B$ (as a fundamental domain), i.e., to all the linear sets of finite or infinite (L) -measure. The measure concept thus obtained is called the *Lebesgue–Stieltjes measure with respect to g* .

8.32 Substituting now μ_g for the μ -measure in the definition of abstract Lebesgue integral, and putting $X = (-\infty, \infty)$, we get the so-called *Lebesgue–Stieltjes* [briefly *(LS)*-] *integral of $f(x)$ with respect to $g(x)$* , relating to a linear set $E \in \tilde{\mathcal{S}}_L$ which is assumed to be (L) -measurable in wider sense. The *(LS)*-integral is usually denoted like the *(RS)*-integral, so that we can write:

$$\int_{(E)} f(x) dg(x) = \int_{(E)} f(x) d\mu_g, \quad (8.3-2)$$

¹ Cf. Lebesgue [8], Kamke [1].

² $g(x)$ must not be bounded, so that its limit for $x \rightarrow \pm\infty$ may be $+\infty$ and $-\infty$, respectively.

where μ_g signifies the (LS) -measure associated with g (i.e., the Lebesgue continuation of the above μ_g^*).¹

Next we show that the (LS) -integral turns into the (RS) -integral, if E is a finite interval, the integrand is continuous, furthermore the values of the integrator are suitably chosen at the end points. So we can say: *the Lebesgue-Stieltjes integral generalizes the concept of Riemann-Stieltjes integral in a certain direction.*

8.3.1 Suppose that $f(x)$ is continuous on a finite segment $[a, b]$, further $g(x)$ is there monotone increasing, left-hand continuous and at the end points on both sides continuous. Then the (LS) - and (RS) -integral of $f(x)$ with respect to $g(x)$ over $[a, b]$ exist and are equal to one another, in symbols:

$${}^{(L)}\int_a^b f dg = {}^{(R)}\int_a^b f dg. \quad (8.3-3)$$

P r o o f. On our assumptions, the existence of the pertinent integrals is assured by 8.2.1 and 5.7.3 (I); therefore we only have to verify the equality (8.3-3).

For that purpose, consider a partition $a = x_0 < x_1 < \dots < x_n = b$ of $[a, b]$ whose every point of subdivision is a point of continuity of $g(x)$. The set of these point systems of decomposition is non-countable: omitting from (a, b) the points of discontinuity of $g(x)$ (which form a countable set by 4.1.2), the points $x_1, x_2, \dots, x_{n-1}$ may arbitrarily be marked out on the remaining part of $[a, b]$, while the points $x_0 = a, x_n = b$ belong to the set of points in question by hypothesis.

Let $I_k = [x_{k-1}, x_k]$ ($k = 1, 2, \dots, n$) be subintervals of an admitted partition, and let us denote by M_k and m_k the maximum and minimum of $f(x)$ in I_k , respectively. Then by fundamental properties of the μ -integral we have:

$$m_k \mu_g(I_k) \leq \int_{(I_k)} f(x) d\mu_g \leq M_k \mu_g(I_k) \quad (k = 1, 2, \dots, n),$$

where $\mu_g(I_k) = g(x_k) - g(x_{k-1})$ [cf. (8.3-1)]; hence by addition:

$$\sum_{k=1}^n m_k [g(x_k) - g(x_{k-1})] \leq {}^{(L)}\int_a^b f dg \leq \sum_{k=1}^n M_k [g(x_k) - g(x_{k-1})]. \quad (8.3-4)$$

On the other hand, the sum (5.7-2) occurring in the definition of ${}^{(R)}\int_a^b f dg$ lies for any partition between the two sums in (8.3-4), and

¹ If the integral on the left-hand side in (8.3-2) exists, we say also that f is (LS) -integrable with respect to g , or it is μ_g -integrable over E .

the common limit of these sums is just the (RS) -integral at issue as $\max(x_k - x_{k-1}) \rightarrow 0$.

Consequently, by applying a sequence of unboundedly refined partitions of $[a, b]$, (8.3-4) becomes indeed (8.3-3).

* * *

8.33 As far as the connection between the (LS) -integral and the ordinary (L) -integral is concerned, first of all we point out: *the (L) -integral is that particular case of the (LS) -integral for which $g(x) = x$.* This is evident because in the mentioned case μ_g becomes the Lebesgue measure μ_L and so the right-hand term of (8.3-2) turns into the (ordinary) (L) -integral of f over E .

In another connection it is an essential circumstance that *under certain restrictions the (LS) -integral can be transformed into an (L) -integral even in two different ways.*

We have hinted at one of these bearings already in 7.72:

8.3.2 *Let $y = g(x)$ be continuous and strictly monotone increasing in a finite interval $[a, b]$. If the (LS) -integral of $f(x)$ with respect to $g(x)$ exists on $[a, b]$, then it can be written as an (L) -integral as follows:*

$$\int_a^b f(x) dg(x) = \int_{g(a)}^{g(b)} f[g_{-1}(y)] dy, \quad (8.3-5)$$

where g_{-1} signifies the inverse function of g . Conversely: the existence of the right-hand side (L) -integral implies that of the left-hand side (LS) -integral and the equality (8.3-5).¹

P r o o f. In our case, the function $y = g(x)$ establishes a one-to-one, continuous mapping between the segments $a \leq x \leq b$ and $g(a) \leq y \leq g(b)$, further the same holds for $x = g_{-1}(y)$. From the fact that the μ_g -measure of an interval $[\alpha, \beta] \subseteq [a, b]$ is nothing else but the ordinary measure (i.e., length) of the corresponding interval $[g(\alpha), g(\beta)] \subseteq [g(a), g(b)]$, we can conclude that, in general, the μ_g -measure of any (L) -measurable set E on the x -axis is identical with the (L) -measure of the image set of E on the y -axis.

Hence it follows for every μ_g -step function $\varphi(x)$ whose μ_g -integral exists in $[a, b]$ and which takes there the values $c_1, c_2, \dots$:

¹ The theorem holds with some slightly more general premises too.

$$\begin{aligned} & \sum_k c_k \mu_g(\{x : x \in [a, b], \varphi(x) = c_k\}) = \\ & = \sum_k c_k \mu_L(\{y : y \in [g(a), g(b)], \varphi[g_{-1}(y)] = c_k\}) \end{aligned}$$

this involving by (8.2-3) the validity of (8.3-5) with $f(x) = \varphi(x)$.

Owing to the definition (8.2-4), we see therefore that (8.3-5) also holds generally on the existence of any of the pertinent two integrals.

* * *

The second integral formula indicated before is a counterpart of (5.8-1) for Lebesgue integrals:

8.3.3 *If $g(x)$ is monotone increasing and absolutely continuous on a finite segment $[a, b]$, further $f(x)$ is there μ_g -integrable, then*

$$\int_a^b f(x) dg(x) = \int_a^b f(x) g'(x) dx. \quad (8.3-6)$$

P r o o f. Let $\varphi(x)$ be a μ_g -step function of the form

$$\varphi(x) = \sum_k c_k \chi_{e_k}(x) \quad (a \leq x \leq b)$$

where $\{c_k\}$ is a countable set and e_k ($k = 1, 2, \dots$) are pairwise disjoint sets with $\bigcup_k e_k = [a, b]$; assume that φ is μ_g -integrable in $[a, b]$.

Then on the one hand, we can write by (8.3-2) and (8.2-3):

$$\int_a^b \varphi(x) dg(x) = \int_a^b \varphi(x) d\mu_g = \sum_k c_k \mu_g(e_k), \quad (8.3-7)$$

and the last series is absolutely convergent by hypothesis. On the other hand, it follows from the absolute continuity of $g(x)$:

$$\begin{aligned} \int_a^b \varphi(x) g'(x) dx &= \sum_k c_k \int_a^b \chi_{e_k}(x) g'(x) dx = \\ &= \sum_k c_k \int_{(e_k)} g'(x) dx = \sum_k c_k \mu_g(e_k). \end{aligned} \quad (8.3-8)$$

Namely, proposition 7.6.6 implies immediately for any interval system $I \subseteq [a, b]$ the formula

$$\int_{(I)} g'(x) dx = \mu_g(I) \quad (8.3-9)$$

[cf. (8.3-1)]; and the validity of (8.3-9) may be extended to an arbitrary (L) -measurable set $E \subseteq [a, b]$ on the basis of the definition of outer (L) -measure.

The comparison of (8.3-7) and (8.3-8) furnishes the connection

$$\int_a^b \varphi(x) dg(x) = \int_a^b \varphi(x) g'(x) dx,$$

whence we get the assertion, if we consider (8.2-4) as well as the definition of (L) -integral.

* * *

8.3.4 Observe that the formula (7.6-24) of *integration by parts* may be written by means of 8.3.3 in the form

$$\int_a^b f(x) dg(x) + \int_a^b g(x) df(x) = [f(x)g(x)]_a^b, \quad (8.3-10)$$

f and g denoting arbitrary absolutely continuous functions. [Cf. (5.7-8).]

Moreover, we get as a very simple consequence of 8.3.2 another important formula which contains the rule of *integration by substitution* for (L) -integrals as a particular case.

8.3.4 (TRANSFORMATION RULE OF LEBESGUE-STIELTJES INTEGRALS) Let $\gamma(t)$ be strictly monotone increasing and continuous in a finite interval $\alpha \leq t \leq \beta$, further let $g(x)$ have the same properties on the corresponding segment $g(\alpha) = a \leq x \leq b = g(\beta)$. Consider a function $f(x)$ whose (LS) -integral with respect to $g(x)$ exists in $[a, b]$.

Then we have

$$\int_a^b f(x) dg(x) = \int_{\alpha}^{\beta} f[\gamma(t)] dg[\gamma(t)], \quad (8.3-11)$$

and especially for $g(x) = x$ and an absolutely continuous $\gamma(t)$ the integral relation:

$$\int_a^b f(x) dx = \int_{\alpha}^{\beta} f[\gamma(t)] \gamma'(t) dt. \quad (8.3-12)$$

P r o o f. Suppose that f, g, γ satisfy the conditions of the theorem, and let us denote the inverse of $y = g(x)$ by $x = g_{-1}(y)$, that of $x = \gamma(t)$ by $t = \gamma_{-1}(x)$; for the sake of brevity write $g[\gamma(t)] = \lambda(t)$.

Since $g(a) = g[\gamma(\alpha)] = \lambda(\alpha)$ and $g(b) = g[\gamma(\beta)] = \lambda(\beta)$, furthermore by $g(\gamma[\lambda_{-1}(y)]) = \lambda[\lambda_{-1}(y)] = y$, we see that

$$g_{-1}(y) = \gamma[\lambda_{-1}(y)];$$

thus applying 8.3.2 twice, we obtain:

$$\begin{aligned} \int_a^b f(x) dg(x) &= \int_{g(a)}^{g(b)} f[g_{-1}(y)] dy = \int_{\lambda(\alpha)}^{\lambda(\beta)} f(\gamma[\lambda_{-1}(y)]) dy = \\ &= \int_{\alpha}^{\beta} f[\gamma(t)] d\lambda(t). \end{aligned}$$

This is the formula (8.3-11).

If $g(x)$ and $\gamma(t)$ is even absolutely continuous in $[\alpha, \beta]$, then 8.3.3 implies

$$\int_{\alpha}^{\beta} f[\gamma(t)] d\lambda(t) = \int_{\alpha}^{\beta} f[\gamma(t)] d\gamma(t) = \int_{\alpha}^{\beta} f[\gamma(t)] \gamma'(t) dt;$$

therefore (8.3-11) indeed yields (8.3-12).

* * *

8.35 The properties of (LS) -integrals follow immediately from the related features of abstract Lebesgue integrals; their treatment can be compared at the same time with the theory of (RS) -integrals. (Cf. 5.7-5.8.) Anyway, let us mention here that 1. in the theory of (LS) -integrals often the words “ g -null set”, “ g -equivalent” etc. are used instead of the terms “ μ_g -null set”, “ μ_g -equivalent” and so on, i.e., one refers merely to the integrator function; 2. since by 4.5.1 any function of bounded variation can be expressed as a difference of monotone increasing functions, the notion of (LS) -integral can also be extended at once to the case of *integrators of bounded variation*; 3. there is no objection even to introducing *higher dimensional (LS) -integrals* (i.e., generalized Radon integrals), in which certain additive interval functions of several dimensions take the part of monotone increasing integrators or those of bounded variation.

The wide-ranging utilization of *Lebesgue-Stieltjes integrals* is based mainly upon the fact that the concept of (LS) -integral assures a unified and concise treatment for large classes of sums and infinite series as well as of ordinary integrals.

In particular, we have to emphasize the close connection between the Lebesgue-Stieltjes measure and the notion of *distribution function* which is of central importance in probability theory: the latter is namely [like $g(x)$] by its definition such a single-valued real function $F(x)$ which is in $(-\infty, \infty)$ monotone increasing, left-hand continuous, furthermore tends to 0 or 1 as $x \rightarrow -\infty$ and $x \rightarrow +\infty$, respectively. Thus it becomes possible to express some *characterizing data of random variables* (expected value, variance, moments) as (LS) -integrals formed from the pertaining distribution functions, and thereby to obtain such representations of the

parameters at issue which hold for the discrete and continuous cases alike. The concept of the so-called *characteristic functions of a variate* (random variable) is a noteworthy special case, which presents at the same time an example of a Fourier-Stieltjes transform. (Cf. 14.62.)

8.4 Measures and integrals on product spaces

8.41 The elementary fact that the Euclidean ν -space E_ν can be considered as the Cartesian product of ν number lines (cf. 1.12), and correspondingly we mean by the measure of a ν -dimensional interval the product of the lengths of the pertaining ν linear intervals, suggests to us the problem of defining a measure concept on set structures arising by "Cartesian multiplication" of arbitrary measure spaces. We can expect that the study of this question will make it possible also to find such a connection between certain pertaining abstract Lebesgue integrals which is a counterpart of the reduction of higher dimensional (R)-integrals to one-dimensional ones.

Let us denote the σ -algebra constituted by the ν -dimensional sets of finite or infinite (L)-measure by $\tilde{\mathcal{S}}_L^{(\nu)}$ and the ν -dimensional Lebesgue measure of a set $S \in \tilde{\mathcal{S}}_L^{(\nu)}$ by $\mu_L^{(\nu)}(S)$. In what follows we shall essentially utilize the chain of ideas which leads from ν one-dimensional measure spaces $(E_1, \tilde{\mathcal{S}}_L^{(1)}, \mu_L^{(1)})$ to the measure space $(E_\nu, \tilde{\mathcal{S}}_L^{(\nu)}, \mu_L^{(\nu)})$. The theme is very simple: we take on each number line E_1 a special system of sets, namely the finite segments, whose measure concept (length) is defined directly and can be transferred (by ordinary multiplication) to the corresponding sets in E_ν : the ν -dimensional intervals. This "*base measure*" will thereafter be gradually extended by the Lebesgue process to all sets $\in \tilde{\mathcal{S}}_L^{(\nu)}$. Hence it is also reasonable to call $(E_\nu, \tilde{\mathcal{S}}_L^{(\nu)}, \mu_L^{(\nu)})$ the *abstract (or "independent") product* of the ν measure spaces mentioned above.¹

8.42 Let now ν complete measure spaces $(X_1, \tilde{\mathcal{S}}_1, \mu_1), (X_2, \tilde{\mathcal{S}}_2, \mu_2), \dots, (X_\nu, \tilde{\mathcal{S}}_\nu, \mu_\nu)$ be given, and let us select from every σ -algebra $\tilde{\mathcal{S}}_k$ such a "substructure" (consisting of sets $\in X_k$) which takes over the role of one-dimensional finite intervals. Since the latter obviously do not form a Boolean ring, their totality not being closed with respect to the union operation, it is advisable to introduce a new (more general) notion.

A system $\mathcal{P}$ of sets belonging to a basic set $X \neq \emptyset$ is called a *semi-ring*, if it contains $\emptyset$, is closed with respect to the meet operation, moreover has the following property: after being prescribed any set $S \in \mathcal{P}$ and a subset $s \in S$, the difference $S - s$ is representable as the union of finitely many sets

¹ The attribute "independent" refers to a certain interpretation in probability theory, relating to so-called stochastically independent events.

$\in \mathcal{P}$, disjoint with s and with one another. It is manifest that every Boolean ring is a semi-ring at the same time, but the converse is not true.

For instance, the totality of left closed and right open finite intervals on the line, supplemented by $\bigcirc$, form such a semi-ring $\mathcal{P}_I^{(1)}$ which can trivially be made a Boolean ring, if we apply finitely many times to the elements of $\mathcal{P}_I^{(1)}$ the union operation and attach to $\mathcal{P}_I^{(1)}$ all the intervals thus obtained. Like 1.5.1, it is easy to see:

8.4.1 *Every semi-ring $\mathcal{P}$ can be widened to a Boolean ring in such a way that we supplement it by the union of any finite number of its elements. The resulting system $\mathcal{R}(\mathcal{P})$ is the narrowest (minimal) Boolean ring which contains $\mathcal{P}$.*

If $\mathcal{P}_I^{(1)}$ denotes the above special semi-ring, then the ν -times iterated Cartesian product of this system by itself which we denote by $\mathcal{P}_I^{(\nu)}$, is again a semi-ring, namely it consists of all the finite ν -dimensional intervals of the type $a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2, \dots, a_\nu \leq x_\nu \leq b_\nu$ and of $\bigcirc$. The importance of the notion of semi-ring lies in the fact that this peculiar "closure with respect to Cartesian multiplication" holds in the case of arbitrary semi-rings as well.

8.4.2 *The Cartesian product of finitely many semi-rings is always a semi-ring.*¹

P r o o f. Let $\mathcal{P}_1, \mathcal{P}_2$ be two semi-rings and let us put $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2$.

1° If $S \in \mathcal{P}$, $T \in \mathcal{P}$, then both sets can be written as Cartesian products of certain elements of $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively:

$$S = S_1 \times S_2 \quad (S_1 \in \mathcal{P}_1, S_2 \in \mathcal{P}_2), \quad T = T_1 \times T_2 \quad (T_1 \in \mathcal{P}_1, T_2 \in \mathcal{P}_2)$$

Consequently also

$$S \cap T = (S_1 \cap T_1) \times (T_1 \cap T_2),$$

and the meets on the right belonging by the definition of semi-ring to $\mathcal{P}_1$ and $\mathcal{P}_2$, we obtain

$$S \cap T \in \mathcal{P}. \quad (8.4-1)$$

2° Consider furthermore a set $S \in \mathcal{P}$ and a subset $s \in \mathcal{P}$ of it. If again $S_1 \in \mathcal{P}_1$, $S_2 \in \mathcal{P}_2$ with $S_1 \times S_2 = S$, then there are such sets $s_1 \subseteq S_1, s_2 \subseteq S_2$ for which $s_1 \times s_2 = s$.

But $\mathcal{P}_1$ and $\mathcal{P}_2$ being semi-rings, S_1 and S_2 allow also decomposition of the form:

$$S_1 = s_1 \cup s_1^{(1)} \cup \dots \cup s_1^{(m)}, \quad S_2 = s_2 \cup s_2^{(1)} \cup \dots \cup s_2^{(n)},$$

¹ The pendant of this assertion does, of course, not hold for Boolean rings.

where $s_1, s_1^{(1)}, \dots, s_1^{(m)}$ and $s_2, s_2^{(1)}, \dots, s_2^{(n)}$ are pairwise disjoint subsets of $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively.

Hence follows:

$$S = S_1 \times S_2 = (s_1 \times s_2) \cup (s_1 \times s_2^{(1)}) \cup \dots \cup (s_1 \times s_2^{(n)}) \cup \\ \cup (s_1^{(1)} \times s_2) \cup (s_1^{(1)} \times s_2^{(1)}) \cup \dots \cup (s_1^{(1)} \times s_2^{(n)}) \cup \\ \dots \dots \dots \cup (s_1^{(m)} \times s_2) \cup (s_1^{(m)} \times s_2^{(1)}) \cup \dots \cup (s_1^{(m)} \times s_2^{(n)}). \quad (8.4-2)$$

Our assumptions imply that all the Cartesian products $s_1^{(i)} \times s_2^{(j)}$ ($i = 0, 1, \dots, m; j = 0, 1, \dots, n; s_1^{(0)} = s_1, s_2^{(0)} = s_2$) belong to $\mathcal{P}$ and are pairwise disjoint sets.

Therefore our assertion can directly be verified by (8.4-1) and (8.4-2) for the case of two semi-rings, while otherwise we get through easily by induction.

* * *

8.43 By way of the concept of semi-ring we possess the general set structure taking over the role of systems of finite intervals. Thus next we have to introduce the analogue of the interval measure: we call *base measure* and denote by $\mu^0(S)$ any such non-negative and finitely additive set function, which is defined on a semi-ring.

Let us select from any one of the above σ -algebras $\tilde{\mathcal{S}}_k (k = 1, 2, \dots, \nu)$ [associated with the prescribed measure spaces $(X_k, \tilde{\mathcal{S}}_k, \mu_k)$] a semi-ring $\mathcal{P}_k$, and consider the base measure $\mu_k^0(S_k)$ ($S_k \in \mathcal{P}_k$) obtained by the specification of μ_k to the sets of $\mathcal{P}_k$. By 8.4.2 it is natural that we attach to the pertinent subsets of

$$X = X_1 \times X_2 \times \dots \times X_\nu, \quad (8.4-3)$$

further to the semi-ring

$$\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2 \times \dots \times \mathcal{P}_\nu, \quad (8.4-4)$$

the product

$$\mu^0(S) = \mu_1^0(S_1)\mu_2^0(S_2) \dots \mu_\nu^0(S_\nu) \quad (8.4-5)$$

as a base measure; the finite additivity of (8.4-5) can be shown easily.¹

¹ For a Cartesian product of more than two factors, holds the *associative law*. Accordingly e.g. $X_1 \times X_2 \times X_3 = (X_1 \times X_2) \times X_3 = X_1 \times (X_2 \times X_3)$, because any element (x_1, x_2, x_3) of the first product is identified with the corresponding elements $((x_1, x_2), x_3)$ and $(x_1, (x_2, x_3))$ of the latter.

Using 8.4.1, the definition of $\mu^0(S)$ can be extended at once to the narrowest Boolean ring $\mathcal{R}(\mathcal{P})$ containing $\mathcal{P}$, so that we get a *premeasure* $\mu^*(S)$ [$S \in \mathcal{R}(\mathcal{P})$]; moreover, the Lebesgue continuation of the latter yields such a totally additive *measure* $\mu(S)$, whose domain of definition is the σ -algebra $\tilde{\mathcal{S}}(\mathcal{P})$ arising from the minimal widening of $\mathcal{R}(\mathcal{P})$. (Cf. 1.5.1, 8.15.)

Summing up, we establish the following theorem:

8.4.3 *In case of any given system of complete and σ -finite measure spaces $(X_k, \tilde{\mathcal{S}}_k, \mu_k)$ ($k = 1, 2, \dots, \nu$), there exists such an abstract measure $\mu(S)$, determined uniquely, which is induced by the base measure (8.4-5) pertaining to the semi-ring (8.4-4) on the σ -algebra $\tilde{\mathcal{S}}(\mathcal{P})$.*

We add some pertinent names: the measure space belonging to the basic set (8.4-3) as well as to $\tilde{\mathcal{S}}(\mathcal{P})$ and $\mu(S)$ is called the *abstract (or independent) product* of $(X_1, \tilde{\mathcal{S}}_1, \mu_1), \dots, (X_\nu, \tilde{\mathcal{S}}_\nu, \mu_\nu)$, briefly the *product space* formed from these measure spaces. The abstract measure $\mu(S)$ is also said to be the *abstract (independent) product* of $\mu_1, \mu_2, \dots, \mu_\nu$, in short, *product measure*, and it is usually denoted by the symbol:

$$\mu = \mu_1 \otimes \mu_2 \otimes \dots \otimes \mu_\nu. \quad (8.4-6)$$

Finally, on the basis of the measure concept (8.4-6) and of the considerations in Section 8.2, we can define without difficulty the *abstract Lebesgue integral of a function* $f(x_1, x_2, \dots, x_\nu)$ ($x_1 \in X_1, x_2 \in X_2, \dots, x_\nu \in X_\nu$) over a set $S \in \tilde{\mathcal{S}}(\mathcal{P})$, provided that f is μ -integrable with respect to the product measure at issue, in a sense similar to the case of (8.2-4). This concept is called briefly *(L)-integral relating to product space* and its symbol is

$$\int_{(S)} f(x_1, x_2, \dots, x_\nu) d\mu. \quad (8.4-7)$$

As regards the calculation of (8.4-7), observe that it is about a strong generalization of the ν -dimensional *(L)-integral*. Thus there is a pertinent transformation rule which will be discussed for $\nu = 2$ in the forthcoming section.

We note still that the ideas about product space and related *(L)-integral* can be extended to the case of *infinitely many measure spaces*, too. When it comes to that, we always need the restriction $\mu(X_1 \times X_2 \times \dots) = 1$, originating from applications in probability theory.

8.5 Fubini's theorem for product spaces; connections with probability theory

8.51 In order to extend suitably Fubini's theorem 7.7.1, we have to premise a lemma, which also has a certain own relevancy concerning Lebesgue continuation of measures (cf. 8.4.1).

8.5.1 Let $\mu(S)$ be a measure attached to a complete and σ -finite measure space $(X, \tilde{S}, \mu)$; let $\mathcal{P}$ denote a semi-ring consisting of subsets of $\tilde{S}$.

Then to any given set $S \in \tilde{S}$ we can find another set T such that

$$T = \bigcap_{n=1}^{\infty} T_n, \quad T_1 \supseteq T_2 \supseteq \dots \supseteq T_n \supseteq \dots,$$

$$T_n = \bigcup_{r=1}^{\infty} T_{nr}, \quad T_{n1} \subseteq T_{n2} \subseteq \dots \subseteq T_{nr} \subseteq \dots,$$

where $T_{nr} \in \mathcal{R}(\mathcal{P})$, furthermore

$$\mu(S) = \mu(T) \quad (S \subseteq T) \quad (8.5-1)$$

holds.

P r o o f. According to the definition of μ -measure, for any fixed natural number n there are countably many sets $s_{np} \in \mathcal{P}$, such that $Z_n = \bigcup_p s_{np} \supseteq S$ and the inequality

$$\mu(Z_n) < \mu(S) + \frac{1}{n} \quad (8.5-2)$$

is fulfilled.

If $T_n = \bigcap_{i=1}^n Z_i$, then T_n can plainly be written as the union of certain sets $\tau_{nq} \in \mathcal{P}$ ($q = 1, 2, \dots$). Thus the choice $T_{nr} = \bigcup_{q=1}^r \tau_{nq} \in \mathcal{R}(\mathcal{P})$ yields at the same time $T_n = \bigcup_{r=1}^{\infty} T_{nr}$, and $\{T_{nr}\}$ is an increasing, $\{T_n\}$ is a decreasing sequence of sets. Considering that for $T = \bigcap_{n=1}^{\infty} T_n$ we have also $T = \bigcap_{n=1}^{\infty} Z_n$, and it is $S \subseteq Z_n$ by the above, the remaining part of the assertion can be obtained from (8.5-2).

* * *

A further subsidiary proposition means in the most elementary case, that the (L) -measure of a planar set can be written as the integral of some linear (L) -measure.

8.5.2 Let $(X_1, \tilde{S}_1, \mu_1)$ and $(X_2, \tilde{S}_2, \mu_2)$ be two complete, σ -finite measure spaces, and let $\mu = \mu_1 \otimes \mu_2$ denote the corresponding product measure.

Then we have for any μ -measurable set $S \in X_1 \times X_2$ the representations:

$$\mu(S) = \int_{(X_1)} \mu_2(S_{x_1}) d\mu_1 = \int_{X_2} \mu_1(S_{x_2}) d\mu_2$$

$$(x_1 \in X_1, x_2 \in X_2), \quad (8.5-3)$$

where

$$S_{x_1} = \{x_2 : (x_1, x_2) \in S\}, \quad S_{x_2} = \{x_1 : (x_1, x_2) \in S\}.^1 \quad (8.5-4)$$

P r o o f. By reasons of symmetry, it suffices to show e.g. that

$$\mu(S) = \int_{(X_1)} M_S(x_1) d\mu_1, \quad (8.5-5)$$

where $M_S(x_1) = \mu_2(S_{x_1})$.

By 8.4.2 and 8.4.3, the μ -measure is the unique extension of the base measure belonging to the Cartesian product $\mathcal{P}$ of certain semi-rings $\mathcal{P}_1$ and $\mathcal{P}_2$, the latter being selected from the σ -algebras $\tilde{S}_1$ and $\tilde{S}_2$, respectively; we know also that the elements of $\mathcal{P}$ can be written in the form $S_{\alpha_1} \times S_{\alpha_2}$ ($\alpha_1 \in X_1, \alpha_2 \in X_2$).

If $S \in \mathcal{P}$, then $M_S(x_1) = \mu_2(S_{\alpha_1})$ or 0 according to $x_1 \in S_{\alpha_2}$ and $x_1 \notin S_{\alpha_2}$, respectively; consequently:

$$\int_{(X_1)} M_S(x_1) d\mu_1 = \mu_2(S_{\alpha_1}) \mu_1(S_{\alpha_2}) = \mu(S_{\alpha_1} \times S_{\alpha_2}) = \mu(S).$$

It is easy to see that (8.5-5) holds for the sets belonging to $\mathcal{R}(\mathcal{P})$ as well, since each of these is the union of finitely many, pairwise disjoint sets $\in \mathcal{P}$. (Cf. 8.4.1.)

In case of any other μ -measurable set S , let us apply lemma 8.5.1, setting out from the following idea: the validity of (8.5-5) being established

¹ If X_1 and X_2 signify the axes of a planar right-handed rectangular Cartesian co-ordinate system, then S_{x_1} means the orthogonal projection of points with abscissa x_1 of the planar set S on the X_2 -axis, while S_{x_2} is the orthogonal projection of points with ordinate x_2 of S on the X_1 -axis.

for the sets $T_{nr} \in \mathcal{R}(\mathcal{P})$, this formula must also hold for any T_n , moreover for T too, owing to the pendant of B. Levi's theorem for μ -integrals. (Cf. 7.5.3.) Namely

$$M_{T_n}(x_1) = \lim_{r \rightarrow \infty} M_{T_{nr}}(x_1), \quad M_{T_{n1}} \leq M_{T_{n2}} \leq \dots$$

and

$$M_T(x_1) = \lim_{n \rightarrow \infty} M_{T_n}(x_1), \quad M_{T_1} \geq M_{T_2} \geq \dots$$

Now, assuming $\mu(S) = 0$ we conclude $\mu(T) = 0$ and μ_1 -almost everywhere $M_S(x_1) = 0$; furthermore by $S \subseteq T$ it follows $M_S(x_1) = 0$ and $\int M_S(x_1) d\mu_1 = 0 = \mu(S)$. Therefore (8.5-5) holds for sets S having the

(X_1) μ -measure zero. If $\mu(S) > 0$, let us put $S = T - V$ where V is a μ -null set because of (8.5-1). Since we know by the above that (8.5-5) holds for $S = T$ and have just shown that it is valid for every μ -null set, the assertion

$$\mu(T) = \int_{(X_1)} M_T(x_1) d\mu_1, \quad \int_{(X_1)} M_V(x_1) d\mu_1 = 0$$

simply by subtraction.

* * *

Let us mention the important particular case where X_2 is a number line, μ_2 is the linear (L)-measure, further S consists of those pairs (x_1, x_2) , for which x_1 belongs to a μ_1 -measurable set s_1 , and $x_2 \in [0, f(x_1)]$, f denoting a given function. Under such circumstances $\mu_2(S_{x_1})$ is equal to $f(x_1)$ for $x_1 \in s_1$ and vanishes otherwise, so that we get

$$\mu(S) = \int_{(S_1)} f(x_1) d\mu_1 \quad [f(x_1) \geq 0]. \quad (8.5-6)$$

This can be regarded as a kind of extension of the geometric interpretation of (L)-integral.

8.52 By the foregoing we can deduce the rule at issue on the transformation of an (L)-integral over product space into a "double" integral in an extremely short way.

8.5.3 (FUBINI'S THEOREM FOR ABSTRACT LEBESGUE INTEGRALS) Let $(X_1, \tilde{S}_1, \mu_1)$, $(X_2, \tilde{S}_2, \mu_2)$ be complete and σ -finite measure spaces, denote by $\mu = \mu_1 \otimes \mu_2$ the pertaining product measure.

If $f(x_1, x_2)$ is μ -integrable on a μ -measurable set $S = S_1 \times S_2$ ($S_1 \in \tilde{S}_1$, $S_2 \in \tilde{S}_2$), then from the integrals

$$F_{x_1} = \int_{(S_{x_1})} f(x_1, x_2) d\mu_2, \quad F_{x_2} = \int_{(S_{x_2})} f(x_1, x_2) d\mu_1 \quad (x_1 \in X_1, x_2 \in X_2) \quad (8.5-7)$$

the first one is μ_1 -integrable on X_1 , the second one is μ_2 -integrable on X_2 , and we have the formula

$$\int_{(S)} f(x_1, x_2) d\mu = \int_{(X_1)} F_{x_1} d\mu_1 = \int_{(X_2)} F_{x_2} d\mu_2. \quad (8.5-8)$$

P r o o f. Consider first the case $f(x_1, x_2) \geq 0$.

Let us start from the Cartesian product $U = X_1 \times X_2 \times E_1$ and the corresponding product measure $\mu_U = \mu_1 \otimes \mu_2 \otimes \mu_L^{(1)} = \mu \otimes \mu_L^{(1)}$, where $\mu_L^{(1)}$ is the linear (L)-measure.

If we denote by s the set of triplets (x_1, x_2, y) for which $x_1 \in S_{x_2}$, $x_2 \in S_{x_1}$ and $y \in [0, f(x_1, x_2)]$, then by 8.5.2 we can write [cf. (8.5-6)]:

$$\mu_U(s) = \int_{(S)} f(x_1, x_2) d\mu. \quad (8.5-9)$$

On the other hand, using again 8.5.2 and applying the notations $\mu_0 = \mu_2 \otimes \mu_L^{(1)}$, $s_{x_1} = \{(x_2, y) : (x_1, x_2, y) \in s\}$ we get:

$$\mu_0(s_{x_1}) = \int_{(S_{x_1})} f(x_1, x_2) d\mu_2 \quad \text{and} \quad \mu_U(s) = \int_{X_1} \mu_0(s_{x_1}) d\mu_1. \quad (8.5-10)$$

The comparison of (8.5-9) and (8.5-10) furnishes the integral relation

$$\int_{(S)} f(x_1, x_2) d\mu = \int_{(X_1)} F_{x_1} d\mu_1,$$

i.e., the first part of our assertion. The second part of (8.5-8) can be obtained simply by changing the notations.

The extension of the result to the case of an integrand with arbitrary sign can be carried out "mechanically" through the decomposition of $f(x_1, x_2)$ into its positive and negative parts.

* * *

It is a remarkable fact that the existence of the double integrals in (8.5-8) does not imply the existence of $\int_{(S)} f(x_1, x_2) d\mu$.

For example, let $X_1 = X_2 = [-1, 1]$ and take

$$f(x_1, x_2) = \begin{cases} \frac{x_1 x_2}{(x_1^2 + x_2^2)^2} & \text{for } (x_1, x_2) \neq (0, 0), \\ 0 & \text{for } (x_1, x_2) = (0, 0). \end{cases} \quad (8.5-11)$$

Then it is obvious that

$$\int_{-1}^1 dx_1 \int_{-1}^1 f(x_1, x_2) dx_2 = \int_{-1}^1 dx_2 \int_{-1}^1 f(x_1, x_2) dx_1 = 0,$$

because f is an odd function both of x_1 and of x_2 . In turn, the (L) -integral of the function (8.5-11) over the square $-1 \leq x_1 \leq 1$, $-1 \leq x_2 \leq 1$ does not exist, since an easy transformation of variables yields the estimation

$$\int_{-1}^1 \int_{-1}^1 |f(x_1, x_2)| dx_1 dx_2 \geq \int_{\varepsilon}^1 d\rho \int_{\varepsilon}^{2\pi} \frac{\sin \vartheta \cos \vartheta}{\rho} d\vartheta = 2 \int_{\varepsilon}^1 \frac{d\rho}{\rho} \quad (0 < \varepsilon < 1)$$

where the last term tends to infinity as $\varepsilon \rightarrow +0$.

We have to add, however, that *the validity conditions of (8.5-8) become essentially simpler, if $f(x_1, x_2) \geq 0$* . In this case namely, as it can be shown by a detailed discussion, the finiteness of the second or third term in (8.5-8) implies the existence of the two other terms, moreover also the equality of all three integrals to one another.

The generalization of Fubini's theorem 8.5.1 for any integral of the type (8.4-7) can be realized in a natural manner: we may integrate successively in arbitrary order with respect to the individual abstract measures attached to the corresponding measure spaces, just like in the classical ν -dimensional case with respect to every single independent variable.

8.53 Finally, we want to give a brief outline on the utilization of the main concepts of this chapter in the axiomatic *theory of probability*, due to A.N. KOLMOGOROV (1933). This building up has been generally adopted by specialists and appliers alike during the last decades.¹

The starting-point of the "calculus" of probabilities: the *concept of event* can be characterized by the notion of set, and thus the *algebra of events* is nothing else but a Boolean algebra in the simplest cases and a σ -algebra in full generality. In the framework of the algebra of events, the role of $\bigcirc$ is played by the *impossible event*, the counterpart of a basic set X is the *certain event*; *opposite events* are associated with complementary sets, *mutually exclusive events* with disjoint sets, and so on. The *mathematical probability* can be considered as an (abstract) measure, i.e., a non-negative and totally additive set function — anyway with the restriction that its domain of values is the interval $[0, 1]$ and that the probability of impossible events shall be 0, that of certain events the unity.

After these, it is understandable that 1. the analytic description of so-called *random experiments* (observations) is carried out by means of measure spaces; 2. the concept of *random variable* signifies a measurable function in wider sense; 3. the *characterizing parameters* of random variables — such as *expected value*, *variance*, *moments* — can all be written in the form of abstract Lebesgue integrals (cf. 8.35). Moreover, the analytic model for *iteration countably many times of stochastically independent experiments* is furnished by the theory of product spaces, by which of course the use of Fubini's theorem is often necessitated, too.

¹ Cf. e.g. Kolmogorov [3]–[4]; Feller [1].

This is the background which has made modern function theoretical apparatus indispensable for probability theory and for the mathematical statistics closely related to it since the pioneering activity of KOLMOGOROV.

10-5.3 Finally, we want to give a brief outline on the utilization of the main concepts of this chapter in the axiomatic theory of probability, due to A. N. KOLMOGOROV (1933). This building has been generally adopted by specialists and appears alike during the last decades.

The starting-point of the "calculus" of probabilities: the concept of event can be characterized by the notion of set, and thus the algebra of events is nothing else but a Boolean algebra in the simplest case and a σ -algebra in full generality. In the framework of the algebra of events, the role of Ω is played by the impossible event, the counterpart of a point set Λ is the certain event, opposite events are associated with complementary sets, mutually exclusive events with disjoint sets, and so on. The mathematical probability can be considered as an (abstract) measure, i.e., a non-negative and locally additive set function—anyway with the restriction that its domain of values is the interval $[0, 1]$ and that the probability of impossible events shall be 0, that of certain events the unity.

After these, it is understandable that 1. the analytic description of so-called random experiments (observations) is carried out by means of measure spaces; 2. the concept of random variable signifies a measurable function in wider sense; 3. the characteristic parameters of random variables—such as expected value, variance, moments—can all be written in the form of abstract Lebesgue integrals (cf. 8.35). Moreover, the analytic model for iteration countably many times of stochastically independent experiments is furnished by the theory of product spaces, by which of course the method of Kolmogorov's theorem is often necessitated, too.

11-5.3 Finally, we want to give a brief outline on the utilization of the main concepts of this chapter in the axiomatic theory of probability, due to A. N. KOLMOGOROV (1933). This building has been generally adopted by specialists and appears alike during the last decades.

The starting-point of the "calculus" of probabilities: the concept of event can be characterized by the notion of set, and thus the algebra of events is nothing else but a Boolean algebra in the simplest case and a σ -algebra in full generality. In the framework of the algebra of events, the role of Ω is played by the impossible event, the counterpart of a point set Λ is the certain event, opposite events are associated with complementary sets, mutually exclusive events with disjoint sets, and so on. The mathematical probability can be considered as an (abstract) measure, i.e., a non-negative and locally additive set function—anyway with the restriction that its domain of values is the interval $[0, 1]$ and that the probability of impossible events shall be 0, that of certain events the unity.

After these, it is understandable that 1. the analytic description of so-called random experiments (observations) is carried out by means of measure spaces; 2. the concept of random variable signifies a measurable function in wider sense; 3. the characteristic parameters of random variables—such as expected value, variance, moments—can all be written in the form of abstract Lebesgue integrals (cf. 8.35). Moreover, the analytic model for iteration countably many times of stochastically independent experiments is furnished by the theory of product spaces, by which of course the method of Kolmogorov's theorem is often necessitated, too.

FUNCTION SPACES AND ORTHOGONAL EXPANSIONS

9.1 Orthogonal series and functional analysis

9.11 The part of analysis dealing with so-called orthogonal series may be considered as one of the most important applications of Lebesgue's modern integration theory.

The most classic example of this series type (which was introduced by L. EULER more than two hundred years ago) is the trigonometric series with real coefficients:

$$A_0 + (A_1 \cos x + B_1 \sin x) + (A_2 \cos 2x + B_2 \sin 2x) + \dots \quad (9.1-1)$$

occurring in real analysis and physics almost as frequently as the power series.¹ The applications are attached to the elementary property of the "base system"

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \dots \quad (9.1-2)$$

that the integral over a given interval of length 2π of any two distinct terms in (9.1-2) is equal to 0.

More generally, some system of functions $\varphi_1(x), \varphi_2(x), \dots$ is said to be orthogonal on a segment (a, b) , if we have

$$\int_a^b \varphi_m(x) \varphi_n(x) dx = 0 \quad (9.1-3)$$

for every pair of indices $m \neq n$.² If $\{\varphi_n(x)\}$ is an orthogonal system, so any linear combination with constant coefficients:

$$C_1 \varphi_1(x) + C_2 \varphi_2(x) + \dots \quad (9.1-4)$$

is called orthogonal series.

¹ It is well-known from complex analysis that the decomposition of any complex power series into real and imaginary parts yields two series of the form (9.1-1).

² It is advisable to take also the restriction that (9.1-3) does not hold for $m = n$, i.e., that $\int_a^b \varphi_n(x)^2 dx \neq 0$ ($n = 1, 2, \dots$).

FOUNDATIONS OF FUNCTIONAL ANALYSIS. THE SPACES l^2 AND L^2

9.1 Orthogonal series and functional analysis

9.11 The part of analysis dealing with so-called orthogonal series may be considered as one of the most important applications of Lebesgue's modern integration theory.

The most classic example of this series type (which was introduced by L. EULER more than two hundred years ago) is the trigonometric series with real coefficients:

$$A_0 + (A_1 \cos x + B_1 \sin x) + (A_2 \cos 2x + B_2 \sin 2x) + \dots, \quad (9.1-1)$$

occurring in real analysis and physics almost as frequently as the power series.¹ The applications are attached to the elementary property of the "base system"

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \quad (9.1-2)$$

that the integral over a given interval of length 2π of any two distinct terms in (9.1-2) is equal to 0.

More generally: some system of functions $\varphi_1(x), \varphi_2(x), \dots$ is said to be orthogonal on a segment (a, b) , if we have

$$\int_a^b \varphi_m(x) \varphi_n(x) dx = 0 \quad (9.1-3)$$

for every pair of indices $m \neq n$.² If $\{\varphi_n(x)\}$ is an orthogonal system, so any linear combination with constant coefficients:

$$C_1 \varphi_1(x) + C_2 \varphi_2(x) + \dots \quad (9.1-4)$$

is called orthogonal series.

¹ It is well-known from complex analysis that the decomposition of any complex power series into real and imaginary parts yields two series of the form (9.1-1).

² It is advisable to take also the restriction that (9.1-3) does not hold for $m = n$, i.e., that $\int_a^b \varphi_n(x)^2 dx \neq 0$ ($n = 1, 2, \dots$).

9.12 It is well known, that orthogonality means perpendicularity in geometry. Thus it is natural to ask: how is it justifiable that one can transfer this term, usually applied only to vectors (or straight lines), to functions in the analysis? For the moment, the answer cannot be other than: the left-hand integral in (9.1-3) is analogous to the coordinate form of the inner (scalar) product of two vectors; and the vanishing of the latter is indeed a characteristic condition of perpendicularity.

The deeper motivation, the cogent establishment of the theory of orthogonal systems and series presupposes transferring the whole geometric terminology of Euclidean spaces to certain sets of functions. We shall see that this not only facilitates the formulation of results as well as their systematization, but also discloses and so to say makes vividly descriptive a number of hidden facts. Let us emphasize: when speaking about orthogonality in the (9.1-3) sense, *Lebesgue's concept of integral will always be used*; if we restricted ourselves to Riemann integrals, some relevant statements of the theory would not keep their validity.

9.13 The development of science led at the beginning of this century to the recognition, that a lot of results relating to different special orthogonal series can be incorporated in such a general (unified) theory which permits significant applications connected with differential and integral equations, probability theory and mathematical statistics, approximations and numerical analysis, linear algebra, quantum theory, etc. Following these considerations, researchers were soon led to such fundamental ideas as *function spaces* and other *abstract spaces*, respectively, furthermore to the foundation of a new mathematical discipline: *functional analysis*. The latter deals — as mentioned already in Section 3.1 — both with properties of functionals and with investigation of other operators; and in its building up, the *theory of orthogonal series* similarly takes a prominent part, as e.g. planar and spatial *analytic geometry* are indispensable to the analysis of real functions of two and three variables.

All the foregoing now prompts us to set forth in detail that natural chain of thought by which we can move from the Euclidean spaces regarded as sets of points or vectors to the most important pertinent abstract spaces, those due to D. HILBERT. Thereafter we want to expose, in its many aspects, the “geometry” of *Hilbert's function space* which means essentially the same as to study expansions in terms of various orthogonal systems.¹

¹ Cf. e.g. Vitali [4].

9.2 Vectorial interpretation of the Euclidean space E_ν

9.21 We know that there is a one-to-one correspondence between the points of line, plane, space and the vectors attached to them. Therefore E_1 , E_2 , E_3 can be considered not only as point sets but also as sets of vectors with a common initial point at the origin — an interpretation often advantageous in analytic geometry.

Turning to an arbitrary Euclidean space E_ν , any element (i.e., a ν -tuple of real numbers) may be called a ν -dimensional vector, if the basic operations in E_ν are introduced like in the case of ordinary vectors. Thus we define the *sum* of two elements $X = (x_1, x_2, \dots, x_\nu)$, $Y = (y_1, y_2, \dots, y_\nu)$ and the *product* of the former by a real number α as follows:

$$\left. \begin{aligned} X + Y &= (x_1 + y_1, x_2 + y_2, \dots, x_\nu + y_\nu), \\ \alpha X &= (\alpha x_1, \alpha x_2, \dots, \alpha x_\nu), \end{aligned} \right\} \quad (9.2-1)$$

by which the definition of *difference* can be obtained at once in the form:

$$X - Y = X + (-1)Y.$$

The origin $O = (0, 0, \dots, 0)$ is also called the *zero vector* of E_ν ; for every

$X \in E_\nu$ we have $X - \overset{\nu}{X} = O$ and $X \cdot O = O \cdot X = O$.

The *inner* (scalar) *product* of two elements $X = (x_1, x_2, \dots, x_\nu)$, $Y = (y_1, y_2, \dots, y_\nu)$ is defined by

$$(X, Y) = x_1 y_1 + x_2 y_2 + \dots + x_\nu y_\nu, \quad (9.2-2)$$

because this turns with $\nu = 3$ into the elementary representation of scalar product in coordinate form.¹ The operations just mentioned obviously possess the properties usual in vector algebra (e.g. distributivity).

In the Euclidean 3-space, (X, Y) becomes the square length of any of the vectors if $X = Y$; this motivates the following definition for the *absolute value* (length, modulus) of $X = (x_1, x_2, \dots, x_\nu)$ [cf. (3.1-6)]:

$$\|X\| = +\sqrt{(X, X)} = (x_1^2 + x_2^2 + \dots + x_\nu^2)^{1/2}. \quad (9.2-3)$$

Let us emphasize that here it is more usual instead of absolute value the denomination *norm*, and we use mostly instead of $|X|$ the notation $\|X\|$. Of course, *unit vector* now means such a ν -dimensional vector whose absolute value (i.e., norm) is 1; example: any ν -tuple containing 1 once and 0 at the

¹ There is no objection to introducing also the *outer* (vector) *product* as a determinant of ν -th order, the application possibilities of this notion being anyway much weaker.

remaining $\nu - 1$ places. By putting in (9.2-3) $X - Y$ for X i.e., $x_k - y_k$ for x_k , we obtain the distance of X, Y as points of E_ν [cf. (2.1-1)]:

$$\|X - Y\| = \left[(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_\nu - y_\nu)^2 \right]^{1/2}, \quad (9.2-4)$$

in accordance with the elementary fact that the absolute value of the difference of two ordinary vectors is equal to the distance of their end points.

9.22 The notion: *inclination angle* of two ν -dimensional vectors can easily be defined by means of the concepts of inner product and norm. For, keeping in view the geometric definition of scalar product, we mean by cosine of the angle of inclination ϑ in case of two vectors $X \neq O, Y \neq O$:

$$\cos \vartheta = \frac{(X, Y)}{\|X\| \|Y\|}. \quad (9.2-5)$$

The existence of an angle $\vartheta \in [0, \pi]$ (defined uniquely) which satisfies (9.2-5) follows from the inequality (2.1-3); hence we can see that the absolute value of the right-hand quotient in (9.2-5) does not exceed 1. In particular, if $\cos \vartheta = 1$, i.e. $(X, Y) = \|X\| \|Y\|$,¹ then we say that X and Y are *unidirectional*; in case of $\cos \vartheta = 0$, i.e., when

$$(X, Y) = 0, \quad (9.2-6)$$

X and Y are called *orthogonal* (perpendicular) to one another.²

(9.2-5) is closely connected also with the notion of *projection* or *component*: since for ordinary vectors the first term of the equation

$$\|X\| \cos \vartheta = \frac{(X, Y)}{\|Y\|} \quad (9.2-7)$$

signifies the projection (with sign) of X relating to Y , the latter is defined by the right-hand ratio in case of ν -dimensional vectors. If the component of X in the direction of Y is denoted briefly by X_Y , we can write:

$$X_Y = \frac{(X, Y)}{\|Y\|} = (X, \|Y\|^{-1}Y) = (X, Y^0), \quad (9.2-8)$$

where $Y^0 = \|Y\|^{-1}Y$ is a unit vector, unidirectional with Y .

¹ It is easy to see, that this is equivalent to the condition $x_k = \lambda y_k$ ($k = 1, 2, \dots, \nu$) where $\lambda \neq 0$ denotes a constant.

² Note that (9.2-6) is trivially fulfilled if any of X, Y equals O ; but orthogonality involves $X \neq O, Y \neq O$ by our assumption.

As a complementary result, we point out: any element of E_ν , distinct from O , becomes a unit vector of the same direction when divided by its norm. This simple operation is the so-called *norming* or *normalization*.

9.23. In classical vector algebra, it is of fundamental importance that any vector can be represented as the linear combination of the unit vectors directed along the coordinate axes; and the occurring coefficients are simply the components of the pertinent vector in the direction of the respective coordinate axes. In case of a right-handed rectangular (Cartesian) coordinate system, the axes are pairwise orthogonal to each other, i.e., they constitute a so-called orthogonal system of vectors; furthermore it is relevant that in the plane at most two, in the space at most three mutually orthogonal directions can be given. By the above, we can immediately introduce the corresponding notions in the space E_ν .

By *orthogonal system* we mean an arbitrary subset of the space whose any two elements are orthogonal to each other in the (9.2-6) sense. An orthogonal system is said to be *normalized*, if it consists of unit vectors; instead of the expression "orthogonal and normalized" the shorter term "*orthonormal*" is in use. (Every orthogonal system can plainly be made orthonormal by norming its elements.) Finally: an orthogonal system of the space is called *complete*, if it cannot be expanded to a wider one, i.e., if there is no element of the space orthogonal to all elements of the system. (In other words: in case of a complete orthogonal system, O is the only element of the space whose inner product by any element of the system equals to zero.)

For instance, consider the ν -dimensional unit vectors

$$U_1 = (1, 0, \dots, 0), U_2 = (0, 1, 0, \dots, 0), \dots, U_\nu = (0, \dots, 0, 1), \quad (9.2-9)$$

mentioned before. Since evidently

$$(U_m, U_n) = \begin{cases} 1 & (m = n), \\ 0 & (m \neq n), \end{cases}$$

this system is *orthonormal*. Let $X = (x_1, x_2, \dots, x_\nu)$ be an arbitrary element of E_ν : its k -th coordinate is nothing else but the inner product of X by the unit vector U_k , in detail:

$$(X, U_1) = 1 \cdot x_1 + 0 \cdot x_2 + \dots + 0 \cdot x_\nu = x_1,$$

$$(X, U_2) = 0 \cdot x_1 + 1 \cdot x_2 + \dots + 0 \cdot x_\nu = x_2,$$

$$\dots\dots\dots$$

$$(X, U_\nu) = 0 \cdot x_1 + 0 \cdot x_2 + \dots + 1 \cdot x_\nu = x_\nu.$$

We see that $(X, U_k) = 0$ ($k = 1, 2, \dots, \nu$) implies $X = (0, 0, \dots, 0)$, which means that the system $\{U_1, U_2, \dots, U_\nu\}$ is *complete*.

It is noteworthy that U_k ($k = 1, 2, \dots, \nu$) is a generalization of the triplet of 3-dimensional unit vectors, and has the characteristic property that the projections of any element $X \in E_\nu$ relating to the vectors U_k furnish the coordinates x_k of X in proper order.

9.24 Thereafter it is to be expected that an arbitrary element $X = (x_1, x_2, \dots, x_\nu) \in E_\nu$ can uniquely be written in the form

$$X = C_1 U_1 + C_2 U_2 + \dots + C_\nu U_\nu, \quad (9.2-10)$$

with suitable real coefficients: furthermore that the k -th coefficient on the right is

$$C_k = (X, U_k) = x_k \quad (k = 1, 2, \dots, \nu). \quad (9.2-11)$$

In fact: from the meaning of $U_1, U_2, \dots, U_\nu$ and the definition of addition we conclude that the k -th coordinate of $x_1 U_1 + \dots + x_\nu U_\nu$ equals x_k ; therefore this sum signifies the ν -dimensional vector $(x_1, x_2, \dots, x_\nu)$. On the other hand: assuming that $X = (x_1, x_2, \dots, x_\nu)$ is of the form (9.2-10), by inner multiplication on both sides we get

$$\begin{aligned} (X, U_k) &= C_1 (U_1, U_k) + \dots + C_\nu (U_\nu, U_k) = \\ &= C_k \|U_k\|^2 = C_k \quad (k = 1, 2, \dots, \nu), \end{aligned} \quad (9.2-12)$$

i.e. C_k cannot differ from the value (9.2-11).

Let us call (for the sake of brevity) any system R consisting of ν elements $\in E_\nu$, the *base* of E_ν , if every element of the Euclidean ν -space can uniquely be represented as a linear combination of elements of R . Then we can say: *the system $\{U_1, U_2, \dots, U_\nu\}$ is the simplest base of E_ν .*

It is clear that any orthogonal system $\{V_1, V_2, \dots, V_\nu\}$ of distinct elements $\in E_\nu$ also constitutes a base of the space; i.e., an arbitrary ν -dimensional vector X permits a representation of the form

$$X = C_1 V_1 + C_2 V_2 + \dots + C_\nu V_\nu. \quad (9.2-13)$$

The above argument yields again a unique determination of the coefficients here, namely

$$C_k = \frac{(X, V_k)}{\|V_k\|^2} \quad (k = 1, 2, \dots, \nu), \quad (9.2-14)$$

because we obtain from (9.2-13) by the orthogonality of V_k ($k = 1, 2, \dots, \nu$):

$$(X, V_k) = \sum_{j=1}^{\nu} C_j (V_j, V_k) = C_k \|V_k\|^2 \quad (k = 1, 2, \dots, \nu). \quad (9.2-15)$$

Is $\{V_1, V_2, \dots, V_\nu\}$ also normalized, i.e., $\|V_k\| = 1$ ($k = 1, 2, \dots, \nu$), then the decomposition (9.2-13) becomes:

$$X = (X, V_1) \cdot V_1 + (X, V_2) \cdot V_2 + \dots + (X, V_\nu) \cdot V_\nu. \quad (9.2-16)$$

This is a counterpart of the representation of an ordinary vector by means of its components relating to a prescribed orthogonal system [cf. (9.2-8)].

Applying inner multiplication by X to both sides of (9.2-16), it results [cf. (9.2-3)]:

$$\|X\|^2 = (X, V_1)^2 + (X, V_2)^2 + \dots + (X, V_\nu)^2, \quad (9.2-17)$$

which can be regarded as an *extension of the Pythagoras theorem*.

We may mention also that in E_ν any system of linearly independent vectors can be utilized as a base too; however, the coefficients of a decomposition with respect to some non-orthogonal base can be determined in a rather complicated manner.

9.3 Transfer of fundamental ideas to the Hilbert space l^2

9.31 Keeping in mind the setting of our objectives given previously, we will now try to extend the concept of E_ν through the limiting process $\nu \rightarrow \infty$; more precisely: we want to study the totality of sequences $X = (x_1, x_2, \dots)$ of real numbers with the restriction $x_1^2 + x_2^2 + \dots < \infty$.¹ This set is called *Hilbert's sequence space* and denoted by l^2 ; any element X — like the case of E_ν — may be considered either as an abstract *point* or as an abstract (infinite-dimensional) *vector*, whose coordinates are x_k ($k = 1, 2, \dots$).²

To begin with the vectorial interpretation, first we establish that the *vectorial operations* (9.2-1)–(9.2-2), as well as the notion of *norm* can be transferred mechanically. Namely, if $X = (x_1, x_2, \dots)$, $Y = (y_1, y_2, \dots)$ denote two arbitrary elements of l^2 , we define

$$\left. \begin{aligned} X + Y &= (x_1 + y_1, \quad x_2 + y_2, \dots), \\ \alpha \cdot X &= (\alpha x_1, \quad \alpha x_2, \dots), \end{aligned} \right\} \quad (9.3-1)$$

$$(X, Y) = \sum_{k=1}^{\infty} x_k y_k, \quad (9.3-2)$$

$$\|X\| = \left(\sum_{k=1}^{\infty} x_k^2 \right)^{1/2}. \quad (9.3-3)$$

It is immediately seen that the condition $\sum x_k^2 < \infty$ figuring in the definition of l^2 expresses the natural requirement: *the norm* (length) of

¹ Note that by $x_k^2 \geq 0$, the *convergence* of the series $\sum x_k^2$ is equivalent to the boundedness of its partial sums; further that $\sum x_k^2 < \infty$ naturally implies $\lim_{k \rightarrow \infty} x_k = 0$.

² Hilbert [1].

any element of the space must be finite. Hence it follows that the partial sums of the series (9.3-2) are bounded, i.e., the series is convergent; namely, Cauchy's inequality (2.1-3) yields for every n the estimation

$$\sum_{k=1}^n |x_k| |y_k| \leq \left(\sum_{k=1}^n x_k^2 \right)^{1/2} \left(\sum_{k=1}^n y_k^2 \right)^{1/2} \leq \|X\| \|Y\|.$$

As a corollary, we get for $n \rightarrow \infty$ the so-called *generalized Cauchy inequality*:

$$\left(\sum_{k=1}^{\infty} x_k y_k \right)^{1/2} \leq \left(\sum_{k=1}^{\infty} x_k^2 \right)^{1/2} \left(\sum_{k=1}^{\infty} y_k^2 \right)^{1/2} \quad (9.3-4)$$

which is, as can easily be shown, equivalent to the following:

$$\left[\sum_{k=1}^{\infty} (x_k + y_k)^2 \right]^{1/2} \leq \left(\sum_{k=1}^{\infty} x_k^2 \right)^{1/2} + \left(\sum_{k=1}^{\infty} y_k^2 \right)^{1/2}. \quad (9.3-5)$$

By (9.3-5), the simultaneous convergence of $\sum x_k^2$ and $\sum y_k^2$ implies also $\sum (x_k + y_k)^2 < \infty$. Thus we may state: if $X \in l^2$ and $Y \in l^2$, the $X + Y$ and (for all real α) $\alpha \cdot X$ belong to l^2 as well.

As regards the related notions, the role of the *zero vector* O is plainly played by the sequence $\{0, 0, \dots\}$ (this is the only element of l^2 whose norm is 0); we call *unit vector* of l^2 any element with norm 1.

Two elements are said to be *equal*, when their difference is O , i.e., if their corresponding coordinates are identical by pairs. The *angle of inclination*, *unidirectionality* and *orthogonality* of two non-zero vectors, further the notion of *projection (component)* can exactly be defined so as in the case of the vectors of E_n . [Cf. (9.2-5)—(9.2-8).] The definitions of *orthogonal system*, *normality* and *completeness* remain unchanged.

In the space l^2 , the *simplest example of an orthonormal, complete system* is naturally countable infinite:

$$U_1 = (1, 0, 0, 0, \dots), U_2 = (0, 1, 0, 0, \dots), U_3 = (0, 0, 1, 0, \dots). \quad (9.3-6)$$

The discussion of this system can be carried out similarly to that of its finite pendant in the foregoing section.

9.32 As mentioned above, any element of l^2 can also be handled as an abstract *point* with the same *coordinates*; the first step in the "point-like" treatment is the introduction of distance.

On the basis of (9.2-4), the concept of *distance* for any two elements $X = (x_1, x_2, \dots)$, $Y = (y_1, y_2, \dots) \in l^2$ is given by the norm of their difference:

$$d[X, Y] = \left[\sum_{k=1}^{\infty} (x_k - y_k)^2 \right]^{1/2}. \quad (9.3-7)$$

The quantity $d[X, Y]$ is evidently non-negative and vanishes if and only if $X = Y$; it is also trivial that $d[X, Y] = d[Y, X]$. As for the so-called “triangle inequality” [corresponding to (2.1-2)]:

$$d[X, Y] \leq d[X, Z] + d[Y, Z], \quad (9.3-8)$$

it is nothing else but a transcription of (9.3-5) with $x_k - z_k$ for x_k and with $z_k - y_k$ for y_k .

Next we come to the transfer of notions connected with convergence in the *theory of point sets* (see Chapter 2). Since the (ordinary) convergence of a real number sequence means that the distance between its generic element and an appropriate real number (the limit) tends to zero, it is advisable to accept the definition:

A sequence of elements of l^2 :

$$\left. \begin{array}{l} X_1 = (x_{11}, x_{12}, \dots) \\ X_2 = (x_{21}, x_{22}, \dots) \\ \dots\dots\dots \\ X_n = (x_{n1}, x_{n2}, \dots) \\ \dots\dots\dots \end{array} \right\} \quad (9.3-9)$$

is said to be *convergent*, if there exists an element $X = (x_1, x_2, \dots) \in l^2$ such that

$$\lim_{n \rightarrow \infty} d[X_n, X] = 0; \quad (9.3-10)$$

X is called the *limit* of $\{X_1, X_2, \dots\}$. It is often important to distinguish between this kind of convergence and the classical one; then we speak about l^2 -convergence and l^2 -limit, respectively.

It is easy to see that any sequence (9.3-9) can possess only one single limit in the (9.3-10) sense; furthermore *Cauchy's convergence principle* (cf. 2.2.1) holds for the l^2 -convergence too: a sequence $X_k \in l^2$ ($k = 1, 2, \dots$) is convergent if and only if to any given $\varepsilon > 0$, however small, we can attach a natural number $n_0 = n_0(\varepsilon)$ such that $d[X_m, X_n] < \varepsilon$ whenever $m \geq n_0$, $n \geq n_0$.

9.33 The relevancy of l^2 -convergence appears at once when the problem is studied: how the *decompositions* (9.2-10)—(9.2-11) and (9.2-13)—(9.2-14), respectively, can be adapted here. Namely, in the space l^2 any complete orthogonal system consists of infinitely many elements [cf. (9.3-6)], and therefore the “ l^2 -pendants” of the representations in question must contain infinite series as well. But the sum of such a series can then only be defined in the usual manner (i.e., as a limit of the partial sums), if we clarify before: what do we mean by the limit of a sequence consisting of elements of l^2 ?

As an illustration we mention e.g.: for any $X = (x_1, x_2, \dots) \in l^2$ we have the expansion

$$X = x_1 U_1 + x_2 U_2 + \dots \quad (l^2) \quad (9.3-11)$$

in the sense that the l^2 -limit of the sequence consisting of partial sums of the right-hand series equals X .¹

In fact, since by (9.3-6)

$$x_1 U_1 + x_2 U_2 + \dots + x_n U_n = (x_1, x_2, \dots, x_n, 0, 0, \dots),$$

so we have, owing to (9.3-7):

$$\left(d \left[\sum_{k=1}^n x_k u_k, X \right] \right)^2 = x_{n+1}^2 + x_{n+2}^2 \dots$$

in the classical sense; and the series on the right-hand side — as a remainder term of the series $\sum x_k^2$ assumed to be convergent — tends to 0 as $n \rightarrow \infty$.

We omit further details concerning l^2 , because we shall utilize in the sequel an extension of it, the space L^2 .

9.4 Further generalization: the function space L^2 as a vector field

9.41 If the generic element x_n of a sequence $(x_1, x_2, \dots)$ is considered as a function of the discrete variable n , then the sequence becomes the totality of certain function values. Thus it is advisable to develop further the concept of l^2 in such a way that functions of a continuous variable t substitute for sequences, appropriate integrals for series and their sums.

Let (a, b) be a given finite or infinite interval; it is irrelevant, whether the end points are included in it or not. A function $\varphi(t)$ is called *square integrable* in (a, b) , if it is measurable there and the Lebesgue integral $\int_a^b \varphi(t)^2 dt$ exists.² (The latter condition obviously corresponds to the restriction $\sum x_k^2 < \infty$.) By 7.4.1, every bounded and (L) -integrable function

¹ Accordingly, the system $\{U_1, U_2, \dots\}$ is the *simplest base* of l^2 .

² For finite (a, b) , the (L) -integrability of $\varphi(t)^2$ and the elementary inequality $|\varphi(t)| \leq \frac{1}{2} [1 + \varphi(t)^2]$ (geometric and arithmetic means !) imply the existence of $\int_a^b |\varphi(t)| dt$ by 7.4.1. Consequently, on a finite interval the square integrability always involves the (L) -integrability. If (a, b) is infinite, then the measurability of $\varphi(t)$ signifies that $\varphi(t)$ is measurable in all finite subintervals of (a, b) simultaneously. (Cf. 6.5.2.)

is square integrable; but an unbounded or a not (L) -integrable function can also be square integrable as e.g. $\log t$ in $(0,1)$ and $1/t$ in the interval $(1, \infty)$, respectively. On the other hand, e.g. $t^{-1/2}$ is (L) -integrable but not square integrable on the segment $0 < t \leq 1$. The totality of functions square integrable in (a,b) is called *Hilbert's function space* and it is denoted by $L^2(a,b)$, briefly by L^2 .

This concept is not only a counterpart of l^2 , but also a *generalization* of it *in a certain sense*. For, any real number sequence $(x_1, x_2, \dots)$ can be regarded as the (ordered) codomain of a function $\varphi(t)$ equalling x_1 on the segment $0 \leq t < 1$, x_2 on the segment $1 \leq t < 2$, etc. This $\varphi(t)$ — as the limit of a sequence of step functions — is plainly measurable and also square integrable in $(0, \infty)$ provided that $x_1^2 + x_2^2 + \dots < \infty$; then we have $\int_0^\infty \varphi(t)^2 dt = \sum_{k=1}^\infty x_k^2$ by 7.1.2. Hence we see that the space $L^2(a,b)$ reduces to l^2 , if $a = 0$, $b = \infty$ and only such square integrable functions are considered for which every interval $[k-1, k)$ ($k = 1, 2, \dots$) is a segment of constancy. Like the case of l^2 , also the elements of L^2 , i.e., the square integrable functions can be interpreted either as abstract *points* or as abstract *vectors*.¹

9.42 We define the *addition* of the elements of $L^2(a,b)$ and their *multiplication by a constant* in the manner usual for functions, further the *inner product* of $\varphi(t) \in L^2(a,b)$, $\psi(t) \in L^2(a,b)$ by the formula

$$(\varphi, \psi) = \int_a^b \varphi(t)\psi(t) dt; \quad (9.4-1)$$

the integral on the right is obviously an analogue and — according to the above — also a generalization of the series (9.3-2). Patterned after (9.3-3), the definition of the *norm* (length) of an element $\varphi(t) \in L^2(a,b)$ is:

$$\|\varphi\| = \left[\int_a^b \varphi(t)^2 dt \right]^{1/2}, \quad (9.4-2)$$

and at the same time from (9.4-1)

$$(\varphi, \varphi) = \int_a^b \varphi(t)^2 dt = \|\varphi\|^2. \quad (9.4-3)$$

These terms and notations can be taken as fully justified only if the pertinent operations can be carried out for arbitrary elements of $L^2(a,b)$,

¹ Cf. Hilbert [1].

further if in case of $\varphi(t) \in L^2(a, b)$, $\psi(t) \in L^2(a, b)$ the functions $\varphi(t) + \psi(t)$, $\alpha\varphi(t)$ (with any real number α) belong to $L^2(a, b)$ as well, finally if the notions just introduced possess the properties suggested by their geometric origin.

In fact, the existence of the norm (9.4-2) is assured by the definition of square integrable functions, while that of the inner product (9.4-1) is a consequence of the inequality between the geometric and arithmetic mean of $\varphi(t)^2$ and $\psi(t)^2$, i.e., of

$$|\varphi(t)\psi(t)| \leq \frac{1}{2}[\varphi(t)^2 + \psi(t)^2], \quad (9.4-4)$$

as well as of the majorant criterion 7.4.1 (valid for an unbounded basic set, too). The well-known *inequality of Buniakowsky and Schwarz*:

$$\left[\int_a^b \varphi(t)\psi(t) dt \right]^2 \leq \left[\int_a^b \varphi(t)^2 dt \right] \left[\int_a^b \psi(t)^2 dt \right] \quad (9.4-5)$$

yields at once the pendant of the elementary estimation of a scalar product by means of the lengths of factors: $|(\varphi, \psi)| \leq \|\varphi\| \|\psi\|$ and gives another easy verification for the existence of (9.4-1).

As for the closure of $L^2(a, b)$ with respect to addition and multiplication by a constant, it is trivial that $\varphi(t) \in L^2(a, b)$ implies $\alpha\varphi(t) \in L^2(a, b)$ for any real number α . On the other hand, from $\varphi(t) \in L^2(a, b)$, $\psi(t) \in L^2(a, b)$ it follows $\varphi(t) + \psi(t) \in L^2(a, b)$ by $[\varphi(t) + \psi(t)]^2 = \varphi(t)^2 + \psi(t)^2 + 2\varphi(t)\psi(t)$. Note that the so called *Minkowski inequality*:

$$\left\{ \int_a^b [\varphi(t) + \psi(t)]^2 dt \right\}^{1/2} \leq \left[\int_a^b \varphi(t)^2 dt \right]^{1/2} + \left[\int_a^b \psi(t)^2 dt \right]^{1/2}, \quad (9.4-6)$$

or in an abbreviated form: $\|\varphi + \psi\| \leq \|\varphi\| + \|\psi\|$ can easily be deduced from (9.4-4), and corresponds to the elementary inequality for the absolute value of a sum.¹

Let us insert the simplest proof of (9.4-5), which consists of one single sentence: on the one hand, we have the obvious identity

$$\left. \begin{aligned} \frac{1}{2} \int_a^b \left\{ \int_a^b [\varphi(t)\psi(u) - \psi(t)\varphi(u)]^2 dt \right\} du = \\ = \left[\int_a^b \varphi(t)^2 dt \right] \left[\int_a^b \psi(t)^2 dt \right] - \left[\int_a^b \varphi(t)\psi(t) dt \right]^2; \end{aligned} \right\} \quad (9.4-7)$$

¹ Observe that (9.4-5) is a counterpart of (9.3-4), while (9.4-6) corresponds to (9.3-5). Cf. Minkowski [1].

on the other hand, the inner integrand on the left, and consequently both the left- and the right-hand side expressions in (9.4-7) are non-negative. Since the left-hand side integral vanishes if and only if one of the functions φ and ψ can be obtained from the other by multiplying the latter by a constant, apart possibly from a null set in (a, b) — this is also the exact condition of equality in (9.4-5).

9.43 We call a *zero vector* or *null element* of $L^2(a, b)$ an arbitrary function vanishing in (a, b) almost everywhere; for this notion we use the notation ω . The motivation of the name is that $\|\varphi\| = 0$ if and only if $\varphi(t) = \omega$ (cf. 7.12). Correspondingly, $\varphi(t) \in L^2(a, b)$ and $\psi(t) \in L^2(a, b)$ are said to be *identical elements* of the space, if $\varphi(t) - \psi(t) = \omega$, i.e., $\varphi(t) = \psi(t)$ for almost every t in (a, b) .¹ Any element satisfying the condition $\|\varphi\| = 1$ is called *unit vector* of $L^2(a, b)$; in the case of a finite (a, b) , a trivial example is furnished by the function $\varphi(t) = (b - a)^{-1/2}$ ($a < t < b$), but there exist of course infinitely many non-constant unit vectors in $L^2(a, b)$, too.

Let $\varphi(t)$ and $\psi(t)$ be two non-zero vectors $\in L^2(a, b)$. Like (9.2-5), the *cosine of the inclination angle* between φ and ψ is defined by the ratio

$$q = \frac{\int_a^b \varphi(t)\psi(t) dt}{\left[\int_a^b \varphi(t)^2 dt \right]^{1/2} \left[\int_a^b \psi(t)^2 dt \right]^{1/2}}. \quad (9.4-8)$$

[Cf. (9.4-1)—(9.4-2).] This name is reasonable, because by (9.4-5) we always have the inequality $|q| \leq 1$. On the basis of the proof of (9.4-5), $q = 1$ holds if and only if $\varphi(t) = K\psi(t)$ (K constant) for a.e. $t \in (a, b)$; in this case we say that $\varphi(t)$ and $\psi(t)$ are *unidirectional*. In turn, the exact condition of $q = 0$ is

$$\int_a^b \varphi(t)\psi(t) dt = 0; \quad (9.4-9)$$

if (9.4-9) is fulfilled, $\varphi(t)$ and $\psi(t)$ are said to be *orthogonal* (perpendicular) to each other. By the *projection* (component) of $\varphi(t)$ relating to $\psi(t)$ we mean the integral

$$\int_a^b \varphi(t)\psi^0 dt, \quad (9.4-10)$$

where $\psi^0 = \|\psi\|^{-1}\psi(t)$ is the unit vector unidirectional with $\psi(t)$ [cf. (9.2-8)]. As can be seen, the division of any element by its norm always yields a unit vector; for this operation, we keep the term *norming* or *normalization*.

¹ Remember that such functions were said to be (L) -equivalent in 7.12.

To systems consisting of elements of $L^2(a, b)$, the notions of orthogonality, normality and completeness can be transferred immediately (cf. 9.23). In particular: a function system $\{\varphi_n(t)\}$ is called *orthogonal in* (a, b) , if the value of the integral

$$\int_a^b \varphi_m(t) \varphi_n(t) dt \quad (9.4-11)$$

is 0 whenever $m \neq n$ and positive for $m = n$; if besides

$$\int_a^b \varphi_n(t)^2 dt = 1 \quad (n = 1, 2, \dots), \quad (9.4-12)$$

then the system is *orthonormal*. We say that $\{\varphi_n(t)\}$ is *complete* in $L^2(a, b)$, if there is no element of $L^2(a, b)$ orthogonal to all elements of the system; in other words: if $f(t) = \omega$ is the only element of $L^2(a, b)$ satisfying the conditions

$$\int_a^b f(t) \varphi_n(t) dt = 0 \quad (n = 1, 2, \dots).^1 \quad (9.4-13)$$

9.44 It is not uninteresting to consider a *simple example of orthonormal systems* which corresponds to (9.3-6) (and generalizes it at the same time in a sense).

Let $(a, b) = \bigcup_{n=1}^{\infty} I_n$ be a decomposition of (a, b) into pairwise disjoint finite subintervals $I_1, I_2, \dots$; we define the system in question as follows:

$$\kappa_n(t) = \begin{cases} |I_n|^{-1/2} & \text{for } t \in I_n \\ 0 & \text{for } t \notin I_n \end{cases} \quad (n = 1, 2, \dots). \quad (9.4-14)$$

Evidently, every term of $\{\kappa_n(t)\}$ is a unit vector and any two distinct terms of the system are orthogonal to each other. We have namely for every n

$$\int_a^b \kappa_n(t)^2 dt = \int_{(I_n)} |I_n|^{-1} dt = 1; \quad (9.4-15)$$

and for $m \neq n$ $\kappa_m(t)$ vanishes at any point where $\kappa_n(t) \neq 0$, so that

$$\int_a^b \kappa_m(t) \kappa_n(t) dt = 0. \quad (9.4-16)$$

Unfortunately, the system $\{\kappa_n(t)\}$ is *not complete* in the space $L^2(a, b)$; it is easy to find such an element of the space, which is different from ω and

¹ Note that the concept of orthogonality has only been defined for *non-zero* elements of $L^2(a, b)$.

orthogonal to any of the functions $\kappa_n(t)$. Such a function can be constructed e.g. in the following manner. Let us divide any of the intervals $I_1, I_2, \dots$ into halves, and denote the left-hand half of I_n by I'_n , the right-hand of it by I''_n ($n = 1, 2, \dots$); let us take a sequence of positive numbers $x_1, x_2, \dots$ for which $x_1^2 + x_2^2 + \dots < \infty$, and consider the function:

$$\Psi(t) = \begin{cases} x_n & \text{if } t \in I'_n \\ -x_n & \text{if } t \in I''_n \end{cases} \quad (n = 1, 2, \dots). \quad (9.4-17)$$

Evidently $\Psi(t) \in L^2(a, b)$ and $\int_a^b \Psi(t)^2 dt = \sum_{n=1}^{\infty} x_n^2 |I_n|$, moreover by (9.4-14) and (9.4-17) we get

$$\int_a^b \Psi(t) \kappa_n(t) dt = |I_n|^{-1/2} \int_{(I_n)} \Psi(t) dt = 0 \quad (n = 1, 2, \dots), \quad (9.4-18)$$

as asserted above.

Our statements about the system $\{\kappa_n(t)\}$ already suggest that the geometry of $L^2(a, b)$ is more complicated than the features of abstract spaces discussed before, and that we need deeper reasoning in considering the former. On the other hand, the example of $\Psi(t)$ shows that orthogonal functions are in general not of constant sign, they take positive and negative values alike.¹

9.5 Introduction of the notions of distance and convergence in space L^2

9.51 Next let us study the space $L^2(a, b)$ as a set of abstract points.

The distance $d[\varphi, \psi]$ of the elements $\varphi(t) \in L^2(a, b)$ and $\psi(t) \in L^2(a, b)$ (as abstract points) is defined on the model of (9.3-7) by the formula:

$$d[\varphi, \psi] = \|\varphi - \psi\| = \left\{ \int_a^b [\varphi(t) - \psi(t)]^2 dt \right\}^{1/2}. \quad (9.5-1)$$

Hence we see that (I) $d[\varphi, \psi] = 0$ holds if and only if $\varphi(t) - \psi(t) = 0$, i.e., φ and ψ are identical elements of the space; (II) otherwise $d[\varphi, \psi] > 0$ and

¹ Quite recently, such a building up of the theory of $L^2(a, b)$ has also been found which is based — corresponding to the ideas of 9.44 — upon certain systems of characteristic functions belonging to certain partitions of $[a, b]$, as well as upon appropriate passages to limit. (Cf. Mikolás [15] and partly Sz.-Nagy [1], [2], [9].)

always $d[\varphi, \psi] = d[\psi, \varphi]$; (III) for any triplet φ, ψ, χ of elements we have the so-called *triangle inequality*

$$d[\varphi, \psi] \leq d[\varphi, \chi] + d[\psi, \chi]. \quad (9.5-2)$$

[Cf. (9.4-6) with $\varphi(t) - \chi(t)$ instead of $\varphi(t)$ and with $\chi(t) - \psi(t)$ instead of $\psi(t)$.]

It is very important to transfer the convergence concept of the theory of point sequences to the space L^2 . The fundamental definition is: a *sequence of functions* $f_n(t) \in L^2(a, b)$ ($n = 1, 2, \dots$) is called $L^2(a, b)$ -convergent, if there exists a function $f(t) \in L^2(a, b)$ whose distance from $f_n(t)$ tends to 0 as $n \rightarrow \infty$.¹ The latter condition is obviously equivalent to

$$\lim_{n \rightarrow \infty} \int_a^b [f_n(t) - f(t)]^2 dt = 0 \quad (9.5-3)$$

or for finite (a, b) to the assumption: the expression

$$(b-a)^{-1/2} d[f_n, f] = \sqrt{\frac{1}{b-a} \int_a^b [f_n(t) - f(t)]^2 dt}, \quad (9.5-4)$$

i.e., the so-called *quadratic integral mean* of $|f_n(t) - f(t)|$ on (a, b) should tend to 0 as $n \rightarrow \infty$.

This " $L^2(a, b)$ -convergence" (in short: " L^2 -convergence") differs from the concepts of ordinary (pointwise) and uniform convergence since here the square integral in (9.5-3) and the integral mean (9.5-4) take over the role of the deviation $|f_n(t) - f(t)|$ and of its maximum over (a, b) , respectively. Thus it is understandable that the terms "*convergent with respect to square integral*" or "*convergent in mean*" (in French: *en moyenne*) on (a, b) are also in use instead of the expression: " $L^2(a, b)$ -convergent". The name of the function f mentioned in the above definition is: the $L^2(a, b)$ -limit of the sequence of functions $f_n(t)$ ($n = 1, 2, \dots$). The specified notations of the $L^2(a, b)$ -convergence are:

$$\lim_{n \rightarrow \infty} f_n = f \quad [L^2(a, b)] \quad (9.5-5)$$

or

$$f_n \rightarrow f \quad \text{as } n \rightarrow \infty \quad [L^2(a, b)]. \quad (9.5-6)$$

According to the foregoing, we can define at once also the $L^2(a, b)$ -convergence of a series of functions $f_1(t) + f_2(t) + \dots$, with $f_n(t) \in L^2(a, b)$ ($n = 1, 2, \dots$). By this we mean the $L^2(a, b)$ -convergence of the sequence

¹ It can easily be concluded on the basis of (9.5-2) that the existence of such a function f is *unique*.

$\{s_n(t)\}$ consisting of partial sums of $\sum f_n(t)$, and the $L^2(a, b)$ -sum of the series means the $L^2(a, b)$ -limit of $\{s_n(t)\}$. The way of writing

$$\sum_{n=1}^{\infty} f_n(t) = s(t) \quad [L^2(a, b)]$$

signifies that the $L^2(a, b)$ -sum of the series $f_1(t) + f_2(t) + \dots$ exists and equals $s(t)$.

9.52 It is to be stressed that for finite (a, b) the L^2 -convergence is certainly a weaker condition than uniform convergence, namely

$$(b-a)^{-1/2} d[f_n, f] \leq \max_{(a,b)} |f_n(t) - f(t)|.$$

In turn, we can simply establish that *there is no immediate connection between the concept of pointwise convergence and the $L^2(a, b)$ -convergence.*

Pertinent examples:

I. As is well-known, the function sequence $1, t, t^2, \dots$ tends to 0 at any point $t \in [0, 1)$ and to 1 at $t = 1$; furthermore it is not uniformly convergent in the closed interval $[0, 1]$. On the other hand, this sequence converges with respect to square integral in $[0, 1]$ and its L^2 -limit is identically 0, because

$$\int_0^1 (t^n)^2 dt = \frac{1}{2n+1} \rightarrow 0 \quad (n \rightarrow \infty).$$

II. Let $\vartheta \geq 1/2$ and regard the sequence of functions:

$$f_n(t) = \begin{cases} n^\vartheta & \text{if } 0 < t < 1/2 \\ 0 & \text{if } 1/n \leq t < 1 \end{cases} \quad (n = 1, 2, \dots).$$

Then at any fixed point $t \in (0, 1)$ we have $f_n(t) = 0$ only if n is sufficiently large; therefore $\{f_n(t)\}$ converges everywhere in $(0, 1)$ to the limit 0. So it could be expected that the function $f(t) \equiv 0$ would also be the L^2 -limit of our sequence.

Unfortunately this is not true, since

$$\int_0^1 f_n(t)^2 dt = \int_0^{1/n} n^{2\vartheta} dt = n^{2\vartheta-1},$$

which is equal to 1 or tends to $+\infty$, according to $\vartheta = 1/2$ and $\vartheta > 1/2$, respectively.

We have to note, however, a trivial "global" connection:

9.5.1 If a sequence of functions $f_n(t) \in L^2(a, b)$ ($n = 1, 2, \dots$) is $L^2(a, b)$ -convergent and at the same time convergent for almost all $t \in (a, b)$ in the classical sense, so its $L^2(a, b)$ -limit and its (ordinary) limit function in (a, b) are identical elements of the space $L^2(a, b)$.¹

P r o o f. Denoting the $L^2(a, b)$ -limit of $\{f_n(t)\}$ by $f(t)$, and the ordinary (pointwise) limit of the sequence by $f^*(t)$, our assumptions imply by Lebesgue's main theorem 7.5.1:

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_a^b [f_n(t) - f(t)]^2 dt &= \int_a^b \lim_{n \rightarrow \infty} [f_n(t) - f(t)]^2 dt = \\ &= \int_a^b (f^* - f)^2 dt = 0. \end{aligned} \quad (9.5-7)$$

Hence we obtain that $f(t) = f^*(t)$ for almost all $t \in (a, b)$.

* * *

9.53 It does not seem to be assured at all, that for this new-type notion of convergence, the $L^2(a, b)$ -convergence, the *pendant of Cauchy's convergence principle* holds good. (By the deviation of two elements, of course, we mean the norm of their difference.) Yet in 1907, F. RIESZ and — independently — the Austrian mathematician E. FISCHER could give a positive answer to this question — by discovering one of the main “pillars” of the theory, the so-called *Riesz-Fischer convergence criterion*:²

9.5.2 A sequence of functions $f_n(t) \in L^2(a, b)$ ($n = 1, 2, \dots$) is $L^2(a, b)$ -convergent, if and only if (*) an index $n_0 = n_0(\varepsilon)$ can be attached to any given number $\varepsilon > 0$ such that $\|f_m - f_n\| < \varepsilon$ whenever $m \geq n_0$, $n \geq n_0$.

The original demonstration of this result is founded on *Fatou's lemma* and *B. Levi's theorem* (cf. 7.5.2—7.5.3), further it utilizes (like most later proofs) some principles of choice, so that we have a *consideration of existence*. In what follows a more recent verification found by the author will be presented, which relies beside the definition of the (L) -integral only on the notion of *absolute continuity* as the characteristic property of (L) -integral functions (cf. 7.6.7), and moreover which also includes an *explicit representation of the $L^2(a, b)$ -limit* looked for.

Since the necessity of the above condition (called simply *Cauchy's condition*) is obvious by the “triangle inequality” (9.5-2), it is enough to deal with the statement:

¹ Cf. 9.43.

² Both mathematicians published their pertinent result in “Comptes Rendus” of the Academy of Paris; RIESZ two month earlier than FISCHER. Cf. F. Riesz [1] and Fischer [1], as well as F. Riesz [3], [9], [19] and Radon [1].

