A History of Infinite Matrices

A Study of Denumerably Infinite Linear Systems as the First Step in the History of Operators Defined on Function Spaces

MICHAEL BERNKOPF

Communicated by M. KLINE

Content

1. Introduction

The present paper traces the development of the theory of infinite matrices and allied theories. These topics are considered as only the first part of a more general history of operators defined on function spaces.

The history of a general theory of infinite matrices begins, as we shall see, with HENRI POINCARÉ in 1884. His interest was excited by two papers, written by others (see p. 316, below), which used infinite matrices and determinants without logical justification, and it was his purpose to provide a rigorous basis for these works. After POINCARÉ, HELGE VON KOCH was the next to take up the study, and by 1893 he had proved all of the "routine" theorems about infinite matrices and their determinants. In 1906, a tremendous impulse was given to the subject when DAVID HILBERT used infinite quadratic forms, which are equivalent to infinite matrices, to solve the integral equation

$$f(s)=\varphi(s)+\lambda\int_a^b K(s,t)\,\varphi(t)\,dt.$$

HILBERT's ideas were taken up by his followers — ERHARD SCHMIDT, ERNST HELLINGER and OTTO TOEPLITZ, among others — and within a few years many of the theorems fundamental to the theory of more abstract operators had been discovered, although they were couched in special matrix terms. Finally, in 1929, JOHN VON NEUMANN showed that the theory of infinite matrices was not the effective tool for the study of operators on function spaces; instead, he demonstrated that an abstract approach was preferable.

It is not difficult to understand why infinite matrices were among the first tools to be considered in the study of function space operators. The earliest of the spaces looked at were all sets of infinite sequences of numbers, and it is obvious to consider sequences as generalizations of n-tuples. Since finite matrices correspond to the natural linear operators on finite dimensional spaces, it is but a short step to conceive of infinite matrices, the analogous extension of finite matrices, as the natural linear operators defined on sequence spaces. We shall see some of the difficulties connected with this approach.

There is another obvious manner in which infinite matrices can be generated, this time by problems from analysis. Consider, for example, the differential equation

$$\text{(i)} \qquad \frac{du}{dz} + u\,f(z) = 0$$

where $f(z)$ has a known Laurent expansion

$$\text{(ii)} \qquad f(z) = \sum_{n=-\infty}^{\infty} f_n z^n$$

valid in some annulus A about the origin. For simplicity, suppose there exists an unknown solution of (i) which has the expansion

$$\text{(iii)} \qquad u = \sum_{n=-\infty}^{\infty} u_n z^n$$

also valid in A. Then, substituting (ii) and (iii) into (i) yields [here we are ignoring all but formal considerations]

$$\sum_{n=-\infty}^{\infty} n\,u_n z^{n-1} + \left(\sum_{n=-\infty}^{\infty} u_n z^n\right)\left(\sum_{n=-\infty}^{\infty} f_n z^n\right) = 0.$$

This is easily transformed into

$$\sum_{n=-\infty}^{\infty} (n+1)\,u_{n+1} z^n + \sum_{n=-\infty}^{\infty}\left(\sum_{k=-\infty}^{\infty} u_k f_{n-k}\right) z^n = 0.$$

Since the coefficients of z^n must now vanish for each n, we are led to the infinite homogeneous system of equations

$$\text{(iv)} \qquad (n+1)\,u_{n+1} + \sum_{k=-\infty}^{\infty} u_k f_{n-k} = 0 \qquad (n = \cdots -1, 0, 1, \ldots).$$

We can consider that (iv) is a matrix equation $MU=0$ where U is the unknown vector $(\ldots, u_{-1}, u_0, u_1, \ldots)$ and M is the known matrix

$$\begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & ⋰ \\ \cdots & f_1 & f_0 & f_{-1} & f_{-2} & f_{-3} & \cdots \\ \cdots & f_2 & f_1 & f_0 & 1+f_{-1} & f_{-2} & \cdots \\ \cdots & f_3 & f_2 & f_1 & f_0 & 2+f_{-1} & \cdots \\ ⋰ & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Some of the earliest infinite matrices were derived from similar considerations. In fact, the first and most important studies of infinite matrices came from problems arising out of analysis, rather than from algebra (RIESZ (1; p. 1))*.

Some remarks about terminology and notation are necessary. As with any young subject, the notation and vocabulary was not standardized in the period under review. Wherever possible, we have adhered to an author's original terminology; changes have been made only to avoid confusion. We have also used the term *theory of infinite linear systems* to mean any or all of the following infinite theories: matrices, linear equations, determinants, bilinear forms, or quadratic forms.

Finally, we note that the paper is intended to be a continuation of an earlier work which outlined the history of the function space concept in some detail (BERNKOPF (1)). However, it is self-contained in the sense that it can be read without reference to the previous paper.