9.5.3 Any sequence $\{f_n\} \subset L^2(a, b)$ satisfying the condition

$$\lim_{m, n \rightarrow \infty} \|f_m - f_n\| = 0 \quad (9.5-8)$$

is $L^2(a, b)$ -convergent, and its $L^2(a, b)$ -limit can be written a.e. in the form:

$$f(t) = \frac{d}{dt} \left[\lim_{n \rightarrow \infty} \int_{\alpha}^t f_n(u) du \right] \quad [\alpha \in (a, b)]. \quad (9.5-9)$$

P r o o f. ¹ From (9.5-8) we conclude by $\| \|f_m\| - \|f_n\| \| \leq \|f_m - f_n\|$ the boundedness and convergence of $\{\|f_n\|\}$; let us denote the least upper bound and limit of this number sequence by B and A , respectively. Let $\alpha \leq t \leq \beta$ be a finite closed subinterval of (a, b) , then on $[\alpha, \beta]$ the sequence of integral functions

$$\Phi_n(t) = \int_{\alpha}^t f_n(u) du \quad (n = 1, 2, \dots)$$

has a finite limit $\Phi(t)$, namely on the basis of the inequality

$$|\Phi_m(t) - \Phi_n(t)| \leq (t - \alpha)^{1/2} \|f_m - f_n\|$$

[cf. (9.4-5)] and of (9.5-8).

The function $\Phi(t)$ thus obtained is absolutely continuous in $[\alpha, \beta]$, because we have for any finite system $\{[\alpha_k, \beta_k]\}$ of non-overlapping subintervals belonging to $[\alpha, \beta]$:

$$\left| \sum_k [\Phi(\beta_k) - \Phi(\alpha_k)] \right| = \lim_{n \rightarrow \infty} \left| \int_I f_n(t) dt \right| \leq A \cdot |I|^{1/2}$$

where $I = \bigcup_k [\alpha_k, \beta_k]$ and the last term tends to 0 as the total length of I becomes arbitrarily small. Therefore some well-known theorems (cf. 4.64, 7.6.7) imply that $\Phi'(t) = f(t)$ exists at a.e. point $t \in [\alpha, \beta]$, further f is measurable, (L) -integrable in $[\alpha, \beta]$, and $\Phi(t) = \int_{\alpha}^t f(u) du$ ($\alpha \leq t \leq \beta$).

It is easy to see that the last relation can be extended to every bounded (L) -measurable set $E \subset (a, b)$, i.e., we have the limit relation:

$$\lim_{n \rightarrow \infty} \int_{(E)} f_n(t) dt = \int_{(E)} f(t) dt. \quad (9.5-10)$$

¹ Mikolás [9].

In fact if E is open, then by 2.4.4 it is the union of finitely many or countably infinitely many, pairwise disjoint open intervals, and accordingly, the integrals in (9.5-10) can be decomposed into the sum of some subintegrals taken on line segments. If the number of these subintegrals is finite, so we get (9.5-10) by termwise passage to the limit immediately. Otherwise we have to split up the integral of $f_n(t) - f(t)$ over E into parts taken on a finite and on an infinite system of subintervals, respectively, where the second term becomes arbitrary small when diminishing the total length of the pertinent subintervals, while the first one (after fixing the number of subintervals) will become as small as we please by choosing the index n large enough. Finally, the case of a non-open set E reduces to that of an open E inasmuch as we put $E = S_1 - S_2$, where S_1 is an open set covering E , and S_2 is a subset of S_1 whose (L) -measure may be arbitrarily small. Then the validity of (9.5-10) appears at once from the estimation [resulting by (7.1-11), 7.43 and (9.4-5)]:

$$\left| \int_{(E)} [f_n(t) - f(t)] dt \right| \leq \left| \int_{(S_1)} [f_n(t) - f(t)] dt \right| + \left| \int_{(S_2)} f(t) dt \right| + B \cdot |S_2|^{1/2}.$$

Turning now to the deviation $\|f_n - f\|$, let us take first on the ordinate axis a "scale" $0 = y_0 < y_1 < \dots$ with $y_k \rightarrow \infty$, $y_k - y_{k-1} \leq \delta y_k^{-1}$ (δ fixed); next to a given interval $i \subset (a, b)$ let us attach the point sets:

$$\begin{aligned} e_{k,n} &= \{t : t \in i, y_{k-1} \leq f(t) - f_n(t) < y_k\} \quad (k > 0); \\ e_{-1,n} &= \{t : t \in i, -y_1 < f(t) - f_n(t) < 0\}; \\ e_{-k,n} &= \{t : t \in i, -y_k < f(t) - f_n(t) \leq -y_{k-1}\} \quad (k > 1). \end{aligned}$$

By (9.4-5) and (9.5-8) there exists an index $n_0 = n_0(\varepsilon)$ (independent of the choice of $\{y_k\}$ and i) such that for any pair p, q of positive integers and with the restriction $|e_{k,n}| \neq 0$:

$$\sum_{k=-p}^q |e_{k,n}|^{-1} \left\{ \int_{(e_{k,n})} [f_m(t) - f_n(t)] dt \right\}^2 \leq \|f_m - f_n\|^2 < \varepsilon \quad (m, n \geq n_0).^1$$

If we apply here first the limiting process $m \rightarrow \infty$, after increasing p and q indefinitely, we obtain by using (9.5-10):

$$\sum_{k=-\infty}^{\infty} |e_{k,n}|^{-1} J_{k,n}^2 \leq \varepsilon \quad (n \geq n_0), \quad (9.5-11)$$

¹ The dash beside the sign of summation means, as usual, that the value $k = 0$ is to be omitted.

where

$$J_{k,n} = \int_{(e_{k,n})} [f(t) - f_n(t)] dt.$$

But the sum (9.5-11) lies by $y_{k-1}|e_{k,n}| \leq J_{k,n} \leq y_k|e_{k,n}|$ and $-y_k|e_{-k,n}| \leq J_{-k,n} \leq -y_{k-1}|e_{-k,n}|$ ($k > 0$) between two "approximating sums of Lebesgue type" (cf. 7.41):

$$\sigma'_L = \sum_{k=1}^{\infty} y_{k-1}^2 (|e_{k,n}| + |e_{-k,n}|), \quad \sum'_L = \sum_{k=1}^{\infty} y_k^2 (|e_{k,n}| + |e_{-k,n}|)$$

with $y_k^2 - y_{k-1}^2 \leq 2\delta$. Therefore (9.5-11) implies that the (L) -integral of $[f(t) - f_n(t)]^2$ over i exists for $n \geq n_0$ and has a value $\leq \varepsilon$. Since this holds for every $i \subset (a, b)$, and ε can be chosen arbitrarily small, we conclude $f(t) \in L^2(a, b)$, further $\|f - f_n\| \rightarrow 0$ ($n \rightarrow \infty$).

Accordingly, the function f constructed above, i.e., the derivative with respect to t of the limit of $\{\Phi_n(t)\}$ yields almost everywhere the $L^2(a, b)$ -limit of the sequence $\{f_n\}$ on the condition (9.5-8), as asserted before.

* * *

9.54 The significance of $L^2(a, b)$ -convergence will turn up several times in the sequel. Now, we want to mention first of all an interesting consequence of this concept: if for finite (a, b) a function sequence $\{f_n\} \subset L^2(a, b)$ is $L^2(a, b)$ -convergent with the $L^2(a, b)$ -limit f , then the number sequence $\int_a^b f_n(t) dt$ ($n = 1, 2, \dots$) obtained by (L) -integration tends to $\int_a^b f(t) dt$. More generally:

$$\lim_{n \rightarrow \infty} \int_a^b \varphi(t) f_n(t) dt = \int_a^b \varphi(t) f(t) dt \quad (9.5-12)$$

holds, $\varphi(t)$ denoting an arbitrary square integrable function. (Therefore f serves as substitute, so to say, for the "classical" limit $\lim_{n \rightarrow \infty} f_n$ at termwise integration also in the case when the latter does not exist; cf. 7.2.1, 7.5.1.) In fact, our assumption imply by (9.4-5):

$$\left| \int_a^b \varphi(t) f_n(t) dt - \int_a^b \varphi(t) f(t) dt \right| \leq \|\varphi\| \cdot \|f_n - f\|,$$

and the bound on the right becomes for $\|f_n - f\| \rightarrow 0$ as small as we please.

Hence it is advisable to introduce two new terms: we say that a sequence $\{f_n\} \subset L^2(a, b)$ converges weakly to some function $f \in L^2(a, b)$, if we have (9.5-12) for every $\varphi \in L^2(a, b)$; on the other hand, the $L^2(a, b)$ -convergence

is often called *strong convergence*. Thus it can be stated that in the space $L^2(a, b)$, any strongly convergent sequence is also weakly convergent.

9.55 Being in possession of the concept of $L^2(a, b)$ -convergence, it is easy to develop further the discussion of the space $L^2(a, b)$ as an abstract point set — only the material on E_ν in Chapter 2 is to be followed.

For instance, the notion of *accumulation point* can be introduced in the following manner: a function $f(t) \in L^2(a, b)$ is said to be an accumulation point of some set S [consisting of elements $\in L^2(a, b)$], if we can select from S a sequence of distinct elements whose $L^2(a, b)$ -limit is $f(t)$. (The function f itself does not necessarily belong to S .) The pendant of the notion “*neighbourhood*” is indispensable as well: by the ε -neighbourhood of an element $\varphi_0(t) \in L^2(a, b)$ we mean the totality of those $\varphi \in L^2(a, b)$ for which $d[\varphi, \varphi_0] < \varepsilon$. [Cf. (9.5–1).]

After these, we can transfer to subsets of $L^2(a, b)$ almost word for word the definitions of *boundedness*, *compactness*, *closure* and *openness*, as well as a number of further fundamental notions connected with *derivate sets*.

9.6 The problem of decomposition and the concept of general Fourier series

9.61 Let us return to the vectorial interpretation of the space $L^2(a, b)$ and deal with the question of central importance: *when and how can be represented a function $f(t) \in L^2(a, b)$ (considered as an abstract vector) as a linear combination with constant coefficients of the elements $\varphi_1(t) \in L^2(a, b), \varphi_2(t) \in L^2(a, b), \dots$ of an orthogonal system, i.e., in the form of an orthogonal series*

$$C_1\varphi_1(t) + C_2\varphi_2(t) + \dots \quad (9.6-1)$$

Note that the corresponding geometrical problems are: decomposition of a given vector to components directed along some vectors pairwise perpendicular to each other, and representation of any position vector with respect to a common base (orthogonal vector system). This is why we also speak generally about *decomposition* or *representation problem*, the orthogonal system $\varphi_1(t), \varphi_2(t), \dots$ being called the *base system* of decomposition.

In the Euclidean ν -space the system of unit vectors (9.2–9), in l^2 the pertinent base system (9.3–6) has played a distinguished role, but in the space $L^2(a, b)$, there is no *complete* analogue of these systems (cf. 9.44). Thus here we shall study the problem for the case of an arbitrary base system, and for the present purely in a formal manner.

9.62 Suppose that $f(t)$ can be represented in the desired form (in a classical sense), i.e., it holds

$$f(t) = C_1\varphi_1(t) + C_2\varphi_2(t) + \dots \quad (9.6-2)$$

at every point $t \in (a, b)$. Then multiplying by $\varphi_n(t)$, after integrating on (a, b) , further carrying out the integration on the right formally term-by-term, we get:

$$\int_a^b f(t)\varphi_n(t) dt = C_1 \int_a^b \varphi_1(t)\varphi_n(t) dt + C_2 \int_a^b \varphi_2(t)\varphi_n(t) dt + \dots \quad (9.6-3)$$

Since the orthogonality of $\{\varphi_n(t)\}$, the only integral on the right differing from zero is

$$\int_a^b \varphi_n(t)\varphi_n(t) dt = \int_a^b \varphi_n(t)^2 dt,$$

(9.6-3) becomes

$$\int_a^b f(t)\varphi_n(t) dt = C_n \int_a^b \varphi_n(t)^2 dt,$$

whence

$$C_n = \frac{\int_a^b f(t)\varphi_n(t) dt}{\int_a^b \varphi_n(t)^2 dt} \quad (n = 1, 2, \dots). \quad (9.6-4)$$

Note that instead of integration of products we might have to speak about inner (scalar) multiplication, instead of saying termwise integration also the expression "reversing the order of inner multiplication and addition" (distributivity!) might have to be used. Furthermore, the result of calculation can also be written simply as follows [cf. (9.2-14)]:

$$C_n = \frac{(f, \varphi_n)}{\|\varphi\|^2} \quad (n = 1, 2, \dots). \quad (9.6-5)$$

9.63 Let us observe: the above argumentation shows, strictly speaking, not more than the series (9.6-1) can represent in (a, b) the function $f(t)$ as sum *only* with the special values (9.6-4)—(9.6-5) of the coefficients, *provided* that the term-by-term integration of the series $\sum_{k=1}^{\infty} C_k \varphi_k(t) \varphi_n(t)$ for any fixed n is permitted (e.g. if the partial sums of any such series are bounded in their totality, cf. 7.2.2). In other words: on the condition just mentioned, the choice of coefficients according to (9.6-4)—(9.6-5) is *necessary*, in order that the representation (9.6-2) should be valid.

It is to be expected, however, that the series obtained by entering values (9.6-4) of the coefficients in (9.6-1), i.e.,¹

$$\sum_{n=1}^{\infty} \left[\int_a^b f(t) \varphi_n(t) dt \right] \left[\int_a^b \varphi_n(t)^2 dt \right]^{-1} \varphi_n(x) = \sum_{n=1}^{\infty} \frac{(f, \varphi_n)}{\|\varphi_n\|^2} \varphi_n(x) \quad (9.6-6)$$

furnishes under some suitable assumption indeed the solution of the decomposition (representation) problem, and exactly this circumstance does assure (9.6-6) a particular relevance among the orthogonal series of the form $\sum C_n \varphi_n(x)$.

(9.6-6) is called *the (general) Fourier series of $f(t)$ in terms of the base system $\{\varphi_n\}$* , while the occurring coefficients

$$\gamma_n = \frac{(f, \varphi_n)}{\|\varphi_n\|^2} \quad (n = 1, 2, \dots) \quad (9.6-7)$$

are called *Fourier constants of $f(t)$ relating to $\{\varphi_n(t)\}$* . Since the generic term of (9.6-6) can also be written in the form: $(f, \varphi_n^0) \varphi_n^0(x)$ [cf. (9.4-10)], so the Fourier series is a pendant of the right-hand sum in (9.2-16). If $\|\varphi_n\| = 1$, then γ_n reduces simply to (f, φ_n) , i.e., it signifies the component of f relating to φ_n .

9.64 The main part of the theory of orthogonal series is constituted — as it is understandable by the above — by considerations concerning Fourier series. In the sequel, (9.6-6) will be studied first in full generality, joined with the theory of $L^2(a, b)$. We shall show amongst other matters that the series (9.6-6) really represents the function $f(t)$ in the case of any *complete* base system $\{\varphi_n\}$, if not the classical but the $L^2(a, b)$ -convergence (introduced in the previous section) is used.

On the other hand, we shall also investigate some special types of Fourier series (with particular regard to the trigonometric ones) on the basis of the classical notion of convergence and of certain extensions of the latter. The results in this field will inevitably be more ramifying and less “elegant”, nevertheless these come naturally into prominence from the point of view of applications. Let us stress e.g. that the *Fourier analysis of vibrations* which is entwined with numerous branches of natural sciences and technology, is based in a more or less direct manner on the material of Chapters 12-15.

To facilitate understanding, we include here a short survey of the *most important fundamental notions and their geometric counterparts*, keeping in mind the course of gradual abstraction by which the former concepts have been derived from the latter ones:

¹ Here the change in the notation of independent variable (x instead of t) is motivated by the fact that t also occurs as a variable of integration.

function space $L^2(a, b)$	ordinary space of points or vector field
(φ, ψ) (inner product)	scalar product of two vectors
$\ \varphi\ $ (norm)	absolute value (length) of a vector
orthogonality	perpendicularity of vectors
orthogonal system	system of mutually perpendicular position vectors
orthonormal system	system of mutually perpendicular unit vectors with common initial point
complete orthogonal system	three mutually perpendicular vectors with common initial point
$d[\varphi, \psi]$ (distance)	distance of two points
$L^2(a, b)$ -convergence	convergence of a point set
orthogonal series	linear combination of mutually perpendicular position vectors
Fourier series	representation of a vector by means of its components relating to mutually perpendicular position vectors
Fourier constants	vector components relating to mutually perpendicular position vectors

$$1, \cos t, \sin t, \cos 2t, \sin 2t, \dots, \cos nt, \sin nt, \dots, \quad (10.1-1)$$

we already pointed out at the beginning of the previous chapter, that it is orthogonal in any line segment $[a, a + 2\pi]$ (a arbitrary real).¹ This implies, as is well-known, that the inner product of every two distinct elements of (10.1-1) vanishes; i.e., in detail:

$$\int_a^{a+2\pi} \cos nt \, dt = \int_a^{a+2\pi} \sin nt \, dt = 0 \quad (n = 1, 2, \dots), \quad (10.1-2)$$

furthermore

$$\int_a^{a+2\pi} \cos mt \cdot \sin nt \, dt = 0 \quad (10.1-3)$$

for all positive integers m and n , as well as

$$\int_a^{a+2\pi} \cos mt \cdot \cos nt \, dt = \int_a^{a+2\pi} \sin mt \cdot \sin nt \, dt = 0 \quad (10.1-4)$$

for any pair $m \neq n$ of natural numbers.

In fact, the formulae (10.1-2) are obvious, while (10.1-3)—(10.1-4) follow from the identities (3.4-16). We remark that any of the functions (10.1-1) is periodic with the period 2π ; thus it is sufficient to handle the

¹ The terms (10.1-1) belong trivially to $L^2(a, a + 2\pi)$, since each of them is continuous in $[a, a + 2\pi]$.

SPECIAL ORTHOGONAL SYSTEMS AND SERIES

10.1 The trigonometric base system

10.11 Before going further in the general theory, we have to deal with some special orthogonal systems and series, which have become important in the course of historical development; this is all the more reasonable, since no actual example of a *complete* orthogonal system in L^2 has been given until now. We shall also become acquainted with a comprehensive so-called orthogonalization process which enables us to construct orthogonal systems from any prescribed sequence of elements in the Hilbert space.

As regards the *base system of trigonometric functions*:

$$1, \cos t, \sin t, \cos 2t, \sin 2t, \dots, \cos nt, \sin nt, \dots, \quad (10.1-1)$$

we already pointed out at the beginning of the previous chapter, that it is *orthogonal in any line segment* $[a, a + 2\pi]$ (a arbitrary real).¹ This implies, as is well-known, that the inner product of every two distinct elements of (10.1-1) vanishes; i.e., in detail:

$$\int_a^{a+2\pi} \cos nt \, dt = \int_a^{a+2\pi} \sin nt \, dt = 0 \quad (n = 1, 2, \dots), \quad (10.1-2)$$

furthermore

$$\int_a^{a+2\pi} \cos mt \cdot \sin nt \, dt = 0 \quad (10.1-3)$$

for all positive integers m and n , as well as

$$\int_a^{a+2\pi} \cos mt \cdot \cos nt \, dt = \int_a^{a+2\pi} \sin mt \cdot \sin nt \, dt = 0 \quad (10.1-4)$$

for any pair $m \neq n$ of natural numbers.

In fact, the formulae (10.1-2) are obvious, while (10.1-3)—(10.1-4) follow from the identities (3.4-16). We remark that any of the functions (10.1-1) is periodic with the period 2π ; thus it is sufficient to handle the

¹ The terms (10.1-1) belong trivially to $L^2(a, a + 2\pi)$, since each of them is continuous in $[a, a + 2\pi]$.

system for a special value of a e.g. for $a = 0$ or $a = -\pi$ (which correspond to the basic intervals $[0, 2\pi]$ and $[-\pi, \pi]$, respectively).

Since the square integrals of the terms are:

$$\int_a^{a+2\pi} 1^2 dt = 2\pi,$$

$$\int_a^{a+2\pi} \cos^2 nt dt = \int_a^{a+2\pi} \sin^2 nt dt = \pi \quad (n = 1, 2, \dots), \quad (10.1-5)$$

so that we have for the norms of the system:

$$\|1\| = \sqrt{2\pi}, \quad \|\cos nt\| = \|\sin nt\| = \sqrt{\pi} \quad (n = 1, 2, \dots),$$

the *normalized form* of (10.1-1) may be written as

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{1}{\sqrt{\pi}} \cos t, \quad \frac{1}{\sqrt{\pi}} \sin t, \dots, \quad \frac{1}{\sqrt{\pi}} \cos nt, \quad \frac{1}{\sqrt{\pi}} \sin nt, \dots \quad (10.1-6)$$

10.12 *The trigonometric base system is complete* in the sense formulated in 9.43; the verification of this assertion, however, also requires some deeper expedients.

We begin by an elementary observation, relevant in itself too:

10.1.1 *If $M(t)$ is continuous in a finite interval $[a, b]$ and such that*

$$\int_a^b M(t)\eta(t) dt = 0 \quad (10.1-7)$$

holds for every continuous function $\eta(t)$ vanishing at $t = a$ and $t = b$, then $M(t)$ is identically zero in $[a, b]$.¹

Proof. Assume that there exists a point $c \in (a, b)$ e.g. with $M(c) > 0$, then it is easy to find an allowed function $\eta(t)$ which does not satisfy (10.1-7). The continuity of $M(t)$ assures the existence of a neighbourhood $(c-\varepsilon, c+\varepsilon) \subset (a, b)$ where $M(t)$ exceeds the value $m = M(c)/2$. If we choose for $\eta(t)$ the function taking at c the value m , vanishing in the intervals $[a, c-\varepsilon]$ and $[c+\varepsilon, b]$, further being linear on $[c-\varepsilon, c]$ as well as on $[c, c+\varepsilon]$, then we get:

$$\int_a^b M(t)\eta(t) dt = \int_{c-\varepsilon}^{c+\varepsilon} M(t)\eta(t) dt > m \int_{c-\varepsilon}^{c+\varepsilon} \eta(t) dt = \varepsilon m^2 > 0,$$

¹ It is easy to see that it suffices for the validity of 10.1.1 to require (10.1-7) only for *continuously differentiable* functions $\eta(t)$. The proposition thus obtained is the so-called *Lagrange's lemma*, often used in the *calculus of variations*.

in contradiction to (10.1-7). One can similarly show, that the condition $M(c) < 0$ [$c \in (a, b)$] is not consistent with the premises of our proposition.

Consequently, the latter implies $M(t) \equiv 0$ ($a < t < b$) which must also hold at the end points by reason of continuity.

* * *

10.13 Next it is advisable to state the completeness of (10.1-1) in the following form:

10.1.2 If a function $f(t) \in L^2(a, a + 2\pi)$ fulfils the conditions

$$\int_a^{a+2\pi} f(t) dt = 0,$$

$$\int_a^{a+2\pi} f(t) \cos nt dt = 0, \quad \int_a^{a+2\pi} f(t) \sin nt dt = 0$$

$$(n = 1, 2, \dots) \quad (10.1-8)$$

so $f(t) = 0$ for almost every point $t \in (a, a + 2\pi)$.¹

P r o o f. Let $\Phi(x) = \int_a^x f(t) dt$; we see at once that

$$F(x) = \Phi(x) - \frac{1}{2\pi} \int_a^{a+2\pi} \Phi(t) dt$$

is the only integral function of $f(t)$ for which

$$\int_a^{a+2\pi} F(t) dt = 0. \quad (10.1-9)$$

This $F(x)$ is continuous in $[a, a + 2\pi]$ by 7.6.4, and its value at the end points are equal:

$$F(a + 2\pi) = F(a) = -\frac{1}{2\pi} \int_a^{a+2\pi} \Phi(t) dt. \quad (10.1-10)$$

By (10.1-8) we can conclude that $F(t)$ is also orthogonal in $[a, a + 2\pi]$ to any element of the system (10.1-1). Namely, (10.1-9) indicates the

¹ 10.1.2 also remains valid in the case of $f \in L(a, a + 2\pi)$, since this condition assures the existence of the integrals (10.1-8) by $|\cos nt| \leq 1$, $|\sin nt| \leq 1$ as well.

orthogonality to 1; furthermore on the basis of (10.1-10), we obtain by integration by parts:

$$\begin{aligned}\int_a^{a+2\pi} F(t) \cos nt \, dt &= \left[F(t) \frac{\sin nt}{n} \right]_a^{a+2\pi} - \frac{1}{n} \int_a^{a+2\pi} f(t) \sin nt \, dt = 0, \\ \int_a^{a+2\pi} F(t) \sin nt \, dt &= \left[F(t) \frac{-\cos nt}{n} \right]_a^{a+2\pi} + \frac{1}{n} \int_a^{a+2\pi} f(t) \cos nt \, dt = 0 \\ (n = 1, 2, \dots).\end{aligned}$$

Thus we can write:

$$\int_a^{a+2\pi} F(t) S(t) \, dt = 0, \quad (10.1-11)$$

$S(t)$ denoting an arbitrary trigonometric polynomial [cf. (3.4-2)].

But owing to 3.4.1 (I), any function $g(t)$ which is continuous in $[a, a+2\pi]$ and takes the same value at points a and $a+2\pi$, is uniformly approximable on $[a, a+2\pi]$ by suitable trigonometric polynomials $S_1(t), S_2(t), \dots$; therefore

$$\int_a^{a+2\pi} F(t) g(t) \, dt = \lim_{n \rightarrow \infty} \int_a^{a+2\pi} F(t) S_n(t) \, dt = 0,$$

or by (10.1-9)

$$\int_a^{a+2\pi} F(t) [g(t) - g(a)] \, dt = 0. \quad (10.1-12)$$

Hence it results from the use of 10.1.1 that $F(t) \equiv 0$ ($a \leq t \leq 2\pi$), i.e., at a.e. point $t \in L^2(a, a+2\pi)$ we have $f(t) = \Phi'(t) = F'(t) = 0$, as asserted above.¹

* * *

If we split up the system (10.1-1) into two parts, and consider the so-called *cosine system* $\cos nt$ ($n = 0, 1, 2, \dots$) and *sine system* $\sin nt$ ($n = 1, 2, \dots$) separately — these sequences of functions also turn out to be orthogonal on any segment of length 2π , but do not possess the property of completeness.

It can simply be shown on the above model that *both systems* $1, \cos t, \cos 2t, \dots$ and $\sin t, \sin 2t, \dots$ are complete and orthogonal in the interval $[0, \pi]$.

¹ The completeness of the system (10.1-1) can also be verified, of course, in another way, e.g. on the basis of a pertinent general criterion due to A. RÉNYI. (Cf. Rényi [1].)

10.2 The ordinary (trigonometric) Fourier series and its complex form

10.21 Let us form the Fourier series pertaining to the base system (10.1-1).

According to (9.6-7) and (10.1-5), the Fourier constants of a function $f(t) \in L^2(a, a+2\pi)$ relating to 1, and $\cos nt$, $\sin nt$ ($n \geq 1$) are:

$$\alpha_0 = \frac{1}{2\pi} \int_a^{a+2\pi} f(t) dt, \quad (10.2-1)$$

$$\left. \begin{aligned} \alpha_n &= \frac{1}{\pi} \int_a^{a+2\pi} f(t) \cos nt dt \\ \beta_n &= \frac{1}{\pi} \int_a^{a+2\pi} f(t) \sin nt dt \end{aligned} \right\} \quad (n = 1, 2, \dots) \quad (10.2-2)$$

$$(10.2-3)$$

These formulae — found heuristically by L. EULER (about 1750) and J. B. FOURIER (1807) — are mostly mentioned in the literature as *Euler-Fourier coefficient representations*.

The coefficients α_n ($n = 0, 1, 2, \dots$) are the so-called *cosine constants*, the coefficients β_n ($n = 1, 2, \dots$) are the *sine constants*; all of $\alpha_0, \alpha_1, \beta_1, \alpha_2, \beta_2, \dots$ are called *Fourier constants* or *Fourier coefficients* of $f(t)$. The short name of the series with coefficients (10.2-1)–(10.2-3) i.e., of

$$\begin{aligned} \alpha_0 + \sum_{n=1}^{\infty} (\alpha_n \cos nx + \beta_n \sin nx) &= \left(\frac{1}{2\pi} \int_a^{a+2\pi} f(t) dt \right) + \\ &+ \sum_{n=1}^{\infty} \left[\left(\frac{1}{\pi} \int_a^{a+2\pi} f(t) \cos nt dt \right) \cos nx + \right. \\ &\left. + \left(\frac{1}{\pi} \int_a^{a+2\pi} f(t) \sin nt dt \right) \sin nx \right] \end{aligned} \quad (10.2-4)$$

is *ordinary Fourier series* of $f(t)$.¹

Note that the logical connection between (10.2-4) and more general notion of trigonometric series is somewhat similar to that between Taylor series and power series; it is expected also from (10.2-4) that it represents the function f on a certain (rather general) condition.

¹ We note that a may be any real number; in the applications the values $a = 0$ or $a = -\pi$ are preferred.

10.22 We know that every trigonometric series $(A_n, B_n \text{ real})$

$$A_0 + \sum_{n=1}^{\infty} (A_n \cos nx + B_n \sin nx) \quad (10.2-5)$$

may be regarded as the real part of a *power series with complex coefficients*, namely of

$$A_0 + \sum_{n=1}^{\infty} (A_n - iB_n)z^n \quad (z = e^{ix}).^1 \quad (10.2-6)$$

If the real part of the last generic term is written as the half of $(A_n - iB_n)z^n + \overline{(A_n - iB_n)z^n}$ or if we apply immediately Euler's formulae in (10.2-5):

$$\cos nx = \frac{1}{2}(e^{inx} + e^{-inx}), \quad \sin nx = \frac{1}{2i}(e^{inx} - e^{-inx}),$$

so the series (10.2-5) becomes

$$A_0 + \sum_{n=1}^{\infty} (A_n e^{inx} + A_{-n} e^{-inx}) = \sum_{k=-\infty}^{\infty} A_k e^{ikx}, \quad (10.2-7)$$

where

$$A_k = \begin{cases} A_0 & \text{for } k = 0 \\ \frac{1}{2}(A_k - iB_k) & \text{for } k > 0, \\ \frac{1}{2}(A_{-k} + iB_{-k}) & \text{for } k < 0. \end{cases} \quad (10.2-8)$$

In case of a Fourier series, i.e., if $A_k = \alpha_k$, $B_k = \beta_k$, then all the coefficients can be given by a single formula [cf. (10.2-1)–(10.2-3)]:

$$A_k = \frac{1}{2\pi} \int_a^{a+2\pi} f(t) e^{-ikt} dt \quad (k = 0, \pm 1, \pm 2, \dots). \quad (10.2-9)$$

Thus (10.2-7) turns into the series

$$\sum_{k=-\infty}^{\infty} \left[\frac{1}{2\pi} \int_a^{a+2\pi} f(t) e^{-ikt} dt \right] e^{ikx}, \quad (10.2-10)$$

which is called the *complex form of the ordinary Fourier series*.

¹ It is also justified to introduce a special name for the imaginary part of (10.2-6), i.e., for the series

$$\sum_{n=1}^{\infty} (A_n \sin nx - B_n \cos nx); \quad (*)$$

this is the so-called *conjugate series*. If (10.2-5) is the Fourier series of a function $f \in L^2(a, a+2\pi)$, then (*) can be regarded as the Fourier series of another function $\tilde{f} \in L^2(a, a+2\pi)$.

10.23 It is often easier to use the latter series, than the real form (10.2-4); besides (10.2-10) realizes a close connection between the notions of ordinary Fourier series and *Laurent's series*. In fact, if some function $G(z)$ of one complex variable is regular in an annulus T whose boundary consists of two circles concentric with $|\zeta - z_0| = R$ and which contains the latter one, then — as is well-known from complex analysis — we have the so-called *Laurent expansion*

$$G(z) = \sum_{k=-\infty}^{\infty} c_k (z - z_0)^k \quad (z \in T) \quad (10.2-11)$$

with the coefficients

$$\begin{aligned} c_k &= \frac{1}{2\pi i} \oint_{|\zeta - z_0| = R} G(\zeta) (\zeta - z_0)^{-(k+1)} d\zeta = \\ &= \frac{1}{2\pi} R^{-k} \int_a^{a+2\pi} G(z_0 + Re^{it}) e^{-ikt} dt \quad (k = 0, \pm 1, \dots) \end{aligned} \quad (10.2-12)$$

(Here a denotes an arbitrary real number.) Since $z_0 = 0$, $R = 1$ the above (10.2-11) transforms into the series (10.2-10) with the notations $G(e^{it}) = f(t)$, $z = e^{ix}$, we may point out: *any ordinary Fourier series is nothing else but a Laurent series taken along the unit circle.*

Separate examination of the ordinary Fourier series does not become superfluous. 1) We can investigate (10.2-10) by complex function theoretical tools only in the case if f had been expanded analytically into a suitable regular function which is a strong restriction. 2) As shown by (10.2-7), the study of the partial sums of a trigonometric series is equivalent to the discussion of “symmetric” sums of the following type: $\sum_{k=-N}^N A_k e^{ikx}$, while

in the theory of Laurent's series the simultaneous convergence of the “main part” $\sum_{k=-1}^{\infty} c_k (z - z_0)^k$ and the “accessory part” is required — i.e., a stronger condition is established.

The investigation of the behaviour of power series, Laurent series along some circle requires the finer theory of trigonometric and ordinary Fourier series, respectively, if the circle in question lies on the *boundary* of the domain of convergence.

10.3 Legendre polynomials

10.31 The sequence of base functions of a power series, i.e.,

$$1, t, t^2, \dots \quad (10.3-1)$$

obviously does not constitute an orthogonal system, whatever may be the fundamental interval. On the other hand, appropriate finite linear combinations (with real coefficients) of these functions, i.e., certain *entire rational functions*

$$p_0(t) = c_{00}, \quad p_1(t) = c_{10} + c_{11}t, \quad p_2(t) = c_{20} + c_{21}t + c_{22}t^2, \dots$$

may already be pairwise orthogonal to each other in some prescribed *finite* interval $[a, b]$; naturally on the condition that $c_{nn} \neq 0$ ($n = 0, 1, 2, \dots$), i.e., that for any n , $p_n(t)$ is a polynomial of the n -th degree. In fact, for the validity of the orthogonality relations

$$\int_a^b p_m(t)p_n(t) dt = 0 \quad (m < n; m, n = 0, 1, 2, \dots), \quad (10.3-2)$$

it is sufficient that we should have for any fixed $n \geq 0$

$$\int_a^b t^k p_n(t) dt = 0 \quad (k = 0, 1, \dots, n-1), \quad (10.3-3)$$

since every integral of the form (10.3-2) is a linear combination of integrals of the type (10.3-3); and the condition (10.3-3) refers to a system of n linear equations with the n unknowns c_{nk}/c_{nn} ($k = 0, 1, \dots, n-1$), because

$$\begin{aligned} \int_a^b t^k p_n(t) dt &= \frac{b^{k+1} - a^{k+1}}{k+1} c_{n0} + \\ &+ \frac{b^{k+2} - a^{k+2}}{k+2} c_{n1} + \dots + \frac{b^{n+k+1} - a^{n+k+1}}{n+k+1} c_{nn}. \end{aligned}$$

10.32 In what follows, we restrict ourselves to the case of the fundamental interval $[-1, 1]$. This is not a significant limitation, since any finite segment $a \leq t \leq b$ can be transferred by means of the linear transformation $x = [2t - (b+a)]/(b-a)$ into the interval $-1 \leq x \leq 1$. Thus, if a system of polynomials $P_n(x)$ ($n = 0, 1, \dots$) is orthogonal in the interval $-1 \leq x \leq 1$, then the polynomials

$$p_n(t) = P_n\left(\frac{2t - b - a}{b - a}\right) \quad (n = 0, 1, 2, \dots)$$

constitute an orthogonal system relating to the segment $a \leq t \leq b$.

It is easy to see that the pendant of (10.3-3), i.e., the equations

$$\int_{-1}^1 x^k P_n(x) dx = 0 \quad (k = 0, 1, \dots, n-1) \quad (10.3-4)$$

are satisfied, if we write instead of $P_n(x)$ the n -th derivative of a polynomial which has zeros of order n at $x = -1$ and $x = 1$; i.e., if $P_n(x)$ is of the form

$$[(x+1)^n(x-1)^n]^{(n)} = [(x^2-1)^n]^{(n)},$$

eventually multiplied by a constant. Namely a $(k+1)$ -fold integration by parts yields:

$$\int_{-1}^1 x^k (x^2-1)^n dx = (-1)^{k+1} \int_{-1}^1 (x^k)^{k+1} [(x^2-1)^n]^{(n-k-1)} dx = 0,$$

considering also that the $(k+1)$ -th derivative of x^k vanishes identically.

Consequently, the polynomials

$$P_n(x) = C_n [(x^2-1)^n]^{(n)} \quad (C_n \neq 0; n = 0, 1, 2, \dots)$$

form an orthogonal system relating to $[-1, 1]$. In the pertinent literature, the coefficients C_n are fixed mostly in such a way, that we require $P_n(1) = 1$ ($n = 0, 1, 2, \dots$), which is fulfilled by

$$[(x+1)^n(x-1)^n]_{x=1}^{(n)} = n! 2^n = 2 \cdot 4 \dots (2n) = (2n)!!$$

(Leibniz formula) in the case

$$C_n = \frac{1}{(2n)!!}.$$

The entire rational functions thus obtained, i.e.,

$$P_n(x) = \frac{1}{(2n)!!} [(x^2-1)^n]^{(n)} \quad (n = 0, 1, 2, \dots), \quad (10.3-5)$$

satisfying the condition $P_n(1) = 1$, are the so-called *Legendre polynomials*.¹

The representation (10.3-5) is usually called — to distinguish it from other ones — the *Rodrigues formula* for $P_n(x)$.²

¹ Cf. Legendre [1].

² Rodrigues [1].

In particular:

$$P_0(x) = 1, P_1(x) = x, P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}, P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x, \dots,$$

and the above definition implies that the coefficient of x^n in $P_n(x)$ has the value:

$$\frac{1}{(2n)!!} 2n(2n-1)\dots(n+1) = \frac{(2n-1)!!}{n!}. \quad (10.3-6)$$

10.33 Next we deal with some *fundamental properties of Legendre polynomials*:

10.3.1 (I) *Apart from a constant factor, $P_n(x)$ is the only entire rational function which is orthogonal in $[-1, 1]$ to all polynomials of degree $< n$.* (II) $P_n(x)$ is an even or an odd function according to an even and odd degree n , respectively. (III) $P_n(x)$ has for $n \geq 1$ n distinct real zeros, all lying in the interval $(-1, 1)$. (IV) The orthogonal system $P_n(x)$ ($n = 0, 1, \dots$) is complete relating to $(-1, 1)$.¹

P r o o f. (I) Suppose that $p_n(x)$ is such a polynomial of the n -th degree which is orthogonal in $[-1, 1]$ to each polynomial of the some less degree; then at the same time:

$$\int_{-1}^1 P_k(x)p_n(x) dx = 0 \quad (k = 0, 1, \dots, n-1). \quad (10.3-7)$$

On the other hand, $p_n(x)$ may be written with suitable coefficients $c_0, c_1, \dots, c_n$ in the form:

$$p_n(x) = c_0P_0(x) + c_1P_1(x) + \dots + c_nP_n(x) \quad (c_n \neq 0), \quad (10.3-8)$$

since the functions $1, x, \dots, x^n$ can also be represented as linear combinations of $P_0(x), P_1(x), \dots, P_n(x)$. Multiplying both sides by $P_k(x)$ and integrating in $(-1, 1)$, we get on the basis of (10.3-7):

$$c_k \int_{-1}^1 P_k(x)^2 dx = \int_{-1}^1 P_k(x)p_n(x) dx = 0, \\ c_k = 0 \quad (k = 0, 1, \dots, n-1).$$

Therefore (10.3-8) yields indeed $p_n(x) = c_nP_n(x)$.

¹ Property (I) is evidently equivalent to the orthogonality of $P_n(x)$ to any of the functions $1, x, \dots, x^{n-1}$ [cf. (10.3-4)]. From (III) we see that every zero of $P_n(x)$ is necessarily simple. Since the zeros are also points of changing sign, we may say that the Legendre polynomials "reflect" $\sin nx$ resp. $\cos nx$ ($n = 0, 1$) to a certain extent.

(II) follows immediately from the form (10.3-5) of $P_n(x)$, because the expansion of $(x^2 - 1)^n$ according to the binomial theorem plainly contains only powers of x with *even* exponents.

Assertion (III) is a consequence of the fact that $(x^2 - 1)^n$ vanishes together with its $n - 1$ derivatives at points $x = \pm 1$. Hence, by iterated application of Rolle's theorem, we can conclude the existence of one zero of $[(x^2 - 1)^n]'$, of two zeros of $[(x^2 - 1)^n]''$ etc., finally of n distinct zeros of $[(x^2 - 1)^n]^{(n)}$ in the interior of $[-1, 1]$.

As regards (IV), it can be reduced to the proposition 10.1.2 on the completeness of (10.1-1). Let $g(x)$ denote any function square integrable in $(-1, 1)$ which is there orthogonal to all polynomials $P_n(x)$ ($n = 0, 1, \dots$), so that we have

$$\int_{-1}^1 g(x)x^n dx = 0 \quad (n = 0, 1, \dots) \quad (10.3-9)$$

Consider now the (L)-integral of $g(x)e^{-iux} = g(x)\cos(ux) - ig(x)\sin(ux)$ on the interval $-1 \leq x \leq 1$, u denoting a real parameter; and expand it in terms of the powers of u . If we use the uniform convergence on any finite segment of the exponential series, further (10.3-9), it results for every real value of u :

$$\begin{aligned} & \int_{-1}^1 g(x)e^{-iux} dx = \\ &= \int_{-1}^1 g(x) dx - \frac{i u}{1!} \int_{-1}^1 g(x)x dx + \frac{u^2}{2!} \int_{-1}^1 g(x)x^2 dx + \dots = 0; \end{aligned}$$

and hence by putting $x = t/\pi$, $u = n\pi$, $f(t) = g(t/\pi)$:

$$\int_{-\pi}^{\pi} f(t)e^{-int} dt = 0 \quad (n = 0, \pm 1, \pm 2, \dots). \quad (10.3-10)$$

Since (10.3-10) is equivalent to the conditions (10.1-8), we obtain by 10.1.2 that $f(t) = 0$ in $(-\pi, \pi)$, i.e., $g(x) = 0$ in $(-1, 1)$ almost everywhere.

* * *

10.34 Beside the explicit definition (10.3-5), we have also an important implicit way of representation for the polynomials $P_n(x)$ ($n = 0, 1, 2, \dots$):

10.3.2 For any pair of values x and u satisfying together to the restriction $|u|^2 + 2|x||u| < 1$ ¹ the expansion

$$(1 - 2xu + u^2)^{-1/2} = \sum_{n=1}^{\infty} P_n(x)u^n \quad (10.3-11)$$

holds, therefore $P_n(x)$ is the coefficient of index n of the Taylor expansion in terms of powers of u of the function on the left. In this sense, $(1 - 2xu + u^2)^{-1/2}$ is the generating function of Legendre polynomials.²

P r o o f. For the sake of brevity write $u^2 - 2xu = Y$. Since our assumption implies $|Y| \leq |u|^2 + 2|x||u| < 1$, $(1 + Y)^{-1/2} = (1 - 2xu + u^2)^{-1/2}$ can be given in terms of Y^n ($n = 0, 1, 2, \dots$), and the binomial series thus obtained may be rearranged in terms of powers of u , by the absolute convergence of the former series.

Consequently we get:

$$\begin{aligned} (1 - 2xu + u^2)^{-1/2} &= \\ &= 1 + \binom{-1/2}{1} Y + \binom{-1/2}{2} Y^2 + \dots = \\ &= R_0(x) + R_1(x)u + R_2(x)u^2 + \dots, \end{aligned} \quad (10.3-12)$$

where (by the binomial theorem) $R_n(x)$ is a polynomial of the n -th degree of x . We have still to verify $R_n(x) \equiv P_n(x)$ ($n = 0, 1, 2, \dots$).

To that end, let us multiply by x^k ($k \geq 0$, integer) in (10.3-12) and integrate from -1 to $+1$. The termwise integration, being permitted by uniform convergence, results:

$$\int_{-1}^1 x^k (1 + Y)^{-1/2} dx = \int_{-1}^1 x^k R_0(x) dx + u \int_{-1}^1 x^k R_1(x) dx + \dots; \quad (10.3-13)$$

and the integral on the left becomes by the substitution $x = -\xi + u(1 - \xi^2)/2$:

$$\int_{-1}^1 \left[\frac{1}{2}u(1 - \xi^2) - \xi \right]^k d\xi$$

which is plainly a polynomial of the k -th degree of the parameter u .

¹ This inequality is obviously fulfilled for any fixed x , if only $|u|$ is sufficiently small.

² As is well-known, we say that some $G(x, u)$ is the *generating function* of a function sequence $\{g_n(x)\}$, if in an expansion of G (mostly in a power series representation of it) the functions $g_1(x), g_2(x), \dots$ occur as successive coefficients.

Hence we see that the series on the right-hand side in (10.3-13) breaks off at the term containing u^k , i.e.,

$$\int_{-1}^1 x^k R_n(x) dx = 0 \quad (n = k + 1, k + 2, \dots);$$

so we conclude by 10.3.1 (I) that for any fixed n , $R_n(x)$ can differ from $P_n(x)$ at most by a constant factor. But the first term of (10.3-12) turns into $(1-u)^{-1}$ for $x = 1$, so that a comparison of coefficients with the geometric series $1 + u + u^2 + \dots$ furnishes $R_n(1) = 1$ ($n = 0, 1, 2, \dots$).

Therefore, the constant factor just-mentioned is equal to 1 [cf. (10.3-5)], i.e., $R_n(x) = P_n(x)$ as asserted above.

* * *

10.35 In view of the foregoing we can deduce easily *two further relevant properties of the Legendre polynomials*:

10.3.3 For any three consecutive elements $P_{n-1}(x)$, $P_n(x)$, $P_{n+1}(x)$ of the system $\{P_n(x)\}$, the recurrence formula holds:

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x) \quad (n = 1, 2, \dots);^1 \quad (10.3-14)$$

furthermore $y = P_n(x)$ ($n = 0, 1, \dots$) satisfies the differential equation of second order:

$$[(1-x^2)y']' = -n(n+1)y. \quad (10.3-15)$$

P r o o f. If the functions on both sides of the identity

$$(2xu-1)\frac{\partial}{\partial u}(1-2xu+u^2)^{-1/2} + x(1-2xu+u^2)^{-1/2} = u\frac{\partial}{\partial u}[u(1-2xu+u^2)^{-1/2}]$$

are expanded in terms of powers of u on the basis of (10.3-11), the formula (10.3-14) is obtained by comparison of coefficients.

Concerning (10.3-15), the shortest way of verification is the use of the orthogonality of $\{P_n(x)\}$. If we show that $[(x^2-1)P'_n(x)]'$ is orthogonal in $(-1, 1)$ to any of the functions x^k ($k = 0, 1, \dots, n-1$), then it follows again by 13.3.1 (I) that the derivative in question is equal to $P_n(x)$, multiplied by some constant.

¹ According to (10.3-14), the knowledge of $P_0(x)$ and $P_1(x)$ enables us to determine $P_2(x)$, from $P_1(x)$ and $P_2(x)$ we obtain immediately $P_3(x)$, etc.

Now, twofold integration by parts yields:

$$\begin{aligned} \int_{-1}^1 [(x^2 - 1)P'_n(x)]' x^k dx &= -k \int_{-1}^1 (x^2 - 1)P'_n(x)x^{k-1} dx = \\ &= k \int_{-1}^1 [(k+1)x^k - (k-1)x^{k-2}] P_n(x) dx = 0 \quad (k = 0, 1, \dots, n-1), \end{aligned}$$

and equalizing the coefficients of the terms of highest degree, we get

$$[(x^2 - 1)P'_n(x)]' = n(n+1)P_n(x).^1$$

* * *

10.36 Let us calculate the *norm* of $P_n(x)$, which has already been utilized before.

Starting from (10.3-5) and integrating by parts n -times, we deduce the formula:

$$\begin{aligned} \int_{-1}^1 P_n(x)^2 dx &= \frac{(-1)^n}{[(2n)!!]^2} \int_{-1}^1 [(x^2 - 1)^n]^{(2n)} (x^2 - 1)^n dx = \\ &= (-1)^n \frac{(2n-1)!!t}{(2n)!!} \int_{-1}^1 (x^2 - 1)^n dx. \end{aligned} \quad (10.3-16)$$

If the last integral is written in the form

$$\int_{-1}^1 (x^2 - 1)^{n-1} x^2 dx - \int_{-1}^1 (x^2 - 1)^{n-1} dx,$$

then the first term is integrated by parts, we see that

$$\int_{-1}^1 (x^2 - 1)^n dx = -\frac{2n}{2n+1} \int_{-1}^1 (x^2 - 1)^{n-1} dx;$$

the repeated application of this conversion yields:

$$\int_{-1}^1 (x^2 - 1)^n dx = (-1)^n \frac{(2n)!!}{(2n+1)!!} \int_{-1}^1 (x^2 - 1)^0 dx = \frac{(-1)^n \cdot 2(2n)!!}{(2n+1)!}.$$

¹ Another proof of (10.3-15) is based upon the $(n+1)$ -fold differentiation of the identity $(x^2 - 1) \left[(x^2 - 1)^n \right]' = 2nx(x^2 - 1)^n$ and upon the use of Leibniz's formula.

Thus finally (10.3-16) becomes:

$$\int_{-1}^1 P_n(x)^2 dx = 2 \frac{(2n-1)!!}{(2n+1)!!} = \frac{2}{(2n+1)!!}, \quad (10.3-17)$$

whence

$$\|P_n(x)\| = \left(\frac{2}{2n+1} \right)^{1/2} \quad (n = 0, 1, 2, \dots). \quad (10.3-18)$$

Correspondingly, the *normalized form of the system of Legendre polynomials* is the following:

$$\frac{1}{\sqrt{2}}P_0(x), \sqrt{\frac{3}{2}}P_1(x), \sqrt{\frac{5}{2}}P_2(x), \dots, \sqrt{\frac{2n+1}{2}}P_n(x), \dots, \quad (10.3-19)$$

while the *Fourier series relating to* $\{P_n(x)\}$ of a function $f(t) \in L^2(-1, 1)$ [cf. (9.6-6)] may be written:

$$\sum_{n=0}^{\infty} \left[\left(n + \frac{1}{2} \right) \int_{-1}^1 f(t) P_n(t) dt \right] P_n(x). \quad (10.3-20)$$

The shorter name of (10.3-20) is *Legendre series* of $f(t)$.

An interesting *minimum property* of Legendre polynomials is connected with the notion of norm:

10.3.4 Among all the real polynomials of the n -th degree whose leading coefficient¹ has the value (10.3-6), $P_n(x)$ is characterized by the least square integral on $[-1, 1]$.

P r o o f. If $Q_n(x)$ is an arbitrary polynomial of n -th degree with leading coefficient $(2n-1)!!/n!$, then it may be written in the form [cf. (10.3-8)] $Q_n(x) = c_0 + c_1 P_1(x) + \dots + c_{n-1} P_{n-1}(x) + P_n(x)$.

Owing to the orthogonality of $\{P_n(x)\}$ we get:

$$\begin{aligned} \int_{-1}^1 Q_n(x)^2 dx &= c_0^2 + c_1^2 \|P_1(x)\|^2 + \dots + c_{n-1}^2 \|P_{n-1}(x)\|^2 + \|P_n(x)\|^2 = \\ &= \sum_{k=0}^{n-1} \frac{2}{2k+1} c_k^2 + \frac{2}{2n+1}. \end{aligned}$$

¹ By this we mean the coefficient of the term of highest degree.

This sum obviously takes its minimal value if and only if

$$c_0 = c_1 = \dots = c_{n-1} = 0, \quad \text{i.e., when } Q_n(x) \equiv P_n(x).$$

* * *

The Legendre polynomials and the orthogonal series associated with them also play a relevant role in mathematical analysis and mathematical physics; their importance is comparable in several regards to that of trigonometric systems and series, respectively.

10.4 Properties of polynomial systems orthogonal with respect to weight functions

10.41 As a natural further development of a problem raised at the beginning of the previous section, now let us look for such a *sequence of polynomials*

$$p_n(x) = c_{n0} + c_{n1}x + \dots + c_{nn}x^n \quad (c_{nn} \neq 0; n = 0, 1, \dots), \quad (10.4-1)$$

which constitutes an orthogonal system only after multiplication by a prescribed square integrable function $g(x)$, i.e., for which

$$g(x)p_0(x), \quad g(x)p_1(x), \dots, \quad g(x)p_n(x), \dots \quad (10.4-2)$$

is orthogonal in (a, b) . (We have to assume the square integrability of $g(x)x^k$ for every positive integer k). Since the above condition means that

$$\int_a^b \rho(x)p_m(x)p_n(x)dx = 0 \quad (m \neq n; m, n = 0, 1, 2, \dots) \quad (10.4-3)$$

with $\rho(x) = g(x)^2 \geq 0$, physical analogies (e.g. calculation of centres of gravity) suggest introducing a particular denomination: we say that the system $\{p_n(x)\}$ is *orthogonal in (a, b) with respect to the (non-negative) weight function $\rho(x)$* ,¹ if (10.4-3) is fulfilled.

It is to be stressed that the fundamental interval (a, b) can be finite or infinite. In the latter case, the advantage of the concept "weight function" is

¹ Instead of "weight", the term "density" is also used. Orthogonality with respect to weight 1 means of course "ordinary" orthogonality. Furthermore: the orthogonality of a system $\{p_n(x)\}$ with respect to some weight $\rho(x)$ can be replaced by the requirement that $\{\rho(x)^{1/2}p_n(x)\}$ should be orthogonal in the same interval.

especially clear: on an infinite interval, no entire rational function is square integrable, so that in this case the orthogonality of a polynomial system in the ordinary sense is also impossible; on the other hand, the integrals (10.4-3) may exist on an infinite (a, b) too — namely if $\rho(x)$ tends sufficiently rapidly to zero as $|x| \rightarrow \infty$.

Note that (10.4-3) is equivalent to the condition:

$$\int_a^b \rho(x) x^k p_n(x) dx = 0 \quad (k = 0, 1, \dots, n-1) \quad (10.4-4)$$

which can also be formulated as: any $p_n(x)$ should be orthogonal in (a, b) to every polynomial of at most $(n-1)$ degree with respect to the weight $\rho(x)$. Namely it is evident that *any polynomial of the k -th degree can uniquely be expressed as a linear combination of arbitrary polynomials $Q_0(x), Q_1(x), \dots, Q_k(x)$, provided that each of the latter is exactly of the degree given by its index.*

On the same basis it can be seen [cf. (10.3-8)], that in the knowledge of one single system of polynomials $\{p_n(x)\}$ having the property (10.4-3), all the other polynomial systems orthogonal in (a, b) with respect to the weight $\rho(x)$ can be written in the form $C_0, C_1 p_1(x), C_2 p_2(x), \dots$, where $C_0, C_1, C_2, \dots$ are arbitrary constants. Briefly speaking: *in case of polynomial systems, the orthogonality in the extended sense is also a property characterizing the system "essentially" (i.e., apart from constant factors).*

10.42 The question is: how is it possible to construct effectively an orthogonal polynomial system associated with an arbitrary fundamental interval and a weight function. Owing to the observation just mentioned, such a polynomial system becomes uniquely determined by prescribing in each polynomial its leading coefficient; it is advisable to take these coefficients equal to 1. The answer is given by the proposition:

10.4.1 *Let a finite or infinite interval (a, b) and a weight function $\rho(x)[x \in (a, b)]$ be given. Assume that $\rho(x)$ is positive almost everywhere and that its product by $|x|^N$ is (L) -integrable in (a, b) , N denoting a however large positive number.¹*

Then the only polynomial system $\{\tilde{p}_n(x)\}$ which is orthogonal in (a, b) with the weight $\rho(x)$ and all the leading coefficients 1, can be presented as follows:

$$\left. \begin{aligned} \tilde{p}_0(x) &= 1, \quad \tilde{p}_1(x) = x - \tilde{I}_0 \quad \text{and} \\ \tilde{p}_{n+1}(x) &= (x - \tilde{I}_n) \tilde{p}_n(x) - \tilde{K}_n \tilde{p}_{n-1}(x) \quad (n = 1, 2, \dots), \end{aligned} \right\} \quad (10.4-5)$$

¹ In the case of a finite segment (a, b) , it is enough to exchange the latter condition for a simpler one: let $\rho(x)$ itself be (L) -integrable.

where

$$\begin{aligned}\tilde{I}_n &= \frac{\int_a^b x \rho(x) \tilde{p}_n(x)^2 dx}{\int_a^b \rho(x) \tilde{p}_n(x)^2 dx} \quad (n = 0, 1, 2, \dots), \\ \tilde{K}_n &= \frac{\int_a^b \rho(x) \tilde{p}_n(x)^2 dx}{\int_a^b \rho(x) \tilde{p}_{n-1}(x)^2 dx} \quad (n = 1, 2, \dots)^1\end{aligned} \quad (10.4-6)$$

P r o o f. First of all we can state that the assumptions on $\rho(x)$ assure the existence of every integral occurring in (10.4-6), as well as the positivity of the denominators.

The linear function $\tilde{p}_1(x)$ is trivially orthogonal to $\tilde{p}_0(x)$ in (a, b) , with respect to the weight $\rho(x)$. Suppose that $\tilde{p}_{n-1}(x)$ and $\tilde{p}_n(x)$ are orthogonal in the above sense to all the terms with lower indices of the system $\tilde{p}_0(x), \tilde{p}_1(x), \dots$. Then we get from the recurrence formula (10.4-5) after multiplication by $\rho(x) \tilde{p}_n(x)$ and integration on (a, b) :

$$\begin{aligned}\int_a^b \rho(x) \tilde{p}_{n+1}(x) \tilde{p}_n(x) dx &= \\ &= \int_a^b \rho(x) (x - \tilde{I}_n) \tilde{p}_n(x)^2 dx - \tilde{K}_n \int_a^b \rho(x) \tilde{p}_{n-1}(x) \tilde{p}_n(x) dx = \\ &= \int_a^b \rho(x) x \tilde{p}_n(x)^2 dx - \tilde{I}_n \int_a^b \rho(x) \tilde{p}_n(x)^2 dx = 0 \quad (n = 1, 2, \dots).\end{aligned} \quad (10.4-7)$$

Writing $\tilde{p}_{n-1}(x)$ instead of $\tilde{p}_n(x)$, we obtain for any positive integer n :

$$\begin{aligned}\int_a^b \rho(x) \tilde{p}_{n+1}(x) \tilde{p}_{n-1}(x) dx &= \\ &= \int_a^b \rho(x) x \tilde{p}_n(x) \tilde{p}_{n-1}(x) dx - \tilde{I}_n \int_a^b \rho(x) \tilde{p}_n(x) \tilde{p}_{n-1}(x) dx - \\ &\quad - \tilde{K}_n \int_a^b \rho(x) \tilde{p}_{n-1}(x)^2 dx = \\ &= \int_a^b \rho(x) \tilde{p}_n(x)^2 dx - \tilde{K}_n \int_a^b \rho(x) \tilde{p}_{n-1}(x)^2 dx = 0.\end{aligned} \quad (10.4-8)$$

¹ It is obvious that the successive application of (10.4-5)—(10.4-6) enables us to calculate any prescribed element of the system $\{\tilde{p}_n(x)\}$.

In the last calculation, we utilized that if we write the polynomial $x\tilde{p}_{n-1}(x)$ as a linear combination of $\tilde{p}_n(x)$, $\tilde{p}_{n-1}(x)$, ..., $\tilde{p}_0(x)$, so the first term has the coefficient 1, and any other term yields 0 when multiplied by $\rho(x)\tilde{p}_n(x)$, further integrated on (a, b) — because of our condition of orthogonality.

Finally for $k < n - 1$, we obtain from (10.4-5) immediately:

$$\begin{aligned} & \int_a^b \rho(x) \tilde{p}_{n+1}(x) \tilde{p}_k(x) dx = \\ & = \int_a^b \rho(x) (x - \tilde{I}_n) \tilde{p}_k(x) \tilde{p}_n(x) dx - \\ & - \tilde{K}_n \int_a^b \rho(x) \tilde{p}_{n-1}(x) \tilde{p}_k(x) dx = 0, \end{aligned} \quad (10.4-9)$$

since the degree of $(x - \tilde{I}_n) \tilde{p}_k(x)$ is at most $n - 1$, while that of $\tilde{p}_k(x)$ is at most $n - 2$.

The formulae (10.4-7)–(10.4-9) imply that also $\tilde{p}_{n+1}(x)$ is orthogonal in (a, b) with respect to $\rho(x)$ to each previous term of the system $\{\tilde{p}_n(x)\}$. Thus we accomplished the proof by induction.

* * *

As far as an *arbitrary orthogonal polynomial system* of the form (10.4-1) [with fundamental interval (a, b) and weight $\rho(x)$] is concerned, we have according to the above:

$$p_n(x) = A_n \tilde{p}_n(x) \quad (n = 0, 1, 2, \dots), \quad (10.4-10)$$

where $A_n = c_{nn} \neq 0$ denotes the leading coefficient of $p_n(x)$.

So introducing the further notations:

$$\begin{aligned} I_n &= \frac{\int_a^b x \rho(x) p_n(x)^2 dx}{\int_a^b \rho(x) p_n(x)^2 dx} \quad (n = 0, 1, 2, \dots), \\ K_n &= \frac{\int_a^b \rho(x) p_n(x)^2 dx}{\int_a^b \rho(x) p_{n-1}(x)^2 dx} \quad (n = 1, 2, \dots), \end{aligned} \quad (10.4-11)$$

the following *recurrence formula* issues for $\{p_n(x)\}$:

$$\begin{aligned} A_n^2 A_{n+1}^{-1} p_{n+1}(x) &= (x - I_n) A_n p_n(x) - K_n A_{n-1} p_{n-1}(x) \\ &\quad (n = 1, 2, \dots). \end{aligned} \quad (10.4-12)$$

Hence, starting from $p_0(x) = A_0$, $p_1(x) = A_1(x - I_1)$, we get $p_2(x)$, next by the use of $p_1(x)$ and $p_2(x)$ we calculate $p_3(x)$, and so on. [Cf. (10.3-14).]

We add that by

$$\left\| \rho(x)^{1/2} p_n(x) \right\|^2 = \int_a^b \rho(t) p_n(t)^2 dt \quad (n = 0, 1, 2, \dots),$$

the *normalized form* of the orthogonal system $\left\{ \rho(x)^{1/2} p_n(x) \right\}$ is:

$$\left[\int_a^b \rho(t) p_n(t)^2 dt \right]^{-1/2} \rho(x)^{-1/2} p_n(x) \quad (n = 0, 1, 2, \dots). \quad (10.4-13)$$

10.43 Turning to the *fundamental properties of general orthogonal polynomials* (10.4-10), we establish [cf. 10.3.1 (III)-(IV)]:

10.4.2 (I) Any polynomial $p_n(x) = A_n \tilde{p}_n(x)$ ($n = 1, 2, \dots$) has n distinct zeros in (a, b) . — (II) The orthogonal system $\left\{ \rho(x)^{1/2} p_n(x) \right\}$ is complete relating to (a, b) , if there exists a number $M > 0$ such that $e^{M \cdot |x|} \rho(x)$ is (L) -integrable in this interval.¹

P r o o f. The polynomial $p_n(x)$ has anyway a simple real zero in the interval (a, b) , because otherwise the integral of $\rho(x) p_n(x)$ on (a, b) would be $\neq 0$, contrasted with (10.4-3). Let us denote by s the number of all the points in (a, b) where $p_n(x)$ has a change of sign, and let $\xi_1, \xi_2, \dots, \xi_s$ be the abscissae of these points. Then in case of $s < n$ we should have

$$\int_a^b \rho(x) p_n(x) (x - \xi_1)(x - \xi_2) \dots (x - \xi_s) dx \neq 0,$$

since the integrand is positive a.e.; and this contradicts (10.4-3). Therefore $s = n$, according to the assertion.²

The statement on completeness can be verified for finite (a, b) in a similar manner as in the case of Legendre polynomials. If (a, b) is infinite, the main concept of argumentation remains the same, but certain minor complementary considerations are necessary.³

* * *

¹ For finite (a, b) , the latter condition is equivalent to $\rho(x) \in L(a, b)$.

² As is well-known, $s > n$ is impossible by the fundamental theorem of algebra.

³ For particulars see e.g. Sz.-Nagy [16].

The *minimum property 10.3.4* can immediately be transferred together with the essence of its proof. The polynomial $p_n(x)$ [cf. (10.4–10)] minimizes square integrals of the form

$$\|\rho(x)^{1/2}Q_n(x)\|^2 = \int_a^b \rho(x)Q_n(x)^2 dx, \quad (10.4-14)$$

where $Q_n(x)$ signifies an arbitrary real polynomial of the n -th degree, the leading coefficient of which has a fixed value A_n .

10.44 We want still to insert some general denominations and notations pertaining to the notion of orthogonality with respect to a given weight function $\rho(x) \geq 0$.

The symbol $L^2_{\rho(x)}(a, b)$ (or briefly L^2_{ρ}) usually denotes the class of functions $\varphi(x)$ for which $\rho(x)^{1/2}\varphi(x)$ is square integrable on a segment (a, b) (cf. 9.4). By the norm of a function $\varphi(x) \in L^2_{\rho(x)}(a, b)$ relating to the weight $\rho(x)$, in short its $\rho(x)$ -norm, we mean the norm of the product $\rho(x)^{1/2}\varphi(x)$ in the space $L^2(a, b)$; this notion is denoted by

$$\|\varphi(x)\|_{\rho(x)} = \|\rho(x)^{1/2}\varphi(x)\| = \left[\int_a^b \varphi(x)^2 \rho(x) dx \right]^{1/2}. \quad (10.4-15)$$

We say that a function sequence $\{\varphi(x)\} \in L^2_{\rho(x)}(a, b)$ is orthogonal in (a, b) with the weight $\rho(x)$, briefly $\rho(x)$ -orthogonal, if $\{\rho(x)^{1/2}\varphi_n(x)\}$ is orthogonal there in the ordinary sense. $\{\varphi(x)\}$ is called $\rho(x)$ -complete, if $\{\rho(x)^{1/2}\varphi_n(x)\}$ is complete in $L^2(a, b)$.

The use of the terms just introduced enables us to formulate more concisely all the results concerning general orthogonal polynomial systems. On this basis, it is also suggested that any occurring (L) -integral of the form

$$\int_a^b \varphi(x)^2 \rho(x) dx$$

should be considered alternatively as a *Lebesgue–Stieltjes integral* of the type

$$\int_a^b \varphi(x)^2 dq(x),$$

$q(x)$ denoting an (L) -integral function of $\rho(x)^1$ (cf. 7.6.7, 8.3.3). Since integrals of the latter form can also exist on more general conditions [e.g. the

¹ The assumption on $\rho(x)$ implies that $q(x)$ is absolutely continuous and monotone increasing, furthermore it holds a.e. the relation $q'(x) = \rho(x)$ (cf. 7.6.4).

continuity of $q(x)$ is not required], so the going over to class $L_q(a, b)$ of the pertinent functions φ assures simultaneously certain formal generalization possibilities.

10.5 Jacobi, Laguerre and Hermite polynomials

10.51 As it appears from the above, several important properties of Legendre polynomials can be extended to arbitrary orthogonal polynomials. But obviously only certain special weight functions render possible an *explicit polynomial representation* of Rodrigues type [cf. (10.3-5)] and the characterization by means of some *differential equation* [cf. (10.3-15)], although these are relevant for applications. Now, we want to deal briefly with a number of special orthogonal polynomials which are distinguishable in such a sense.

The simplest pertinent examples: the so-called *Tchebyshev polynomials* can be generated from the systems $\{\cos nt\}$, $\{\sin nt\}$ (both orthogonal in $[0, \pi]$, cf. 10.13) essentially by changes of notation.

Namely the orthogonality relations

$$\int_0^\pi \cos mt \cdot \cos nt \, dt = \int_0^\pi \sin(m+1)t \cdot \sin(n+1)t \, dt = 0$$

$$(m \neq n; m, n = 0, 1, 2, \dots) \quad (10.5-1)$$

yield by the substitution $\cos t = x$:

$$\int_{-1}^1 T_m(x) T_n(x) (1-x^2)^{-1/2} dx =$$

$$= \int_{-1}^1 U_m(x) U_n(x) (1-x^2)^{1/2} dx = 0$$

$$(m \neq n; m, n = 0, 1, 2, \dots), \quad (10.5-2)$$

where $T_n(x)$ signifies $\cos nt$ as a function of $\cos t = x$, and $U_n(x)$ denotes the quotient $\sin(n+1)t / \sin t$ as a function of $\cos t = x$.

Both $T_n(x)$ and $U_n(x)$ are polynomials of the n -th degree, since by Moivre's formula and the binomial theorem, respectively,¹ we have

$$\cos nt + i \sin nt =$$

$$\cos^n t + i \binom{n}{1} \cos^{n-1} t \cdot \sin t - \binom{n}{2} \cos^{n-2} t \cdot \sin^2 t + \dots, \quad (10.5-3)$$

¹ Observe that the right-hand sum in (10.5-3) consists of $n+1$ terms.

so that we get by comparison of the real and imaginary parts:

$$\left. \begin{aligned} \cos nt &= (\cos t)^n - \binom{n}{2} (\cos t)^{n-2} (1 - \cos^2 t) + \dots, \\ \sin nt &= (\sin t) \left[\binom{n}{1} (\cos t)^{n-1} - \binom{n}{3} (\cos t)^{n-3} (1 - \cos^2 t) + \dots \right] \end{aligned} \right\} \quad (10.5-4)$$

i.e.,

$$\left. \begin{aligned} T_n(x) &= x^n - \binom{n}{2} x^{n-2} (1 - x^2) + \binom{n}{4} x^{n-4} (1 - x^2)^2 - \dots, \\ U_n(x) &= (n+1)x^n - \binom{n+1}{3} x^{n-2} (1 - x^2) + \binom{n+1}{5} x^{n-4} (1 - x^2)^2 - \dots \end{aligned} \right\} \quad (10.5-5)$$

In particular:

$$\begin{aligned} T_0(x) &= 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1, \quad T_3(x) = 4x^3 - 3x, \dots, \\ U_0(x) &= 1, \quad U_1(x) = 2x, \quad U_2(x) = 4x^2 - 1, \quad U_3(x) = 8x^3 - 4x, \dots, \end{aligned}$$

and by collecting the terms containing x^n in (10.5-5), we see that the leading coefficient of $T_n(x)$, $U_n(x)$ is 2^{n-1} and 2^n , respectively. The name of $\{T_n(x)\}$ is: *Tchebyshev polynomials of the first kind*, that of $\{U_n(x)\}$: *Tchebyshev polynomials of the second kind*. Their common interval of orthogonality is $(-1, 1)$, but their weight functions are different: according to (10.5-2), $\rho(x) = 1/\sqrt{1-x^2}$ for $\{T_n(x)\}$ and $\rho(x) = \sqrt{1-x^2}$ for $\{U_n(x)\}$.

It is to be emphasized that the significance of the Tchebyshev polynomials originates from certain *extremal properties* of them, different from that of Legendre polynomials (cf. 10.3.4). Let us mention only the following characterization of $\{T_n(x)\}$:

10.5.1 Among all the real polynomials $Q_n(x)$ of the degree n and with the leading coefficient 2^{n-1} , for $T_n(x)$ the functional

$$M[Q_n] = \max_{-1 \leq x \leq 1} |Q_n(x)| \quad (10.5-6)$$

attains its minimal value.¹

¹ This minimum is found to be 1; so for all polynomials in question, the inequality $M[Q_n] \geq 1$ holds.

Hence originates the concept of the so-called *Tchebyshev* (or *extended uniform*) *approximation*: we say that some function F is best approached by another function f on a prescribed set E in Tchebyshev's sense, if there the maximum of the deviation between f and F is possibly minimal.

10.52 On the basis of (10.5-2) it is justified to study the polynomials $^{(\alpha)}P_n(x)$ ($n = 0, 1, 2, \dots$) — as a common generalization of the system $\{T_n(x)\}$, $\{U_n(x)\}$ — which are orthogonal in $(-1, 1)$ with weight $\rho(x) = (1 - x^2)^\alpha$ ($\alpha > -1$). The system of these so-called *ultraspherical* (or *Gegenbauer*) *polynomials* contains the Legendre polynomials as well ($\alpha = 0$).

Going still further, we can consider the so-called *Jacobi polynomials* $\{P_n^{(\alpha, \beta)}(x)\}$, orthogonal on $(-1, 1)$ with respect to any weight function of the form $\rho(x) = (1 - x)^\alpha(1 + x)^\beta$ ($\alpha, \beta > -1$),¹ defined explicitly by the *generalized Rodrigues formula*:

$$P_n^{(\alpha, \beta)}(x) = \frac{(-1)^n}{(2n)!!} (1 - x)^{-\alpha} (1 + x)^{-\beta} \left[(1 - x)^{\alpha+n} (1 + x)^{\beta+n} \right]^{(n)} \quad (n = 0, 1, 2, \dots). \quad (10.5-7)$$

By means of (10.5-7), we can discuss the properties of arbitrary Jacobi polynomials in the same way, as happened in the case of Legendre polynomials through (10.3-5). We have

$$P_n^{(\alpha, \alpha)}(x) = {}^{(\alpha)}P_n(x), \quad P_n^{(0, 0)}(x) = P_n(x), \quad (10.5-8)$$

moreover $P^{(-1/2, -1/2)}(x)$ differs from $T_n(x)$ only by a constant factor, and the same holds for $P_n^{(1/2, 1/2)}(x)$, $U_n(x)$.²

By the use of (10.5-7), it can be verified like (10.3-15) without difficulty that $y = P_n^{(\alpha, \beta)}(x)$ for all non-negative integers n satisfies the *differential equation* of second order:

$$\begin{aligned} & \left[(1 - x)^{\alpha+1} (1 + x)^{\beta+1} y' \right]' = \\ & = -n(n + \alpha + \beta + 1) (1 - x)^\alpha (1 + x)^\beta y \\ & \quad (-1 < x < 1) \end{aligned} \quad (10.5-9)$$

which becomes (10.3-15) as $\alpha = \beta = 0$.

¹ The conditions for the parameters α, β are necessary only to assure the (L) -integrability of the weight $(1 - x)^\alpha(1 + x)^\beta$ in $(-1, 1)$.

² The variant of $P_n^{(\alpha, \beta)}(x)$ transformed to the fundamental interval $[0, 1]$ also has an alternative name: *hypergeometric polynomial* of the n -th degree, because it may be considered as a particular case of the *hypergeometric series* (cf. Jacobi [1]): $G(x; A, B, C) = 1 + \frac{A \cdot B}{1 \cdot C} x + \frac{A(A+1)}{1 \cdot 2} \cdot \frac{B(B+1)}{C(C+1)} x^2 + \dots$ where A, B, C are given constants.

Hence we obtain by a straightforward calculation that the derivative of any Jacobi polynomial is again of the same type, aside from a constant factor; more precisely:

$$\left[P_n^{(\alpha, \beta)}(x)\right]' = \frac{\alpha + \beta + n + 1}{2} P_{n-1}^{(\alpha+1, \beta+1)}(x) \quad (n = 1, 2, \dots). \quad (10.5-10)$$

For instance, the Tchebyshev polynomials of the first and second kind are interlinked by the simple relation:

$$T'(x) = nU_{n-1}(x). \quad (10.5-11)$$

10.53 Some special orthogonal polynomials are also attached to certain infinite fundamental intervals. We mention first the so-called *ordinary Laguerre polynomials* $\{L_n(x)\}$, for which $(a, b) = (0, \infty)$ and $\rho(x) = e^{-x}$ as well as their extensions, the *general Laguerre polynomials*, defined by the Rodrigues formula:

$$({}^{(\alpha)}L_n(x) = \frac{1}{n!} x^{-\alpha} e^x \frac{d^n}{dx^n} (x^{n+\alpha} e^{-x}) \quad (\alpha > -1; n = 0, 1, \dots)^1 \quad (10.5-12)$$

The latter constitute an orthogonal system on $(0, \infty)$ with respect to the weight $\rho(x) = x^\alpha e^{-x}$, and turn into $L_n(x)$ for $\alpha = 0$. By the representation (10.5-12), the main properties of $y = ({}^{(\alpha)}L_n(x))$ are the validity of the differential equation

$$(x^{\alpha+1} e^{-x} y')' = -n x^\alpha e^{-x} y \quad (0 < x < \infty) \quad (10.5-13)$$

and of the relation

$$\left[({}^{(\alpha)}L_n(x)\right]' = ({}^{(\alpha+1)}L_{n-1}(x)). \quad (10.5-14)$$

Finally, the whole number line is the fundamental interval of another very usual orthogonal system, the so-called *Hermite polynomials* $\{H_n(x)\}$.² These are orthogonal on $(-\infty, \infty)$ with the weight function $\rho(x) = e^{-x^2}$ and can also be given explicitly by a formula of Rodrigues type:

$$H_n(x) = \frac{(-1)^n}{n!} e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \quad (n = 0, 1, 2, \dots). \quad (10.5-15)$$

It is easy to see that $y = H_n(x)$ is a particular solution of the differential equation

$$(e^{-x^2} y')' = 2n e^{-x^2} y \quad (-\infty < x < \infty) \quad (10.5-16)$$

furthermore we may point out that the differentiation process essentially does not take out of the system $\{H_n(x)\}$, because

$$H'_n(x) = -2H_{n-1}(x). \quad (10.5-17)$$

¹ Laguerre [1].

² Hermite [1].

10.54 All the above systems $\{P_n^{(\alpha,\beta)}(x)\}$, $\{({}^{(\alpha)}L_n(x)\}$, $\{H_n(x)\}$, and the polynomial sequences arising from them when we apply multiplication by constants as well as arbitrary linear transformations $x = At + B$ ($A \neq 0$) are together called *classical orthogonal polynomial systems* and will briefly be denoted by COP.

An orthogonal polynomial system $y = p_n(x)$ ($n = 0, 1, 2, \dots$) belongs to COP if and only if its elements satisfy a *differential equation* of the form

$$[\sigma(x)y']' = -\lambda_n \rho(x)y \quad (a < x < b) \quad (10.5-18)$$

where λ_n is a non-negative constant and the functions $\rho(x)$, $\sigma(x)$ are positive in the fundamental interval (a, b) . [Cf. (10.5-9), (10.5-13), (10.5-16).] Another characteristic property of the system is: in order to have a "classical" orthogonal polynomial system $\{p_n(x)\}$, it is necessary and sufficient that the *system of derivatives* $p'_n(x)$ should again be orthogonal on the same interval. [Cf. (10.5-10), (10.5-14), (10.5-17)].¹

Observe that the coefficients $\rho(x)$, $\sigma(x)$ in the differential equation (10.5-18) always yield the *weight functions* of $\{p_n(x)\}$ and $\{p'_n(x)\}$, respectively; further, every system $\in$ COP is $\rho(x)$ -complete in its fundamental interval by 10.4.2 (II). (Cf. 10.44.)

Summing up, the main data of the above "basic" COP are:

Name	Symbol	(a, b)	$\rho(x)$	$\sigma(x)$
<i>Jacobi polynomials</i> ($\alpha > -1, \beta > -1$)	$P_n^{(\alpha,\beta)}(x)$	$(-1, 1)$	$(1-x)^\alpha(1+x)^\beta$	$(1-x)^{\alpha+1}(1+x)^{\beta+1}$
Ultraspherical ($\alpha = \beta$)	$({}^{(\alpha)}P_n(x)$	$(-1, 1)$	$(1-x^2)^\alpha$	$(1-x^2)^{\alpha+1}$
Tchebyshev I. ($\alpha = \beta = -1/2$)	$T_n(x)$	$(-1, 1)$	$1/\sqrt{1-x^2}$	$\sqrt{1-x^2}$
Tchebyshev II. ($\alpha = \beta = 1/2$)	$U_n(x)$	$(-1, 1)$	$\sqrt{1-x^2}$	$(1-x^2)^{3/2}$
Legendre ($\alpha = \beta = 0$)	$P_n(x)$	$(-1, 1)$	1	$1-x^2$
<i>Laguerre polynomials</i> ($\alpha > -1$)	$({}^{(\alpha)}L_n(x)$	$(0, \infty)$	$x^\alpha e^{-x}$	$x^{\alpha+1} e^{-x}$
Ordinary Laguerre ($\alpha = 0$)	$L_n(x)$	$(0, \infty)$	e^{-x}	$x e^{-x}$
<i>Hermite polynomials</i>	$H_n(x)$	$(-\infty, \infty)$	e^{-x^2}	e^{-x^2}

¹ For details and further developments from these results, see Mikolás [2], [4] and W. Hahn [1].

10.55 Any of the polynomial systems in this table can be derived also from a *generating function* (cf. 10.34).

For every fixed x and for sufficiently small $|u|$ it holds:

$$2^{\alpha+\beta} D(x, u)^{-1} [1 - u + D(x, u)]^{-\alpha} [1 + u + D(x, u)]^{-\beta} = \sum_{n=0}^{\infty} P_n^{(\alpha, \beta)}(x) u^n \quad (10.5-19)$$

with $D(x, u) = \sqrt{1 - 2xu + u^2}$.

Furthermore, we have:

$$(1 - u)^{-(\alpha+1)} e^{-xu(1-u)} = \sum_{n=0}^{\infty} {}^{(\alpha)}L_n(x) u^n \quad (|u| < 1) \quad (10.5-20)$$

and

$$e^{u(2x-u)} = \sum_{n=0}^{\infty} H_n(x) u^n \quad (-\infty < u < \infty). \quad (10.5-21)$$

We mention without proof as well, that the pertinent *square norms* associated with the corresponding weights can also be given in a closed form [cf. (10.4-15)]:

$$\begin{aligned} & \left\| (1-x)^{-\alpha/2} (1+x)^{-\beta/2} P_n^{(\alpha, \beta)}(x) \right\|^2 = \\ & = \frac{2^{\alpha+\beta+1}}{2n + \alpha + \beta + 1} \frac{\Gamma(n + \alpha + 1) \Gamma(n + \beta + 1)}{\Gamma(n + 1) \Gamma(n + \alpha + \beta + 1)}, \end{aligned} \quad (10.5-22)$$

$$\left\| x^{\alpha/2} e^{-x/2} {}^{(\alpha)}L_n(x) \right\|^2 = \Gamma(\alpha + 1) \binom{\alpha + n}{n}, \quad (10.5-23)$$

$$\left\| e^{-x^2/2} H_n(x) \right\|^2 = \frac{2^n}{n!} \sqrt{\pi}, \quad (10.5-24)$$

n denoting an arbitrary non-negative integer.

In (10.5-22) and (10.5-23), $\Gamma(x)$ signifies the well-known *gamma-function* (generalized factorial), whose most natural definition, due to Euler, is as follows:

$$\Gamma(x) = \int_0^{\infty} u^{x-1} e^{-u} du \quad (x > 0), \quad (10.5-25)$$

and whose most elementary properties [from (10.5-25) on the basis of integration by parts] are:

$$\Gamma(x + 1) = x \Gamma(x) \quad (x > 0), \quad (10.5-26)$$

$$\Gamma(n + 1) = n! \quad (n = 0, 1, 2, \dots). \quad (10.5-27)$$

10.56 By the use of (10.5-22)—(10.5-24), also the *Fourier series of an arbitrary function* $\in L^2_{\rho(x)}$ in terms of *Jacobi, Laguerre and Hermite polynomials* can be written. [Cf. (10.3-20).] We obtain especially simple series in case of Tchebyshev polynomials and for the system of ordinary Laguerre polynomials, since

$$\begin{aligned} \left\| (1-x^2)^{-1/4} T_n(x) \right\|^2 &= \int_{-1}^1 T_n(x)^2 (1-x^2)^{-1/2} dx = \\ &= \int_0^\pi \cos^2 nt \, dt = \begin{cases} \frac{\pi}{2} & \text{if } n = 0, \\ \frac{\pi}{2} & \text{if } n = 1, 2, \dots; \end{cases} \end{aligned} \quad (10.5-28)$$

$$\begin{aligned} \left\| (1-x^2)^{1/4} U_n(x) \right\|^2 &= \int_{-1}^1 U_n(x)^2 \sqrt{1-x^2} dx = \\ &= \int_0^\pi \sin^2(n+1)t \, dt = \frac{\pi}{2} \quad (n = 0, 1, 2, \dots); \end{aligned} \quad (10.5-29)$$

$$\left\| e^{-x/2} L_n(x) \right\|^2 = \int_0^\infty L_n(x)^2 e^{-x} dx = 1 \quad (n = 0, 1, 2, \dots). \quad (10.5-30)$$

10.6 Orthogonal systems arising from differential equations of Sturm–Liouville type

10.61 The elements of the classical orthogonal polynomial systems do not “by chance” satisfy relations of the form (10.5-18), but in case of *every* differential equation of the mentioned type, *any* system of particular solutions pertaining to suitable boundary conditions does necessarily constitute an orthogonal system.

In order to deal with this assertion, it is advisable to extend (10.5-18) slightly and to set out from the so-called *general differential equation of Sturm–Liouville type*:

$$[\sigma(x)y']' = [\tau(x) - \lambda\rho(x)]y \quad [a < x < b; y = y(x)]. \quad (10.6-1)$$

Here (a, b) is a finite or infinite fundamental interval, λ denotes a non-negative (real) parameter; from the given functions $\rho(x)$, $\sigma(x)$, $\tau(x)$ the first two are assumed to be positive for $a < x < b$, furthermore $\tau(x)$ and $\rho(x)$ must be continuous, $\sigma(x)$ must be continuous and smooth (continuously differentiable) in (a, b) . We look for functions $y = y(x)$ twice continuously

differentiable at any point $x \in (a, b)$, for which at the boundary points $x = a$ and $x = b$ one of the following four pairs of conditions is fulfilled:

$$\left. \begin{array}{ll} \text{(I)} & Y_a = Y_b = 0 \quad \text{with} \quad Y'_a, Y'_b \text{ finite,} \\ \text{(II)} & Y'_a = Y'_b = 0 \quad \text{with} \quad Y_a, Y_b \text{ finite,} \\ \text{(III)} & Y_a = Y'_b = 0 \quad \text{with} \quad Y_b, Y'_a \text{ finite or} \\ & Y_b = Y'_a = 0 \quad \text{with} \quad Y_a, Y'_b \text{ finite,} \\ \text{(IV)} & Y_a = Y_b \text{ finite and} \quad Y'_a = Y'_b \text{ finite.} \end{array} \right\} \quad (10.6-2)$$

The abbreviations used in (10.6-2) are: $Y_a = [\sigma(x)^{1/2}y(x)]_{x=a}$, $Y_b = [\sigma(x)^{1/2}y(x)]_{x=b}$, $Y'_a = [\sigma(x)^{1/2}y'(x)]_{x=a}$, $Y'_b = [\sigma(x)^{1/2}y'(x)]_{x=b}$, and we mean by each value at a boundary point the corresponding (finite) limit from the interior of (a, b) .

Now, it is a fundamental fact in the theory of second-order boundary problems — a very important one also from point of view of the physical applications — that the so-called *Sturm-Liouville boundary value problem* given by (10.6-1)—(10.6-2) is in general solvable if and only if the parameter λ equals any of certain discrete real values: the *eigenvalues* (characteristic numbers) of the problem. (It can be shown that the latter must be non-negative.) Let $\lambda_1 < \lambda_2 < \dots$ be the increasing sequence of all eigenvalues; then to any λ_n at most two linearly independent solutions of the problem — called *eigenfunctions* — are attached such that their $\rho(x)$ -norms are finite. If $y_n = y_n(x)$ denotes one of the eigenfunctions mentioned, so it may happen that every eigenfunction associated with λ_n can be written in the form $\text{const} \cdot y_n(x)$. In this case, λ_n is called a *simple eigenvalue*; otherwise (i.e., if two linearly independent eigenfunctions belong to λ_n , and an arbitrary eigenfunction pertaining to λ_n is a linear combination of the former ones) λ_n is said to be a *twofold (double) eigenvalue*. The totality of eigenvalues is the so-called *spectrum* of the problem.¹

10.62 After these, we get in a few lines a general theorem on the *orthogonality of eigenfunctions*:

10.6.1 Let $y_1(x), y_2(x), \dots$ be a sequence of eigenfunctions of the Sturm-Liouville equation (10.6-1), satisfying one of the pair of conditions (10.6-2) (I)–(IV), and such that all terms of the sequence are associated with different eigenvalues. Then the system $\{y_n(x)\}$ is orthogonal in (a, b) with the weight function $\rho(x)$.

¹ The pertinent classical results were published by J. STURM and J. LIOUVILLE during the 1830's, in the first volumes of "Journal de mathématiques pures et appliquées". (Cf. Sturm [1]–[2] and Liouville [1]–[2].)

As regards recent development of the theory, see e.g. Atkinson [1], Mikolás [2], Sz.-Nagy [16].

P r o o f. Consider two arbitrary eigenfunctions $y_m(x)$, $y_n(x)$ ($m \neq n$) of the Sturm-Liouville boundary problem, and denote the corresponding distinct non-negative eigenvalues by λ_m and λ_n , respectively. Then substituting in (10.6-1), we obtain the identities

$$\begin{aligned} [\sigma(x)y'_m(x)]' &= [\tau(x) - \lambda_m\rho(x)]y_m(x), \\ [\sigma(x)y'_n(x)]' &= [\tau(x) - \lambda_n\rho(x)]y_n(x), \end{aligned}$$

which yield after multiplication by $y_n(x)$ and $y_m(x)$, respectively, and subtraction of corresponding terms the formula:

$$\begin{aligned} (\lambda_m - \lambda_n)\rho(x)y_m(x)y_n(x) &= \\ &= y_n(x)[\sigma(x)y'_m(x)]' - y_m(x)[\sigma(x)y'_n(x)]' = \\ &= [\sigma(x)(y'_m(x)y_n(x) - y_m(x)y'_n(x))]' \quad (a < x < b). \end{aligned}$$

Integrating on both sides and using the boundary conditions, we can write:

$$\begin{aligned} (\lambda_m - \lambda_n) \int_a^b \rho(x)y_m(x)y_n(x) dx &= [\sigma(x)(y'_m(x)y_n(x) - y'_n(x)y_m(x))]_{x=a}^{x=b} = \\ &= [\sigma(x)^{1/2}y'_m(x) \cdot \sigma(x)^{1/2}y_n(x)]_{x=a}^{x=b} - [\sigma(x)^{1/2}y'_n(x) \cdot \sigma(x)^{1/2}y_m(x)]_{x=a}^{x=b} \end{aligned}$$

Therefore by $\lambda_m \neq \lambda_n$ it follows

$$\int_a^b \rho(x)y_m(x)y_n(x) dx = 0 \quad (m \neq n),$$

i.e., the assertion.

* * *

10.63 Note that (10.6-1) turns for $\rho(x) = \sigma(x) = 1$, $\tau(x) = 0$ and with the notation $\lambda = \omega^2 > 0$ into

$$y'' = -\omega^2 y, \quad (10.6-3)$$

which is the differential equation of *harmonic oscillation*. The well-known general solution of this equation is:

$$y = C_1 \sin(\omega x) + C_2 \cos(\omega x), \quad (10.6-4)$$

C_1 and C_2 denoting arbitrary constants. We can see that the Sturm-Liouville equation (10.6-1) may also be regarded as a strong generalization of (10.6-3).

If in our case e.g. $[0, \pi]$ is taken as a fundamental interval, then easy calculation shows that the eigenvalues associated with the boundary conditions $y(0) = y(\pi) = 0$ are: $\lambda_n = \omega_n^2 = n^2$ ($n = 0, 1, 2, \dots$), while the corresponding eigenfunctions can be written in the form: $y_n(x) = \text{const} \cdot \sin nx$. On the other hand, to the boundary conditions $y'(0) = y'(\pi) = 0$ the same eigenvalues and the eigenfunctions $y_n(x) = \text{const} \cdot \cos nx$ are to be attached.

We can establish similarly that the trigonometric base system (10.1-1) is an example of eigenfunctions belonging to boundary conditions of the (10.6-2) (IV) type; furthermore, all the classical orthogonal polynomial systems (COP, cf. 10.54) present again examples of eigenfunctions associated with boundary conditions of the (10.6-2) (I) and (II) type. In all these cases, the pertaining eigenvalues are simple, except the system (10.1-1) having double eigenvalues; for the latter, the totality of eigenfunctions corresponding to the eigenvalue $\lambda_n = n^2$ is given by the linear combinations of $\sin nx$ and $\cos nx$.

10.64 It is also useful to present here such a special orthogonal system of Sturm-Liouville type, which satisfies boundary conditions of the form (10.6-2) (III) with $\tau(x) \neq 0$; moreover, which originates from the family of higher transcendental functions applied most frequently: the so-called *Bessel functions*.

The *Bessel function of order μ of the first kind*¹ — denoted by $J_\mu(t)$ — is the simplest particular solution of the following linear homogeneous differential equation of second order:

$$t^2 y'' + ty' + (t^2 - \mu^2)y = 0 \quad (y = y(t)) \quad (10.6-5)$$

where μ is a given real (not necessarily integral-valued) parameter.

Its explicit definition is yielded by a generalized power series convergent everywhere:

$$J_\mu(t) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma(\mu + k + 1)} \left(\frac{t}{2}\right)^{\mu+2k} \quad (-\infty < t < \infty), \quad (10.6-6)$$

$\Gamma(x)$ being the gamma-function introduced in the previous section [cf. (10.5-25)–(10.5-27)]. In particular, by $\Gamma(1/2) = \sqrt{\pi}$ we have

$$J_{-1/2}(t) = \sqrt{\frac{2}{\pi t}} \cos t, \quad J_{1/2}(t) = \sqrt{\frac{2}{\pi t}} \sin t; \quad (10.6-7)$$

¹ Cf. Bessel [1], Watson [1]. Note that the so-called *Bessel function of order μ of the second kind* is defined as the simplest particular solution of (10.6-5) which is linearly independent of $J_\mu(t)$.

and for $\mu > -1$, the asymptotic behaviour of $\sqrt{t}J_\mu(t)$ reflects the functions sine and cosine, as $t \rightarrow \infty$. Correspondingly, every zero of $J_\mu(t)$ ($\mu > -1$) is real and, apart from the trivial zero ($t = 0$) occurring for $\mu > 0$, each of them has the multiplicity 1; furthermore, all the zeros lie symmetrically by pairs with respect to the origin.

Let $\omega_1, \omega_2, \dots$ be all positive zeros of $J_\mu(t)$ ($\mu > -1$) in increasing order. Owing to the above, $y = J_\mu(\omega_n x)$ (as a function of x) satisfies the differential equation

$$x^2 y'' + xy' + (\omega_n^2 x^2 - \mu^2)y = 0,$$

which, after dividing by x , can be written also in the form:

$$(xy')' = \left(\frac{\mu^2}{x} - \omega_n^2 x \right) y.$$

Hence it is obvious that the function $y = J_\mu(\omega_n x)$ ($0 \leq x \leq 1$) may also be considered as an eigenfunction attached to the eigenvalue $\lambda = \omega_n^2$ arising from the Sturm-Liouville boundary value problem:

$$\left. \begin{aligned} (xy')' &= \left(\frac{\mu^2}{x} - \lambda x \right) y, \\ y(1) &= 0, \quad y'(0) \text{ finite.} \end{aligned} \right\} \quad (10.6-8)$$

A comparison with (10.6-1)—(10.6-2) shows that in this case $(a, b) = (0, 1)$, $\rho(x) = \sigma(x) = x$, $\tau(x) = 1/x$, $Y_b = (x^{1/2}y)_{x=1} = y(1) = 0$, $Y'_a = (x^{1/2}y')_{x=0} = 0$; moreover all the eigenvalues are different and simple.

Thus we can conclude by 10.6.1 at once:

10.6.2 *If $\mu > -1$, then the sequence of functions*

$$J_\mu(\omega_1 x), \quad J_\mu(\omega_2 x), \dots$$

constitutes an orthogonal system in the interval $0 < x < 1$ with respect to the weight function $\rho(x) = x$, i.e.,

$$\int_0^1 x J_\mu(\omega_m x) J_\mu(\omega_n x) dx = 0 \quad (m \neq n). \quad (10.6-9)$$

Among other properties of Bessel functions, we mention the validity of certain recurrence formulae, analogous to (10.4-5). These relations render it possible to evaluate the integral (10.6-9) for $m = n$ too, i.e., to get the following expression for the pertinent square norm:

$$\|\sqrt{x}J_\mu(\omega_n x)\|^2 = \frac{1}{2}[J_{\mu+1}(\omega_n)]^2. \quad (10.6-10)$$

By means of (10.6–10), we can form immediately the expansion of a function $f(t) \in L_x^2(0, 1)$ in terms of the system $\{J_\mu(\omega_n x)\}$, i.e., the so-called *Fourier-Bessel series*.

10.65 The above material is typical concerning the fundamental properties of *eigenfunctions associated with Sturm-Liouville differential equations*. The pertinent general investigations depend essentially on whether the basic interval (a, b) is finite and if the boundary limits of the function $\sigma(x)$ figuring in (10.6–1) are positive.

On the assumption of both conditions, that is in the so-called *regular case*, the examples with $\rho(x) = \sigma(x) = 1$, $\tau(x) = 0$ can be used as models when we are interested in the existence and asymptotic behaviour of eigenvalues, in the distribution of zeros of the corresponding eigenfunctions. Then, according to the so-called *theorem of oscillation* which is simple to prove: 1) the set of eigenvalues is countable and can be arranged in a monotone increasing sequence tending to $+\infty$; 2) on the boundary conditions (10.6–2) (I)–(III) all eigenvalues are simple; 3) if $y_1(x), y_2(x), \dots$ is a sequence of pertaining eigenfunctions, so $y_n(x)$ has exactly n zeros in the interval (a, b) .

In turn, in the case of an infinite fundamental interval (a, b) or of a function $\sigma(x)$ vanishing at a or b or at both boundary points (so-called *singular case*), similar statements can be made only with simultaneous restriction about $\sigma(x)$, moreover each zero of $\sigma(x)$ has a certain effect on the range of values of eigenfunctions. (See e.g. the case of classical orthogonal polynomial systems or of $\{J_\mu(\omega_n x)\}$.)

We have to remark still: the system containing *all the Sturm-Liouville eigenfunctions* pairwise linearly independent of each other is always $\rho(x)$ -complete in the regular case and has this property only on suitable additional conditions in the singular case. [Cf. 10.4.2 (II) and 10.54.]

The Fourier series in terms of a complete system of eigenfunctions belonging to a boundary value problem of the type (10.6–1)—(10.6–2) is usually called for short *Sturm-Liouville series*. We shall see later that the discussion of this type of orthogonal series is closely connected with the theory of ordinary Fourier series of the form (10.2–4).

10.7 The Gram-Schmidt orthogonalization procedure

10.71 We saw in Section 10.3, that the sequence of power functions of non-negative integral exponents (orthogonal in no interval) can be transformed into a system orthogonal in some prescribed finite interval (a, b) by appropriate linear combinations with constant coefficients, i.e., polynomials. It may thus be possible for any arbitrary non-orthogonal system of

functions to make it orthogonal — *orthogonalize* — in such a way that we consider linear combinations of the first two, first three etc. terms of the given sequence with suitable coefficients.

Let us start with a sequence $g_n(t) \in L^2(a, b)$ ($n = 1, 2, \dots$) satisfying the natural restrictions that it does not contain the null element of $L^2(a, b)$ (cf. 9.43), further that none of $g_n(t)$ depends linearly on the terms with lower indices of the sequence. Then we can find at once a real constant K such that $g_2(t) + K\hat{g}_1(t)$ should be orthogonal to $\hat{g}_1(t)$, where

$$\hat{g}_1(t) = \|g\|^{-1}g_1(t) \quad (10.7-1)$$

denotes the normed form of $g_1(t)$. The condition

$$\int_a^b [g_2(t) + K\hat{g}_1(t)] \hat{g}_1(t) dt = (g_2, \hat{g}_1) + K\|\hat{g}_1\|^2 = 0$$

yields by $\|\hat{g}_1\| = 1$ the value

$$K = -(g_2, \hat{g}_1).$$

Hence it is obvious that the difference

$$G_2(t) = g_2(t) - (g_2, \hat{g}_1)\hat{g}_1(t) \quad (10.7-2)$$

or its normalized form

$$\hat{g}_2(t) = \|G_2\|^{-1}G_2(t) \quad (10.7-3)$$

satisfy the relation of orthogonality:

$$(\hat{g}_2, \hat{g}_1) = 0. \quad (10.7-4)$$

Proceeding similarly, let us determine two constants K_1, K_2 such that $g_3(t) + K_1\hat{g}_1(t) + K_2\hat{g}_2(t)$ should be orthogonal both to $\hat{g}_1(t)$ and to $\hat{g}_2(t)$ in (a, b) ; from the equations

$$\begin{cases} \int_a^b [g_3(t) + K_1\hat{g}_1(t) + K_2\hat{g}_2(t)] \hat{g}_1(t) dt = 0, \\ \int_a^b [g_3(t) + K_1\hat{g}_1(t) + K_2\hat{g}_2(t)] \hat{g}_2(t) dt = 0, \end{cases}$$

we get by $\|\hat{g}_1\| = \|\hat{g}_2\| = 1$ and (10.7-4) immediately:

$$K_1 = -(g_3, \hat{g}_1), \quad K_2 = -(g_3, \hat{g}_2).$$

Consequently, if we attach to $\hat{g}_1(t)$ and $\hat{g}_2(t)$ the function resulting from

$$G_3(t) = g_3(t) - [(g_3, \hat{g}_1)\hat{g}_1(t) + (g_3, \hat{g}_2)\hat{g}_2(t)] \quad (10.7-5)$$

by normalization, denoted expediently by

$$\hat{g}_3(t) = \|G_3\|^{-1}G_3(t), \quad (10.7-6)$$

we obtain an orthonormal system of three elements, composed from mutually linearly independent functions g_1, g_2, g_3 .

It is clear that this procedure can be continued indefinitely. At the n -th step ($n \geq 2$) we conclude that the function

$$G_n(t) = g_n(t) - \sum_{k=1}^{n-1} (g_n, \hat{g}_k)\hat{g}_k(t) \quad (10.7-7)$$

is orthogonal to any of the functions $\hat{g}_1(t), \hat{g}_2(t), \dots, \hat{g}_{n-1}(t)$, where

$$\hat{g}_k(t) = \|G_k\|^{-1}G_k(t) \quad (k \geq 2) \quad (10.7-8)$$

is defined recursively by $\hat{g}_1, \dots, \hat{g}_{k-1}$. Thus the sequence

$$\hat{g}_1(t), \hat{g}_2(t), \dots, \hat{g}_n(t), \dots, \quad (10.7-9)$$

whose generic element $\hat{g}_n$ is a linear combination of $g_1, g_2, \dots, g_n$ by the above construction, is in (a, b) orthogonal and normal.

This is the *Gram-Schmidt orthogonalization*; it was discovered towards the end of the last century.¹

10.72 Our argumentation shows: the connection between the orthonormal system $\{\hat{g}_n(t)\}$ and the original sequence $\{g_n(t)\}$ is *reversible* in a sense, inasmuch as any function g_n can also be written as a linear combination of the functions $\hat{g}_1, \hat{g}_2, \dots, \hat{g}_n$.

The fact that any partial sum of the Fourier series of a function $f(t) \in L^2(a, b)$ in terms of $\hat{g}_n(t)$ ($n = 1, 2, \dots$) is at the same time a linear combination of the pertaining elements of $\{g_n(t)\}$ as well, clarifies the theoretical and practical significance of orthogonalization: this enables us to approximate *arbitrary* functions by means of the terms of a function system *prescribed arbitrarily*.

¹ See Gram [1], Schmidt [1].

10.8 Applications: systems of Haar's type, Rademacher and Walsh functions

10.81 By the orthogonalization procedure just presented, *any* sequence of square integrable functions can be made orthonormal, provided that the eventual linear connections between elements of the original system have been previously eliminated. The resulting orthogonal system $\{\hat{g}_n(t)\}$ is of course not complete in general. It can be shown, however, that $\{\hat{g}_n(t)\}$ is always complete whenever the original sequence $\{g_n(t)\}$ is *closed* in the sense that every function continuous in the fundamental interval is there uniformly approximable by suitable finite linear combinations of the functions $g_1(t), g_2(t), \dots$ (cf. 3.42).¹

The most classical pertinent example is the system of power functions $1, t, t^2, \dots$ which is closed on any finite segment $[a, b]$ by the I. approximation theorem of Weierstrass [see 3.4.1 (I)] and in which the linear independence of the terms is obvious. If we orthogonalize on $[-1, 1]$ the function sequence $g_n(t) = t^{n-1}$ ($n = 1, 2, \dots$), then we find that in this case

$$\hat{g}_1(t) = 1/\sqrt{2} = \sqrt{\frac{1}{2}}P_0(t), \quad \hat{g}_2(t) = \sqrt{\frac{3}{2}}t = \sqrt{\frac{3}{2}}P_1(t),$$

$$\hat{g}_3(t) = \sqrt{\frac{5}{2}}\left(\frac{3}{2}t^2 - \frac{1}{2}\right) = \sqrt{\frac{5}{2}}P_2(t),$$

$$\hat{g}_4(t) = \sqrt{\frac{7}{2}}\left(\frac{5}{2}t^3 - \frac{3}{2}t\right) = \sqrt{\frac{7}{2}}P_3(t),$$

and in general — as can be easily seen by mathematical induction —

$$\hat{g}_n(t) = \sqrt{\frac{2n-1}{2}} \cdot \frac{[(t^2-1)^{n-1}]^{(n-1)}}{(2n-2)!!} = \sqrt{\frac{2n-1}{2}}P_{n-1}(t),$$

where $P_0(t), P_1(t), \dots$ are the Legendre polynomials introduced in 10.32.

The system of Legendre polynomials is simply the orthogonalized form relating to $[-1, 1]$ of the power functions with non-negative integral exponents; further the completeness of $\{P_n(t)\}$ is a consequence of the fact that the system $1, t, t^2, \dots$ is closed in the above sense.

¹ The connection between the concepts of closure and completeness becomes still closer, if the definition of the former is modified in such a way that the approximated functions are required to be square integrable instead of to be continuous, and if the norm of the corresponding difference is accepted as a measure of approximation. In the space L^2 , any orthogonal system is complete if and only if it is closed. (The notion of completeness can also be extended to non-orthogonal systems.)

A system of polynomials orthogonal in an arbitrary finite segment $[a, b]$ with respect to a prescribed weight function $\rho(t)$ can similarly be generated by orthogonalization of the base functions $g_n(t) = \rho(t)^{1/2} t^{n-1}$ ($n = 1, 2, \dots$) on $[a, b]$. (Cf. 10.41.)

10.82 As another important example, we consider the *system of characteristic functions* associated with a sequence of decompositions refining indefinitely a finite interval $[a, b]$ (cf. 1.14).¹ On the one hand, the theorem on uniform continuity implies at once that any function continuous in $[a, b]$ is there uniformly approximable by step functions pertaining to such a sequence of partitions, on the other hand, each step function is a linear combination of characteristic functions belonging to the occurring subintervals; so this system of characteristic functions is trivially closed. (Note that this property can be transferred from the class of continuous functions essentially to that of measurable functions as well; just this is on what the F. RIESZ theory of Lebesgue integral is based. Cf. 6.43.)

Restricting ourselves to the simplest case, let us examine first the characteristic functions of the dyadic partitions of $[0, 1]$:

$$\begin{aligned} & \chi[0, 1]; \chi[0, 1/2]; \chi[1/2, 1]; \\ & \chi[0, 1/4]; \chi[1/4, 1/2]; \chi[1/2, 3/4]; \chi[3/4, 1]; \\ & \dots \end{aligned}$$

The subsequence relating to the n -th decomposition has the form:

$$\chi[0, 2^{-n}]; \chi[2^{-n}, 2 \cdot 2^{-n}]; \dots, \chi[(2^n - 1) \cdot 2^{-n}, 1],$$

and evidently constitutes an orthogonal system which is not closed and not complete [cf. (9.4-14)]; in turn, the whole sequence is closed but not orthogonal in $[0, 1]$.

In order that we can apply the Gram-Schmidt method to the latter, we have first to eliminate the linear connections between the terms:

$$\begin{aligned} & \chi[0, 2^{-n}] + \chi[2^{-n}, 2 \cdot 2^{-n}] = \chi[0, 2^{-(n-1)}], \\ & \chi[2 \cdot 2^{-n}, 3 \cdot 2^{-n}] + \chi[3 \cdot 2^{-n}, 4 \cdot 2^{-n}] = \chi[2^{-(n-1)}, 2 \cdot 2^{-(n-1)}], \\ & \dots \end{aligned}$$

If in any decomposition every second characteristic function is omitted accordingly, then the sequence thus obtained:

$$\left. \begin{aligned} & \chi[0, 1]; \chi[0, 1/2]; \chi[0, 1/4]; \chi[1/2, 3/4]; \\ & \chi[0, 1/8]; \chi[1/4, 3/8]; \chi[1/2, 5/8]; \chi[3/4, 7/8]; \\ & \dots \end{aligned} \right\} \quad (10.8-1)$$

¹ The "natural" representation of Haar's system described below, and particulars of some orthonormal systems of similar type, also obtained by orthogonalization are to be found in Mikolás [12].

already consists only of linearly independent terms, and at the same time (10.8-1) remains closed with respect to $[0,1]$.

Orthogonalization yields the following step functions:¹

$$\chi_0^{(0)}(t) = 1; \quad \chi_0^{(1)}(t) = \begin{cases} 1 & \text{for } t \in (0, 1/2); \\ -1 & \text{for } t \in (1/2, 1); \end{cases} \quad (10.8-2)$$

$$\chi_1^{(1)}(t) = \begin{cases} \sqrt{2} & \text{for } t \in (0, 1/4) \\ -\sqrt{2} & \text{for } t \in (1/4, 1/2); \end{cases} \quad \chi_1^{(2)}(t) = \begin{cases} 0, & \text{for } t \in (0, 1/2) \\ \sqrt{2} & \text{for } t \in (1/2, 3/4); \\ -\sqrt{2} & \text{for } t \in (3/4, 1) \end{cases}$$

In general, to the partition of $[0,1]$ into 2^{m+1} equal subintervals we attach the following 2^m functions ($m \geq 1$):

$$\chi_m^{(k)}(t) = \begin{cases} 2^{m/2} & \text{for } (k-1)2^{-m} < t < (k-\frac{1}{2})2^{-m}, \\ -2^{m/2} & \text{for } (k-\frac{1}{2})2^{-m} < t < k \cdot 2^{-m}, \\ 0 & \text{otherwise,} \end{cases} \quad (10.8-3)$$

(see Fig. 6).

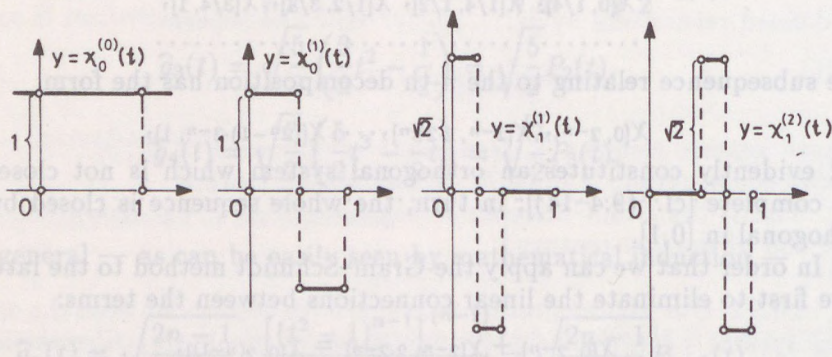


Fig. 6

10.83 The orthonormal system (10.8-2)—(10.8-3) was first studied in 1909 by A. HAAR (1885–1933), a professor of international reputation at the University in Szeged (Hungary) and discoverer a.e. of the abstract Haar's measure; $\{\chi_m^{(k)}(t)\}$ has played an essential role in the literature of orthogonal series.² As we shall see later, the corresponding Fourier series

¹ As a rule, it is advisable not to fix the values of Haar's functions at the points of discontinuity.

² See Haar [1], further e.g. Franklin [1], Mikolás [11]–[12], Nelson [1], Price [1]–[2], Schipp [1], Splettstösser–Wagner [1], Sz.-Nagy [13], Wade [1].

possesses very favourable convergence properties. At present, we shall deal only with one of the most elementary features of the so-called *Haar system*.

10.8.1 *Haar's orthonormal function system is complete in its fundamental interval $[0, 1]$.*

P r o o f. Let $f(t) \in L^2(0, 1)$, and assume $f(t)$ to be orthogonal to all the terms of (10.8-2)—(10.8-3); consider the integral function

$$F(x) = \int_0^x f(t) dt \quad (0 \leq x \leq 1).$$

Since $F(0) = 0$ and also

$$F(1) = \int_0^1 f(t) dt = \int_0^1 f(t) \chi_0^{(0)}(t) dt = 0,$$

we can write:

$$\begin{aligned} \int_0^1 f(t) \chi_0^{(1)}(t) dt &= \int_0^{1/2} f(t) dt - \int_{1/2}^1 f(t) dt = \\ &= F(1/2) - [F(1) - F(1/2)] = 2F(1/2) = 0; \end{aligned}$$

therefore

$$F(1/2) = 0.$$

We get similarly by the vanishing of $\int_0^1 f(t) \chi_1^{(1)}(t) dt$ and $\int_0^1 f(t) \chi_1^{(2)}(t) dt$ on the basis of the foregoing first $F(1/4) = F(3/4) = 0$; after, because all the remaining Fourier constants of f are zero as well, it follows successively:

$$F(k/2^m) = 0 \quad (k = 1, 2, \dots, 2^{m-1})$$

for every m .

Thus $F(x)$ takes the value 0 on an everywhere dense subset of $[0, 1]$, which implies by the continuity of its integral function that $F(x) \equiv 0$ ($0 \leq x \leq 1$). Hence we may conclude (cf. 7.6.4)

$$f(x) = F'(x) = 0$$

for almost all $x \in [0, 1]$.

* * *

10.84 Quite recently, the above "natural" construction of Haar's system has suggested to the author to introduce in the theory certain much more general related orthogonal systems, the $\{\mu_{k_1 \dots k_\nu}^{(r)}\}$ family of orthogonal functions. (Cf. Mikolás [16].)

Starting from an arbitrary finite basic interval $I_0 = [a, b]$, let us take a partition of I_0 into mutually disjoint subintervals (not necessarily of equal length), say

$$I_0 = \bigcup_{k=1}^N I_k.$$

Next divide further each I_k into union of some disjoint I_{kl} , each subinterval I_{kl} into union of some disjoint I_{klm} , and so on; expressed by formulae:

$$I_k = \bigcup_{l=1}^{N_k} I_{kl} \quad (1 \leq k \leq N), \quad I_{kl} = \bigcup_{m=1}^{N_{kl}} I_{klm} \quad (1 \leq k \leq N, 1 \leq l \leq N_k), \dots$$

Let us assume that the number of subintervals is always > 1 , and that the partitions

$$I_{k_1 \dots k_\nu} = \bigcup_{k_{\nu+1}=1}^{N_{k_1 \dots k_\nu}} I_{k_1 \dots k_\nu k_{\nu+1}}$$

satisfy the condition $\max |I_{k_1 \dots k_{\nu+1}}| \rightarrow 0$.

Now we define the elements of the above-mentioned function family in $[a, b]$ as follows:

$$\left. \begin{aligned} \mu_0^{(0)}(t) &\equiv 1; \\ \mu_0^{(r)}(t) &= \sum_{p=r+1}^N [|I_p| \chi_{I_\nu}(t) - |I_r| \chi_{I_p}(t)] \quad (1 \leq r \leq N-1) \\ \mu_k^{(r)}(t) &= \sum_{p=r+1}^{N_k} [|I_{kp}| \chi_{I_{k_r}}(t) - |I_r| \chi_{I_{kp}}(t)] \\ &\quad \left(\begin{array}{l} 1 \leq k \leq N \\ 1 \leq r \leq N_k - 1 \end{array} \right); \\ \dots\dots\dots \\ \dots\dots\dots \end{aligned} \right\} \quad (10.8-4)$$

Without using characteristic functions, this definition becomes:

$$\mu_{k_1 \dots k_\nu}(t) = \begin{cases} \sum_{p=r+1}^{N_{k_1 \dots k_\nu}} |I_{k_1 \dots k_\nu p}| & \text{for } t \in I_{k_1 \dots k_\nu r}, \\ -|I_{k_1 \dots k_\nu r}| & \text{for } t \in \bigcup_{p=r+1}^{N_{k_1 \dots k_\nu}} I_{k_1 \dots k_\nu p}, \\ 0 & \text{for other } t \in I_0. \end{cases} \quad (10.8-5)$$

We see that each term of the system (10.8-4)—(10.8-5) has a *very simple geometric meaning*: $\mu_{k_1 \dots k_\nu}^{(r)}(t)$ ($r \geq 1$) is a step function with at most three intervals of constancy and with a zero-valued integral on I_0 . (Fig. 7; the rectangles over and below the t -axis are congruent.)

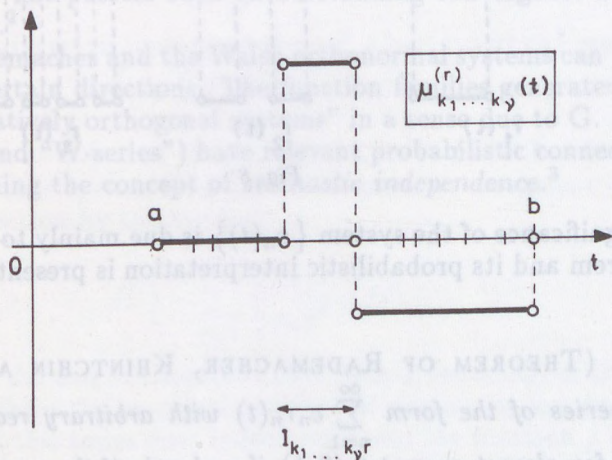


Fig. 7

It can be shown that the normalized version of $\{\mu_{k_1 \dots k_\nu}^{(r)}(t)\}$ is such a *generalization of Haar's orthonormal system*, whose orthogonality and completeness relate to $I_0 = [a, b]$ and which is associated with a Fourier series, like the "dyadic" Haar series. The problem of an expansion in terms of the function family $\{\mu_{k_1 \dots k_\nu}^{(r)}(t)\}$ will also be examined in Chapter 12.

10.85 A further interesting sequence *orthonormal in $(0, 1)$* and interlinked with Haar's functions is the so-called *Rademacher system* $\{r_n(t)\}$, found in 1922. Each term of this sequence can be generated by addition of the Haar functions with identical subscripts and by normalization, i.e.,

$$r_0(t) = \chi_0^{(0)}(t), \quad r_1(t) = \chi_0^{(1)}(t), \quad r_{m+1}(t) = 2^{-m/2} \sum_{k=1}^{2^m} \chi_m^{(k)}(t). \\ (m = 1, 2, \dots)$$

Hence we see that the Rademacher function $r_n(t)$ takes the values $+1$ and -1 alternately in the subintervals $(k/2^n, (k+1)/2^n)$ ($k = 0, 1, \dots, 2^n - 1$), and it can also be written apart from the points of discontinuity in the form:

$$r_n(t) = \text{sgn}(\sin 2^n \pi t) \quad (n = 0, 1, 2, \dots) \quad (10.8-6)$$

(Fig. 8).

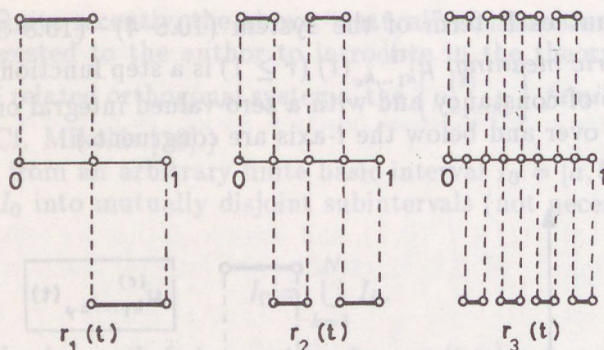


Fig. 8

The significance of the system $\{r_n(t)\}$ is due mainly to a useful convergence theorem and its probabilistic interpretation is presented here without proof:

10.8.2 (THEOREM OF RADEMACHER, KHINTCHIN AND KOLMOGOROV)¹ A series of the form $\sum_{n=1}^{\infty} c_n r_n(t)$ with arbitrary real coefficients is convergent for almost every $t \in (0, 1)$ if and only if the condition

$$\sum_{n=1}^{\infty} c_n^2 < \infty \quad (10.8-7)$$

is fulfilled.

Considering that the generic term of $\sum c_n r_n(t)$ is a.e. of the form $\pm c_n$, the terminology of *probability theory* enables us to establish the last statement also as follows:

10.8.3 If we choose the signs in the series $\sum (\pm c_n)$ in a random manner, we obtain with probability 1 a convergent or divergent series according to the cases $\sum c_n^2 < \infty$ and $\sum c_n^2 = \infty$, respectively.

10.86 It should be noted: $\{r_n(t)\}$ is obviously *not complete*, since e.g. the function $r_1(t)$ is orthogonal to any element of the system on the segment $0 < t < 1$.

To the work of J. L. WALSH (1923) is due the "natural" widening of Rademacher's sequence to a complete orthonormal system in the interval $[0, 1]$. Its definition can be written in the form:

$$w_n(t) = r_{\alpha_1+1}(t) r_{\alpha_2+1}(t) \dots r_{\alpha_\nu+1}(t) \quad (0 < t < 1; n = 0, 1, \dots) \quad (10.8-8)$$

¹ We stress that the pendant of this theorem is *not valid for general orthogonal series*. Cf. Khintchin-Kolmogorov [1], Rademacher [1].

where $\alpha_1, \alpha_2, \dots, \alpha_\nu$ denote the digits of the dyadic representation $n = 2^{\alpha_1} + 2^{\alpha_2} + \dots + 2^{\alpha_\nu}$. The Walsh system $\{w_n(t)\}$ turns into $\{r_n(t)\}$ for $n = 2^m$ ($m = 0, 1, \dots$).

The Fourier expansion in terms of Walsh functions $w_n(t)$ has wide-ranging applications in modern computer technics, though this so-called Walsh series does not possess such an outstanding convergence behaviour as the Haar series.²

Both the Rademacher and the Walsh orthonormal systems can be developed further in certain directions. The function families generated in such a way ("multiplicatively orthogonal systems" in a sense due to G. ALEXITS or "W-systems" and "W-series") have relevant probabilistic connections, in particular concerning the concept of stochastic independence.³ Denote this connection by the symbol:

$$f(t) \sim \sum_{n=1}^{\infty} \frac{(f, \varphi_n)}{\|\varphi_n\|^2} \varphi_n(t). \quad (11.1-1)$$

As was shown at the end of Chapter 9, the formal deduction of the concept of Fourier series from the so-called problem of decomposition does not yet assure at all, that this series does indeed represent the function f in a sense, so the clarification of the problem demands further detailed study.

In view of the geometric background of the theory of Hilbert space, it is to be expected that the symbol of "associating" $\sim$ may be replaced by the sign of equality in (11.1-1), if the base system $\{\varphi_n\}$ is complete, and we mean by convergence and sum of the right-hand side Fourier series the pertinent notions "made to measure" for $L^2(a, b)$, as introduced in 9.51.

This chapter is devoted to the treatment of the general expansion theorem above and of other related fundamental facts. The base system $\{\varphi_n\}$ will always be taken as normal, i.e., we premise $\|\varphi_n\| = 1$ ($n = 1, 2, \dots$). This is no restriction of generality, because any orthogonal system can be made orthonormal by normalizing, i.e., multiplying each term by a suitable constant factor.

11.12 Our starting-point is a simple, but very useful formula:

11.1.1 (FUNDAMENTAL IDENTITY FOR PARTIAL SUMS OF ORTHOGONAL SERIES) For any function $f(t) \in L^2(a, b)$ and for any system of functions $\varphi_n(t) \in L^2(a, b)$ ($n = 1, 2, \dots$), orthonormal in (a, b) , holds:

$$\left\| f(t) - \sum_{n=1}^N c_n \varphi_n(t) \right\|^2 = \|f\|^2 - \sum_{n=1}^N \pi_n^2 + \sum_{n=1}^N (c_n - \pi_n)^2 \quad (11.1-2)$$

($N = 1, 2, \dots$)

² Cf. Walsh [1] and e.g. Wagner [1].

³ See Alexits [1]-[2], as well as Alexits-Tandori [1].

PROPERTIES OF ARBITRARY ORTHOGONAL SYSTEMS AND SERIES IN HILBERT'S FUNCTION SPACE

11.1 Minimum property of the partial sums of a Fourier series

11.11 Let us attach to a function $f(t) \in L^2(a, b)$ its Fourier series in terms of a given base system $\{\varphi_n(t)\}$, orthogonal in a segment (a, b) , and denote this connection by the symbol:

$$f(t) \sim \sum_{n=1}^{\infty} \frac{(f, \varphi_n)}{\|\varphi_n\|^2} \varphi_n(t). \quad (11.1-1)$$

As was shown at the end of Chapter 9, the formal deduction of the concept of Fourier series from the so-called problem of decomposition does not yet assure at all, that this series does indeed represent the function f in a sense, so the clarification of the problem demands further detailed study.

In view of the geometric background of the theory of Hilbert space, it is to be expected that the symbol of "associating" $\sim$ may be replaced by the sign of equality in (11.1-1), if the base system $\{\varphi_n\}$ is *complete*, and we mean by convergence and sum of the right-hand side Fourier series the pertinent notions "made to measure" for $L^2(a, b)$, as introduced in 9.51.

This chapter is devoted to the treatment of the general expansion theorem above and of other related fundamental facts. The base system $\{\varphi_n\}$ will always be taken as *normal*, i.e., we premise $\|\varphi_n\| = 1$ ($n = 1, 2, \dots$). This is no restriction of generality, because any orthogonal system can be made orthonormal by norming, i.e., multiplying each term by a suitable constant factor.

11.12 Our starting-point is a simple, but very useful formula:

11.1.1 (FUNDAMENTAL IDENTITY FOR PARTIAL SUMS OF ORTHOGONAL SERIES) For any function $f(t) \in L^2(a, b)$ and for any system of functions $\varphi_n(t) \in L^2(a, b)$ ($n = 1, 2, \dots$), orthonormal in (a, b) , holds:

$$\left\| f(t) - \sum_{n=1}^N c_n \varphi_n(t) \right\|^2 = \|f\|^2 - \sum_{n=1}^N \gamma_n^2 + \sum_{n=1}^N (c_n - \gamma_n)^2 \quad (11.1-2)$$

$(N = 1, 2, \dots)$

where $c_1, c_2, \dots, c_N$ are arbitrary real numbers and $\gamma_n = (f, \varphi_n)$ denotes the n -th Fourier constant of f relating to $\{\varphi_n\}$.

In fact, our assumptions involve:

$$\begin{aligned} & \left\| f(t) - \sum_{n=1}^N c_n \varphi_n(t) \right\|^2 = \\ &= \int_a^b \left\{ f(t)^2 - 2f(t) \sum_{n=1}^N c_n \varphi_n(t) + \left[\sum_{n=1}^N c_n \varphi_n(t) \right]^2 \right\} dt = \\ &= \|f\|^2 - 2 \sum_{n=1}^N c_n (f, \varphi_n) + \sum_{n=1}^N c_n^2 = \|f\|^2 + \sum_{n=1}^N (c_n - \gamma_n)^2 - \sum_{n=1}^N \gamma_n^2. \end{aligned}$$

* * *

11.1.3 Consider now the square integral $\left\| f - \sum_{n=1}^N c_n \varphi_n \right\|^2$ as a function of *variable coefficients* $(c_1, c_2, \dots, c_N)$, the number N , further f and $\{\varphi_n\}$ being fixed. Since the expression on the right in (11.1-2) becomes minimal evidently for $c_n = \gamma_n$ ($n = 1, 2, \dots, N$) when it attains the value $\|f\|^2 - \sum_{n=1}^N \gamma_n^2$, we may conclude:

11.1.2 (MINIMUM PROPERTY OF PARTIAL SUMS OF FOURIER SERIES)

The Fourier series $\sum_{n=1}^N (f, \varphi_n) \varphi_n(t)$ is characterized among all orthogonal series of the form $\sum_{n=1}^{\infty} c_n \varphi_n(t)$ by the property that its N -th partial sum, for any fixed N , minimizes in (a, b) the square integral on the left of (11.1-2).

We remark that this important approximative feature may also be utilized to define the concept of Fourier series, and it renders possible a similarly elegant treatment of the theory as our present basis.

The special case of (11.1-2) which arises by substitution of the minimizing N -tuple $c_1 = (f, \varphi_1)$, $c_2 = (f, \varphi_2)$, ..., $c_N = (f, \varphi_N)$ is applied frequently, when the *square deviation* of f from the N -th partial sum of its Fourier series is estimated. Thus the formula in question:

$$\left\| f(t) - \sum_{n=1}^N (f, \varphi_n) \varphi_n(t) \right\|^2 = \|f\|^2 - \sum_{n=1}^N (f, \varphi_n)^2 \quad (N = 1, 2, \dots) \quad (11.1-3)$$

merits particular attention; its name is: *Bessel's identity*.

11.1.4 In general, on the basis of the minimum property 11.1.2 and the connection (11.1-3), the *best approximation with respect to square integral*

(i.e., the “continuous” counterpart of approaching according to the *principle of least squares*) is called briefly also *Bessel’s approximation*. (Cf. 9.51.) Beside the uniform or Tchebyshev approximation (cf. 10.51), this is the main expedient of the *general theory of approximation*, and of the so-called *constructive theory of functions* whose object is the simple approximative characterization of arbitrary functions.¹

The attachment of this latter approximation method, further of the pertinent identity (11.1–3) and its consequences to the name of F. W. BESSEL is motivated by the fact, that he first applied it in 1815 on the discrete pendant of 11.1.2 concerning trigonometric polynomials, in order to minimize systems of random errors at observations about periodical phenomena of nature.²

11.2 Bessel’s inequality and the Parseval–Hurwitz formula

11.21 Taking into consideration that the left-hand term of (11.1–3) is non-negative, we get for $N \rightarrow \infty$ the so-called *Bessel’s inequality*:

$$\sum_{n=1}^N (f, \varphi_n)^2 \leq \|f\|^2. \quad (11.2-1)$$

Hence we see that the square sum of all Fourier constants relating to $\{\varphi_n\}$ of a function $f \in L^2(a, b)$ constitutes a *convergent* series in case of *any* system $\{\varphi_n\}$ orthonormal in (a, b) .

Since the analogue of the left-hand series (11.2–1) in an Euclidean space is the square sum of all rectangular components of some vector, while the right-hand term in (11.2–1) corresponds to square of the length of the same vector, furthermore the latter two quantities are *equal* to one another, the question may be raised: whether the sign $\leq$ in (11.2–1) may be replaced by the sign of equality — it being assumed of course that $\{\varphi_n\}$ is also *complete*, just like the system of all unit vectors in E_ν ?

11.22 The answer is affirmative, inasmuch as we have:

11.2.1 (PARSEVAL–HURWITZ FORMULA) *If $\varphi_n(t)$ ($n = 1, 2, \dots$) is a complete orthonormal system in $L^2(a, b)$, then for every $f \in L^2(a, b)$ the representation holds:*

$$\sum_{n=1}^{\infty} (f, \varphi_n)^2 = \|f\|^2. \quad (11.2-2)$$

¹ See e.g. Jackson [1] as well as Akhieser [1] and Natanson [1]–[2]. We refer still to Erdős–Turán [1].

² Cf. Burkhardt [2], p. 647–648 and Zygmund [1], Vol. II, p. 331.

Before verifying 11.2.1, we should note that (11.2-2) implies also the (formally more general) *extended Parseval formula*:

$$\sum_{n=1}^{\infty} (f, \varphi_n)(g, \varphi_n) = (f, g), \quad (11.2-3)$$

This is valid for arbitrary $f \in L^2(a, b)$, $g \in L^2(a, b)$, and evidently corresponds to the elementary representation of the scalar product of two vectors by means of the pertaining vector components. The identity

$$\|f + g\|^2 = \|f\|^2 + \|g\|^2 + 2(f, g)$$

yields immediately the formula (11.2-3), if we apply (11.2-2).

As regards the proof of 11.2.1, in view of (11.1-3) the statement is equivalent to the limit relation:

$$\lim_{N \rightarrow \infty} \left\| f(t) - \sum_{n=1}^N (f, \varphi_n) \varphi_n(t) \right\|^2 = 0 \quad (11.2-4)$$

The above Parseval formulae are thus reduced to an expansion theorem on Fourier series of square integrable functions, which will be discussed — by its central importance — in the forthcoming section.

11.23 Some *historical comments* about the origin of the above results.

A. PARSEVAL was the first mathematician, who (about simultaneously with S. F. LACROIX) between 1799 and 1806 has dealt with the product integral of two functions represented by complex trigonometric series, whereby he formally obtained the simplest particular case of (11.2-3).¹

A. HURWITZ gave in his papers [1] and [2] at the beginning of this century the first exact foundations for the formulae (11.2-2)—(11.2-3) in the case of the trigonometric base system. At the same time, he presented a relevant application, namely an elegant analytic solution of the so-called *isoperimetric problem* on the maximalization of plane figures with prescribed perimeter.²

¹ Cf. Burkhardt [1], p. 947 and Hardy-Rogosinski [1], p. 94.

² As became well-known earlier through pioneering work of J. STEINER and others, the solution is the circle. See also e.g. Courant-Hilbert [1], p. 82-83 and Sz.-Nagy [16], Ch. VII, § 2.4.

11.3 L^2 -convergence of general Fourier series

11.31 We are going to treat such a representation of the elements of Hilbert's function space, which strictly corresponds to the decomposition of ordinary vectors into components directed along the rectangular coordinate axes. (Cf. 9.51.)

11.3.1 (GENERAL EXPANSION THEOREM)¹ *The Fourier series of every function $f \in L^2(a, b)$ in terms of any prescribed complete orthonormal base system $\{\varphi_n\}$ is in (a, b) convergent with respect to square integral (i.e., L^2 -convergent) and represents f in this sense. In short:*

$$\sum_{n=1}^{\infty} (f, \varphi_n) \varphi_n(t) = f(t) \quad [L^2(a, b)]. \quad (11.3-1)$$

Proof. 1° Let $\gamma_n = (f, \varphi_n)$ ($n = 1, 2, \dots$) denote the Fourier constants of f relating to the given base system $\{\varphi_n\}$.

According to the Riesz–Fischer convergence criterion 9.5.2, the Fourier series in question is $L^2(a, b)$ -convergent if and only if

$$\lim_{M, N \rightarrow \infty} \|\gamma_{M+1} \varphi_{M+1}(t) + \dots + \gamma_N \varphi_N(t)\| = 0 \quad (M < N);$$

this condition is fulfilled in our case, because (11.2-1) implies $\sum_{n=1}^{\infty} \gamma_n^2 < \infty$ and so

$$\left\| \sum_{n=M+1}^N \gamma_n \varphi_n(t) \right\|^2 = \int_a^b \left[\sum_{n=M+1}^N \gamma_n \varphi_n(t) \right]^2 dt = \sum_{n=M+1}^N \gamma_n^2$$

becomes as small as we please when M, N are sufficiently large.

Hence we conclude the existence of a function $f^* \in L^2(a, b)$ for which:

$$\sum_{n=1}^{\infty} (f, \varphi_n) \varphi_n(t) = f^*(t) \quad [L^2(a, b)]. \quad (11.3-2)$$

2° In order to show that $F^*(t) = f(t)$ for a.e. $t \in (a, b)$, let us estimate the deviation of (f^*, φ_n) and $\gamma_n = (f, \varphi_n)$.

By the inequality (9.4-5), it holds for any fixed n and $N \geq n$:

$$\begin{aligned} |(f^*, \varphi_n) - \gamma_n| &= |(f^*, \varphi_n) - (\gamma_1 \varphi_1 + \dots + \gamma_N \varphi_N, \varphi_n)| = \\ &= \left| \left(f^* - \sum_{k=1}^N \gamma_k \varphi_k, \varphi_n \right) \right| \leq \left\| f^* - \sum_{k=1}^N \gamma_k \varphi_k \right\| \cdot \|\varphi_n\| = \\ &= \|f^*(t) - [\gamma_1 \varphi_1(t) + \dots + \gamma_N \varphi_N(t)]\|; \end{aligned}$$

¹ Cf. F. Riesz [4].

since, owing to (11.3-2), the last term tends to 0 as $N \rightarrow \infty$, we obtain

$$|(f^*, \varphi_n) - \gamma_n| = 0, \quad \text{i.e.} \quad (f^*, \varphi_n) = \gamma_n.$$

But then we can write:

$$(f^* - f, \varphi_n) = (f^*, \varphi_n) - \gamma_n = 0 \quad (n = 1, 2, \dots),$$

that is to say: $f^* - f$ is orthogonal to *every* element of $\{\varphi_n\}$. The completeness of the base system being also required, this is possible only if

$$f^* - f = 0, \quad f^*(t) = f(t)$$

almost everywhere in (a, b) , which was to be proved.

* * *

11.3.2 Let us formulate two interesting consequences of theorem 11.3.1!

(I) In the previous chapter, we treated a number of classical examples of *complete* orthonormal systems in $L^2(a, b)$. If $\{\varphi_n(t)\}$ signifies one of these systems and f is any prescribed element of $L^2(a, b)$, so $s_n(t) = (f, \varphi_1)\varphi_1(t) + \dots + (f, \varphi_n)\varphi_n(t)$ ($n = 1, 2, \dots$) yields a concrete sequence consisting of elements $\in L^2(a, b)$ whose $L^2(a, b)$ -limit is f by 11.3.1. Hence we can see: all the partial sums of an orthogonal series of the form $c_1\varphi_1(t) + c_2\varphi_2(t) + \dots$ ($\sum c_n^2 < \infty$) constitute a *set dense everywhere* in $L^2(a, b)$. (Cf. 9.55.)

(II) The second part of the above proof implies: if we know all Fourier constants of a function $f(t) \in L^2(a, b)$ relating to some complete orthonormal system $\varphi_n(t) \in L^2(a, b)$ ($n = 1, 2, \dots$) or — what means the same — the corresponding Fourier series of f , then any other square integrable function having the same Fourier constants or the same Fourier series relating to $\{\varphi_n\}$, cannot differ from f , except the points of a null set. Briefly:

11.3.2 (LEMMA OF UNIQUENESS) *Every square integrable function is determined essentially uniquely (i.e. apart from a null set) by its Fourier constants or Fourier series relating to any complete orthonormal system in $L^2(a, b)$.*

11.4 The Riesz–Fischer theorem: isomorphism of Hilbert's function space and sequence space

11.41 According to our last statement, an arbitrary element $f \in L^2(a, b)$ can also be characterized by *countably many data* (namely by the Fourier constants just mentioned); and this is very surprising because the

range of values $f(t)$ — like the fundamental interval (a, b) — is a non-countable set.

But then some sort of *hidden structural likeness* between Hilbert's sequence space (l^2) and the function space $L^2(a, b)$ can be expected. The Bessel inequality (11.2-1) implies that the sequence of Fourier constants $\gamma_n = (f, \varphi_n)$ ($n = 1, 2, \dots$), in case of any $f \in L^2(a, b)$ and of a given orthonormal base system $\{\varphi_n\}$ belongs surely to l^2 ; and if one could show of every number sequence $\in l^2$ that its terms can *always* be regarded as Fourier constants of a suitable square integrable function, then hereby a "bridge" would be built, so to say, connecting l^2 with $L^2(a, b)$.

11.42 The solution to the problem was found by F. RIESZ and E. FISCHER in 1907, together with the convergence criterion 9.5.2 named after them. This result has since been of fundamental importance in investigations on orthogonal series.

11.4.1 (FUNDAMENTAL THEOREM OF RIESZ AND FISCHER FOR ORTHOGONAL SERIES)¹ *Version I.* To a real number sequence $c_1, c_2, \dots$ we can find a function $f(t) \in L^2(a, b)$ such that the terms of $\{c_n\}$ are just the Fourier constants of f relating to some complete orthonormal system $\varphi_n(t) \in L^2(a, b)$ ($n = 1, 2, \dots$), if and only if the series $c_1^2 + c_2^2 + \dots$ is convergent.

Version II. An orthogonal series $c_1\varphi_1(t) + c_2\varphi_2(t) + \dots$ with a complete orthonormal base system $\{\varphi_n\} \subset L^2(a, b)$ may be considered as the Fourier series of some $f \in L^2(a, b)$ in terms of $\{\varphi_n\}$ if and only if $c_1^2 + c_2^2 + \dots < \infty$.

In both versions, the function $f(t)$ is given for almost all t by the $L^2(a, b)$ -sum of the series $c_1\varphi_1(t) + c_2\varphi_2(t) + \dots$.

P r o o f ² Since the above two forms of the assertion are obviously equivalent to each other, it is sufficient to deal e.g. with version (I).

1° (Necessity.) If there exists a function $f(t) \in L^2(a, b)$ such that $(f, \varphi_n) = c_n$ ($n = 1, 2, \dots$), then (11.2-1) assures the convergence of the series $\sum c_n^2$, besides we know by 11.2.1 also that the sum of $c_1 + c_2 + \dots$ is $\|f\|^2$.

2° (Sufficiency.) Let $\{c_n\}$ be a sequence with $\sum c_n^2 < \infty$, and consider a complete orthonormal base system $\{\varphi_n\} \subset L^2(a, b)$.

Using Chauchy's convergence criterion, we get that

$$\sum_{n=M+1}^N c_n^2 = \left\| \sum_{n=M+1}^N c_n \varphi_n(t) \right\|^2$$

¹ F. Riesz [1], [3], [9] and E. Fischer [1].

² Note that the brevity of the proof is due only to the fact that 9.5.2 has already been treated.

becomes arbitrarily small for M and N large enough. But then applying 9.5.2 to the sequence

$$S_N(t) = c_1\varphi_1(t) + c_2\varphi_2(t) + \dots + c_N\varphi_N(t) \quad (N = 1, 2, \dots),$$

we get that this converges with respect to square integral to a function $f(t) \in L^2(a, b)$.

If we estimate the deviation of (f, φ_n) from c_n on the basis of the inequality (9.4-5), we can write:

$$\begin{aligned} |(f, \varphi_n) - c_n| &= \left| (f, \varphi_n) - \left(\sum_{k=1}^N c_k \varphi_k, \varphi_n \right) \right| = \\ &= \left| \left(f - \sum_{k=1}^N c_k \varphi_k, \varphi_n \right) \right| \leq \left\| f(t) - \sum_{k=1}^N c_k \varphi_k(t) \right\|; \end{aligned}$$

and by the definition of f , the last norm tends to 0 as $N \rightarrow \infty$.

Consequently we have for every n

$$(f, \varphi_n) = c_n,$$

i.e., the function f possesses the property required in the assertion.

The uniqueness of the solution follows from the completeness of $\{\varphi_n\}$, just like at the end of the verification of 11.3.1.

* * *

It is to be emphasized: the kernel of theorem 11.4.1, i.e., the statement that *to any number sequence $\{c_n\}$ with $\sum c_n^2 < \infty$ we can find a square integrable function whose Fourier constants relating to a prescribed orthonormal base system are just $c_1, c_2, \dots$* — this proposition is plainly the counterpart of the existence in E_3 of a vector having arbitrarily prescribed rectangular components.¹ We have also to point out: while the situation in E_3 degenerates to a triviality because of the finite-dimensionality of the space, the “geometry” of Hilbert’s function space is much more complicated, and thus the sufficiency of the (extremely simple) condition $\sum c_n^2 < \infty$ cannot be expected at all.

11.43 The primary significance of the Riesz–Fischer fundamental theorem is that by means of it, the elements of the function space L^2 can be mapped bijectively to the elements of the sequence space l^2 , in such a way that applying the operations of addition, subtraction and the multiplication

¹ Here the number of components is finite and so the convergence of their square sum is self-evident.

by a fixed number first in one of these spaces and after to the corresponding elements of the other space, we get mutually corresponding elements in the case of each fundamental operation separately; furthermore [cf. (11.2-2)—(11.2-3)] the norms of mutually corresponding elements as well as the inner products of any corresponding pairs of elements are identical. Using the terminology of modern algebra, this correspondence can simply be expressed by saying that *there is an isomorphic connection (isomorphy) between the spaces L^2 and l^2* . The advantages of isomorphy for research and applications are evident; on such a basis, any result concerning elements of one space can be transferred directly to the corresponding elements of the other space.

As regards version (II) of 11.4.1, it furnishes a *simple necessary and sufficient condition in order that an orthogonal expansion in terms of a complete orthonormal system should be a Fourier series*. In view of theorem 11.3.1, this result means a useful expedient as well.

Finally, we want to draw attention to the fact that *the Riesz-Fischer theorems 9.5.2 and 11.4.1 are valid only if Lebesgue's concept of integral is taken for basis*. Just this has furnished one of the most striking and convincing evidence for the significance of Lebesgue's new theory of integral, a few years after its discovery in the early 1900's.

11.5 Convergence almost everywhere of arbitrary orthogonal series

11.51 The main finds of theorems 11.3.1 and 11.4.1 can be expressed in contracted form as follows:

11.5.1 *Every orthogonal series $\sum c_n \varphi_n(t)$ for which $\sum c_n^2 < \infty$ and the orthonormal base system $\{\varphi_n(t)\} \subset L^2(a, b)$ is also complete, is $L^2(a, b)$ -convergent and can be considered as the Fourier series of its $L^2(a, b)$ -sum with respect to $\{\varphi_n\}$.*

This theorem — based upon the very appropriate notion of L^2 -convergence — yields the maximum, if such a theory of general orthogonal series is aimed at, which shows a parallelism in all respects with the geometric model serving as a background. However, no treatment of this field may omit considerations connected with the *classical concept of convergence*; all the less, since the Riesz-Fischer theorem can be used only limitedly for effective determination of the L^2 -limit in question. (Cf. also 9.5.3.)

11.52 As far as the ordinary convergence of an orthogonal series is concerned, we can try the further development of 11.5.1 in two directions. On the one hand, which condition regarding the coefficients $\{c_n\}$ assures the *convergence almost everywhere* of the series $\sum c_n \varphi_n(t)$ in case of an

arbitrary (not necessarily complete) orthonormal base system $\{\varphi_n(t)\}$? On the other hand, maintaining the "natural" coefficient condition $\sum c_n^2 < \infty$ and fixing the complete orthonormal system $\{\varphi_n(t)\}$, we can study the convergence of $\sum c_n \varphi_n(t)$ on a prescribed set of t -values. (It is obvious that in the first case, the condition of convergence everywhere would be entirely inadequate, since the base system remains orthonormal when the definition of the functions $\varphi_n(t)$ is modified on some null sets, but such alterations could lead to divergence of $\sum c_n \varphi_n(t)$ at certain points.)

About the convergence problem mentioned first, it is easy to see that the restriction $\sum c_n^2 < \infty$ does not involve in general the p.p. convergence of $\sum c_n \varphi_n(t)$; this is true only for some special base systems such as $\{r_n(t)\}$ and $\{\chi_m^{(k)}(t)\}$. (Cf. 10.8.) The result is not surprising at all, if we remember the fact that L^2 - and pointwise convergence are concepts of dissimilar nature (cf. 9.52); thus it may not be expected that a condition on $\{c_n\}$ would be of the same effectiveness in case of both kinds of convergence in question.

Anyway, the following elementary proposition holds:

11.5.2 If (a, b) is finite, then the condition $\sum_{n=1}^{\infty} |c_n| < \infty$ involves the absolute convergence of $\sum_{n=1}^{\infty} c_n \varphi_n(t)$ for a.e. $t \in (a, b)$, with respect to an arbitrary orthonormal base system $\{\varphi_n\} \subset L^2(a, b)$.

Proof. By using the Buniakovsky-Schwarz inequality (9.4-5), we can write:

$$\begin{aligned} \sum_{n=1}^{\infty} \int_a^b |c_n \varphi_n(t)| dt &\leq \sum_{n=1}^{\infty} |c_n| \cdot (b-a)^{1/2} \left[\int_a^b \varphi_n(t)^2 dt \right]^{1/2} = \\ &= \sqrt{b-a} \sum_{n=1}^{\infty} |c_n| < \infty, \end{aligned}$$

provided that the series $\sum c_n$ is absolutely convergent.

Consequently, the application of B. Levi's theorem 7.5.3 to the partial sums of the series $\sum c_n \varphi_n(t)$ yields immediately the assertion.

* * *

11.53 The pertinent sharpest result was obtained by H. RADEMACHER and D. MENSHOV in 1922-23 independently of each other.¹ Their basic theorem can be formulated in the following manner:

¹ Cf. Menshov [1]-[3], Rademacher [1].

11.5.3 If the condition

$$\sum_{n=1}^{\infty} c_n^2 \log^2 n < \infty \quad (11.5-1)$$

is fulfilled, then the series $\sum_{n=1}^{\infty} c_n \varphi_n(t)$ is convergent a.e. in case of an arbitrary orthonormal base system $\varphi_n(t) \in L^2(a, b)$ ($n = 1, 2, \dots$).

The assumption (11.5-1) cannot be improved in the sense that if we replace in it $\log^2 n$ by any positive, monotone increasing factor $\lambda(n)$ which tends to infinity more slowly than $\log^2 n$, then there exist a sequence of coefficients $c_1, c_2, \dots$ and a special base system $\{\varphi_n(t)\}$, such that

$$\sum_{n=1}^{\infty} c_n^2 \lambda(n) < \infty, \text{ but } \sum_{n=1}^{\infty} c_n \varphi_n(t) \text{ is divergent for every } t \in (a, b).$$

The deep theorem 11.5.3 has numerous interesting consequences. Instead of its full verification, we can treat here — as an illustration — only a relevant part of the assertion, namely, that on the condition (11.5-1) we can always select from partial sums of the series $\sum c_n \varphi_n(x)$ a subsequence convergent almost everywhere, e.g. those partial sums in which the number of terms is a power of 2.

To show this, let us put $S_N(t) = \sum_{n=1}^N c_n \varphi_n(t)$ ($N = 1, 2, \dots$) and consider the numerical series:

$$\sum_{\nu=0}^{\infty} \int_a^b \nu^2 [S_{2^{\nu+1}}(t) - S_{2^{\nu}}(t)]^2 dt = \sum_{\nu=0}^{\infty} \nu^2 (c_{2^{\nu+1}}^2 + \dots + c_{2^{\nu+1}}^2). \quad (11.5-2)$$

Owing to the elementary inequality $\nu < (\log 2)^{-1} \log(2^{\nu} + h)$ ($h > 0$), any partial sum of the right-hand series in (11.5-2) can be estimated as follows:

$$\begin{aligned} & \sum_{\nu=0}^m \nu^2 (c_{2^{\nu+1}}^2 + \dots + c_{2^{\nu+1}}^2) < \\ & < \sum_{\nu=0}^m \left[\frac{\log^2(2^{\nu} + 1)}{\log^2 2} c_{2^{\nu+1}}^2 + \dots + \frac{\log^2 2^{\nu+1}}{\log^2 2} c_{2^{\nu+1}}^2 \right] = \\ & = (\log 2)^{-2} \sum_{\nu=0}^m \sum_{k=2^{\nu+1}}^{2^{\nu+1}} c_k^2 \log^2 k = (\log 2)^{-2} \sum_{k=2}^{2^{m+1}} c_k^2 \log^2 k. \end{aligned} \quad (11.5-3)$$

Since m may be taken however large, by (11.5-2)—(11.5-3) the condition (11.5-1) involves immediately the convergence of the left-hand series in (11.5-2). Hence, using again theorem 7.5.3, we get that the function series

$$\sum_{\nu=0}^{\infty} \nu^2 [S_{2^{\nu+1}}(t) - S_{2^{\nu}}(t)]^2$$

converges at almost all points $t \in (a, b)$.

Now, by using Chauchy's inequality (2.1-3), it may be written:

$$\begin{aligned} [S_{2^{m+p}}(t) - S_{2^m}(t)]^2 &= \\ &= \left\{ \sum_{\nu=m}^{m+p-1} \frac{1}{\nu} \cdot \nu [S_{2^{\nu+1}}(t) - S_{2^\nu}(t)] \right\}^2 < \\ &< \left(\sum_{\nu=1}^{\infty} \frac{1}{\nu^2} \right) \cdot \sum_{\nu=m}^{\infty} \nu^2 [S_{2^{\nu+1}}(t) - S_{2^\nu}(t)]^2. \end{aligned} \quad (11.5-4)$$

In the last upper bound, the first factor is a convergent series of numbers, while the second factor is the remainder (after $m-1$ terms) of a function series converging a.e. Both factors being independent of p , we see from (11.5-4) that for any $\varepsilon > 0$, however small, we have

$$|S_{2^{m+p}}(t) - S_{2^m}(t)| < \varepsilon \quad (p = 1, 2, \dots),$$

if only m is sufficiently large.

According to Cauchy's criterion of convergence this just means that the "dyadic" subsequence

$$S_{2^\nu}(t) \quad (\nu = 0, 1, 2, \dots)$$

is convergent for almost every $t \in (a, b)$, as asserted above.¹

11.6 The convergence and summation problem of general Fourier series

11.61 Let us also deal with the other fundamental problem of convergence, raised briefly in the previous section: given a complete orthonormal base system $\{\varphi_n(t)\} \subset L^2(a, b)$ and an (L) -measurable set $E \subset (a, b)$, what can be said about the convergence of the series $\sum_{n=1}^{\infty} c_n \varphi_n(t)$ on E , provided

that $\sum_{n=1}^{\infty} c_n^2 < \infty$?

First of all we want to establish: by 11.5.1 the last-mentioned condition ensures the existence of a function $f(t) \in L^2(a, b)$ such that $\sum c_n \varphi_n(t)$ is

¹ The second part of theorem 11.5.3 (relating to divergence) has recently been greatly sharpened in the series of articles [2] by K. TANDORI, furthermore by the paper [1] of L. LEINDLER. The latter author showed e.g. that every special orthonormal system occurring in the divergence theorems may be taken for some suitable sequence of *entire rational functions*, too.

the *Fourier series* of f in terms of $\{\varphi_n(t)\}$. Therefore, the orthogonal series in question is necessarily of the form

$$\sum_{n=1}^{\infty} (f, \varphi_n) \varphi_n(t), \quad (11.6-1)$$

whose N -th partial sum, taken at a point $t = x$, may be written as:

$$S_N(x) = \sum_{n=1}^N \varphi_n(x) \int_a^b f(t) \varphi_n(t) dt = \int_a^b f(t) K_N(t, x) dt. \quad (11.6-2)$$

Here

$$K_N(t, x) = \sum_{n=1}^N \varphi_n(t) \varphi_n(x) \quad (11.6-3)$$

is the so-called N -th *kernel function* of (11.6-1).

Thus our problem demands the examination of the limit

$$\lim_{N \rightarrow \infty} \int_a^b f(t) K_N(t, x) dt \quad (x \in E). \quad (11.6-4)$$

Consequently, the convergence behaviour of (11.6-1) depends both on the choice of f and on the properties of the base system $\{\varphi_n(t)\}$.

11.62 Evidently, the determination of (11.6-4) in general cannot be carried out in the trivial way, i.e., by performing the limiting process $N \rightarrow \infty$ directly under the integral sign. Instead of this, one needs mostly certain subtle considerations and estimations, respectively; and during the examination, it can be a particularly favourable circumstance, if $K_N(t)$ has a representation in *closed form*. It is also understandable that the estimation of $|S_N(x)|$ is connected essentially with the so-called *Lebesgue functions*, defined by

$$\mathcal{L}_N(x) = \int_a^b |K_N(t, x)| dt \quad (N = 1, 2, \dots)^1 \quad (11.6-5)$$

The situation is wholly analogous, when we want to study not the partial sums $S_N(x)$ ($N = 1, 2, \dots$) themselves, but the arithmetic means or other appropriate linear expressions of the latter ("linear summation processes"). In each pertinent case, the discussion of a so-called *singular integral* [of

¹ For instance, let us mention the following theorem (Kaczmarz [1], Kaczmarz-Steinhaus [1]): if a subsequence of the Lebesgue functions (11.6-5) satisfies on a linear set E the inequality $\mathcal{L}_{\nu_n}(x) \leq \text{const} \cdot \lambda_{\nu_n}$ where $\{\lambda_{\nu_n}\}$ is a strictly increasing sequence of positive numbers, then we have also the estimation $|S_{\nu_n}(x)| \leq \text{const} \cdot \sqrt{\lambda_{\nu_n}}$ ($x \in E$; $n = 1, 2, \dots$). Concerning applications, we refer to Tandori [1].

the type (11.6-2)] is required, whose integrand contains — as factor of the function f — the kernel characteristic of the summation method in examination.¹

11.63 From the foregoing we see: actual local results about the convergence or summation problem of a Fourier series can be obtained only in such cases where the properties of the pertaining base system $\{\varphi_n\}$ as well as of the corresponding kernel function (11.6-3) are known in a satisfactory manner. This programme will be realized in the next chapters.

It should be noted that the rapid development of the convergence theory of special and general orthogonal series during the last decades is also connected with the contributions of several mathematicians belonging to Hungarian analytic schools, first of all to those of L. FEJÉR, F. and M. RIESZ, A. HAAR, O. SZÁSZ, G. SZEGŐ and G. ALEXITS.²

¹ For the general definition of the concept of singular integral see e.g. Lebesgue [7]. It will be not needed here.

² Cf. e.g. Alexits [2], Fejér [3], Haar [1], Hilb-M. Riesz [1], Plessner [1], F. Riesz-Sz.-Nagy [1], O. Szász [1], Szegő [1]-[2].

CONVERGENCE OF SPECIAL FOURIER SERIES

12.1 Elementary convergence theorems on ordinary Fourier series; the Riemann–Lebesgue lemma

12.11 If $\{\varphi_n(t)\}$ is the trigonometric base system (10.1–1), then the Fourier series (11.1–1) turns for $t = x$ into the *ordinary or trigonometric Fourier series*

$$\alpha_0 + \sum_{n=1}^{\infty} (\alpha_n \cos nx + \beta_n \sin nx), \quad (12.1-1)$$

where

$$\left. \begin{aligned} \alpha_0 &= \frac{1}{2\pi} \int_a^{a+2\pi} f(t) dt, \\ \alpha_n &= \frac{1}{\pi} \int_a^{a+2\pi} f(t) \cos nt dt \\ \beta_n &= \frac{1}{\pi} \int_a^{a+2\pi} f(t) \sin nt dt \quad (n = 1, 2, \dots) \end{aligned} \right\} \quad (12.1-2)$$

[cf. (10.2–1)–(10.2–3)]. Since for all n we have $|\cos nt| \leq 1$, $|\sin nt| \leq 1$, so the *ordinary Fourier constants* (12.1–2) can be formed not only for square integrable functions, but for *any* (L)-integrable function f as well. Thus we shall assume that $f \in L[a, a + 2\pi]$ (a fixed) when speaking about the series (12.1–1) with (12.1–2).

Furthermore, it will be supposed in this context steadily (also without saying it explicitly) that $f(t)$ is *periodic with period* 2π . In the trigonometric base system each term has the period 2π , the sum of the series (12.1–1) possesses plainly the same property whenever it does exist. We have also to stress that it is possible to construct a function with some prescribed period from any single-valued real function defined in a finite interval, by so-called “*periodic continuation*”. On the other hand, no restriction lies in the condition that the period of $f(t)$ should possess just the value 2π , since an arbitrary function $F(u)$ with the period $p > 0$ becomes periodic with 2π if the coordinate transformation $u = (p/2\pi)t$ is applied.

The last assumption implies that $f(t) = F(tp/2\pi)$ has the period 2π , so that the ordinary Fourier series of $F(u)$ pertaining to the fundamental

interval $[a, a + p]$ taken at a point $u = x$ may be written in the form

$$A_0 + \sum_{n=1}^{\infty} \left[A_n \cos \left(\frac{2n\pi}{p} x \right) + B_n \sin \left(\frac{2n\pi}{p} x \right) \right], \quad (12.1-3)$$

where the coefficients are:

$$\begin{aligned} A_0 &= \frac{1}{p} \int_a^{a+p} F(u) du \\ A_n &= \frac{2}{p} \int_a^{a+p} F(u) \cos \left(\frac{2n\pi}{p} u \right) du \\ B_n &= \frac{2}{p} \int_a^{a+p} F(u) \sin \left(\frac{2n\pi}{p} u \right) du \quad (n = 1, 2, \dots). \end{aligned} \quad (12.1-4)$$

The situation that the interval of integration in the coefficient formulae (12.1-2) may be an arbitrary line segment of length 2π , since the Fourier constants do not depend on a because of the periodicity (with 2π) of $f(t)$, is useful mainly in the case of an even or odd function f . Namely, if $f(t)$ is e.g. even, $f(t) \sin nt$ is odd, consequently putting $a = -\pi$ we get for every n :

$$\beta_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt dt = \frac{1}{\pi} \left(\int_{-\pi}^0 + \int_0^{\pi} \right) = 0.$$

Similarly, $\alpha_n = 0$ ($n = 0, 1, 2, \dots$) whenever f is an odd function. Thus we may point out: the ordinary Fourier series of an even function is of the form $\sum_{n=0}^{\infty} \alpha_n \cos nx$, i.e., it is a so-called *cosine series*; in turn, the ordinary Fourier series of an odd function is always of the type $\sum_{n=1}^{\infty} \beta_n \sin nx$, i.e., a so-called *sine series*.

12.12 The convergence properties follow in certain simple cases immediately by estimating the Fourier constants α_n, β_n ; so first it is advisable to study the coefficients (12.1-2).

As the most trivial case may be considered that of a function $f(t)$ which is k -times continuously differentiable at each point of $[a, a + 2\pi]$ ($k \geq 1$, fixed). Then k -fold integration by parts yields the formulae:

$$\alpha_n = \frac{(-1)^k}{\pi n^k} \int_a^{a+2\pi} f^{(k)}(t) \cos \left(nt - k \frac{\pi}{2} \right) dt$$

$$(n = 1, 2, \dots),$$

$$\beta_n = \frac{(-1)^k}{\pi n^k} \int_a^{a+2\pi} f^{(k)}(t) \sin \left(nt - k \frac{\pi}{2} \right) dt \quad (12.1-5)$$

whence by the notation

$$D_k = \frac{1}{\pi} \int_a^{a+\pi} |f^{(k)}(t)| dt \quad (12.1-6)$$

we obtain the estimates

$$|\alpha_n| \leq \frac{D_k}{n^k}, \quad |\beta_n| \leq \frac{D_k}{n^k}. \quad (12.1-7)$$

Since D_k is independent of n according to (12.1-6), we can state: if $f^{(k)}(t)$ exists and is continuous everywhere, then the ordinary Fourier constants of $f(t)$ tend to zero for $n \rightarrow \infty$ at least as rapidly as $1/n^k$.

Taking $k = 2$, we see that (12.1-1) is majorizable for all x by the convergent numerical series $|\alpha_0| + 2D_2 \sum_{n=1}^{\infty} n^{-2}$, and therefore the Fourier series is uniformly convergent in $(-\infty, \infty)$ by the well-known criterion of Weierstrass. Thus any series resulting from (12.1-1) by multiplication by some term of the trigonometric base system may be integrated termwise, so that (12.1-1) is in this case necessarily the ordinary Fourier series of its sum function. (Cf. 9.63.) Besides, this sum must be identical everywhere with $f(x)$ for following reasons: 1) the trigonometric base system is complete in virtue of 10.1.2; 2) consequently, any two functions having the same Fourier constants can differ from one another at most on a null set E (cf. 11.32); 3) in our case, $E = \emptyset$ since it is about a sum function which is continuous at any point x by the above.

Summing up, we got an "elementary" expansion theorem:

12.1.1 If a periodic function f is twice continuously differentiable everywhere, then its ordinary Fourier series converges uniformly to $f(x)$ for $-\infty < x < \infty$.

12.13 It is noteworthy that (12.1-5)—(12.1-7) remain valid also on more general conditions about f . For instance, integrating by parts by

means of formula (7.6-24), we see that *instead of the existence and continuity of $f^{(k)}(t)$ it suffices to require the existence and absolute continuity of $f^{(k-1)}(t)$.*¹

Another possibility of sharpening is opened by the transformation of Fourier constants into *Stieltjes integrals*. If $f(t)$ is assumed to be of bounded variation, we can write on the basis (5.7-8), (5.8-1) and (5.7-16):

$$\left. \begin{aligned} \pi|\alpha_n| &= \left| \int_a^{a+2\pi} f(t) \cos nt \, dt \right| = \frac{1}{n} \left| \int_a^{a+2\pi} \sin nt \, df(t) \right| \leq \frac{1}{n} V_f, \\ \pi|\beta_n| &= \left| \int_a^{a+2\pi} f(t) \sin nt \, dt \right| = \frac{1}{n} \left| \int_a^{a+2\pi} \cos nt \, df(t) \right| \leq \frac{1}{n} V_f; \end{aligned} \right\} \quad (12.1-8)$$

V_f denoting the total variation of the function f in $[a, a + 2\pi]$.

Thus we may establish: *if $f(t)$ is of bounded variation on a line segment of length 2π , then its ordinary Fourier constants tend to 0 in the order $O(1/n)$ as $n \rightarrow \infty$.*²

Correspondingly, the premise of proposition 12.1.1 can be made less strict: *the twofold continuous differentiability of f may be replaced by the conditions that $f(t)$ should be continuous everywhere and $f'(t)$ should be a function of bounded variation in the interval $[a, a + 2\pi]$.*³

In particular, using geometric wording:

12.1.2 *Every function f continuous everywhere and of the period 2π has an ordinary Fourier series uniformly convergent in $(-\infty, \infty)$ and is expressible at any point as the sum of the latter, provided that 1) the diagram of f on some finite interval of length 2π consists of a finite number of line segments or convex or concave curve arcs; 2) the semi-tangents of the latter pertaining to the end points are not perpendicular to the axis of abscissae.*

12.14 The foregoing considerations suggest at once the question: whether the Fourier constants of any square integrable function tend to 0 for $n \rightarrow \infty$ as well?

The answer is affirmative and can easily be obtained from the *Parseval-Hurwitz formula with respect to the trigonometric base system*.⁴ The ortho-

¹ This involves, as we know, the existence and (L) -integrability a.e. of $f^{(k)}(t)$. (Cf. 7.6.6.) Applying lemma 12.1.3 (to be treated soon) to the integrals in (12.1-5), it turns out also that $n^k \alpha_k$ and $n^k \beta_k$ tend to 0 as $n \rightarrow \infty$, i.e. $\alpha_n = o(n^{-k})$, $\beta_n = o(n^{-k})$. In the case $k = 1$, the above condition becomes simpler by $f^{(0)}(t) = f(t)$.

² Littlewood [2].

³ F. Riesz [14].

⁴ Cf. (11.2-2) and 11.23.

normality and completeness of the system (10.1-6) (cf. 10.1.2) implies for an arbitrary function $f(t) \in L^2(a, a + 2\pi)$ by 11.2.1:

$$2\alpha_0^2 + \sum_{n=1}^{\infty} (\alpha_n^2 + \beta_n^2) = \frac{1}{\pi} \int_a^{a+2\pi} f(t)^2 dt; \quad (12.1-9)$$

and from $\sum_{n=1}^{\infty} (\alpha_n^2 + \beta_n^2) < \infty$ we conclude the limit relations:

$$\lim_{n \rightarrow \infty} \alpha_n = 0 \quad \lim_{n \rightarrow \infty} \beta_n = 0. \quad (12.1-10)$$

Now, we want to show that this result can still be developed further, as it is true also for arbitrary (L) -integrable functions, moreover, it may be extended to the case of a continuous parameter μ instead of the discrete variable n , and to an arbitrary finite fundamental interval (a, b) instead of $(a, a + 2\pi)$. The lemma is simple but of basic significance for the whole convergence theory of trigonometric Fourier series.

12.1.3 (RIEMANN-LEBESGUE LEMMA)¹ *If $F(t)$ is (L) -integrable in a finite interval (a, b) , then the integrals*

$$\int_a^b F(t) \cos \mu t dt, \quad \int_a^b F(t) \sin \mu t dt \quad (12.1-11)$$

tend to 0 as $|\mu| \rightarrow \infty$.

Proof. Suppose $F(t) \in L(a, b)$ and $F(t) = 0$ ($t \notin [a, b]$). Furthermore let $\mu > 0$, this being plainly no restriction of generality. The integral

$$I_\mu = \int_a^b F(t) e^{i\mu t} dt \quad (12.1-12)$$

turns by the substitution $t = x + \frac{\pi}{\mu}$ into

$$I_\mu = - \int_a^b F\left(x + \frac{\pi}{\mu}\right) e^{i\mu x} dx - \int_{a-\frac{\pi}{\mu}}^a F\left(x + \frac{\pi}{\mu}\right) e^{i\mu x} dx, \quad (12.1-13)$$

¹ The assertion does not hold for non-absolutely convergent improper integrals, which is understandable by the geometric background too. [There is a kind of "marring" sign effect in the integrands (12.1-11) as $\mu \rightarrow \infty$.] The limit relations in question were verified by RIEMANN for (R) -integrable functions in 1854, while the extension to (L) -integrable functions is due to LEBESGUE (1905). (Cf. Riemann [1], Lebesgue [5]-[6].)

so that we get from (12.1-12) and (12.1-13) by addition:

$$2I_\mu = \int_a^b \left[F(t) - F\left(t + \frac{\pi}{\mu}\right) \right] e^{i\mu t} dt + \int_a^{a+\frac{\pi}{\mu}} f(t) e^{i\mu t} dt. \quad (12.1-14)$$

Accordingly

$$2|I_\mu| \leq \int_a^b \left| F\left(t + \frac{\pi}{\mu}\right) - F(t) \right| dt + \int_a^{a+\frac{\pi}{\mu}} |F(t)| dt; \quad (12.1-15)$$

and hence we see by 7.5.4 that I_μ as well as its real and imaginary parts [i.e. both integrals (12.1-11)] tend to 0 as $\mu \rightarrow \infty$.

* * *

12.15 Note that in the case of ordinary Fourier constants α_n, β_n ($n \geq 1$), the pendants of (12.1-14)–(12.1-15) are:

$$\left. \begin{aligned} \alpha_n &= \frac{1}{2\pi} \int_a^{a+2\pi} \left[f(t) - f\left(t + \frac{\pi}{n}\right) \right] \cos nt \, dt \\ \beta_n &= \frac{1}{2\pi} \int_a^{a+2\pi} \left[f(t) - f\left(t + \frac{\pi}{n}\right) \right] \sin nt \, dt \end{aligned} \right\} \quad (12.1-16)$$

and

$$\left\{ \begin{aligned} |\alpha_n| \\ |\beta_n| \end{aligned} \right\} \leq \frac{1}{2\pi} \int_a^{a+2\pi} \left| f\left(t + \frac{\pi}{n}\right) - f(t) \right| dt, \quad (12.1-17)$$

respectively. The inequalities (12.1-17) may be useful when we want to know: *how strongly* tend to 0 the ordinary Fourier constants of some actual integrable function f ?

We have still to mention that the proof of lemma 12.1.3 enables us to supplement the assertion in a sense: if $F(t)$ is (L) -integrable in a finite or infinite interval (A, B) , then to every given number $\eta > 0$, however small, we can find a number $\mu_o > 0$, depending on η only, such that the inequalities

$$\left| \int_a^b F(t) \cos \mu t \, dt \right| < \eta, \quad \left| \int_a^b F(t) \sin \mu t \, dt \right| < \eta$$

hold for $|\mu| > \mu_0$ and for all subintervals $(a, b) \subset (A, B)$, simultaneously. This is briefly expressed by saying that the integrals (12.1-11) converge *uniformly* to zero as $|\mu| \rightarrow \infty$, with respect to a and b .

12.2 Dirichlet's formula and Riemann's localization theorem

12.21 We now study the *local convergence behaviour* of ordinary Fourier series.

In case of (12.1-1), the N -th partial sum can be written in closed integral form ($a = -\pi$, $N = 0, 1, 2, \dots$):

$$\begin{aligned} S_N(x) &= \alpha_0 + \sum_{n=1}^N (\alpha_n \cos nx + \beta_n \sin nx) = \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) dt + \frac{1}{\pi} \sum_{n=1}^N \left[\int_{-\pi}^{\pi} f(t) \cos nt \cdot \cos nx dt + \right. \\ &\quad \left. + \int_{-\pi}^{\pi} f(t) \sin nt \cdot \sin nx dt \right] = \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left[1 + 2 \sum_{n=1}^N (\cos nt \cdot \cos nx + \sin nt \cdot \sin nx) \right] dt = \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left[1 + 2 \sum_{n=1}^N \cos n(t-x) \right] dt = \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \frac{\sin(2N+1)\frac{t-x}{2}}{\sin \frac{t-x}{2}} dt. \end{aligned}$$

Here the well-known trigonometric sum formula:

$$\begin{aligned} 1 + 2 \cos \Theta + 2 \cos 2\Theta + \dots + 2 \cos n\Theta &= \\ = \frac{\sin(2N+1)\frac{\Theta}{2}}{\sin \frac{\Theta}{2}} \quad (\Theta \neq 0, \pm 2\pi, \pm 4\pi, \dots) \end{aligned} \quad (12.2-1)$$

was utilized.

Applying the substitution $t = x + u$ and splitting up the fundamental interval $[-\pi, \pi]$ in two (symmetric) parts, it follows:

$$S_N(x) = \frac{1}{\pi} \int_0^{\pi} \varphi_x(u) D_N(u) du = \frac{2}{\pi} \int_0^{\pi/2} \varphi_x(2u) \frac{\sin(2N+1)u}{\sin u} du, \quad (12.2-2)$$

where

$$\varphi_x(u) = \frac{f(x+u) + f(x-u)}{2} \quad \text{and} \quad D_N(u) = \frac{\sin(2N+1)\frac{u}{2}}{\sin \frac{u}{2}}. \quad (12.2-3)$$

Putting in (12.2-2) $f(u) \equiv 1$, $\varphi_x(u) \equiv 1$ and introducing the notation

$$\psi_x(u) = \varphi_x(u) - f(x) = \frac{f(x+u) + f(x-u)}{2} - f(x), \quad (12.2-4)$$

we get by subtraction:

$$\begin{aligned} S_N(x) - f(x) &= \frac{1}{\pi} \int_0^{\pi} \psi_x(u) D_N(u) du = \\ &= \frac{2}{\pi} \int_0^{\pi/2} \psi_x(2u) \frac{\sin(2N+1)u}{\sin u} du. \end{aligned} \quad (12.2-5)$$

The name of the (equivalent) representations (12.2-2) or (12.2-5) is *Dirichlet's formula*, while the factor $D_N(u)$ independent of the choice of f is called *Dirichlet's kernel function* of the index N [cf. (11.6-2)—(11.6-3)].¹ As regards the latter, it is relevant that $|D_N(u)|$ has for any fixed $u \neq 0$ an upper bound independent of N (namely $1/\sin u$), but for $u = 0$ by (12.2-1) we have

$$D_N(0) = 2N + 1 \rightarrow \infty \quad (N \rightarrow \infty).$$

Therefore, point $u = 0$ is an exceptional (singular) locus for the integrand which motivates that the integrals in (12.2-2) and (12.2-5) are considered also as examples of *singular integrals*. (Cf. 11.62.) Another important property of $D_N(u)$ has already been used implicitly before, when (12.2-5) was deduced by means of (12.2-1); it is:

$$\int_0^{\pi} D_N(u) du = \pi \quad (N = 0, 1, 2, \dots). \quad (12.2-6)$$

¹ The introduction of the formulae just mentioned is due to P. G. L. DIRICHLET, who in a pioneering work of the year 1829 first deduced some fundamental results of the convergence theory of ordinary Fourier series.

12.22 As far as the behaviour of the so-called *Dirichlet's integral* (12.2-2) or (12.2-5) for the limiting process $N \rightarrow \infty$ is concerned, the passage to the limit cannot be carried out under the integral sign by the above, i.e., we have to look for certain special ways of discussion. Our starting point is furnished by a quite natural idea: since the only singularity of $D_N(u)$ lies at the left-hand boundary of the integration interval, it seems advisable to "separate" the right-hand neighbourhood of the origin and to discuss the Dirichlet integral decomposed into two corresponding parts.