2. A Perspective of the Place of Infinite Matrices in the History of Operator Theory

In this section we shall briefly sketch the researches of mathematicians who were working in the theory of operators defined on spaces other than Hilbert sequence spaces. This is not intended to provide a complete history, but rather an orientation so the reader may locate historically the work on infinite matrices more precisely. The events related here have been summarized from BERNKOPF (1).

In the discussion of a history of operators defined on function spaces, it is necessary to establish a working definition of the term "operator." To illustrate the difficulty, consider the matter of integration and integral equations. The process of integration can be considered as a mapping from one set of functions into another. Thus, the problem of finding solutions to a given integral equation may be considered as the problem of determining whether or not a given function lies in the range of a particular transformation. On the other hand, the same equation could be considered as an entity in itself, and explicit methods of solving it (such as an iteration scheme) could be sought.

We shall say that a given work comes under the theory of operators if its primary concern is with the questions of the first type. That is, an *operator* is defined to be a transformation (or a mapping), usually linear, from one function space into another, and the question of whether or not a particular paper is or is not concerned with operator theory becomes a matter of considering the point of view of the author of that paper. The reader is warned, however, that the matter is not clear cut in many cases, and that a certain number of arbitrary decisions have been made.

With this definition of operator in mind, it appears that the theory of operators had its beginnings in the calculus of variations. As early as 1879 KARL WEIERSTRASS (1815—1897) defined an ε neighborhood of a function, and this concept was used by VITO VOLTERRA (1860—1940) to develop his "theory of functions of lines." Especially noteworthy was VOLTERRA's introduction of the notions of

* See the bibliography at the end for references.

continuity and differentials for functionals.* Also working in the same area were GIULIO ASCOLI (1843—1896) and CESARE ARZELÀ (1847—1912). They hoped to generalize CANTOR's theory of point sets to a theory which would include sets of functions, and then to apply the results principally to the calculus of variations (SANGER (1), LÉVY (1)). At the beginning of this century, J. HADAMARD (1865—1963) and MAURICE FRÉCHET (1878—) investigated further into the nature of functionals, obtaining some representational theorems (FRÉCHET (1)).

In a different direction, SALVATORE PINCHERLE (1853—1936) introduced, before 1906, a primitive theory of spaces of analytic functions as represented by their power series. He was concerned with linear operators defined in such spaces which he considered in an abstract manner, and particularly with determining under what conditions the equation $A\alpha=\varphi$ would have solutions, where A is a linear operator, φ is a known power series, and α is to be determined (PINCHERLE (1) and (2)).

We mention also a somewhat later (1908) theory developed by E. H. MOORE (1862—1924). MOORE, struck by similarities between HILBERT's work on integral equations [see below, section 6] and the theory behind the solution of (finite or infinite) systems of linear equations, was led to attempt a generalization which would include all these theories. He looked at families of real valued functions, defined a generalized form of convergence for sequences of these functions, and then considered functionals whose domains were those sets of functions. It is evident from the form of MOORE's work that he was consciously working in the realm of operator theory (MOORE (1)).

In spite of all this diverse activity, it cannot be said that the total impact of any of the above theories was very great. Thus the modern theory of operators starts with FRÉCHET's famous thesis of 1906 (FRÉCHET (2)). FRÉCHET, led by an interest in the calculus of variations, developed the general concepts of the abstract metric space. In so doing, he provided a setting in which the abstract operator point of view could be made more meaningful.

Starting only with the concept of an abstract set on which a limit is defined, FRÉCHET was able to generalize for such sets many of the results of GEORGE CANTOR's (1845—1918) point set theory. Then, by adding more hypotheses, FRÉCHET showed that it was possible to define a metric on a collection of objects which were not points in the then usually accepted sense.

FRÉCHET also worked with functionals defined on his metric spaces. He showed that concepts such as continuity, equicontinuity, completeness, *etc.*, were meaningful for his functionals, and he was able to generalize certain theorems from classical analysis such as ARZELÀ's theorem.

At this point the history of operator theory splits into two fairly distinct schools. One, which I shall call the German school (whose history will be discussed in subsequent sections of this paper), arose as a direct outgrowth of HILBERT's work. Its members were concerned with infinite matrices defined on square summable sequence spaces. Even the Riesz-Fischer theorem of 1907, which showed the isometric-isomorphism between the space of Lebesgue square integrable

* A functional is a real or complex valued transformation whose domain is a function space.

functions and the space of square summable sequences, was largely ignored.* The other school, which could be called the Fréchet school, followed his lead, and worked with an ever increasing degree of abstraction.