We use the second representation (12.2-5), and according to the last remark we write:

$$S_N(x) - f(x) = \frac{2}{\pi} \int_0^{\delta} \frac{\psi_x(2u)}{\sin u} \sin(2N+1)u \, du + \frac{2}{\pi} \int_{\delta}^{\pi/2} \frac{\psi_x(2u)}{\sin u} \sin(2N+1)u \, du, \quad (12.2-7)$$

δ denoting an arbitrary fixed positive number less than $\pi/2$.

In the second term on the right, the factor $\psi_x(2u)/\sin u$ is (L) -integrable (by the boundedness of $1/\sin u$), so that the application of the Riemann-Lebesgue lemma with the notations $a = \delta$, $b = \pi/2$, $F(u) = \psi_x(2u)/\sin u$, $\mu = 2N+1$ yields:

$$\lim_{N \rightarrow \infty} \int_{\delta}^{\pi/2} \frac{\psi_x(2u)}{\sin u} \sin(2N+1)u \, du = 0. \quad (12.2-8)$$

Hence we see that the limit relation

$$\lim_{N \rightarrow \infty} [S_N(x) - f(x)] = 0,$$

i.e., the convergence of the ordinary Fourier series of f at a point x to the pertaining value of the function depends only thereupon, whether the first subintegral in (12.2-7) tends to 0 for $N \rightarrow \infty$ as well; and by the definition of $\psi_x(u)$, the integrand of this term relating to $[0, \delta]$ contains only such values of f which are associated with the interval $x - 2\delta \leq u \leq x + 2\delta$. Since $\delta > 0$ is as small as we please, we come finally to the conclusion:

12.2.1 (RIEMANN'S LOCALIZATION THEOREM)¹ *The convergence behaviour of the ordinary Fourier series of some function f at a point x is*

¹ Riemann [1].

wholly determined by the values of f in a however small neighbourhood of this point.

The surprising property of the series (12.1-1) implied by 12.1.1 indicates that any modification of the definition of function f outside a δ -neighbourhood of x does not affect the sum of the ordinary Fourier series of f at the point x .

Note that the sharpening about uniform convergence of 12.1.3 involves a complement of similar nature in case of 12.2.1, too. If $f_1(t)$, $f_2(t)$ are two functions with period 2π , (L) -integrable in $(-\pi, \pi)$ and identical in an interval (a, b) of length $< 2\pi$, and denoting by $S_N^{(1)}(x)$, and $S_N^{(2)}(x)$ the N -th partial sum of the ordinary Fourier series of f_1 and f_2 at point x , respectively, then the limit relation

$$\lim_{N \rightarrow \infty} [S_N^{(1)}(x) - S_N^{(2)}(x)] = 0$$

holds uniformly in every inner subinterval of (a, b) .

12.23 Investigations on the first subintegral in (12.2-7) may be replaced by that of the simpler integral

$$I_N^{(\delta)}[f; x] = \int_0^{\delta} \frac{\psi_x(2u)}{u} \sin(2N+1)u \, du. \quad (12.2-9)$$

On the basis of L'Hospital's rule we get

$$\frac{1}{\sin u} - \frac{1}{u} = \frac{u - \sin u}{u \sin u} \rightarrow 0 \quad (u \rightarrow 0); \quad (12.2-10)$$

thus the product of this difference by $\psi_x(2u)$ is (L) -integrable in $(0, \delta)$, and by lemma 12.1.3 we have:

$$\lim_{N \rightarrow \infty} \int_0^{\delta} \psi_x(2u) \left(\frac{1}{\sin u} - \frac{1}{u} \right) \sin(2N+1)u \, du = 0. \quad (12.2-11)$$

Herewith the following necessary and sufficient criterion is deduced:

12.2.2 A function f may be written as the sum of its ordinary Fourier series at point x if and only if to any given $\varepsilon > 0$, however small, there exists a positive number δ less than $\pi/2$ such that for N large enough

$$|I_N^{(\delta)}[f, x]| < \varepsilon. \quad (12.2-12)$$

12.3 Convergence tests of Dini, Dirichlet, Jordan and Lipschitz

12.3.1 Our last result yields a number of important *sufficient* conditions of convergence, among which criterion 12.3.1 — found for (R) -integrable functions in 1880 — can be verified simply.

We introduce here an abbreviating term: a function f is said to be *Fourier expansible* at a point x , if its ordinary Fourier series converges there to the function value $f(x)$ in the classical sense.

12.3.1 (DINI'S CRITERION)¹ If $\psi_x(2u)/u$ (as a function of u) is (L) -integrable in a right-hand neighbourhood of the point $u = 0$, then the function $f(t)$ is Fourier expansible at x .

P r o o f. As assumed, the integral (12.2-9) exists for some suitable $\delta > 0$ and it tends to 0 as $N \rightarrow \infty$ on the basis of 12.1.3. Thus 12.2.2 yields directly the assertion.

* * *

Another quite general condition (from 1881) may be formulated as follows:

12.3.2 (JORDAN'S CRITERION)² If $\psi_x(u)$ is of bounded variation in a right-hand neighbourhood of point $u = 0$ and $\psi_x(+0) = 0$, then function $f(t)$ is Fourier expansible at $t = x$.

P r o o f. We need the elementary inequality:

$$\left| \int_0^x \frac{\sin t}{t} dt \right| < 3 \quad (x > 0). \quad (12.3-1)$$

For $x \in (0, 1]$ this is trivial by

$$\left| \int_0^x \frac{\sin t}{t} dt \right| < \int_0^1 1 dt = 1 \quad (0 < x \leq 1).$$

¹ Dini [1]. For the definition of the function ψ_x , see (12.2-3)—(12.2-4).

² C. Jordan [1].

On the other hand, in case of $x > 1$ we can write:

$$\left| \int_1^x \frac{\sin t}{t} dt \right| = \left| \left[-\frac{\cos t}{t} \right]_1^x - \int_1^x \frac{\cos t}{t^2} dt \right| \leq \\ \leq 1 + \frac{1}{x} + \int_1^x \frac{dt}{t^2} = 2,$$

so we obtain for the integral in question:

$$\left| \int_0^x \frac{\sin t}{t} dt \right| \leq \left| \int_0^1 \frac{\sin t}{t} dt \right| + \left| \int_1^x \frac{\sin t}{t} dt \right| < 3 \quad (x > 1).$$

By (12.3-1), the absolute value of any integral of the type

$$j_N(u) = \int_0^u \frac{\sin(2N+1)v}{v} dv = \int_0^{(2N+1)u} \frac{\sin t}{t} dt \\ (u > 0; \quad N = 0, 1, 2, \dots) \quad (12.3-2)$$

is also less than 3.

Suppose now that $\psi_x(u)$ is of bounded variation in an interval $0 \leq u \leq 2\delta$. Then $\psi_x(u)$ has one-sided limits at every point $u \in (0, 2\delta)$, further $\psi_x(+0)$ and $\psi_x(2\delta-0)$ exist as well (cf. 4.5.2). Consider the integral (12.2-9) in the form:

$$I_N^{(\delta)}[f; x] = \int_0^\delta [\psi_x(2u) - \psi_x(2\delta-0)] \frac{\sin(2N+1)u}{u} du + \\ + \psi_x(2\delta-0) \int_0^\delta \frac{\sin(2N+1)u}{u} du.$$

The first term on the right-hand side can be estimated on the basis of integration by parts and by means of (12.3-2) as follows [cf. (5.7-16)]:

$$\left| \int_0^\delta [\psi_x(2u) - \psi_x(2\delta-0)] \frac{\sin(2N+1)u}{u} du \right| = \\ = \left| \int_0^\delta j_N(u) d\psi_x(2u) \right| < 3V_{\psi_x}(0, 2\delta),$$

the last factor denoting the total variation of $\psi_x(u)$ on $[0, 2\delta]$.

On the condition $\psi_x(+0) = 0$, the second term of $I_N^{(\delta)}[f; x]$ satisfies the inequality [cf. (4.4-2) and (12.3-2)]:

$$\left| \psi_x(2\delta - 0) \int_0^\delta \frac{\sin(2N+1)u}{u} du \right| = |\psi_x(2\delta - 0) - \psi_x(+0)| |j_N(\delta)| < < 3V_{\psi_x}(0, 2\delta).$$

Therefore we get by addition:

$$\left| I_N^{(\delta)}[f; x] \right| < 6V_{\psi_x}(0, 2\delta),$$

and the upper bound on the right-hand side becomes arbitrarily small when δ is chosen suitably. This implies by 12.2.2 the desired result.

* * *

12.32 From the above tests, we deduce easily such simple criteria which concern instead of the auxiliary function $\psi_x(u)$ the function $f(t)$ itself, and so are applicable directly.

We begin by a denomination: if both the right-hand and the left-hand limit of $f(t)$ at point x exist and the function value at x is defined by $f(x) = \frac{1}{2}[f(x+0) + f(x-0)]$, then we say that x is a *regular point* of f . (Accordingly, each point of continuity is a regular point of the function, while any point of discontinuity of the first kind (cf. 1.61) becomes regular first by an appropriate choice of the pertaining function value.)

The outstanding role of regular points in connection with convergence properties of ordinary Fourier series may be illustrated by the condition $\psi_x(+0) = 0$, i.e., by the limit relation

$$\lim_{u \rightarrow +0} \frac{f(x+u) + (f(x-u))}{2} = f(x), \quad (12.3-3)$$

in 12.3.2. This restriction can be fulfilled in a most natural way by assuming the existence of the one-sided limits $f(x+0)$, $f(x-0)$ and by equalizing $f(x)$ with the arithmetic mean of the latter.

As a corollary of 12.3.1, we now establish a further very useful test, discovered in 1864.

12.3.3 (CRITERION OF LIPSCHITZ)¹ Let x be a regular point of $f(t)$. If there exist a right-hand and a left-hand neighbourhood of x such that for some constant $K > 0$ and exponent $\alpha > 0$ the inequalities

¹ Cf. Lipschitz [1].

$$|f(x+l) - f(x+0)| \leq Kl^\alpha \quad (12.3-4)$$

$$(0 < l \leq \delta)$$

$$|f(x-l) - f(x-0)| \leq Kl^\alpha \quad (12.3-5)$$

hold simultaneously, then $f(t)$ is Fourier expansible at the point $t = x$.

P r o o f. On the basis of (12.3-3) as well as of (12.3-4)—(12.3-5), we have for $0 < u < 2\delta$:

$$\left| \frac{\psi_x(2u)}{u} \right| = \frac{1}{2u} |f(x+2u) + f(x-2u) - [f(x+0) + f(x-0)]|$$

$$\leq \frac{1}{2u} (|f(x+2u) - f(x+0)| + |f(x-2u) - f(x-0)|) \leq$$

$$\leq 2K(2u)^{\alpha-1}. \quad (12.3-6)$$

Since $u^{\alpha-1}$ ($\alpha > 0$) is (L) -integrable in the interval $0 < u < 2\delta$, by (12.3-6) we can say the same about $\psi_x(2u)/u$ as a function of u , so that 12.3.1 leads to the assertion.

* * *

The content of (12.3-4), (12.3-5) is expressed by saying that $f(t)$ satisfies a pointwise *right or left Lipschitz condition of order α with constant K* at $t = x$; the corresponding notations are:

$$f(t) \in \text{Lip}_K \alpha \quad (t \rightarrow x+0) \quad \text{and} \quad f(t) \in \text{Lip}_K \alpha \quad (t \rightarrow x-0),$$

respectively.¹

Especially, if x is a point of continuity and

$$|f(x+l) - f(x)| \leq K|l|^\alpha \quad (-\delta < l < \delta) \quad (12.3-7)$$

holds, then we say that $f(t)$ fulfils a *bilateral Lipschitz condition of order α with the constant K* at the point x ; i.e.: $f(t) \in \text{Lip}_K \alpha$ ($t = x$). [It is clear that (12.3-7) involves both of (12.3-4) and (12.3-5).]

If $\alpha > 1$, (12.3-4) and (12.3-5) yield that the right-hand and left-hand difference quotient, respectively, of $f(t)$ at x tends to 0 as $l \rightarrow +0$; further

¹ These denominations and notations are often transferred also to some *interval* as follows: we mean by $f(t) \in \text{Lip}_K \alpha$ ($t \in [a, b]$) that for any two points u and x of $[a, b]$ the inequality $|f(u) - f(x)| \leq K|u - x|^\alpha$ holds.

(12.3-7) implies in this case that $f'(x) = 0$. Hence we see that only the orders with $0 < \alpha \leq 1$ are of significance. In particular, the *Lipschitz conditions of order 1* have a remarkably simple geometric meaning: the *boundedness* of a unilateral or the bilateral difference quotient of f in the corresponding neighbourhood of point x . (Cf. 4.62.)

Observe that each of the above Lipschitz conditions implies more than the corresponding (unilateral or bilateral) kind of pointwise continuity, but less than the common derivability; these conditions are situated "between" continuity and differentiability.

Thus we arrive from 12.3.3 at a local test particularly convenient for practice:

12.3.4 *The Fourier expansibility of $f(t)$ is assured at any regular point where the image curve of the function possesses a half-tangent both from the right and from the left, further the latter are not perpendicular to the axis of abscissae. The assertion holds specially at any point where $f(t)$ is differentiable.*

12.33 The other practically most significant test was historically the first one (1829), but can be obtained also as an immediate consequence of 12.3.2.

12.3.5 (DIRICHLET'S CRITERION)¹ *If to a regular point x of $f(t)$ we can find a right-hand neighbourhood $(x, x + \delta)$ and a left-hand one $(x - \delta, x)$ such that the function is monotonic in both of these,² then $f(t)$ is Fourier expansible at $t = x$.*

P r o o f. The assumptions clearly imply $\psi_x(+0) = 0$ [cf. (12.3-3)], moreover the decomposition

$$\psi_x(u) = \frac{1}{2} [f(x+u) - f(x+0)] + \frac{1}{2} [f(x-u) - f(x-0)]$$

shows that $\psi_x(u)$ as the sum of two monotonic functions is of bounded variation in the interval $0 < u < \delta$ (cf. 4.4.1).

Therefore the premise of Jordan's criterion is fulfilled.

* * *

¹ See Dirichlet [1].

² Monotony of f is not required to be of the same sense on both sides of the point x ; for instance, the function is permitted to be monotone increasing on the right, monotone decreasing on the left of x .

We have to stress that the last two tests show features of function $f(t)$ quite different in nature. So it is not uninteresting to show in *examples* that none of them is a consequence of the other.¹

I. The function

$$y = \begin{cases} t^\alpha \sin \frac{1}{t} & \text{if } t \neq 0 \\ 0 & \text{if } t = 0 \end{cases} \quad (0 < \alpha \leq 1) \quad (12.3-8)$$

satisfies bilateral Lipschitz condition of order α at point $t = 0$, but it does not fulfil Dirichlet's criterion in any neighbourhood of the origin. [Cf. (4.1-1).]

II. In the case of the function

$$y = \begin{cases} 1/\log(1/|t|) & \text{if } t \neq 0, \\ 0 & \text{if } t = 0, \end{cases} \quad (12.3-9)$$

we have an opposite situation: this is monotonic in any right or left neighbourhood of $t = 0$, but does not satisfy a Lipschitz condition of the type (12.3-7) at the origin, however the order α may be chosen.

It must be added that both of the functions in question are obviously continuous but not differentiable at $t = 0$; furthermore by putting $\alpha = 1$ in (12.3-8), we get the most elementary example of a function continuous but not of bounded variation near to the origin.

12.34 As regards the sufficient criteria discussed, in each variant of them relating to a *finite closed interval*, *uniform convergence* may also be asserted, if we require at the same time the *continuity* of the pertinent function. (This follows partly from Heine's theorem, partly from the sharpened form of the Riemann-Lebesgue lemma.)

For example, Dirichlet's criterion 12.3.5 allows the following "global" supplementary statement:

12.3.6 *The ordinary Fourier series of $f(t)$ (periodic with 2π) converges everywhere uniformly to the function, if f is continuous for all t and its diagram on a finite line segment of length 2π can be decomposed into a finite number of monotone increasing or monotone decreasing curve arcs.*

The last condition may also be formulated in such a way that $f(t)$ should be in $[0, 2\pi]$ a continuous function having finitely many local maxima or minima. This is plainly satisfied by any graph which can be drawn at all.

We add still: since the function value $f(x)$ in (12.2-4) and (12.2-5) may be replaced by any arbitrary real number s , an easy rewording of the convergence criteria discussed hitherto yields conditions in order that the sum of the ordinary Fourier series of a function f at a given point x should be

¹ The consequences of all the above convergence tests are investigated thoroughly e.g. in Hardy [4].

just s and [assuming the existence of $f(x \pm 0)$] the mean $\frac{1}{2} [f(x+0) + f(x-0)]$, respectively.

Finally, the most general local convergence criterion of the literature, due to LEBESGUE (1905)¹ is mentioned here. According to this test — which implies of course both 12.3.1 and 12.3.2 — *the simultaneous validity of the limit relations*

$$\lim_{l \rightarrow +0} \frac{1}{l} \int_0^l \varphi_x(u) du = s, \quad (12.3-10)$$

$$\lim_{l \rightarrow +0} \int_l^\delta \frac{|\varphi_x(u+l) - \varphi_x(u)|}{u} du = 0 \quad (\delta > 0 \text{ fixed}) \quad (12.3-11)$$

[cf. (12.2-3)] assures that the ordinary Fourier series of $f(t)$ at the point $t = x$ is convergent and its sum equals s .

12.4 Example of a continuous function with ordinary Fourier series divergent at some point

12.41 Now we see that the local Fourier expansibility of function f can only slightly be brought into connection with the simplest function properties based on limit concept: the continuity and differentiability.

Trivial examples show that continuity of a function is *not necessary* to its Fourier expansibility at the same point. Thus the functions

$$C(t) = \begin{cases} -\pi/4 & \text{if } -\pi < t < 0, \\ 0 & \text{if } t = 0, \\ \pi/4 & \text{if } 0 < t < \pi, \end{cases} \quad C(t+2\pi) = C(t), \quad (12.4-1)$$

and

$$S(t) = \begin{cases} \frac{\pi-t}{2} & \text{if } 0 < t < 2\pi, \\ 0 & \text{if } t = 0, \end{cases} \quad S(t+2\pi) = S(t) \quad (12.4-2)$$

can be given either on the basis of 12.3.2 or of 12.3.5 at any point, and also at their points of discontinuity $t = k\pi$ ($k = 0, \pm 1, \dots$) and $t = 2k\pi$ ($k = 0, \pm 1, \dots$), respectively, as the sums of the associated ordinary Fourier series. Calculating the corresponding Fourier constants, we get for $-\infty < x < \infty$

¹ See Lebesgue [5].

$$C(x) = \sin x + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \dots \quad (12.4-3)$$

and

$$S(x) = \sum_{n=1}^{\infty} \frac{1}{n} \sin nx. \quad (12.4-4)$$

(Cf. Fig. 9.)

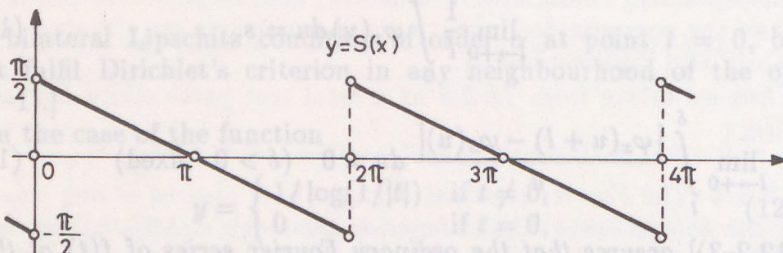


Fig. 9

12.42 Otherwise, in any closed interval of continuity of the functions $C(t)$ or $S(t)$ both expansions are valid *uniformly* too; furthermore their partial sums are *uniformly bounded* for all x , which can be verified e.g. in case of (12.4-4) as follows:

Let $0 < x \leq \pi$, $N \geq 2$, and consider the decomposition

$$\sum_{n=1}^N \frac{\sin nx}{n} = \sum_{n=1}^{\nu-1} \frac{\sin nx}{n} + \sum_{n=\nu}^N \frac{\sin nx}{n},$$

where ν denotes the greatest integer for which $(\nu - 1)x \leq \pi$, so that $\nu = \left[\frac{\pi}{x}\right] + 1$. On the one hand, by $0 \leq \sin nx \leq nx$ we have

$$\sum_{n=1}^{\nu-1} \frac{\sin nx}{n} < (\nu - 1)x \leq \pi.$$

On the other hand, the diagram of $\sin x$ being concave (from below) in $[0, \frac{\pi}{2}]$, the inequality $\sin t \geq \frac{2}{\pi}t$ holds there, consequently we can write:

$$\begin{aligned} \sum_{n=\nu}^N \frac{\sin nx}{n} &= \left(2 \sin \frac{x}{2}\right)^{-1} \sum_{n=\nu}^N \frac{1}{n} \left\{ \cos \left(n - \frac{1}{2}\right)x - \cos \left(n + \frac{1}{2}\right)x \right\} = \\ &= \left(2 \sin \frac{x}{2}\right)^{-1} \left\{ \frac{1}{\nu} \cos \left(\nu - \frac{1}{2}\right)x - \frac{1}{N} \cos \left(N + \frac{1}{2}\right)x + \right. \\ &\quad \left. + \sum_{n=\nu}^{N-1} \left(\frac{1}{n+1} - \frac{1}{n}\right) \cos \left(n + \frac{1}{2}\right)x \right\}, \end{aligned}$$

$$\left| \sum_{n=\nu}^N \frac{\sin nx}{n} \right| \leq \left(2 \sin \frac{x}{2} \right)^{-1} \left\{ \frac{1}{\nu} + \frac{1}{N} + \sum_{n=\nu}^{N-1} \left(\frac{1}{n} - \frac{1}{n+1} \right) \right\} = \\ = \left(\sin \frac{x}{2} \right)^{-1} \cdot \frac{1}{\nu} \leq \frac{\pi}{x} \cdot \frac{x}{\pi} = 1.$$

As a final result, we get by addition [cf. (12.3-1)]:

$$\left| \sum_{n=1}^N \frac{\sin nx}{n} \right| \leq \pi + 1; \quad (12.4-5)$$

and this is valid at the same time for any real x as well as for any positive integer N , since $\sin nx$ is an odd function with period 2π .¹

12.43 Differentiability is a *sufficient* condition of the Fourier expansibility according to 12.3.4, but — as is easy to see — it means a too strong condition of convergence, inadequate to the subject. There exist *nowhere differentiable* functions, which are *Fourier expansible everywhere*. Pertinent examples are yielded by the *Weierstrass functions* (3.7-1):

$$W_k(x) = \sum_{n=0}^{\infty} q^n \cos((2k+1)^n \pi x) \quad (k = 1, 2, \dots),$$

where $0 \leq q \leq 1$ and $(2k+1)q > 1 + \frac{3}{2}\pi$; the right-hand series being uniformly convergent for every x is necessarily the ordinary Fourier series of its sum.

The continuity everywhere of $W_k(x)$, and the circumstance that the Lipschitz condition (12.3-7) quasi approaches the concept of continuity as $\alpha \rightarrow +0$, suggest that maybe pointwise continuity should imply the Fourier expansibility as well, in view of the fact that ordinary Fourier series of any function occurring in practice converges actually at every point of continuity to the pertaining function value.

12.44 P. DU BOIS-REYMOND was the first mathematician who elucidated in negative sense the above problem of theoretical importance, by proving the *existence of a function continuous everywhere whose ordinary Fourier series is divergent at some point* (1873). Much simpler relevant counterexamples are due to L. FEJÉR (1909); these were based upon certain uniformly bounded cosine polynomials.²

¹ We remark that $\sin x + \frac{\sin 2x}{2} + \dots + \frac{\sin Nx}{N}$ is positive in $(0, \pi)$ and negative in $(-\pi, 0)$. Estimations like (12.4-5) hold for the partial sums of any sine or cosine series with monotone decreasing coefficients.

² This phenomenon is called in the literature *Du Bois-Reymond's singularity*. Cf. Du Bois-Reymond [1] and Fejér [4], [6], respectively.

Here an example of H. A. SCHWARZ is presented by which the function with required properties is given not as the sum of a trigonometric series but directly — so to say “diagrammatically.”¹

Let $g(t)$ be an even function with period 2π which is defined in $[0, \pi]$ in the following manner:

$$g(t) = \begin{cases} k^{-2} \sin(3^{k^4} t) & \text{if } 3^{-k^4} \pi \leq t \leq 3^{-(k-1)^4} \pi \quad (k = 1, 2, \dots) \\ 0 & \text{if } t = 0. \end{cases} \quad (12.4-6)$$

Accordingly, the diagram of $g(t)$ consists of sine waves with rapidly increasing frequencies and relatively slowly decreasing amplitudes in the intervals $[\pi/3, \pi]$, $[\pi/3^{2^4}, \pi/3^{1^4}]$, It is evident that $g(t)$ is continuous everywhere and e.g. by the criterion of Lipschitz, it is Fourier expansible at every point $t \neq 0$ of $[-\pi, \pi]$.

We assert:

12.4.1 The ordinary Fourier series of the continuous function (12.4-6) is divergent at the origin.

P r o o f. The basic integral (12.2-9) becomes in our case [$x = 0$, $\delta = \pi/6$, $\psi_0(t) = g(t)$]:

$$\int_0^{\pi/6} \frac{g(2u)}{u} \sin(2N+1)u \, du = \int_0^{\pi/3} \frac{g(t)}{t} \sin\left(N + \frac{1}{2}\right)t \, dt;$$

and the deviation of this integral type from the simpler one:

$$\int_0^{\pi/3} \frac{g(t)}{t} \sin Nt \, dt \quad (N = 1, 2, \dots) \quad (12.4-7)$$

tends to 0 as $N \rightarrow \infty$ by 12.1.3.

Thus it is sufficient to show that from the sequence of integrals (12.4-7) we can always select an unbounded subsequence. For that purpose, put $N = 3^{r^4}$ ($r = 2, 3, \dots$) and study the behaviour of the integral

$$\begin{aligned} J_r &= \int_0^{\pi/3} \frac{g(t)}{t} \sin(3^{r^4} t) \, dt = \\ &= \sum_{k=2}^{\infty} \frac{1}{k^2} \int_{\pi/3^{k^4}}^{\pi/3^{(k-1)^4}} t^{-1} \sin(3^{k^4} t) \sin(3^{r^4} t) \, dt = \sum_{k=2}^{\infty} \frac{1}{k^2} I_{kr} \end{aligned} \quad (12.4-8)$$

at the limiting process $r \rightarrow \infty$.

¹ See Hobson [1], vol. II, p. 545. The basic idea of the example has been published by H. A. SCHWARZ through the dissertation of one of his pupils; cf. Sachse [1], 271-274.

Let us write for $k \neq r$:

$$2I_{kr} = \int_{\pi/3^{k^4}}^{\pi/3^{(k-1)^4}} t^{-1} \cos\left(\left|3^{k^4} - 3^{r^4}\right|t\right) dt - \int_{\pi/3^{k^4}}^{\pi/3^{(k-1)^4}} t^{-1} \cos\left[\left(3^{k^4} + 3^{r^4}\right)t\right] dt,$$

then substituting $t = 3^{-k^4}u$ and $t = 3^{-r^4}u$ into both of the last integrals we obtain:

$$\int_{\pi}^{m\pi} \frac{\cos(cu)}{u} du \quad (12.4-9)$$

where m is a positive integer and $c > 2/3$. Hence, applying integration by parts and using the inequality (12.3-1), we get:

$$|I_{kr}| < \frac{3}{\pi} \quad (k \neq r). \quad (12.4-10)$$

When $k = r$, we have:

$$\begin{aligned} I_{rr} &= \frac{1}{2} \int_{\pi/3^{r^4}}^{\pi/3^{(r-1)^4}} t^{-1} \left[1 - \cos\left(3^{r^4} \cdot 2t\right)\right] dt = \\ &= \left[r^4 - (r-1)^4\right] \frac{\log 3}{2} - \frac{1}{2} \int_{2\pi}^{2\pi \cdot 3^{r^4} - (r-1)^4} \frac{\cos u}{u} du. \end{aligned}$$

Since integration by parts yields the elementary inequality:

$$\left| \int_{2\pi}^x \frac{\cos u}{u} du \right| < \frac{1}{x} + \int_{2\pi}^x \frac{1}{u^2} du < \frac{1}{\pi} \quad (x > 2\pi),$$

it follows:

$$I_{rr} > r^2(2r-1) \frac{\log 3}{2} - \frac{1}{2\pi} > 0 \quad (r = 2, 3, \dots). \quad (12.4-11)$$

Finally, from the decomposition

$$J_r = \frac{1}{r^2} I_{rr} + \sum_{\substack{k \geq 2 \\ k \neq r}} \frac{1}{k^2} I_{kr}$$

we can see that the first term on the right-hand side tends to infinity by (12.4-11), while the second term remains bounded by (12.4-10) and $\sum k^{-2} < \infty$. Consequently, $\lim_{r \rightarrow \infty} J_r = \infty$, which was to be verified.

* * *

12.45 The function $g(t)$ just discussed or other pertinent examples enable us to construct easily — by linear transformation and by use of the so-called superposition principle — such an everywhere continuous function, whose ordinary Fourier series diverges on a countably infinite set, or on a null set having the continuum power. A. N. KOLMOGOROV found in 1926 an example of (L)-integrable function, whose ordinary Fourier series is *divergent everywhere*; however, this function is a quite *discontinuous* one.¹

It was uncertain until recent times, whether the ordinary Fourier series of a continuous or at least square integrable function may diverge everywhere. (Namely, Kolmogorov's example does not belong to the class L^2 .) In 1966, however, L. CARLESON gave in negative sense a complete solution of the problem (utilizing very deep considerations) by proving an old conjecture of LUSIN:

12.4.2 *The ordinary Fourier series of any square integrable function is convergent almost everywhere.*²

This implies, that the divergence points of ordinary Fourier series of any continuous function form a null set as well.

There is another (indirect) way to deduce the result 12.4.1: application of the Lebesgue functions of (10.1-1), i.e., of the so-called *Lebesgue constants*:

$$\lambda_N = \frac{1}{2\pi} \int_0^{2\pi} |D_N(t)| dt = \frac{2}{\pi} \int_0^{\pi/2} \left| \frac{\sin(2N+1)u}{\sin u} \right| du$$

($N = 0, 1, 2, \dots$); (12.4-12)

more precisely, of the asymptotic formula

$$\lambda_N \simeq \frac{4}{\pi^2} \log N \quad (N \rightarrow \infty). \quad (12.4-13)$$

¹ Cf. Kolmogorov [1]–[2], Zeller [1]. It is much easier to give such a *trigonometric series diverging everywhere* whose coefficients tend to zero but the series is not an ordinary Fourier series. Example: $\sum_{n=2}^{\infty} (\log k)^{-1} \cos k(x - \log \log k)$. See Steinhaus [1], Zygmund [3].

² The pertinent results of Carleson [1] have been developed further by Hunt [1], showing that Lusin's conjecture holds also with an extended premise. (Cf. Lusin [2], Kahane–Katznelson [1].)

[Cf. (11.6-5).] This method has the advantage that it may be applied to Fourier series of non-trigonometric orthogonal systems, too.¹

12.5 The problem of uniqueness; term-by-term integrability of ordinary Fourier series

12.51 With respect to the connection of the general concept of Fourier series with the so-called decomposition problem (cf. 9.6), it is reasonable to deal briefly also with specifying the relationship between the coefficients and the sum function of convergent trigonometric series.

(I) If for the sum $f(x)$ of a trigonometric series

$$A_0 + \sum_{n=1}^{\infty} (A_n \cos nx + B_n \sin nx), \quad (12.5-1)$$

convergent at each point of an interval $(a, a + 2\pi)$, the "minimal" requirement is made that it should be integrable on the line segment $(a, a + 2\pi)$, whether hence we may conclude that the coefficients of (12.5-1) are the Fourier constants in (12.1-2) ?

We have to consider separately the particular case where $f(x)$ vanishes identically:

(II) If the sum of a series of the form (12.5-1) equals 0 everywhere in an interval of length 2π , will this imply the vanishing of all the coefficients A_n, B_n ?

The so-called *problem of uniqueness* may immediately be reduced to the latter:

(III) If the sum functions of two trigonometric series are identical at every point of a segment of length 2π , whether does this involve the equality by pairs of corresponding coefficients, i.e., that the series in question are identical with each other ? (In other words: the sum of a trigonometric series convergent everywhere does determine *uniquely* the coefficients of the series?)

12.52 The answer to any of these questions is affirmative, as it has been shown in case of (II) and (III) by G. CANTOR, in case of (I) by P. DU BOIS-REYMOND (for Riemann integrable functions), further by CH. J. DE LA VALLÉE-POUSSIN (for Lebesgue integrable functions), respectively.²

The most relevant result of the theory of uniqueness is [cf. (I)]:

¹ See Fejér [5], Szegő [1]; moreover Hardy [6].

² Cf. Cantor [1], Du Bois-Reymond [2] and De la Vallée-Poussin [1].

12.5.1 If a trigonometric series converges on a segment of length 2π to an (L) -integrable function f , then it is necessarily the ordinary Fourier series of f .¹

12.5.2 (CANTOR-LEBESGUE THEOREM)² In order that a trigonometric series (12.5-1) should converge on a point set E of positive (L) -measure, it is necessary that $A_n \rightarrow 0$, $B_n \rightarrow 0$ as $n \rightarrow \infty$.

12.5.3 (CANTOR'S THEOREM OF UNIQUENESS) In order that all coefficients of a series (12.5-1), convergent everywhere, should be 0, the vanishing of its sum function at all but countably many points is sufficient.

From the last proposition we see that to the conclusion $A_n = 0$, $B_n = 0$ (for every n), the vanishing of the sum for all x is not necessary: the existence of an exceptional set s (where the sum function may be $\neq 0$) is permitted. In general, a linear set $s \subset (0, 2\pi)$ is called a *set of uniqueness* of (12.5-1), if the convergence of this series to 0 for every $x \in (0, 2\pi) - s$ implies the vanishing of all of its coefficients. Such sets can be of quite complicated nature (e.g. the triadic set of Cantor — cf. 2.4.5 — is also a set of uniqueness), and their full characterization has not been achieved yet. G. Cantor discovered the theory of sets through the investigation of sets of uniqueness in the 1870's.

The most essential common tool of deep research about problems of uniqueness is a method found by RIEMANN (1854). The basic idea of Riemann's theory is as follows: the convergent trigonometric series is integrated twice formally term-by-term, and from the everywhere continuous sum function thus obtained the properties of the original series can be inferred — by applying a suitable generalized derivation procedure.³

12.53 The problem of termwise integration of trigonometric series is closely related to the foregoing. We ask: if a series of the form (12.5-1) converges on a segment of length 2π to some (L) -integrable function f , then its term-by-term integration on an arbitrary subinterval (x_1, x_2) does always yield a series converging to $\int_{x_1}^{x_2} f(t) dt$?

Theorem 12.5.1 shows that the series to be integrated may be taken a priori as a Fourier series — without loss of generality. In our case, the result

¹ Example of a trigonometric series which is though convergent almost everywhere, but its sum function is not (L) -integrable, so that the series is no Fourier series:

$$\sum_{n=2}^{\infty} (\log n)^{-1} \sin nx.$$

² Cf. Lebesgue [2]. Since the premise of 12.5.2 implies $\lim_{n \rightarrow \infty} (A_n \cos nx + B_n \sin nx) = 0$ ($x \in E$), it is enough to show that the latter condition involves $A_n \rightarrow 0$, $B_n \rightarrow 0$.

³ See Riemann [1], Zygmund [1].

will be more in a sense than it could be expected heuristically: it turns out that *any ordinary Fourier series may be integrated term-by-term*, i.e., the formal integration of some divergent series of the form (12.1-1)—(12.1-2) yields also convergent series. (This is possible, since integrating the generic term a multiplier $1/n$ is applied, by which the convergence behaviour of the series is improved.)

12.5.4 ¹If $f(t) \sim \alpha_0 + \sum_{n=1}^{\infty} (\alpha_n \cos nt + \beta_n \sin t)$ [cf. (11.1-1)], then for any $[x_1, x_2] \subset (0, 2\pi)$ the integration formula holds:

$$\int_{x_1}^{x_2} f(t) dt = \sum_{n=1}^{\infty} \left[\frac{\alpha_n - 2\alpha_0}{n} (\sin nx_2 - \sin nx_1) - \frac{\beta_n}{n} (\cos nx_2 - \cos nx_1) \right]. \quad (12.5-2)$$

For arbitrary x_1 and x_2 :

$$\int_{x_1}^{x_2} f(t) dt = \alpha_0(x_2 - x_1) + \sum_{n=1}^{\infty} \left[\frac{\alpha_n}{n} (\sin nx_2 - \sin nx_1) - \frac{\beta_n}{n} (\cos nx_2 - \cos nx_1) \right]. \quad (12.5-3)$$

At any point $x \in [0, 2\pi]$:

$$\int_0^x f(t) dt = (\alpha_0 x + \alpha_0^*) + \sum_{n=1}^{\infty} \frac{1}{n} (\alpha_n \sin nx - \beta_n \cos nx), \quad (12.5-4)$$

where

$$\alpha_0^* = \frac{1}{2\pi} \int_0^{2\pi} f(t)(\pi - t) dt. \quad (12.5-5)$$

P r o o f. The ordinary Fourier constants of the integral function $F(x) = \int_0^x f(t) dt$ can be connected through integration by parts with the Fourier constants of $f(t)$ [cf. (7.6-22)]:

¹ Cf. W. H. Young [4], Hardy [5].

$$A_0 = \frac{1}{2\pi} \int_0^{2\pi} F(t) dt = \frac{1}{2\pi} \left\{ [F(t)(t - \pi)]_0^{2\pi} - \int_0^{2\pi} f(t)(t - \pi) dt \right\} = \\ = \pi\alpha_0 + \frac{1}{2\pi} \int_0^{2\pi} f(t)(\pi - t) dt,$$

$$A_n = \frac{1}{\pi} \int_0^{2\pi} F(t) \cos nt dt = \frac{1}{\pi} \left\{ \left[F(t) \frac{\sin nt}{n} \right]_0^{2\pi} - \frac{1}{n} \int_0^{2\pi} f(t) \sin nt dt \right\} = \\ = -\frac{1}{n} \beta_n \quad (n = 1, 2, \dots),$$

$$B_n = \frac{1}{n} \int_0^{2\pi} F(t) \sin nt dt = \frac{1}{\pi} \left\{ \left[F(t) \frac{-\cos nt}{n} \right]_0^{2\pi} + \frac{1}{n} \int_0^{2\pi} f(t) \cos nt dt \right\} = \\ = \frac{1}{n} \alpha_n - \frac{2}{n} \alpha_0 \quad (n = 1, 2, \dots).$$

Since $F(t)$ is absolutely continuous and thus also of bounded variation (cf. 7.6.4), this integral function is Fourier expansible in $(0, 2\pi)$ on the basis of criteria 12.3.2—12.3.5. Consequently:

$$F(x) = A_0 + \sum_{n=1}^{\infty} (A_n \cos nx + B_n \sin nx) = \\ = \pi\alpha_0 + \frac{1}{2\pi} \int_0^{2\pi} f(t)(\pi - t) dt + \sum_{n=1}^{\infty} \left(-\frac{\beta}{n} \cos nx + \frac{\alpha_n - 2\alpha_0}{n} \sin nx \right) \\ (0 < x < 2\pi). \quad (12.5-6)$$

If in (12.5-6) x is replaced by x_2 then by x_1 , so it results by subtraction in the formula (12.5-2); and for $x_1 \neq 0$, $x_2 \neq 2\pi$, the latter turns into (12.5-3) by using (12.4-4).

The extension of (12.5-3)—(12.5-5) to the end points of $[0, 2\pi]$, i.e., the verification of the formula:

$$\sum_{n=1}^{\infty} \frac{\beta_n}{n} = \frac{1}{2\pi} \int_0^{2\pi} f(t)(\pi - t) dt \quad (12.5-7)$$

is still to be done. Introducing to the right-hand side the series $\sum_{n=1}^{\infty} \frac{\sin nx}{n}$ [cf. (12.4-4)], instead of $(\pi - t)/2$, using 7.5.1 and (12.4-5), we may integrate termwise. Hence (12.5-7) follows.

* * *

12.54 The *termwise differentiation* of ordinary Fourier series — unlike the case of power series — is *not permissible in general*, since derivation increases the order of magnitude of Fourier constants. A pertinent trivial example is the formal term-by-term differentiation of the series $\sum_{n=1}^{\infty} n^{-1} \sin nx$ (which converges for all x) yielding the totally divergent series $\sum_{n=1}^{\infty} \cos nx$.¹

Nevertheless, it is easy to obtain a (rather strong) sufficient condition for the termwise derivability of Fourier series by applying 12.5.4 to a suitable integral function:

12.5.5 If $F(x) \in L(0, 2\pi)$, $F(0) = F(2\pi) = 0$ and F can be written as the (L) -integral function of such a continuous function f which is Fourier expansible everywhere, then the ordinary Fourier series of F converges to this function and may be differentiated term-by-term at every point $x \in [0, 2\pi]$. That is, in our case:

$$F(x) = A_0 + \sum_{n=1}^{\infty} (A_n \cos nx + B_n \sin nx) \quad (0 \leq x \leq 2\pi), \quad (12.5-8)$$

$$F'(x) = f(x) = \sum_{n=1}^{\infty} (-nA_n \sin nx + nB_n \cos nx) \quad (0 \leq x \leq 2\pi) \quad (12.5-9)$$

A_n and B_n denoting the ordinary Fourier constants of F .

12.6 The conjugate series

12.61 It has been mentioned before (cf. 10.22) that if we consider an arbitrary trigonometric series of the form (12.5-1) as the real part of a complex power series, then another trigonometric series also occurs, namely

$$\sum_{n=1}^{\infty} (A_n \sin nx - B_n \cos nx), \quad (12.6-1)$$

¹ Cf. Fejér [2].

as imaginary part of the same power series. Substituting in (12.6-1) A_n and B_n for the ordinary Fourier constants (12.1-2) of a function $f(t)$, we get

$$\sum_{n=1}^{\infty} (\alpha_n \sin nx - \beta_n \cos nx), \quad (12.6-2)$$

whose name is the *conjugate series* of the Fourier series (12.1-1). The investigation of the convergence of (12.6-2) also requires examination of the partial sums, like the treatment of (12.1-1).

Thus we express first the N -th partial sum of (12.6-2) in closed (integral) form:

$$\begin{aligned} \tilde{s}_N(x) &= \sum_{n=1}^N (\alpha_n \sin nx - \beta_n \cos nx) = \\ &= \frac{1}{\pi} \sum_{n=1}^N \left[\left(\int_{-\pi}^{\pi} f(t) \cos nt \, dt \right) \sin nx - \left(\int_{-\pi}^{\pi} f(t) \sin nt \, dt \right) \cos nx \right] = \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cdot \sum_{n=1}^N (\sin nx \cos nt - \cos nx \sin nt) \, dt = \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sum_{n=1}^N \sin n(x-t) \, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \tilde{D}_N(x-t) \, dt, \end{aligned}$$

where

$$\begin{aligned} \tilde{D}_N(u) &= 2 \sum_{n=1}^N \sin nu = \cot \frac{u}{2} - \frac{\cos(2N+1)\frac{u}{2}}{\sin \frac{u}{2}} \\ &\quad (u \neq 0, \pm 2\pi, \pm 4\pi, \dots) \end{aligned} \quad (12.6-3)$$

is the so-called N -th *conjugate Dirichlet kernel function*.

Hence by the substitution $t = x - u$ and applying transformations analogous to those in (12.2-2), we obtain:

$$\begin{aligned} \tilde{s}_N(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-u) \tilde{D}_N(u) \, du = \\ &= \frac{1}{2\pi} \int_0^{\pi} [f(x-u) - f(x+u)] \tilde{D}_N(u) \, du. \end{aligned} \quad (12.6-4)$$

This is usually called *conjugate Dirichlet's integral*.

12.62 Now, we can easily deduce the counterparts of condition 12.3.1 for the conjugate series:

12.6.1 (PRINGSHEIM'S CRITERION)¹ If for some fixed x , the quotient $[f(x+u) - f(x-u)]/u$ (as a function of u) is (L) -integrable in a right-hand neighbourhood of the point $u = 0$, then the series (12.6-2) is convergent for this x and has there the sum:

$$\tilde{f}(x) = \frac{1}{2\pi} \int_0^\pi [f(x-u) - f(x+u)] \cot \frac{u}{2} du. \quad (12.6-5)$$

P r o o f. The above premise implies immediately [cf. (12.2-10)], that $[f(x+u) - f(x-u)]/\sin \frac{u}{2}$ is also (L) -integrable in the interval $0 < u \leq \pi$. Therefore by (12.6-3)—(12.6-4), the integral (12.6-5) exists as well, and we have for every N the decomposition:

$$\tilde{s}_N(x) = \tilde{f}(x) - \frac{1}{2\pi} [f(x-u) - f(x+u)] \left(\sin \frac{u}{2}\right)^{-1} \cos(2N+1)\frac{u}{2} du.$$

But the last integral tends to 0 as $N \rightarrow \infty$ according to the Riemann-Lebesgue lemma 12.1.3, so we get the limit relation

$$\lim_{N \rightarrow \infty} \tilde{s}_N(x) = \tilde{f}(x). \quad (12.6-6)$$

* * *

12.63 As can be seen from the above, the “natural” sum of the series (12.6-2) is not $f(x)$ but $\tilde{f}(x)$, provided that the integral (12.6-5) exists.

Anyway it is obvious that $\tilde{f}(x)$ does not exist at such points x , where the limit

$$\delta_x = \lim_{u \rightarrow +0} [f(x+u) - f(x-u)] \quad (12.6-7)$$

exists and $\delta_x \neq 0$ (i.e., where the function f has a jump); then, since $\lim_{u \rightarrow +0} \frac{u}{2} \cot \frac{u}{2} = 1$, for u sufficiently small the inequality

$$|f(x+u) - f(x-u)| \cot \frac{u}{2} > \frac{1}{2} |\delta_x| u^{-1}$$

holds. Accordingly, we may expect that at any point of discontinuity, $\tilde{s}_N(x)$ tends to infinity as $N \rightarrow \infty$, i.e., the conjugate series is *divergent*.

¹ Pringsheim [1].

F. LUKÁCS proved in 1918 not only the latter statement, but also succeeded in characterizing the asymptotic behaviour of $\tilde{s}_N(x)$ around the jumping points of f . His theorem establishes that

$$\tilde{s}_N(x) \simeq -\frac{1}{\pi} \delta_x \log N \quad (N \rightarrow \infty) \quad (12.6-8)$$

at any point x where δ_x exists; consequently, the total jump (cf. 4.11) pertaining to such a point can be written in the form:¹

$$|\delta_x| = \lim_{N \rightarrow \infty} \frac{\pi |\tilde{s}_N(x)|}{\log N}. \quad (12.6-9)$$

We remark that by theorem 12.5.1, the conjugate series (12.6-2) may be considered as the ordinary Fourier series of $\tilde{f}(x)$ in the case, when $\tilde{f}(x)$ exists and is (L) -integrable in an interval of length 2π .

12.7 Ordinary Fourier series of some special functions and harmonic analysis

12.71 Let us regard some important examples for the utilization of the foregoing general results.

As it is well-known, the so-called *fractional part of x* , denoted by $\{x\}$, belongs to the periodic functions occurring most often in practice. Its definition is: $\{x\} = x - [x]$ where $[x]$ ("*integral part of x* ") means the greatest integer less than or equal to x . Thus we have for every real x : $0 \leq \{x\} < 1$ (with equality only if x is an integer), furthermore the *primitive period* of $\{x\}$ is 1. The function $\{x\}$ has jumps at points $x = 0, \pm 1, \pm 2, \dots$, while between any two integers it is represented by a straight line segment with slope 1; so the diagram of the function looks like saw-teeth (Fig. 10).

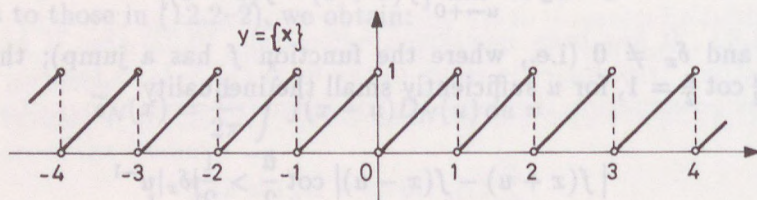


Fig. 10

¹ Lukács [1]; cf. also Fejér [7]-[8].

$[0, 1]$ being taken as the fundamental interval, the ordinary Fourier constants of $\{x\}$ are, according to the formulae (12.1-4) ($a = 0$, $p = 1$):

$$a_0 = \int_0^1 u \, du = \frac{1}{2},$$

$$\begin{cases} a_n = 2 \int_0^1 u \cos 2n\pi u \, du = 0 \\ b_n = 2 \int_0^1 u \sin 2n\pi u \, du = -\frac{1}{n\pi} \end{cases} \quad (n = 1, 2, \dots).$$

Considering that function $\{x\}$ is differentiable for every non-integral value of x , by 12.3.4 we obtain:¹

$$\{x\} = \frac{1}{2} - \frac{1}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin 2n\pi x \quad (x \neq 0, \pm 1, \pm 2, \dots). \quad (12.7-1)$$

In particular, inside the fundamental interval the expansion

$$x - \frac{1}{2} = - \sum_{n=1}^{\infty} \frac{\sin 2n\pi x}{n\pi} \quad (0 < x < 1), \quad (12.7-2)$$

holds, which follows also from (12.4-4) by substituting x for $2\pi x$.

12.72 If we integrate both sides of (12.7-2) repeatedly (cf. 12.5.4), we get important series representations for the so-called *Bernoulli polynomials*, defined successively by the formulae:²

$$B_0(x) \equiv 1, \quad (12.7-3)$$

$$\left. \begin{aligned} B'_r(x) &= B_{r-1}(x) \\ \int_0^1 B_r(x) \, dx &= 0 \end{aligned} \right\} \quad (r = 1, 2, \dots).^2 \quad (12.7-4)$$

¹ Since $x = 0, \pm 1, \pm 2, \dots$ are not regular points of $\{x\}$, so the function is not Fourier expandable at these points. (Cf. 12.32, 13.31.)

² The first few Bernoulli polynomials can easily be written in explicit form: $B_1(x) = x - \frac{1}{2}$, $B_2(x) = \frac{1}{2}x(x-1) + \frac{1}{12}$, $B_3(x) = \frac{1}{6}x(x - \frac{1}{2})(x-1)$, $B_4(x) = \frac{1}{24}x^2(x-1)^2 - \frac{1}{720}, \dots$

Since $B_1'(x) = 1$ and $\int_0^1 B_1(x) dx = 0$, the first term of (12.7-2) must be identical with $B_1(x)$; and $B_2'(x) = B_1(x)$ and $\int_0^1 B_2(x) dx = 0$, we see from (12.7-2) that

$$B_2(x) = \sum_{n=1}^{\infty} \frac{2 \cos 2n\pi x}{(2n\pi)^2} \quad (0 < x < 1). \quad (12.7-5)$$

Hence by integrating again, observing that by (12.7-4) the occurring constant of integration vanishes:

$$B_3(x) = \sum_{n=1}^{\infty} \frac{2 \sin 2n\pi x}{(2n\pi)^3} \quad (0 < x < 1); \quad (12.7-6)$$

and so on.

We deduce from (12.7-2) by $(r-1)$ -fold integration, on the basis of conditions (12.7-3)—(12.7-4) the formula:

$$B_r(x) = - \sum_{n=1}^{\infty} \frac{2}{(2n\pi)^r} \cos \left(2n\pi x - \frac{\pi r}{2} \right) \quad (0 < x < 1; r = 1, 2, \dots); \quad (12.7-7)$$

and this may be extended for arbitrary non-integral real values of x , if $B_r(x)$ is replaced by the corresponding function with period 1, i.e., by $B_r(\{x\})$.

The Bernoulli polynomials and the pertaining so-called *Bernoulli numbers*:

$$B_r = r! B_r(0) \quad (r = 0, 1, 2, \dots) \quad (12.7-8)$$

have many applications in analysis; they were introduced in a work of JAKOB BERNOULLI (1713).

The usefulness of (12.7-7) can already be shown by putting in it $x = 0$ for an arbitrary even index r . This yields the sum of the "hyperharmonic" series in question, namely¹

$$\sum_{n=1}^{\infty} \frac{1}{n^{2\nu}} = (-1)^{\nu-1} \frac{B_{2\nu}}{2(2\nu)!} (2\pi)^{2\nu} \quad (\nu = 1, 2, \dots).^1 \quad (12.7-9)$$

For $\nu = 1$ and $\nu = 2$, by $B_2 = 1/6$, $B_4 = -1/30$ we obtain:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}, \quad \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}. \quad (12.7-10)$$

¹ Formulae (12.7-9), (12.7-10) were discovered by L. EULER in 1736 (in another way). Note that the summation of the series $\sum_{n=1}^{\infty} 1/n^{2\nu+1}$ ($\nu = 1, 2, \dots$), i.e., the case of odd indices r is a problem of quite different nature. Cf. Mikolás [1].

From the more general application possibilities of Bernoulli's polynomials and numbers, we mention the so-called *Euler-Maclaurin sum formula*:

$$\begin{aligned} \sum_{m=0}^{M-1} f(m) &= \\ &= \int_0^M f(x) dx + \sum_{r=1}^N \frac{B_r}{r!} \left(f^{(r-1)}(M) - f^{(r-1)}(0) \right) + \\ &+ (-1)^{N-1} \int_0^M B_r(\{x\}) f^{(N)}(x) dx \quad (N = 1, 2, \dots) \end{aligned} \quad (12.7-11)$$

as well as *Poisson's summation formula* to be discussed later [cf. (14.5-24)]. These hold for arbitrary functions f sufficiently smooth, and are indispensable in deeper investigations on asymptotics of certain series.

The Fourier series representations (12.7-7) plainly enable us to determine also the sums of the corresponding *conjugate series* in closed integral form. (Cf. 12.61-12.62.) Specially for $r = 1$, by separating the real and imaginary parts in the complex logarithmic series.¹

$$\sum_{n=1}^{\infty} \frac{z^n}{n} = \text{Log} \frac{1}{1-z} \quad (|z| \leq 1, z \neq 1), \quad (12.7-12)$$

and using also (12.7-2) we get the formula

$$\sum_{n=1}^{\infty} \frac{\cos 2n\pi x}{n} = -\log(2 \sin \pi x) \quad (0 < x < 1). \quad (12.7-13)$$

12.73 In the next chapter, we shall need a.o. the so-called *Hurwitz zeta-function* $\zeta(\Theta, x)$, which is closely related to the system of Bernoulli polynomials and is important from several points of view, too.

The definition of $\zeta(\Theta, x)$ for $\Theta > 0$, $\Theta \neq 1$ and $0 < x < 1$ can be written as follows:

$$\zeta(\Theta, x) = \lim_{M \rightarrow \infty} \left[\sum_{m=0}^M (x+m)^{-\Theta} - \frac{M^{1-\Theta}}{1-\Theta} \right]. \quad (12.7-14)$$

¹ Here "Log" signifies the principal value of the complex logarithm.

Thus for $\Theta > 1$ it is simply

$$\zeta(\Theta, x) = \sum_{m=0}^{\infty} (x+m)^{-\Theta}, \quad (12.7-15)$$

while for $0 < \Theta < 1$ we have also

$$\zeta(\Theta, x) = \sum_{m=0}^{\infty} \left[(x+m)^{-\Theta} - \frac{(m+1)^{1-\Theta} - m^{1-\Theta}}{1-\Theta} \right]. \quad (12.7-16)$$

Since for any fixed $\Theta > 0$, $\Theta \neq 1$ (12.7-14) is obviously equivalent to the representation

$$\zeta(\Theta, x) - x^{-\Theta} = \frac{1}{\Theta-1} + \sum_{m=1}^{\infty} \left[(x+m)^{-\Theta} - \frac{(m+1)^{1-\Theta} - m^{1-\Theta}}{1-\Theta} \right], \quad (12.7-17)$$

and the last series is uniformly convergent in $0 < x \leq 1$ by

$$\begin{aligned} \left| (x+m)^{-\Theta} - \frac{(m+1)^{1-\Theta} - m^{1-\Theta}}{1-\Theta} \right| &= \left| (x+m)^{-\Theta} - \int_0^1 (t+m)^{-\Theta} dt \right| \leq \\ &\leq \int_0^1 \left| (x+m)^{-\Theta} - (t+m)^{-\Theta} \right| dt < m^{-\Theta} - (m+1)^{-\Theta} \\ &\quad (m=1, 2, \dots), \end{aligned} \quad (12.7-18)$$

so we can establish: $\zeta(\Theta, x)$ as a function of x is continuous and monotone increasing for $0 < x \leq 1$, further tends to infinity for $x \rightarrow +0$ in the order $x^{-\Theta}$.

As regards the variable Θ , it is easy to see that for any fixed x , $\zeta(\Theta, x)$ is continuous at any point $\Theta > 0$, $\Theta \neq 1$. Namely, if we fix $\Theta' > 0$ and $\Theta'' > \max\{1, \Theta'\}$, then it holds for every $m > 2\Theta''$:

$$\begin{aligned} m^{-\Theta} - (m+1)^{-\Theta} &= m^{-\Theta} \left[1 - \left(1 + \frac{1}{m} \right)^{-\Theta} \right] < \\ &< \frac{\Theta}{m^{\Theta+1}} \left(1 + \frac{\Theta+1}{2} \cdot \frac{1}{m} + \frac{\Theta+1}{2} \frac{\Theta+2}{3} \cdot \frac{1}{m^2} + \dots \right) < \\ &< \frac{\Theta''}{m^{\Theta'+1}} \left[1 + \frac{\Theta''}{m} + \left(\frac{\Theta''}{m} \right)^2 + \dots \right] < \frac{2\Theta''}{m^{\Theta'+1}}. \end{aligned}$$

It follows from (12.7-18) that the series (12.7-17) is uniformly convergent on the segment $\Theta' < \Theta < \Theta''$ too, consequently its sum is there a continuous function of Θ .

We have to stress that the particular case $\zeta(\Theta, 1) = \zeta(\Theta)$, is the so-called *zeta-function of Riemann* which is defined for $\Theta > 1$ as the sum of the series $\sum_{n=1}^{\infty} n^{-\Theta}$, and plays a prominent role in the *analytic theory of numbers*. In that field it is essential to consider Θ as a *complex* variable. The representations (12.7-14)—(12.7-17), further the continuity of $\zeta(\Theta, x)$ with respect to Θ remain valid for all complex values of Θ with appropriate real parts, besides $\zeta(\Theta, x)$ is an *analytic* function of Θ at any point $\Theta \neq 1$ of the complex plane.

Now, we restrict ourselves to the values $0 < \Theta < 1$. According to (12.7-18), the modulus of every partial sum of (12.7-16) is less than $x^{-\Theta} + (1 - \Theta)^{-1} + 1$, which $\in L(0, 1)$ as a function of x . So the same holds on $\zeta(\Theta, x)$ by 7.5.1. If we consider that function of x with the period 1 which is identical with $\zeta(\Theta, x)$ for $0 < x \leq 1$ [this may clearly be written in the form $\zeta(\Theta, \{x\})$], then the “periodic continuation” — as a piecewise continuous and piecewise monotonic function of x — is Fourier expansible for any non-integral x on the basis of 12.3.5. Let us calculate its ordinary Fourier constants pertaining to the fundamental interval (0,1) by means of the formulae (12.1-4) as well as of an easy consequence of (10.5-25):¹

$$\int_0^{\infty} t^{\vartheta-1} e^{-i\lambda t} dt = \Gamma(\vartheta)(\lambda i)^{-\vartheta} \quad (\lambda \neq 0 \text{ real}, \quad 0 < \vartheta < 1). \quad (12.7-19)$$

Using Lebesgue's theorem 7.5.1, the coefficients looked for can be evaluated as follows:

$$\begin{aligned} a_0 &= \int_0^1 \zeta(\Theta, u) du = \sum_{m=0}^{\infty} \int_0^1 \left[(u+m)^{-\Theta} - \frac{(m+1)^{1-\Theta} - m^{1-\Theta}}{1-\Theta} \right] du = 0; \\ a_n &= \int_0^1 \zeta(\Theta, u) \cos 2n\pi u du = \\ &= 2 \sum_{m=0}^{\infty} \int_0^1 \left[(u+m)^{-\Theta} - \frac{(m+1)^{1-\Theta} - m^{1-\Theta}}{1-\Theta} \right] \cos 2n\pi u du = \\ &= 2 \sum_{m=0}^{\infty} \int_m^{m+1} t^{-\Theta} \cos 2n\pi t dt = \int_0^{\infty} t^{-\Theta} (e^{2n\pi i t} + e^{-2n\pi i t}) dt = \end{aligned}$$

¹ Here and later any complex power is to be taken with its principal value.

$$\begin{aligned}
&= \Gamma(1-\Theta)[(-2n\pi i)^{\Theta-1} + (2n\pi i)^{\Theta-1}] = \\
&= \Gamma(1-\Theta)(2n\pi)^{\Theta-1} \cdot 2 \sin \frac{\pi\Theta}{2} \quad (n = 1, 2, \dots);
\end{aligned}$$

and similarly:

$$\begin{aligned}
b_n &= 2 \int_0^1 \zeta(\Theta, u) \sin 2n\pi u \, du = i \int_0^\infty t^{-\Theta} (e^{2n\pi i t} - e^{-2n\pi i t}) \, dt = \\
&= i\Gamma(1-\Theta)[(-2n\pi i)^{\Theta-1} - (2n\pi i)^{\Theta-1}] = \\
&= \Gamma(1-\Theta)(2n\pi)^{\Theta-1} \cdot 2 \cos \frac{\pi\Theta}{2} \quad (n = 1, 2, \dots).
\end{aligned}$$

Summing up, we obtain the validity for $0 < \Theta < 1$ and $0 < x < 1$ of the representation:

$$\zeta(\Theta, x) = 2\Gamma(1-\Theta) \sum_{n=1}^{\infty} (2n\pi)^{\Theta-1} \sin \left(2n\pi x + \frac{\pi\Theta}{2} \right), \quad (12.7-20)$$

which is usually called *Hurwitz's formula*.¹

Note that the right-hand series also converges in the case when the exponent $\Theta - 1$ is not greater than -1 , i.e., $\Theta \leq 0$. (Cf. 12.41 and 12.42.) Therefore, the connection (12.7-20) can be extended (by analytic continuation with respect to Θ) to every complex value of Θ whose real part is < 1 ; furthermore (12.7-20) also remains true for any non-integral real value of x with $\zeta(\Theta, \{x\})$ instead of $\zeta(\Theta, x)$ on the left. In the case $x = 1/2$, Hurwitz's formula furnishes the classical *functional equation of* $\zeta(\Theta)$, due to *Riemann* (1859).²

$$\zeta(\Theta) = 2(2\pi)^{\Theta-1} \Gamma(1-\Theta) \sin \frac{\pi\Theta}{2} \zeta(1-\Theta). \quad (12.7-21)$$

In fact, by the definitions we have for the Θ -values mentioned before:

$$\zeta(\Theta, 1/2) = (2^\Theta - 1) \zeta(\Theta) \quad \text{and} \quad \sum_{m=1}^{\infty} (-1)^m m^{-\Theta} = (2^{1-\Theta} - 1) \zeta(\Theta).$$

On the other hand, by the notation $\Theta = 1 - \vartheta$ (12.7-20) may be written as

$$\zeta(1 - \vartheta, x) = \Gamma(\vartheta) \sum_{n=1}^{\infty} \frac{2}{(2n\pi)^\vartheta} \cos \left(2n\pi x - \frac{\pi\vartheta}{2} \right); \quad (12.7-22)$$

¹ Cf. e.g. Whittaker-Watson [1], where (12.7-20) is deduced by complex function theoretical methods. The above simple verification has been given by Mikolás [5].

² See e.g. Titchmarsh [1], [3].

this holds for $\vartheta > 0$ and $0 < x < 1$. Since the last series is obviously a generalization of (12.7-7), the formula (12.7-22) becomes for $\vartheta = r$ ($r = 1, 2, \dots$), owing to $\Gamma(r) = (r-1)!$ [cf. (10.5-27)]:

$$\zeta(1-r, x) = -(r-1)! B_r(x) \quad (r = 1, 2, \dots). \quad (12.7-23)$$

Hence we see that Hurwitz's function $\zeta(\Theta, x)$ turns for $\Theta = 0, -1, -2, \dots$ substantially into $B_1(x), B_2(x), \dots$, i.e., the former may also be considered as a common *extension of the Bernoulli polynomials*.

12.74 Let v be a fixed *non-integral* real number and let $K_v(t)$ denote that real function with period 2π which is in $[-\pi, \pi]$ identical with $\cos(vt)$. This function is plainly even and continuous everywhere, furthermore it satisfies the conditions of theorem 12.1.2. Therefore $K_v(t)$ is Fourier expansible at any point $t = x$, and its ordinary Fourier series must be a cosine series (cf. 12.11), uniformly convergent on the whole number line (Fig. 11).

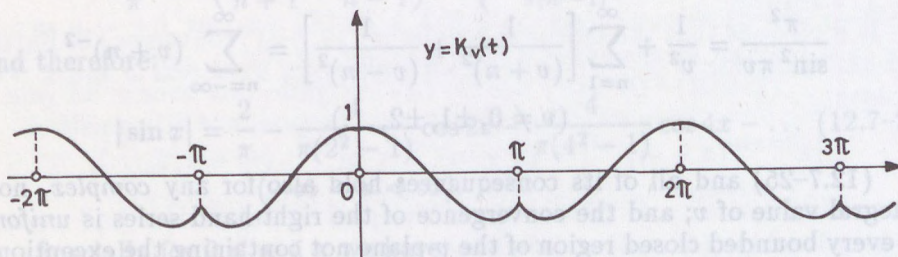


Fig. 11

The Fourier constants of $K_v(t)$ can also be calculated very easily [cf. (12.1-2)]:

$$\alpha_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos vt \, dt = \frac{1}{\pi} \left[\frac{\sin vt}{v} \right]_{t=0}^{t=\pi} = \frac{\sin \pi v}{\pi v},$$

$$\begin{aligned} \alpha_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos vt \cdot \cos nt \, dt = \frac{1}{\pi} \int_0^{\pi} [\cos(v+n)t + \cos(v-n)t] \, dt = \\ &= \frac{1}{\pi} (-1)^n \sin(\pi v) \left(\frac{1}{v+n} + \frac{1}{v-n} \right) \quad (n = 1, 2, \dots). \end{aligned}$$

Thus

$$K_v(x) = \frac{\sin \pi v}{\pi v} + \frac{\sin \pi v}{\pi} \sum_{n=1}^{\infty} (-1)^n \left(\frac{1}{v+n} + \frac{1}{v-n} \right) \cos nx$$

($-\infty < x < \infty$)

holds, and for the points of the fundamental interval it becomes:

$$\cos vx = \frac{\sin \pi v}{\pi v} + \frac{\sin \pi v}{\pi} \sum_{n=1}^{\infty} (-1)^n \left(\frac{1}{v+n} + \frac{1}{v-n} \right) \cos nx$$

($-\pi \leq x \leq \pi$). (12.7-24)

The main significance of (12.7-24) is that it can yield most simply an important expansion of the cotangent function. Namely, putting in (12.7-24) $x = \pi$ and after multiplying on both sides by $\pi/\sin(\pi v)$, we get:

$$\pi \cot \pi v = \frac{1}{v} + \sum_{n=1}^{\infty} \left(\frac{1}{v+n} + \frac{1}{v-n} \right) \quad (v \neq 0, \pm 1, \pm 2, \dots) \quad (12.7-25)$$

i.e., the *decomposition of cotangent into partial fractions*, the counterpart of the expression of a fractional rational function by means of its partial fractions.

From formula (12.7-25) we can derive by elementary transformations a number of related results, from which we want to mention here only the decomposition obtained by differentiation with respect to v :

$$\frac{\pi^2}{\sin^2 \pi v} = \frac{1}{v^2} + \sum_{n=1}^{\infty} \left[\frac{1}{(v+n)^2} + \frac{1}{(v-n)^2} \right] = \sum_{n=-\infty}^{\infty} (v+n)^{-2} \quad (v \neq 0, \pm 1, \pm 2, \dots). \quad (12.7-26)$$

(12.7-25) and all of its consequences hold also for any *complex*, non-integral value of v ; and the convergence of the right-hand series is *uniform* in every bounded closed region of the v -plane not containing the exceptional points $v = 0, \pm 1, \dots$

12.75 Like the above, an easy calculation leads to the ordinary Fourier expansion of $y = |\sin x|$. This may be regarded as the *even* function with period 2π which is identical with $\sin x$ in the fundamental interval $0 \leq x \leq \pi$.

Accordingly, $|\sin x|$ is continuous everywhere and besides differentiable apart from the points $x = k\pi$ ($k = 0, \pm 1, \pm 2, \dots$). (The latter are obviously break points of the function diagram.) Thus by 12.1.2, 12.3.4, the function in question is Fourier expansible — moreover uniformly — in $(-\infty, \infty)$; and its ordinary Fourier series is a cosine series owing to the evenness of $|\sin x|$, i.e. $\beta_n = 0$ for all n . (Cf. 12.11 and Fig. 12.)

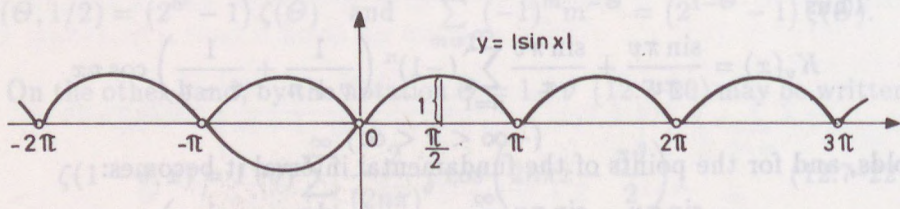


Fig. 12

As for the remaining Fourier constants, we can write:

$$\alpha_0 = \frac{1}{\pi} \int_0^{\pi} \sin t \, dt = \frac{2}{\pi},$$

$$\alpha_1 = \frac{2}{\pi} \int_0^{\pi} \sin t \cos t \, dt = \frac{1}{\pi} \int_0^{\pi} \sin 2t \, dt = 0,$$

$$\begin{aligned} \alpha_n &= \frac{2}{\pi} \int_0^{\pi} \sin t \cos nt \, dt = \frac{1}{\pi} \int_0^{\pi} [\sin(n+1)t - \sin(n-1)t] \, dt = \\ &= \frac{1 + (-1)^n}{\pi} \left(\frac{1}{n+1} - \frac{1}{n-1} \right) = \begin{cases} 0, & \text{if } n = 3, 5, 7, \dots, \\ -\frac{4}{\pi(n^2-1)}, & \text{if } n = 2, 4, 6, \dots, \end{cases} \end{aligned}$$

and therefore:

$$|\sin x| = \frac{2}{\pi} - \frac{4}{\pi(2^2-1)} \cos 2x - \frac{4}{\pi(4^2-1)} \cos 4x - \dots \quad (12.7-26)$$

$$(-\infty < x < \infty).$$

Specially for $0 \leq x \leq \pi$ we have:

$$\sin x = \frac{2}{\pi} - \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\cos 2nx}{(2n-1)(2n+1)}, \quad (12.7-27)$$

whence for $x = 0$ or $x = \pi$ it follows:

$$\frac{1}{1 \cdot 3} + \frac{1}{3 \cdot 5} + \frac{1}{5 \cdot 7} + \dots = \frac{1}{2}, \quad (12.7-28)$$

while the case $x = \pi/2$ yields:

$$\frac{\pi}{4} = \frac{1}{2} + \left(\frac{1}{1 \cdot 3} - \frac{1}{3 \cdot 5} + \frac{1}{5 \cdot 7} - \frac{1}{7 \cdot 9} + \dots \right). \quad (12.7-29)$$

From the latter we may infer immediately the well-known classical series form of $\pi/4$:

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots = \frac{\pi}{4}, \quad (12.7-30)$$

the so-called *Leibniz series*.

Let us observe that (12.4-3) or (12.4-4) with $x = \pi/2$ imply (12.7-30) as well; nevertheless we remark that the convergence of (12.7-29) is better than that of the Leibniz series.

We insert here a self-evident, but useful statement: *the ordinary Fourier expansion of any trigonometric polynomial $S_N(x)$ is identical with the polynomial itself.* [Cf. (3.4-2).] Hereby is meant that all non-zero coefficients of $S_N(x)$ yield at the same time the corresponding Fourier constants of $S_N(x)$, and all the other (infinitely many) Fourier coefficients have the value 0.

12.76 There are some important special functions among those represented by trigonometric series where the *complex way of writing* is favoured (cf. 10.22). We are going to deal now with *Lerch's* (generalized) *zeta-function*, defined for $\Theta > 0$, $0 < x < 1$, $0 < \omega < 1$ by the formula:¹

$$\zeta(\Theta; x, \omega) = \sum_{m=0}^{\infty} \frac{e^{2m\pi i \omega}}{(x+m)^{\Theta}} \quad (12.7-31)$$

[Cf. (12.7-15).]

It is obvious that the right-hand series in (12.7-31) converges absolutely and uniformly for $\Theta \geq 1 + \varepsilon$ ($\varepsilon > 0$) with arbitrary fixed real ω and $x > 0$, because

$$|e^{2m\pi i \omega} (x+m)^{-\Theta}| = (x+m)^{-\Theta} < m^{-\Theta} \quad (m = 1, 2, \dots).$$

Hence the continuity of $\zeta(\Theta; x, \omega)$ with respect to variables x and ω , and the relationship $\zeta(\Theta; x, 1) = \zeta(\Theta, x)$ are shown.

If besides we apply the elementary identities:

$$\begin{aligned} \sum_{k=0}^m e^{2k\pi i \omega} &= \frac{e^{2(m+1)\pi i \omega} - 1}{e^{2\pi i \omega} - 1} \quad (m = 0, 1, 2, \dots), \\ \sum_{m=0}^M \frac{e^{2m\pi i \omega}}{(x+m)^{\Theta}} &= x^{-\Theta} + \frac{e^{2\pi i \omega}}{e^{2\pi i \omega} - 1} \left[(x+M)^{-\Theta} e^{2M\pi i \omega} - (x+1)^{-\Theta} \right] + \\ &+ \frac{e^{2\pi i \omega}}{e^{2\pi i \omega} - 1} \sum_{m=1}^{M-1} e^{2m\pi i \omega} \left[(x+m)^{-\Theta} - (x+m+1)^{-\Theta} \right], \end{aligned}$$

then by considerations similar to those in 12.73 it can be verified: whatever $x \in (0, 1)$, $\omega \in (0, 1)$ may be, the convergence of the series (12.7-31) holds also for any complex Θ with positive real part, and $\zeta(\Theta; x, \omega)$ is an *analytic* function of Θ in this half-plane.

Let us have $0 < \Theta < 1$ and let us consider for fixed Θ and $\omega \in (0, 1)$ the following function with period 1:

$$F(u) = \zeta(\Theta; \{u\}, \omega) e^{2\pi i \omega \{u\}}. \quad (12.7-32)$$

¹ The sum of the right-hand series, whenever it exists, is a *periodic* function of ω with period 1.

We want to show that the complex form of the ordinary Fourier expansion of $F(u)$ yields a relevant functional equation for Lerch's zeta-function, which is related to Hurwitz's formula (12.7-20).

12.7.1 The function (12.7-31) satisfies the three-term relation¹

$$\begin{aligned} (2\pi)^{\vartheta} \Gamma(\vartheta)^{-1} \zeta(1-\vartheta; x, \omega) &= \\ &= e^{-i\pi\vartheta/2} \cdot e^{2\pi i(1-\omega)x} \zeta(\vartheta; 1-\omega, x) + \\ &+ e^{i\pi\vartheta/2} \cdot e^{2\pi i\omega(-x)} \zeta(\vartheta; \omega, 1-x) \\ (0 < \vartheta < 1, 0 < x < 1, 0 < \omega < 1).^1 \end{aligned} \quad (12.7-33)$$

P r o o f. Since by the above, $u^{\Theta} |F(u)|$ remains bounded for fixed Θ , ω as $u \rightarrow +0$, we have $|F(u)| \in L(0, 1)$, and all complex Fourier coefficients of the function (12.7-32) relating to the fundamental interval $0 < u < 1$ may be formed according to (12.1-4) and (10.2-8)—(10.2-9).

Using still the integral formula (12.7-19), we get:

$$\begin{aligned} c_k &= \int_0^1 F(u) e^{-2k\pi i u} du = \int_0^1 \zeta(\Theta; u, \omega) e^{2\pi i u(\omega-k)} du = \\ &= \sum_{m=0}^{\infty} \int_m^{m+1} t^{-\Theta} e^{2\pi i [(t-m)(\omega-k)+m\omega]} dt = \\ &= \int_0^1 t^{-\Theta} e^{2\pi i t(\omega-k)} dt = \Gamma(1-\Theta) [2\pi i(k-\omega)]^{\Theta-1} \\ &\quad (k = 0, \pm 1, \pm 2, \dots), \end{aligned} \quad (12.7-34)$$

the reversal of order of summation and integration being justified on the basis of 7.5.1.

But the uniform convergence of $\sum_{m=1}^{\infty} (x+m)^{-\Theta} e^{2m\pi i \omega}$ in $0 < x < 1$ for fixed Θ , ω involves the differentiability of $\zeta(\Theta; x, \omega)$ in the same interval; and this by 12.3.4 assures the Fourier expansibility of $F(u)$ at any point $u = x$. Thus we can write for $x \in (0, 1)$:

$$\begin{aligned} \zeta(\Theta; x, \omega) e^{2\pi i \omega x} &= \sum_{k=-\infty}^{\infty} c_k e^{2k\pi i x} = \\ &= \Gamma(1-\Theta) \sum_{k=-\infty}^{\infty} [2\pi i(k-\omega)]^{\Theta-1} e^{2k\pi i x}, \end{aligned} \quad (12.7-35)$$

¹ See Mikolás [10]. This relation was proved by M. LERCH in 1887 by complicated tools of complex analysis (contour integration, calculus of residues). See Lerch [1].

whence the decomposition $\sum_{-\infty}^{\infty} = \sum_1^{\infty} + \sum_{-\infty}^0$ and the change of notation $\Theta = 1 - \vartheta$ lead to (12.7-33).

* * *

Note that Hurwitz's formula (12.7-22) is an easy consequence of (12.7-33), if we transform the latter into the form [cf. (12.7-34), (12.7-35)]:

$$e^{2\pi i x \omega} \zeta(1 - \vartheta; x, \omega) - \int_0^1 \zeta(1 - \vartheta; u, \omega) e^{2\pi i u \omega} du = \\ = \frac{\Gamma(\vartheta)}{(2\pi)^\vartheta} \sum_{n=1}^{\infty} \left[e^{2n\pi i x - \frac{i x \vartheta}{2}} (n - \omega)^{-\vartheta} + e^{-2n\pi i x + \frac{i \pi \vartheta}{2}} (n + \omega)^{-\vartheta} \right]$$

and apply the limiting process $\omega \rightarrow +0$. The fact that the real and imaginary parts of (12.7-31) i.e., the series

$$\sum_{m=0}^{\infty} (x + m)^{-\Theta} \cos(2m\pi\omega) \quad \text{and} \quad \sum_{m=1}^{\infty} (x + m)^{-\Theta} \sin(2m\pi\omega)$$

can be regarded as simplest common generalizations of the series given in (12.7-7) for the Bernoulli polynomials, as well as of the series occurring in Hurwitz's formula (12.7-22), renders understandable the special significance of $\zeta(\Theta; x, \omega)$ in connection with several problems of analysis, geometry and theory of numbers. We add that *Lerch's relation* (12.7-33) is also a natural starting-point of the theory of so-called *zeta-functions of higher order*.¹

12.77 It is well-known that trigonometric Fourier expansions of concrete functions are used in physics and engineering mostly in connection with the so-called *fundamental problem of harmonic analysis*, i.e., when decomposition of arbitrary periodic processes into harmonic vibrations (in other words, their representations in terms of so-called fundamental and accessorial harmonic frequencies) are looked for.

From this aspect, some preceding examples may be interpreted in various ways: 1) the expansion (12.4-3) of (12.4-1) illustrates (apart from linear transformations) the voltage distribution between two poles of a *Hertz oscillator* after discharge by means of sinusoidal components and at the same time the harmonic analysis of so-called *rectangle signs*; 2) the representation (12.7-1) of $\{x\}$ means so-called *saw-tooth vibrations* in electrotechnics;

¹ This function family was defined and discussed first by Epstein [1], towards the beginning of the century.

3) the expansion (12.7-26) of $|\sin x|$ describes the splitting up of *sinusoidal waves* into fundamental and further harmonics at *two-sided rectification*; etc.

On the other hand, the results of Section 12.7 can also show the utility of harmonic analytical methods *from a purely mathematical point of view* as well: Fourier expansions of some appropriate functions open up frequently some elegant new paths — sometimes the only possible way — to establish important relations, functional equations or to evaluate the sums of certain series in closed form, and so on.

Let us refer finally to the fact that trigonometric series are applied in the literature also for constructing *simple generating functions*. (Cf. 10.34.) A typical pertinent example: the *Bessel functions* $J_k(t)$ of integral order (see 10.64) can conveniently be defined by the complex Fourier expansion:

$$e^{it \sin x} = \sum_{k=-\infty}^{\infty} J_k(t) e^{ikx}, \quad (12.7-36)$$

whence we obtain almost without any calculation all the elementary properties of $J_k(t)$, moreover, the famous integral formula

$$J_n(t) = \frac{1}{\pi} \int_0^{\pi} \cos(t \sin u - nu) du \quad (n = 0, 1, 2, \dots), \quad (12.7-37)$$

found by BESSEL in 1824.¹

Note that it is easy to deduce from (12.7-37) the “standard” power series representation of $J_n(t)$ too [cf. (10.6-6)], namely by expanding the integrand in terms of powers of $t \sin u$ with termwise integration for $n = 0$, $n = 1$, and after by induction.

12.8 Series of Haar's type; equiconvergence theorems for Sturm-Liouville series

12.81 The existence of continuous functions with locally divergent ordinary Fourier series suggests us the question: whether we can find such a special type of Fourier series which represents the function at *every* point of continuity, i.e. for which continuity is a sufficient condition of “Fourier expansibility”?

A. HAAR solved this problem affirmatively in 1909, inasmuch as he showed that the Fourier series in terms of the base system (10.8-2) possesses the property just mentioned (plainly very valuable for the applica-

¹ Cf. Bessel [1], Watson [1].

tions). Moreover, the Fourier series pertaining to Haar's orthonormal system $\{\chi_m^{(k)}(t)\}$ of any (L) -integrable function f :

$$\gamma_0^{(0)} \chi_0^{(0)}(x) + \sum_{m=0}^{\infty} \sum_{k=1}^{2^m} \gamma_m^{(k)} \chi_m^{(k)}(x) \left[\gamma_m^{(k)} = \int_0^1 f(t) \chi_m^{(k)}(t) dt \right] \quad (12.8-1)$$

yields as regards convergence also "globally" the maximum to be expected.

Formulating in detail:

12.8.1 *Haar's series (12.8-1) converges for almost all x to $f(x)$. This holds in particular at every point x where f is continuous, and besides uniform in every closed interval of continuity if at the end points f is assumed to be continuous from outside as well.¹*

12.82 The usual verification of theorem 12.8.1 depends on the closed form of partial sums of Haar's series (i.e., on the pendant of Dirichlet's integral) and is rather complicated. In what follows we shall give such a *new and simple proof* whose background lies in the connection of the orthonormal system (10.8-2) with the characteristic functions (10.8-1).²

First of all, let us establish that $\chi_0^{(0)}(t) = \chi_{[0,1]}$ and generally:

$$\chi_m^{(k)}(t) = 2^{m/2} \left(\chi_{d'_{km}} - \chi_{d''_{km}} \right) \quad (m \geq 0, k \geq 1), \quad (12.8-2)$$

where

$$d'_{km} = \left[(k-1)2^{-m}, \left(k - \frac{1}{2} \right) 2^{-m} \right), \quad d''_{km} = \left[\left(k - \frac{1}{2} \right) 2^{-m}, k \cdot 2^{-m} \right).$$

(The definition of Haar's function is extended in a "natural" manner to the origin and to all points of discontinuity.)

Now, the partial sums of the series can be rearranged as follows:

$$\begin{aligned} s_0^{(0)} &= \gamma_0^{(0)} \chi_0^{(0)}(x) = \chi_{[0,1]} \cdot \int_0^1 f(t) dt, \\ s_0^{(1)}(x) &= \gamma_0^{(0)} \chi_0^{(0)}(x) + \gamma_0^{(1)} \chi_0^{(1)}(x) = \\ &= 2 \left[\chi_{[0,1/2]} \int_0^{1/2} f(t) dt + \chi_{[1/2,1]} \int_{1/2}^1 f(t) dt \right], \end{aligned}$$

¹ Haar [1].

² Cf. Mikolás [11].

and so on. In general, we get by induction:

$$\begin{aligned}
 s_{\nu}^{(r)}(x) &= \gamma_0^{(0)} \chi_0^{(0)}(x) + \sum_{m=1}^{\nu-1} \sum_{k=1}^{2^m} \gamma_m^{(k)} \chi_m^{(k)}(x) + \sum_{k=1}^r \gamma_{\nu}^{(k)} \chi_{\nu}^{(k)}(x) = \\
 &= 2^{\nu+1} \sum_{k=1}^r \left[\chi_{d'_{k\nu}} \cdot \int_{(d'_{k\nu})} f(t) dt + \right. \\
 &\quad \left. + \chi_{d''_{k\nu}} \cdot \int_{(d''_{k\nu})} f(t) dt \right] + 2^{\nu} \sum_{k=r+1}^{2^{\nu}} \chi_{d_{k\nu}} \cdot \int_{(d_{k\nu})} f(t) dt,
 \end{aligned}$$

where $d_{k\nu} = d'_{k\nu} \cup d''_{k\nu} = [(k-1)2^{-\nu}, k \cdot 2^{-\nu})$ and $1 \leq r \leq 2^{\nu}$, furthermore, the last sum is to be replaced by 0 for $r = 2^{\nu}$.

Accordingly, each partial sum may be written in the form:

$$s_{\nu}^{(r)}(x) = \sum_{n=1}^N f_{I_n} \cdot \chi_{I_n}(x), \quad (12.8-3)$$

where $I_1, I_2, \dots, I_N$ are all the left closed, right open and pairwise disjoint subintervals belonging to a partition of $[0, 1)$ and

$$f_{I_n} = (I_n)^{-1} \int_{(I_n)} f(t) dt. \quad (12.8-4)$$

Thus there is one single subinterval in the system $\{I_n\}$ which covers the fixed point $x \in [0, 1)$; let us denote it by $i_x = [x_1, x_2)$. Since then $\chi_{i_x} = 1$ and the other characteristic functions in (12.8-3) vanish at point x , the utilization of the integral function $\Phi(u) = \int_0^u f(t) dt$ yields:

$$s_{\nu}^{(r)}(x) = f_{i_x} = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} f(t) dt = \frac{\Phi(x_1) - \Phi(x_2)}{x_1 - x_2}. \quad (12.8-5)$$

The limiting process $\nu \rightarrow \infty$ involves in our case that $\max |I_n| \rightarrow 0$ and $x_1 \rightarrow x - 0$, $x_2 \rightarrow x + 0$ (i.e., the line segment i_x quasi "contracts" into the point x). Consequently, we infer from (12.8-5) by (3.7-2):

$$\lim_{\nu \rightarrow \infty} s_{\nu}^{(r)}(x) = \Phi'(x) = f(x), \quad (12.8-6)$$

provided that x is a Lebesgue point of $f \in L(0, 1)$. But as we know, almost every point $x \in [0, 1)$ and specially any point of continuity of $f(t)$ is a Lebesgue point. (Cf. 7.33 and 7.6.5.)