After FRÉCHET, the main outlines of the abstract theory became even more discernible. FRIEDRICH RIESZ (1880—1956) in 1910 announced the discovery of L^p $(p>1)$ spaces (RIESZ (2)); these are spaces of functions whose p^{th} powers are Lebesgue integrable. He also showed that the set of continuous linear functionals defined on L^p can be identified in a natural way with L^q where $1/p+1/q=1$. In this same paper, RIESZ also developed the concept of an operator on L^p; *i.e.*, an operator whose domain and range is L^p. In addition, he introduced the concept of the adjoint** of such an operator and found necessary and sufficient conditions for the existence of operator inverses. He then used these concepts to solve the eigenvalue problem for the equation $\varphi(x)-\lambda K(\varphi(x))=f(x)$ in L^2. Here f is a known element of L^2, φ is unknown, λ a scalar, and K a given bounded linear transformation on L^2; this is an obvious generalization of an integral equation of the second kind, see equation (40), p. 327 below. These results were extended by RIESZ in 1918 (RIESZ (3)) to spaces of continuous functions, and in so doing he introduced many of the underlying concepts as well as much of the vocabulary for Banach space theory which first appeared in 1922.

However, one year before the publication of BANACH's well known paper on abstract spaces, EDUARD HELLY investigated the nature of linear functionals defined on sequence spaces. Of particular significance is his introduction of a semi-norm onto the set of such functionals (HELLY (1)). This concept was sharpened to a true norm by HANS HAHN (1879—1934) who used the results to obtain certain integral representation theorems (HAHN (1)).

Nevertheless, it was STEFAN BANACH (1892—1945) who, in 1922, gave the theory of abstract operators defined on rather general spaces its final form. (Later, as we shall see, VON NEUMANN considered abstract operators defined on Hilbert spaces. This a somewhat special case of the general theory.) BANACH (BANACH (1)) listed the axioms of Banach spaces***, and established many of their fundamental properties. He then went on to consider operators on these spaces, and proved many important theorems about them, such as an early form of the principle of uniform boundedness, the contracting mapping theorem, and a spectral radius theorem. The importance of this paper is that BANACH worked entirely in an abstract setting without specific reference to any realizations of his spaces or operators.

In a few years HAHN again became interested in the study of linear functionals (HAHN (2)). Starting with the concept of a Banach space, he was able to show that the set of bounded linear functionals defined on such a space is also a Banach space [such a space is called the adjoint space]. He also proved one form of the Hahn-Banach theorem, which states that a bounded linear functional defined

* In conversation with the author, February, 1967, Professor KURT O. FRIEDRICHS said that the Riesz-Fischer theorem was considered to be a theorem chiefly concerned with Fourier series.

** The *adjoint* of an operator T on L^p is an operator on L^q which can be defined in terms of T. See, for example, DUNFORD & SCHWARTZ (1).

*** A Banach space is a complete normed vector space. See, for example, DUNFORD & SCHWARTZ (1) for a definition; there it is called a B space.

on a proper closed subspace of a Banach space can be extended to the whole space with its norm preserved.

Some two years later, in 1929, BANACH himself considered functionals (BANACH (2)). He first obtained all of HAHN'S results and then went on to prove a more general version of the Hahn-Banach theorem. This theory was then utilized by BANACH to prove a generalized alternative theorem concerning the solvability of the equation $U(x)=y$, where U is a bounded linear operator from a Banach space R into another, S. This then is where the theory of abstract linear operators of the Fréchet school stood before 1930. We note in passing that there was also a theory of non-linear operators being developed; see GRAVES (1) for some results and a bibliography.

Thus work on the abstract theory of operators defined on spaces more general than a Hilbert space was being actively pursued from 1906 to 1930. Meanwhile, the German school, encouraged by HILBERT'S success in using infinite matrices to solve integral equations, continued to consider operators defined on Hilbert spaces from the point of view of infinite matrix theory.

3. Origins and Prehistory

When one consults the earliest works in which systems of infinitely many linear equations in infinitely many unknowns appear, it is evident that in none of these papers is there a general theory under consideration. Instead, each set of equations is taken up on an *ad hoc* basis as a tool for use in the solution of a single particular problem.

Typically, in the seventeenth and eighteenth centuries, infinite systems arose in connection with attempts to obtain series solutions for differential equations (RIESZ (1)). The technique was to suppose that a series solution existed, substitute the series with unknown coefficients into the given equation, and then use the conditions imposed by the original equation to solve for the desired coefficients. This process would, in general (at least for those equations which were successfully dealt with), lead to a set of infinitely many linear equations in the infinitely many unknown coefficients. The early workers were then usually able to develop some type of recursive relation for the coefficients, but in any case they only found it necessary to solve finite systems of equations with finitely many unknowns, albeit infinitely often.

The first to solve infinite systems for which no recursive relation was available, was JOSEPH FOURIER (1768–1830). In 1822 FOURIER published his famed *Théorie Analytique de la Chaleur*, and it is in this work that the solution of an infinite system which used more sophisticated methods than those outlined above was attempted.

From his investigations into the propagation of heat, FOURIER was led to the determination of the coefficients $\{a_n\}$ [we do not follow FOURIER'S notation] for the series

$$\sum_{n=1}^{\infty} a_n \cos[(2n-1)x], \tag{1}$$

so that the function represented by this series would be constant for $-\pi/2 \leqq x \leqq \pi/2$ (FOURIER (1; pp. 187, ff.)). However, he immediately generalized the problem to one involving the representation of "any" function in a series of "multiple arcs of sines and cosines" similar to expression (1). This generalization was not introduced solely for the purposes of generalizing, but because FOURIER felt it necessary "... in order to integrate conveniently the equations of the propagation of heat."

Actually, FOURIER did not tackle the most general problem which he had set for himself. In fact, he considered just those analytic functions whose Maclaurin expansions contain only odd powers of x, and he further restricted himself to developing such functions in series involving nothing but $\sin kx$, $k=1, 2, \ldots$. That is, he considered

$$(2)\qquad \begin{aligned} f(x) &= \sum_{n=1}^{\infty} (-1)^{n+1} \frac{A_{2n-1}\, x^{2n-1}}{(2n-1)!} \\ &= A_1 x - A_3 \frac{x^3}{3!} + A_5 \frac{x^5}{5!} - A_7 \frac{x^7}{7!} + \cdots, \end{aligned}$$

where the set $\{A_{2n-1}\}$ is given. He then supposed that

$$(3)\qquad \begin{aligned} f(x) &= \sum_{n=1}^{\infty} a_n \sin(n x) \\ &= a_1 \sin x + a_2 \sin 2x + a_3 \sin 3x + \cdots, \end{aligned}$$

and proposed to solve for the a_n. Now, if one takes the derivatives of (2) and (3) and sets $x=0$, one sees (ignoring, as FOURIER did, all but formal considerations) that $A_1 = \sum_{n=1}^{\infty} n a_n$. Taking third derivatives yields $A_3 = \sum_{n=1}^{\infty} n^3 a_n$, and in general $A_{2k-1} = \sum_{n=1}^{\infty} n^{2k-1} a_n$ for any positive integer k. Thus FOURIER was led to the system of infinitely many equations

$$(4)\qquad A_{2k-1} = \sum_{n=1}^{\infty} n^{2k-1} a_n, \qquad k = 1, 2, 3, \ldots,$$

for the infinitely many unknowns $\{a_n\}$.

His method for solving these systems was to suppress all but the first m equations and the first m unknowns. The solutions, say $\{a_n^{(m)};\ n=1, 2, \ldots, m\}$, for this $m \times m$ system are then found, where it is clear that the coefficient matrix, $\{c_{jn}\colon c_{jn} = n^{2j-1}\}$, is non-singular, but of course these solutions depend on m. FOURIER now set for himself the task of determining $\lim_{m\to\infty} a_n^{(m)}$, $n=1, 2, \ldots$. We need not follow the tortuous path taken by FOURIER to show that

$$\begin{aligned} \frac{1}{2} a_1 &= A_1 - A_3\left(\frac{\pi^2}{3!} - 1\right) + A_5\left(\frac{\pi^4}{5!} - \frac{\pi^2}{3!} + 1\right) - \cdots, \\ -\frac{2}{2} a_2 &= A_1 - A_3\left(\frac{\pi^2}{3!} - \frac{1}{2^2}\right) + A_5\left(\frac{\pi^4}{5!} - \frac{1}{2^2}\frac{\pi^2}{3!} + \frac{1}{2^4}\right) - \cdots, \\ \frac{3 a_3}{2} &= A_1 - A_3\left(\frac{\pi^2}{3!} - \frac{1}{3^2}\right) + A_5\left(\frac{\pi^4}{5!} - \frac{1}{3^2}\frac{\pi^2}{3!} + \frac{1}{3^4}\right) - \cdots, \end{aligned}$$

and similar expressions for the other unknowns. His reasoning required lengthy calculations which are not very illuminating, and which occasionally needed some patching. For example, at one point FOURIER (1; p. 191) writes down a series of fractions each of whose denominators is infinite but whose numerators are constant.*

FOURIER's brief treatment of infinite matrices seems, to the modern reader, to be incredibly naive. The entire discussion teems with unasked convergence questions. In particular, the possibility of rapid divergence of each of the series in expression (4) would immediately call for the establishment of the existence of solutions. Yet FOURIER's intuition was so perceptive that for his purposes the treatment worked. Equally, we are forced to observe that a meaningful development of a theory for infinite matrices had not yet begun.

It would seem at first glance that FOURIER's scheme of suppressing all but the first m equations and m unknowns, solving the resulting equations, and then seeing if the limits as m becomes infinite exist and are meaningful as solutions would be a promising one. However, RIESZ (1; p. 8), who calls this technique the *principe des réduites*, points out that this method will only work in a limited number of cases which call for extremely restrictive hypotheses. Yet this fact was not discovered for another sixty years.