Finally, if f is continuous in a subinterval $[a, b] \subset [0, 1)$ and also continuous from outside at a as well as at b , then to any given number $\varepsilon > 0$ we can assign a $\delta > 0$ such that in $[a, b]$ we have $|f(t) - f(x)| < \varepsilon$ whenever $|t - x| < \delta$. Therefore by (12.8-5) for any point $x \in [a, b]$: the inequality

$$|s_{\nu}^{(r)}(x) - f(x)| \leq \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} |f(t) - f(x)| dt < \varepsilon,$$

holds if ν is large enough. This completes the proof.

* * *

12.8.3 The above argumentation can essentially be transferred to arbitrary Fourier series of Haar's type, whose base system originates from characteristic functions belonging to a suitable partition sequence of some finite interval by orthogonalization. In this way, we obtain the following counterpart of 12.8.1 for expansions in term of the family $\{\mu_{k_1 \dots k_{\nu}}^{(r)}(t)\}$ of orthogonal functions, discussed in 10.84:

12.8.2 (GENERALIZED EXPANSION THEOREM FOR SERIES OF TYPE $\mu_{k_1 \dots k_{\nu}}^{(r)}$)¹ *The Fourier expansion of any function $f \in L(a, b)$ in terms of the system (10.8-5) represents f at every point $x \in (a, b)$ where the function equals the derivative of its indefinite integral. So this series converges to $f(x)$ almost everywhere and, in particular, at all the points of continuity. Besides the convergence is uniform in any closed subinterval of (a, b) where (the end points included) f is from both sides continuous.*

P r o o f. Let $S_M = S_M(x)$ be a partial sum of the Fourier series in question, and let $i_1, i_2, \dots, i_{\kappa}$ denote the subintervals of $I_0 = (a, b)$ in which $S_M(x)$ equals one single constant. Then S_M (as a step function) may clearly be written in the form

$$S_M(x) = \sum_{n=1}^{\kappa} c_n \chi_{i_n}(x) \quad (12.8-7)$$

with coefficients depending on the function f .

If $f \in L^2(I_0)$, the minimum property of partial sums of an arbitrary Fourier series (cf. 11.1.2) yields immediately

$$c_n = |i_n|^{-1} \int_{(i_n)} f(t) dt \quad (n = 1, 2, \dots, \kappa); \quad (12.8-8)$$

¹ See Mikolás [6].

and this holds also in case of $f \in L(I_0)$, $f \notin L^2(I_0)$, since every such function is the $L(I_0)$ -limit of appropriate functions $\{f_q\} \subset L^2(I_0)$ so that each Fourier coefficient of f is the limit for $q \rightarrow \infty$ of the corresponding Fourier coefficient of f_q .

(12.8-7) and (12.8-8) imply that there is a subinterval i_x (determined uniquely) among i_n ($n = 1, 2, \dots, \kappa$), for which $x \in i_x$ and

$$S_M(x) = |i_x|^{-1} \int_{(i_x)} f(t) dt.$$

Therefore, $S_M(x) \rightarrow f(x)$ as $\max |I_{k_1 \dots k_{\nu+1}}| \rightarrow 0$ ($\nu \geq 2$) provided that

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+h} f(t) dt = f(x) \quad (h \neq 0),$$

which is just the main point of our assertion.

The remaining part of the theorem can be got in the same manner as its analogue for the classic Haar expansion.

* * *

12.84 About some *other special types of Fourier series* in terms of further orthogonal systems studied in Chapter 10, we restrict ourselves here to certain comments of expository character.

The general theory of "regular" Sturm-Liouville eigenvalue problems (outlined in 10.65) enables us to connect closely the asymptotic behaviour of pertaining eigenfunctions with certain facts related to the trigonometric base system. (We mean the so-called LIOUVILLE-HOBSON asymptotic formula.)¹ On this basis, the convergence problems concerning the class of *Sturm-Liouville series* can fully be reduced to the convergence theory of ordinary Fourier series. Namely, it can be shown:

12.8.3 (EQUICONVERGENCE THEOREM OF HAAR) *In the regular case, at some point, any Fourier series of Sturm-Liouville type converges or diverges with the trigonometric Fourier series of the same function, transformed suitably. In the case of convergence, also the sums of the two series are equal.*

Going still further, the classical results about uniqueness may also be transferred to the Sturm-Liouville series just mentioned. (Cf. 12.52)²

¹ Cf. Atkinson [1].

² All these theorems are contained in the doctoral dissertation of A. HAAR and have been published in the Haar papers [1].

As regards the so-called "*singular*" case — e.g. expansions in terms of Jacobi, Laguerre and Hermite polynomials or in terms of Bessel functions — equiconvergence theorems like 12.8.3 can be obtained only much more intricately and often on strong restrictions. Thus the more subtle investigation of these kinds of special Fourier series cannot omit some "ad hoc" methods founded on characteristic properties of the pertinent base systems; anyway, the main expedients of the theory and the function classes needed for convergence criteria are essentially the same as those given for ordinary Fourier series. For instance: *a piecewise monotonic function can be represented at any regular point by its Legendre or Fourier-Bessel series in the classical sense.* (Cf. 10.36 and 10.64.)

As a particularly important field, the convergence theory of *most general Fourier series in terms of weighted orthogonal polynomials* (cf. 10.4) has been developed further also very intensively and fruitfully during the last decades.¹

12.8.2 (GENERALIZED EXPANSION THEOREM FOR SERIES OF TYPE

(μ)^(*) The Fourier expansion of any function $f \in L(a, b)$ in terms of the further orthogonal systems (12.8.1) is uniformly convergent at every point x of the interval (a, b) where f is continuous. (Cf. 10.65) (The generalization of the theorem to the case of non-trigonometric base systems is contained in 10.65.)

12.8.3 (EQUICONVERGENCE THEOREM OF HAAR) In the regular case, at some point, any Fourier series of Sturm-Liouville type converges or diverges with the trigonometric Fourier series of the same function, transformed suitably. In the case of continuous functions, the series are uniformly convergent. (Cf. 12.52)

¹ Cf. e.g. Sz.-Nagy [11], Sansone [1], Szegő [2], Freud [1].

SUMMATION OF ORDINARY FOURIER SERIES

13.1 Linear summation methods and problems of Tauber's type

13.11 The treatment of convergence behaviour of ordinary Fourier series in the preceding chapter — though the pertinent criteria assure the Fourier expansibility of functions belonging to wide classes — also revealed certain defects of convergence theory. On the one hand, as regards *global* properties of the series (12.1-1), it turned out that this Fourier series may diverge everywhere; on the other hand, the punctual continuity of a function does not involve the *local* convergence of its ordinary Fourier expansion, in spite of the fact that (12.1-1) can be formed in the case of any (L) -integrable function. (Cf. 12.11.)

Thus may be raised the natural question: whether the convergence in Cauchy's sense of a series $c_0 + c_1 + c_2 + \dots$ i.e., the existence of

$$s = \lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} (c_0 + c_1 + \dots + c_n) \quad (13.1-1)$$

can be exchanged for such an immediate extension of this concept which enables us to eliminate the above-mentioned inadequacies?

The answer is affirmative: L. FEJÉR (1880-1959), who created one of the Hungarian analytical schools of greatest international influence during his professorial activity in Budapest from 1911 till his death, stated in 1900 that already "averaging" the partial sums, more precisely forming the limit of arithmetic mean of the first $n + 1$ partial sums

$$\begin{aligned} \sigma &= \lim_{n \rightarrow \infty} \frac{1}{n+1} (s_0 + s_1 + \dots + s_n) = \\ &= \lim_{n \rightarrow \infty} \left[c_0 + \left(1 - \frac{1}{n+1}\right) c_1 + \dots + \left(1 - \frac{n}{n+1}\right) c_n \right] \end{aligned} \quad (13.1-2)$$

is a very useful expedient for the indicated purpose. His theorem has been a turning-point not only in the theory of Fourier series, but has given a decisive impulse to a new field of analysis as well, namely to the general theory of so-called *summation processes* which has nowadays an ample literature.¹

13.12 The common fundamental idea of various summation methods is simple: "averaging" finitely or infinitely many partial sums of the series in

¹ Cf. e.g. Fejér [10]-[11]; Hardy [7], Knopp [2].

question by an appropriate transformation of its terms yields an expression, whose limit obtained in a suitable sense leads to a unique generalization of (13.1-1). Hence it can be seen that every method of summation must satisfy the following two basic requirements: 1) in the case of any convergent series it may not furnish a sum different from the value (13.1-1) (*principle of permanence*);¹ 2) it may not yield when applied to any divergent series a sum other than that obtained by any other summation process (*principle of consistency*). Beside these, we still require self-evidently the existence of at least one divergent series having a finite sum with the pertinent summation method.

The actual summation processes are usually denoted by Roman capital letters [(A)-method, (B)-method, etc.], while the results of summation are indicated mostly by the pertaining letter before the summation sign [$^{(A)}\sum$, $^{(B)}\sum$, ...] and called briefly in corresponding manner [(A)-sum, (B)-sum, etc.]. If there exists the sum of some series in the just-mentioned sense, then the series is said to be *summable* by the process in question; thus we speak about (A)-summability, (B)-summability and so on.²

13.13 In the following, besides the Cesàro summation of the first order which was defined by (13.1-2) and has also the shorter name “(C, 1)-method”, we shall need also a strong generalization of this, the *Cesàro summation of order r* [briefly: (C, r)-method], defined for any positive integer r by means of the so-called (C, r)-*mean* of the partial sums $s_0, s_1, \dots, s_n$ as follows:

$$\begin{aligned} {}^{(C,r)}\sum_{k=0}^{\infty} c_k &= \lim_{n \rightarrow \infty} \binom{n+r}{r}^{-1} \sum_{k=0}^n \binom{n+r-k-1}{r-1} s_k = \\ &= \lim_{n \rightarrow \infty} \binom{n+r}{r}^{-1} \sum_{k=0}^n \binom{n+r-k}{r} c_k. \end{aligned} \quad (13.1-3)$$

Note that the latter form (containing the terms c_k) of the (C, r)-sum turns into the ordinary sum $s = \lim_{n \rightarrow \infty} s_n$ for $r = 0$, so that the convergence (in the classical sense) may be called also “(C, 0)-summability”. Going still further, the transformation

$$\binom{n+r-k}{r} = \frac{(n+r-k)!}{r!(n-k)!} = \Gamma(n+1-k)^{-1} \frac{\Gamma(n+r+1-k)}{\Gamma(r+1)} \quad (13.1-4)$$

¹ It is anyway not necessary that the method should be applicable to *all* convergent series.

² It is evident that each summation process has a counterpart relating to *sequences* by which to a divergent sequence we can attach some value replacing its limit in a classical sense. (Namely, we have simply to substitute the partial sums of a series for the terms of a sequence.) In case of sequences the expressions “*limiting method*”, “*limitability*” are used instead of “*summation method*” and “*summability*”, respectively.

[cf. (10.5-25)–(10.5-27)] enables us to extend the definition of (C, r) -method to any non-integral $r > -1$, and to introduce the notion: “ (C, r) -summation of fractional order”.

It can be proved that for $p < q$, any (C, p) -summable series is (C, q) -summable with the same sum; on the other hand, we can always find a series whose (C, q) -sum exists but its (C, p) -sum does not. Therefore, the effectiveness of the (C, r) -method increases with the summability order r . In particular, for $-1 < r \leq 0$ the (C, r) -summability does not imply the convergence; hence we see that the (C, r) -method may be regarded as a proper summation process only if $r > 0$.

It should be noted that for $r = 2, 3, \dots$, the (C, r) -summation is equivalent to the so-called *Hölder summation of order r* [denoted by (H, r)]; in the definition of the latter, the expression after the first limiting sign of (13.1-3) is simply replaced by the r -times iterated arithmetic mean $\sigma_n^{(r)}$ of $s_0, s_1, \dots, s_n$. This is a result due to K. KNOPP and W. SCHNEE (1907-1909).¹

As far as the basic requirements are concerned, it can be immediately seen that every convergent series is Cesàro summable of arbitrary order $r > 0$, its (C, r) -sum equalling its ordinary sum.² Moreover we establish: if a divergent series has a Cesàro sum of some order $p > 0$, then its (C, q) -sum for any $q > p$ does also exist and is identical with the former. This means that all the Cesàro summation processes of positive order are permanent and consistent with each other.

Another very important classical summation process is the so-called *Abel-Poisson summation* [briefly (A) -method]. In case of a series $\sum_{k=0}^{\infty} c_k$, it may be defined by the formula:

$$(A) \sum_{k=0}^{\infty} c_k = \lim_{u \rightarrow 1-0} (1-u) \sum_{k=0}^{\infty} s_k u^k = \lim_{u \rightarrow 1-0} \sum_{k=0}^{\infty} c_k u^k \quad (13.1-5)$$

where we have to assume that the convergence radius of $c_0 + c_1 u + c_2 u^2 + \dots$ is ≥ 1 . [In other words: the last power series must converge for any $u \in (0, 1)$.] According to a well-known *limit theorem of Abel*, the existence of the sum $s = c_0 + c_1 + c_2 + \dots$ implies that of the second limit in (13.1-5) and the

¹ For example, $\sigma_m^{(1)} = \frac{s_0 + s_1 + \dots + s_m}{m+1}$, $\sigma_n^{(2)} = \frac{\sigma_0^{(1)} + \sigma_1^{(1)} + \dots + \sigma_n^{(1)}}{n+1}$, etc. The definition of (H, r) -summation is plainly easier to formulate than that of the (C, r) -method, but the former is practically much more difficult to handle than the latter, since the recursive way of representation of $\sigma_n^{(r)}$ so to say “conceals” the connection between the (H, r) -sum and the terms of the series. Cf. Hölder [1], further Knopp [2], Schnee [1].

² For $r = 1$, the last statement becomes an elementary *limit theorem of Cauchy*: the existence of $\lim_{n \rightarrow \infty} s_n = s$ involves also $(s_0 + s_1 + \dots + s_n)/(n+1) \rightarrow s$ as $n \rightarrow \infty$. Cf. Cauchy [1].

equality in $(A) \sum c_k = s$; so the (A) -summation satisfies the above principle of permanence.

Furthermore, every divergent series $\sum_{k=0}^{\infty} c_k$ which is (C, r) -summable of some order $r > 0$, is simultaneously (A) -summable with the same sum. (The special case $r = 1$ of this assertion is sometimes cited in the literature as "theorem of Frobenius".) Considering that there exist (A) -summable series not (C, r) -summable for any $r > 0$, we may conclude: the (A) -summation is consistent with all the Cesàro summation processes and exceeds any of them in efficiency.

The Abel-Poisson method has plainly the advantage that it is closely related to the theory of power series and thus often allows easy calculation of the sum. For instance, the series $1 - 1 + 1 - + \dots$ is trivially (A) -summable to the sum $1/2$ by

$$1 - u + u^2 - + \dots = \frac{1}{1 + u} \quad (0 < u < 1). \quad (13.1-6)$$

For the discussion of (A) -summability, however, *infinite series* are necessary, unlike the "finite" character of (C, r) -summation.

13.14 Some historical comments: The name of Abel-Poisson summation refers partly to the above-mentioned limit theorem of ABEL, partly to the almost simultaneous investigations of POISSON about the so-called *harmonic series of functions*, the latter being naturally interlinked with the (A) -summability of ordinary Fourier series.¹

The foundation of summation theory proper was laid during the last two decades of the previous century: we have to mention first of all the extension of Abel's theorem for iterated arithmetic means, due to HÖLDER, as well as CESÀRO's theorem on the $(C, 1)$ -summability of Cauchy products.² At the end of the century these results were followed by the discovery of a new procedure by E. BOREL:

$$(B) \sum_{k=0}^{\infty} c_k = \lim_{u \rightarrow \infty} \left(e^{-u} \sum_{k=0}^{\infty} s_k \frac{u^k}{k!} \right) \quad (13.1-7)$$

[called in short (B) -method] which is particularly suitable for summing power series; moreover, in the early 1900's the sudden development initiated by L. FEJÉR's theorem which influenced, along with the theory of Fourier series, mainly the research on the so-called *generalized Dirichlet series*. As regards the latter field, a new summation process introduced by M. RIESZ

¹ Cf. Abel [1], Poisson [1]-[2].

² Hölder [1], Cesàro [1].

in 1909 has turned out to be of great importance; this is based on the iteration of *weighted arithmetic means* and on transforming the multiple sums obtained into simple integrals.¹

A few years later, O. TOEPLITZ (1911) and O. PERRON (1920) showed that each of the summation processes occurring in practice is of linear character, i.e., the sums in question originate from the partial sums of the series by linear operations. The general form of the definition of *linear summation methods* can be written as:

$${}^{(T)} \sum_{k=0}^{\infty} c_k = \lim_{u \rightarrow u_0} \sum_{k=0}^{\infty} \mu_k(u) s_k, \quad (13.1-8)$$

where $\{\mu_k(u)\}$ denotes a sequence of coefficients chosen suitably, depending on the parameter u . The passage to the limit $u \rightarrow u_0$ may be either a "discrete" or a "continuous" one; in the first case, the values of u run over a numerical sequence.²

A theorem of TOEPLITZ yields a necessary and sufficient criterion in order that some summation process of the type (13.1-8) should attach to every convergent series the limit of its partial sums as (T) -sum. (Such a (T) -method is called to be *regular*.) The condition in question consists of three parts:

1) it should be $\lim_{u \rightarrow u_0} \mu_k(u) = 0$ for any fixed index k ;

2) we should have the limit relation $\lim_{u \rightarrow u_0} \sum_{k=0}^{\infty} \mu_k(u) = 1$;

3) the sequence of all partial sums of the series $\sum_{k=0}^{\infty} |\mu_k(u)|$ must be uniformly bounded from above.

The last restriction may be omitted if $\mu_k(u) \geq 0$ for each k . This is obviously the case e.g. for the (A) -method [with $\mu_k(u) = (1-u)u^k$, $0 < u < 1$, $u_0 = 1$]; for the (B) -method [with $\mu_k(u) = e^{-u} \cdot u^k/k!$, $0 < u < \infty$, $u_0 = \infty$]; furthermore for any (C, r) -method of order $r > 0$ where $u_0 = \infty$ and

$$\mu_k(u) = \begin{cases} r \frac{n!}{(n-k)!} \cdot \frac{\Gamma(n+r-k)}{\Gamma(n+r+1)} & \text{if } 0 \leq k \leq n, \\ 0 & \text{if } k > n \end{cases} \quad (13.1-13)$$

with $u = n = 0, 1, 2, \dots$

13.15 A relevant part of deeper theorems of the modern summation theory deals with so-called *problems of Tauber's type* whose object is to infer

¹ Borel [1], M. Riesz [1]. If $\{\mu_k\}$ denotes the sequence of pertaining weights, the usual notation of the Riesz method is: (R, μ_k) . For deeper historical background of this subject see also Hardy-M. Riesz [1] and Mikolás [20].

² Cf. Toeplitz [1], Perron [2].

convergence from summability on some suitable supplementary condition regarding the series.¹ The last-mentioned accessory restriction is plainly necessary, since the convergence always implies summability according to the permanence principle, but the converse is not true.) The name is motivated by the fact that the first proposition of this type was found by A. TAUBER in 1897, related to (A) -summation.² The counterpart of his result about the $(C, 1)$ -summation follows from the identity

$$\begin{aligned} s_n - \frac{s_0 + s_1 + \dots + s_n}{n+1} &= \frac{(s_n - s_0) + (s_n - s_1) + \dots + (s_n - s_{n-1})}{n+1} = \\ &= \frac{(c_1 + c_2 + \dots + c_n) + (c_2 + \dots + c_n) + \dots + c_n}{n+1} = \\ &= \frac{c_1 + 2c_2 + \dots + nc_n}{n+1}. \end{aligned} \quad (13.1-9)$$

Since the last fraction tends to 0 by a well-known limit theorem of Cauchy when

$$\lim_{k \rightarrow \infty} (kc_k) = 0, \quad \text{i.e., } c_k = o(1/k), \quad (13.1-10)$$

the existence of

$$\lim_{n \rightarrow \infty} \frac{s_0 + s_1 + \dots + s_n}{n+1} = \sigma \quad (13.1-11)$$

which, together with (13.1-10) involves also the existence of

$$\lim_{n \rightarrow \infty} s_n = s, \quad (13.1-12)$$

i.e., the convergence of the series $c_0 + c_1 + c_2 + \dots$ and the equality $s = \sigma$.

An improvement of the proposition just verified is, however, not obvious. Here the condition $c_k = O(1/k)$ (i.e., the boundedness of kc_k) is used instead of the o -condition (13.1-10); it will be applied later.

13.1.1 (THEOREM OF HARDY AND LANDAU)³ If a series $\sum_{k=0}^{\infty} c_k$ is $(C, 1)$ -summable and its terms satisfy the condition

$$|c_k| \leq \frac{Q}{k} \quad (Q > 0 \text{ constant, } k \geq k_0 \geq 1), \quad (13.1-13)$$

then this series is convergent, and its sum equals to its $(C, 1)$ -sum.

¹ More generally: suppose that some summation process (T_1) is more efficient than another method of summation (T_2) , that is, the set of series summable by the (T_2) -method forms a proper subset of the totality of series summable by the (T_1) -method. Then any theorem is said to be of *Tauber's type* which enables us to conclude (T_2) -summability from (T_1) -summability.

² Tauber [1].

³ Hardy [1], Landau [1].

P r o o f. Let us put for brevity

$$\sigma_n = \frac{s_0 + s_1 + \dots + s_n}{n+1} \quad (n = 0, 1, 2, \dots) \quad (13.1-14)$$

and

$$\lim_{n \rightarrow \infty} \sigma_n = \sigma. \quad (13.1-15)$$

1° Starting from the identity, valid for all $m \geq 1, n \geq m$:

$$\begin{aligned} (n-m+1)s_n - [(n+1)\sigma_n - m\sigma_{m-1}] &= \\ &= (s_n - s_m) + (s_n - s_{m+1}) + \dots + (s_n - s_{n-1}) = c_{m+1} + 2c_{m+2} + \dots + (n-m)c_n, \end{aligned}$$

we get by a simple calculation:

$$\begin{aligned} s_n - \sigma &= \frac{n+1}{n+1-m}(\sigma_n - \sigma) - \frac{m}{n+1-m}(\sigma_{m-1} - \sigma) + \\ &+ \frac{1}{n+1-m} \sum_{k=1}^{n-m} kc_{m+k}. \end{aligned} \quad (13.1-16)$$

Owing to (13.1-13), the latter sum can be estimated for $m \geq k_0 - 1$ by $k^{-1}(m+k) = 1 + \frac{m}{k} \geq 1 + \frac{m}{n-m} = \frac{n}{n-m}$, i.e., by $k/(m+k) \leq (n-m)/n$ as follows:

$$\left| \sum_{k=1}^{n-m} kc_{m+k} \right| \leq Q \sum_{k=1}^{n-m} \frac{k}{m+k} \leq Q \frac{(n-m)^2}{n}. \quad (13.1-17)$$

Since (13.1-15) is also assumed to exist, thus to any prescribed positive number $\varepsilon < 1$ we can find an integer $m_0 = m_0(\varepsilon) \geq 1$ such that for $m > m_0 + 1$ and $n > m$ the inequalities hold:

$$|\sigma_{m-1} - \sigma| < \varepsilon^2, \quad |\sigma_n - \sigma| < \varepsilon^2. \quad (13.1-18)$$

Using now (13.1-17)-(13.1-18), from (13.1-16) we obtain:

$$|s_n - \sigma| < \frac{n+1+m}{n+1-m} \varepsilon^2 + Q \frac{n+1-m}{n}, \quad (13.1-19)$$

where

$$0 < \varepsilon < 1; \quad m > \max(k_0 - 1, m_0(\varepsilon) + 1) = \nu_\varepsilon; \quad n > m.$$

2° On the other hand, let the natural number n be as large that we have simultaneously $n > 6/\varepsilon$ and $n > (1 + 2\varepsilon)\nu_\varepsilon$ i.e., $n > N_\varepsilon = \max(\frac{6}{\varepsilon}, (1 + 2\varepsilon)\nu_\varepsilon)$ holds. Moreover, choose m in such a way that

$$1 + \varepsilon < \frac{n}{m} < 1 + 2\varepsilon \quad (13.1-20)$$

and therefore

$$\frac{n}{1 + 2\varepsilon} < m < \frac{n}{1 + \varepsilon}. \quad (13.1-21)$$

This is possible, since

$$\frac{n}{1 + \varepsilon} - \frac{n}{1 + 2\varepsilon} = n \frac{\varepsilon}{(1 + \varepsilon)(1 + 2\varepsilon)} > \frac{n\varepsilon}{6} > 1.$$

We see that the first part of (13.1-21) implies $m > n/(1 + 2\varepsilon) > \nu_\varepsilon$, further from (13.1-20) $n > (1 + \varepsilon)m > m$; hence (13.1-19) follows for any n and m in question.

But in our case:

$$\frac{n + 1 + m}{n + 1 - m} = 1 + \frac{2}{(m^{-1}n - 1) + m^{-1}} < 1 + \frac{2}{\varepsilon}$$

and

$$\frac{n + 1 - m}{n} = \left(1 - \frac{m}{n}\right) + \frac{1}{n} < \frac{2\varepsilon}{1 + 2\varepsilon} + \frac{\varepsilon}{6} < 3\varepsilon,$$

so by (13.1-9) we can write:

$$|s_n - \sigma| < \left(1 + \frac{2}{\varepsilon}\right) \varepsilon^2 + Q \cdot 3\varepsilon < 3(Q + 1)\varepsilon \quad (13.1-22)$$

whenever $n > N_\varepsilon$.

The least upper bound becomes with ε as small as we please, which implies indeed $s_n \rightarrow \sigma$ ($n \rightarrow \infty$).

* * *

Note that 13.1.1 may be extended to series which are (C, r) -summable of an arbitrary positive order r and to (A) -summable series, too.¹

¹ Cf. Hardy [1] and Littlewood [1], respectively.

13.2 $(C, 1)$ -summation of ordinary Fourier series; Fejér's fundamental and approximation theorem

13.21 Now we deal with the application of the $(C, 1)$ -method in the case of ordinary Fourier series, which opened up new perspectives of the development at the beginning of the century, as mentioned above.

Recalling the introduction of investigations about convergence in section 12.2, we have first to represent an arbitrary arithmetic mean of the partial sums $s_n(x)$ ($n = 0, 1, 2, \dots$) of the trigonometric Fourier series (12.1-1), i.e.

$$\sigma_N = \frac{1}{N+1} \sum_{n=0}^N s_n(x)$$

in a *closed form*. To that end, it is advisable to transform slightly Dirichlet's integral formula (12.2-2):

$$\begin{aligned} s_n(x) &= \frac{1}{\pi} \int_0^\pi \varphi_x(u) D_n(u) du = \frac{1}{\pi} \int_0^\pi \varphi_x(u) \frac{2 \sin \frac{u}{2} \sin(2n+1)\frac{u}{2}}{2 \sin^2 \frac{u}{2}} du = \\ &= \frac{1}{\pi} \int_0^\pi \varphi_x(u) \frac{\cos nu - \cos(n+1)u}{1 - \cos u} du \quad (n = 0, 1, 2, \dots). \end{aligned}$$

On this basis we get:

$$\begin{aligned} \sigma_N(x) &= \frac{1}{\pi(N+1)} \int_0^\pi \varphi_x(u) \sum_{n=0}^N \frac{\cos nu - \cos(n+1)u}{1 - \cos u} du = \\ &= \frac{1}{\pi(N+1)} \int_0^\pi \varphi_x(u) \frac{1 - \cos(N+1)u}{1 - \cos u} du = \\ &= \frac{1}{\pi(N+1)} \int_0^\pi \varphi_x(u) \left[\frac{\sin(N+1)\frac{u}{2}}{\sin \frac{u}{2}} \right]^2 du. \end{aligned}$$

Briefly, it holds for $N = 0, 1, 2, \dots$:

$$\sigma_N(x) = \frac{1}{\pi} \int_0^\pi \varphi_x(u) H_N(u) du =$$

$$= \frac{2}{\pi} \int_0^{\pi/2} \varphi_x(2u) \cdot \frac{1}{N+1} \left[\frac{\sin(N+1)u}{\sin u} \right]^2 du, \quad (13.2-1)$$

where

$$H_N(u) = \frac{1}{N+1} \frac{1 - \cos(N+1)u}{1 - \cos u} = \frac{1}{N+1} \left[\frac{\sin(N+1)\frac{u}{2}}{\sin \frac{u}{2}} \right]^2 \quad (13.2-2)$$

is the so-called N -th *Fejér kernel function*. We have to stress that this latter function possesses an important property lacking in the case of Dirichlet's kernel, namely the *non-negativity*:

$$H_N(u) \geq 0 \quad (0 \leq u \leq \pi). \quad (13.2-3)$$

Observe that Fejér's kernel satisfies at any point $u \in (0, \pi]$ the inequality $|H_N(u)| \leq (N+1)^{-1} (\sin \frac{u}{2})^{-2}$ so that

$$\lim_{N \rightarrow \infty} H_N(u) = 0 \quad (0 < u \leq \pi), \quad (13.2-4)$$

but

$$H_N(0) = \frac{1}{N+1} \sum_{n=0}^N D_n(0) = \frac{1}{N+1} \sum_{n=0}^N (2n+1) = N+1 \rightarrow \infty \quad (13.2-5)$$

as $N \rightarrow \infty$. Furthermore, since for $f(u) \equiv 1$, $\varphi_x(u) \equiv 1$ we have $s_n(x) \equiv 1$ ($n = 0, 1, \dots$) and $\sigma_N(x) \equiv 1$ ($N = 0, 1, \dots$), the formula (13.2-1) yields:

$$\int_0^\pi H_N(u) du = \pi \quad (N = 0, 1, 2, \dots), \quad (13.2-6)$$

in accordance with (12.2-6).

We shall need still the following trivial consequence of (13.2-1) and (13.2-6), valid for an arbitrary real number σ :

$$\begin{aligned} \sigma_N(x) - \sigma &= \int_0^\pi [\varphi_x(u) - \sigma] H_N(u) du = \\ &= \frac{2}{\pi} \int_0^{\pi/2} [\varphi_x(2u) - \sigma] \cdot \frac{1}{N+1} \left[\frac{\sin(N+1)u}{\sin u} \right]^2 du \\ &\quad (N = 0, 1, \dots). \end{aligned} \quad (13.2-7)$$

The integral in (13.2-1) or (13.2-7) is called *Fejér's integral*. By (13.2-4)—(13.2-5), this is also an example of *singular integral*, whose exceptional (singular) point lies at $u = 0$. The fact that Fejér's integral is

much easier to handle than Dirichlet's one, is due primarily — as will be shown — to the property (13.2-3) of the pertaining kernel.¹

13.22 Using the above, we can deduce simply:

13.2.1 (FEJÉR'S FUNDAMENTAL THEOREM AND APPROXIMATION THEOREM)²

(I) The ordinary Fourier series of a function $f(t)$ is $(C,1)$ -summable at any point x of continuity or of discontinuity of the first kind, its $(C,1)$ -sum equalling $f(x)$ and $\frac{1}{2}[f(x+0) + f(x-0)]$, respectively. In particular, we have at any regular point of f :

$$\sigma_N(x) \rightarrow f(x) \quad (N \rightarrow \infty).$$

(II) If the function f is continuous at every interior point of an interval (a,b) and on both sides at its end points a and b , then $\sigma_N(x)$ ($N = 0, 1, \dots$) tends uniformly to $f(x)$ in the whole interval $[a,b]$.

P r o o f. 1° Suppose that x is a point of continuity or a point of discontinuity of the first kind of f and put $\sigma = \frac{1}{2}[f(x+0) + f(x-0)]$; consider a decomposition of the first integral (13.2-7) into two parts, taken on the subintervals $[0, \delta]$ and $[\delta, \pi]$ with $0 < \delta < \pi$.

Since the definition of $\varphi_x(u)$ implies

$$\varphi_x(u) - \sigma = \frac{1}{2}[(f(x+u) + f(x+0)) + \frac{1}{2}[f(x-u) - f(x-0)]] \rightarrow 0 \quad (u \rightarrow +0),$$

having $\varepsilon > 0$ given, there exists a positive number $\delta = \delta(\varepsilon)$ such that $|\varphi_x(u) - \sigma| < \frac{\varepsilon}{2}$ ($0 \leq u \leq \delta$).

After fixing δ in this manner, using (13.2-3) and (13.2-6) we can write:

$$\left| \frac{1}{\pi} \int_0^\delta [\varphi_x(u) - \sigma] H_N(u) du \right| < \frac{\varepsilon}{2} \cdot \frac{1}{\pi} \int_0^\pi H_N(u) du = \frac{\varepsilon}{2};$$

and using (13.2-4) the inequality is obtained:

$$\begin{aligned} \left| \frac{1}{\pi} \int_0^\pi [\varphi_x(u) - \sigma] H_N(u) du \right| &\leq \\ &\leq \frac{1}{\pi} \frac{1}{(N+1) \sin^2(\delta/2)} \int_0^\pi |\varphi_x(u) - \sigma| du < \frac{\varepsilon}{2}, \end{aligned}$$

provided that N is large enough, say $N \geq N_0(\varepsilon)$.

¹ Cf. Fejér [12].

² Fejér [1], [3].

Thus, from (13.2-7) we obtain:

$$|\sigma_N(x) - \sigma| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad (n \geq N_0(\varepsilon)),$$

which involves assertion (I), since $\varepsilon > 0$ may be chosen arbitrarily small.

2° If function f is continuous in a segment $a \leq x \leq b$ including the conditions $f(a-0) = f(a)$ and $f(b+0) = f(b)$, then, by Heine's theorem, another number $\delta > 0$ can be attached to any $\varepsilon > 0$ so that

$$|f(x \pm u) - f(x)| < \frac{\varepsilon}{2} \quad (a \leq x \leq b, \quad 0 \leq u \leq \delta).$$

Consequently, for the values x and u we have:

$$\begin{aligned} |\psi_x(u)| &= |\varphi_x(u) - f(x)| \leq \\ &\leq \frac{1}{2} [|f(x+u) - f(x)| + |f(x-u) - f(x)|] < \frac{\varepsilon}{2}. \end{aligned}$$

Now, replacing $\varphi_x(u) - \sigma$ by $\psi(u)$ [cf. (12.2-4)] in the preceding argumentation and utilizing

$$\begin{aligned} \int_0^\pi |\psi_x(u)| \, du &\leq \frac{1}{2} \int_0^\pi [|f(x+u)| + |f(x-u)|] \, du + \pi |f(x)| \leq \\ &\leq \frac{1}{2} \int_{-\pi}^\pi |f(t)| \, dt + \pi \max_{a \leq x \leq b} |f(x)|, \end{aligned}$$

we have $|\sigma_N(x) - f(x)| < \varepsilon$ at all points $x \in [a, b]$ simultaneously, if only N is sufficiently large.

Part (II) of the theorem about uniform approximation is verified too.

* * *

13.3 Some consequences of Fejér's theorem

13.31 Let us deal with direct applications of Fejér's theorem!

1) For any convergent series being also $(C, 1)$ -summable and its sum in the classical sense equalling its $(C, 1)$ -sum, it follows that: *if the ordinary Fourier series of a function f is convergent in some interval, then its sum at a point of continuity cannot differ from $f(x)$, and at a point of discontinuity of the first kind it has necessarily the value $\frac{1}{2} [f(x+0) + f(x-0)]$.*

We note that the analogue of this property does not hold for power series: *if a real Taylor series is convergent in some neighbourhood of the*

expansion point, its sum there must not be equal to the corresponding value of the expounded function.¹ The simplest pertinent example was found, as is well-known, by CAUCHY: the Maclaurin series of the function

$$g(x) = \begin{cases} e^{-1/x^2} & (x \neq 0) \\ 0 & (x = 0) \end{cases} \quad (13.3-1)$$

has the form $0 + 0 \cdot x + 0 \cdot x^2 + \dots$ and represents the function identically zero; therefore this sum differs from $g(x)$ at any point $x \neq 0$.

2) If a function $f(x)$ (having the period 2π) is bounded on a closed interval of length 2π , and its range of values is contained in a segment $[m, M]$; then the same holds for the arithmetic means $\sigma_N(x)$ ($N = 0, 1, 2, \dots$) associated with the ordinary Fourier series of f .

The assumption $m \leq f(x) \leq M$ ($-\pi \leq x \leq \pi$) involves also $m \leq \varphi_x(u) \leq M$ ($0 \leq u \leq \pi$). Thus by (13.2-1), (13.2-3), (13.2-6) and using the first mean value theorem of integral calculus we get:

$$\frac{1}{\pi} \int_0^\pi m \cdot H_N(u) du \leq \frac{1}{\pi} \int_0^\pi \varphi_x(u) H_N(u) du \leq \frac{1}{\pi} \int_0^\pi M \cdot H_N(u) du,$$

i.e.

$$m \leq \sigma_N(x) \leq M \quad (-\pi \leq x \leq \pi). \quad (13.3-2)$$

13.32 It is to be stressed that the partial sums $s_N(x)$ of the Fourier series do not possess the property in question: some bound of $f(x)$ is not necessarily a bound of $s_N(x)$ ($N = 0, 1, 2, \dots$). This is clear already in the case of the elementary example (12.4-1): on the one hand we have $-\pi/4 \leq C(x) \leq \pi/4$ for all x , on the other hand e.g. the second partial sum of the ordinary Fourier series of $C(x)$: $s_1(x) = \sin x$ takes values over $\pi/4$ and under $-\pi/4$ alike, as a.o. ± 1 . [The above argumentation cannot be transferred to the case of Dirichlet's integral, because the Dirichlet kernel is not of constant sign, as contrasted with $H_N(u)$.]

At this point, we have to mention a sort of "irregularity" in the behaviour of partial sums $s_N(x)$ as $N \rightarrow \infty$, called *Gibbs phenomenon* after its discoverer (1898).² According to this, the partial sums $s_N(x)$ in the case of a function of bounded variation always "overreach themselves" near to any discontinuity of the first kind in the sense, that $\lim_{N \rightarrow \infty} M_N$ is larger, $\lim_{N \rightarrow \infty} m_N$ is smaller than the limit of the function from the corresponding

¹ We know that the sum function of a power series is not only continuous but also derivable, besides however many times differentiable inside its domain of convergence.

² Cf. Gibbs [1], 258-260.

side, M_N and m_N denoting that local maximum or minimum of $s_N(x)$ which is next to the actual point of discontinuity. An illustration for the above: for $C(x)$ in the vicinity of the origin we have $C(\pm 0) = \pm \pi/4$, and on the other hand:

$$\lim_{N \rightarrow \infty} M_N = \frac{1}{2} \int_0^{\pi} \frac{\sin x}{x} dx > \frac{\pi}{4}$$

and

$$\lim_{N \rightarrow \infty} m_N = -\frac{1}{2} \int_0^{\pi} \frac{\sin x}{x} dx < -\frac{\pi}{4}$$

(Fig. 13.)

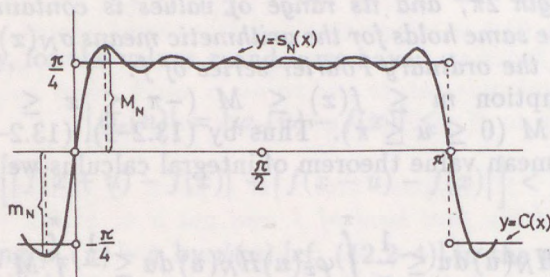


Fig. 13

13.33 Further pertinent observations:

3) On the basis of proposition 13.1.1, we can conclude from Fejér's theorem the convergence of Fourier series, provided that some additional coefficient conditions are known. Thus e.g. Jordan's criterion 12.3.2 or Dirichlet's criterion 12.3.5 can easily be obtained from 13.2.1 by using the estimates (12.1-8).

Generally, recent literature often presents such building up of the convergence theory of trigonometric Fourier series where Fejér's elegant result is the starting point of the treatment, and the classical convergence criteria are deduced therefrom by means of certain ideas of Tauberian type.

4) The approximation theorems of Weierstrass (cf. 3.4.1) are simple corollaries of 13.2.1 (II).

Application of Fejér's approximation theorem in the case of a function f with period 2π and of the fundamental interval $[0, 2\pi]$ yields immediately

$$|f(x) - \sigma_N(x)| < \varepsilon \quad (-\infty < x < \infty), \quad (13.3-3)$$

$\varepsilon > 0$ being as small as we please and $N = N(\varepsilon)$ denoting a sufficiently large integer. Since the mean $\sigma_N(x)$ is plainly a trigonometric polynomial of at most N -th order, herewith we proved not only the assertion 3.4.1 (II) but also a little more, since direct procedure for constructing the approximation polynomial in question has also been presented.

As regards an approximation by entire rational functions, this case can be reduced without difficulty to the former: if $F(u)$ is a given function which

is continuous for $u \in [a, b]$, then we construct therefrom an everywhere continuous, even and periodic function f satisfying $f(x) = F\left(a + \frac{b-a}{\pi}x\right)$ ($0 \leq x \leq \pi$), $f(-x) = f(x)$, $f(x + 2\pi) = f(x)$ ($-\infty < x < \infty$). (The range of values of f relating to $[0, \pi]$ is obviously identical with that of F relating to $[a, b]$.)

After these, we consider a trigonometric polynomial $\sigma_N(x) = A_0 + (A_1 \cos x + B_1 \sin x) + \dots + (A_N \cos Nx + B_N \sin Nx)$ of order N such that

$$|\sigma_N(x) - f(x)| < \frac{\varepsilon}{2} \quad (-\infty < x < \infty). \quad (13.3-4)$$

If the Maclaurin expansion of $\sigma_N(x)$ is denoted by

$$\sigma_N(x) = c_0^{(N)} + c_1^{(N)}x + \dots + c_\nu^{(N)}x^\nu + \dots \quad (-\infty < x < \infty), \quad (13.3-5)$$

and $Q_\nu^{(N)}(x)$ signifies the ν -th partial sum of the right-hand series, then we have for every x and ν sufficiently large:

$$|Q_\nu^{(N)}(x) - \sigma_N(x)| < \frac{\varepsilon}{2}. \quad (13.3-6)$$

Connecting now (13.3-4) and (13.3-6), we obtain

$$\begin{aligned} |Q_\nu^{(N)}(x) - f(x)| &\leq |Q_\nu^{(N)}(x) - \sigma_N(x)| + |\sigma_N(x) - f(x)| < \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad (-\infty < x < \infty) \end{aligned}$$

and in particular

$$\left| F\left(a + \frac{b-a}{\pi}x\right) - Q_\nu^{(N)}(x) \right| < \varepsilon \quad (0 \leq x \leq \pi),$$

if only the integers N and ν are chosen large enough. Hence by the notation $x = \pi(u - a)/(b - a)$ we get

$$\left| F(u) - Q_\nu^{(N)}\left(\pi \frac{u-a}{b-a}\right) \right| < \varepsilon \quad (a \leq u \leq b), \quad (13.3-7)$$

which implies 3.4.1 (I), because the subtrahend on the left is an entire rational function of u .

13.4 Lebesgue's summation theorem and other generalizations

13.41 A few years after Fejér's fundamental results, LEBESGUE succeeded in extending them considerably on the basis of his own new integral theory. First of all, he showed that the $(C, 1)$ -summability of ordinary Fourier series holds not only at points of continuity or of discontinuity of the first kind, but more generally at each point where the so-called absolute integral mean of the function exists.

The pertinent definitions are: $f(t)$ has a (local) *absolute integral mean* at a point $t = x$ and this is equal to a number σ , if the limit condition

$$\lim_{l \rightarrow +0} \frac{1}{2l} \int_{x-l}^{x+l} |f(t) - \sigma| dt = 0 \quad (13.4-1)$$

holds. Is (13.4-1) valid without the sign of absolute value, i.e., in case of the existence of

$$\lim_{l \rightarrow +0} \frac{1}{2l} \int_{x-l}^{x+l} f(t) dt = \sigma, \quad (13.4-2)$$

f has at x an *ordinary integral mean*, equalling σ . [Accordingly, the condition (13.4-1) implies (13.4-2) but the converse is not true; furthermore, both of the mentioned local integral means exist at points x where f is continuous or has a discontinuity of the first kind, when $\sigma = f(x)$ and $\sigma = \frac{1}{2}[f(x+0) + f(x-0)]$, respectively.

The above main result of Lebesgue has the straightforward and important consequence, that the *ordinary Fourier series* is $(C, 1)$ -summable to $f(x)$ *almost everywhere*. Therefore we see: the "anomaly" that a trigonometric Fourier series may diverge on a linear set of positive (L) -measure or everywhere (cf. 12.45), can wholly be eliminated by substituting partial sums for their arithmetic means.

13.42 Formulating in detail, we are going to treat the following proposition found in 1905:

13.4.1 (LEBESGUE'S SUMMATION THEOREM)¹ If a function $f \in L(0, 2\pi)$ satisfies (13.4-1) at some point x , then there the ordinary Fourier series of f is $(C, 1)$ -summable and its $(C, 1)$ -sum has the value σ . Specially for $\sigma = f(x)$, (13.4-1) is fulfilled almost everywhere.

¹ See Lebesgue [5]-[6].

P r o o f. ¹ Since for any $h > 0$

$$\begin{aligned} \frac{1}{h} \int_0^h |\varphi_x(2u) - \sigma| du &\leq \frac{1}{4h} \int_0^{2h} [|f(x+u) - \sigma| + |f(x-u) - \sigma|] du = \\ &= \frac{1}{4h} \int_{x-2h}^{x+2h} |f(t) - \sigma| dt, \end{aligned}$$

the condition (13.4-1) implies that

$$\frac{1}{h} \int_0^h |\varphi_x(2u) - \sigma| du \rightarrow 0 \quad (h \rightarrow +0). \quad (13.4-3)$$

1° From the representation (13.2-7) of Fejér's integral, we conclude by the elementary inequalities $\sin u \geq 2u/\pi$ ($0 \leq u \leq \pi/2$) and $(1+u)|\sin u| \leq 2u$ ($u \geq 0$) for every positive integer N :

$$\begin{aligned} |\sigma_{N-1}(x) - \sigma| &\leq \frac{2}{\pi} \int_0^{\pi/2} |\varphi_x(2u) - \sigma| \cdot \frac{\pi^2 N}{(1+Nu)^2} du = \\ &= 2\pi \int_0^{\pi/2} |\varphi_x(2u) - \sigma| \frac{N}{(1+Nu)^2} du. \end{aligned}$$

But integration by parts yields:

$$\begin{aligned} &\int_0^{\pi/2} |\varphi_x(2u) - \sigma| \frac{N}{(1+Nu)^2} du = \\ &= \Psi_x\left(\frac{\pi}{2}\right) \cdot \frac{N}{(1+N\frac{\pi}{2})^2} + 2 \int_0^{\pi/2} \Psi_x^*(u) \frac{N}{(1+Nu)^3} du, \end{aligned} \quad (13.4-10)$$

where

$$\Psi_x(u) = \int_0^u |\varphi_x(2\tau) - \sigma| d\tau, \quad \Psi_x^*(u) = \Psi_x(u)/u.$$

¹ We present here a verification due to L. FEJÉR, which is much simpler than the original proof of Lebesgue. (Cf. Fejér [9].)

Hence we get by

$$\frac{N^2 u}{(1 + Nu)^3} = \frac{N}{(1 + Nu)^2} \cdot \frac{Nu}{1 + Nu} < \frac{N}{(1 + Nu)^2} \quad (u \geq 0).$$

the inequality

$$\begin{aligned} |\sigma_{N-1}(x) - \sigma| &< \frac{8}{\pi N} \int_0^{\pi/2} |\varphi_x(2\tau) - \sigma| d\tau + \\ &+ 4\pi \int_0^{\pi/2} \Psi_x^*(u) \frac{N}{(1 + Nu)^2} du. \end{aligned} \quad (13.4-4)$$

Now, the first term on the right obviously tends to 0 as $N \rightarrow \infty$, so to any $\varepsilon > 0$ we shall have

$$\frac{8}{\pi N} \int_0^{\pi/2} |\varphi_x(2\tau) - \sigma| d\tau < \varepsilon, \quad (13.4-5)$$

if N is sufficiently large. For estimating the second subintegral in (13.4-4), let us split up it into two parts, taken on $[0, \delta]$ and $[\delta, \pi/2]$.

Concerning these, we begin by establishing that $\lim_{u \rightarrow +0} \Psi_x^*(u) = 0$ owing to (13.4-3). Thus $\Psi_x^*(u) < \varepsilon$ ($0 < u \leq \delta$) holds for δ small enough and also

$$\int_0^\delta \Psi_x^*(u) \frac{N}{(1 + Nu)^2} du < \varepsilon \int_0^{\pi/2} \frac{N}{(1 + Nu)^2} du < \varepsilon. \quad (13.4-6)$$

On the other hand, after fixing δ as above, we can write:

$$\begin{aligned} \int_\delta^{\pi/2} \Psi_x^*(u) \frac{N}{(1 + Nu)^2} du &< \int_\delta^{\pi/2} \Psi_x^*(u) \frac{N}{(Nu)^2} du < \\ &< \frac{1}{N\delta^2} \max_{u \in (0, \pi/2]} \Psi_x^*(u) < \varepsilon. \end{aligned} \quad (13.4-7)$$

From (13.4-4)—(13.4-7) we see that $\lim_{N \rightarrow \infty} |\sigma_{N-1}(x) - \sigma| = 0$, which is equivalent to the main part of the assertion.

2° If we have at a point x :

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+h} |f(t) - f(x)| dt = 0 \quad (h \neq 0), \quad (13.4-8)$$

i.e., a *Lebesgue point* of f is concerned [cf. (7.3-6)], then by

$$\begin{aligned} & \frac{1}{2l} \int_{x-l}^{x+l} |f(t) - f(x)| dt = \\ &= \frac{1}{2} \left(\frac{1}{l} \int_x^{x+l} |f(t) - f(x)| dt + \frac{1}{-l} \int_x^{x-l} |f(t) - f(x)| dt \right) \quad (l > 0) \end{aligned}$$

the function f has an absolute integral mean at x , namely $f(x)$.

But we know that (13.4-8) is fulfilled for a.a. x (cf. 7.6.5); consequently the condition (13.4-1) [assuring by the above the $(C, 1)$ -summability to $f(x)$ of the ordinary Fourier series] holds almost everywhere as well.

* * *

13.43 Note that both theorems 13.2.1 and 13.4.1 are valid also for arbitrary (C, r) ($r > 0$) summation instead of $(C, 1)$ -summation, as was shown in the period 1909–1913 partly by M. RIESZ, partly by G. H. HARDY and S. CHAPMAN.¹ Since the case $r = 0$ may be identified with the classical convergence, we can say that the ordinary Fourier series is a.e. “nearly convergent” in a sense.

A quite different way of further development is the so-called “*strong Cesàro summation*” of Fourier series, introduced by G. H. HARDY and J. E. LITTLEWOOD in 1913. The pertinent main result states that for $N \rightarrow \infty$ not only the difference

$$\sigma_N(x) - f(x) = \frac{1}{N+1} \sum_{n=0}^N [s_n(x) - f(x)], \quad (13.4-9)$$

but also the arithmetic mean of the corresponding deviations, i.e.,

$$\frac{1}{N+1} \sum_{n=0}^N |s_n(x) - f(x)| \quad (13.4-10)$$

tends to 0 almost everywhere and in particular at any point of continuity of f .² Moreover, this assertion may be extended to trigonometric Fourier series

¹ Cf. M. Riesz [1]–[3]; Chapman [1], Hardy [2].

² This makes clear that the $(C, 1)$ -summability of ordinary Fourier series depends primarily on a certain “levelling” of deviations caused by averaging — independently to some extent of sign alterations, too.

of functions f for which $|f|^p$ ($p > 1$) is (L) -integrable, and the deviations in (13.4-10) are replaced by $|s_n(x) - f(x)|^p$ ($n = 0, 1, \dots, N$).¹

During the last decades, other "finite" methods of the type (13.1-8) have also often been utilized in harmonic analysis. Concerning this direction, let us refer e.g. to the papers [4], [7], [8], [10] of B. SZÖKEFALVI-NAGY.

13.5 Summation of ordinary Fourier series by the Abel-Poisson method

13.51 The application of the Abel-Poisson summation to Fourier series is also relevant both from theoretical and practical points of view.

According to the definition (13.1-5), let us start from the so-called harmonic function series

$$\alpha_0 + \sum_{n=1}^{\infty} \rho^n (\alpha_n \cos nx + \beta_n \sin nx) \quad (0 < \rho < 1), \quad (13.5-1)$$

α_n ($n = 0, 1, \dots$) and β_n ($n = 1, 2, \dots$) signifying the ordinary Fourier coefficients of a function f on $[0, 2\pi]$. Since $\alpha_n \rightarrow 0$, $\beta_n \rightarrow 0$ by the Riemann-Lebesgue lemma, the inequality holds for n large enough and for every real x

$$|\rho^n (\alpha_n \cos nx + \beta_n \sin nx)| \leq \rho^n (|\alpha_n| + |\beta_n|) \leq \rho^n.$$

Thus (13.5-1) is uniformly convergent in $-\infty < x < \infty$ for a fixed $\rho \in (0, 1)$; its sum function will be denoted by $U(\rho, x)$.

Entering the value of Fourier constants and using the identities

$$\left. \begin{aligned} 1 + 2 \sum_{n=1}^{\infty} \rho^n e^{in\Theta} &= \frac{2}{1 - \rho e^{i\Theta}} - 1 = \frac{1 + \rho e^{i\Theta}}{1 - \rho e^{i\Theta}}, \\ 1 + 2 \sum_{n=1}^{\infty} \rho^n \cos n\Theta &= \operatorname{Re} \frac{1 + \rho e^{i\Theta}}{1 - \rho e^{i\Theta}} = \frac{1 - \rho^2}{1 - 2\rho \cos \Theta + \rho^2}, \end{aligned} \right\} \quad (13.5-2)$$

we get the closed form of $U(\rho, x)$:

$$\begin{aligned} U(\rho, x) &= \frac{1}{2\pi} \int_0^{2\pi} f(t) dt + \sum_{n=1}^{\infty} \rho^n \cdot \frac{1}{\pi} \int_0^{2\pi} f(t) (\cos nt \cdot \cos nx + \sin nt \cdot \sin nx) dt = \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(t) \left[1 + 2 \sum_{n=1}^{\infty} \rho^n \cos n(x-t) \right] dt = \end{aligned}$$

¹ See e.g. Hardy-Littlewood [1]-[3] and Carleman [1], Marcinkiewicz [1].

$$= \frac{1}{2\pi} \int_0^{2\pi} f(t) \frac{1 - \rho^2}{1 - 2\rho \cos(x - t) + \rho^2} dt.$$

This becomes by the substitution $u = x - t$:

$$\begin{aligned} U_{\rho, x} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \frac{1 - \rho^2}{1 - 2\rho \cos(x - t) + \rho^2} dt = \\ &= \frac{1}{\pi} \int_0^{\pi} \varphi_x(u) P_{\rho}(u) du, \end{aligned} \quad (13.5-3)$$

where

$$P_{\rho}(u) = \frac{1 - \rho^2}{1 - 2\rho \cos u + \rho^2} \quad (13.5-4)$$

is the so-called *Poisson's kernel function*.

The integral in (13.5-3) is named *Poisson's integral*. This is another example of *singular integral* with singular point at the origin. Since $P_{\rho}(u) \rightarrow 0$ ($\rho \rightarrow 1 - 0$) if $0 < u \leq \pi$, but

$$P_{\rho}(0) = \frac{1 - \rho^2}{(1 - \rho)^2} = \frac{1 + \rho}{1 - \rho} \rightarrow +\infty \quad (\rho \rightarrow 1 - 0).$$

The most important property of the kernel (13.5-4) is — like Fejér's kernel function — *positivity*, namely its denominator may be written in the form:

$$1 - 2\rho \cos u + \rho^2 = (1 - \rho)^2 + 4\rho \left(\sin \frac{u}{2} \right)^2 > 0. \quad (13.5-5)$$

Furthermore, (13.5-3) yields for $f(t) = \varphi_x(t) \equiv 1$ the integral formula:

$$\int_0^{\pi} P_{\rho}(u) du = \pi. \quad (13.5-6)$$

Finally, by (13.5-5) we see from (13.5-4) that $P_{\rho}(u)$ is a *monotone decreasing function* of u in the interval $0 < u \leq \pi$.

13.52 From the foregoing we can deduce a pendant of the Fejér-Lebesgue theorem 13.4.1 which contains, however, the more general condition (13.4-2):

13.5.1 If the limit σ given by (13.4-2) exists for a function $f \in L(0, 2\pi)$ at some point x , then there the ordinary Fourier series of f is (A)-summable and its (A)-sum equals σ .

This condition (13.4-2) is fulfilled specially at any point where the derivative of the integral function of f exists and has the value σ ; therefore it holds in case of $\sigma = f(x)$ almost everywhere.

P r o o f. Let us observe first that by

$$\frac{1}{l} \int_0^l \varphi_x(u) du = \frac{1}{2l} \left[\int_0^l f(x+u) du + \int_0^l f(x-u) du \right] = \frac{1}{2l} \int_{x-l}^{x+l} f(t) dt$$

(13.4-2) is equivalent to

$$\lim_{l \rightarrow +0} \frac{1}{l} \int_0^l \varphi_x(u) du = \sigma. \quad (13.5-7)$$

Thus the function

$$\Phi_x(u) = \int_0^u [\varphi_x(\tau) - \sigma] d\tau$$

multiplied by u^{-1} tends to 0 as $u \rightarrow +0$; therefore to any given $\varepsilon > 0$, however small, we can find a number $\delta \in (0, \pi)$ such that

$$|\Phi_x(u)| \leq \frac{\varepsilon}{2} u \quad (0 \leq u \leq \delta). \quad (13.5-8)$$

This δ being fixed, consider now on the basis of (13.5-3) and (13.5-6) the decomposition

$$\begin{aligned} U(\rho, x) - \sigma &= \frac{1}{\pi} \int_0^\pi [\varphi_x(u) - \sigma] P_\rho(u) du = \\ &= \frac{1}{\pi} \int_0^\delta [\varphi_x(u) - \sigma] P_\rho(u) du + \frac{1}{\pi} \int_\delta^\pi [\varphi_x(u) - \sigma] P_\rho(u) du = \\ &= I_1 + I_2. \end{aligned} \quad (13.5-9)$$

Integrating by parts, the first term here becomes:

$$I_1 = \frac{1}{\pi} \left\{ \Phi_x(\delta) P_\rho(\delta) + \int_0^\delta \Phi_x(u) [-P'_\rho(u)] du \right\},$$

where by $P'_\rho(u)$ we mean $dP_\rho(u)/du$, and $-P'_\rho(u) = |P'_\rho(u)|$ since $P_\rho(u)$ is monotone decreasing. Using still the transformation

$$\int_0^\delta u[-P'_\rho(u)] du = -\delta P_\rho(\delta) + \int_0^\delta P_\rho(u) du,$$

we obtain by (13.5-8) and (13.5-6) the inequality:

$$|I_1| \leq \frac{\varepsilon}{2\pi} \int_0^\delta P_\rho(u) du < \frac{\varepsilon}{2\pi} \int_0^\pi P_\rho(u) du = \frac{\varepsilon}{2} \quad (13.5-10)$$

On the other hand, the fundamental properties of the kernel function imply:

$$|I_2| < \frac{1}{\pi} \int_0^\pi |\varphi_x(u) - \sigma| P_\rho(\delta) du \leq \frac{1}{\pi} P_\rho(\delta) \int_0^\pi |\varphi_x(u) - \sigma| du < \frac{\varepsilon}{2}, \quad (13.5-11)$$

provided that $1 - \rho$ is sufficiently small.

The formulae (13.5-9)—(13.5-11) together yield:

$$|U(\rho, x) - \sigma| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

i.e., it follows that

$$\lim_{\rho \rightarrow 1-0} U(\rho, x) = \sigma.$$

The remaining part of the assertion is a consequence of the trivial inequality

$$\left| \frac{1}{2l} \int_{x-l}^{x+l} f(t) dt - \sigma \right| \leq \frac{1}{2} \left[\left| \frac{1}{l} \int_x^{x+l} f(t) dt - \sigma \right| + \left| \frac{1}{-l} \int_x^{x-l} f(t) dt - \sigma \right| \right],$$

and the fact that any (L) -integrable function equals a.e. the derivative of its indefinite (L) -integral. (Cf. 7.6.4.)

* * *

13.53 We have to emphasize that 13.5.1 can be generalized considerably: the existence of the ordinary integral mean (13.4-2) assures not only

the (A) -summability of the series (12.1-1) but also its Cesàro-summability for any order $r > 1$.¹

Note still that the implication $(C, r) (r > 0) \Rightarrow (A)$ (cf. 13.13) renders it possible to transfer all the results about Cesàro-summation of Fourier series — such as Fejér's theorem — to the case of (A) -summation. (The proposition obtained so can mostly be verified without difficulty in a direct manner, namely by use of Poisson's integral, too.)²

The most important application field of the (A) -method is the theory of *harmonic functions*, and a classical boundary problem of mathematical physics, the so-called *Dirichlet's problem* regarding the Laplace equation in the plane. Here we refer to the well-known fact from complex analysis that the generic term in the representation (13.5-1) of $U(\rho, x)$, where ρ and x are to be considered as polar coordinates, is at the same time a *harmonics with respect to the rectangular coordinates* $\xi = \rho \cos x$, $\eta = \rho \sin x$ in the sense that it satisfies the two-dimensional Laplace equation.

13.6 Application of the recent (D) -method

13.61 The Cesàro and Abel–Poisson summation methods enable us to eliminate the insufficiencies of the convergence theory of ordinary Fourier series, mentioned in this chapter; moreover, the results obtained assure the summability of the series not only at any regular point of the function, but almost everywhere, too. The processes in question are *Fourier effective* in the sense that they assign to the Fourier series a.e. the corresponding function value as a sum.

In spite of all these, the summation theory of ordinary Fourier series may by no means be regarded as finished. On the one hand, there exist rather simple real or complex Fourier series which are neither by (C, r) -methods, nor by the (A) -process summable at some single points. An example is yielded by the series

$$\sum_{n=1}^{\infty} (n^{-1-i\tau} \cos nx + b_n \sin nx) \quad (\tau \neq 0 \text{ real, } \sum b_n^2 < \infty), \quad (13.6-1)$$

becoming for $x = 2\nu\pi$ ($\nu = 0, \pm 1, \pm 2, \dots$) the well-known divergent series $\sum n^{-(1+i\tau)}$. The latter cannot be either Cesàro summable of any positive

¹ The $(C, 1)$ -summability, however, does not follow: to that end, the stronger (13.4-1) condition should be necessary. Some historical notes on this topic: the analogue of 13.5.1 concerning Cesàro summation is to be found for the case $r = 2$ also in the paper [3] of H. LEBESGUE, while the extension to any order $r > 1$ is due to W. H. YOUNG [5]. (See further B. SZ.-NAGY [12].)

² For any *square integrable function*, the general Fourier series in terms of an arbitrary orthonormal base system is simultaneously (A) -summable and (C, r) -summable, whatever $r > 0$ may be. (Cf. Zygmund [1]–[2].)

order or Abel summable, otherwise the proposition 13.1.1 and its extension due to HARDY and LITTLEWOOD would imply the convergence of the series.¹

On the other hand, each of the criteria dealt with, even the most general among them namely that given by (13.4-2), presents merely a *sufficient but not necessary* condition of summability; furthermore it is easy to construct a step function whose local integral mean as a limit [cf. (13.4-2)] does not exist at some point. For instance, it is easy to see that if we consider the indefinite (L)-integral of the function defined on $(0, \pi)$ by

$$\gamma(x) = \begin{cases} -1 & \text{for } 2\pi/3^{k+1} < x \leq \pi/3^k \quad (k = 0, 1, 2, \dots), \\ +1 & \text{for } \pi/3^k < x \leq 2\pi/3^k \quad (k = 1, 2, \dots) \end{cases} \quad (13.6-2)$$

and hence continued to be even and periodic with the period 2π , so this integral function has neither unilateral derivatives nor a (symmetrical) derivative of the type (13.4-2) at the points $x = 2\nu\pi$ ($\nu = 0, \pm 1, \dots$) (Fig. 14).²

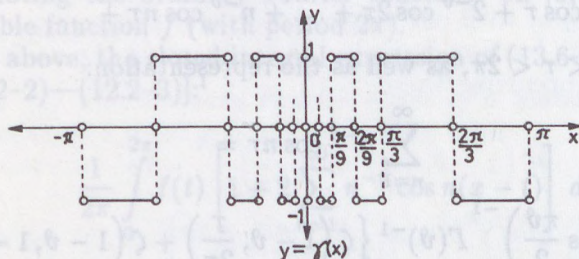


Fig. 14

13.62 All these justify the investigation of a subtle summation process (quasi "made to measure" for Fourier series) which 1) *exceeds in efficiency* all the methods of summation discussed hitherto; 2) permits a very simple, *necessary and sufficient* criterion of summability. This process will be called (D)-method, because its definition contains an ordinary Dirichlet series of the form $\sum_{k=1}^{\infty} c_k \cdot k^{-\vartheta}$ ($\vartheta > 0$); the (D)-sum of a series $\sum_{k=0}^{\infty} c_k$ is given by the limit formula:³

$$({}^D) \sum_{k=0}^{\infty} c_k = c_0 + \lim_{\vartheta \rightarrow +0} \sum_{k=1}^{\infty} c_k \cdot k^{-\vartheta}, \quad (13.6-3)$$

where we have to assume the convergence of the right-hand series for ϑ sufficiently small.

¹ Note that there exist also nowhere (A)-summable orthogonal series whose coefficients are positive and monotone decreasing. Cf. Alexits [2], p. 143; Tandori [2]/II.

² See Mikolás [7].

³ This summation method was first applied to trigonometric series by the author; cf. Mikolás [6]–[8]. For further special literature references see Zeller [2].

We shall need the notion of the so-called *Hurwitz zeta-function* introduced in 12.73 and the *Hurwitz's formula* in the form¹

$$\zeta(1-\vartheta, x) = 2(2\pi)^{-\vartheta} \Gamma(\vartheta) \sum_{n=1}^{\infty} n^{-\vartheta} \cos \left(2n\pi x - \frac{\pi\vartheta}{2} \right) \\ (0 < \vartheta < 1, \quad 0 < x < 1); \quad (13.6-4)$$

furthermore an estimate following from the definition of $\zeta(\Theta, x)$:

$$|\zeta(\Theta, x)| \leq x^{-\Theta} + (1-\Theta)^{-1} + 1 \\ (0 < \Theta < 1, \quad 0 < x \leq 1). \quad (13.6-5)$$

If we substitute x for $1-x$ in (13.6-4) and combine the resulting formula with the original one, then we get with the notation $x = \tau/2\pi$ the convergence of the fundamental cosine series

$$\cos \tau + 2^{-\vartheta} \cos 2\tau + \dots + n^{-\vartheta} \cos n\tau + \dots$$

for $0 < \vartheta < 1$, $0 < \tau < 2\pi$, as well as the representation:

$$\sum_{n=1}^{\infty} n^{-\vartheta} \cos n\tau = \\ = \frac{1}{4}(2\pi)^{\vartheta} \left(\cos \frac{\pi\vartheta}{2} \right)^{-1} \Gamma(\vartheta)^{-1} \left\{ \zeta \left(1-\vartheta, \frac{\tau}{2\pi} \right) + \zeta \left(1-\vartheta, 1-\frac{\tau}{2\pi} \right) \right\}.$$

Hence it is evident that the sum of $\sum n^{-\vartheta} \cos n\tau$ is continuous in the interval $0 < \tau < 2\pi$ and majorizable by an expression of the form $\text{const.} \{ \tau^{\vartheta-1} + (2\pi-\tau)^{\vartheta-1} \}$, consequently also (L) -integrable. We assert the same for *all* partial sums of the series with a *common* constant factor (depending on ϑ only).

In fact, for $0 < \tau \leq \pi$ we can write by the abbreviation $\nu = \left[\frac{\pi}{\tau} \right] + 1$ immediately:

$$\left| \sum_{n=1}^{\nu-1} n^{-\vartheta} \cos n\tau \right| < \int_0^{\nu-1} t^{-\vartheta} dt = \frac{(\nu-1)^{1-\vartheta}}{1-\vartheta} \leq \frac{\pi^{1-\vartheta}}{1-\vartheta} \tau^{\vartheta-1};$$

further, a consideration wholly analogous to that preceding the inequality (12.4-5) leads to

$$\left| \sum_{n=\nu}^N n^{-\vartheta} \cos n\tau \right| \leq \left(\sin \frac{\tau}{2} \right)^{-1} \nu^{-\vartheta} < \pi^{1-\vartheta} \cdot \tau^{\vartheta-1}.$$

¹ Let us note the fact that (13.6-4) is actually a particular case of *Lerch's relation* (12.7-33).

Since the case $\pi \leq \tau < 2\pi$ may be reduced by substitution $\tau = 2\pi - t$ to the former, we obtain finally for $N \geq 2$, $0 < \tau < 2\pi$:

$$\left| \sum_{n=1}^N n^{-\vartheta} \cos n\tau \right| < \left(1 + \frac{1}{1-\vartheta} \right) \pi^{1-\vartheta} \{ \tau^{\vartheta-1} + (2\pi - \tau)^{\vartheta-1} \},$$

and the indicated result follows.

13.63 Now, having in mind the definition (13.6-3) of the (D) -method, let us form the trigonometric series

$$\alpha_o + \sum_{n=1}^{\infty} n^{-\vartheta} (\alpha_n \cos nx + \beta_n \sin nx) \quad (0 < \vartheta < 1, \quad 0 \leq x < 2\pi), \quad (13.6-6)$$

α_n, β_n denoting the ordinary Fourier constants on $[0, 2\pi]$ of a bounded (L) -integrable function f (with period 2π).

By the above, the closed integral expression of (13.6-6) may be written as [cf. (12.2-2)—(12.2-3)]:¹

$$\begin{aligned} & \frac{1}{2\pi} \int_0^{2\pi} f(t) \left[1 + 2 \sum_{n=1}^{\infty} n^{-\vartheta} \cos n(x-t) \right] dt = \\ & = \frac{1}{\pi} \int_0^{\pi} \varphi_x(u) du + \frac{1}{\pi} \int_0^{2\pi} f(x-\tau) \left(\sum_{n=1}^{\infty} n^{-\vartheta} \cos n\tau \right) d\tau = \\ & = \frac{1}{2\pi} \int_0^{2\pi} \varphi_x(u) du + (2\pi)^{\vartheta-1} \left(\cos \frac{\pi\vartheta}{2} \right)^{-1} \Gamma(\vartheta)^{-1} \int_0^{2\pi} \varphi_x(v) \zeta \left(1-\vartheta, \frac{v}{2\pi} \right) dv = \\ & = \frac{1}{2\pi} \int_0^{2\pi} \varphi_x(v) Z_{\vartheta}(v) dv, \end{aligned} \quad (13.6-7)$$

¹ Note that by decomposition in two terms and suitable substitution we have:

$$\begin{aligned} & \int_0^{2\pi} f(x-\tau) \left[\zeta \left(1-\vartheta, \frac{\tau}{2\pi} \right) + \zeta \left(1-\vartheta, 1 - \frac{\tau}{2\pi} \right) \right] d\tau = \\ & \int_0^{2\pi} \left[f(x-v) + f(x+v) \right] \zeta \left(1-\vartheta, \frac{v}{2\pi} \right) dv. \end{aligned}$$

where

$$Z_{\vartheta}(v) = 1 + (2\pi)^{\vartheta} \left(\cos \frac{\pi\vartheta}{2} \right)^{-1} \Gamma(\vartheta)^{-1} \zeta \left(1 - \vartheta, \frac{v}{2\pi} \right). \quad (13.6-8)$$

As regards the inversion of integration and summation, we have to stress: 1) the existence of the integral (13.6-7) is always assured by the boundedness of $f(t)$ and because the sum of the series $\sum_{n=1}^{\infty} n^{-\vartheta} \cos n\tau$ as a function of τ belongs to $L(0, 2\pi)$; 2) the foregoing estimate of the partial sums $\sum_{n=1}^N n^{-\vartheta} \cos n\tau$ justifies termwise integration in the first term on the basis of 7.5.1; 3) this involves simultaneously the convergence of the series (13.6-6) for all values of ϑ and x in consideration.

Taking into account the properties of $\zeta(\Theta, u)$, we see that (13.6-7) is again *singular integral* with one single exceptional point at $v = 0$. Namely we can write on the basis of formula (12.7-17), using the uniform convergence of the series occurring there:

$$\begin{aligned} \vartheta \zeta \left(1 - \vartheta, \frac{v}{2\pi} \right) &= \vartheta \left(\frac{v}{2\pi} \right)^{\vartheta-1} - 1 + \\ &+ \sum_{m=1}^{\infty} \left\{ \vartheta \cdot \left(\frac{v}{2\pi} + m \right)^{\vartheta-1} - [(m+1)^{\vartheta} - m^{\vartheta}] \right\} \rightarrow -1 \quad (\vartheta \rightarrow +0). \end{aligned}$$

It follows for the *kernel function* $Z_{\vartheta}(v)$ at any fixed $v \in (0, 2\pi)$:

$$\begin{aligned} \lim_{\vartheta \rightarrow +0} Z_{\vartheta}(v) &= \lim_{\vartheta \rightarrow +0} \left[1 + (2\pi)^{\vartheta} \left(\cos \frac{\pi\vartheta}{2} \right)^{-1} \Gamma(\vartheta+1)^{-1} \cdot \vartheta \zeta \left(1 - \vartheta, \frac{v}{2\pi} \right) \right] = \\ &= 1 + \lim_{\vartheta \rightarrow +0} \left[\vartheta \zeta \left(1 - \vartheta, \frac{v}{2\pi} \right) \right] = 0. \end{aligned}$$

On the other hand, $Z_{\vartheta}(v)$ is not defined at the origin according to (12.7-16)—(12.7-18); more precisely, $Z_{\vartheta}(v)$ becomes infinite for $v \rightarrow +0$ in the order of $v^{\vartheta-1}$. Furthermore, $Z_{\vartheta}(v)$ is a *continuous and monotone decreasing function of v* in the interval $0 < v \leq 2\pi$.

The most important fact is that after subtracting an appropriate term bearing the “singularity” at $v = 0$, the remaining part of the kernel function already tends uniformly to 0 in $0 \leq v \leq 2\pi$ as $\vartheta \rightarrow +0$. Owing to (12.7-17)—(12.7-18) and using (10.5-26), we get the inequality:

$$\begin{aligned}
& \left| Z_{\vartheta}(v) - 2\pi \left(\cos \frac{\pi \vartheta}{2} \right)^{-1} \Gamma(\vartheta)^{-1} v^{\vartheta-1} \right| = \left| 1 + (2\pi)^{\vartheta} \left(\cos \frac{\pi \vartheta}{2} \right)^{-1} \Gamma(\vartheta)^{-1} \right. \\
& \quad \left. \left\{ -\frac{1}{\vartheta} + \sum_{m=1}^{\infty} \left[\left(\frac{v}{2\pi} + m \right)^{\vartheta-1} - \frac{(m+1)^{\vartheta} - m^{\vartheta}}{\vartheta} \right] \right\} \right| \leq \\
& \leq \left| 1 - (2\pi)^{\vartheta} \left(\cos \frac{\pi \vartheta}{2} \right)^{-1} \Gamma(\vartheta+1)^{-1} \right| + \\
& \quad + (2\pi)^{\vartheta} \left(\cos \frac{\pi \vartheta}{2} \right)^{-1} \frac{\vartheta}{\Gamma(\vartheta+1)}; \tag{13.6-9}
\end{aligned}$$

and both terms of the last bound (independent of v) tend to 0 with ϑ .

13.64 After these premises, we can treat the pertinent summation theorem of the author¹:

13.6.1 (I) *The ordinary Fourier series of a bounded function $f(t)$ is (D)-summable at a point x if and only if the limit*

$$f\langle\langle x \rangle\rangle = \lim_{\vartheta \rightarrow +0} \left[\vartheta \int_0^{\delta} \varphi_x(v) v^{\vartheta-1} dv \right] \tag{13.6-10}$$

exists where $\varphi_x(v) = [f(x+v) + f(x-v)]/2$ and δ denotes a fixed positive number, as small as we please. The value (13.6-10) does not depend on δ , and in case of its existence, it yields the (D)-sum of the Fourier series in question.

(II) In particular, $f\langle\langle x \rangle\rangle = \frac{1}{2}[f(x+0) + f(x-0)]$ holds at any point x where both unilateral limits of the function exist, furthermore the (D)-summability is uniform in each closed continuity interval of the function (at the end points bilateral continuity being meant).

(III) There exist ordinary Fourier series which are not Cesàro-summable of any positive order, not even (A)-summable at certain points, but whose (D)-sum exists on the same set.

P r o o f. 1° Let $0 < \vartheta < 1$, $0 < x \leq 2\pi$ and for brevity write $K \sup_{0 \leq t \leq 2\pi} |f(t)|$.

¹ Cf. Mikolás [8].

Consider a decomposition of the integral representation (13.6-7) of the series (13.6-6) into three parts:

$$\begin{aligned} & \frac{1}{2\pi} \int_0^{2\pi} \varphi_x(v) Z_{\vartheta}(v) dv = \\ &= \frac{1}{2\pi} \int_0^{2\pi} \varphi_x(v) \left[Z_{\vartheta}(v) - 2\pi \left(\cos \frac{\pi\vartheta}{2} \right)^{-1} \Gamma(\vartheta)^{-1} v^{\vartheta-1} \right] dv + \\ & \quad + \left(\cos \frac{\pi\vartheta}{2} \right)^{-1} \Gamma(\vartheta)^{-1} \int_0^{\delta} \varphi_x(v) v^{\vartheta-1} dv + \\ & \quad + \left(\cos \frac{\pi\vartheta}{2} \right)^{-1} \Gamma(\vartheta)^{-1} \int_{\delta}^{2\pi} \varphi_x(v) v^{\vartheta-1} dv = J_1 + J_2 + J_3 \quad (13.6-11) \end{aligned}$$

where $0 < \delta < 1$.

To any given $\varepsilon > 0$ we can attach by (13.6-9) another positive number $\vartheta'_\varepsilon < 1$ such that

$$|J_1| \leq \frac{1}{2\pi} \int_0^{2\pi} |\varphi_x(v)| \cdot \varepsilon dv < \varepsilon \cdot K \quad (13.6-12)$$

if $\vartheta < \vartheta'_\varepsilon$.

As far as the second subintegral is concerned, we obtain by $\Gamma(\vartheta)^{-1} = \vartheta \Gamma(\vartheta + 1)^{-1}$ and $\Gamma(\vartheta + 1) \cos(\pi\vartheta/2) \rightarrow 1$ ($\vartheta \rightarrow 0$):

$$\begin{aligned} & \left| J_2 - \vartheta \int_0^{\delta} \varphi_x(v) v^{\vartheta-1} dv \right| = \\ &= \vartheta \left| \left(\cos \frac{\pi\vartheta}{2} \right)^{-1} \Gamma(\vartheta + 1)^{-1} - 1 \right| \left| \int_0^{\delta} \varphi_x(v) v^{\vartheta-1} dv \right| < \\ &< \left| \left(\cos \frac{\pi\vartheta}{2} \right)^{-1} \Gamma(\vartheta + 1)^{-1} - 1 \right| \cdot K \vartheta \int_0^1 v^{\vartheta-1} dv = \\ &= K \left| \left(\cos \frac{\pi\vartheta}{2} \right)^{-1} \Gamma(\vartheta + 1)^{-1} - 1 \right| < K\varepsilon, \quad (13.6-13) \end{aligned}$$

provided that $\vartheta < \vartheta''_\varepsilon$.

Finally, there exists a number $\vartheta_\varepsilon''' > 0$ such that

$$\begin{aligned} |J_3| &< \left(\cos \frac{\pi\vartheta}{2}\right)^{-1} \Gamma(\vartheta)^{-1} \cdot K(2\pi)^{\vartheta-1} \int_{\delta}^{2\pi} \left(\frac{v}{2\pi}\right)^{-1} dv < \\ &< \vartheta \left(\cos \frac{\pi\vartheta}{2}\right)^{-1} \Gamma(\vartheta+1)^{-1} K \cdot 2\pi \log(2\pi/\delta) < K\varepsilon, \end{aligned} \quad (13.6-14)$$

if $\vartheta < \vartheta_\varepsilon'''$.

Summing up, (13.6-11)—(13.6-14) yield together

$$\left| \frac{1}{2\pi} \int_0^{2\pi} \varphi_x(v) Z_\vartheta(v) dv - \vartheta \int_0^\delta \varphi_x(v) v^{\vartheta-1} dv \right| < 3K\varepsilon$$

for ϑ small enough; and this is equivalent to the conclusion that either both of the limits

$$\lim_{\vartheta \rightarrow +0} \left[\frac{1}{2\pi} \int_0^{2\pi} \varphi_x(v) Z_\vartheta(v) dv \right], \quad \lim_{\vartheta \rightarrow +0} \left[\vartheta \int_0^\delta \varphi_x(v) v^{\vartheta-1} dv \right]$$

exist or none of them, and these limits are equal in the former case.

2° Assuming the existence of $f(x \pm 0)$, we can write for $0 < \eta < \delta < 1$:

$$\begin{aligned} &\left| \vartheta \int_0^\delta \varphi_x(v) v^{\vartheta-1} dv - \frac{1}{2} [f(x+0) + f(x-0)] \right| \leq \\ &\leq \frac{\vartheta}{2} \int_0^\delta [|f(x+v) - f(x+0)| + |f(x-v) - f(x-0)|] v^{\vartheta-1} dv + \\ &\quad + \frac{1}{2} |f(x+0) + f(x-0)| \cdot \left| \vartheta \int_0^\delta v^{\vartheta-1} dv - 1 \right| \leq \\ &\leq \frac{1}{2} \left[\sup_{0 \leq v \leq \eta} \{ |f(x+v) - f(x+0)| \} + \sup_{0 \leq v \leq \eta} \{ |f(x-v) - f(x-0)| \} \right] + \\ &\quad + \frac{\vartheta}{2} \int_\eta^\delta [|f(x+v) - f(x+0)| + |f(x-v) - f(x-0)|] v^{-1} dv + \\ &\quad + \frac{1}{2} |f(x+0) + f(x-0)| (1 - \delta^\vartheta). \end{aligned}$$

The last bound becomes plainly as small as we please, if first η , then (after fixing η) ϑ are chosen sufficiently small.

The assertion on *uniform* (D) -summability follows essentially from the fact that the above upper bounds for the subintegrals J_1, J_2, J_3 are independent of x .

3° It is easy to see that already the series $\sum_{n=1}^{\infty} n^{-(1+i\tau)}$ ($\tau \neq 0$) as a special case of (13.6-1) for $x = 2\nu\pi$ ($\nu = 0, \pm 1, \pm 2, \dots$) presents an example of (D) -summable but non- (A) -summable [and *a fortiori* non- (C, r) -summable] series.

Namely, $\zeta(\Theta, u)$ being a continuous function of Θ for $\Theta \neq 1$, we get for $\tau \neq 0$:

$$\begin{aligned} \lim_{\vartheta \rightarrow +0} \sum_{n=1}^{\infty} n^{-(1+i\tau)} \cdot n^{-\vartheta} &= \\ &= \lim_{\vartheta \rightarrow +0} \zeta(1 + \vartheta + i\tau) = \zeta(1 + i\tau, 1) = \zeta(1 + i\tau). \end{aligned} \quad (13.6-15)$$

Thus the verification of the theorem is completed.

* * *

13.65 It may be noted that 13.6.1 holds also on a slightly extended condition: since another form of (13.6-10) is

$$f\langle\langle x \rangle\rangle = \lim_{\vartheta \rightarrow +0} \left[\vartheta \int_{x-\delta}^{x+\delta} f(t) |t-x|^{\vartheta-1} dt \right], \quad (13.6-16)$$

instead of the boundedness of $f(t)$ it is sufficient to require that its product by any positive power of $|t-x|$ should be bounded in some bilateral symmetric neighbourhood of the point $t=x$.

About the *effectiveness* of the (D) -method we have to stress: *for any bounded f the existence of $f\langle\langle x \rangle\rangle$ implies Lebesgues's condition (13.4-2) and the equality $f\langle\langle x \rangle\rangle = \sigma$, but the converse is not true.*

One of the simplest pertinent examples is yielded by the function (13.6-2): in case of this $\gamma(x)$ the limit (13.4-2) does not exist at the points $x = 0, \pm 2\pi, \pm 4\pi, \dots$; in turn we have

$$\begin{aligned} \vartheta \int_0^{\pi} \gamma(v) v^{\vartheta-1} dv &= \vartheta \left(\int_{2\pi/3}^{\pi} + \int_{\pi/3}^{2\pi/3} + \int_{2\pi/9}^{\pi/3} + \int_{\pi/9}^{2\pi/9} + \dots \right) = \\ &= -\pi^{\vartheta} + 2(2\pi)^{\vartheta} \sum_{k=1}^{\infty} 3^{-k\vartheta} - 2\pi^{\vartheta} \sum_{k=1}^{\infty} 3^{-k\vartheta} = -\pi^{\vartheta} + 2\pi^{\vartheta} \frac{2^{\vartheta} - 1}{3^{\vartheta} - 1} \quad (0 < \vartheta < 1), \end{aligned}$$

whence by L'Hospital's rule it follows for every integer ν :

$$\gamma\langle\langle 2\nu\pi\rangle\rangle = \lim_{\vartheta \rightarrow +0} \left(-\pi^{\vartheta} + 2\pi^{\vartheta} \frac{2^{\vartheta} - 1}{3^{\vartheta} - 1} \right) = \frac{\log 4}{\log 3} - 1. \quad (13.6-17)$$

It has been shown in Mikolás [8] by means of some deeper expedients (e.g. on the basis of certain theorems of Tauberian type) that among the ordinary Fourier series of bounded functions, the totality of (D) -summable series is wider than that of (A) -summable (or all Cesàro-summable) series; moreover *consistency* and *permanence* exist. We may state the *Fourier effectiveness* of (D) -method, too.

The results of this section are closely connected with the recent theory of *generalized integro-differential* (briefly "*fractional*") *operators* which has many applications in classical and modern analysis, and furthermore, in mathematical physics and technics.¹

As an illustration, *Weyl's fractional integral* is mentioned, whose definition may be written for a function $f \in L(0, 1)$ with $\int_0^1 f(t) dt = 0$ and with the period 1 in the form:²

$$f_{(\vartheta)}(x) = \sum_{n=1}^{\infty} (2n\pi)^{-\vartheta} \left[a_n \cos \left(2n\pi x - \frac{\pi\vartheta}{2} \right) + b_n \sin \left(2n\pi x - \frac{\pi\vartheta}{2} \right) \right],$$

where $0 < \vartheta < 1$ and a_n, b_n are ordinary Fourier coefficients of f relating to $[0, 1]$. [The expression "integral of fractional order" originates from the fact that in the case $p = 1, 2, \dots$ the simplest p -th integral functions of $\cos 2n\pi x$, $\sin 2n\pi x$ are $(2n\pi)^{-p} \cos(2n\pi x - \frac{p\pi}{2})$ and $(2n\pi)^{-p} \sin(2n\pi x - \frac{p\pi}{2})$, respectively.]

Correspondingly: if we define the (W_+) - and (W_-) -sum of a trigonometric series periodic with 1 and of constant term 0 by

$$\begin{aligned} (W_{\pm}) \sum_{n=1}^{\infty} (a_n \cos 2n\pi x + b_n \sin 2n\pi x) &= \\ &= \lim_{\vartheta \rightarrow +0} \sum_{n=1}^{\infty} (2n\pi)^{-\vartheta} \left[a_n \cos \left(2n\pi x \pm \frac{\pi\vartheta}{2} \right) + \right. \\ &\quad \left. + b_n \sin \left(2n\pi x \pm \frac{\pi\vartheta}{2} \right) \right], \end{aligned} \quad (13.6-18)$$

¹ Cf. e.g. Al-Bassam [1], Kalla [1], Love [1]–[3], Mikolás [14], Nishimoto [1], M. Riesz [4], Ross [1], Samko–Kilbas–Marichev [1], Srivastava [1].