For a half century FOURIER's work on infinite matrices went almost unnoticed, even in France (see p. 317, below). According to RIESZ (1; p. 8), there was only a single paper, published in 1828, which acknowledged using his method.** However, two other authors, EDUARD FÜRSTENAU and TH. KÖTTERITZSCH worked independently of each other and of FOURIER on infinite systems during this period. Of these, KÖTTERITZSCH's paper (1) is by far the more interesting of the two.

KÖTTERITZSCH's paper contains some points of interest and also some curiosities. As an example of the latter, he first considers the finite system [we do not follow his notation]

$$\sum_{i=1}^{n} a_{ik} x_i = u_k, \quad k = 1, 2, \ldots, n. \tag{5a}$$

He then observes that a solution, $\{x_k\}$, can be written as

$$x_k = \sum_{i=1}^{n} \frac{A_{ik}}{|A|} u_i, \quad k = 1, 2, \ldots, n. \tag{5b}$$

Here $|A|$ is the determinant of the $n \times n$ coefficient matrix $A = \{a_{ik}\}$ and A_{ik} is the ik^{th} minor of A. This is, of course, just CRAMER's rule. Next, he notes that the form of (5b) is not changed if another variable is added to each equation in (5a) and another equation is also added, thus making (5a) an $(n+1) \times (n+1)$ system. From this, he argues (KÖTTERITZSCH (1; p. 2)), "If the system [5a] ... is so constituted that on it the number [n] of equations grows infinite in exactly the same manner as the number of unknowns, then ...

$$x_k = \frac{A_{1k}}{R} u_1 + \frac{A_{2k}}{R} u_2 + \cdots, \tag{5c}$$

* The "annotated" English translation by FREEMAN (FOURIER (2)) lets this and other equally remarkable statements by FOURIER pass without comment. Compare, for example, DARBOUX (FOURIER (1); p. 191) with FREEMAN, p. 172.

** This was by GABRIO PIOLA (1) who published actively from 1822 to 1856. I have not been able to see this paper and hence cannot verify the reference.

where R is the determinant of $\lim_{n=\infty} n^2$ elements of the given system of infinitely many equations and A_{ik} is the coefficient of a_{ik} in R." No further comments or definitions are made to give these new concepts meaning, and at no time is convergence even acknowledged as being a matter for discussion.

However, KÖTTERITZSCH's paper does make some advance. He first gets an explicit solution (again, ignoring convergence questions) for the special *upper triangular* case of

$$\sum_{k=1}^{\infty} a_{ik} x_k = \alpha_i, \qquad i = 1, 2, \ldots, \tag{5 d}$$

where $a_{ik} = 0$ for $i > k$. He then shows that an arbitrary system of the type of (5 d) can be converted, by Gaussian elimination, to the upper triangular case under the assumption that the diagonal minors do not vanish; that is, if $A_n = \{a_{ij}: i, j = 1, 2, \ldots, n\}$, then $\det A_n \neq 0$ for $n = 1, 2, \ldots$. Under this hypothesis, system (5 d) is reduced to

$$b_{11} x_1 + \sum_{k=2}^{\infty} b_{1k} x_k = \beta_1,$$

$$b_{22} x_2 + \sum_{k=3}^{\infty} b_{2k} x_k = \beta_2,$$

$$\vdots$$

and it is easy to see that $b_{nn} = \det A_n \neq 0$. Now, to solve for x_n, the x_{n+p} ($p = 1, 2, \ldots$) are eliminated from all but the first $n-1$ equations which gives

$$x_n = \sum_{k=n+1}^{\infty} B_{nk} \beta_k$$

where the B_{nk} are functions of the b_{nk}. KÖTTERITZSCH points out that the technique has special importance in the application of the method of undetermined coefficients, particularly for Fourier series.

The significance of this work is that, for the first time, a general system of equations is under consideration. However, KÖTTERITZSCH seems to be unaware that he has done anything remarkable in extending the concepts of determinant, minor, *etc.*, to infinite matrices, and in particular, that there were convergence questions to consider. It is interesting to note that KÖTTERITZSCH alludes to Fourier series by name, and so must have been aware of FOURIER's work, yet makes no mention of the Frenchman's discussion of infinite systems.

4. Poincaré and the Beginning of a General Theory

There were two papers which triggered the modern theory of infinitely many equations in infinitely many unknowns. One, by the French mathematician PAUL APPELL (1855—1930), was published in 1884; the other was written by the American astronomer G. W. HILL (1838—1914) in 1877 and first appeared in Europe in February, 1886. Each provided inspiration for HENRI POINCARÉ (1854—1912), and it is with POINCARÉ that the modern theory begins.