² See e.g. Zygmund [5], vol. II. 133–142.

supposing the convergence of the series under the limit sign for ϑ small enough, then the pendant of the criterion (13.6-10) in case of

$a_n = 2 \int_0^1 f(t) \cos nt \, dt$, $b_n = 2 \int_0^1 f(t) \sin nt \, dt$ will be the existence of

$$f\langle x \pm 0 \rangle = \lim_{\vartheta \rightarrow +0} \left[\vartheta \int_0^\delta f(x \pm t) t^{\vartheta-1} \, dt \right] \quad [\delta \in (0, 1) \text{ fixed}]^1 \quad (13.6-19)$$

Developing further his research on the above (D)- and ($W_\pm$)-procedures, the author also introduced a few years ago another general summation process for arbitrary series of functions $\varphi_n(x) \in L(x_0, x_1)$ ($n = 1, 2, \dots$), defined by the formula

$$(\mathfrak{M}) \quad \sum_{n=1}^{\infty} \varphi_n(x) = \lim_{\vartheta \rightarrow +0} \sum_{n=1}^{\infty} x_0 I_x^\vartheta \varphi_n. \quad (13.6-20)$$

Here the functional on the right is a so-called *Riemann-Liouville fractional integral*, i.e.,

$$x_0 I_x^\vartheta f = \Gamma(\vartheta)^{-1} \int_{x_0}^x f(t) (x-t)^{\vartheta-1} \, dt \quad (x_0 < x < x_1, \vartheta > 0) \quad (13.6-21)$$

with $f \in L(x_0, x_1)$; we see that (13.6-21) is nothing else but a natural extension of *Cauchy's integral solution of the initial value problem*

$$\left. \begin{aligned} y^{(m)}(x) &= f(x) \quad (x_0 < x < x_1) \\ y(x_0) &= y'(x_0) = \dots = y^{(m-1)}(x_0) = 0. \end{aligned} \right\} \quad (13.6-22)$$

Recent investigations indicate wide utilization possibilities of this $\mathfrak{M}$ -method e.g. to boundary asymptotics of power series in Hadamard's sense.²

¹ Cf. Mikolás [6]-[7].

² Cf. for instance Mikolás [14], and [18]-[22].

FOURIER INTEGRALS

14.1 Foundations of the theory of integral transforms of Fourier and related type

14.11 A formula utilized before:

$$\begin{aligned}\int_0^{\infty} t^{\vartheta-1} e^{-i\lambda t} dt &= \int_0^{\infty} t^{\vartheta-1} \cos \lambda t dt - i \int_0^{\infty} t^{\vartheta-1} \sin \lambda t dt = \\ &= \Gamma(\vartheta)(\lambda i)^{-\vartheta} \quad (0 < \vartheta < 1, \lambda \neq 0 \text{ real})\end{aligned}$$

[cf. (12.7–19)] yields simple examples of so-called *trigonometric integrals* which generally may be written as

$$\int_{-\infty}^{\infty} f(t) e^{-iut} dt = \int_{-\infty}^{\infty} f(t) \cos ut dt - i \int_{-\infty}^{\infty} f(t) \sin ut dt. \quad (14.1-1)$$

Here u is a real parameter and the occurring integrals are to be taken not always in Lebesgue's sense, but often as so-called *Cauchy's principal values*, i.e., of the type $\lim_{T \rightarrow \infty} \int_{-T}^T$. The latter is justified since the *analogue of*

(14.1-1) in the “discrete” case is yielded by the *complex form of trigonometric series* for which the only advisable definition of the sum implies symmetric partial sums. (Cf. 10.22–10.23.)

If the integrals (14.1-1) are considered as abstract functions (operators) which transfer the elements of the function space into certain elements $\Phi(u)$, $\operatorname{Re} \Phi(u)$, $\operatorname{Im} \Phi(u)$ of another space of functions, then we speak about *Fourier transformation*. In more detail: the first term in (14.1-1) is called the (ordinary or exponential) *Fourier transform* of $f(t)$, while for the second and third integral there we use the names (Fourier's) *cosine* and *sine transform* of the function, respectively.

Owing to the wide application possibilities in mathematics and physics, investigations on Fourier transformation have led since the beginning of the past century (mainly since Fourier's research on heat conduction) to the formation of an extensive new branch of analysis, which bears, however, a close relationship partly to ordinary Fourier series, partly to the pendants of power and Laurent series, i.e., to the theory of so-called one- and two-sided *Laplace transforms*:

$$\int_0^{\infty} f(t)e^{-zt} dt \quad \text{and} \quad \int_{-\infty}^{\infty} f(t)e^{-zt} dt \quad (z = x + iy) \quad (14.1-2)$$

or some other similar integral transforms.¹

14.12 When instead of the counterpart of the general notion of trigonometric series the integral analogue of ordinary Fourier series as a particular case is looked for, we have to start from the pertinent “problem of decomposition” (cf. 9.6) and determine from a given function f some “coefficient functions” $A(u)$, $B(u)$ and $C(u) = A(u) - iB(u)$ such that there should hold the formulae:

$$f(x) = \int_{-\infty}^{\infty} C(u)e^{ixu} du \quad (14.1-3)$$

and

$$f(x) = \int_0^{\infty} [A(u) \cos xu + B(u) \sin xu] du, \quad (14.1-4)$$

respectively, possibly at any point x .

The corresponding “discrete” problems have the solutions by 0.22, on suitable conditions concerning f : in the case of:

$$f(x) = \sum_{n=-\infty}^{\infty} C_n e^{inx} \quad (14.1-5)$$

we conclude

$$C_n = \frac{1}{2\pi} \int_a^{a+2\pi} f(t)e^{-int} dt \quad (n = 0, \pm 1, \pm 2, \dots);$$

and in the case of the “real way of writing”

$$f(x) = \sum_{n=0}^{\infty} (A_n \cos nx + B_n \sin nx) \quad (14.1-6)$$

we get

$$2A = \frac{1}{\pi} \int_a^{a+2\pi} f(t) dt,$$

$$A_n = \frac{1}{\pi} \int_a^{a+2\pi} f(t) \cos nt dt, \quad B_n = \frac{1}{\pi} \int_a^{a+2\pi} f(t) \sin nt dt$$

$$(n = 1, 2, \dots),$$

a denoting an arbitrary fixed real number.

Thus it can be expected heuristically that the equations (14.1-3) and (14.1-4) can be satisfied, with a suitable restriction concerning f , by the Fourier transforms

¹ Cf. for instance Doetsch [1], Sneddon [1], Titchmarsh [2], Wiener [1].

$$C(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-iut} dt \quad (14.1-7)$$

and

$$A(u) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \cos ut dt, \quad B(u) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \sin ut dt, \quad (14.1-8)$$

respectively.

In other words: we may assume that on certain general conditions the representations

$$\begin{aligned} f(x) &= \int_{-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-iut} dt \right) e^{iux} du = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t) e^{iu(x-t)} dt du \end{aligned} \quad (14.1-9)$$

or

$$\begin{aligned} f(x) &= \int_0^{\infty} \left[\left(\frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \cos ut dt \right) \cos xu + \right. \\ &\quad \left. + \left(\frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \sin ut dt \right) \sin xu \right] du = \\ &= \frac{1}{\pi} \int_0^{\infty} \int_{-\infty}^{\infty} f(t) \cos u(x-t) dt du \end{aligned} \quad (14.1-10)$$

are valid.

In the literature, (14.1-9) and (14.1-10) are called *Fourier's integral theorem*; the occurring integral analogues of ordinary Fourier series are named *Fourier integrals*.¹ The content of these fundamental connections may also be expressed as the Fourier transform

¹ Sometimes we speak pointedly about Fourier's *double integrals*. Observe that (14.1-9)—(14.1-10) enable us to express also *aperiodic* functions f everywhere as "*continuous superpositions*" of trigonometric functions.

$$\mathcal{F}(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-iut} dt \quad (14.1-11)$$

can be inverted in the form

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{F}(u) e^{ixu} du, \quad (14.1-12)$$

so for (14.1-11)—(14.1-12) together, the term *Fourier's inversion formulae* is used as well. This latter manner of writing has an evident advantage: *symmetry*.

Note that the exponential factors on the right in (14.1-11) and (14.1-12) can be substituted for $\cos ut$ or $\sin ut$, respectively, if f is *even*, while for $-i \sin ut$ or $i \sin xu$ if f is an *odd* function; in both cases, $1/\sqrt{2/\pi}$ is to be replaced by $\sqrt{2/\pi}$ and the interval of integration will be $(0, \infty)$ instead of $(-\infty, \infty)$. Consequently, f and $\mathcal{F}$ must be simultaneously even or simultaneously odd.

14.2 Fourier transformation in Hilbert's function space

14.21 We are getting around to the exact study of Fourier transformation, which presupposes first of all the extension of the concepts of (L) -integrability and square integrability (in Lebesgue's sense) to functions taking *complex* values. It is to be stressed in advance that the treatment of Fourier integrals requires particular precaution owing to the infinite fundamental interval. For instance, from the measurability of some function $f(t)$ for $-\infty < t < \infty$ we may conclude the same property of $f(t)e^{-iut}$, and similarly, the (L) -integrability of $f(t)$ in $(-\infty, \infty)$ implies that $|f(t)e^{-iut}| = |f(t)|$ is also (L) -integrable in $(-\infty, \infty)$ [i.e., $\mathcal{F}(u)$ exists] for any fixed (real) value of the parameter u ; but from the square integrability of $f(t)$ in $(-\infty, \infty)$ the conclusion that $\mathcal{F}(u)$ exists for *every* real u may not be drawn.

Let (a, b) denote a given finite or infinite interval, and let us call a complex-valued function defined in (a, b) measurable or (L) -integrable on a subset of the latter, if there both the real and imaginary parts of the function have the pertinent property. The totality of complex-valued functions (L) -integrable in (a, b) will be denoted by $L(a, b)$. Correspondingly, the so-called *complex $L^2(a, b)$ -space* (Hilbert space) is constituted by those in (a, b) complex-valued, measurable functions $\varphi(t)$ for which $|\varphi(t)|^2$ is (L) -integrable on $a < t < b$. In the case $a = -\infty$, $b = +\infty$, i.e., when

it is about $L^2(-\infty, \infty)$, so any integral over $(-\infty, \infty)$ is to be meant as a Cauchy principal value.

All the basic notions introduced in Chapter 9 for Hilbert's function space can be transferred mechanically to the complex $L^2(a, b)$ -space, but the definition (9.4-1) of the *inner (scalar) product* is to be modified as follows:

$$(\varphi, \psi) = \int_a^b \varphi(t) \overline{\psi(t)} dt \quad [\varphi, \psi \in L^2(a, b)], \quad (14.2-1)$$

$\overline{\psi(t)}$ signifying the complex conjugate of $\psi(t)$; furthermore, the norm in $L^2(a, b)$ is given by the non-negative number

$$\|\varphi\| = (\varphi, \varphi)^{1/2} = \left[\int_a^b |\varphi(t)|^2 dt \right]^{1/2}, \quad (14.2-2)$$

resulting from (14.2-1) for $\varphi = \psi$. On this basis, every statement of Chapter 9 about *orthogonality* essentially remains true also for complex-valued square integrable functions; we have only to take into account that any product of two Fourier constants — by (14.2-1) — is always to be substituted for the product of the first factor by the conjugate of the second factor, as well as the connection $\overline{(\varphi, \psi)} = (\psi, \varphi)$ holds.

Example: the functions e^{int} ($n = 0, \pm 1, \pm 2, \dots$) form an orthogonal system in the complex $L^2(a, a + 2\pi)$ -space (a denoting an arbitrary real number), since by $\overline{e^{int}} = e^{-int}$ we can write

$$(e^{imt}, e^{int}) = \int_a^{a+2\pi} e^{i(m-n)t} dt = \begin{cases} 0 & \text{if } m \neq n, \\ 2\pi & \text{if } m = n. \end{cases}$$

This is the so-called *complex trigonometric base system*, turning by normalization into

$$(2\pi)^{-1/2} e^{int} \quad (n = 0, \pm 1, \pm 2, \dots). \quad (14.2-3)$$

The *complex form of ordinary Fourier series* [cf. (10.2-10)] is nothing else but the Fourier series in terms of the complete orthonormal system (14.2-3) in the sense (11.1-1).

At this point, it is noteworthy that the *Parseval-Hurwitz formula* pertaining to (14.2-3) has the form [cf. (11.2-2), (12.1-7)]:

$$\sum_{n=-\infty}^{\infty} \frac{1}{2\pi} \left| \int_a^{a+2\pi} f(t) e^{-int} dt \right|^2 = \int_a^{a+2\pi} |f(t)|^2 dt, \quad (14.2-4)$$

where $f \in L^2(a, a + 2\pi)$ and $\sum_{n=-\infty}^{\infty}$ is to be taken as $\lim_{N \rightarrow \infty} \sum_{n=-N}^N$.

We shall need still a “continuous” extension of the concept “ L^2 -limit”, introduced for function sequences in Section 9.5. Let ξ be a real parameter, $\xi_0 \in [-\infty, \infty]$; consider a family of functions φ_ξ , depending on ξ , in the complex $L^2(a, b)$ -space. We say that $\varphi_\xi(t)$ has an $L^2(a, b)$ -limit $\varphi(t)$ for $\xi \rightarrow \xi_0$, if there exists an element $\varphi \in L^2(a, b)$ such that for the limiting process in question we have $\|\varphi_\xi - \varphi\| \rightarrow 0$, i.e.,

$$\lim_{\xi \rightarrow \xi_0} \int_a^b |\varphi_\xi(t) - \varphi(t)|^2 dt = 0. \quad (14.2-5)$$

The fact that to $\varphi_\xi(t)$ the $L^2(a, b)$ -limit $\varphi(t)$ is attached, will be denoted — like the L^2 -convergence — as follows:

$$\left. \begin{aligned} \lim_{\xi \rightarrow \xi_0} \varphi_\xi &= \varphi \quad [L^2(a, b)] \quad \text{or} \\ \varphi_\xi(t) &\rightarrow \varphi(t) \quad \text{for } \xi \rightarrow \xi_0 \quad [L^2(a, b)].^1 \end{aligned} \right\} \quad (14.2-6)$$

[Cf. (9.5-5)—(9.5-6).]

It is easy to see that generally no immediate relation holds between the pointwise existence in (a, b) of the ordinary limit $\lim_{\xi \rightarrow \xi_0} \varphi_\xi(t)$ and the existence

of (14.2-6). Anyway, in the case where a function $\varphi_\xi(t) \in L^2(a, b)$ for $\xi \rightarrow \xi_0$ possesses a (finite) $L^2(a, b)$ -limit and also its limit in the classical sense exists and belongs to $L^2(a, b)$ for a.a. fixed $t \in (a, b)$, then these limits are equal to each other almost everywhere. (Cf. 9.52.) By reduction to sequences of functions we get the counterpart of the Riesz–Fischer convergence criterion 9.5.1:

14.2.1 *To the existence of the $L^2(a, b)$ -limit (14.2-6) it is necessary and sufficient that $\|\varphi_{\xi_1} - \varphi_{\xi_2}\|$ should become arbitrarily small whenever ξ_1 , and ξ_2 are sufficiently near to ξ_0 .²*

On the basis of the inequality

$$|\|\varphi_\xi\| - \|\varphi\|| \leq \|\varphi_\xi - \varphi\|$$

[cf. (9.4-6)], the existence of (14.2-6) involves that

$$\lim_{\xi \rightarrow \xi_0} \int_a^b |\varphi_\xi(t)|^2 dt = \int_a^b |\varphi(t)|^2 dt. \quad (14.2-7)$$

¹ The alternative notation: l.i.m. $\varphi_\xi(t) = \varphi(t)$ (“*limes in medio*”) is in use as well.

² For $\xi_0 = \pm\infty$, the latter condition means simply that $|\xi_1|$ and $|\xi_2|$ are large enough.

14.22 From the above we see that the double integral (14.1-9) is the “continuous” analogue of the ordinary Fourier series written in complex form, in which the role of the “discrete” parameter n is taken over by the “continuous” variable u . Correspondingly, instead of summing over the set of all integers n we have to integrate with respect to u from $-\infty$ to ∞ ; and also the integration with respect to t relates to $(-\infty, \infty)$ instead of a finite interval of length 2π . So it is clear that the pendant of formula (14.2-4) in the theory of Fourier transformation must be of the form

$$\int_{-\infty}^{\infty} \frac{1}{2\pi} \left| \int_{-\infty}^{\infty} f(t) e^{-iut} dt \right|^2 du = \int_{-\infty}^{\infty} |f(t)|^2 dt; \quad (14.2-8)$$

therefore the coefficient of index n of the Fourier series in terms of the system (14.2-3) is to be substituted formally in the “continuous” case for the complex Fourier transform $\mathcal{F}(u)$, given by formula (14.1-11).

Let us discuss now the exact validity conditions of the connection just mentioned!

14.2.2 (PARSEVAL'S FORMULA FOR FOURIER INTEGRALS) It holds for any $f \in L^2(-\infty, \infty)$

$$\int_{-\infty}^{\infty} |\mathcal{F}(u)|^2 du = \int_{-\infty}^{\infty} |f(t)|^2 dt \quad (14.2-9)$$

in the sense that we mean by $\mathcal{F}(u)$ the $L^2(-\infty, \infty)$ -limit relating to the limiting process $T \rightarrow \infty$ of

$$\mathcal{F}_T(u) = (2\pi)^{-1/2} \int_{-T}^T f(t) e^{-iut} dt, \quad (14.2-10)$$

further each integral in (14.2-9) is to be taken with its Cauchy principal value.¹

P r o o f. The basic idea of the argumentation stands to reason: the “subintegral” $\int_{-T}^T f(t) e^{-iut} dt$ ($T > 0$) is approached by sums of the form

$$S_N^{(\lambda)}(u) = \sum_{k=-N}^N \gamma_k e^{-i \frac{k}{\lambda} u} \left[\gamma_k = \int_{k/\lambda}^{(k+1)/\lambda} f(t) dt, \quad \lambda > 0 \right], \quad (14.2-11)$$

and thus the proposition is reduced to the “discrete” case in a sense.

¹ Cf. 14.11 and 14.21. The first complete verification of the assertion is due to Plancherel [1], hence the denomination “Parseval-Plancherel formula” occurs too.

1° Let us assume $f(t)$ to be square integrable in $(-\infty, \infty)$, let T and λ denote positive parameters. Moreover let $\lambda T \geq 1$ and put $N = [\lambda T] - 1$, i.e., suppose $(N+1)/\lambda \leq T < (N+2)/\lambda$.

By these assumptions and notations we can write:

$$\begin{aligned} \int_{-T}^T f(t) e^{-iut} dt &= \sum_{k=-N}^N \int_{k/\lambda}^{(k+1)/\lambda} f(t) e^{-iut} dt + \int_{-T}^{-N/\lambda} f(t) e^{-iut} dt + \\ &\quad + \int_{(N+1)/\lambda}^T f(t) e^{-iut} dt, \\ S_N^{(\lambda)}(u) - \int_{-T}^T f(t) e^{-iut} dt &= \sum_{k=-N}^N \int_{k/\lambda}^{(k+1)/\lambda} f(t) \left(e^{-i\frac{k}{\lambda}u} - e^{-iut} \right) dt - \\ &\quad - \int_{-T}^{-N/\lambda} f(t) e^{-iut} dt - \int_{(N+1)/\lambda}^T f(t) e^{-iut} dt, \end{aligned}$$

so that owing to the inequality

$$\begin{aligned} \left| e^{-i\frac{k}{\lambda}u} - e^{-iut} \right| &= 2 \sin \frac{1}{2} \left(t - \frac{k}{\lambda} \right) |u| \leq \left(t - \frac{k}{\lambda} \right) |u| \leq \frac{|u|}{\lambda} \\ &\quad \left(\frac{k}{\lambda} \leq t \leq \frac{k+1}{\lambda} \right) \end{aligned}$$

and by $[-N/\lambda, (N+1)/\lambda] \subset [-T, T]$ we obtain the estimate:

$$\begin{aligned} \left| S_N^{(\lambda)}(u) - \int_{-T}^T f(t) e^{-iut} dt \right| &\leq \frac{|u|}{\lambda} \int_{-T}^T |f(t)| dt + \int_{-N/\lambda}^{-N/\lambda} |f(t)| dt + \\ &\quad + \int_{(N+1)/\lambda}^{(N+2)/\lambda} |f(t)| dt. \end{aligned}$$

But the last bound tends to 0 as $\lambda \rightarrow \infty$ and besides uniformly on any finite segment of the u -axis; consequently, we have for any fixed T , uniformly in an arbitrary interval $-U \leq u \leq U$ ($U > 0$) the limit relation

$$\lim_{\lambda \rightarrow \infty} S_N^{(\lambda)}(u) = \int_{-T}^T f(t) e^{-iut} dt. \quad (14.2-12)$$

2° In turn, for $U \leq \pi\lambda$ the inequalities of Bessel and Schwarz-Buniakowsky yield:

$$\begin{aligned} \int_{-U}^U |S_N^{(\lambda)}(u)|^2 du &\leq \int_{-\pi\lambda}^{\pi\lambda} S_N^{(\lambda)}(u) \overline{S_N^{(\lambda)}(u)} du \leq \\ &\leq 2\pi\lambda \sum_{k=-N}^N \frac{1}{\lambda} \int_{k/\lambda}^{(k+1)/\lambda} |f(t)|^2 dt = \\ &= 2\pi \int_{-N/\lambda}^{(N+1)/\lambda} |f(t)|^2 dt \leq 2\pi \int_{-T}^T |f(t)|^2 dt; \end{aligned}$$

and this combined with (14.2-12), applying the limiting processes $\lambda \rightarrow \infty$ and next $U \rightarrow \infty$ leads to

$$\int_{-\infty}^{\infty} \left| \int_{-T}^T f(t) e^{-iut} dt \right|^2 du \leq 2\pi \int_{-T}^T |f(t)|^2 dt.$$

Therefore it holds for any $T > 0$ [cf. (14.2-10)]:

$$\int_{-\infty}^{\infty} |\mathcal{F}_T(u)|^2 du \leq \int_{-T}^T |f(t)|^2 dt. \quad (14.2-13)$$

Consider now an optional subinterval (T_1, T_2) of $[0, T]$ and apply (14.2-13) to the (square integrable) function, which is identical with $f(t)$ on the segments (T_1, T_2) and $(-T_2, -T_1)$, otherwise vanishes everywhere. In this particular case the inequality (14.2-13) becomes

$$\int_{-\infty}^{\infty} |\mathcal{F}_{T_2}(u) - \mathcal{F}_{T_1}(u)|^2 du \leq \int_{-T_2}^{-T_1} |f(t)|^2 dt + \int_{T_1}^{T_2} |f(t)|^2 dt,$$

whose right-side is by $f \in L^2(-\infty, \infty)$ as small as we please if only T_1 and T_2 are chosen large enough.

In such a way, the criterion 14.2.1 assures the existence of

$$\lim_{T \rightarrow \infty} \mathcal{F}_T(u) = \mathcal{F}(u) \quad [L^2(-\infty, \infty)], \quad (14.2-14)$$

and considering also (14.2-7), (14.2-13) implies for $T \rightarrow \infty$:

$$\int_{-\infty}^{\infty} |\mathcal{F}(u)|^2 du \leq \int_{-\infty}^{\infty} |f(t)|^2 dt. \quad (14.2-15)$$

(Both integrals here are obviously defined as Cauchy's principal values.)

3° It still remains to show that (14.2-15) holds also with the sign of equality. With that end in view, let us state first of all: if we repeat the preceding argumentation for the function $\mathcal{F}(u)$ [given by (14.2-14)] instead of f , then it results

$$\int_{-\infty}^{\infty} |f^*(x)|^2 dx \leq \int_{-\infty}^{\infty} |\mathcal{F}(u)|^2 du, \quad (14.2-16)$$

where

$$f^*(u) = \lim_{U \rightarrow \infty} (2\pi)^{-1/2} \int_{-U}^U \mathcal{F}(u) e^{ixu} du \quad [L^2(-\infty, \infty)].$$

This means that we have only to verify the equality of f and f^* at a.e. point x [which is suggested already by the heuristic result (14.1-12)], since then (14.2-9) would follow simply from comparison of (14.2-15) and (14.2-16).

In fact, on the one hand by the definition of f^* :

$$\begin{aligned} \int_0^x f^*(t) du &= \lim_{U \rightarrow \infty} (2\pi)^{-1/2} \int_{-U}^U \mathcal{F}(u) \int_0^x e^{itu} dt du = \\ &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} \mathcal{F}(u) \frac{e^{ixu} - 1}{iu} du, \end{aligned}$$

on the other hand the formula (14.1-12) yields:

$$\begin{aligned} \int_0^x f(t) dt &= \int_0^x \left[\frac{1}{\sqrt{2\pi}} \lim_{U \rightarrow \infty} \int_{-U}^U \mathcal{F}(u) e^{itu} du \right] dt = \\ &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} \mathcal{F}(u) \frac{e^{ixu} - 1}{iu} du. \end{aligned}$$

[the integrals on $(-\infty, \infty)$ are principal values].

The reverse order of integration and passage to the limit under integration sign is justified in our case by the general theory of Lebesgue integrals as well as by the properties of L^2 -limits.¹

¹ The fact that formal integration of Fourier's integral formulae is always permissible if the result so obtained has a sense, corresponds plainly to 12.5.2.

As a final result, we get $\int_0^x f^*(t) dt = \int_0^x f(t) dt$ for every x , i.e., $f^*(x) = f(x)$ almost everywhere, which was to be proved.

* * *

According to (11.2-3), we can easily deduce a formal extension of (14.2-9), when we apply this formula first to $f + g$, after to $f + ig$ where f and g are arbitrary complex-valued functions of the class $L^2(-\infty, \infty)$. By using the two relationships obtained in the above manner, furthermore considering the identities (holding for arbitrary complex numbers c_1, c_2):

$$\begin{cases} |c_1 + c_2|^2 = |c_1|^2 + |c_2|^2 + 2 \operatorname{Re}(c_1 \overline{c_2}), \\ |c_1 + ic_2|^2 = |c_1|^2 + |c_2|^2 + 2 \operatorname{Im}(c_1 \overline{c_2}), \end{cases}$$

we get after some abbreviations the formula

$$\int_{-\infty}^{\infty} \mathcal{F}(u) \mathcal{G}(u) du = \int_{-\infty}^{\infty} f(t) \overline{g(t)} dt \quad (14.2-17)$$

with $f, g \in L^2(-\infty, \infty)$ and

$$\mathcal{F}(u) = \frac{1}{\sqrt{2\pi}} \lim_{T \rightarrow \infty} \int_{-T}^T f(t) e^{-iut} dt \quad [L^2(-\infty, \infty)] \quad (14.2-18)$$

$$\mathcal{G}(u) = \frac{1}{\sqrt{2\pi}} \lim_{T \rightarrow \infty} \int_{-T}^T g(t) e^{-iut} dt \quad [L^2(-\infty, \infty)], \quad (14.2-19)$$

the integrals on $(-\infty, \infty)$ being again Cauchy's principal values.

The result (14.2-17) — (14.2-19) is the *generalized Parseval formula for Fourier integrals*. Let us remark that by the notions of inner product and norm [cf. (14.2-1) — (14.2-2)], (14.2-9) and (14.2-17) may also be expressed in short as follows:

$$\|\mathcal{F}\|^2 = \|f\|^2 \quad \text{and} \quad \|\mathcal{F}\mathcal{G}\|^2 = \|fg\|^2,$$

respectively.

14.23 A property of central importance of the Fourier series in Hilbert space is, as we know, that it always represents *in the sense of "L²-convergence"* the function $f \in L^2$ from which the Fourier coefficients are formed. (Cf. 11.3.1.) Since in our case this "*general expansion theorem*"

corresponds formally to *Fourier's integral theorem* (14.1-9), by the above we may expect that in the complex Hilbert space the following elegant result holds:

14.2.3 (THEOREM OF PLANCHEREL)¹ *Every complex-valued function $f \in L^2(-\infty, \infty)$ can be represented in the form:*

$$f(x) = \lim_{U \rightarrow \infty} (2\pi)^{-1/2} \int_{-U}^U \mathcal{F}(u) e^{ixu} du \quad [L^2(-\infty, \infty)], \quad (14.2-20)$$

where $\mathcal{F}(u)$ denotes the $L^2(-\infty, \infty)$ -limit (14.2-18). In other words: for any element of the space $L^2(-\infty, \infty)$, Fourier's inversion formulae (14.1-1)–(14.1-2) are valid, provided that the convergence of the occurring integrals over $(-\infty, \infty)$ is meant according to 14.21.

P r o o f. The assertion is a corollary of some arguments contained in the verification of 14.2.2. On the one hand, (14.2-9) implies that $\mathcal{F}(u)$ exists and $\in L^2(-\infty, \infty)$, so it is at the same time (L) -integrable in any finite interval of the u -axis; on the other hand, the last part of the preceding proof involves also the validity of (14.2-20).

* * *

Note that we can get in the same way the representations holding almost everywhere:

$$f(x) = (2\pi)^{-1/2} \frac{d}{dx} \int_{-\infty}^{\infty} \mathcal{F}(u) \frac{e^{iux} - 1}{iu} du, \quad (14.2-21)$$

furthermore

$$\mathcal{F}(u) = (2\pi)^{-1/2} \frac{d}{du} \int_{-\infty}^{\infty} \frac{e^{-itu} - 1}{-it} dt. \quad (14.2-22)$$

We want to stress the following common consequences of Plancherel's theorem as well as of the generalized Parseval formula: I. the complex $L^2(-\infty, \infty)$ space contains together with its any element f also the Fourier transform of the latter taken in the (14.2-18) sense, and if $\mathcal{F}$ is the transform of f , then f is the transform of $\mathcal{F}$ as well (*reciprocity*); II. in $L^2(-\infty, \infty)$, the

¹ Cf. Plancherel [1]. It is understandable that 14.2.3 was discovered in 1910, i.e., soon after the first papers on the Riesz-Fischer theorem had appeared. For other verifications see e.g. F. Riesz [18], Titchmarsh [2], Wiener [1].

connection between all the subsets E and corresponding Fourier transforms $\widehat{E}$ is an *isomorphy* just like the mapping between Hilbert's sequence space and function space on the basis of the Riesz-Fischer theorem;¹ III. especially in the case of any sequence of functions $\{\varphi_n\} \subset L^2(-\infty, \infty)$ constituting an orthonormal system on $(-\infty, \infty)$, the corresponding Fourier transforms form also an orthonormal system in the interval $(-\infty, \infty)$, moreover, if $\{\varphi_n\}$ is complete, the sequence of pertaining Fourier transform has the same property (*invariance of orthogonality*).

It is a relevant fact which is unfortunately disregarded in the pertinent literature, that *Plancherel's theorem can be considered also as pendant of the Riesz-Fischer theorem*. Namely: to a given function f for which the limit of

$$\int_{-T}^T |f(t)|^2 dt \text{ as } T \rightarrow \infty \text{ exists, we can find such a function } \mathcal{F} \in L^2(-\infty, \infty)$$

by 14.2.3, whose (generic) Fourier coefficient in the 14.22 sense equals f ; furthermore this $\mathcal{F}$ cannot be else apart from a null set but the Fourier transform of the type (14.2-18) of f , defined as an $L^2(-\infty, \infty)$ -limit. And the latter assertion corresponds word for word to the main point of theorem 11.4.1 — with the only difference that in our case the role of the “original space” l^2 and of the “image space” L^2 is played by the space $L^2(-\infty, \infty)$. So we have the conclusion: *is the Fourier transformation defined by the formula (14.2-18), then it describes by the foregoing a linear and isometric (i.e., “norm-preserving”) map of the complex Hilbert space $L^2(-\infty, \infty)$ onto itself.*²

We saw before that the *existence* of $\mathcal{F}(u)$ for a.e. u is assured in the sense of $L^2(-\infty, \infty)$ -convergence, moreover by (14.2-22) there holds also a useful explicit representation of $\mathcal{F}(u)$. Nevertheless, a natural question has been unanswered until recently, whether the symbol $[L^2(-\infty, \infty)]$ in (14.2-18) may be omitted, i.e., whether for any function $f \in L^2(-\infty, \infty)$ a.e. the formula

$$\mathcal{F}(u) = (2\pi)^{-1/2} \lim_{T \rightarrow \infty} \int_{-T}^T f(t) e^{-iut} dt \quad (14.2-23)$$

¹ Therefore, if the Fourier transform of an arbitrary element $\varphi \in E$ is denoted by $\mathcal{T}[\varphi]$ and α is some constant, then $\alpha\varphi \in E$ implies $\mathcal{T}[\alpha\varphi] = \alpha\mathcal{T}[\varphi]$, further $\|\mathcal{T}[\varphi]\| = \|\varphi\|$; besides for $\varphi \in E$, $\psi \in E$ we have always $\mathcal{T}[\varphi + \psi] = \mathcal{T}[\varphi] + \mathcal{T}[\psi]$ and $(\mathcal{T}[\varphi], \mathcal{T}[\psi]) = (\varphi, \psi)$. Cf. 11.43.

² In the theory of Hilbert space, such operators are called *unitary transformations*. We recall here a remark made early in 4.21 according to which in this section $L^2(-\infty, \infty)$ signifies the totality of complex-valued functions φ for which the limit of $\int_{-T}^T |\varphi(t)|^2 dt$ exists as $T \rightarrow \infty$; and that in $L^2(-\infty, \infty)$ also any inner product is a principal value.

holds in the *classical* sense.¹ Now it is known that the answer is affirmative, since the problem is equivalent to Lusin's conjecture (mentioned in 12.45) which was proved by Carleson in 1966. This new complementary result does not affect our subject above.

Finally, concerning applications we have to mention an interesting characterization theorem due to R. E. PALEY and N. WIENER [1] on functions of a real variable x given in the form

$$\int_{-T}^T f(t)e^{ixt} dt \quad [f \in L^2(-T, T), T > 0 \text{ fixed}]. \quad (14.2-24)$$

Their result is: a function $F(x)$ may be written as an integral of the type (14.2-24) if and only if it belongs to $L^2(-\infty, \infty)$, furthermore it can be continued analytically to a complex-valued function $F(z)$ ($z = x + iy$) which is regular on the whole complex plane and satisfies everywhere an inequality of the type $|F(z)| \leq Ke^{T|z|}$ (K, T const.) [Therefore $F(z)$ must be a so-called *entire function of exponential type*, T denoting its "type number".]

The above theorem is an important tool in a physical application of Fourier transformation: in the *spectral analysis of electromagnetic waves*.²

14.3 Properties of absolutely convergent Fourier transforms

14.31 As was shown in the previous section, the theory of Fourier transformation relating to the complex $L^2(-\infty, \infty)$ -space can be developed in a quite elegant way, and this presentation may be closely compared with the results on orthogonal systems and series treated in Chapter 11. But the theorems and connections obtained hitherto are based upon the concept of $L^2(-\infty, \infty)$ -convergence which is more abstract and not immediately utilizable from the point of view of applications; therefore now we deal with the convergence of Fourier transforms in the classical sense, and with their corresponding properties.

The simplest condition assuring the existence of the integral (14.1-11) is plainly $f(t) \in L(-\infty, \infty)$ i.e., $\int_{-\infty}^{\infty} |f(t)| dt < \infty$; let us emphasize that the (L) -integrability of $f(t)$ on $(-\infty, \infty)$ has the same meaning as the *absolute convergence* of the Fourier transform

¹ Cf. Zygmund [5], 251.

² See for instance Manolitsakis [1].

$$\mathcal{T}[f] = \mathcal{T}[f(t); x] = (2\pi)^{-1/2} \int_{-\infty}^{\infty} f(t)e^{-ixt} dt \quad (14.3-1)$$

for every real x .¹

In this case the estimate

$$|\mathcal{T}[f; x]| \leq (2\pi)^{-1/2} \int_{-\infty}^{\infty} |f(t)| dt \quad (-\infty < x < \infty) \quad (14.3-2)$$

implies that $\mathcal{T}[f; x]$ as a function of x is in $(-\infty, \infty)$ bounded and continuous, moreover uniformly continuous, too.

Note that we have by the Riemann-Lebesgue lemma extended to infinite intervals (cf. 12.14-12.15):

$$\lim_{x \rightarrow \pm\infty} \mathcal{T}[f; x] = 0. \quad (14.3-3)$$

Easy calculation yields the formulae:

$$\mathcal{T}[f(t - t_0); x] = e^{it_0x} \mathcal{T}[f(t); x] \quad (t_0 \text{ const.}), \quad (14.3-4)$$

$$\mathcal{T}[f(ct); x] = \frac{1}{|c|} \mathcal{T}\left[f(t); \frac{x}{c}\right] \quad (c \neq 0 \text{ const.}). \quad (14.3-5)$$

Other elementary, but fundamental properties of the Fourier transformation are:

14.3.1 (I) $\mathcal{T}[f]$ is a linear operator (cf. 3.1) i.e., for arbitrary constants c_1, c_2 and functions $f_1, f_2 \in L(-\infty, \infty)$ the relation:

$$\mathcal{T}[c_1 f_1 + c_2 f_2] = c_1 \mathcal{T}[f_1] + c_2 \mathcal{T}[f_2] \quad (14.3-6)$$

holds.

(II) If $f(t)$ and $tf(t)$ belong to $L(-\infty, \infty)$, then

$$\mathcal{T}[tf(t); x] = i \frac{d}{dx} \mathcal{T}[f(t); x]. \quad (14.3-7)$$

(III) Let $f(t)$ be absolutely continuous in $(-\infty, \infty)$ and $f(t), f'(t) \in L(-\infty, \infty)$; then

$$\mathcal{T}[f'(t); x] = ix \mathcal{T}[f(t); x]. \quad (14.3-8)$$

¹ The way of writing (14.3-1) which is in accordance with the general symbology of operators (cf. 3.11), is very suitable for indicating the dependence on the function f and on the parameter x simultaneously. In recent times, the abbreviated notation

$\hat{f}(x) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} f(t)e^{-ixt} dt$ has also more and more frequently been used. (See e.g.

Zygmund [5].)

P r o o f. 1° The property (14.3-6) is a trivial consequence of the definition of $\mathcal{T}[f]$.

2° Consider the difference quotient

$$\begin{aligned} & \frac{1}{l} \left\{ \mathcal{T}[f(t); x+l] - \mathcal{T}[f(t); x] \right\} = \\ & = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) \frac{e^{-it(x+l)} - e^{-itx}}{l} dt \quad (l \neq 0). \end{aligned} \quad (14.3-9)$$

Since for any fixed t we can write by Euler's formulae:

$$\begin{aligned} \frac{e^{-it(x+l)} - e^{-itx}}{l} &= -it e^{-it(x+\frac{l}{2})} \frac{\sin(lt/2)}{lt/2} \rightarrow \\ &\rightarrow -it e^{-itx} \quad (l \rightarrow 0), \end{aligned}$$

and the integrand in (14.3-9) is majorized in $(-\infty, \infty)$ by the (L) -integrable function $t|f(t)|$ independently of l , so the passage to the limit $l \rightarrow 0$ may be carried out under the integral sign on the basis of theorem 7.5.1.

But then (14.3-9) becomes for $l \rightarrow 0$ indeed (14.3-7).

3° The absolute continuity of $f(t)$ in $(-\infty, \infty)$ involves the existence of the limits $\lim_{t \rightarrow \pm\infty} f(t)$ owing to the boundedness of the function; and both of the latter have necessarily the value 0, because otherwise it would follow $f \notin L(-\infty, \infty)$.

Integration by parts leads to the expected result:

$$\begin{aligned} \mathcal{T}[f'(t); x] &= [f(t)e^{-itx}]_{t=-\infty}^{t=+\infty} + ix \int_{-\infty}^{\infty} f(t)e^{-itx} dt = \\ &= ix \mathcal{T}[f(t); x]. \end{aligned}$$

* * *

Iterating the formula (14.3-8) in case of higher derivatives, we get

$$\mathcal{T}[f^{(m)}(t); x] = (ix)^m \mathcal{T}[f(t); x] \quad (14.3-10)$$

on the assumptions that $f^{(m-1)}(t)$ exists and is absolutely continuous [especially if $f^{(m)}(t)$ exists and is continuous] in $(-\infty, \infty)$, and when $f^{(k)}(t) \in L(-\infty, \infty)$ ($k = 1, 2, \dots, m$).

It can similarly be deduced from (14.3-7) by mathematical induction:

$$\frac{d^m}{dx^m} \mathcal{T}[f(t); x] = (-i)^m \mathcal{T}[t^m f(t); x] \quad (m = 1, 2, \dots), \quad (14.3-11)$$

provided that $t^m f(t) \in L(-\infty, \infty)$.

14.32 In the theory of Fourier transformation the notion of *convolution* is of fundamental importance defined for a pair of functions $f_1(t)$, $f_2(t)$ by the formula:

$$f_1 \star f_2 = (f_1 \star f_2)(x) = \int_{-\infty}^{\infty} f_1(t) f_2(x-t) dt \quad (-\infty < x < \infty)^1 \quad (14.3-12)$$

About the existence of (14.3-12) we establish:

14.3.2 If f_1 and $f_2 \in L(-\infty, \infty)$, so $f_1 \star f_2$ exists for almost every x and belongs to $L(-\infty, \infty)$ as well. If at least one of f_1 and f_2 is also bounded or both of them belong to $L^2(-\infty, \infty)$, then (14.3-12) exists and is bounded for every x .

P r o o f. Suppose the (L) -integrability of f_1, f_2 in $(-\infty, \infty)$.

1° Let us introduce the auxiliary function

$$g_1(t) = \begin{cases} |f_1(t)| & \text{if } f_1(t) \text{ is finite,} \\ 0 & \text{if } |f_1(t)| = \infty, \end{cases}$$

and study the integral

$$\int_{-\infty}^{\infty} (|f_1| \star |f_2|)(x) dx = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} |f_1(t)| |f_2(x-t)| dt.$$

If we apply Fubini's theorem 8.5.1 for non-negative integrands to the last term, then its existence follows immediately from

$$\int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} |f_1(t)| |f_2(x-t)| dx = \left[\int_{-\infty}^{\infty} g_1(t) dt \right] \cdot \left[\int_{-\infty}^{\infty} |f_2(u)| du \right],$$

¹ The particular case $\int_a^x f_1(t) f_2(x-t) dt$ plays a fundamental role e.g. in the so-called *operational calculus* of J. MIKUSINSKI (1950-52). The "discrete" counterpart of (14.3-12) is clearly the generic coefficient of the Cauchy product of two power series or Laurent series. (For $\sum a_n z^n$ and $\sum b_n z^n$, it is about $\sum_k a_k b_{n-k}$.)

and we obtain

$$\int_{-\infty}^{\infty} (|f_1| \star |f_2|)(x) dx = \left[\int_{-\infty}^{\infty} |f_1(t)| dt \right] \cdot \left[\int_{-\infty}^{\infty} |f_2(t)| dt \right].$$

Thus $f_1 \star f_2$ is also (L) -integrable and a.e. finite in $(-\infty, \infty)$, as asserted above.

2° If we assume still the validity of an inequality $|f_1(t)| \leq C$ (C const., $-\infty < t < \infty$) or the existence of the integrals

$$F_1 = \int_{-\infty}^{\infty} |f_1(t)|^2 dt \quad \text{and} \quad F_2 = \int_{-\infty}^{\infty} |f_2(t)|^2 dt,$$

so we can arrive at the estimates

$$|(f_1 \star f_2)(x)| \leq C \int_{-\infty}^{\infty} |f_2(u)| du \quad (-\infty < x < \infty)$$

and [cf. (9.4-5)]

$$\begin{aligned} |f_1 \star f_2(x)|^2 &\leq \left[\int_{-\infty}^{\infty} |f_1(t)|^2 dt \right] \cdot \left[\int_{-\infty}^{\infty} |f_2(u)|^2 du \right] = \\ &= F_1 F_2 \quad (-\infty < x < \infty), \end{aligned}$$

respectively. Hence we see the existence of $f_1 \star f_2$ and its boundedness in $-\infty < x < \infty$.

* * *

Elementary properties of (14.3-12) are:

14.3.3 For (L) -integrable functions, the forming of convolution is an almost everywhere commutative and associative operation. In particular, if one of f_1 and f_2 is bounded and the other belongs to $L(-\infty, \infty)$ or both of them $\in L^2(-\infty, \infty)$, then (14.3-12) is uniformly continuous in $-\infty < x < \infty$.

P r o o f. 1° If f_1 and $f_2 \in L(-\infty, \infty)$, so $f_1 \star f_2$ exists a.e. by the preceding theorem, and the substitution $x - t = u$ leads to

$$(f_1 \star f_2)(x) = \int_{-\infty}^{\infty} f_1(x - u) f_2(u) du = (f_2 \star f_1)(x).$$

Let us consider three functions: f_1 , f_2 and f_3 belonging to the class $L(-\infty, \infty)$. Then using again theorem 14.3.2, the existence of both $(f_1 \star f_2) \star f_3$ and $f_1 \star (f_2 \star f_3)$ is assured for almost every x ; moreover, we get at any such point x on the basis of commutativity:

$$\begin{aligned} f_1 \star (f_2 \star f_3) &= \int_{-\infty}^{\infty} f_1(t) \left[\int_{-\infty}^{\infty} f_2((x-t)-u) f_3(u) du \right] dt = \\ &= \int_{-\infty}^{\infty} f_3(u) \left[\int_{-\infty}^{\infty} f_1(t) f_2((x-u)-t) dt \right] du = (f_1 \star f_2) \star f_3. \end{aligned}$$

2° Let now $f_1(t)$ be bounded: $|f_1(t)| \leq C$ ($-\infty < t < \infty$) and $f_2(t) \in L(-\infty, \infty)$. Then by 14.3.2 $f_1 \star f_2$ exists everywhere, and the absolute value of the difference

$$(f_1 \star f_2)(x+l) - (f_1 \star f_2)(x) = \int_{-\infty}^{\infty} f_1(t) [f_2(x+l-t) - f_2(x-t)] dt$$

can be majorized by

$$C \int_{-\infty}^{\infty} |f_2(u+l) - f_2(u)| du.$$

Owing to 7.5.4, the last term becomes as small as we please, provided that $|l|$ is small enough; hence it follows that $f_1 \star f_2$ is uniformly continuous in $(-\infty, \infty)$.

In the case $f_1(t) \in L^2(-\infty, \infty)$, $f_2(t) \in L^2(-\infty, \infty)$ we can argue similarly, observing that [cf. (9.4-5)]:

$$\begin{aligned} &|(f_1 \star f_2)(x+l) - (f_1 \star f_2)(x)|^2 \leq \\ &\leq \left[\int_{-\infty}^{\infty} |f_1(t)|^2 dt \right] \cdot \left[\int_{-\infty}^{\infty} |f_2(u+l) - f_2(u)|^2 du \right]. \end{aligned}$$

* * *

14.33 After these, it is easy to verify one of the most important features of the theory, namely that by means of Fourier transformation the operation of "convolution forming" (i.e., a "transcendental" operation) can be reduced to ordinary ("algebraic") multiplication.¹

¹ The "discrete" analogue of the property in question is obviously the fact, that the forming of a sum of the type $\sum_k a_k b_{n-k}$ can be substituted for ordinary multiplication by using the corresponding power series or Laurent series.

Therefore in the class of (L) -integrable functions the functional $\mathcal{T}[f]$ (and its extension: the Laplace transformation) plays an analogous simplifying role in connection with convolution forming as the logarithm function with respect to the multiplication of positive numbers.

14.3.4 (CONVOLUTION THEOREM) *If $f_1, f_2 \in L(-\infty, \infty)$, then the relation holds*

$$\mathcal{T}[f_1 \star f_2] = (2\pi)^{-1/2} \mathcal{T}[f_1] \mathcal{T}[f_2] \quad (14.3-13)$$

P r o o f. Since $f_1 \star f_2 \in L(-\infty, \infty)$ by 14.3.2, all three Fourier transforms occurring in (14.3-13) exist.

Thus the application of Fubini's theorem on double integration yields:

$$\begin{aligned} \mathcal{T}[f_1 \star f_2] &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-ixt} \left[\int_{-\infty}^{\infty} f_1(u) f_2(t-u) du \right] dt = \\ &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} f_1(u) \left[\int_{-\infty}^{\infty} f_2(\tau) e^{-ix(u+\tau)} d\tau \right] du = \\ &= (2\pi)^{-1/2} \left[\int_{-\infty}^{\infty} f_1(u) e^{-ixu} du \right] \cdot \left[\int_{-\infty}^{\infty} f_2(\tau) e^{-ix\tau} d\tau \right] = \\ &= (2\pi)^{-1/2} \mathcal{T}[f_1] \mathcal{T}[f_2], \end{aligned}$$

as was to be proved.

* * *

We remark that from (14.3-13) we get by suitable change in notation the generalized Parseval formula (14.2-17)–(14.2-19) for functions belonging simultaneously to $L(-\infty, \infty)$ and $L^2(-\infty, \infty)$.

14.4 Fourier's integral theorem on classical convergence conditions

14.41 The formulae (14.1-9)–(14.1-10), as well as (14.1-11)–(14.1-12) which form the basis of an independent theory of Fourier integrals were deduced by FOURIER in his famous monograph [1] in a pure *formal* manner, namely by expanding $f(x)$ in ordinary Fourier series relating to an arbitrary fundamental interval $(-l, l)$ and next by passage to the limit $l \rightarrow \infty$. Then the Fourier coefficients turn into integrals taken on

the whole number line, and the summation corresponds to integration over $(-\infty, \infty)$.¹

Let us now examine the validity of Fourier's integral theorem on the classical conditions in 14.31. *Discussion of the problem can be carried out on the model of the convergence theory of ordinary Fourier series.* (Cf. Chapter 12.) In order to facilitate the formulation of results, we adopt for this whole section the convention that any Fourier integral is to be taken with its Cauchy principal value.

The pendant of Riemann's criterion 12.2.2 is:

14.4.1 For $f(t) \in L(-\infty, \infty)$, Fourier's integral theorem holds at some point x if and only if to any given $\varepsilon > 0$ we can find certain positive numbers δ and U_0 such that

$$\left| \int_0^\delta \frac{\psi_x(\tau)}{\tau} \sin(U\tau) d\tau \right| < \varepsilon \quad (14.4-1)$$

with $\psi_x(\tau) = \frac{1}{2} [f(x + \tau) + f(x - \tau)] - f(x)$, whenever $U > U_0(\varepsilon)$.

P r o o f. Since $\int_{-\infty}^{\infty} |f(t)| dt < \infty$ implies the uniform convergence of $\int_{-\infty}^{\infty} f(t)e^{-iut} dt$ for $-\infty < u < \infty$, and

$$\mathcal{F}(u) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} f(t)e^{-iut} dt$$

is continuous everywhere [cf. (14.3-2)], we can write for any $U > 0$:

$$\begin{aligned} f_U(x) &= (2\pi)^{-1/2} \int_{-U}^U \mathcal{F}(u)e^{ixu} du = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) \int_{-U}^U e^{iu(x-t)} du dt = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) \frac{e^{iU(x-t)} - e^{-iU(x-t)}}{i(x-t)} dt = \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \frac{\sin U(x-t)}{x-t} dt. \end{aligned} \quad (14.4-2)$$

¹ Fourier's original chain of ideas from 1822 (with slight modifications in terms and notations) is to be found e.g. on the first pages of Titchmarsh's standard work [2]. The way of writing (14.1-9) is due to Cauchy.

Let $\delta \in (0, 1)$ be fixed and $T = |x| + 1$; decompose the last integral as follows:

$$\int_{-\infty}^{\infty} = \int_{-\infty}^{-T} + \int_{-T}^{x-\delta} + \int_{x-\delta}^{x+\delta} + \int_{x+\delta}^T + \int_T^{\infty}. \quad (14.4-3)$$

The first and fifth subintegral on the right can trivially be estimated by $|x - t| < 1$:

$$\left| \int_{-\infty}^{-T} \right| \leq \int_{-\infty}^{-T} |f(t)| dt, \quad \left| \int_T^{\infty} \right| \leq \int_T^{\infty} |f(t)| dt,$$

and owing to the above premise, these bounds tend to 0 as $T \rightarrow \infty$. Thus after prescribing any $\varepsilon > 0$, T can be fixed in a such way that for every U simultaneously we have

$$\left| \frac{1}{\pi} \int_{-\infty}^{-T} + \frac{1}{\pi} \int_T^{\infty} \right| < \frac{\varepsilon}{3}. \quad (14.4-4)$$

The second and fourth subintegrals in (14.4-3), can be brought by substitution $x - t = v$ to the forms:

$$\int_{-T}^{x-\delta} = \int_{\delta}^{x+T} \frac{f(x-v)}{v} \sin Uv dv, \quad \int_{x+\delta}^T = \int_{x-T}^{-\delta} \frac{f(x-v)}{v} \sin Uv dv.$$

Since $f(x-v)/v$ is an (L) -integrable function of v in the intervals $(\delta, x+T)$ and $(x-T, -\delta)$, we get by the Riemann-Lebesgue lemma 12.1.3 that the integrals tend to 0 as $U \rightarrow \infty$. Accordingly,

$$\left| \frac{1}{\pi} \int_{-T}^{x-\delta} + \frac{1}{\pi} \int_{x+\delta}^T \right| < \frac{\varepsilon}{3}, \quad (14.4-5)$$

if only $U \geq U_1(\varepsilon)$.

Finally, the third subintegral in (14.4-3) becomes through suitable transformations [cf. (12.2-4)]:

$$\begin{aligned} \int_{x-\delta}^{x+\delta} &= \int_{-\delta}^{\delta} f(x-v) \frac{\sin Uv}{v} dv = \\ &= \int_0^{\delta} [f(x+\tau) + f(x-\tau)] \frac{\sin U\tau}{\tau} d\tau = \\ &= 2 \int_0^{\delta} \frac{\psi_x(\tau)}{\tau} \sin U\tau d\tau + 2f(x) \int_0^{\infty} \frac{\sin U\tau}{\tau} d\tau - 2f(x) \int_{\delta}^{\infty} \frac{\sin U\tau}{\tau} d\tau; \end{aligned}$$

and here by $\int_0^\infty t^{-1} \sin t \, dt = \pi/2$ the middle integral will be for every U :

$$\int_0^\infty \frac{\sin U\tau}{\tau} d\tau = \frac{\pi}{2},$$

further

$$\int_\delta^\infty \frac{\sin U\tau}{\tau} d\tau = \int_{U\delta}^\infty \frac{\sin t}{t} dt \rightarrow 0 \quad (U \rightarrow \infty).$$

Consequently:

$$\left| \frac{1}{\pi} \int_{x-\delta}^{x+\delta} - \left[\frac{2}{\pi} \int_0^\delta \frac{\psi_x(\tau)}{\tau} \sin U\tau \, d\tau + f(x) \right] \right| < \frac{\varepsilon}{3}, \quad (14.4-6)$$

if $U \geq U_2(\varepsilon)$.

Summing up the results (14.4-2)—(14.4-6), we obtain for U sufficiently large:

$$\left| f_U(x) - \frac{2}{\pi} \int_0^\delta \frac{\psi_x(\tau)}{\tau} \sin U\tau \, d\tau - f(x) \right| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

i.e.

$$\lim_{U \rightarrow \infty} \left[f_U(x) - \frac{2}{\pi} \int_0^\delta \frac{\psi_x(\tau)}{\tau} \sin U\tau \, d\tau \right] = f(x).$$

This means that in our case for the validity of

$$\lim_{U \rightarrow \infty} (2\pi)^{-1/2} \int_{-U}^U \mathcal{F}(u) e^{ixu} \, du = f(x) \quad (14.4-7)$$

it is necessary and sufficient that (14.4-1) should be satisfied.

* * *

Comparing this criterion with the condition in 12.2.2, we can state: the only difference is that (12.2-9) contains the "discrete" parameter $2N+1$, while in (14.4-1) the "continuous" limiting process $U \rightarrow \infty$ occurs. Clearly either Dini's or Jordan's test or other familiar tests deduced from these assure the fulfilment of condition (14.4-1).

For instance, we get the following theorem simply as a corollary (cf. 12.32):

14.4.2 *Fourier's integral theorem (14.1-9) — (14.1-10) holds and the Fourier transform (14.1-11) is invertible in the form (14.1-12) on the conditions: $f(t)$ is (L) -integrable in $(-\infty, \infty)$ and of bounded variation in some neighbourhood of the point x , the latter being a regular point of the function f .*

14.42 The analogy between Fourier series and Fourier integrals suggests that the *summation theory* of the former — studied in the previous chapter — can also be *transferred to Fourier transforms*. This may allow simplification and further development of the convergence theory of Fourier integrals in several respects.

Keeping in view the applications, the pertinent main results will be discussed here: pendants of Fejér's summation theorem 13.2.1 and its extension of the type 13.4.1. It is necessary first of all the extension of the "method of arithmetic means" $[(C, 1)$ -summation] to integrals over infinite intervals.

An integral $\int_0^\infty \varphi(u) du$ is said to be $(C, 1)$ -summable, if the integral mean

$$U^{-1} \int_0^U \left[\int_0^U \varphi(\tau) d\tau \right] du = \int_0^U \left(1 - \frac{u}{U} \right) \varphi(u) du \quad (U > 0)$$

has a (finite) limit as $U \rightarrow \infty$, called $(C, 1)$ -sum of the integral in question. In symbols briefly:

$$\begin{aligned} {}^{(C,1)} \int_0^\infty \varphi(u) du &= \lim_{U \rightarrow \infty} \frac{1}{U} \int_0^U \left[\int_0^u \varphi(\tau) d\tau \right] du = \\ &= \lim_{U \rightarrow \infty} \int_0^U \left(1 - \frac{u}{U} \right) \varphi(u) du. \end{aligned} \quad (14.4-8)$$

Evidently, the expression (14.4-8) is a full analogue (and also generalization in a sense) of the notion " $(C, 1)$ -sum of series" which was defined in 13.11 as follows:

$$\begin{aligned} {}^{(C,1)} \sum_{n=1}^\infty a_n &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \sum_{k=1}^n a_k = \\ &= \lim_{N \rightarrow \infty} \sum_{n=1}^N \left(1 - \frac{n}{N} \right) a_n. \end{aligned} \quad (14.4-9)$$

It is noticeable that in (14.4-8) the role of the "discrete" variable n is taken over by the "continuous" variable u , the role of the sequence a_n ($n = 1, 2, \dots$) by the function $\varphi(u)$ ($0 < u < \infty$); N is replaced by U and summation by integration.

It is easy to see that the summation method (14.4-8) satisfies the fundamental requirements mentioned in 13.12 (valid for arbitrary processes of summation), provided that the convergence of the integral is meant in

Cauchy's manner, i.e., according to the formula $\int_0^\infty \varphi(u) du = \lim_{U \rightarrow \infty} \int_0^U \varphi(u) du$.

A trivial example of non-convergent (divergent) but $(C, 1)$ -summable integral

is: $\int_0^\infty \cos u du$. The limit of $\int_0^U \cos u du = \sin U$ does not exist for $U \rightarrow \infty$, but we have

$$\begin{aligned} {}^{(C,1)} \int_0^\infty \cos u du &= \lim_{U \rightarrow \infty} \int_0^U \left(1 - \frac{u}{U}\right) \cos u du = \\ &= \lim_{U \rightarrow \infty} \frac{1 - \cos U}{U} = 0. \end{aligned} \quad (14.4-10)$$

If in the last term in (14.4-8) the factor $1 - \frac{u}{U}$ is replaced by $\left(1 - \frac{u}{U}\right)^\alpha$ ($\alpha \geq 0$), the usual definition of (C, α) -summation for integrals is obtained:

$${}^{(C,\alpha)} \int_0^\infty \varphi(u) du = \lim_{U \rightarrow \infty} \int_0^U \left(1 - \frac{u}{U}\right)^\alpha \varphi(u) du. \quad (14.4-11)$$

The process (14.4-11) introduced by G. H. HARDY towards the beginning of the century¹, may be considered as a common extension of the classical convergence ($\alpha = 0$) and of the simple "integral mean method" ($\alpha = 1$).

For the family of these (C, α) -methods with $0 \leq \alpha \leq \infty$, all the relevant properties of Cesàro methods of series maintain their validity. (Cf. 13.13.)

14.43 Theorem 14.4.1 can be transferred to the case of $(C, 1)$ -summation in a slightly extended form as follows:

14.4.3 For any $f(t) \in L(-\infty, \infty)$, the Fourier integrals (14.1-9) or (14.1-10) are $(C, 1)$ -summable at a point x to some prescribed value σ , if and only if to any given $\varepsilon > 0$ there exists a number $\eta > 0$ such that for U large enough the inequality

$$\left| \int_0^\eta [\varphi_x(2t) - \sigma] \cdot \frac{1}{U} \left[\frac{\sin(Ut)}{t} \right]^2 dt \right| < \varepsilon \quad (14.4-12)$$

¹ Cf. e.g. Hardy [7].

with

$$\varphi_x(\tau) = \frac{1}{2} [f(x + \tau) + f(x - \tau)]$$

holds.

P r o o f. Arguing like the verification of 14.4.1, we invert first the order of integrations with respect to t and u , then apply the formula obtained from integration by parts:

$$\int_0^U \left(1 - \frac{u}{U}\right) \cos u(x - t) du = \frac{1}{2U} \left[\frac{\sin \left(\frac{U(x-t)}{2} \right)}{\frac{x-t}{2}} \right]^2.$$

Hence we see that the $(C, 1)$ -sum of (14.1-10) is the limit of the integral

$$J_U[f; x] = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \cdot \frac{1}{2U} \left[\frac{\sin \left(\frac{U(x-t)}{2} \right)}{\frac{x-t}{2}} \right]^2 dt \quad (14.4-13)$$

as $U \rightarrow \infty$.

If now $J_U[f; x]$ will be decomposed according to the model of (14.4-3) and the occurring subintegrals are estimated in appropriate manner, we get for some fixed $\delta > 0$:

$$\lim_{U \rightarrow \infty} \left\{ J_U[f; x] - \frac{1}{\pi} \int_0^{\delta} [\varphi_x(\tau) - \sigma] \cdot \frac{1}{U} \left[\frac{\sin(U\tau/2)}{\tau/2} \right]^2 d\tau \right\} = \sigma. \quad (14.4-14)$$

This limit relation involves the assertion immediately.

* * *

Using the criterion 14.4.3, we can deduce substantially on the analogy of 13.42 the most important result of the summation theory of Fourier transformation:

14.4.4 (FEJÉR-LEBESGUE THEOREM FOR FOURIER INTEGRALS) *If a function $f \in L(-\infty, \infty)$ satisfies at some point x the condition (13.4-1), then the Fourier integrals (14.1-9) or (14.1-10) are at x $(C, 1)$ -summable with the $(C, 1)$ -sum σ .*

In particular, the last statement holds for almost all x , if $\sigma = f(x)$; furthermore the Fourier integrals in question are $(C, 1)$ -summable to $f(x)$ and $\frac{1}{2}[f(x+0) + f(x-0)]$, depending on whether x is a point of continuity or discontinuity of the first kind of f , respectively.

P r o o f. 1° We know from 13.42 that the condition (13.4-1), i.e.,
 $(2l)^{-1} \int_{x-l}^{x+l} |f(t) - \sigma| dt \rightarrow 0 \quad (l \rightarrow +0)$ implies

$$\lim_{h \rightarrow +0} h^{-1} \int_0^h |\varphi_x(2t) - \sigma| dt = 0 \quad (14.4-15)$$

with $\varphi_x(\tau) = \frac{1}{2} [f(x+\tau) + f(x-\tau)]$. In order to verify the first part of the assertion, we have only to show that (14.4-12) is a consequence of the limit relation (14.4-15).

Suppose that some function $f \in L(-\infty, \infty)$ fulfils at a point x the condition (14.4-15) with a given value σ , and try to connect the latter with the integral (14.4-12). Applying the elementary inequality $|\sin u| \leq 2u/(1+u)$ ($u \geq 0$) and the notations $\Psi_x(t) = \int_0^t |\varphi_x(2\tau) - \sigma| d\tau$, $\Psi_x^*(t) = \Psi_x(t)/t$, integration by parts yields:

$$\begin{aligned} & \left| \int_0^\eta [\varphi_x(2t) - \sigma] \cdot \frac{1}{U} \left[\frac{\sin(Ut)}{t} \right]^2 dt \right| \leq \\ & \leq 4 \int_0^\eta |\varphi_x(2t) - \sigma| \cdot \frac{U}{(1+Ut)^2} dt = \\ & = 4\Psi_x(\eta) \frac{U}{(1+U\eta)^2} + 8 \int_0^\eta \Psi_x^*(t) \frac{U^2 t}{(1+Ut)^3} dt < \\ & < \frac{4}{\eta^2 U} \int_0^\eta |\varphi_x(2t) - \sigma| dt + 8 \int_0^\eta \Psi_x^*(t) \frac{U}{(1+Ut)^2} dt. \quad (14.4-16) \end{aligned}$$

Since by our assumption $\lim_{h \rightarrow +0} \Psi^*(h) = 0$, to an arbitrarily prescribed number $\varepsilon > 0$ another number $\eta = \eta_\varepsilon > 0$ can be attached such that $\Psi_x^*(t) < \varepsilon/16$ ($0 < t \leq \eta$). Then the last term in formula (14.4-16) satisfies the inequality:

$$8 \int_0^\eta \Psi_x^*(t) \frac{U}{(1+Ut)^2} dt < \frac{\varepsilon}{2} \left[\frac{1}{1+Ut} \right]_{t=0}^{t=\eta} < \frac{\varepsilon}{2} \quad (14.4-17)$$

for every $U > 0$. On the other hand, η being fixed in the just-mentioned manner, U can be chosen as large that

$$\frac{4}{\eta^2 U} \int_0^\eta |\varphi_x(2t) - \sigma| dt < \frac{\varepsilon}{2}. \quad (14.4-18)$$

holds.

Collecting the results (14.4-16)—(14.4-18), we obtain that in our case the first term of (14.4-12) is less than $\frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, if η is fixed sufficiently small and $U \geq U_0(\varepsilon)$. This is just what we had first to prove.

2° As far as the remaining part of the assertion is concerned, we have only to consider some findings of 13.41–13.42. Namely, it has been established that in case of any function $f \in L(-\infty, \infty)$ the condition (13.4-1) with $\sigma = f(x)$ is satisfied for almost all x and specially at any point of continuity x of f , furthermore with $\sigma = \frac{1}{2}[f(x+0) + f(x-0)]$ the same is true at any discontinuity of the first kind of the function.

* * *

Concerning the background of criteria 14.4.1 and 14.4.3 we want to stress: the proof of the former is based on the integral of “Dirichlet’s type”

$$\Delta(x, U) = \frac{1}{\pi} \int_{x-\delta}^{x+\delta} f(t) \frac{\sin U(x-t)}{x-t} dt, \quad (14.4-19)$$

while the second depends decisively on the behaviour of

$$\begin{aligned} \Phi(x, U) &= \frac{1}{U} \int_0^U \Delta(x, u) du = \\ &= \frac{1}{\pi} \int_{x-\delta}^{x+\delta} f(t) \cdot \frac{1}{2U} \left[\frac{\sin \left(U \frac{x-t}{2} \right)}{\frac{x-t}{2}} \right]^2 dt, \end{aligned} \quad (14.4-20)$$

i.e., of an integral of “Fejér’s type” for $U \rightarrow +\infty$. (Cf. sections 12.2 and 13.2.) Note still that 14.4.4 allows transferring a remarkable property of Fourier series to Fourier integrals: it may converge at points of continuity only to the corresponding value of the function, and at points of discontinuity of the first kind only to the arithmetic mean of the pertaining unilateral limits. (See 13.31.)

As a generalization possibility, we mention that the summation theorem 14.4.4 — like 13.4.1 in the theory of ordinary Fourier series — is valid also for every method (C, α) ($\alpha > 0$) in the (14.4-11) sense.¹

¹ Cf. e.g. Titchmarsh [2].

14.5 Non-absolutely convergent Fourier integrals

14.51 In Sections 14.3–14.4 the absolute convergence of any Fourier transform has been assumed, because this condition guarantees most conveniently the existence and numerous useful properties of $\mathcal{T}[f]$. We have to observe, however, that a lot of *non-absolutely convergent Fourier transforms* often occurs on diverse application fields, as e.g. integrals of the type $\int_0^\infty t^\vartheta-1 e^{-i\lambda t} dt$ ($0 < \vartheta < 1$, $\lambda \geq 0$) which have taken a prominent part in Sections 12.7 and 13.6. Here a survey will be given on such recent results of the theory, which aim at further development in the above direction and have been found by the author during the last decades. (See Mikolás [13] and [15].)

Throughout the present section, only the “minimal” restriction will be made on f , namely that it should be *(L)-integrable in every finite interval*. A sequence or series of functions will be called *dominatedly* (majorizedly) *convergent* in an interval (a, b) , if 1) its limit or sum function exists in (a, b) almost everywhere; 2) its terms or partial sums have in (a, b) a common *(L)-integrable majorant*.¹

It is advisable to study here the validity of Fourier’s integral theorem with respect to the fundamental system $e^{2\nu\pi ix}$ ($\nu = 0, \pm 1, \pm 2, \dots$) of 1-period functions, further in a slightly modified form. We look for the exact connection between the double limit taken at a fixed point x :

$$\lim_{M, N \rightarrow \infty} \int_{-(M+\frac{1}{2})}^{M+\frac{1}{2}} e^{2\pi i x u} \left[\int_{\xi-(N+\frac{1}{2})}^{\xi+(N+\frac{1}{2})} f(t) e^{-2\pi i u t} dt \right] du, \quad (14.5-1)$$

and the sum (defined as Cauchy’s principal value) of the corresponding ordinary Fourier series of f relating to $\xi - \frac{1}{2} < t < \xi + \frac{1}{2}$:

$$\lim_{M \rightarrow \infty} \sum_{\nu=-M}^M e^{2\nu\pi ix} \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} f(t) e^{-2\nu\pi it} dt; \quad (14.5-2)$$

in these formulae, M and N run over all the non-negative integers, further ξ denotes such a real number for which $x \in (\xi - \frac{1}{2}, \xi + \frac{1}{2})$.

We shall see: the existence of the limit (14.5-1) always implies that of (14.5-2) and the equality of both limits, but the converse conclusion is right only on certain restrictions about f .

¹ According to Lebesgue’s theorem 7.5.1, these conditions involve termwise integrability. It is obvious that *any sequence or series which converges uniformly in a finite segment (a, b) to an *(L)-integrable function, is in (a, b) also dominatedly convergent.**

14.52 The following peculiar theorem is valid:

14.5.1 (EQUICONVERGENCE THEOREM FOR FOURIER'S DOUBLE INTEGRAL)¹ In order that the double limit (14.5-1) should exist and be equal to a number f_x , it is necessary that the Fourier series (14.5-2) should converge to f_x .

The latter condition is also sufficient, if the function f is chosen in such a way that the series

$$\sum_{n=1}^{\infty} (-1)^{n-1} \left[\frac{f(\tau+n)}{\tau-x+n} + \frac{f(\tau-x)}{\tau-x-n} \right] \quad \left[x \in \left(\xi - \frac{1}{2}, \xi + \frac{1}{2} \right) \right] \quad (14.5-3)$$

converges dominantly for $\xi - \frac{1}{2} < \tau < \xi + \frac{1}{2}$; and this is assured, whenever f is bounded, (L) -integrable in $(\xi - \frac{1}{2}, \xi + \frac{1}{2})$ and 1-periodic.

P r o o f. Let us start with a kind of conversion of the inner integral in (14.5-1):

$$\begin{aligned} F_N(u) &= \int_{\xi-(N+\frac{1}{2})}^{\xi+(N+\frac{1}{2})} f(t) e^{-2\pi i u t} dt = \\ &= \sum_{n=-N}^N \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} f(\tau+n) e^{-2\pi i u (\tau+n)} d\tau = \\ &= \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} e^{-2\pi i u \tau} \varphi_N(\tau, -u) d\tau, \end{aligned} \quad (14.5-4)$$

where

$$\varphi_N(\tau, \omega) = \sum_{n=-N}^N f(\tau+n) e^{2\pi i \omega n} \quad (N = 0, 1, 2, \dots). \quad (14.5-5)$$

Since $F_N(u)$ is (L) -integrable in each finite interval, on the basis of Fubini's theorem 7.7.1 we can write:

¹ See Mikolás [13], 157-161.

$$\begin{aligned}
& \int_{-(M+\frac{1}{2})}^{M+\frac{1}{2}} e^{2\pi i x u} F_N(u) du = \\
&= \sum_{\nu=-M}^M \int_{\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} e^{2\pi i x u} \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} e^{-2\pi i u \tau} \varphi_N(\tau, -u) d\tau du = \\
&= \sum_{\nu=-M}^M \int_{-1/2}^{1/2} e^{2\pi i x(\nu-\omega)} \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} e^{2\pi i(\omega-\nu)\tau} \varphi_N(\tau, \omega) d\tau d\omega = \\
&= \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} g_{N,x}(\tau) \sum_{\nu=-M}^M e^{2\nu\pi i(x-\tau)} d\tau, \quad (14.5-6)
\end{aligned}$$

introducing for brevity the notation

$$g_{N,x}(\tau) = \int_{-1/2}^{1/2} \varphi_N(\tau, \omega) e^{2\pi i \omega(\tau-x)} d\omega.$$

But this latter integral can be calculated at once:

$$\begin{aligned}
g_{N,x}(\tau) &= \sum_{n=-N}^N f(\tau+n) \int_{-1/2}^{1/2} e^{2\pi i \omega(\tau-x+n)} d\omega = \\
&= \begin{cases} f(x) & \text{for } \tau = x, \\ \frac{\sin \pi(\tau-x)}{\pi} \sum_{n=-N}^N (-1)^n \frac{f(\tau+n)}{\tau-x+n}, & \text{if } \tau \neq x \text{ and } \tau \in (\xi - \frac{1}{2}, \xi + \frac{1}{2}). \end{cases} \quad (14.5-7)
\end{aligned}$$

Thus we get

$$\begin{aligned}
& \frac{\pi g_{N,x}(\tau)}{\sin \pi(\tau-x)} = \frac{f(\tau)}{\tau-x} + \\
&+ \sum_{n=1}^N (-1)^n \left[\frac{f(\tau+n)}{\tau-x+n} + \frac{f(\tau-n)}{\tau-x-n} \right] \quad (\tau-x \text{ non-integer}); \quad (14.5-8)
\end{aligned}$$

and writing

$$G_{N,x}(\tau) = \sum_{n=1}^N (-1)^{n-1} \left[\frac{f(\tau+n)}{\tau-x+n} + \frac{f(\tau-n)}{\tau-x-n} \right], \quad (14.5-9)$$

it follows by formula (12.2-1):

$$\begin{aligned}
 & \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} g_{N,x}(\tau) \sum_{\nu=-M}^M e^{2\nu\pi i(x-\tau)} d\tau = \\
 & = \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} g_{N,x}(\tau) \frac{\sin(2M+1)\pi(x-\tau)}{\sin\pi(x-\tau)} d\tau = \\
 & = \frac{1}{\pi} \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} \frac{f(\tau)}{x-\tau} \sin(2M+1)\pi(x-\tau) d\tau + \\
 & + \frac{1}{\pi} \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} G_{N,x}(\tau) \sin(2M+1)\pi(x-\tau) d\tau. \quad (14.5-10)
 \end{aligned}$$

Accordingly, the difference of the double integral (14.5-1) and the sum (14.5-2) can be brought to the form [cf. (14.5-6)]:

$$\begin{aligned}
 & \int_{-(M+\frac{1}{2})}^{M+\frac{1}{2}} e^{2\pi i x u} F_N(u) du - \sum_{\nu=-M}^M e^{2\nu\pi i x} \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} f(\tau) e^{-2\nu\pi i \tau} d\tau = \\
 & = \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} f(\tau) \left[\frac{1}{\pi(x-\tau)} - \frac{1}{\sin\pi(x-\tau)} \right] \sin(2M+1)\pi(x-\tau) d\tau + \\
 & + \frac{1}{\pi} \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} G_{N,x}(\tau) \sin(2M+1)\pi(x-\tau) d\tau; \quad (14.5-11)
 \end{aligned}$$

and in the last two integrands the factors before $\sin[(2M+1)\pi(x-\tau)]$ are (L) -integrable functions of τ , so that we may apply to both of them the Riemann-Lebesgue lemma as M tends to infinity. (See 12.14.)

1° (*Necessity.*) Suppose now that (14.5-1) exists and equals f_x . Then after giving $\varepsilon > 0$ (however small), choose M and N as large that

$$\left| \int_{-(M+\frac{1}{2})}^{M+\frac{1}{2}} e^{2\pi i x u} F_N(u) du - f_x \right| < \frac{\varepsilon}{2}.$$

The integer N being fixed in such a way, by further increase of M we can also realize owing to 12.1.3 the inequality:

$$\left| \sum_{\nu=-M}^M e^{2\nu\pi ix} \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} f(t) e^{-2\nu\pi it} dt - \int_{-(M+\frac{1}{2})}^{M+\frac{1}{2}} e^{2\pi i x u} F_N(u) du \right| < \frac{\varepsilon}{2}.$$

Addition yields:

$$\left| \sum_{\nu=-M}^M e^{2\nu\pi ix} \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} f(t) e^{-2\nu\pi it} dt - f_x \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

if only $M \geq M_0(\varepsilon)$; i.e., the limit (14.5-2) exists and equals f_x as well.

2° (*Sufficiency*.) To verify the other part of the assertion, assume that the Fourier series (14.5-2) converges to f_x (in Cauchy's sense), and that at the same time (14.5-3) is convergent dominatedly for $\tau \in (\xi - \frac{1}{2}, \xi + \frac{1}{2})$.

In this case, the latter series of functions and its product by the bounded function $\sin[(2M+1)\pi(x-t)]$ may be integrated term-by-term in $(\xi - \frac{1}{2}, \xi + \frac{1}{2})$ on the basis of 7.5.1. More precisely, the sum of (14.5-3), i.e., [cf. (14.5-9)]

$$\begin{aligned} G_x(\tau) &= \lim_{N \rightarrow \infty} G_{N,x}(\tau) = \\ &= \sum_{n=1}^{\infty} (-1)^{n-1} \left[\frac{f(\tau+n)}{\tau-x+n} + \frac{f(\tau-n)}{\tau-x-n} \right] \end{aligned} \quad (14.5-12)$$

is (L) -integrable for $\xi - \frac{1}{2} < \tau < \xi + \frac{1}{2}$, furthermore for any fixed non-negative integer M we have the limit relation:

$$\begin{aligned} \lim_{N \rightarrow \infty} \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} G_{N,x}(\tau) \sin(2M+1)\pi(x-\tau) d\tau &= \\ &= \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} G_x(\tau) \sin(2M+1)\pi(x-\tau) d\tau. \end{aligned}$$

Thus in (14.5-11) doing first the limiting process $N \rightarrow \infty$, after carrying out $M \rightarrow \infty$, and considering that the left-hand sum tends to f_x by hypothesis while the right-hand integrals $\rightarrow \infty$ by the Riemann-Lebesgue lemma, we obtain indeed

$$\lim_{M, N \rightarrow \infty} \int_{-(M+\frac{1}{2})}^{M+\frac{1}{2}} e^{2\pi i x u} F_N(u) du = f_x.$$

3° The final statement about the auxiliary series (14.5-3) can be proved quite simply.

(12.7-24) yields in the case $x = 0$ another important decomposition into partial fractions:

$$\frac{\pi}{\sin \pi v} = \frac{1}{v} + \sum_{n=1}^{\infty} (-1)^n \left(\frac{1}{v+n} + \frac{1}{v-n} \right) \quad (14.5-13)$$

which holds uniformly on any finite segment of the number line without integers; so the condition $f(\tau + \nu) = f(\tau)$ ($\tau = 0, \pm 1, \dots$) implies that the limit relation

$$\lim_{N \rightarrow \infty} G_{N,x}(\tau) = f(\tau) \left[\frac{1}{\tau - x} - \frac{\pi}{\sin \pi(\tau - x)} \right] \quad (14.5-14)$$

is valid uniformly in the interval $\xi - \frac{1}{2} < \tau < \xi + \frac{1}{2}$ as well. [(Cf. (14.5-9).]

Since the right-hand term of (15.4-14) is by our assumption (L) -integrable in the last-mentioned interval, we see that the uniform existence of the limit of $G_{N,x}(\tau)$ for $N \rightarrow \infty$ assures also the dominated convergence of this function sequence in $(\xi - \frac{1}{2}, \xi + \frac{1}{2})$.

* * *

14.53 The theorem just verified furnishes plainly the widest possible information on the connection of (14.5-1) and (14.5-2) in a sense, because the terms of the Fourier series (14.5-2) are functions of x with *period* 1, consequently a complete equiconvergence with the limit (14.5-1) can be realized merely on suitable restrictions concerning f . It turns out at the same time: the significance 14.5.1 is that by means of this criterion the existence problem of the double limit (14.5-1) can be reduced to the study of the Fourier series (14.5-2) and the function series (14.5-3).

For instance, we get in such manner the following result: if there is a positive function $\mu(\tau) \in L(\xi - \frac{1}{2}, \xi + \frac{1}{2})$ such that for every N the inequality

$$\left| \sum_{n=-N}^N (-1)^n f(\tau + n) \right| \leq \mu(\tau) \quad \left(\xi - \frac{1}{2} < \tau < \xi + \frac{1}{2} \right) \quad (14.5-15)$$

holds, then the limit (14.5-1) exists at every convergence point x of the Fourier series (14.5-2) and equals the corresponding sum of the latter.

Now we want to show that a "natural" extension of the condition (14.5-15) also assures the existence of $F(u) = \lim_{N \rightarrow \infty} F_N(u)$ [cf. (14.5-4)], thus giving a *further criterion* — a considerably sharper one — to the *invertibility of Fourier transformation*.¹

¹ Cf. Mikolás [13], 162-164.

14.5.2 If the symmetric partial sums of the series

$$\sum_{n=-\infty}^{\infty} f(\tau + n)e^{2n\pi i\omega} \quad (14.5-16)$$

converge dominatedly for every fixed ω in some interval $\xi - \frac{1}{2} < \tau < \xi + \frac{1}{2}$, then the Fourier transform

$$F(u) = \lim_{N \rightarrow \infty} \int_{\xi - (N + \frac{1}{2})}^{\xi + (N + \frac{1}{2})} f(t)e^{-2\pi i u t} dt \quad (14.5-17)$$

exists and is bounded in $-\infty < u < \infty$, furthermore is (L) -integrable on any finite segment.

On the above condition, the inversion formula

$$f(x) = \lim_{M \rightarrow \infty} \int_{-(M + \frac{1}{2})}^{M + \frac{1}{2}} F(u)e^{2\pi i x u} du \quad (14.5-18)$$

holds if and only if the ordinary Fourier series of f pertaining the fundamental interval $(\xi - \frac{1}{2}, \xi + \frac{1}{2})^1$ represents the corresponding value of the function at x .

P r o o f. 1° The stipulation on the series (14.5-16) involves the dominated convergence of the sequence $e^{-2\pi i u \tau} \varphi_N(\tau, -u)$ ($N = 0, 1, 2, \dots$) for any fixed u in the interval $\xi - \frac{1}{2} < \tau < \xi + \frac{1}{2}$ [cf. (14.5-5)]; consequently it implies by 7.5.1 also the existence of $F(u)$ as well as the representation

$$F(u) = \int_{\xi - \frac{1}{2}}^{\xi + \frac{1}{2}} \varphi(\tau, -u)e^{-2\pi i u \tau} d\tau \quad (14.5-19)$$

where

$$\varphi(\tau, \omega) = \lim_{N \rightarrow \infty} \varphi_N(\tau, \omega) = \sum_{n=-\infty}^{\infty} f(\tau + n)e^{2n\pi i \omega}. \quad (14.5-20)$$

According to our premise, we can find a positive function $\mu(\tau) \in L(\xi - \frac{1}{2}, \xi + \frac{1}{2})$ such that for arbitrary real values of ω

$$|\varphi_N(\tau, \omega)| \leq \mu(\tau) \quad \left(N = 0, 1, 2, \dots; \xi - \frac{1}{2} < \tau < \xi + \frac{1}{2} \right) \quad (14.5-21)$$

¹ Hereby we mean of course the trigonometric Fourier series of the 1-periodic function, which is identical with f in $(\xi - \frac{1}{2}, \xi + \frac{1}{2})$.

and therefore

$$|\varphi(\tau, \omega)| \leq \mu(\tau) \left(\xi - \frac{1}{2} < \tau < \xi + \frac{1}{2} \right).$$

Thus on the basis of (14.5-4) and (14.5-19) every element of the sequence $\{F_N(u)\}$, further also its limit $F(u)$ can be majorized in absolute value for $-\infty < u < \infty$ by the integral $\int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} \mu(\tau) d\tau$. But from the “uniform” boundedness of the functions $F_N(u)$ ($N = 0, 1, \dots$) and from their (L) -integrability in every finite interval, we can infer the latter property of $F(u)$.

2° Using Lebesgue’s limit theorem 7.5.1 in case of the sequence $\{F_N(u)e^{2\pi i x u}\}$ and of the interval $-M - \frac{1}{2} < u < M + \frac{1}{2}$ we obtain:

$$\begin{aligned} \lim_{N \rightarrow \infty} \int_{-(M+\frac{1}{2})}^{M+\frac{1}{2}} F_N(u) e^{2\pi i x u} du &= \\ &= \int_{-(M+\frac{1}{2})}^{M+\frac{1}{2}} F(u) e^{2\pi i x u} du. \end{aligned} \quad (14.5-22)$$

Besides this formula we have also the representation (14.5-11), owing to the fact that the dominated convergence of the series (14.5-16) in $(\xi - \frac{1}{2}, \xi + \frac{1}{2})$ assures by the above the same property of (14.5-3).

If we let first N , after M tend to infinity in (14.5-11), then it follows that the difference

$$\int_{-(M+\frac{1}{2})}^{M+\frac{1}{2}} F(u) e^{2\pi i x u} du - \sum_{\nu=-M}^M e^{2\nu\pi i x} \int_{\xi-\frac{1}{2}}^{\xi+\frac{1}{2}} f(t) e^{-2\nu\pi i t} dt$$

tends to 0 as $N \rightarrow \infty$.

This means the equiconvergence of the limit (14.5-18) to the Fourier series (14.5-2), i.e., the second part of our assertion is also true.

* * *

14.54 The premise of 14.5.2 can naturally be replaced by some other — perhaps not so “fine” but anyway simpler — sufficient conditions.

For instance, the application of certain estimates about cosine and sine polynomials of monotone decreasing coefficients leads to the corollary:

14.5.3 The statements contained in proposition 14.5.2 preserve their validity, if instead of the dominated convergence of the series (14.5-16) we require that $f(\xi \pm T)$ with $T > 0$ large enough should tend decreasingly to 0 as $T \rightarrow \infty$.

It is still noteworthy that the condition

$$\int_{-\infty}^{\infty} |f(t)| dt < \infty \quad (14.5-23)$$

involves immediately by B. Levi's theorem 7.5.3 the absolute convergence almost everywhere of the series (14.5-16) and simultaneously its dominated convergence too. Hereby it is clarified unambiguously that the premise of 14.5.2 is weaker than the classical specification (14.5-23).

Lastly we can deduce from the above an important expedient for asymptotic investigations, if we observe that the series (14.5-16) becomes in the case $\tau = \omega = 0$ the first term of Poisson's summation formula:

$$\sum_{n=-\infty}^{\infty} f(n) = \lim_{M \rightarrow \infty} \sum_{\nu=-M}^M \int_{-\infty}^{\infty} f(t) e^{-2\nu\pi it} dt. \quad (14.5-24)$$

This is utilized in the literature, as well-known, mostly with the requirements: f should satisfy (14.5-23), be of bounded variation in $(-\infty, \infty)$; furthermore any point of the t -axis should be a regular point of f in the sense given in 12.32.¹

In order to extend (14.5-24) in a suitable way, let us consider, for some fixed real ω , the 1-periodic function $P_{\omega}(\tau)$ which is identical with $e^{2\pi i \tau \omega} \varphi(\tau, \omega)$ in the interval $\xi - \frac{1}{2} \leq \tau < \xi + \frac{1}{2}$. [Cf. (14.5-20).] The complex Fourier coefficients of this function pertaining to $[\xi - \frac{1}{2}, \xi + \frac{1}{2})$ are [cf. (14.5-19)]:

$$\int_{\xi - \frac{1}{2}}^{\xi + \frac{1}{2}} P_{\omega}(t) e^{-2\nu\pi it} dt = F(\nu - \omega) \quad (\nu = 0, \pm 1, \dots); \quad (14.5-25)$$

thus we have at every point $\tau \in [\xi - \frac{1}{2}, \xi + \frac{1}{2})$ where $P_{\omega}(\tau)$ is Fourier expansible (cf. 12.31) the connection

¹ Cf. e.g. Zygmund [5], Vol. I, p. 68. The mentioned conditions assure the absolute convergence not only for the Fourier transform in the second term of (14.5-24), but also for the series $\sum_{n=-\infty}^{\infty} f(n)$.

$$\varphi(\tau, \omega) e^{2\pi i \tau \omega} = \sum_{\nu=-\infty}^{\infty} F(\nu - \omega) e^{2\nu \pi i \tau},$$

i.e., in a more detailed form:

$$\begin{aligned} \lim_{N \rightarrow \infty} \sum_{n=-N}^N f(n + \tau) e^{2n \pi i \omega} &= \\ &= \lim_{M \rightarrow \infty} \sum_{\nu=-M}^M F(\nu - \omega) e^{2\pi i \tau (\nu - \omega)}. \end{aligned} \quad (14.5-26)$$

It can at once be seen that (14.5-26) contains indeed (14.5-24) as a special case ($\tau = \omega = 0$), so the former will be called *generalized Poisson sum formula*. The relevance of our result is given by the fact that the coefficients on the left and on the right are formed from the function f and from its Fourier transform F in a very similar manner. If for brevity we introduce the notations

$$\varphi(x, y) = \sum_{n=-\infty}^{\infty} f(n + x) e^{2n \pi i y}, \quad (14.5-27)$$

$$\Phi(x, y) = \sum_{n=-\infty}^{\infty} F(n + x) e^{2n \pi i y} \quad (14.5-28)$$

and adopt the convention that we mean by the sums of the right-hand series the limits of the corresponding symmetric partial sums, then the formula may be written after multiplication by $e^{\pi i \tau \omega}$ as a simple *functional equation* connecting φ and Φ :

$$\varphi(\tau, \omega) e^{\pi i \tau \omega} = \Phi(-\omega, \tau) e^{\pi i (-\omega) \tau}. \quad (14.5-29)$$

On the basis of the results just mentioned, in particular by putting $f(t) = |t|^{-z}$ ($0 < \operatorname{Re} z < 1$) the author succeeded not long ago in finding some new relations of higher transcendental functions.¹

14.6 Fourier-Stieltjes transforms

14.61 To close the present chapter, we want to deal in brief with the pendant of Fourier transformation written as an (LS) -integral, the so-called *Fourier-Stieltjes transformation*.

¹ Cf. Mikolás [10].

The (complex) *Fourier-Stieltjes transform* of a function of bounded variation $w(t)$ is defined and denoted by

$$\begin{aligned}\mathcal{T}[dw] &= \mathcal{T}[dw; x] = \\ &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-ixt} dw(t) \quad (-\infty < x < \infty).\end{aligned}\quad (14.6-1)$$

The last term can be expressed as a difference:

$$\int_{-\infty}^{\infty} e^{-ixt} dw_1(t) - \int_{-\infty}^{\infty} e^{-ixt} dw_2(t),$$

$w_1(t)$ and $w_2(t)$ denoting monotone increasing, bounded and everywhere left-hand continuous functions for which $w(t) = w_1(t) - w_2(t)$ (cf. 4.5.1 and 8.3); further we have:

$$\begin{aligned}\int_{-\infty}^{\infty} e^{-ixt} dw_{1,2}(t) &= \\ &= \int_{-\infty}^{\infty} \cos(xt) dw_{1,2}(t) - i \int_{-\infty}^{\infty} \sin(xt) dw_{1,2}(t).\end{aligned}$$

$\mathcal{T}[dw; x]$ as an (LS)-integral exists obviously for every x and is bounded; namely the inequalities $\lim_{t \rightarrow \infty} w_{1,2} \leq B < \infty$, $\lim_{t \rightarrow \infty} w_{1,2}(t) \geq A > -\infty$ imply

$$\left| \int_{-\infty}^{\infty} e^{-ixt} dw_{1,2}(t) \right| \leq \int_{-\infty}^{\infty} dw_{1,2}(t) \leq B - A \quad (-\infty < x < \infty).$$

It is also easy to see that $\mathcal{T}[dw; x]$ is uniformly continuous in $(-\infty, \infty)$, and especially by 8.3.3 it holds

$$\mathcal{T}[dw; x] = \mathcal{T}[f(t); x] \quad (14.6-2)$$

for any absolutely continuous function $w(t)$ satisfying $w'(t) = f(x)$ almost everywhere. [The latter condition means in other words that $w(t)$ is the (L)-integral function of a suitable $f(t) \in L(-\infty, \infty)$; cf. 7.6.7.]

As regards the inversion of (14.6-1), Fubini's theorem 8.5.3 yields the representation

$$\begin{aligned}w(b) - w(a) &= \\ &= (2\pi)^{-1/2} \lim_{T \rightarrow \infty} \int_{-T}^T \frac{e^{ibx} - e^{iax}}{ix} \mathcal{T}[dw; x] dx,\end{aligned}\quad (14.6-3)$$

by means of which $w(t)$ can be determined from $\mathcal{T}[dw; x]$ essentially uniquely.

In view of the general role of Lebesgue–Stieltjes integrals in modern analysis (cf. 8.35), it is clear that *the whole convergence and summation theory of ordinary Fourier series can formally be incorporated in the treatment of Fourier–Stieltjes transformation*, i.e., the properties of Fourier series can also be deduced from those of Fourier–Stieltjes transforms. This way is quasi the reverse of the heuristic chain of ideas which leads from ordinary Fourier series through formal limiting processes to the Fourier transformation (cf. 14.41), and it reflects both from the theoretical and practical side more properly the connection between Fourier series and Fourier integrals.

14.62 Among the *application possibilities*, let us stress mainly the use of Fourier–Stieltjes transforms in *probability theory*, especially the connection between the concept of characteristic functions (pertaining to a random variable) and of probability distribution functions, as well as the problems of the so-called *central limit theorem*.¹

If we choose as $w(t)$ in (14.6–1) the *distribution function* $F(t) = P(\xi < t)$ of a random variable ξ (cf. 8.35), it yields essentially (apart from a constant factor) just the definition of the characteristic function of ξ . More precisely: we mean by the *characteristic function of ξ* the Fourier–Stieltjes transform

$$\varphi(x) = \varphi_\xi(x) = \int_{-\infty}^{\infty} e^{-ixt} dF(t) \quad (-\infty < x < \infty)^2 \quad (14.6-4)$$

This quantity can also be interpreted as the *expected value* of $e^{-ix\xi}$, and in the case of a continuous ξ it may be replaced by the (L) -integral $\int_{-\infty}^{\infty} e^{-ixt} f(t) dt$

where $f(t)$ denotes the *probability density* of ξ .

(14.6–4) shows immediately that $\varphi(x)$ is continuous everywhere, furthermore $|\varphi(x)| \leq 1$ and $\varphi(0) = 1$.

By the foregoing, not only $\varphi(x)$ is determined by $F(t)$ but also F can be calculated uniquely at any point of continuity if φ is known. This interdependence may further be deepened by the following limit theorem which we insert here for orientation:

14.6.1 Let $F_1, F_2, \dots$ be a sequence of probability distribution functions and let $\varphi_1, \varphi_2, \dots$ denote, in turn, the corresponding characteristic functions.

(I) If F_n tends to a distribution function F at every continuity of the latter and φ is the characteristic function of F , then $\varphi_n(x) \rightarrow \varphi(x)$ ($n \rightarrow \infty$) holds uniformly in any finite interval.

¹ Cf. e.g. Rényi [2].

² Here we omit (on historical grounds) the factor $1/\sqrt{2\pi}$.

(II) Assuming — conversely — the convergence of $\{\varphi_n(x)\}$ for every x to some $\varphi(x)$ as well as the continuity of the latter function at $x = 0$, we can conclude the convergence of $\{F_n(x)\}$ to such a distribution function $F(x)$ whose characteristic function is $\varphi(x)$.

The proposition 14.6.1 proves to be fundamental to the asymptotic investigation of the sum of a large number of random variables and to the theoretic foundation of the particularly frequent occurrence of *normal (Gaussian) distributions*.

15.1 Complex L^p -spaces of F. Riesz

15.11 The discussion of Fourier integrals has shown on the one hand that the extension of Lebesgue's integral theory to complex-valued functions is also useful. On the other hand, some important theorems of the preceding chapter (cf. 14.3.2, 14.3.3) contain such independent conditions for pairs of functions — e.g. boundedness + (L)-integrability or square integrability of both functions — which seem to be of heterogeneous nature and yet have the same consequences.

These observations may suggest to develop further the concept of complex $L^2(a, b)$ -space in such a direction that the more general function class thus obtained should include also $L(a, b)$ as a special case, i.e., it should quasi "interpolate" between $L(a, b)$ and $L^2(a, b)$. It is to be expected that in the framework of such a more general theory, the pairs of conditions above mentioned will fuse in a single, more elegant form.

15.12 F. RIESZ who — as is well-known — contributed considerably to the foundations of complex Hilbert space, in the first two decades of this century also elaborated a far-reaching extension of Hilbert's complex theory, in the above "interpolative" aspect. These function classes are called *complex L^p -spaces of Riesz*¹.

Let p be a given positive number (integer or non-integer). The complex $L^p(a, b)$ -class consists of all the complex-valued functions φ measurable in (a, b) for which the integral

$$\int_a^b |\varphi(t)|^p dt$$

exists.² The latter raised to the power $1/p$ has the denomination " $L^p(a, b)$ -norm of φ " and is denoted by

$$\|\varphi\|_p = \left[\int_a^b |\varphi(t)|^p dt \right]^{1/p} \quad (15.1-1)$$

¹ F. Riesz [8]–[9].

² Here and later, the fundamental interval (a, b) can always be replaced by some (L)-measurable set in any Euclidean space.

FURTHER ABSTRACT SPACES, APPLICATIONS

15.1 Complex L^p -spaces of F. Riesz

15.11 The discussion of Fourier integrals has shown on the one hand that the extension of Lebesgue's integral theory to *complex-valued* functions is also useful. On the other hand, some important theorems of the preceding chapter (cf. 14.3.2, 14.3.3) contain such independent conditions for pairs of functions — e.g. boundedness + (L) -integrability or square integrability of both functions — which seem to be of heterogeneous nature and yet have the same consequences.

These observations may suggest to develop further the concept of complex $L^2(a, b)$ -space in such a direction that the more general function class thus obtained should include also $L(a, b)$ as a special case, i.e., it should quasi “interpolate” between $L(a, b)$ and $L^2(a, b)$. It is to be expected that in the framework of such a more general theory, the pairs of conditions above mentioned will fuse in a single, more elegant form.

15.12 F. RIESZ who — as is well-known — contributed considerably to the foundations of complex Hilbert space, in the first two decades of this century also elaborated a far-reaching extension of Hilbert's complex theory, in the above “interpolatoric” aspect. These function classes are called *complex L^p -spaces of Riesz*.¹

Let p be a given positive number (integer or non-integer). The complex $L^p(a, b)$ -class consists of all the complex-valued functions φ measurable in (a, b) for which the integral

$$\int_a^b |\varphi(t)|^p dt$$

exists.² The latter raised to the power $1/p$ has the denomination “ $L^p(a, b)$ -norm of φ ” and is denoted by

$$\|\varphi\|_p = \left[\int_a^b |\varphi(t)|^p dt \right]^{1/p}. \quad (15.1-1)$$

¹ F. Riesz [8]–[9].

² Here and later, the fundamental interval (a, b) can always be replaced by some (L) -measurable set in any Euclidean space.

The limiting case $p \rightarrow \infty$ is important as well. Since for finite (a, b) $\lim_{p \rightarrow \infty} \|\varphi\|_p$ equals plainly the so-called effective least upper bound of $|\varphi(t)|$ i.e., the minimal number $K > 0$ for which $|\varphi(t)| \leq K$ in (a, b) almost everywhere, we adopt the following definitions: the class $L^\infty(a, b)$ means the totality of functions *measurable and essentially* (i.e., apart from a null set) *bounded* in (a, b) , furthermore the pertaining norm is simply

$$\|\varphi\|_\infty = \sup_{\text{a.e. } t \in (a, b)} \{|\varphi(t)|\}. \quad (15.1-2)$$

Specially in the class of functions $\varphi(t)$ *continuous* on a finite segment $[a, b]$ which is usually denoted by $C[a, b]$, the right-hand bound in (15.1-2) becomes the maximum of $|\varphi(t)|$ for $a \leq t \leq b$, so that it is justified to write:

$$\|\varphi\| = \max_{t \in [a, b]} |\varphi(t)| \quad (\varphi \in C[a, b]). \quad (15.1-3)$$

Note that *within any class $L^p(a, b)$, the elements taking the same values in (a, b) apart from a null set are to be considered as identical.*

For $1 \leq p \leq \infty$ it is easy to see that 1) the basic linear operations introduced in the space $L^2(a, b)$ (addition, multiplication by a number) do not take out even of the class $L^p(a, b)$; 2) the concept of norm has the properties corresponding to its geometric origin [cf. (3.1-6)]; 3) from the norm we get similarly the pendants of the notions “distance” and “convergence” as in the case of $L^2(a, b)$; furthermore 4) the Riesz–Fischer criterion 9.5.2 is also transferable to L^p -convergence, so that each sequence $\{\varphi_n\} \subset L^p(a, b)$ satisfying the condition $\|\varphi_m - \varphi_n\|_p \rightarrow 0$ ($m, n \rightarrow \infty$) possesses necessarily an $L^p(a, b)$ -limit.¹

Thus it is natural enough to use the denomination “function space” in case of $L^p(a, b)$ ($p \in [1, \infty]$). Putting $p = 1$ we obtain $L^1(a, b) = L(a, b)$, while for $p = 2$ we get the class $L^2(a, b)$ of square integrable functions; this means that the concept of $L^p(a, b)$ does indeed “interpolate” between $L(a, b)$ and $L^2(a, b)$.

The Hilbert space $L^2(a, b)$ has a special feature: it is the only function space in the set $\{L^p(a, b) : 1 \leq p \leq \infty\}$ admitting the operation (9.4-1) of “inner scalar multiplication” too, so that for $p \neq 2$ the notion of norm cannot be derived from any kind of inner (scalar) product. In fact, for L^p the pendant of (9.4-5) is the so-called Hölder inequality:

$$\left| \int_a^b \varphi(t) \overline{\psi(t)} dt \right| \leq \|\varphi\|_p \|\psi\|_{p'}, \quad \left(\frac{1}{p} + \frac{1}{p'} = 1 \right), \quad (15.1-4)$$

¹ After the first verification of this important proposition published by F. Riesz in his paper [9], several independent proofs were given. Cf. e.g. G.C. Young–W.H. Young [2], Hobson [2], Mikolás [9].

and as a consequence of (15.1-4), (9.4-6) corresponds to the following “*generalized Minkowski inequality*”

$$\|\varphi + \psi\|_p \leq \|\varphi\|_p + \|\psi\|_p; \quad (15.1-5)$$

but $p' = p/(p-1)$ on the right of (15.1-4) equals p (i.e. $\|\varphi\|_p$ and $\|\psi\|_{p'}$ refer to the *same* function space) if and only if $p = 2$.¹

Since Hölder's inequality is an important tool of the theory, independently of the inner product, here it is advisable to introduce some new denominations: two indices $p \in [1, \infty]$, $p' \in [1, \infty]$ as well as the pertaining function spaces $L^p(a, b)$, $L^{p'}(a, b)$ are said to be *dual* (or *conjugate*) of each other, if $p^{-1} + (p')^{-1} = 1$. Accordingly (with the convention $1/\infty = 0$), the dual space of $L(a, b)$ is $L^\infty(a, b)$ and that of $L^2(a, b)$ is itself; Hilbert's function space is the only L^p -space having this latter property. By the above we see for instance that the parts of theorems 14.3.2-14.3.3 holding for all x may be extended to arbitrary pairs of functions f_1, f_2 which belong to dual L^p -spaces.

Is (a, b) a finite interval, then the so-called *p-th integral mean* of a function $\varphi \in L^p(a, b)$

$$\mathfrak{M}_p[\varphi] = \left[\frac{1}{b-a} \int_a^b |\varphi(t)|^p dt \right]^{1/p} = (b-a)^{-1/p} \|\varphi\|_p \quad (15.1-6)$$

increases steadily as $p \rightarrow \infty$. Therefore (cf. 7.4.1) the existence of $\int_a^b |\varphi(t)|^{p_2} dt$ involves that of $\int_a^b |\varphi(t)|^{p_1} dt$ whenever $1 \leq p_1 < p_2 \leq \infty$, i.e., on the latter condition we have

$$L^{p_1}(a, b) \supset L^{p_2}(a, b). \quad (15.1-7)$$

Our result can also be expressed by saying: *in the case of a finite fundamental interval (a, b) , the space $L^p(a, b)$ becomes narrower as p is increased.*

Examples: the power function t^μ belongs to any space $L^p(0, 1)$ ($1 \leq p \leq \infty$) if $\mu \geq 0$, but $t^\mu \notin L^p(0, 1)$ for every $p \in [1, \infty]$ if $\mu \leq -1$; furthermore in the case $-1 < \mu < 0$, for the validity of $t^\mu \in L^p(0, 1)$ it is necessary and sufficient that $1 \leq p < -1/\mu$. Note that specially $1/\sqrt{t}$ belong or does not belong to $L^p(0, 1)$, when $1 \leq p < 2$ and $p \geq 2$, respectively.

It is useful for applications that the notion of L^p -convergence allows transferring immediately to $L^p(a, b)$ ($p \in [1, \infty]$) the concept of *accumulation point*, together with all the related basic ideas of point set theory. (Cf. 9.5.) In particular, we have to stress that any given $L^p(a, b)$ with fixed

¹ Cf. e.g. Hardy-Littlewood-Pólya [1].

$p \in [1, \infty]$ contains the elements of every space $L^r(a, b)$ ($r > p$) and also those of $C(a, b)$ *densely*, i.e., that in each neighbourhood (however small) of whichever element $\in L^p(a, b)$ we can find some elements of $L^r(a, b)$ ($r > p$) and $C(a, b)$ — “neighbourhood” being naturally meant in the $L^p(a, b)$ -sense.

15.13 Deeper considerations on Riesz spaces concern first of all the so-called *bounded linear functionals* (cf. 3.12). Their starting-point is the observation that any (L) -integral of the form

$$\mathcal{L}[\varphi] = \int_a^b \varphi(t) \overline{\psi(t)} dt, \quad (15.1-8)$$

where $\varphi(t) \in L^p(a, b)$ and $\psi(t)$ is a fixed element of the dual space $L^{p'}(a, b)$, satisfies the specifications in the definition of the above operators, provided that the corresponding norm $\|\varphi\|_p$ is utilized.

In 1910 F. RIESZ discovered that there exist no further bounded linear functionals in $L^p(a, b)$ ($1 \leq p < \infty$). Formulating in detail:¹

15.1.1 *If $1 \leq p < \infty$, then to any given bounded linear functional $\mathcal{L}[\varphi]$ of the space $L^p(a, b)$ we can determine uniquely an element $\psi(t)$ of the dual space $L^{p'}(a, b)$ [$p' = p/(p-1)$] such that $\mathcal{L}[\varphi]$ can be written in the form (15.1-8). Moreover we have:*

$$\|\mathcal{L}\| = \|\psi\|_{p'}. \quad (15.1-9)$$

This does not hold for the space $L^\infty(a, b)$. In the special case $C[a, b]$, another famous theorem of Riesz from 1909 allows representation as a Lebesgue–Stieltjes integral (cf. 8.3):²

15.1.2. *In the space of functions φ continuous on a finite segment $[a, b]$, an arbitrary bounded linear functional $\mathcal{L}[\varphi]$ can be written as follows:*

$$\mathcal{L}[\varphi] = \int_a^b \varphi(t) dw(t), \quad (15.1-10)$$

$w(t)$ denoting a suitable function of bounded variation.

The latter proposition also has a certain historical significance. Namely, it has drawn attention to the usefulness of Stieltjes integral (a method hardly known at the turn of the century) in the framework of Lebesgue’s new integral theory.

¹ Cf. F. Riesz [9], § 11.

² F. Riesz [7] and [12].

15.14 It should still be noted that the above chain of ideas leading from L^2 to the more general concept of L^p , can be adapted to the case of Hilbert's sequence space l^2 too (cf. 9.3). In the *complex Riesz sequence space* l^p ($p \geq 1$) so obtained,¹ the definition of the norm of an element $X = (x_1, x_2, \dots)$ is:

$$\|X\|_p = \left(\sum_{k=1}^{\infty} |x_k|^p \right)^{1/p}, \quad (15.1-11)$$

and any statement concerning L^p has a pendant relating to l^p (integrals being replaced by series). The space l^p may also be regarded as a particular case of $L^p(0, \infty)$, restricted to certain step functions.

Some application possibilities suggest occasionally substituting any (L)-integral occurring in the definition of $L^p(a, b)$ ($1 \leq p \leq \infty$) for a suitable (LS)-integral containing an integrator $g(x)$ (cf. 8.3, 10.44). The associated function spaces are usually denoted by $L_g^p(a, b)$ ($1 \leq p < \infty$).

15.2 Linearity, metrics and completeness; Banach's fixed point theorem

15.21 Looking over the results on abstract spaces, we see that certain common properties of these spaces take a prominent part in building up the theory. These properties are the following:

I. *Linear (or vectorial) structure*. Hereby we mean that the so-called linear operations: addition, multiplication by a scalar (constant) are defined in the space, and the elementary rules of vector algebra (commutativity, associativity, distributivity), hold, furthermore, in connection with addition zero element and inverse exist.² Thus it is reasonable to consider the elements of the space as *abstract vectors*.

II. *Metric structure* (existence of *metrics*). We mean by this term that the distance $d[x_1, x_2]$ of any two elements x_1, x_2 of the space is defined, and it fulfils the general requirements for an arbitrary concept of distance. In detail, $d[x_1, x_2]$ is always non-negative and vanishes only for $x_1 = x_2$; further we have $d[x_1, x_2] = d[x_2, x_1]$ and for any three elements x_0, x_1, x_2 the so-called "triangle inequality":

$$d[x_1, x_2] \leq d[x_1, x_0] + d[x_0, x_2].$$

On this basis, it is justified to call the elements of the space *abstract points* as well.

¹ Cf. F. Riesz [7] and [12].

² This is briefly expressed by saying that the elements of the space form a *commutative* (or *Abelian*) *group* with respect to addition.

III. *Completeness*. Hereby is meant that each sequence of elements $\{x_n\}$ satisfying the condition $d[x_m, x_n] \rightarrow 0$ ($m, n \rightarrow \infty$), i.e., any so-called *Cauchy sequence* converges to a limit belonging to the space, provided that the convergence is defined according to the corresponding notion of distance. This property concerns the fact: the elements of the space are *dense everywhere* in a sense mentioned earlier. (Cf. e.g. 2.33.)

15.22 Owing to the above, it is advisable to introduce some new definitions: any set of abstract elements possessing property I., property II. or property III. separately, will be called *linear* (or *vector*) *space*, *metric space* and *complete space*, respectively. Therefore all the spaces E_ν , Hilbert's spaces l^2 and L^2 , more generally any space l^p, L^p ($p \geq 1$) as well as the space C are linear, metric and complete spaces.

These three notions are obviously connected with each other through the concept of *norm* which may be regarded as length of an element (as an abstract vector) or also as distance of the element in question (as an abstract point) from the zero element (i.e., origin) of the space. As far as completeness is concerned, its dependence on the metrics is shown by the definition; moreover, a pertinent deep theorem establishes that *every metric space can be extended to a complete metric space* by means of the convergence concept pertaining to the given space. For studying completeness, another related notion is still useful: an abstract space is said to be *separable*, if it contains a countable subset dense everywhere in the space.¹

The general theory of linear spaces — intertwined with the foundations of higher linear algebra (linear dependence and independence, linear basis, dimension, linear subspaces etc.) — is well-known. The general concept of metric space allows much more far-reaching conclusions; after completion we can easily transfer to such a space the whole theory based on the notions of neighbourhood, convergence and accumulation point — having special regard to compactness. (See Chapter 2.)

15.23 Dealing with arbitrary complete metric spaces, an operator theorem published by the Polish mathematician S. BANACH in 1932 is of utmost importance. It may be formulated as follows:

15.2.1 (BANACH'S FIXED POINT THEOREM)² *Let Ω be a complete metric space and consider a mapping $\mathcal{T}[x] = \mathcal{T}x$ of Ω in itself such that any two elements $x_1, x_2 \in \Omega$ for suitable $q \in (0, 1)$ satisfy the inequality*

$$d[\mathcal{T}x_1, \mathcal{T}x_2] \leq q \cdot d[x_1, x_2].^1 \quad (15.2-1)$$

¹ Example: the set of points with rational coordinates in E_ν . All the complete spaces mentioned hitherto possess the property of separability.

² Cf. Banach [1].

Then the operator equation

$$\mathcal{T}[x] = x \quad (15.2-2)$$

has exactly one solution, in other terms the operator $\mathcal{T}$ has one single fixed point.

P r o o f. We need a procedure similar to the *iteration* used for approximate solving ordinary equations, which yields the only solution of (15.2-2) looked for as “ Ω -limit” of a sequence of elements $\in \Omega$.

Let us start from an optional point $x^{(0)} \in \Omega$ and form the sequence in Ω :

$$x^{(1)} = \mathcal{T}x^{(0)}, x^{(2)} = \mathcal{T}x^{(1)}, \dots, x^{(m)} = \mathcal{T}x^{(m-1)}, \dots \quad (15.2-3)$$

By successive application of the condition (15.2-1) we get:

$$d[x^{(m)}, x^{(m+1)}] \leq q^m d[x^{(0)}, x^{(1)}] \quad (m = 1, 2, \dots),$$

hence by the triangle equality, in $n - m$ steps:

$$\begin{aligned} d[x^{(m)}, x^{(n)}] &\leq (q^m + q^{m+1} + \dots + q^{n-1}) d[x^{(0)}, x^{(1)}] \leq \\ &\leq \frac{q^m}{1 - q} d[x^{(0)}, x^{(1)}]. \end{aligned}$$

Therefore the limit relation

$$\lim_{m, n \rightarrow \infty} d[x^{(m)}, x^{(n)}] = 0,$$

holds, i.e., (15.2-3) is a *Cauchy sequence* of the space Ω .

But Ω is assumed to be complete; consequently, the sequence has a limit $\in \Omega$ in the sense of convergence according to the metrics of Ω . This means the existence of an element $x^* \in \Omega$ such that

$$\lim_{m \rightarrow \infty} d[x^{(m)}, x^*] = 0,$$

and then by (15.2-1) we have

$$\lim_{m \rightarrow \infty} d[\mathcal{T}x^{(m)}, \mathcal{T}x^*] = 0.$$

Thus the definition of $x^{(m)}$ implies that x^* satisfies the operator equation (15.2-2). The uniqueness of this solution can be verified in the following

¹ A map having this property is usually called *contraction* in Ω .

way: if $x^{**} \in \Omega$ fulfils (15.2-2) as well, then we obtain on the one hand $d[\mathcal{T}x^*, \mathcal{T}x^{**}] = d[x^*, x^{**}]$, on the other hand by (15.2-1) the inequality $d[\mathcal{T}x^*, \mathcal{T}x^{**}] \leq q \cdot d[x^*, x^{**}]$. Owing to $0 < q < 1$, this is possible only if $d[x^*, x^{**}] = 0$, whence $x^* = x^{**}$.

* * *

In the recent literature, Banach's theorem 15.2.1 and its generalizations are often used in connection with such questions of algebra and analysis which imply various iterative algorithms. In particular, let us mention some problems of existence or uniqueness in the theory of finite and infinite systems of linear equations as well as of differential and integral equations, furthermore the computation of solutions.

15.3 Abstract Hilbert space and Neumann's foundation of quantum mechanics

15.31 We know that the fundamental notions and statements of solid analytic geometry are all transferable to such abstract spaces which are simultaneously linear, metric and complete. Thus it is justified that in functional analysis and its several application fields two particular classes of abstract spaces possessing the above three properties take a prominent part, namely those for which the metric is generated in the most natural manner: through the notion of *inner* (scalar) *product* or immediately by introduction of a *norm*. Next we are going to study the concept of abstract space resulting from the first possibility; its exact definition is:

A set H of arbitrary elements is called *abstract Hilbert space* if I. H is linear; II. to any two elements $\varphi, \psi \in H$ a (real or complex) number is attached, as well as their so-called inner (scalar) product (φ, ψ) , with the properties: $(\varphi, \psi) = \overline{(\psi, \varphi)}$, $(\alpha\varphi, \psi) = \alpha(\varphi, \psi)$ for every constant α , further $(\varphi_1 + \varphi_2, \psi) = (\varphi_1, \psi) + (\varphi_2, \psi)$, $(\varphi, \varphi) > 0$ ($\varphi \neq 0$) and $(\varphi, \varphi) = 0$ ($\varphi = 0$); III. norm and distance are defined in H as follows: $\|\varphi\| = (\varphi, \varphi)^{1/2}$ and $d[\varphi, \psi] = \|\varphi - \psi\|$, respectively; IV. H is complete with respect to the $d[\varphi, \psi]$ -metrics; V. H is "infinite dimensional", i.e., there exist n linearly independent elements in H for every positive integer n ; VI. H is separable. Specially, we speak about a *real* or *complex* space H depending on whether the constants implied by conditions I. and II. are real or complex numbers.¹

The concept of abstract Hilbert space is obviously a common generalization of Hilbert's sequence space l^2 and Hilbert's function space L^2 ; we

¹ The terminology of the field is not unified: the restrictions V-VI. in the above definition are omitted by some authors. A linear space with inner multiplication is usually called *unitary*. Condition II. involves $(\varphi, \psi) = (\psi, \varphi)$ for *real* Hilbert space, i.e., the commutativity of scalar product.

see that such an extension is wholly justified by the isomorphy of l^2 and L^2 , in accordance with the fundamental theorem 11.4.1 of Riesz and Fischer. Accordingly, every abstract Hilbert space is isomorphic to l^2 , this involving the isomorphy of any two abstract Hilbert spaces.

By means of the notion of inner (scalar) product, we can easily extend not only the (metric) theory of point sets, but also the ideas and results connected with orthogonality to abstract Hilbert spaces (cf. 9.3–9.6, 10.7, 11.1–11.6). Thus e.g. Bessel's inequality (11.2–1), Parseval's formula (11.2–2), the general expansion theorem 11.3.1 etc. preserve their validity — naturally in the sense that L^2 -convergence is everywhere to be substituted by convergence according to the metrics of H . Note that all the statements regarding *strong and weak convergence* may directly be transferred to an arbitrary Hilbert space H (cf. 9.54); furthermore, the existence of a *complete orthogonal system* consisting of elements of H is a simple consequence of the separability of H and the Gram–Schmidt orthogonalization process. (In H , any complete orthogonal system necessarily contains countably infinitely many elements.)

15.32 As regards the linear operators of H , the isomorphy of this space to l^2 and L^2 yields a proposition due to M. FRÉCHET and F. RIESZ (1907):¹

15.3.1 Every bounded linear functional defined in the abstract Hilbert space H may be written in the form:

$$\mathcal{L}[\varphi] = (\varphi, \psi). \quad (15.3-1)$$

Here ψ is an element of H , determined by $\mathcal{L}$ uniquely, for which $\|\mathcal{L}\| = \|\psi\|$.

Beside the bounded linear functionals of H , the *bounded linear operators* of the space mapping H into itself are of great importance too (cf. 3.13). For such an operator $\mathcal{A}[\varphi] = \mathcal{A}\varphi$ ($\varphi \in H, \mathcal{A}\varphi \in H$), the *norm* is defined — as it is well-known — on the model of $\|\mathcal{L}\|$; trivially $\|\varphi_n - \varphi\| \rightarrow 0$ ($n \rightarrow \infty$) implies

$$\|\mathcal{A}\varphi_n - \mathcal{A}\varphi\| = \|\mathcal{A}(\varphi_n - \varphi)\| \leq \|\mathcal{A}\| \|\varphi_n - \varphi\| \rightarrow 0.$$

(This property is called *simple continuity* of $\mathcal{A}[\varphi]$.)

If in addition it is still true that the weak convergence of any sequence $\{\varphi_n\} \subset H$ to some element $\varphi \in H$ involves the strong convergence of $\mathcal{A}\varphi_n$ to $\mathcal{A}\varphi$, then the operator $\mathcal{A}$ is said to be *totally* (completely) *continuous*.

In H , a relevant notion connected with the inner (scalar) product is that of *bilinear functional*. Hereby we mean an abstract function $\mathcal{B}[\varphi, \psi]$ defined for certain pairs of elements $\varphi \in H, \psi \in H$, which for fixed ψ becomes a bounded linear functional of φ , and for fixed φ such a functional of ψ .

¹ Cf. F. Riesz [4], Fréchet [1], furthermore Fréchet [2] and F. Riesz [24].

A bilinear functional is called *Hermitian*, if $\mathcal{B}[\varphi, \psi] = \mathcal{B}[\psi, \varphi]$ holds. We conclude from 15.3.1 that $\mathcal{B}[\varphi, \psi]$ may also be written in the form:

$$\mathcal{B}[\varphi, \psi] = (\varphi, \mathcal{A}\psi) = (\mathcal{A}^*\varphi, \psi), \quad (15.3-2)$$

where $\mathcal{A}$ and $\mathcal{A}^*$ denote such bounded linear operators for which $\|\mathcal{A}\| = \|\mathcal{A}^*\|$. $\mathcal{A}$ and $\mathcal{A}^*$ are so-called *adjoint operators*; $\mathcal{A}$ is said to be *self-adjoint* if $\mathcal{A} = \mathcal{A}^*$. The *quadratic functional*

$$\mathcal{Q}[\varphi] = \mathcal{B}[\varphi, \varphi] \quad (15.3-3)$$

is a particular case of the “Hermitian bilinear functional” and can also be written as

$$\mathcal{Q}[\varphi] = (\varphi, \mathcal{A}\varphi) = (\mathcal{A}^*\varphi, \varphi).$$

A central subject of the pertinent theory is to study the operator equation

$$\mathcal{A}\varphi = \lambda\varphi \quad (\varphi \neq 0, \varphi \in H), \quad (15.3-4)$$

$\mathcal{A}$ denoting a totally continuous and self-adjoint, bounded linear operator. It can be shown that in our case all the values of the parameter λ for which (15.3-4) is solvable — these are called *eigenvalues* — must be real and constitute a countable set Λ in their entirety. Moreover, we always can form from the solutions (“*eigenvectors*”) corresponding to Λ (“*spectrum*”) a complete orthogonal system $\varphi_n (n = 1, 2, \dots)$; and if λ_n denotes an eigenvalue to which φ_n is attached, we have the following

15.3.2 (SPECTRAL DECOMPOSITION THEOREM) *Any bounded linear operator $\mathcal{A}\varphi \in H$ which is self-adjoint and totally continuous, may be expressed in terms of the eigenvalues $\{\lambda_n\}$ and eigenvectors $\{\varphi_n\}$ in the form:*

$$\mathcal{A}\varphi = \sum_{n=1}^{\infty} \lambda_n (\varphi, \varphi_n) \varphi_n \quad (\lim_{n \rightarrow \infty} \lambda_n = 0).^1 \quad (15.3-5)$$

15.3.3 Now let us sum up the *linear algebraic interpretation* of the above concepts and results, which corresponds to the case of the simplest Hilbert space, i.e. to l^2 . First we recall the fact that several-dimensional vectors can also be characterized by row or column matrices and *the totality of arbitrary matrices of the same type forms a linear space*.

¹ Cf. F. Riesz [21], [22], [26]. Note that the generalized Fourier series in (15.3-5) may also be regarded as a *Stieltjes integral*. For a comprehensive treatment of the spectral theory of Hilbert's space we refer to the classical monograph [3] of SZÖKEFALVI-NAGY, furthermore to a recent book by him and FOIAS entitled “Harmonic analysis of operators on Hilbert space”, as well as to the following papers of the first mentioned author: [5], [14], [15], [17], [18].

If we consider the elements of the complex space l^2 as infinite column matrices and the inner product of a pair of elements

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \end{bmatrix} \quad \left(\sum_{k=1}^{\infty} |x_k|^2 < \infty, \quad \sum_{k=1}^{\infty} |y_k|^2 < \infty \right)$$

is written correspondingly as a *matrix product*:

$$(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{\infty} x_k \bar{y}_k = \mathbf{x}^* \cdot \bar{\mathbf{y}}$$

(where the asterisk denotes transposition), then we see what and how is to be translated into the "language" of matrix theory.

In space l^2 , a bounded linear functional $\mathcal{L}[\mathbf{x}]$ is plainly a homogeneous linear function of all the coordinates of $\mathbf{x}$, and this puts into evidence the representation $\mathcal{L}[\mathbf{x}] = (\mathbf{x}, \mathbf{y})$ at certain $\mathbf{y} \in l^2$. Similarly: in the case of a bounded linear operator $\mathbf{y} = \mathbf{A}\mathbf{x}$, any coordinate y_m must depend in a homogeneous and linear manner on all the coordinates of $\mathbf{x}$, i.e.

$$y_m = \sum_{n=1}^{\infty} a_{mn} x_n \quad (m = 1, 2, \dots). \quad (15.3-6)$$

Thus we can write:

$$\mathbf{y} = [a_{mn}] \cdot \mathbf{x}, \quad (15.3-7)$$

therefore the application of the operator $\mathbf{A}$ signifies a left-multiplication by the infinite matrix

$$\mathbf{A} = [a_{mn}] = \begin{bmatrix} a_{11} & a_{12} & \dots \\ a_{21} & a_{22} & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix} \quad (15.3-8)$$

The same way of writing yields for any $\mathbf{x}, \mathbf{y} \in l^2$:

$$\overline{\mathcal{B}[\mathbf{x}, \mathbf{y}]} = \bar{\mathbf{x}}^* \cdot (\mathbf{A} \mathbf{y}) = \sum_{m,n} a_{mn} \bar{x}_m y_n, \quad (15.3-9)$$

$$\mathcal{Q}[\mathbf{x}] = \bar{\mathbf{x}}^* \cdot (\mathbf{A} \mathbf{x}) = \sum_{m,n} a_{mn} x_m \bar{x}_n \quad (\mathbf{A} = \bar{\mathbf{A}}^*), \quad (15.3-10)$$

which are the classical algebraic representations of the *bilinear* and *quadratic forms*.

On the other hand, by (15.3-2) the operator $\mathcal{A}^*$ has the matrix $\overline{\mathbf{A}}^*$, i.e., forming an adjoint operator means “conjugation” and “transposition” for the pertaining matrix. In particular: an operator $\mathcal{A}$ is self-adjoint if $\mathbf{A} = \overline{\mathbf{A}}^*$, i.e., if $\mathbf{A}$ is a so-called *Hermitian matrix*.¹

Finally we establish: to the operator problem (15.3-4) the matrix eigenvalue problem

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad (\mathbf{x} \neq \mathbf{0}, \mathbf{x} \in l^2; \mathbf{A} = \overline{\mathbf{A}}^*) \quad (15.3-11)$$

is attached. Besides we see: *the spectral representation (15.3-5) is nothing else but a strong extension of the fundamental formula about the transcription of an Hermitian matrix into diagonal form.*

15.34. The description of the theory of l^2 through matrices (which was developed by D. HILBERT soon after the turn of this century — this is why his name was attached to the space l^2) takes a prominent part in the so-called *matrix mechanics*, i.e., in the treatment of quantum mechanics due to W.K. HEISENBERG, M. BORN and P. JORDAN (1925). The basic problem of the latter theory is the “diagonalization” of the matrix pendant of Hamilton’s function (known from classical mechanics), in close connection with an eigenvalue problem of the type (15.3-11). It is due to certain physical realities that only infinite Hermitian matrices may be taken into consideration.

Wave mechanics published by E. SCHRÖDINGER in 1926, yields the same results though it starts from a boundary problem pertaining to a partial differential equation (the wave equation) and correspondingly operates by analytic methods. Thus the *equivalence* of the two theories became evident and was soon observed by Schrödinger, too. The deeper background of equivalence, however, was found first by J. NEUMANN in 1927: the solutions of the fundamental problem of wave mechanics (the so-called wave functions) are required to be square integrable, i.e., they form an L^2 -space, and this is isomorphic to the l^2 -space used in matrix mechanics according to the theorem 11.4.1 of F. RIESZ and E. FISCHER.

J. NEUMANN also succeeded in giving a new, mathematically *rigorous establishment of a unified quantum mechanics*, by developing further the pertinent results of F. RIESZ and by elaborating the theory of abstract Hilbert space in detail. Heisenberg’s and Schrödinger’s presentations may be regarded as equivalent realizations of Neumann’s theory which are very different in terminology but entirely similar structurally — like the interconnection between the spaces l^2 , L^2 and the concept of H .²

¹ If all elements of $\mathbf{A}$ are real, then the last-mentioned condition becomes simply $\mathbf{A} = \mathbf{A}^*$, that is $\mathbf{A}$ must be a *symmetric matrix*.

² Cf. F. Riesz [11], [13], [22], [24]; J. Neumann [1]; furthermore H. Weyl [1] and Hille [1], Sz.-Nagy-Foias-Gehér [1].

During the twenties P. DIRAC and (independently) P. JORDAN tried to develop a common treatment of matrix and wave mechanics in quite different directions, through substitution of the differential operator in wave equation by a suitable integral operator. If, namely, such a transcription can be carried out, we can utilize the close analogy and connection between integral equations and linear equations of infinitely many variables (occurring in matrix mechanics), and Hilbert's extensive theory about eigenvalue problems of integral equations.¹

15.35 Unfortunately, the transformation of differential operators to integral ones is generally impossible, or rather it can be done only at the price of certain apparently absurd conditions. Such an assumption is e.g. in our case that an operator equation of the type (15.3-4) can be solved for any admitted A ; further that there exists a "function" $\delta(t)$ for which

$$\delta(t) = \begin{cases} 0 & (t \neq 0) \\ \infty & (t = 0) \end{cases} \quad \text{and} \quad \int_{-\infty}^{\infty} \delta(t) dt = 1. \quad (15.3-12)$$

This so-called "*Dirac-impulse*" δ which has been very often used till now (together with its formal derivatives) in physical and electrotechnical publications, may by no means be considered as a function in the classical sense of the word; its symbol refers merely to certain (unadmitted) operations under the integral sign. For instance, the formula

$$u(x) = \int_{-\infty}^{\infty} u(t) \delta(x-t) dt = u \star \delta, \quad (15.3-13)$$

utilized particularly in engineering works on linear systems, means nothing else but the representation

$$u(x) = \int_{-\infty}^{\infty} u(t) dE(x-t) \quad (15.3-14)$$

where

$$E(t) = \begin{cases} 0 & (t < 0) \\ 1/2 & (t = 0) \\ 1 & (t > 0) \end{cases} \quad (15.3-15)$$

is HEAVISIDE'S *unit jump function*; the symbolic integral with respect to t in (15.3-13) comes from the (exactly defined) Stieltjes integral in (15.3-14) by writing $dE/dt = \delta(t)$, i.e., carrying out formally the differentiation of $E(t)$ under the integral sign. (Cf. 8.3.3.)

Nevertheless, the development of the theory of integral transformations and of some other application fields has suggested recently the extension of

¹ See Dirac [1], Hilbert [1].

the classical concept of function in such a way that also $\delta(t)$ and its symbolic derivatives have obtained certain direct (abstract) interpretation. Since the fifties the modern *theory of generalized functions* has been established, with two main branches, L. SCHWARTZ's *distribution theory* and J. MIKUSINSKI's *operational calculus*.¹

15.4 Banach space; concept of general topological space

15.41 As was mentioned prior to the introduction of the concept of abstract Hilbert space, another very important class of linear, metric and complete spaces can be generated by ensuring the metrics immediately by giving norm. Those linear, normed [cf. (3.1-6)] abstract spaces which are also complete with respect to the notion of distance $d[\varphi, \psi] = \|\varphi - \psi\|$ are called *Banach spaces* (cf. 3.12, 15.22). Their prototypes are clearly the Riesz spaces L^p , and their first systematic treatment is found in a famous monograph of S. BANACH (1932).²

The concept of Banach space is a *generalization* of abstract Hilbert space in a certain direction, since the latter plainly satisfies all the requirements in the definition of Banach space, but the conditions characterizing a Banach space do not allow of any kind of conclusion about dimension number and separability. A further feature is that the notion of Hilbert space includes the existence of inner (scalar) product while in case of Banach space this is replaced by the norm; so it is natural to ask: is it always possible to derive the norm in a Banach space from an appropriate notion of scalar product, namely as the square root of the scalar product of an element by itself?

The answer is negative: the introduction of inner (scalar) product in a Banach space can be carried out only on a certain condition, and this restriction is fulfilled only by some proper subclasses of Banach spaces. Formulating in detail:

15.4.1 (THEOREM OF J. NEUMANN AND P. JORDAN) *The notion of norm in a Banach space B can be given by means of a notion of inner product constructed in the same space if and only if any two elements φ and ψ of B satisfy the so-called parallelogram identity.*³

$$\|\varphi + \psi\|^2 + \|\varphi - \psi\|^2 = 2(\|\varphi\|^2 + \|\psi\|^2). \quad (15.4-1)$$

¹ Cf. L. Schwartz [1] and Mikusinski [1].

² See Banach [1].

³ The denomination refers to the geometric meaning of (15.4-1) in the Banach space E_2 : the square sum of diagonals of a parallelogram equals the square sum of its sides. Cf. P. Jordan-J. Neumann [1].

In fact, the necessity of this condition follows from the fact that the notion of scalar product (φ, ψ) ($\varphi \in B$, $\psi \in B$) allows to write:

$$\begin{aligned} \|\varphi + \psi\|^2 + \|\varphi - \psi\|^2 &= (\varphi + \psi, \varphi + \psi) + (\varphi - \psi, \varphi - \psi) = \\ &= \|\varphi\|^2 + \|\psi\|^2 + (\varphi, \psi) + \overline{(\varphi, \psi)} + \|\varphi\|^2 + \|\psi\|^2 - (\varphi, \psi) - \overline{(\varphi, \psi)} = \\ &= 2\|\varphi\|^2 + 2\|\psi\|^2; \end{aligned}$$

and the sufficiency of (15.4-1) can also be obtained in a much more lengthy but elementary way.

* * *

It is easy to see that L^2 is the only Riesz space L^p ($1 \leq p \leq \infty$) satisfying (15.4-1); thus L^p ($p \neq 2$) yields example of Banach spaces which are not abstract Hilbert spaces as well. A similar statement holds on $C[a, b]$, i.e., the space of functions continuous in $[a, b]$, and the same is true for $V[a, b]$, i.e., the functions are of bounded variation on the segment $[a, b]$; in the latter space, the natural definition of norm is (cf. 4.41):

$$\|\varphi\| = V_\varphi(a, b). \quad (15.4-2)$$

Let us observe that every special space mentioned in 15.2 — among them also any Euclidean space or l^p -space — belongs to the family of Banach spaces, all of them being normed too.

Evidently, all our earlier results regarding the structure of metric, linear and complete spaces (cf. 15.21) hold good also for Banach spaces as sets of abstract points or vectors. We have to realize, however, that without inner product the general geometry of a Banach space is devoid also of such notions as inclination angle, orthogonality etc., so that even such fundamental representation theorems as 15.3.1 and 15.3.2 have no pendants in the theory of Banach spaces. Yet we add that the study of bounded linear functionals and operators yields a number of significant, well applicable propositions.

15.42 If the totality of bounded linear functionals $\mathcal{L} = \mathcal{L}[\varphi]$ defined in some Banach space B is denoted by B^* , then the linearity of this set is trivial, and the pertaining notion of norm may be written in the form (cf. 3.12):

$$|\mathcal{L}| = \sup_{\substack{\varphi \in B^* \\ \|\varphi\| > 0}} \{ \|\varphi\|^{-1} |\mathcal{L}[\varphi]| \}. \quad (15.4-3)$$

(15.4-3) fulfils the requirements (3.1-6), and the completeness of B^* with respect to the "distance" $d[\mathcal{L}_1, \mathcal{L}_2] = \|\mathcal{L}_1 - \mathcal{L}_2\|$ follows easily. Therefore

we can state: *the bounded linear functionals of a Banach space B constitute another Banach space themselves; this space B^* is called the adjoint of B .*¹

Since in the case of an abstract Hilbert space H , a bijective mapping can be established between the elements of H and H^* by (15.3-1), for which the norms of corresponding elements are equal and the linear operations carried out on corresponding elements yield again connected elements, so H^* is isomorphic to H and we may conclude: *H is a structurally self-adjoint space.*

Similar considerations show that the adjoint of any Euclidean space E_ν is isomorphic to E_ν itself, the adjoint of an L^p -space is in isomorphic connection to the dual space $L^{p'}$ [$p' = p/(p-1)$], the adjoint of l^2 to its dual $l^{p'}$ [$p' = p/(p-1)$], finally the adjoint of space C is isomorphic to space V (cf. 15.1.1-15.1.2). These results provide deeper background to the correspondence of dual spaces.

Let us now pay attention to arbitrary *bounded* linear operators $\mathcal{T} = \mathcal{T}[\varphi]$ of a Banach space B (cf. 3.13) mapping B onto another Banach space $\tilde{B}$. Like the above, the totality of these operators $\mathcal{T}$ forms again a Banach space, if the linear operations and norm are introduced as usual; the standard notation of the above operator set is $L((B, \tilde{B}))$.

15.43 The most frequent special case is $\tilde{B} = B$, i.e., mappings of B onto itself. The following relevant result concerns the corresponding Banach space $L((B, B))$ and an *operator expansion* which may be regarded as a *strong extension of the geometric series*.

15.4.2 Let B be a Banach space and denote by I the unit operator of $L((B, B))$.

Then $I - \mathcal{T}$ has a unique inverse for any bounded linear operator $\mathcal{T} \in L(B, B)$ with $\|\mathcal{T}\| < 1$, and the so-called *Neumann's expansion* holds

$$(I - \mathcal{T})^{-1} = \sum_{n=0}^{\infty} \mathcal{T}^n. \quad (15.4-4)$$

The convergence of the series on the right is to be taken according to $L((B, B))$ -norm.

P r o o f. Let us start from the operator equation

$$y = (I - \mathcal{T})x \quad (x \in B) \quad (15.4-5)$$

which is equivalent to

$$x = y + \mathcal{T}x = \mathcal{T}_y[x]. \quad (15.4-6)$$

¹ In the literature, the terms "dual" or "conjugate" are also usual.

(15.4-5) shows that $I - \mathcal{T}$ maps space B uniquely onto itself; i.e., our assertion that $(I - \mathcal{T})^{-1}$ exists and is one-valued, can also be expressed as (15.4-6) is solvable for any prescribed $y \in B$ uniquely.

In fact, considering that the estimation

$$\|\mathcal{T}_y x_1 - \mathcal{T}_y x_2\| = \|\mathcal{T}x_1 - \mathcal{T}x_2\| \leq \|\mathcal{T}\| \|x_1 - x_2\|$$

holds for arbitrary pairs of elements $x_1 \in B$, $x_2 \in B$, by $\|\mathcal{T}\| < 1$ we can apply to (15.4-6) Banach's fixed point theorem 15.2.1.

As regards the expansion of $(I - \mathcal{T})^{-1}$, first of all we note that the partial sums

$$\mathcal{U}_N = \sum_{n=0}^N \mathcal{T}^n \quad (N = 0, 1, 2, \dots)$$

form a Cauchy sequence in $L((B, B))$, because for $M < N$

$$\begin{aligned} \|\mathcal{U}_N - \mathcal{U}_M\| &= \left\| \sum_{n=M+1}^N \mathcal{T}^n \right\| \leq \\ &\leq \sum_{n=M+1}^N \|\mathcal{T}\|^n < \frac{\|\mathcal{T}\|^{M+1}}{1 - \|\mathcal{T}\|} \rightarrow 0 \end{aligned}$$

as M and $N \rightarrow \infty$.

Now, $L((B, B))$ is itself a Banach space and thus complete, too; consequently, the sequence $\{\mathcal{U}_N\}$ must converge to some element $\mathcal{U} \in L((B, B))$ with respect to the $L((B, B))$ -norm. It remains to be verified that $\mathcal{U} = (I - \mathcal{T})^{-1}$, i.e.,

$$(I - \mathcal{T}) \mathcal{U} = I. \quad (15.4-7)$$

For that purpose, let us estimate the norm deviation of the left- and right-hand term of (15.4-7). Using the identity

$$(I - \mathcal{T}) \mathcal{U}_N = \sum_{n=0}^N \mathcal{T}^n - \sum_{n=1}^{N+1} \mathcal{T}^n = I - \mathcal{T}^{N+1},$$

we can write:

$$\|(I - \mathcal{T}) \mathcal{U} - I\| \leq \|I - \mathcal{T}\| \|\mathcal{U} - \mathcal{U}_N\| + \|\mathcal{T}\|^{N+1}.$$

Since here $\|\mathcal{U} - \mathcal{U}_N\|$ and $\|\mathcal{T}\|^{N+1}$ tend to 0 as $N \rightarrow \infty$, we get the condition

$$\|(I - \mathcal{T}) \mathcal{U} - I\| = 0,$$

which implies (15.4-7).

* * *

We want to stress that the *Neumann series*¹ (15.4-4) is a powerful tool in various application fields concerning Hilbert space, especially in the theory of integral equations.

15.44 Among the principal trends in the development of functional analysis, we have to mention the concept of *general topological space*, which has become during the last decades a common basis of the study of abstract spaces and of several other mathematical disciplines. The foundations of this field were created in some pioneering works of F. RIESZ and M. FRÉCHET on metric spaces around 1910.² Riesz himself and a number of other mathematicians (F. HAUSDORFF, further P.S. ALEXANDROV, P.S. URYSON et al.) established soon afterwards that the transfer of most important notions and results of the elementary point set theory to abstract sets is possible also *without metrics*, if we provide certain subsets with properties corresponding to *open sets*, as well as *neighbourhoods* of Euclidean spaces. The denomination “topological spaces” refers on the one hand to the lack of metrics, on the other hand to the convention that the spaces in question are regarded as identical if they are “topologically equivalent” (cf. 2.13).

The notion of neighbourhood attached to some element (point) x of a basic set X is usually characterized by the following axioms: I. every neighbourhood of x contains x ; II. if τ is a neighbourhood of x , then the same holds for any subset of X containing τ ; III. if τ_1 and τ_2 are neighbourhoods of x , then $\tau_1 \cap \tau_2$ is as well; IV. to any neighbourhood τ_1 of x we can find another neighbourhood τ_2 of the same point such that τ_1 is also a neighbourhood of every point of τ_2 . If we attach to every point of X the subsets with these properties I-IV. as neighbourhoods, then we say that a *topology* is defined on X , by which X becomes a so-called *general topological space*.

Note that the pertinent “*metrization problem*” was only solved in the early fifties; i.e., during that period the topological spaces generated by some prescribed concept of distance were successfully characterized.³ At about the same time, a unified theory of the topological space (in the above sense) as well as of other related “continuity structures” was elaborated.⁴

15.5 Extended applications concerning ordinary differential equations

15.51 In Part II, we have often dealt with differential equations, orthogonal systems or series originating from eigenvalue problems, furthermore the

¹ See C.G. Neumann [1].

² Cf. e.g. F. Riesz [5], M. Fréchet [2].

³ This result is due to R.H. BING, J. NAGATA and YU.M. SMIRNOV (1950-51).

⁴ Cf. Á. Császár [1].

connection between the theory of various abstract spaces and differential and integral operators. For instance, the classical orthogonal polynomial systems (those of Jacobi, Laguerre and Hermite), the system $\{J_\mu(\omega_n x)\}$ ($x \in [0, 1]$) arising from Bessel functions, as well as the corresponding Fourier series are all generated by boundary problems of Sturm–Liouville type (10.3, 10.6); Haar's equiconvergence theorem concerns Fourier series in terms of eigenfunctions of an arbitrary regular Sturm–Liouville problem (cf. 12.8); the quantum theoretic applications of Hilbert's space and certain properties of linear operators in Banach spaces are interlinked with considerations regarding some kinds of differential and integral operators (cf. 15.3, 15.4).

Here and in the forthcoming sections we want to develop further and extend the above results. Result 10.6.1 on the orthogonality of Sturm–Liouville functions can strongly be extended in the case $\tau(x) = 0$ to so-called self-adjoint eigenvalue problems.

15.52 Let us denote by

$$\mathcal{L}[y] = \sum_{k=0}^{\nu} (-1)^k \left[f_k(x) \cdot y^{(k)}(x) \right]^{(k)} \quad (15.5-1)$$

a so-called *ordinary homogeneous linear differential form of order 2ν* , where $f_0, f_1, \dots, f_{\nu-1}$ and $f_\nu \neq 0$ are given functions, $y = y(x)$ is a variable function, furthermore all the occurring derivatives are supposed to exist in some fixed interval $a \leq x \leq b$. We assume still, that y and its derivatives up to order $2\nu - 1$ satisfy the following homogeneous linear boundary conditions with constant coefficients:

$$\sum_{k=0}^{2\nu-1} \left[A_{jk} y^{(k)}(a) + B_{jk} y^{(k)}(b) \right] = 0 \quad (j = 1, 2, \dots, 2\nu), \quad (15.5-2)$$

and that these 2ν stipulations are linearly independent from one another.

The form (15.5-1) is called *self-adjoint* in $[a, b]$, if it fulfils the integral condition:

$$\int_a^b (u\mathcal{L}[v] - v\mathcal{L}[u]) dx = 0^1 \quad (15.5-3)$$

for any two elements $u = u(x)$, $v = v(x)$ of the above set of functions $\{y(x)\}$. Since the integral on the left can be evaluated by a *transformation formula of Dirichlet*:

$$\int_a^b (u\mathcal{L}[v] - v\mathcal{L}[u]) dx = \quad (15.5-4)$$

¹ More generally, two differential forms $\mathcal{M}[y]$ and $\mathcal{N}[y]$ of the same type (15.5-1) and of order 2ν are said to be *adjoint to each other* in $[a, b]$, if the integral of $u\mathcal{M}[v] - v\mathcal{N}[u]$ on this segment vanishes for every admitted pair of functions $u(x), v(x)$. The notion can plainly be extended also to functions with several variables.

$$\left[\sum_{k=1}^{\nu} \sum_{l=0}^{k-1} (-1)^{k+l} \left\{ u^{(l)}(f_k \cdot v^{(k)})^{(k-1-l)} - v^{(l)}(f_k \cdot u^{(k)})^{(k-1-l)} \right\} \right]_{x=a}^{x=b},$$

it may be pointed out: $\mathcal{L}[y]$ is self-adjoint in $a \leq x \leq b$ if and only if the last expression in (15.5-4) vanishes for any two admitted functions u and v .

15.53 Consider now a homogeneous linear differential equation of order $2r$:

$$\mathcal{L}_1[y] = \lambda \mathcal{L}_2[y], \quad (15.5-5)$$

where $\mathcal{L}_1$ and $\mathcal{L}_2$ are differential operators of the type (15.5-1), both self-adjoint in $[a, b]$ with $y = y(x)$ satisfying boundary conditions of form (15.5-2); $2r$ means the greater of the degrees of $\mathcal{L}_1, \mathcal{L}_2$, while λ is a real parameter.

Is λ_m, λ_n a pair of distinct λ -values for which (15.5-5) is solvable on the boundary conditions given, i.e., if λ_m and λ_n are two different *eigenvalues* of the problem, then $y_m(x)$ and $y_n(x)$ denoting a pair of corresponding *eigenfunctions*, we can deduce like the proof of 10.6.1 the so-called *generalized orthogonality relations*:

$$\begin{cases} \int_a^b y_n \mathcal{L}_2[y_n] dx = 0, \\ \int_a^b y_m \mathcal{L}_1[y_m] dx = 0. \end{cases} \quad (15.5-6)$$

In particular, for $\mathcal{L}_2[y] = -\varrho(x)y [\varrho(x) \geq 0]$ we get:

$$\int_a^b \varrho(x) y_m(x) y_n(x) dx = 0 \quad (\lambda_m \neq \lambda_n), \quad (15.5-7)$$

which signifies that the system of eigenfunctions $y_1(x), y_2(x), \dots$ is orthogonal in (a, b) with the weight function $\varrho(x)$.

Specializing further by putting

$$\mathcal{L}_1[y] = [\sigma(x)y']' \quad (\sigma(x) \geq 0),$$

we come to the case of a Sturm-Liouville eigenvalue problem associated with the differential equation

$$[\sigma(x)y']' = -\lambda \varrho(x)y \quad (a < x < b). \quad (15.5-8)$$

[Cf. (10.6-1)—(10.6-2).]

15.54 In addition, we have also to touch upon a relevant and very useful connection between eigenvalue problems pertaining to self-adjoint linear

differential operators of a single variable and so-called *integral equations* of FREDHOLM's type. The latter are usually written in the form:

$$y(x) = \lambda \int_a^b K(x, t)y(t) dt + f(x),$$

$K(x, t)$ and $f(x)$ being certain known functions, $y(x)$ denoting the unknown function and λ a real parameter.

If a differential equation of the form

$$\mathcal{L}[y] = q(x) - \lambda y \quad [y = y(x); a \leq x \leq b]$$

is given together with some boundary conditions, where $q(x)$ is known and $\mathcal{L}[y]$ is an operator of the type (15.5-1) self-adjoint in $[a, b]$, we can always construct a suitable auxiliary function $K(x, t)$ — the so-called *Green function* — by means of which the eigenvalue problem in question can be transformed into an integral equation of the above form, equivalent to the problem. This $K(x, t)$ fulfils for fixed t the boundary conditions and satisfies for $x \neq t$ also the homogeneous equation $\mathcal{L}[y] = 0$; moreover it is symmetric with respect to x and t , i.e., $K(x, t) = K(t, x)$.

15.6 Partial differential equations and orthogonal series

15.61 For applications in natural sciences and engineering, the *boundary problems relating to partial differential equations* and the orthogonal systems, as well as series connected with them are also of most pressing importance. Here a short survey of some typical problems and results will be given.

From a historical point of view, the most classical pertinent example is undoubtedly the "*problem of vibrating string*", which was first investigated by D. BERNOULLI, J.D'ALEMBERT and L.EULER towards the middle of the 18th century: determine the deviation of any point of abscissa x along a *homogeneous* string of length l at an arbitrary point of time t , if the string is in transversal plane vibration with fixed end points, further the initial deviation and velocity are known at all of its points.

Is one of the end points of the string at the origin and the initial point of time is taken $t = 0$, then we get by mechanical considerations for the deviation $u = u(x, t)$ (at least approximately) the partial differential equation of second order

$$\frac{\partial^2 u}{\partial t^2} = \gamma^2 \frac{\partial^2 u}{\partial x^2} \quad (\gamma \neq 0 \text{ constant}), \quad (15.6-1)$$

connected with the *boundary conditions*:

$$u(0, t) = u(l, t) = 0 \quad (\text{for every } t > 0) \quad (15.6-2)$$

and the *initial conditions*:

$$u(x, 0) = u_0(x), \quad u'_t(x, 0) = u_1(x) \quad (0 \leq x \leq l), \quad (15.6-3)$$

where $u_0(x)$ and $u_1(x)$ denote prescribed functions, continuous and piecewise monotonic in $[0, l]$.

D'Alembert and Euler first found the general solution of equation (15.6-1) by the formula:

$$u = \xi(x + \gamma t) + \eta(x - \gamma t),$$

where $\xi = \xi(\tau)$, $\eta = \eta(\tau)$ denote arbitrary functions of one variable (1747-1748). Another solving method published by D. Bernoulli a few years later, however, corresponds better to the nature of the problem. This is the well-known "*separation process*" whose basic ideas are to look for particular solutions written as products of two factors depending on x and t separately, i.e., in the form

$$u = \Phi(x) \cdot \Psi(t); \quad (15.6-4)$$

and after to construct on the basis of the homogeneity of the differential equation such an infinite linear combination with constant coefficients of the latter which satisfies not only (15.6-1) but also the accessory conditions (15.6-2)–(15.6-3). The second requirement will be called "*superposition principle*".

Since by substituting (15.6-4) into (15.6-1) and by separation of the variables we get the equation

$$\frac{\Psi''(t)}{\Psi(t)} = \gamma^2 \frac{\Phi''(x)}{\Phi(x)},$$

which can hold for arbitrary $t > 0$ and $x \in [0, l]$ only if the left- and right-hand term are both equal to some constant $\omega \neq 0$, it yields for both Ψ and Φ the differential equation of harmonic oscillation:

$$\Psi''(t) = -\omega^2 \Psi(t), \quad \Phi''(x) = -\left(\frac{\omega}{\gamma}\right)^2 \Phi(x).$$

The general solutions of these equations can be given at once e.g. by the expressions:

$$\Psi(t) = C_1 \sin(\omega t) + C_2 \cos(\omega t)$$

and

$$\Phi(x) = c_1 \sin\left(\frac{\omega}{\gamma} x\right) + c_2 \cos\left(\frac{\omega}{\gamma} x\right),$$

respectively, where C_1, C_2, c_1, c_2 are arbitrary constants; hence any function of the form

$$u = [C_1 \sin(\omega t) + C_2 \cos(\omega t)] [c_1 \sin(\frac{\omega}{\gamma} x) + c_2 \cos(\frac{\omega}{\gamma} x)]$$

satisfies (15.6-1).

Lastly, using the superposition principle and Dirichlet's criterion (cf. 12.33) we see that the problem of vibrating string has in our case a unique solution, namely the function in question can be represented as the sum of an ordinary Fourier series.¹

15.62 A natural extension of the preceding subject is the problem of vibration of the inhomogeneous string, whose deeper discussion was accomplished in the thirties of the 19th century, i.e., in the period after Fourier's pioneering activity.

If the material of a vibrating string is not uniformly distributed, but it can be characterized locally by a continuous density function $\rho(x)$ and elasticity (or flexibility) of the string is described by another (smooth) function $\sigma(x)$, then the deviation u of any point as a function of x and t satisfies the partial differential equation

$$[\sigma(x) u'_x]'_x = \rho(x) u''_{tt} \quad (x \in [0, l]; t \geq 0); \quad (15.6-5)$$

we have to solve (15.6-5) together with boundary and initial conditions corresponding to the homogeneous case.

The equation obtained by the substitution

$$u = \varphi(x) \psi(t)$$

"separates" on the basis of a reasoning like the above into the ordinary differential equations

$$\psi''(t) = -\omega^2 \psi(t), \quad [\sigma(x) \varphi'(x)]' = -\omega^2 \rho(x) \varphi(x);$$

from which the second is of Sturm-Liouville type, while the first one is especially again the equation describing the harmonic oscillation. Thus the fulfillment of the accessory conditions necessitates Sturm-Liouville expansions, and the series form of the unique solution is also written in terms of Sturm-Liouville eigenfunctions. (Cf. 12.84.)²

¹ This result could not be perceived clearly in the 18th century while the exact foundations of analysis were still lacking. For instance, it was very heatedly disputed, whether a trigonometric series can represent a piecewise linear function.

² Cf. for further investigations: Sz.-Nagy [6] and F. Riesz—Sz.-Nagy [1], Ch. VI.

15.63 Laplace's differential equation (or potential equation) in plane:

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad [u = u(x, y)] \quad (15.6-6)$$

together with its 3-dimensional pendant occurs extraordinarily often in diverse branches of analysis and in numerous fields of applications; its solutions are called *harmonics* or *potential functions*.¹

Observing that, on the basis of the Cauchy-Riemann relations, the real and imaginary parts of any complex power function $w = \zeta^n$ ($\zeta = x + iy = re^{i\vartheta}$; $N = 0, 1, 2, \dots$) are potential functions, by the superposition principle it is easy to see: any convergent series of the form

$$U = \sum_{n=0}^{\infty} r^n (a_n \cos n\vartheta + b_n \sin n\vartheta) \quad (15.6-7)$$

i.e., every so-called *harmonic function series* (cf. 13.53) satisfies formally the variant of (15.6-6) arising by introduction of polar coordinates, that is, the differential equation:

$$\frac{\partial^2 U}{\partial r^2} + \frac{1}{r} \frac{\partial U}{\partial r} + \frac{1}{r^2} \frac{\partial^2 U}{\partial \vartheta^2} = 0 \quad [U = U(r, \vartheta)]. \quad (15.6-8)$$

The same result is yielded by separation $U = r^n V(\vartheta)$.

Note that by substituting in (15.6-7) the coefficients a_n, b_n for the ordinary Fourier constants of a function $f(\vartheta)$ ($0 \leq \vartheta \leq 2\pi$), we can deduce easily the closed form of this series, namely *Poisson's integral* which is known already from 13.51.

The most remarkable boundary problem of the *theory of harmonic functions* (*potential theory*), the *Dirichlet problem*² may be formulated for the unit circle as follows: determine a function $u = u(x, y)$ which is harmonic on the unit disk $x^2 + y^2 < 1$, and at the points of $x^2 + y^2 = 1$ it turns continuously into prescribed values given by a continuous function.

¹ See Laplace [1].

² A classical variation problem called similarly which has important physical aspects, consists in minimalizing the double integral

$$J[u] = \iint_{x^2+y^2 \leq 1} \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right] dx dy$$

with prescribed boundary values of $u = u(x, y)$. (Cf. Courant-Hilbert [1], p.153.) Note that the theory of so-called *subharmonic functions*, due to F. RIESZ, has opened new perspectives to the development of potential theory. (Cf. e.g. F. Riesz [17], Rado [1].)

If we denote the latter function by $f(\vartheta)$ ($0 \leq \vartheta \leq 2\pi$), the problem can essentially be solved through (A)-summation of the ordinary Fourier series of $f(\vartheta)$, more precisely by applying a suitable limiting process to the Poisson integral according to the considerations of 13.52.

15.64 In case of Laplace's equation (potential equation) for the Euclidean 3-space¹ which is usually written as

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0 \quad [u = u(x, y, z)], \quad (15.6-9)$$

the most important particular solutions can be deduced by the coordinate transformation

$$\begin{cases} x = r \sin \vartheta \cos \varphi \\ y = r \sin \vartheta \sin \varphi \\ z = r \cos \vartheta \end{cases}$$

i.e., by introduction of spherical coordinates instead of rectangular Cartesian coordinates, furthermore by applying to the resulting equation for $u = U(r, \vartheta, \varphi)$ the separation $U = r^n V(\vartheta, \varphi)$, also applying the superposition principle.

We get solutions of the form:

$$U = \sum_{n=0}^{\infty} r^n Y_n(\vartheta, \varphi), \quad (15.6-10)$$

where $Y_n(\vartheta, \varphi)$ signifies a linear combination of the $2n + 1$ basic functions

$$\begin{cases} P_n(\cos \vartheta), \\ P_{n,k}(\cos \vartheta) \cos k\varphi, \quad P_{n,k}(\cos \vartheta) \sin k\varphi \\ (k = 1, 2, \dots, n), \end{cases} \quad (15.6-11)$$

$P_n(x)$ denoting the Legendre polynomial of degree n and

$$P_{n,k}(x) = (1 - x^2)^{k/2} [P_n(x)]^{(k)} \quad (15.6-12)$$

the so-called *associated Legendre functions of order k* related to $P_n(x)$. (Cf. 10.32.)

It is not difficult to see that any product $U_n = r^n Y_n(\vartheta, \varphi)$ which satisfies by its origin the equation (15.6-9), i.e., is a potential function, must always be a homogeneous entire rational function of degree n of the variables (x, y, z) .

Owing to the wide-ranging applications in physics, it is justified to introduce some pertinent names: the products U_n are called (ordinary)

¹ Laplace [1].

spherical harmonics of order n , the linear combinations $Y_n(\vartheta, \varphi)$ are the so-called *spherical surface harmonics* (or *Laplace's spherical harmonics*) of order n , and $P_n(\cos \vartheta)$ are called *zonal spherical harmonics* of order n .¹

Note that every Laplace spherical harmonic can be obtained from the simplest potential function

$$u = \frac{1}{r} = (x^2 + y^2 + z^2)^{-1/2}$$

by repeated partial differentiations. (This is the so-called *Maxwell-Sylvester representation*.²) As a further important condition, we mention that the integral over the surface of unit sphere of any two linearly independent Laplace spherical harmonics equals 0; this property is usually expressed by saying that the sequence $\{Y_n(\vartheta, \varphi)\}$ forms a complete orthogonal system on the sphere $r=1$.

On the analogy of the concept of ordinary Fourier series, the expansion of a function $f(\vartheta, \varphi)$, defined on the surface of unit sphere, is realized in a series of the form $\sum_{n=0}^{\infty} Y_n(\vartheta, \varphi)$ with certain surface integrals as coefficients.

Such a "Fourier series in space" is called *Laplace series* of $f(\vartheta, \varphi)$; and its convergence and summation theories, based upon the properties of Legendre polynomials, may be compared with the topics of Chapters 12-13.

Evidently, *Dirichlet's problem* together with its solution can also be transferred to the three-dimensional case.

15.65 The differential equation of linear heat conduction:

$$\frac{\partial u}{\partial t} = \kappa^2 \frac{\partial^2 u}{\partial x^2} \quad (\kappa \neq 0 \text{ constant}) \quad (15.6-13)$$

describes the temperature distribution $u = u(x, t)$ along a thin metal wire of length l , which has at both end points a fixed temperature of 0, x denoting the abscissa of an arbitrary point of the wire and t the time.

In this case we have beside the above boundary conditions:

$$u(0, t) = u(l, t) = 0 \quad (0 \leq t < \infty) \quad (15.6-14)$$

one single initial condition of the form:

$$u(x, 0) = u_0(x) \quad (0 \leq x \leq l), \quad (15.6-15)$$

with a given function u_0 on the right.

¹ Cf. e.g. Courant-Hilbert [1], 441-451.

² Maxwell [1], Sylvester [1].

Since the separation $u(x, t) = F(x) G(t)$ yields the elementary differential equations of second or first order:

$$F''(x) = -\omega^2 F(x), \quad G'(t) = -(\kappa\omega)^2 G(t)$$

$$(\omega \neq 0 \text{ constant}),$$

and for $F \neq 0$ the value of ω is necessarily $n\pi/l$ ($n = 1, 2, \dots$), an easy calculation shows that the solution of the problem may be written as follows:

$$u(x, t) = \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi}{l}x\right) e^{-(\kappa n\pi/l)^2 t} \quad (15.6-16)$$

— if the coefficients a_n are chosen in accordance with (15.6-15), i.e., the representation

$$u_0(x) = \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi}{l}x\right) \quad (0 \leq x \leq l) \quad (15.6-17)$$

holds good.

It should be noted that heat conduction — among similar physical problems — was treated heuristically by FOURIER in his famous work entitled “Théorie analytique de la chaleur” (1822) which gave a very significant impulse to the development of the theory of trigonometric series.¹ In the same framework, we can find a counterpart of the problem which corresponds to the limiting process $l \rightarrow \infty$ and for which the boundary condition belonging to the right-hand end point is to be replaced by the boundedness of $|u(x, t)|$. The solution of this more difficult problem can be presented by means of *Fourier's integral theorem 14.4.2*, considering that $u_0(x)$ is an odd function in the latter case. The result can be expressed by the double integral formula:

$$u(x, t) = \frac{2}{\pi} \int_0^{\infty} \int_0^{\infty} u_0(\tau) e^{-(\kappa\omega)^2 \tau} \sin(\omega\tau) \sin(\omega x) d\omega d\tau. \quad (15.6-18)$$

Finally we point out that (15.6-13) together with the accessory conditions (15.6-14)—(15.6-15) also characterizes the simplest cases of *diffusion* as well as of *soil consolidation*.

15.66 A further related differential equation is that of a *vibrating membrane*:

$$\Delta u = \frac{1}{\gamma^2} \frac{\partial^2 u}{\partial t^2} \quad (\gamma \neq 0 \text{ constant}), \quad (15.6-19)$$

¹ At this point, let us refer also to the introduction of Section 14.4, furthermore to the work [1] of H. POINCARÉ.

where $u = u(x; y; t)$ denotes the deviation of an arbitrary point as a function of its plane rectangular coordinates and time.

If we introduce polar coordinates (r, ϑ) instead of (x, y) , then substituting u by U the equation becomes:

$$\frac{\partial^2 U}{\partial r^2} + \frac{1}{r} \frac{\partial U}{\partial r} + \frac{1}{r^2} \frac{\partial^2 U}{\partial \vartheta^2} = \frac{1}{\gamma^2} \frac{\partial^2 U}{\partial t^2} \quad [U = U(r, \vartheta; t)]. \quad (15.6-20)$$

Here it is advisable to apply the method of separation twice: write first $U = \Phi(r, \vartheta) \cdot \Psi(t)$, then in the differential equation obtained for Φ we put $\Phi(r, \vartheta) = F(r) \cdot G(\vartheta)$. The resulting particular solutions contain beside trigonometric functions also *Bessel functions*. (Cf. 10.64.) The accessory conditions are: the boundary of the membrane is fixed, and the initial deviation and velocity of any inner point are characterized by certain prescribed functions of two variables.

We have still to emphasize that a three-dimensional analogue of (15.6-19), the so-called *wave equation*:

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} = \frac{1}{\mu^2} \frac{\partial^2 \Psi}{\partial t^2}, \quad (15.6-21)$$

where Ψ is a function of the rectangular Cartesian coordinates (x, y, z) as well as of time t and $\mu \neq 0$ is a suitable constant, occurs frequently in optics, electro- and hydrodynamics, general theory of oscillations, quantum theory; moreover that (15.6-21) turns by the substitution

$$\Psi = \psi(x, y, z) e^{-2\pi i Et/h} \quad (15.6-22)$$

into the so-called *SCHRÖDINGER equation*:

$$\Delta \psi + \frac{8\pi^2 m_0}{h^2} (E - V) \psi = 0 \quad (\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}), \quad (15.6-23)$$

m_0 denoting the mass of electron, h being Planck's constant, E the intensity of field and $V = V(x, y, z)$ the potential energy. As is well-known, (15.6-23) forms the basis of wave mechanics. (Cf. 15.34.)

* * *

Three of the above examples: the equation of a homogeneous vibrating string, Laplace's equation in the plane and the equation of linear heat conduction are of the following common type:

$$a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial^2 u}{\partial x \partial y} + c \frac{\partial^2 u}{\partial y^2} + a_0 \frac{\partial u}{\partial x} + b_0 \frac{\partial u}{\partial y} + c_0 u = 0 \quad (15.6-24)$$

$$[u = u(x, y)],$$

i.e., they are linear homogeneous partial differential equations of second order, with constant coefficients.

Applying the usual classification of differential equations of the form (15.6-24), based on the corresponding geometric interpretation, we see that (15.6-1) is *hyperbolic* ($b^2 - 4ac > 0$), while (15.6-6) represents the *elliptic* ($b^2 - 4ac < 0$) and (15.6-13) the *parabolic* ($b^2 - 4ac = 0$) class. This theoretical background enables us to extend the main characteristic features of the pertaining solutions to wider classes of multidimensional equations of second or fourth order.

15.7 Solving integral equations by operator methods

15.71 The gradual formation of functional analysis ran in the early 1900's parallel with the intensive development of the theory of integral equations. (D. HILBERT, E. SCHMIDT, V. VOLTERRA, Scandinavian analytic school). The connection between these fields has been just as close as the interdependence of functional analytic methods with the general theory of orthogonal series (cf. 9.1). Hence the investigation of integral equations also yield in recent times an attractive application possibility of modern operational tools.¹

As a relevant and typical example, we are going to discuss the *linear integral equation* due to I. FREDHOLM (1903),² which was already mentioned in 15.54:

$$y(x) = \lambda \int_a^b K(x, t) y(t) dt + f(x) \quad (a \leq x \leq b). \quad (15.7-1)$$

Here we need following specifications:

1. The so-called *kernel function* $K(x, t)$ is *square integrable* as a function of two variables for $(x, y) \in [a, b] \times [a, b]$, moreover $K(x, t)$ has the same property in $[a, b]$ for any fixed x as a function of t as well as for any fixed t as a function of x .

¹ Cf. Hilbert [1], Schmidt [2], Volterra [2].

² Of course, Fredholm has used in his work [1] merely the apparatus of classical linear algebra, combined with certain limiting processes.

2. The given function $f(x)$ belongs also to $L^2(a, b)$.

(15.7-1) is said to be *homogeneous* or *inhomogeneous* according to $f = 0$ and $f \neq 0$, respectively. Throughout this section, all the integrals are to be taken in Lebesgue's sense, and the unknown function y is supposed to be square integrable as well.

We start from the observation: (15.7-1) may also be regarded as an equation relating to the *linear integral operator*

$$\mathcal{K}y = \mathcal{K}_x[y] = \int_a^b K(x, t)y(t)dt \quad [y \in L^2(a, b)], \quad (15.7-2)$$

which plainly maps the Hilbert space $L^2(a, b)$ onto itself. Since by the Buniakowsky-Schwarz inequality we have the estimate:

$$\|\mathcal{K}_x[y]\|^2 \leq \|y\|^2 \int_a^b \int_a^b [K(x, t)]^2 dx dt,$$

so the operator exists, is *bounded* and has the *norm*:

$$\|\mathcal{K}\| \leq \left\{ \int_a^b \int_a^b [K(x, t)]^2 dx dt \right\}^{1/2} < \infty \quad (15.7-3)$$

(cf. 3.13). Using $\mathcal{K}$ and the unit operator I , (15.7-1) can be briefly written in the form

$$(I - \lambda\mathcal{K})y = f. \quad (15.7-4)$$

Let us call

$$K^*(x, t) = K(t, x) \quad (15.7-5)$$

the *adjoint* (or *transposed*) of the kernel function $K(x, t)$, and transfer this denomination and notation also to the corresponding linear operators (17.5-2). Thus $\mathcal{K}$ is *self-adjoint* ($\mathcal{K} = \mathcal{K}^*$) if and only if the kernel function is *symmetric* with respect to x and t , i.e., $K(x, t) = K(t, x)$.

15.72 On the basis of Banach's fixed point theorem 15.2.1 and the expansion theorem 15.4.2, it is now easy to show that our conditions assure the existence of a uniquely determined, square integrable solution $y(x)$ of (15.7-1) for $|\lambda|$ small enough, moreover that this solution can effectively be given in the form of an L^2 -convergent series. (Cf. 9.51.)

15.7.1 (GENERALIZED EXISTENCE AND EXPANSION THEOREM OF C. NEUMANN)¹ If the above conditions 1-2. are fulfilled and λ is chosen in such a way that

$$|\lambda| < \left\{ \int_a^b \int_a^b [K(x, t)]^2 dx dt \right\}^{-1/2}, \quad (15.7-6)$$

¹ Cf. C.G. Neumann [1]. The right-hand series in (15.7-7) is called usually *Neumann's series*.

then there exists a function $y(x) \in L^2(a, b)$ satisfying the integral equation (15.7-1).

This solution is uniquely represented by the expansion:

$$y(x) = f(x) + \sum_{n=1}^{\infty} \lambda^n \int_a^b K^{[n]}(x, t) f(t) dt \quad [L^2(a, b)], \quad (15.7-7)$$

where the so-called n -th iterated kernel $K^{[n]}$ is defined by the recurrence formula:

$$\left. \begin{aligned} K^{[1]}(x, t) &= K(x, t), \\ K^{[n+1]}(x, t) &= \int_a^b K(x, t) K^{[n]}(x, t) dt \\ (n &= 1, 2, \dots). \end{aligned} \right\} \quad (15.7-8)$$

P r o o f. The first part of the verification of 15.4.2 implies that by 15.2.1 the form (15.7-4) of (15.7-1) is solvable uniquely, if any pair of functions $y_1 \in L^2(a, b)$, $y_2 \in L^2(a, b)$ fulfils the estimate

$$\|\lambda \mathcal{K} y_1 - \lambda \mathcal{K} y_2\| \leq q \cdot \|y_1 - y_2\|$$

with a suitable factor $q \in (0, 1)$.

The latter condition is obviously satisfied whenever (15.7-6) holds, because from the Buniakowsky-Schwarz inequality we get

$$\|\lambda \mathcal{K} y_1 - \lambda \mathcal{K} y_2\| \leq |\lambda| \left\{ \int_a^b \int_a^b [K(x, t)]^2 dx dt \right\}^{1/2} \cdot \|y_1 - y_2\|,$$

and the factor before $\|y_1 - y_2\|$ is < 1 in our case.

Hence we see that $\mathcal{T} = \lambda \mathcal{K}$ is such a bounded linear mapping of the Banach space $L^2(a, b)$ onto itself whose norm is smaller than 1; thus we may apply to this operator 15.4.2 and obtain as the solution of (15.7-4):

$$\begin{aligned} y &= (I - \lambda \mathcal{K})^{-1} = \left(\sum_{n=0}^{\infty} \lambda^n \mathcal{K}^n \right) f = \\ &= f(x) + \sum_{n=1}^{\infty} \lambda^n \mathcal{K}^n [f] \quad [L^2(a, b)]. \end{aligned} \quad (15.7-9)$$

But by (15.7-8), the n -th iterated kernel can be written as $\mathcal{K} [K^{[n-1]}]$ where $K^{[n-1]}$ is to be regarded as a function of t ($n = 2, 3, \dots$). Therefore induction leads to the formula:

$$\mathcal{K}^n[f] = \int_a^b K^{[n]}(x, t) f(t) dt \quad (15.7-10)$$

$$(n = 1, 2, \dots),$$

by means of which (15.7-9) becomes (15.7-7).

* * *

As far as the interconnection between solutions of

$$(I - \lambda \mathcal{K})y = f \quad (f \neq 0) \quad (15.7-11)$$

(inhomogeneous case) and of

$$(I - \lambda \mathcal{K})y = 0, \quad \text{i.e.} \quad y = \lambda \mathcal{K}y \quad (15.7-12)$$

(homogeneous case) is concerned, for orientation we insert the pertinent deep theorem, which is wholly analogous to the corresponding classical results on systems of linear equations.

15.7.2 (FREDHOLM'S ALTERNATIVE)¹ (I) If the equation (15.7-12) attached to (15.7-11) has only trivial solution ($y = 0$), then (15.7-11) is uniquely solvable in case of any given $f \neq 0$.

(II) Has the equation (15.7-12) attached to (15.7-11) a non-trivial solution, then the number of its all linearly independent solutions is finite and equal to the number of linearly independent solutions of the equation containing the adjoint kernel, i.e., of

$$y = \lambda \mathcal{K}^* y. \quad (15.7-13)$$

Furthermore, in this case the equation (15.7-11) is solvable if and only if the function f on the right is orthogonal to every solution of (15.7-13).

15.7.3 The above operational method renders possible a particularly elegant discussion of homogeneous linear integral equations, if the kernel function is *symmetric*. (This is called *Hilbert-Schmidt theory*.) This is due to the fact that the condition $K(x, t) = K(t, x)$ involves not only the *self-adjointness* but also the *total continuity of the operator* $\mathcal{K}$ in the sense defined in 15.32; and this circumstance is plainly advantageous from point of view of the *eigenvalue problem* associated with the operator equation (15.7-12).

¹ Cf. F. Riesz [11], [13].

In fact, after writing the integral equation

$$y(x) = \lambda \int_a^b K(x, t) y(t) dt \quad [K \neq 0, y \in L^2(a, b)] \quad (15.7-14)$$

in the form

$$\mathcal{K}y = \frac{1}{\lambda} y \quad (\lambda \neq 0), \quad (15.7-15)$$

we may apply immediately the statements about the equation (15.3-4) in Hilbert's space, together with the spectral decomposition theorem 15.3.2.

Thus we get in the case

$$K(x, t) = K(t, x) \quad (x \in [a, b], t \in [a, b]): \quad (15.7-16)$$

I. (15.7-14) has countably many distinct real *eigenvalues*, say $\lambda_1, \lambda_2, \dots$ with $|\lambda_n| \uparrow \infty$;

II. the *eigenfunctions* pertaining to different eigenvalues are *orthogonal* by pairs, so that we can form of them a complete orthonormal system $y_n(x)$ ($n = 1, 2, \dots$).