As was the case with the earlier men who had worked with infinite systems of equations, APPELL was brought to consider such systems by his interest in analysis. His particular problem (APPELL (1)) was to find an "elementary" method for determining the coefficients of the power series of certain elliptic (doubly periodic) functions. In his solution he used the technique of equating of coefficients; this led him to an infinite set of linear equations to which he applied the *principe des réduites*.

Today it is a matter of conjecture as to how much contemporary interest was generated by APPELL'S work. We do know that a scant two weeks after its presentation POINCARÉ was sufficiently impressed by what he considered to be the usefulness of APPELL'S method to give a general treatment of it. He says (POINCARÉ (1; p. 19)), "As equations of the same form can be encountered in other problems, it is important to inquire into what cases one can legitimately use the method [of *principe des réduites*] which was handled so well by M. APPELL ... ." Thus for the first time, infinite linear systems were solved abstractly without *prior* reference to any particular problem. That is, a general solution is first constructed and then applied to the special case of APPELL'S problem.

[As an aside, we note that FOURIER and his *Théorie Analytique de la Chaleur* receive no mention either by APPELL or POINCARÉ in the two papers cited above. Yet, it is almost impossible to believe that neither one of them had read it. On the other hand the passage quoted immediately above, as well as another remark made in connection with HILL'S work (see below), would indicate that at the very least, POINCARÉ had not seen the section of FOURIER'S work dealing with infinite linear systems.]

We outline POINCARÉ'S paper (POINCARÉ (1)) in some detail. He started by considering an infinite sequence of complex numbers $\{a_n\}$ with $|a_{n+1}|>|a_n|$ and $\lim\limits_{n\to\infty}|a_n|=\infty$, and he wished to find a solution sequence $\{A_n\}$ with

$$\sum_{n=1}^{\infty} A_n a_n^p = 0, \qquad p=0, 1, 2, \dots . \tag{6}$$

That is, the p^{th} equation has p^{th} powers of $\{a_n\}$ as coefficients. This particular type of infinite system was similar to the system considered by APPELL. In general, system (6) will not have a solution; more hypotheses are needed. In order to supply the missing hypotheses, POINCARÉ points out that by a theorem of WEIERSTRASS there exists an entire function F which has simple zeros precisely at the a_n's. For simplicity, POINCARÉ assumes that F can be written as

$$F(x)=\prod_{n=1}^{\infty}\left(1-\frac{x}{a_n}\right). \tag{7}$$

Now let $\{c_n\}$ be a sequence of concentric circles centered at the origin, so that the radius r_n of c_n satisfies $|a_{n-1}|<r_n<|a_n|$. The hypothesis POINCARÉ needed can now be stated in terms of the function F. *The infinite system* (6) *has a solution*, $\{A_n\}$, *if*

$$\lim_{n\to\infty}\oint_{c_n}\frac{x^p}{F(x)}\,dx=0 \tag{8}$$

for every p.

From (8) it is now easy to see that (6) has a solution, for if A_i is the residue of $(F(x))^{-1}$ at a_i, then $\{A_i\}$ is a set of solutions for (6). In fact, the solutions are quickly computed:

$$A_i = \frac{-a_i}{\prod\limits_{\substack{n=1\\ n\neq i}}^{\infty}\left(1-\frac{a_i}{a_n}\right)}. \tag{9}$$

Unfortunately, as POINCARÉ pointed out, the solution $\{A_i\}$ given by (9) may not be unique. Let

$$S_p = \sum_{n=1}^{\infty} |A_n a_n^p|, \tag{10}$$

and let $\{\lambda_p\}$ be such that

$$\sum_{p=0}^{\infty} \lambda_p S_p \tag{11}$$

converges absolutely. When these conditions are satisfied, then $\{B_i\}$ will also be a solution for (6) where

$$B_i = A_i\left(\sum_{p=1}^{\infty} \lambda_p a_i^p\right).$$

In fact, it is not too hard to see that under some circumstances (*cf.* RIESZ (1; p. 17)) any set $\{B_i\}$ will be a solution of (6).

After this treatment of equation (6), POINCARÉ then generalized his discussion of infinite systems, first to a set of homogeneous equations generated by a coefficient matrix $\{\alpha_{ij}\colon i, j=0, 1, 2, \ldots\}$, and then [in order to get APPELL's result] to a system generated by a given sequence $\{a_n\colon n=0, \pm 1, \pm 2, \ldots\}$ where $|a_{n+1}| > |a_n|$ and $\lim\limits_{n\to\infty} |a_n| = \infty$ and $\lim\limits_{n\to-\infty} |a_n| = 0$. In each case hypotheses and results are analogous to the case considered above.

One year later, POINCARÉ was inspired to return to the study of infinite systems by a paper of G. W. HILL (HILL (1)). In his astronomical investigations, HILL was led to the differential equation

$$D^2 w = \theta\, w \tag{12}$$

where D [in HILL's original notation] denotes the differential operator $-i\, d/d\tau$. Suppose that in (12)

$$\theta = \sum_{k=-\infty}^{\infty} \theta_k \zeta^{2k} \tag{13}$$

where $\zeta = e^{\tau i}$ and $\theta_{-k} = \theta_k$, $k = 1, 2, \ldots$, and furthermore suppose there exists a solution for (12) of the form

$$w = \sum_{k=-\infty}^{\infty} b_k \zeta^{c+2k} \tag{14}$$

in which the b_k are all constants. Then after substituting (13) and (14) into (12), HILL constructed the following infinite system of homogeneous equations:

$$\begin{array}{c} \vdots \\ \cdots[-2]\,b_{-2}-\theta_{-1}b_{-1}-\theta_2 b_0-\theta_3 b_1-\theta_4 b_2-\cdots=0, \\ \cdots-\theta_1 b_{-2}+[-1]\,b_{-1}-\theta_1 b_0-\theta_2 b_1-\theta_3 b_2-\cdots=0, \\ \cdots-\theta_2 b_{-2}-\theta_1 b_{-1}+[0]\,b_0-\theta_1 b_0-\theta_2 b_2-\cdots=0, \\ \vdots \end{array} \tag{15}$$

where $[k]=(c+2k)^2-\theta_0$, $k=0, \pm 1, \pm 2, \ldots$.

Concerning system (15) he said (HILL (1; p. 18)), "These conditions determine the ratios of all the coefficients b, to one of them, as b_0, which then may be regarded as an arbitrary constant." Observe, not one word as to how this is to be done. Further on (p. 19) he continues, "If, from this group of equations, infinite in number, and in number of terms in each equation also infinite, we eliminate all the b's except one, we get a symmetrical determinant involving c, which, equated to zero, determines this quantity." Still further along, on p. 26, in connection with another determinant, he adds, "*The question of convergence*, so to speak, of a determinant, consisting of an infinite number of constituents, *has nowhere, so far as I am aware, been discussed* [the emphasis has been supplied]. All such determinants must be regarded as having a central constituent; when, in computing in succession the determinants formed from the 3^2, 5^2, 7^2, etc., constituents symmetrically situated with respect to the central constituent, we approach, without limit a determinate magnitude, the determinant may be called convergent, and the determinate magnitude is its value. In the present case, there can scarcely be a doubt that as long as $\sum \theta_k \zeta^{2k}$ [see (13)] is a legitimate expansion of θ, the determinant ... must be regarded as convergent." No further comments were made by HILL on the new concept of infinite determinants which he had just introduced.

We let POINCARÉ give a contemporary reaction to HILL's work (POINCARÉ (3; p. xiii)).

> The solution adopted by M. HILL is as original as it is bold .... Did one have the right to set the determinant of these equations equal to zero? M. HILL ventured to do so and it was a very daring thing to do; until then an infinite number of linear equations had never been considered [*sic*!]; determinants of infinite order had never been studied; no one even knew how to define them, and it was not certain that it was possible to give a precise meaning to this notion. I must add, however, for sake of completeness, that M. KÖTTERITZSCH had touched on the subject .... But his paper was hardly known in the scientific world and in any case was not known to M. HILL. ...
>
> But it is not enough to be daring; daring must be justified by success. M. HILL successfully avoided all the traps that surrounded him; and let no one say that in proceeding this way he exposed himself to the most glaring errors; no, if the method had not been legitimate, he would have been immediately warned, because he would have arrived at a numerical result completely different from that given by observations.

These words were written in 1905, but they still reflect the excitement that POINCARÉ must have felt when he first read HILL's paper [probably in 1884 or 1885]. Thus, it is no wonder that in 1886 the Frenchman once again took up the study of infinite systems; he felt it was necessary to tidy up after HILL by mathematically justifying the assumptions made by the astronomer.

After repeating the results of his earlier paper, POINCARÉ considered the infinite matrix (POINCARÉ (2)) $\{a_{ij}: i, j=0, \pm 1, \pm 2, \ldots\}$, with $a_{ii}=1$. He then set

$$\Delta_n = \det \begin{vmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & 1 & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} \cdots & a_{n,n-1} & 1 \end{vmatrix}.$$

Next, he defined the determinant Δ of the tableau T to be $\lim_{n\to\infty} \Delta_n$, if this limit exists.

He then showed that the determinant will exist whenever

$$\sum_{\substack{n,p=-\infty \\ n\neq p}}^{\infty} |a_{np}| < \infty.$$

After this, a general theorem about infinite determinants was proved. *Let* $\{x_i: i=0, \pm 1, \pm 2, \ldots\}$ *be a bounded sequence, and let* T' *be the matrix obtained by replacing one row of* T *by* $\{x_i\}$. *Then if* Δ *exists so will* Δ', *the determinant of* T'. The balance of this paper used these results to derive those of HILL. POINCARÉ finished by adding (2; p. 90), "After the above development, I believe that there can be no further objections to the fine method of M. HILL."