If the eigenvalue belonging to $y_n(x)$ is denoted by λ_n , it results finally as a corollary:

15.7.3 (HILBERT—SCHMIDT EXPANSION THEOREM)¹ *Any integral operator with symmetric kernel function:*

$$\mathcal{K}_x[y] = \int_a^b K(x, t) y(t) dt \quad [y \in L^2(a, b)] \quad (15.7-17)$$

can be represented by a Fourier series in terms of the eigenfunctions $\{y_n\}$ according to L^2 -convergence, i.e., we have

$$\mathcal{K}_x[y] = \sum_{n=1}^{\infty} \gamma_n y_n(x) \quad [L^2(a, b)], \quad (15.7-18)$$

where the coefficients depend also on the eigenvalues $\lambda_n \neq 0$:

$$\gamma_n = \frac{1}{\lambda_n} \int_a^b y(t) y_n(t) dt \quad (n = 1, 2, \dots). \quad (15.7-19)$$

¹ Cf. Hilbert [1], Schmidt [1]–[3], further F. Riesz [2] and [11]. For an exhaustive treatment of the theory of integral equations let us refer e.g. to the functional analysis [1] of F. Riesz and B. Sz. Nagy, while for the main applications of the theory we mention S. G. Mikhlin [1].

BIBLIOGRAPHY

(*) A list of standard works which may be particularly recommended to readers for further studying the subjects touched on here:

- ALEXITS, G. *Convergence Problems of Orthogonal Series*, 2.ed. New York-Oxford-London-Paris: Pergamon Press, 1961.
- COLLATZ, L. *Funktionalanalysis und numerische Mathematik*, Berlin-Göttingen-Heidelberg: Springer, 1964.
- DIEUDONNÉ, J. *Foundations of Modern Analysis*, 7.ed. New York-London: Academic Press, 1968.
- DUNFORD, N.-SCHWARTZ, J.T. *Linear Operators* I. (General theory), New York, 1958. II. (Spectral theory), New York-London, 1963. Interscience Publ.
- FOIAS, C.-SZŐKEFALVI-NAGY, B. *Harmonic Analysis of Operators on Hilbert Space*, Amsterdam-London: North-Holland Publ., 1970.
- HALMOS, P.R. *Measure Theory*, 10.ed. Princeton: Van Nostrand, 1965.
- HEWITT, E.-STROMBERG, K. *Real and Abstract Analysis*, 2.ed. Berlin-Heidelberg-New York: Springer, 1969.
- KOLMOGOROV, A.N.-FOMIN, S.V. *Elements of the Theory of Functions and Functional Analysis* I. (Metric and normed spaces), New York: Graylock, 1957. II. (Measure, Lebesgue integrals and Hilbert space), New York-London: Academic Press, 1961.
- KURATOWSKI, C. *Introduction to Set Theory and Topology*, Oxford-Paris-Frankfurt, Pergamon Press and Reading (Mass.): Addison-Wesley, 1962.
- NEUMANN, J. *Functional Operators* I. (Measures and integrals), II. (The geometry of orthogonal spaces), Princeton: University Press, 1950.
- RIESZ, F.-SZŐKEFALVI-NAGY, B. *Functional Analysis*, 3.ed. New York: Ungar, 1960.
- SAKS, S. *Theory of Integral*, 2.ed. New York: Dover Publ., 1964.
- SANSONE, G. *Orthogonal Functions*, 2.ed. New York: Interscience Publ. 1959.
- TITCHMARSH, E.C. *The Theory of Functions*, 2.ed. Oxford: University Press, 1952.
- TRICOMI, E.G. *Vorlesungen über Orthogonalreihen*, 2.ed. Berlin-Heidelberg-New York: Springer, 1970.
- YOSIDA, K. *Functional Analysis*, 2.ed. Berlin-Heidelberg-New York: Springer, 1968.
- ZYGMUND, A. *Trigonometric Series* I.-II., 2.ed. Cambridge: University Press, 1968.

(**) Original papers and memoirs referred to in the text:

- ABEL, N.H. [1] Untersuchungen über die Binomialreihe. *Journal f. reine u. angew. Math.*, 1 (1826), 311-339. (Oeuvres I, 219-250.)
- AKHIESER, N. [1] *Theory of Approximation* (New York, 1956).
- AL-BASSAM, M.A. [1] Some applications of fractional calculus to differential equations. *Research Notes Math.*, Vol. 138, Pitman (Boston-London-Melbourne, 1985), 1-11.
- ALEXITS, G. [1] Sur le sommabilité des séries orthogonales. *Acta Math. Acad. Sci. Hung.*, 4 (1953), 181-188.

- [2] *Convergence Problems of Orthogonal Series* (New York-Oxford-London-Paris, 1961).
- ALEXITS, G.-TANDORI, K. [1] Über das Konvergenzverhalten einer Klasse von Orthogonalreihen. *Annales Univ. Sci. Budapest. Sectio Math.*, 3-4 (1960-1961), 15-18.
- ARZELÀ, C. [1] Sulla integrazione per serie. *Rend. Accad. Lincei Roma*, 1 (1885), 532-537, 566-569.
- ATKINSON, F.V. [1] *Discrete and Continuous Boundary Problems* (New York-London, 1964).
- BAIRE, R. [1] Sur les fonctions de variables réelles. *Annali mat. pure e appl.*, (3) 3 (1899), 1-122.
- BANACH, S. [1] *Théorie des opérations linéaires* (Warsaw, 1932).
- BENDIXSON, I. [1] Quelques théorèmes de la théorie des ensembles de points. *Acta Math.*, 2 (1883), 414-429.
- BESSEL, F.V. [1] Untersuchung des Teils der planetarischen Störungen, welcher aus der Bewegung der Sonne entsteht. *Abhandl. Akad. Berlin*, 1824, 1-52. (Werke I, 84-109.)
- BOREL, E. [1] Sur la sommation de séries divergentes. *Comptes Rendus Acad. Sci. Paris*, 121 (1895), 1125-1127.
[2] *Leçons sur la théorie des fonctions* (Paris, 1898; 2.ed. 1914).
- BRUCKNER, A.M. [1] The differentiability properties of typical functions in $C[a, b]$. *Amer. Math. Monthly*, 80 (1973), 679-683.
- BURKHARDT, H. [1] *Trigonometrische Reihen und Integrale bis etwa 1850*. (Enzykl. math. Wiss. II, A 12, 819-1354.)
[2] *Trigonometrische Interpolation*. (Enzykl. math. Wiss. II, A 9a, 642-693.)
- BURKILL, J.C. [1] Functions of intervals. *Proc. London Math. Soc.*, (2) 22 (1924), 275-310.
- CAMPOS, L.M.B.C. [1] On the properties of diffintegrations and branch-point derivatives, including annihilation. *Proc. III. Internat. Conf. "Fractional Calculus"* (Tokyo, 1989), Descartes Press, 1990, 25-33.
- CANTOR, G. [1] Beweis, dass eine für jeden Wert von x durch eine trigonometrische Reihe gegebene Funktion $f(x)$ sich nur auf eine einzige Weise in dieser Form darstellen lässt. *Journal f. reine u. angew. Math.*, 72 (1870), 139-143.
[2] Über lineare Punktmannigfaltigkeiten. *Math. Annalen*, 21 (1883), 564-569.
[3] Sur divers théorèmes de la théorie des ensembles de points ... *Acta Math.*, 2 (1883), 409-413.
- CARATHÉODORY, C. [1] *Vorlesungen über reelle Funktionen* (Leipzig-Berlin, 2.ed. 1927).
- CARLEMAN, T. [1] A theorem concerning Fourier series. *Proc. London Math. Soc.*, (2) 21 (1923), 483-492.
- CARLESON, L. [1] On convergence and growth of partial sums of Fourier series, *Acta Math.*, 116 (1966), 135-157.
- CAUCHY, A.L. [1] *Cours d'analyse de l'École Polytechnique, I.: Analyse algébrique* (Paris, 1821.)
- CESÀRO, E. [1] Sur la multiplication des séries. *Bull. des sciences math.*, (2) 14 (1890), 114-120.
- CHAPMAN, S. [1] On non-integral orders of summability of series and integrals. *Proc. London Math. Soc.*, (2) 9 (1911), 369-409.
- COHEN, P.J. [1] *Set Theory and the Continuum Hypothesis* (New York, 1966).
- COURANT, R.-HILBERT, D. [1] *Methoden der mathematischen Physik, I.* (Berlin-Heidelberg-New York, 3.ed. 1968).
- CSÁSZÁR, Á. [1] *Foundations of General Topology* (Oxford-London-New York-Paris, 1963).
- DANIELL, P.J. [1] A general form of integral. *Annals of Math.*, 19 (1917-18), 279-294.

- DE LA VALLÉE-POUSSIN, CH. J. [1] Sur l'unicité du développement trigonométrique. *Bull. Acad. Royale de Belgique*, 1912, 702–718.
- DENJOY, A. [1] Une extension de l'intégrale de M. Lebesgue. *Comptes Rendus Acad. Sci. Paris*, **154** (1912), 859–862.
[2] Mémoire sur les nombres dérivés des fonctions continues. *Journal de Math.*, (7) **1** (1915), 105–240.
- DIEUDONNÉ, J. [1] Sur une fonction continue sans dérivée. *Mathesis*, **47** (1933), 277–279.
- DINI, U. [1] *Serie di Fourier e altre rappresentazioni analitiche delle funzioni di una variabile reale* (Pisa, 1880).
- DIRAC, P. [1] *Principles of Quantum Mechanics* (Oxford, 2.ed. 1936).
- DIRICHLET, L. G. [1] Sur la convergence des séries trigonométriques. *Journal f. reine u. angew. Math.*, **4** (1829), 157. (Werke I, 129.)
- DOETSCH, G. [1] *Handbuch der Laplace-Transformation*. I–III. (Basel, 1950–56.)
- DU BOIS-REYMOND, P. [1] Untersuchungen über die Konvergenz und Divergenz der Fourierschen Darstellungsformeln. *Abhandl. Akad. München*, **12** (1876), 1–103.
[2] Beweis, dass die Koeffizienten der trigonometrischen Reihe die Fourierschen Werte haben, jedesmal, wenn diese Integrale endlich und bestimmt sind. *Abhandl. Akad. München*, **12** (1876), 117–166.
- EGOROV, D. T. [1] Sur les suites de fonctions mesurables. *Comptes Rendus Acad. Sci. Paris*, **152** (1911), 244–246.
- EPSTEIN, P. [1] Zur Theorie allgemeiner Zetafunktionen, I–II. *Math. Annalen*, **56** (1903), 615–644; **63** (1907), 205–216.
- ERDŐS, P.–TURÁN, P. [1] On interpolation I–III. *Annals of Math.*, **38** (1937), 142–155; **39** (1938), 703–724; **41** (1940), 510–553.
- FABER, G. [1] Einfaches Beispiel einer stetigen nirgends differenzierbaren Funktion. *Jahresber. Deutsch. Math. Verein.*, **16** (1907), 538–540.
[2] Über stetige Funktionen II. *Math. Annalen*, **69** (1910), 372–443.
- FATOU, P. [1] Séries trigonométriques et séries de Taylor. *Acta Math.*, **30** (1906), 335–400.
- FEJÉR, L. [1] Sur les fonctions bornées et intégrables. *Comptes Rendus Acad. Sci. Paris*, **131** (1900), 984–987.
[2] Sur la différentiation de la série de Fourier. *Comptes Rendus Acad. Sci. Paris*, **134** (1902), 762–765.
[3] Untersuchungen über Fouriersche Reihen. *Math. Annalen*, **58** (1904), 501–569.
[4] Beispiele stetiger Funktionen mit divergenter Fourierreihe. *Journal f. reine u. angew. Math.* **137** (1909), 1–5.
[5] Lebesguesche Konstanten und divergente Fourierreihen, *Journal f. reine u. angew. Math.*, **138** (1910), 22–53.
[6] Sur les singularités de la série de Fourier des fonctions continues. *Annales Éc. Norm. Sup.* **28** (1911), 64–103.
[7] Über die Bestimmung des Sprunges der Funktion aus ihrer Fourierreihe. *Journal f. reine u. angew. Math.*, **142** (1913), 165–188.
[8] Über konjugierte trigonometrische Reihen. *Journal f. reine u. angew. Math.*, **144** (1914), 48–56.
[9] Über die arithmetischen Mittel erster Ordnung der Fourierreihe. *Göttinger Nachrichten*, **1925**, 13–17.
[10] On new properties of arithmetical means of the partial sums of Fourier series. *Journal of Math. and Phys.* **13** (1934), 1–17.
[11] Zur Summabilitätstheorie der Fourierschen und Laplaceschen Reihe. *Proc. Cambridge Phil. Soc.*, **34** (1938), 503–509.
[12] Intégrales singulières à noyau positif. *Comm. Math. Helv.* **23** (1949), 177–199.
- FELLER, W. [1] *An Introduction to Probability Theory and its Applications* (New York–London, 1950).

- FISCHER, E. [1] Sur la convergence en moyenne. *Comptes Rendus Acad. Sci. Paris*, 144 (1907), 1022-1024, 1148-1150.
- FOURIER, J.B. [1] *Théorie analytique de la chaleur* (Paris, 1822).
- FRAENKEL, A.A. [1] *Abstract Set Theory* (Amsterdam, 2.ed. 1961).
- FRANKLIN, P. [1] A set of continuous orthogonal functions. *Math. Annalen*, 100 (1928), 522-529.
- FRÉCHET, M. [1] Sur les ensembles de fonctions et les opérations linéaires. *Comptes Rendus Acad. Sci. Paris*, 144 (1907), 1414-1416.
[2] Essai de géométrie analytique à une infinité des coordonnées. *Nouvelles Annales de math.*, (4) 8 (1908), 97-116; 289-317.
- FREDHOLM, I. [1] Sur une classe d'équations fonctionnelles. *Acta Math.*, 27 (1903), 365-390.
- FREUD, G. [1] *Orthogonale Polynome* (Budapest, 1969).
- FUBINI, G. [1] Sugli integrali multipli. *Rend. Accad. Lincei Roma*, 16 (1907), 608-614.
[2] Sulla derivazione per serie. *Rend. Accad. Lincei Roma*, 24 (1915), 204-206.
- GIBBS, J.W. [1] *Scientific Papers*, II. (New York-Bombay, 1906).
- GRAM, J.P. [1] Über die Entwicklung reeller Funktionen in Reihen mittelst der Methode der kleinsten Quadrate. *Journal f. reine u. angew. Math.*, 94 (1883), 21-73.
- HAAR, A. [1] Zur Theorie der orthogonalen Funktionensysteme, I.-II. *Math. Annalen*, 69 (1910), 331-371; 71 (1912), 33-53.
[2] Der Massbegriff in der Theorie der kontinuierlichen Gruppen. *Annals of Math.*, 34 (1933), 147-169.
- HAHN, H. [1] Über stetige Funktionen ohne Ableitung. *Jahresber. Deutsch. Math. Verein.*, 26 (1918), 281-284.
- HAHN, W. [1] Über die Jacobischen Polynome und zwei verwandte Polynomklassen. *Math. Zeitschrift*, 39 (1935), 634-638.
- HARDY, G.H. [1] Theorems relating to the summability and convergence of slowly oscillating series. *Proc. London Math. Soc.* (2) 8 (1910), 301-320.
[2] On the summability of Fourier series. *Proc. London Math. Soc.* (2) 13 (1913), 13-28.
[3] Weierstrass's non-differentiable function. *Trans. Amer. Math. Soc.*, 17 (1916), 301-325.
[4] On certain criteria for the convergence of the Fourier series of a continuous function. *Messenger of Math.*, 49 (1920), 149-155.
[5] On the integration of Fourier series. *Messenger of Math.*, 51 (1922), 186-192.
[6] Note on Lebesgue's constants in the theory of Fourier series. *Journal London Math. Soc.*, 17 (1942), 4-13.
[7] *Divergent Series* (Oxford, 1949).
- HARDY, G. H.-LITTLEWOOD, J.E. [1] Sur la série de Fourier d'une fonction à carré sommable. *Comptes Rendus Acad. Sci. Paris*, 156 (1913), 1307-1309.
[2] Some new convergence criteria for Fourier series. *Journal London Math. Soc.*, 7 (1932), 252-256.
[3] On the strong summability of Fourier series. *Fundamenta Math.*, 25 (1935) 162-189.
- HARDY, G.H.-LITTLEWOOD, J.E.-PÓLYA, G. [1] *Inequalities* (Cambridge, 2.ed. 1954).
- HARDY, G.H.-RIESZ M. [1] *The General Theory of Dirichlet's Series* (Cambridge, 1915; 2.ed. 1952).
- HARDY, G.H.-ROGOSINSKI, W.W. [1] *Fourier Series* (Cambridge, 2.ed. 1950).
- HERMITE, C.H. [1] Sur un nouveau développement en série de fonctions. *Comptes Rendus Acad. Sci. Paris*, 58 (1864), 93-100, 266-273. (Oeuvres II, 293-308.)
- HILB, E.-RIESZ, M. [1] *Neuere Untersuchungen über trigonometrische Reihen*. (Enzykl. math. Wiss. II. C 10, 1189-1228.)

- HILBERT, D. [1] *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen* (Leipzig, 1912).
- HILLE, E. [1] *Functional Analysis and Semi-groups* (New York, 1948).
- HOBSON, E.W. [1] *The Theory of Functions of a Real Variable I-II* (Cambridge, I: 3.ed. 1927, II: 2.ed. 1926).
- [2] Generalization of a theorem of F. Riesz. *Journal London Math. Soc.*, 1 (1926), 211–218.
- HÖLDER, O. [1] Grenzwerte von Reihen an der Konvergenzgrenze. *Math. Annalen*, 20 (1882), 535–549.
- HUNT, R.A. [1] On the convergence of Fourier series. *Proc. Conf. "Orthogonal expansions and their continuous analogues"* (Edwardsville, Ill., 1967), 235–255.
- HURWITZ, A. [1] Sur quelques applications géométriques des séries de Fourier. *Annales Éc. Norm. Sup.*, (3) 19 (1902), 357–408.
- [2] Über die Fourierschen Konstanten integrierbarer Funktionen. *Math. Annalen*, 57 (1903), 425–446.
- JACKSON, D. [1] *The Theory of Approximation* (New York, 1930; last ed. 1968).
- JACOBI, C.G. [1] Untersuchungen über die Differentialgleichung der hypergeometrischen Reihe. *Journal f. reine u. angew. Math.*, 56 (1859), 149–165.
- JORDAN, C. [1] Sur la série de Fourier. *Comptes Rendus Acad. Sci. Paris*, 92 (1881), 228–230.
- [2] *Cours d'analyse de l'École Polytechn.*, I. (Paris, 2.ed. 1893).
- JORDAN, P.–NEUMANN, J. [1] On inner products in linear metric spaces. *Annals of Math.*, 36 (1935), 719–723.
- KACZMARZ, S. [1] Sur la convergence et la sommabilité des développements orthogonaux. *Studia Math.*, 1 (1929), 87–121.
- KACZMARZ, S.–STEINHAUS, H. [1] *Theorie der Orthogonalreihen* (New York, 2.ed. 1951).
- KAHANE, J.P.–KATZNELSON, Y. [1] Sur les ensembles de divergence des séries trigonométriques. *Studia Math.*, 26 (1966), 305–306.
- KALLA, S.L. [1] Operators of fractional integration. *Lecture Notes in Math.*, Vol. 798, Springer-Verlag (Berlin–Heidelberg, New York, 1980), 258–280.
- KALLA, S.L.–KIRYAKOVA, V.S. [1] A generalized fractional calculus dealing with H -functions. *Proc. III. Internat. Conf. "Fractional Calculus"* (Tokyo, 1989), Descartes Press, 1990, 62–69.
- KAMKE, E. [1] *Das Lebesgue–Stieltjes-Integral* (Leipzig, 1956).
- KHINTCHIN, J.–KOLGOMOROV, A. [1] Konvergenz der Reihen deren Glieder durch den Zufall bestimmt werden. *Recueil Math. Moscou (Mat. Sbornik)*, 22 (1925), 668–677.
- KNOPP, K. [1] Ein einfaches Verfahren zur Bildung stetiger nirgends differenzierbarer Funktionen. *Math. Zeitschrift*, 2 (1918), 1–26.
- [2] *Theorie und Anwendung der unendlichen Reihen* (Berlin–Göttingen–Heidelberg, 4.ed. 1947).
- KOLMOGOROV, A.N. [1] Une série de Fourier–Lebesgue divergente presque partout. *Fundamenta Math.*, 4 (1923), 324–328.
- [2] Une série de Fourier–Lebesgue divergente partout. *Comptes Rendus Acad. Sci. Paris*, 183 (1926), 1327–1328.
- [3] Beiträge zur Masstheorie. *Math. Annalen*, 107 (1932), 351–366.
- [4] Grundbegriffe der Wahrscheinlichkeitsrechnung. *Ergebnisse der Math.* II/3 (Berlin, 1933).
- LAGUERRE, E. [1] Sur l'intégrale $\int_x^\infty t^{-1} e^{-t} dt$. *Bull. Soc. Math. de France*, 7 (1879), 72–81.
- LANDAU, E. [1] Über die Bedeutung einiger neuen Grenzwertsätze der Herren Hardy und Axer. *Prace Mat.–Fiz.*, 21 (1910), 107.
- LAPLACE, P.S. [1] *Mécanique céleste* (Paris, 1825).

- LEBESGUE, H. [1] Sur une généralisation de l'intégrale définie. *Comptes Rendus Acad. Sci. Paris*, **132** (1901), 1025-1028.
- [2] Un théorème sur les séries trigonométriques. *Comptes Rendus Acad. Sci. Paris*, **134** (1902), 584-587.
- [3] *Intégrale, longueur, aire* (thèse, Paris, 1902); *Annali mat. pure e appl.*, (3) **7** (1902), 231-359.
- [4] *Leçons sur l'intégration et la recherche des fonctions primitives* (Paris, 1904).
- [5] Recherches sur la convergence des séries de Fourier. *Math. Annalen*, **61** (1905), 251-280.
- [6] *Leçons sur les séries trigonométriques* (Paris, 1906).
- [7] Sur les intégrales singulières. *Annales de Toulouse*, (3) **1** (1909), 25-117.
- [8] Sur l'intégrale de Stieltjes et sur les opérations linéaires. *Comptes Rendus Acad. Sci. Paris*, **150** (1910), 68-88.
- [9] Sur l'intégration des fonctions discontinues. *Annales Éc. Norm. Sup.* (3) **27** (1910), 361-450.
- [10] Une fonction continue sans dérivée. *Enseignement Math.*, **38** (1942), 212-213.
- LEGENDRE, A.M. [1] Recherches sur l'attraction des sphéroides homogènes. *Mém. math. phys. Acad. Sci. Paris*, **10** (1785), 411-434.
- LEINDLER, L. [1] Über die orthogonalen Polynomsysteme. *Acta Sci. Math. Szeged*, **21** (1960), 19-46.
- LERCH, M. [1] Note sur la fonction $K(w; x.s)$. *Acta Math.*, **11** (1887), 19-24.
- LEVI, B. [1] Sopra integrazione delle serie. *Rend. Inst. Lombardo Sci. e Lett.*, (2) **39** (1906), 775-780.
- LINDELÖF, E. [1] Remarques sur un théorème fondamental de la théorie des ensembles. *Acta Math.*, **29** (1905), 183-190.
- LIOUVILLE, J. [1] Sur le développement des fonctions... *Journal de Math.*, **1** (1836), 253-265.
- [2] Démonstration d'un théorème dû à M. Sturm. *Journal de Math.*, **1** (1836), 269-277.
- LIPSCHITZ, R. [1] De explicatione per series trigonometricas... *Journal f. reine u. angew. Math.*, **63** (1864), 296-308.
- LITTLEWOOD, J.E. [1] The converse of Abel's theorem on power series. *Proc. London Math. Soc.*, (2) **9** (1910), 434-448.
- [2] On Fourier coefficients of functions of bounded variation. *Quarterly Journal of Math.*, **7** (1936), 219-226.
- LOVE, E.R. [1] Fractional derivatives of imaginary order. *Journal London Math. Soc.*, (3) **3** (1971), 241-259.
- [2] A third index law for fractional integrals and derivatives. *Research Notes Math.*, **38** (1985), 63-74.
- [3] Two theorems on Riesz-type fractional integrals. *Proc. III. Internat. Conf. "Fractional Calculus"* (Tokyo, 1989). Descartes Press, 1990., 85-93.
- LOVE, E.R.-HUNTER, M.N. [1] Expansions in series of Legendre functions. *Proc. London Math. Soc.*, (3) **64** (1992), 579-601.
- LUKÁCS, F. [1] Über die Bestimmung des Sprunges einer Funktion aus ihrer Fourierreihe. *Journal f. reine u. angew. Math.*, **150** (1920), 107-112.
- LUSIN, N.N. [1] Sur les propriétés des fonctions mesurables. *Comptes Rendus Acad. Sci. Paris*, **154** (1912), 1688-1690.
- [2] Sur une propriété des fonctions à carré sommable. *Bull. Calcutta Math. Soc.*, **20** (1930), 139-154.
- MANOLITSAKIS, I. [1] Two-dimensional scattered fields: A description in terms of the zeros of entire functions. *J. Math. Phys.*, **23** (1982), 2291-2298.
- MARCINKIEWICZ, J. [1] Sur la sommabilité forte de séries de Fourier. *Journal London Math. Soc.*, **14** (1939), 162-168.

- MAXWELL, J.C. [1] *A Treatise on Electricity and Magnetism, I.* (Oxford, 2.ed. 1881).
- MENSHOV, D. [1] Sur les séries de fonctions orthogonales I-III. *Fundamenta Math.*, 4 (1923), 82–105; 8 (1926), 56–108; 10 (1927), 375–420.
- [2] Sur les séries de fonctions orthogonales bornées dans leur ensemble. *Recueil Math. Moscou (Mat. Sbornik)*, 3 (1938), 103–118.
- [3] Sur les multiplicateurs de convergence pour les séries de polynômes orthogonaux. *Recueil Math. Moscou (Mat. Sbornik)*, 6 (1939), 27–51.
- MIKHLIN, S.G. [1] *Vorlesungen über lineare Integralgleichungen* (Berlin, 1962).
- MIKOLÁS, M. [1] Über die Beziehung zwischen der Gammafunktion und den trigonometrischen Funktionen. *Acta Math. Acad. Sci. Hung.*, 4 (1953), 145–190.
- [2] Über gewisse Eigenschaften orthogonaler Systeme der Klasse L^2 und die Eigenfunktionen Sturm-Liouvillescher Differentialgleichungen. *Acta Math. Acad. Sci. Hung.*, 6 (1955), 147–190.
- [3] Construction des familles de fonctions partout continues, non dérivables. *Acta Sci. Math. Szeged*, 17 (1956), 49–62.
- [4] On the common characterization of Jacobi, Laguerre and Hermite polynomials. (In Hungarian.) *Mat. Lapok*, 7 (1956), 238–246.
- [5] A simple proof of the functional equation of the Riemann zeta-function and a formula of Hurwitz. *Acta Sci. Math. Szeged*, 18 (1957), 261–263.
- [6] Sur la sommation des séries de Fourier au moyen de l'intégration d'ordre fractionnaire. *Comptes Rendus Acad. Sci. Paris*, 251 (1960), 837–839.
- [7] Application d'une nouvelle méthode de sommation aux séries trigonométriques et de Dirichlet. *Acta Math. Acad. Sci. Hung.*, 11 (1960), 317–334.
- [8] Über die Dirichlet-Summation Fourierscher Reihen. *Annales Univ. Sci. Budapest, Sectio Math.*, 3–4 (1960–61), 189–195.
- [9] Über den Riesz-Fischerschen Satz und die Vollständigkeit der Funktionenräume L^p . *Annales Univ. Sci. Budapest, Sectio Math.*, 8 (1965), 159–162.
- [10] New proof and extension of the functional equation of Lerch's zeta-function. *Annales Univ. Sci. Budapest, Sectio Math.*, 14 (1971), 111–116.
- [11] Über die Approximation durch Treppenfunktionen in L^p . *Proc. Conf. "Constructive Theory of Functions"* (Budapest, 1972), 307–314.
- [12] Über die Charakterisierung orthogonaler Systeme von Haarschem Typ. *Publicationes Math. Debrecen*, 19 (1972), 239–249.
- [13] Über die Umkehrung der Fouriertransformation in Falle nicht-absoluter Konvergenz. *Annales Univ. Sci. Budapest, Sectio Math.*, 16 (1973), 155–165.
- [14] On the recent trends in the development, theory and applications of fractional calculus. *Lecture Notes in Math.*, Vol 457. Springer-Verlag (Berlin-Heidelberg-New York, 1975), 357–375.
- [15] On an extension of Fourier's integral theorem. *Coll. Math. Soc. Bolyai*. 19/II (Amsterdam-Oxford-New York, 1978), 537–542.
- [16] Orthogonal systems of characteristic functions and Hilbert space. *Resultate d. Math.-Mathematical Results*, 1 (1978), 195–206.
- [17] On the convergence of generalized Haar expansions. *Periodica Polytechn. (Civil Eng.)*, 27 (1983), 31–38.
- [18] Integro-differential operators and theory of summation. *North-Holland Mathematics Studies*, Nr. 86 (Amsterdam-Oxford-New York, 1984), 245–258.
- [19] A new method of summation, based on fractional integration and generalized zeta-functions. *Proc. III. Internat. Conf. "Fractional Calculus"* (Tokyo, 1989), Descartes Press, 1990., 106–109.
- [20] Historical remarks about the works of M. Riesz in the theory of generalized integro-differential operators and their applications. *Proc. III. Internat. Conf. "Fractional Calculus"* (Tokyo, 1989), Descartes Press, 1990., 110–111.

- [21] Einige neuere Aspekte und analytisch-technische Anwendungen Diophantischer Approximationen. *Results in Mathematics-Resultate der Mathematik*, 18 (1990), 298-305.
- [22] Some recent applications of the kernel function of generalized Weyl fractional integrals. *Periodica Polytechn. (Civil Eng.)*, 36 (1992), 393-406.
- MIKUSINSKI, J. [1] *Operational Calculus* (New York-Oxford-London-Paris, 1959).
- MINKOWSKI, H. [1] *Geometrie der Zahlen*, I. (Leipzig, 1896).
- NATANSON, I.P. [1] *Constructive Theory of Functions* (in Russian), (Moscow-Leningrad, 1949).
- [2] *Theory of Functions of a Real Variable* (New York, 1955).
- NELSON, P.I. [1] A class of orthogonal series related to martingales. *Annals of Math. Stat.*, 41 (1970), 1684-1694.
- NEUMANN, C.G. [1] *Untersuchungen über das logarithmische und Newtonsche Potential* (Leipzig, 1877).
- NISHIMOTO, K. [1] *An Essence of Fractional Calculus*, (Koriyama, Japan, 1991)
- NISHIMOTO, K.-S.T. TU [1] *Applications of fractional calculus to ordinary and partial differential equations of Gauss type*, Proc. III. Internat. Conf. "Fractional Calculus" (Tokyo, 1989), Descartes Press, 1990, 159-165.
- ORLICZ, W. [1] Sur les fonctions continues non dérivables. *Fundamenta Math.*, 34 (1947), 45-60.
- PALEY, R.E.-WIENER, N. [1] *Fourier transforms in the complex domain* (New York, 1934).
- PARSEVAL, A. [1] Intégration générale et complète de deux équations importantes dans la mécanique des fluides. *Mém. prés. Inst. Sci. Paris*, 1 (1805), 524-545.
- PERRON, O. [1] Über den Integralbegriff. *Sitzungsber. Akad. Heidelberg*, 16 (1914), 1-16.
- [2] Beiträge zur Theorie der divergenten Reihen. *Math. Zeitschrift*, 6 (1920), 286-310.
- PLANCHEREL, M. [1] Contribution à l'étude de la représentation d'une fonction arbitraire par des intégrales définies. *Rendiconti Circ. Mat. Palermo*, 30 (1910), 289-335.
- PLESSNER, A. [1] *Trigonometrische Reihen* (I/3, in Pascal's Repertorium d. höheren Analysis, 1325-1396).
- POINCARÉ, H. [1] *Théorie analytique de la propagation de la chaleur* (Paris, 1895).
- POISSON, S.D. [1] Mémoire sur la manière d'exprimer les fonctions par les séries de quantités périodiques. *Journal Éc. Polytechn.*, 18 (11), (1820), 417-489.
- [2] Addition au mémoire précédent et au "Mémoire sur la manière d'exprimer..." *Journal Éc. Polytechn.*, 19 (12), (1823), 145-162.
- PRICE, I.I. [1] Orthonormal sets with non-negative Dirichlet kernels I-II. *Trans. Amer. Math. Soc.*, 95 (1960), 256-262; 100 (1961), 153-161.
- [2] An algebraic characterization of certain orthonormal systems. *Proc. Amer. Math. Soc.*, 19 (1968), 268-273.
- PRINGSHEIM, A. [1] Über das Verhalten von Potenzreihen auf dem Konvergenzkreise. *Münchener Sitzungsber.*, 30 (1900), 37-100.
- RADEMACHER, H. [1] Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen. *Math. Annalen*, 87 (1922), 112-138.
- RADO, T. [1] *Subharmonic Functions*. Ergebnisse der Math. V/1 (Berlin, 1937).
- RADON, J. [1] Theorie und Anwendungen der absolut additiven Mengenfunktionen. *Sitzungsber. Akad. Wien*, 122 /IIa (1913), 1295-1438.
- RAJCHMAN, A.-SAKS, S. [1] Sur la dérivabilité des fonctions monotones. *Fundamenta Math.*, 4 (1923), 204-213.
- RÉNYI, A. [1] On a conjecture of Steinhaus. *Annales Soc. Math. Polonaise*, 25 (1952), 279-287.
- [2] *Probability Theory* (Budapest-Amsterdam, 1970).

- RIEMANN, B. [1] *Über die Darstellbarkeit einer Funktion durch eine trigonometrische Reihe.* (Habilitationsschrift, Göttingen, 1854; Werke, 2.ed. 1892.)
- RIESZ, F. [1] Sur les systèmes orthogonaux de fonctions. *Comptes Rendus Acad. Sci. Paris*, **144** (1907), 615–619.
- [2] Sur les systèmes orthogonaux de fonctions et l'équation de Fredholm. *Comptes Rendus Acad. Sci. Paris*, **144** (1907), 734–736.
- [3] Über orthogonale Funktionensysteme. *Göttinger Nachrichten*, 1907, 116–122.
- [4] Sur une espèce de géométrie analytique des systèmes de fonctions sommables. *Comptes Rendus Acad. Sci. Paris*, **144** (1907), 1409–1411.
- [5] Stetigkeitsbegriff und abstrakte Mengenlehre. *Atti IV. Congr. Internaz. Mat. Roma*, **2** (1908), 18–24.
- [6] Sur les suites de fonctions mesurables, *Comptes Rendus Acad. Sci. Paris* **148** (1909), 1303–1305.
- [7] Sur les opérations fonctionnelles linéaires. *Comptes Rendus Acad. Sci. Paris*, **149** (1909), 974–977.
- [8] Sur certains systèmes d'équations fonctionnelles et l'approximation des fonctions continues. *Comptes Rendus Acad. Sci. Paris*, **150** (1910), 674–677.
- [9] Untersuchungen über Systeme integrierbarer Funktionen. *Math. Annalen* **69** (1910), 449–497.
- [10] Sur quelques points de la théorie des fonctions sommables. *Comptes Rendus Acad. Sci. Paris*, **154** (1912), 641–643.
- [11] *Les systèmes d'équations linéaires à une infinité d'inconnues* (Paris, 1913).
- [12] Démonstration nouvelle d'un théorème concernant les opérations fonctionnelles linéaires. *Annales Éc. Norm. Sup.*, **31** (1914), 9–14.
- [13] Über lineare Funktionalgleichungen. *Acta Math.*, **41** (1918), 71–98.
- [14] Über die Fourierkoeffizienten einer stetigen Funktion von beschränkter Schwankung. *Math. Zeitschrift*, **2** (1918), 312–315.
- [15] Sur l'intégrale de Lebesgue. *Acta Math.*, **42** (1920), 191–205.
- [16] Sur le théorème de M. Egoroff et sur les opérations fonctionnelles linéaires. *Acta Sci. Math. Szeged*, **1** (1922–23), 18–28.
- [17] Sur les fonctions subharmoniques et leur rapport à la théorie du potentiel I-II. *Acta Math.*, **48** (1926), 329–343; **54** (1930), 321–360.
- [18] Sur la formule d'inversion de Fourier. *Acta Sci. Math. Szeged*, **3** (1927), 235–241.
- [19] Sur la convergence en moyenne I-II. *Acta Sci. Math. Szeged*, **4** (1928–29), 58–64; 182–185.
- [20] Elementarer Beweis des Egoroff'schen Satzes. *Monatshefte f. Math. u. Phys.*, **35** (1928), 243–248.
- [21] Sur la décomposition des opérations fonctionnelles linéaires. *Atti Congr. Internaz. Mat. Bologna* (1928) **3** (1930), 143–148.
- [22] Über die linearen Transformationen des komplexen Hilbertschen Raumes. *Acta Sci. Math. Szeged*, **5** (1930–32), 23–54.
- [23] Sur l'existence de la dérivée des fonctions monotones et sur quelques problèmes qui s'y rattachent. *Acta Sci. Math. Szeged*, **5** (1930–32), 208–221.
- [24] Zur Theorie des Hilbertschen Raumes, *Acta Sci. Math. Szeged*, **7** (1934–35), 34–38.
- [25] Sur l'intégrale de Lebesgue comme l'opération inverse de la dérivation. *Annali di Pisa*, (2) **5** (1936), 191–212.
- [26] Sur quelques notions fondamentales dans la théorie générale des opérations linéaires. *Annals of Math.*, **4** (1940), 176–206.
- [27] L'évolution de la notion d'intégrale de Lebesgue. *Annales de l'Institut Fourier*, **1** (1949), 29–42.
- RIESZ, F.–SZÓKEFALVI-NAGY, B. [1] *Leçons d'analyse fonctionnelle* (Budapest, 1952).

- RIESZ, M. [1] Sur les séries de Dirichlet et les séries entières. *Comptes Rendus Acad. Sci. Paris*, **149** (1909) 309–312.
 [2] Sur la représentation analytique des fonctions définies par des séries de Dirichlet. *Acta Math.*, **35** (1912), 253–270.
 [3] Sur la sommation des séries de Fourier. *Acta Sci. Math. Szeged*, **1** (1922–23), 104–113.
 [4] L'intégrale de Riemann–Liouville et le problème de Cauchy. *Acta Math.*, **81** (1949), 1–223.
- RODRIGUES, O. [1] Mémoire sur l'attraction des sphéroïdes. *Corr. Éc. Royale Polytechn.*, **III** (1816).
- ROSS, B. [1] A brief history and exposition of fractional calculus. *Lecture Notes in Math.*, Vol. 457, Springer-Verlag (Berlin–Heidelberg–New York, 1975), 1–36.
- SACHSE, A. [1] Versuch einer Geschichte der Darstellung willkürlicher Funktionen einer Variablen durch trigonometrische Reihen. *Zeitschrift f. Math. u. Phys.* **25** (1880), 230–276.
- SAKS, S. [1] Sur les nombres dérivés des fonctions. *Fundamenta Math.* **5** (1924), 98–104.
- SAMKO, S.G.–KILBAS, A.A.–MARICHEV, O.I. [1] *Fractional Integrals and Derivatives*, Gordon and Breach Sci. Publ. (New York, 1992).
- SANSONE, G. [1] *Orthogonal Functions* (New York–London, 1959).
- SCHIPP, F. [1] On a generalization of the Haar system. *Acta Math. Acad. Sci. Hung.*, **33** (1979), 183–188.
- SCHMIDT, E. [1] Entwicklung willkürlicher Funktionen nach Systemen Vorgeschriebener. *Math. Annalen*, **63** (1907), 433–476.
 [2] Auflösung der allgemeinen linearen Integralgleichung. *Math. Annalen*, **64** (1907), 161–174.
 [3] Über die Auflösung linearer Gleichungen mit abzählbar unendlich vielen Unbekannten. *Rend. Circ. Mat. Palermo*, **25** (1908), 53–77.
- SCHNEE, W. [1] Die Identität des Cesàroschen und Hölderschen Grenzwertes. *Math. Annalen*, **67** (1909), 110–125.
- SCHWARTZ, L. [1] *Théorie des distributions* I–II (Paris, 1950–51).
- SIERPINSKI, W. [1] Un lemme métrique. *Fundamenta Math.*, **4** (1923), 201–203.
- SNEDDON, I.N. [1] *Fourier Transforms* (New York, 1951).
- SPLETTSTÖSSER, W.–WAGNER, H.J. [1] Eine dyadische Infinitesimalrechnung für Haar-Funktionen. *Zeitschrift Angew. Math. Mech.*, **57** (1977), 527–544.
- SRIVASTAVA, H.M. [1] The use of fractional calculus and other operational techniques in obtaining generating functions. *Proc. III. Internat. Conf. "Fractional Calculus"* (Tokyo, 1989). Descartes Press, 1990, 253–261.
- STEINHAUS, H. [1] A divergent trigonometrical series. *Journal London Math. Soc.* **4** (1929), 86–88.
- STIELTJES, T.J. [1] Recherches sur les fractions continues. *Annales de Toulouse*, (1) **8** (1894), 1–122.
- STONE, M.H. [1] The generalized Weierstrass approximation theorem. *Math. Mag.* **21** (1947–48), 167–184, 237–254.
- STURM, J. [1] Sur les équations différentielles du second ordre. *Journal de Math.*, **1** (1836), 106–186.
 [2] Sur une classe d'équations à différences partielles. *Journal de Math.*, **1** (1836), 373–444.
- SYLVESTER, J.J. [1] Note on spherical harmonics. *Philos. Mag.*, **2** (1876), 291–307, 400.
- SZÁSZ, O. [1] Über die Fourierschen Reihen gewisser Funktionenklassen. *Math. Annalen*, **100** (1928), 530–536.
- SZEGŐ, G. [1] Über die Lebesgueschen Konstanten bei den Fourierreihen. *Math. Zeitschrift*, **9** (1921), 163–166.
 [2] *Orthogonal Polynomials* (New York, 1939; 2.ed. 1959).

- SZÖKEFALVI-NAGY, B. [1] Über die Gesamtheit der charakteristischen Funktionen im Hilbertschen Funktionenraum. *Acta Sci. Math. Szeged*, **8** (1936–37), 166–176.
- [2] On the set of positive functions in L_2 . *Annals of Math.*, **39** (1938), 1–13.
- [3] *Spektraldarstellung linearer Transformationen des Hilbertschen Raumes*. Ergebnisse der Math. V/5 (Berlin, 1942; 2.ed. 1967).
- [4] Approximation der Funktionen durch die arithmetischen Mittel ihrer Fourierschen Reihen. *Acta Sci. Math. Szeged*, **11** (1946), 71–84.
- [5] On uniformly bounded linear transformations in Hilbert space. *Acta Sci. Math. Szeged*, **11** (1947), 152–157.
- [6] Vibrations d'une corde non homogène. *Bull. Soc. Math. France*, **75** (1947), 193–208.
- [7] Sur une classe générale de procédés de sommation pour les séries de Fourier. *Hung. Acta Math.*, **1** (1948), 14–52.
- [8] Séries et intégrales de Fourier des fonctions monotones non bornées. *Acta Sci. Math. Szeged*, **13** (1949–50), 118–135.
- [9] Une caractérisation affine de l'ensemble des fonctions positives dans l'espace L^2 . *Acta Sci. Math. Szeged*, **12A** (1950), 228–239.
- [10] Méthodes de sommation des séries de Fourier I. *Acta Sci. Math. Szeged*, **12B** (1950), 204–210.
- [11] Über die Konvergenz von Reihen orthogonaler Polynome. *Math. Nachrichten*, **4** (1951), 50–55.
- [12] Sur l'ordre de l'approximation d'une fonction par sa intégrale de Poisson. *Acta Math. Acad. Sci. Hung.*, **1** (1950), 183–187.
- [13] Approximation properties of orthogonal expansions. *Acta Sci. Math. Szeged*, **15** (1953), 31–37.
- [14] Sur les contractions de l'espace de Hilbert. *Acta Sci. Math. Szeged*, **15** (1953), 87–92.
- [15] Un calcul fonctionnel pour les opérateurs linéaires de l'espace hilbertien et certaines de ses applications. *Studia Math.* (spec. ser.) **1** (1963), 119–127.
- [16] *Introduction to Real Functions and Orthogonal Expansions* (Budapest–Oxford–New York, 2.ed. 1964).
- [17] Models of Hilbert space operators. *Proc. Royal Irish Acad.*, **74/A** (1974), 263–270.
- [18] Quasi-similarity of Hilbert space operators. *Proc. Int. Conf. on Differential Equations* (Uppsala, 1977), 179–188.
- SZÖKEFALVI-NAGY, B.–FOIAȘ, C.–GEHÉR, L. [1] On the permutability condition of quantum mechanics. *Acta Sci. Math. Szeged*, **21** (1960), 78–89.
- TANDÓRI, K. [1] Über die Cesàrosche Summierbarkeit der orthogonalen Reihen. *Acta Sci. Math. Szeged*, **14** (1951–52), 85–95.
- [2] Über die orthogonalen Funktionen I–V. *Acta Sci. Math. Szeged*, **18** (1957), 57–130, 149–168, 169–178; **19** (1958), 18–25; **20** (1959), 1–13.
- TAUBER, A. [1] Ein Satz aus der Theorie der unendlichen Reihen. *Monatshafte f. Math. u. Phys.*, **8** (1897), 273–277.
- TCHEBYSHEV, P.L. [1] Sur les questions de minima qui se rattachent à la représentation approximative des fonctions. *Mém. Acad. Sci. St. Petersburg*, **9** (1859), 199–291.
- [2] Sur le développement des fonctions à une seule variable. *Bull. math.-phys. Acad. Sci. St. Petersburg*, **1** (1859), 193–200.
- TITCHMARSH, E.C. [1] *The Theory of Functions* (Oxford, 1932).
- [2] *Introduction to the Theory of Fourier Integrals* (Oxford, 1937).
- [3] *The Theory of the Riemann Zeta-function* (Oxford, 1951).
- TOEPLITZ, O. [1] Über lineare Mittelbildungen. *Prace Mat.-fiz.*, **22** (1911), 113–119.
- VAN DER WAERDEN, B.L. [1] Ein einfaches Beispiel einer nicht-differenzierbaren stetigen Funktion. *Math. Zeitschrift*, **32** (1930), 474–475.

- VITALI, G. [1] Sui gruppi di punti. *Rend. Circ. Mat. Palermo*, 18 (1904), 116-126.
 [2] *Sui problema della misure dei gruppi di punti di una retta* (Bologna, 1905.)
 [3] Sui gruppi di punti e sulle funzioni di variabili reali. *Atti Acad. Sci. Torino*, 43 (1908), 75-92.
- VITALI, G.-SANSONE, G. [1] *Moderna teoria delle funzioni di variabile reale* (Bologna, 3.ed. 1951).
- VOLTERRA, V. [1] Sui principii del calcolo integrale. *Giornale Mat.*, 19 (1881), 333-372.
 [2] Sulla inversione degli integrali definiti I-III. *Rend. Accad. Lincei Roma*, 5 (1896), 177-185; *Annali di Mat.*, (2) 25 (1897), 139-178.
- WADE, W.R. [1] Sets of uniqueness for Haar series. *Acta Math. Acad. Sci. Hung.*, 30 (1977), 265-281.
- WAGNER, H.J. [1] *Ein Differential- und Integralkalkül in der Walsh-Fourier-Analysis mit Anwendungen* (Köln-Opladen, 1973).
- WALSH, J.L. [1] A closed set of normal orthogonal functions. *American Journal Math.*, 45 (1923), 5-24.
- WATSON, G.N. [1] *A Treatise on the Theory of Bessel Functions* (Cambridge, 2.ed. 1944).
- WEIERSTRASS, K. [1] Über kontinuierliche Funktionen eines reellen Arguments, die für keinen Wert des letzteren einen bestimmten Differentialkoeffizienten besitzen. *Werke*, II, 71-74.
- WEYL, H. [1] *Gruppentheorie und Quantummechanik* (Leipzig, 2.ed. 1931).
- WHITTAKER, E.N.-WATSON, G.N. [1] *Modern Analysis* (Cambridge, 4.ed. 1952).
- WIENER, N. [1] *The Fourier Integral* (Cambridge, 1933).
- YOUNG, G.C. [1] On the derivatives of a function. *Proc. London Math. Soc.*, (2) 15 (1916), 360-384.
- YOUNG, G.C.-YOUNG, W.H. [1] On the existence of a differential coefficient. *Proc. London Math. Soc.*, (2) 9 (1910), 325-335.
 [2] On the theorem of Riesz-Fischer. *Quarterly Journal of Math.*, 44 (1913), 49-88.
- YOUNG, L.C. [1] *The Theory of Integration* (Cambridge, 1927).
- YOUNG, W.H. [1] On sequences of sets of intervals containing a given set of points. *Proc. London Math. Soc.*, (2) 1 (1904), 262-266.
 [2] Open sets and the theory of content. *Proc. London Math. Soc.*, (2) 2 (1904), 16-51.
 [3] On a new method in the theory of integration. *Proc. London Math. Soc.* (2) 9 (1910), 15-50.
 [4] On the integration of Fourier series. *Proc. London Math. Soc.*, (2) 9 (1910), 449-462.
 [5] The usual convergence of a class of trigonometrical series. *Proc. London Math. Soc.*, (2) 13 (1914), 13-28.
- ZELLER, K. [1] Über Konvergenzmengen der Fourierreihen. *Archiv der Math.*, 6 (1955), 335-340.
 [2] *Theorie der Limitierungsverfahren* (Berlin-Göttingen-Heidelberg, 1958).
- ZYGMUND, A. [1] Sur l'application de la première moyenne arithmétique dans la théorie des séries orthogonales. *Fundamenta Math.*, 10 (1927), 356-362.
 [2] Sur la sommation des séries de fonctions orthogonales. *Bull. Inst. Acad. Cracovie, série A*, 1927, 293-308.
 [3] *Trigonometrical Series* (Warsaw, 1935).
 [4] Smooth functions. *Duke Math. Journal*, 12 (1945), 47-76.
 [5] *Trigonometric Series I-II*. 2.ed., (Cambridge, 1968).

INDEX

A

- ABEL, N.H. 364
 - Poisson summation 363
 - 's theorem for power series 363
- Abelian group 441
- absolutely continuous functions 114
- abstract spaces 230, 235, 239, 437
- accumulation point 250, 439
- additivity 74, 109, 121, 140
 - total (σ) 152, 176
- adherence (of set): see closure
- adjointness
 - of Banach spaces 452
 - of differential operators 455
 - of integral operators 466
 - of linear operators 446
- AKHIESER, N. 301, 471
- alephs (cardinals) 34
- ALEXANDROV, P.S. 454
- ALEXITS, G. 297, 312, 385
 - 's multiplicatively orthogonal systems 297
- algebra, Boolean 40
 - of events 225
 - of sets 28
- algebraic number 32
- almost everywhere (p.p.) 102
 - (μ)- 209
- analysis situs: see topology
- angle of inclination (for abstract vectors) 232, 236, 241
- approximation
 - Bessel's 300, 301
 - Diophantine 478
 - Tchebyshev's 278, 301
 - uniform 83, 84
- approximation theory 301
- arc length 108, 109

- ARZELÀ, C. 178, 472
 - 's theorem 178
- ATKINSON, F.V. 283, 359

B

- BAIRE, R. 90, 472
 - 's classification of functions 90
- BANACH, S. 442, 450
 - 's fixed point theorem 442, 443
 - space 450
- base 234
- base measure 220
- BENDIXSON, J. 62
- BERNOULLI, D. 457
 - 's method 458
- BERNOULLI, J. 344
- Bernoullian
 - numbers 344
 - polynomials 343
- BERNSTEIN, F. 35
- BESSEL, F.W. 285, 355
 - functions, of the first and second kind 285
 - 's approximation 300, 301
 - 's differential equation 285
 - 's identity 300
 - 's inequality 301
- bijjective mapping 30
- bilinear form 446, 447
- bilinear functional 445, 446
- BING, R.H. 454
- BOLZANO, B. 94
 - 's theorem 82
 - Weierstrass theorem 53
- Boolean algebra 40
- Boolean ring 40
- BOREL, E. 149
 - 's covering theorem 65
 - 's measure 206

- BOREL, E.
 's ring (σ -ring) 40
 's set 206
 's summation 364
- BORN, M. 448
- bound (of linear set) 48
 greatest lower (g.l.b.) 48
 least upper (l.u.b.) 48
- boundary (of point set) 48
 conditions 283, 456, 458
- boundary problem
 Dirichlet's 384
 for partial differential equations 457, 458, 459
 for self-adjoint ordinary differential operators 456
 Sturm-Liouville 283
- boundedness
 of operator 74
 of point function 76
- BROUWER, L.E.J. 37
- BRUCKNER, A.M. 95
- Buniakowsky-Schwarz inequality 240
- BURKHARDT, H. 301, 302, 472
- BURKILL, J.C. 148
 integral 148, 472
- C**
- calculus of variations 256, 460
- CAMPOS, L.M.B.C. 472
- canonical decomposition, 195
- CANTOR, G. 35, 74, 335
 -Bendixson theorem 62, 63
 -Bernstein equivalence theorem 35
 's diagonal processes 31, 33
 's meet theorem 64
 's ternary set 61
 's theorem of uniqueness 336
 's theorem on set operators 74
- CARATHÉODORY, C. 472
 's extension theorem 208
- cardinality relations 34
- cardinals
 finite, transfinite 36, 37, 38
- CARLEMAN, T. 472
- CARLESON, L. 334
 's theorem 472
- Cartesian product (direct product) 29
- CAUCHY, A. 119, 472
 -Dirichlet method 136
 -Riemann equations 460
 's convergence principle 50, 237
 's example 377
 's inequality 47
 's initial value problem 396
 's limit theorem 363
 's mean-value theorem 94
 's principal value 395
 's product 364
 's sequence 442
- central limit theorem 434
- CESÀRO, E. 364
 's summation theorem 362, 364
- CHAPMAN, S. 379
- characteristic function
 of set 30
 of random variable 434
- classical orthogonal polynomial systems (COP) 280
- closed curve and surface (in E_n) 49, 50
- closed set 54, 58
- closure
 of set 55
 of set operation 40
 of system of functions 290, 291
- clustering point 51
- cluster value
 of point function 77
 of sequence 51, 52
- COHEN, P.J. 39
- COLLATZ, L. 471
- common part of sets: see meet
- compactness
 of group 206
 of metric space 250
 of set 53
- complement: see difference of sets
- completeness,
 of abstract space 206, 442
 of orthogonal system 233, 236, 242
- component: see projection
- condensation point 53
- conjugate series 260, 340
- connectedness of set 59
- constructive theory of functions 301
- content (of set): see Jordan's measure theory
- continuity,
 absolute 114, 118
 $L(a, b)$ - 191
 left-hand, right-hand 81
 of linear operators 445
 of mapping 49
 of point function 81
 semi (lower, upper) 81, 82
 uniform 82

- continuum 33, 59
 - hypothesis 39
 - power of 34
- contraction 443
- convergence, dominated (majorized) 423
 - in complete metric space 442
 - in E_p 50
 - l^2 - 237
 - $L^2(a, b)$ - 244
 - $L^p(a, b)$ - 438
 - pointwise 88
 - quasi-uniform 90
 - strong 249, 445
 - uniform 89
 - weak 249, 445
 - with respect to (L) -measure 178
- convergence criterion
 - Cauchy's 50, 237
 - Dini's 323
 - Dirichlet's 327, 473
 - Jordan's 323, 475
 - Lebesgue's 329, 476
 - Lipschitz's 325, 326, 476
 - Pringsheim's 341, 478
 - Riemann's 321, 479
 - Riesz-Fischer's 246, 271, 438, 479
 - Weierstrass' 89
- convex set 58
- convolution 411
 - theorem 414
- cosine series 314
- cosine transform 395
- cotangent's decomposition into partial
 - fractions 350
- countability 30, 31, 32
- COURANT, R. 302, 460, 472
- covering theorem 65
- CSÁSZÁR, Á. 454, 472
- curve in E_p 49
- D**
- D'ALEMBERT, J. 457
- DANIELL, P.J. 472
 - integral 210
- DARBOUX, G.
 - 's lower and upper integral 122, 123
 - 's lower and upper sum 122, 123
 - 's theorem 94
- decomposition problem 250
- definiteness 119
- DE LA VALLÉE-POUSSIN, CH. J. 335
 - 's theorem 473
- De Morgan's formulae 28
- DENJOY, A. 106
 - integral 201, 466
 - Young-Saks theorem 106, 107
- density 147
- derivate numbers 91, 92
 - Dini's 92
- derivate set 54
- derivative
 - left-hand and right-hand 93
 - lower and upper 92
 - partial 91
 - with respect to region 147
- diameter (of set) 64
- DIEUDONNÉ, J. 471, 473
- difference of sets 28
 - symmetric 29
- differentiability
 - left-hand and right-hand 93
 - partial 93
 - repeated 93
- differential
 - partial 91
 - total 91
- differential coefficient: see derivative
- differential equation
 - Bessels' 286
 - elliptic 465
 - hyperbolic 465
 - Laplace (in plane, in space) 460, 461, 462
 - of diffusion 462, 463
 - of harmonic oscillation 288, 289
 - of Hermite polynomials 279
 - of homogeneous vibrating string 457, 458
 - of inhomogeneous vibrating string 459
 - of Jacobi polynomials 279
 - of Laguerre polynomials 279
 - of Legendre polynomials 277
 - of linear heat conduction 462
 - of membrane 463
 - of soil consolidations 463
 - parabolic 465
 - Sturm-Liouville 282, 456
 - with self-adjoint differential operators 456
- differential operator 75, 455, 456
- differentiation
 - inverting of 133, 134, 183, 201
 - of Fourier series 339
 - of (L) -integral function 181, 196
 - of (R) -integral function 132

- differentiation
 - of series of monotonic functions 107
- diffusion 463
- DINI, U. 323, 472
 - 's derivate numbers 92, 93
 - 's test for Fourier series 323
- DIRAC, P. 449, 473
 - impulse (" δ -function") 449
- direct product: see Cartesian product
- DIRICHLET, G. 319, 460, 462
 - 's function 42
 - 's integral 320
 - 's kernel 320
 - 's problem (in plane, in space) 460, 462
 - 's series 364, 385
 - 's test for Fourier series 327
 - 's transformation formula 455
- discontinuities, points of jump 91
- disjoint sets 29
- distance
 - in abstract spaces 47, 236, 243, 438, 442
 - of sets 62
- distribution (function) 216
- distribution theory, 450
- DOETSCH, G. 473
- duality (of functional spaces) 439
- duality principle (for sets) 29, 59
- DU BOIS-REYMOND, P. 331, 335, 473
 - 's example 331
 - 's singularity 473
 - 's theorem 335
- DUNFORD, N. 471
- E
- EGOROV, D.T. 169, 473
 - 's theorem 169
- eigenfunction, eigenvalue 283, 446, 456, 469
- eigenvector: see eigenfunction
- entire function of exponential type 408, 476
- envelope (function), lower and upper 167
- EPSTEIN, P. 354, 473
- equiconvergence 359, 424
- equivalence,
 - g - 216
 - (L)- 174
 - μ - 209
 - of sets 34
 - of wave and matrix mechanics 448
 - theorem of Cantor and Bernstein 35
- ERDŐS, P. 301, 473
- Euclidean spaces 34, 47, 231
- EULER, L. 259, 344, 457, 458
 - Maclaurin summation formula 345
 - 's formulae for Fourier coefficients 259
 - 's integral 281
 - 's relations 260
- expected value 225, 434
- exterior measure: see outer measure
- F
- FABER, G. 95, 473
- FATOU, P. 473
 - 's lemma 189
- FEJÉR, L. 364, 371, 372, 376, 473
 - Lebesgue theorem for Fourier integrals 420
 - 's example 331
 - 's fundamental theorem and approximation theorem 371
 - 's integral 370
 - 's kernel 370
- FELLER, W. 473
- field of sets: see Boolean ring
- FISCHER, E. 305, 473, 474
- fixed point 442, 443
- fluctuation value, lower and upper 78
- FOIAŞ, C. 471, 481
- FOMIN, S.V. 471
- FOURIER, J.B. 252, 259, 310, 369, 395
- Fourier analysis 255, 357
- Fourier coefficients (Fourier constants) 252, 359
 - order of magnitude of 315, 316
- Fourier effectivity 384
- Fourier expansibility 323
- Fourier integrals 395
 - 's integral theorem 397
 - 's inversion formulae 398
- Fourier series
 - general 252
 - ordinary (trigonometric) 259
 - in complex form 260
- Fourier series
 - classical convergence tests for 323, 325, 327
 - convergence in mean (L^2 -convergence) of 303
 - equiconvergence theorems for Sturm-Liouville series 359
 - Fourier-Bessel series 287, 360
 - Fourier expansions with respect to generalized Haar functions 358
 - Laplace series 462

- Fourier series
 - Legendre series 462
 - minimum property of the partial sums 300
 - summation of 369, 376, 380, 384
 - term-by-term integration of 337
- Fourier-Stieltjes transformation 432
- Fourier transforms
 - classical convergence theory of 414
 - general properties 408
 - L^2 -convergence theory of 398
 - non-absolute convergence of 424
 - summation of 419, 420
- fractional calculus: see fractional integrals and derivatives
- fractional integrals and derivatives (generalized integro-differential operators)
 - of Riemann-Liouville and Weyl type 393, 394, 476, 477
 - of M. Riesz type 479
- fractional part (function) 77, 342
- FRAENKEL, A.A. 39, 474
- FRANKLIN, P. 292, 474
- FRÉCHET, M. 445, 454
 - Riesz theorem 445
- FREDHOLM, I. 456, 457
 - 's alternative 468
 - 's integral equation 465
- FREUD, G. 474
- Frobenius' theorem 364
- FUBINI, G. 221
 - 's theorem for abstract (L) -integrals 223
 - 's theorem for double (L) -integrals 203
 - 's theorem on series of monotonic functions 107
- function
 - abstract 73
 - Bessel's 285, 355
 - continuous: see continuity
 - derivable: see differentiability
 - entire rational 83
 - with integral coefficients 32
 - everywhere continuous, non-differentiable 94
 - generalized 449, 450
 - Green's 457
 - Hamilton's 448
 - harmonic 384
 - Heaviside's 449
 - integrable: see integrability
 - jump of 100
 - Lebesgue's 305
 - function
 - Legendre's (associated) 461
 - lower and upper envelope of 167
 - measurable: see measurability
 - monotonic 99, 100, 118
 - of bounded variation 108, 109
 - periodic 313
 - positive and negative part of 157
 - spherical harmonic 384
 - step 165
 - generalized (μ) - 208
 - subharmonic 478, 479
 - summable (in Lebesgue's sense) 190
 - Weierstrass' 331
 - functional (operator) 74
 - bilinear and Hermitian 445, 446
 - bounded 445
 - linear 445
 - quadratic 446
 - functional analysis 76, 230
 - functional transformation 75
 - Laplace 395
 - unitary 444
 - function spaces 438
 - C 438
 - L 186, 438
 - L^2 (Hilbert's) 238, 239, 438
 - L^p (Riesz's) 438
 - V 457
 - fundamental identity for partial sums of orthogonal series 299
- G
 - gamma function 281
 - GAUSS, C.F. 39
 - GEHÉR L. 448
 - general topological space 454
 - generating function 266
 - GIBBS, J.W. 373
 - Gibbs phenomenon 373
 - GRAM, J.P. 289
 - Schmidt orthogonalization 289, 474, 480
 - group, commutative (Abelian) 441
 - locally compact topological 206
- H
 - HAAR, A. 206, 292, 359
 - 's equiconvergence theorem 359
 - 's generalized expansion theorem 358
 - 's measure 206

HAAR, A.
 's orthonormal system 292
 's series 356
 HADAMARD, J. 394
 HAHN, H. 95
 HAHN, W. 474
 HALMOS, P.R. 471
 Hamilton's function 448
 HARDY, G.H. 95, 302, 328, 335, 337, 361,
 365, 368, 379, 380, 419, 439
 -Landau theorem 366
 -Littlewood theorem 368
 harmonic analysis 354
 fundamental problem of 460
 harmonic function series 460
 harmonic oscillation (vibration) 284
 HAUSDORFF, F. 454
 heat conduction 460
 HEAVISIDE, C. 449
 height (of polynomial) 32
 Heine's theorem 82
 HEISENBERG, W.K. 448
 's matrix mechanics 448
 HERMITE, CH. 279
 Hermite's differential equation 279
 functional 446
 Hermitian matrix 448
 polynomials 279
 Hertz's oscillator 354
 HEWITT, E. 471
 HILB, E. 312, 474
 HILBERT, D. 230, 448, 475
 's abstract space 444
 's space of functions 239
 's space of sequences 235
 theorem of Hilbert and Schmidt 469
 HILLE, E. 448
 HOBSON, E.W. 332, 438
 homeomorphy 49
 homogeneity (of operators) 74
 HÖLDER, C.
 's inequality 438
 's summation 363
 HUNT, R.A. 334, 475
 HUNTER, M.N. 476
 HURWITZ, A. 301, 316, 348
 's formula 348
 's zeta-function 345
 hypergeometric series 278

I

idempotency 28
 indefinite integral: see primitive
 independence, linear 287
 stochastic 225
 inequality
 of Bessel 301
 of Buniakowsky and Schwarz 240
 of Cauchy 47, 236
 of Hölder 438
 of Minkowski 439
 triangle 47, 237, 244, 441
 infimum: see greatest lower bound
 initial conditions 458
 inner measure
 (J)- 120
 (L)- 150
 inner product (in abstract spaces) 231, 235
 239
 integrability
 (L)- 156, 160, 173
 (LS)- 211
 μ - 209
 (R)- 123, 124
 (RS)- 139, 140
 integral
 abstract Lebesgue (μ)- 209, 210
 area (double) 136, 137
 Burkill 148
 Daniell 210
 Darboux's lower and upper 123, 124
 Denjoy 201
 Dirichlet's 321
 Euler's 281
 Fejér's 370
 Fourier's 397, 398
 improper (L)- 186
 improper (R)- 136, 137
 Lebesgue 156, 160, 173, 184
 Lebesgue-Stieltjes 209, 210
 line 138
 of several dimensions, (L)- 203
 of several dimensions, (R)- 136
 on product space 223, 224
 on region 137
 Perron 201
 Poisson 381
 Radon 147
 Riemann 123, 124
 Riemann-Stieltjes 139
 singular 311
 surface 138

- integral
 - trigonometric 395
 - volume (triple) 137
 - Young's lower and upper 154, 155
- integral equation
 - of Fredholm's type 456, 457
 - of Hilbert-Schmidt type 469
- integral function
 - (L)- 180, 181
 - (R)- 132
- integral mean 138, 144, 175
 - local (absolute and ordinary) 376
 - of index p 439
 - quadratic 244
- integral operator 75, 456
- integral part (function) 337
- integral transformation 76
- integration
 - by parts 141, 202, 215
 - by substitution 202, 215
 - of derivative 133, 134, 201
 - of Fourier series 337, 338
 - successive 137, 203, 204, 222
 - term-by-term (L)- 177, 187
 - term-by-term (R)- 130
- integrator 139, 210
- integro-differentiation 393, 394
- interior measure: see inner measure
- interval function 74, 147
- isomorphy 307
- isoperimetric problem 302
- iteration 443
- J**
 - JACKSON, D. 301, 475
 - JACOBI, C.G. 276
 - polynomials 276
 - differential equation of 278
 - JORDAN, C. 49, 111, 119
 - 's curve 49
 - 's curve theorem 49
 - 's measure theory 119
 - 's representation theorem 111
 - 's test for Fourier series 323
 - JORDAN, P. 449
 - jump (saltus) 100
 - jumping function, 101
- K**
 - KACZMARZ, S. 311, 475
 - KAHANE, J.P. 334, 475
 - KALLA, S.L. 393, 475
 - KAMKE, E. 475
 - KATZNELSON, Y. 334, 475
 - kernel function 311
 - adjoint and self-adjoint 472
 - Dirichlet's 320
 - Fejér's 370
 - iterated 467
 - of integral transform 466
 - Poisson's 381
 - symmetric 466
 - z_θ - 388
 - KHINTCHIN, J. 296, 475
 - KILBAS, A.A. 393, 480
 - KIRYAKOVA, V.S. 475
 - KNOPP, K. 95, 361
 - Schnee theorem 363
 - KOLMOGOROV, A.N. 471, 475
 - 's example 334
 - 's foundation of probability theory 296
 - KURATOWSKI, C. 471
- L**
 - LACROIX, S.F. 302
 - LAGRANGE, J. 256
 - 's identity 48
 - 's lemma 256
 - 's mean value theorem 94
 - LAGUERRE, E.N. 279
 - polynomials 279
 - differential equation of 279
 - LANDAU, E. 366
 - LAPLACE, P.S. 460, 461, 475
 - 's equation (in plane, in space) 460, 461
 - 's series 462
 - 's spherical harmonic 462
 - transformation 395
 - lattice theory 42
 - Laurent's series 261
 - leading coefficient 269
 - LEBESGUE, H. 95, 103, 126, 149, 154, 177, 183, 187, 191, 312, 317, 329, 336, 376, 384
 - constants: see Lebesgue functions
 - continuation 208
 - functions 311
 - integral 156, 160, 173, 184
 - measurable set 120, 150, 205
 - method 158
 - point 181
 - 's convergence test for Fourier series 329

LEBESGUE, H.
 's criterion for (R) -integrability 125
 's limit theorem 177
 's lower and upper sum 158
 's main theorem on term-by-term integration 187
 's measure 150
 's μ -integral [abstract (L) -integral] 209, 210
 's theorem of summation 376
 's theorem on functions of bounded variation 112
 's theorem on monotonic functions 103
 -Stieltjes integral 204, 211
 -Stieltjes measure 211
 sum (generalized) 183
 LEGENDRE, A.M. 476
 functions (associated) 461
 polynomials 263
 differential equation of 267
 's series 269
 Leibniz series 351
 Leibniz-Newton formula, generalized 131, 179, 194
 LEINDLER L. 310, 476
 LERCH, M. 476
 's relation 353
 's zeta-function 352
 level set 160
 LEVI, B. 476
 's theorem 189, 246
 limit, in complete metric space 442
 lower and upper 51
 l^2 - 237
 $L^2(a, b)$ - 247, 248
 $L^p(a, b)$ - 438
 of point function 76
 of sequence 50, 51, 52
 one-sided 77, 78
 limiting process (passage to the limit)
 under the (L) -integral sign, 187
 under the (R) -integral sign, 130
 LINDELÖF, E. 54
 's theorem 54
 LINDEMANN, F. 33
 linear algebra (abstract) 446
 linear operator 75
 linear space 442
 LIOUVILLE, J. 33, 282, 283, 359
 -Hobson asymptotic formula 359
 LIPSCHITZ, R. 182
 's condition ordinary, of order α 114, 115, 325, 326

LITTLEWOOD, J.E. 316, 368, 379, 439
 logical functional calculus 45
 LOVE, E.R. 393, 476
 LUKÁCS, F. 342
 LUSIN, N.N. 170, 476

M

MANOLITSAKIS, I. 408, 476
 mapping
 bijective 30
 continuous 49
 homeomorphic 49
 isometric 407
 isomorphic 307
 neighbourhood-preserving 49
 norm-preserving 407
 MARCINKIEWICZ, J. 380, 476
 MARICHEV, O.I. 393, 480
 mathematical logic 45
 mathematical statistics: see probability theory
 matrix mechanics 448
 matrix product 447
 MAXWELL, J.C. 462
 -Sylvester representation 462
 mean
 Cesàro's 362
 Hölder's 363
 weighted arithmetic 365
 weighted integral 138
 mean-value theorem, Cauchy's 94
 first (for integrals) 138, 144, 175
 Lagrange's 94
 second (for integrals) 175
 measurability, of function 159, 183
 of set 120, 150, 205
 measure
 abstract (μ) - 205
 Borel's 206
 Haar's 206, 292
 inner (J) - 120
 inner (L) - 150
 invariance of 119
 Jordan's 120
 Lebesgue's 150
 Lebesgue-Stieltjes 211
 outer (J) - 66
 outer (L) - 66
 measure space 205
 closure of 206

- measure theory 119
 - basic axioms of 119
 - general 205
 - Jordan's 120
 - Lebesgue's 150
- meet of sets: see intersection
- MENSHOV, D. 308, 477
- metric space 441
- metrization 454
- MIKHLIN, S.G. 477
- MIKOLÁS, M. 95, 243, 247, 280, 283, 291, 292, 348, 353, 356, 365, 385, 389, 393, 424, 428, 432, 438
 - characterization of Sturm-Liouville eigenfunctions 280, 283
 - new proof of the Riesz-Fischer convergence criterion 246, 438
 - recent applications of Poisson's sum formula 432
 - 's equiconvergence theorem for Fourier's double integral 422
 - 's extension of Haar's system and series 294, 358
 - 's methods for constructing everywhere continuous, non-differentiable functions 96, 97, 98
 - 's summing procedures based on fractional integration 389-394
 - 's tests for non-absolutely convergent Fourier integrals 428
- MIKUSINSKI, J. 405, 444
- MINKOWSKI, H. 240
 - 's inequality 240
 - generalized 439
- monotony, of functions 99, 118
 - of measure 121
 - piecewise 109
- N
- NAGATA, J. 454
- NATANSON, I.P. 301
- neighbourhood (in abstract spaces) 48, 250, 442, 454
- NELSON, P.I. 292
- NEUMANN, C.G. 478
 - 's existence and expansion theorem 466, 467
 - 's series 452
- NEUMANN, J. 448
 - 's foundation of quantum mechanics 471
 - P.Jordan theorem 450, 475
- NISHIMOTO, K. 393, 478
 - norm 76, 231, 235, 239, 437, 442, 450, 466
 - normal (Gauss) distribution 435
 - normalization 233, 241
 - normed bounded set 75, 450
 - null element: see zero element
 - null set 67, 157
 - with respect to functions 142
 - g - 215
 - μ - 209
- null vector: see zero vector
- O
- open set 58, 454
- operational calculus (of Minkowski) 478
- operator 73
 - adjoint and self-adjoint 446, 455, 468
 - bounded 74
 - composed 73
 - continuous and totally continuous 445
 - differential and integral 75
 - fractional (generalized integro-differential) 393, 394
 - inverse of 73
 - linear 75
- operator equation 443
- ordinal number and order type 39
- ordinate set 122
- ORLICZ, W. 95, 478
- orthogonal functions of the family
 - $\left\{ \begin{matrix} (r) \\ \mu_{k_1} \dots \mu_{k_r} \end{matrix} \right\}$ 358, 359
- orthogonality 229, 232, 236, 242
 - of eigenfunctions 283, 456, 469
 - weighted 270, 271
- orthogonalization and orthonormality 288
- orthogonal polynomials, general 271, 272
 - main properties of 273, 274
- orthogonal systems and series of Sturm-Liouville's type 282, 283, 285, 286
- oscillation sum of Riemann 125, 126
- oscillation sum of Young's type 156
- oscillation theorems for Sturm-Liouville series 287, 288
- outer measure, associated to measure space 207
 - (J)- 66
 - (L)- 66, 149, 150
- outer product (in abstract spaces) 231

P

PALEY, R.E. 408, 478
 -Wiener theorem 408

parallelogramma identity 450

parametric representation
 of curve 49
 of surface 50

PARSEVAL, A. 302, 478
 -Hurwitz formula 301
 for trigonometric base system 316
 generalized 302
 's formula for Fourier integrals 401
 generalized 405

PERRON, O. 201, 365
 integral 201

PLANCHEREL, M. 406, 478
 's theorem 406

PLESSNER, A. 312, 478

POINCARÉ, H. 463, 478

POISSON, S.D. 364, 380, 478
 's integral 381
 's kernel 381
 's summation formula 31
 generalized 432

PÓLYA, G. 439, 474

potential equation: see Laplace equation

power-set 38

power series (complex) and trigonometric
 base system 260

premeasure 207

PRICE, J.J. 292

primitive (function) 131

principle of consistency 362

principle of permanence 362

PRINGSHEIM, A. 341
 's criterion 478

probability density 434

probability distribution 434

product measure 220

product space 220

projection (in abstract spaces) 236, 241

Q

quadratic form 447

quadratic functional 446

quantum mechanics and functional
 analysis 448

R

RADEMACHER, H. 478
 -Khintchin-Kolmogorov theorem 296
 -Menshov theorem 308
 's orthonormal system 295, 478

RADO, T. 460, 478

RADON, J. 246, 478
 's integral 147, 217

RAJCHMAN, A. 106, 478

random variable 225
 characteristic function of 434
 main data of 434

rectifiability 108

recurrence formula 267, 273

region, connectedness 59

regular point 325

regularity
 of eigenvalue problem 287
 of linear summation method 364

RÉNYI, A. 258, 434

RIEMANN, B. 119, 317, 321, 336, 479
 integrability 126
 integral 123, 124
 -Lebesgue lemma 317
 -Liouville fractional integral 394
 's criterion for Fourier series 322
 's integrability criterion 126
 's localization theorem 321
 's sum 124
 's summation method 336
 -Stieltjes integral 139
 's zeta-function 347

RIESZ, F. 106, 139, 165, 169, 178, 201, 246,
 307, 312, 316, 406, 437, 438, 440,
 441, 445, 446, 448, 454, 459, 460,
 468, 469
 -Fischer convergence criterion 246
 -Fischer fundamental theorem 305
 -Fréchet theorem 445
 's integral theory 165
 's L^p -space 437
 's sequence spaces 441
 's theorem on linear functionals 440

RIESZ, M. 312, 364, 365, 379, 393, 394,
 479, 480
 's summation methods 364, 365
 's theorem on Fourier series 379

ring
 Boolean 40
 sigma- 40

- RODRIGUES, O. 263, 480
 's formula 263
 generalized 278
- ROGOSINSKI, W.W. 302, 474
- ROSS, B. 393, 480
- S**
- SACHSE, A. 332, 480
- SAKS, S. 106, 480
- SAMKO, S.G. 393, 480
- SANSONE, G. 360, 471, 480
- scalar product: see inner product
- SCHIPP, F. 292, 480
- SCHMIDT, E. 289, 480
- SCHNEE, W. 363, 480
- SCHRÖDINGER, E. 448
 's equation 464
- SCHWARTZ, L. 450, 480
- SCHWARZ, H.A. 94, 332
 's example 332
 's inequality 240
 's theorem 94
- self-adjointness
 of Banach space 451
 of eigenvalue problem 456
 of operator 446, 456, 468
- semi-ring 217
- separability 442
- separation 458
- set theory 27
- SIERPINSKI, W. 69
 's lemma 480
- sigma-additivity: see total additivity
- sigma-algebra: see Boolean algebra
- sigma-ring: see Boolean ring
- sine series 314
- sine transform 395
- singular integral 311
- singularity
 Du Bois-Reymond's 341
 of eigenvalue problem 287
- SMIRNOV, Y.M. 454
- SNEDDON, I.N. 396, 480
- space
 abstract 231, 437
 Banach, 450
 complete 206, 442
 Euclidean 36, 47, 231
 Hilbert 235, 238, 437, 444
 linear 441
 measure 205
 metric 441
- Riesz 440
 separable 442
 topological 454
 unitary 444
- spectral analysis of waves 446
- spectral decomposition 446
- spectrum 446
- SPLETTSTÖSSER, W. 292, 480
- SRIVASTAVA, H.M. 393, 480
- STEINER, J. 302
- STEINHAUS, H. 334, 480
- STIELTJES, T.J. 138, 139, 211, 212, 480
 integral 139, 212
 's approximating sum 139
- STONE, M.H. 83, 84, 480
 's theorem 84
- STROMBERG, K. 471
- STURM, I. 283, 480
 -Liouville differential equation and
 boundary problem 283
 -Liouville eigenfunctions 287
 -Liouville series 287
- subadditivity 121
- sublimit: see clustering point
- summability
 of integrals 418
 of series 362, 364, 365
 strong 379
- summation methods 361, 362
 Abel-Poisson $[(A-)]$ 363
 Borel's $[(B-)]$ 364
 Cesàro's $[(C, r-)]$ 362
 (D)- 385
 (M)- 394
 Hölder's $[(H, r-)]$ 363
 linear $[(T-)]$ 365
 Riemann's 336
 Riesz's $[(R, \mu_k-)]$ 364
 (W $_{\pm}$)- 393
- superposition principle 458
- supremum: see upper bound
- SYLVESTER, J.J. 480
- SZÁSZ, O. 312, 480
- SZEGŐ, G. 312, 335, 360, 480
- SZŐKEFALVI-NAGY, B. 106, 165, 274, 283,
 292, 302, 312, 360, 384, 446, 459
 's spectral theory 481
 functional analysis (joint with F. Riesz)
 471
 harmonic analysis on Hilbert's space
 (joint with C. Foias) 471

T

- TANDORI, K. 297, 310, 311, 385, 481
 TAUBER, A. 366, 481
 Tauberian theorems 366
 TCHEBYSHEV, P.L. 276
 polynomials of the first and second kind 277
 's approximation 278
 tensor 74
 TITCHMARSH, E.C. 348, 396, 406, 415, 422
 TOEPLITZ, O. 365, 481
 's summation theorem 365
 topology 49
 topological mapping 49
 generalized 454
 total variation 109
 transcendental number 32
 transfinite induction 39
 triangle inequality 47, 237, 244, 441
 TRICOMI, F.G. 471
 trigonometric integral 395
 trigonometric polynomial 83
 trigonometric series 229
 TU, S.T. 478
 TURÁN, P. 301, 473

U

- uniqueness problem (for trigonometric series) 335
 unit operator 74
 unit vector 231, 236, 241
 URYSON, P.S. 454

V

- VAN DER WAERDEN, B.L. 95, 481
 vector operations (in abstract spaces) 231, 232, 235, 236, 441
 VITALI, G. 71, 149, 152, 230
 's covering theorem 71
 VOLTERRA, V. 134, 465, 482
 's example 134

W

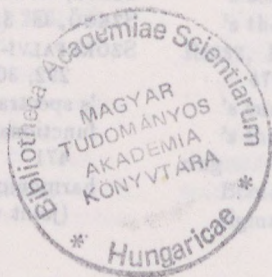
- WADE, W.R. 292
 WAGNER, H.J. 292, 297
 WALSH, J. 297, 482
 's orthonormal system 482
 WATSON, G.N. 285, 348, 355
 wave equation 448, 464
 wave mechanics 448
 WEIERSTRASS, K. 53, 83, 95
 's approximation theorems 95
 's criterion 89
 's example 94, 95
 's fundamental theorem 82
 weight function 138, 270
 well-ordering of sets 39
 WEYL, H. 482
 's fractional integral 393
 WHITTAKER, E.N. 348, 482
 WIENER, N. 390, 406

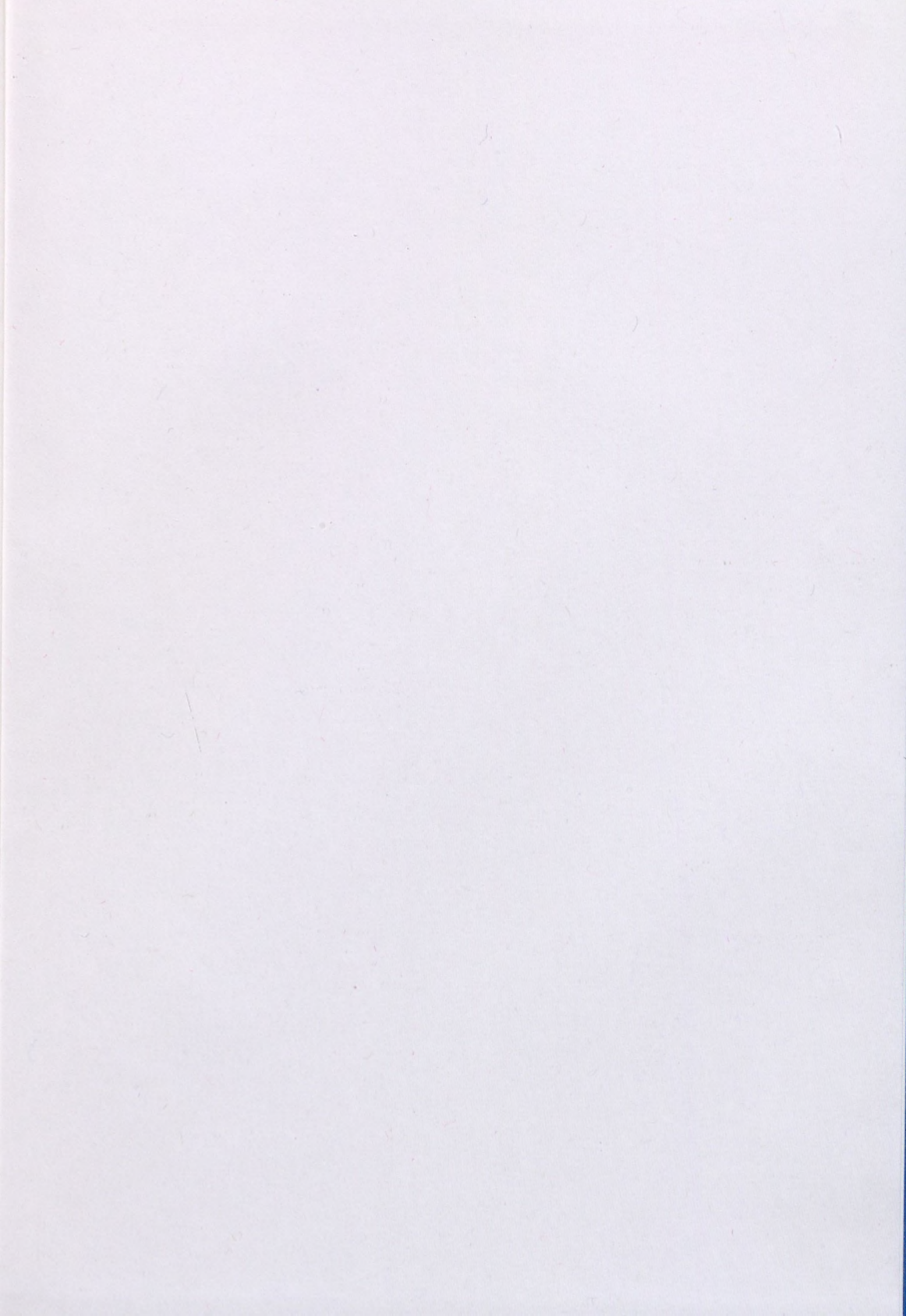
Y

- YOSIDA, K. 471
 YOUNG, G.C. 482
 YOUNG, L.C. 482
 YOUNG, W.H. 90, 106, 149, 154, 337, 384, 438
 's classification of functions 90
 's lower and upper integral 154
 's theorem 94

Z

- ZELLER, K. 334, 385
 ZERMELO, E. 39, 152
 -Fraenkel axiom system 39
 zero element (in abstract spaces) 231, 236, 241
 zeta-functions
 Hurwitz's 345, 346
 Lerch's 352
 Riemann's 347, 348
 ZYGMUND, A. 95, 301, 334, 336, 384, 393, 408, 409, 431





5740



Ideas and results arising from the modern theory of real functions are indispensable today to the more profound bases of probability theory, discussion of generalized functions (distributions), theory of approximation, to solving partial and ordinary differential equations and integral equations, in the various branches of mathematical physics and engineering. Lebesgue's theory of measure and integration, furthermore the related modern Fourier analysis — including functional analytic expedients, the convergence and summation theory of orthogonal series — play an especially prominent role in applications.

As this book was written primarily for graduate students, didactical points of view are taken into consideration throughout. The first half focuses on surveying the theory of real functions, while the second half deals with general and special orthogonal series, both extensively relying on pertinent results obtained from the Hungarian analytical school. Most recent among these topics are e.g. new connections between integro-differential operators and summation methods found by the author.

Readership: graduate students of mathematics, physics, electrical and mechanical engineering, as well as researchers and practitioners in the above fields.

0516

REAL FUNCTIONS,
MIKLÓS MIKOLÁS • ABSTRACT SPACES AND ORTHOGONAL SERIES