The results included in these two papers of POINCARÉ's are disappointing. One would have expected a deeper analysis from him, once he got started on the study of infinite systems. Still, these works are significant in that they represent the beginning of a rigorous treatment of the subject. Two particular points should be noted. First, even at this stage the pathological properties of infinite matrices have appeared, as is seen from the possible plethora of solutions to system (6). Second, and perhaps more significant, is the introduction of analysis into what at first seems to be a purely algebraic problem (see (7) and (8)). As we now know, analytical considerations became even more pronounced as the subject evolved into abstract operator theory, until the techniques of analysis were dominant.

5. Helge von Koch

The first mathematician to attempt a broad and extensive theory of infinite matrices was HELGE VON KOCH (1870—1924) beginning in 1891. His investigation began as a by-product of an interest in FUCHS' equation (VON KOCH (1)).

Consider then FUCHS' equation, given by

$$P(y) = \frac{d^n y}{dx^n} + P_2(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + P_n(x)\,y = 0, \tag{16}$$

where for $r=2, 3, \ldots, n$ each $P_r(x)$ can be represented by a Laurent expansion,

$$P_r(x) = \sum_{\lambda=-\infty}^{\infty} \alpha_{r\lambda} x^{\lambda} \tag{17}$$

valid in the same annulus A centered about the origin. It was already known that a solution

$$y = \sum_{\lambda=-\infty}^{\infty} g_\lambda x^{\lambda+\varrho} \tag{18}$$

existed which converged throughout A. VON KOCH's problem was to calculate a general formula for both the coefficients g_λ and the exponent ϱ of (18). Heretofore, this had only been done for special cases. The computations led VON KOCH (by a series of transformations) to an infinite matrix of the type considered by POINCARÉ. He was able to use POINCARÉ's theory to get explicit representations for the g_λ and for ϱ, but only under certain restrictive hypotheses.

In order to remove these restrictions VON KOCH returned to the subject a year later (VON KOCH (2)). This time he was forced to extend considerably the theory of infinite matrices in order to obtain the results he wanted. Although VON KOCH regretted that little had been done to develop a general theory, he still limited himself to deriving only that much of the theory as he required for his own work in differential equations.

VON KOCH began by considering the infinite array $A = \{A_{ik};\, i, k = \cdots, -1, -2, 0, 1, 2, \ldots\}$, and set

$$D_m = \det\{A_{ik};\, i, k = -m, \ldots, m\}. \tag{19}$$

Then the *determinant* D of A is $\lim D_m$ if this limit exists and is finite; otherwise the determinant of A is said to diverge. The *main diagonal* of A was $\{A_{ii};\, i = -\infty, \ldots, \infty\}$; rows and columns of A were defined as expected. A_{00} was called the *origin*. It is at once clear that the same infinite array can give rise to denumerably many infinite matrices, all with the same main diagonal, and the determinant will not be fixed until an origin has been selected. Thus, VON KOCH's first task was to show that if D existed for one particular choice of origin, it would exist and be the same for any origin; that is, D is a function of the array A itself and does not depend upon the particular enumeration used.

To establish that convergence (alone) was independent of the choice of origin, VON KOCH proved the following:

Theorem. *Let D be an infinite determinant. Then in order that D converge, it is sufficient that the product of the elements on the main diagonal converge absolutely, and that the (double) sum of the elements off the diagonal also converge absolutely.*

Proof. Construct a_{ik} by setting

$$A_{ik} = \delta_{ik} + a_{ik} \qquad (i, k = -\infty, \ldots, \infty). \tag{20}$$

Then by hypothesis [and from the theory of infinite products*],

$$\sum_{i=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |a_{ik}| < \infty. \tag{21}$$

* A necessary and sufficient condition that $\prod_{j=1}^{\infty} (1 + b_j)$ converge absolutely is that $\sum_{j=1}^{\infty} b_j$ converge absolutely.

Consequently, again from the theory of infinite products,

$$\bar{P} = \prod_{i=-\infty}^{\infty} \left(1 + \sum_{k=-\infty}^{\infty} |a_{ik}|\right) \tag{22}$$

converges. Now from

$$P_m = \prod_{i=-m}^{m} \left(1 + \sum_{k=-m}^{m} a_{ik}\right) \tag{23}$$

and

$$\bar{P}_m = \prod_{i=-m}^{m} \left(1 + \sum_{k=-m}^{m} |a_{ik}|\right), \tag{24}$$

von Koch then showed that

$$|D_{m+p} - D_m| \leqq \bar{P}_{m+p} - \bar{P}_m. \tag{25}$$

But the convergence of (22) is just the convergence of $\{\bar{P}_m\}$, which gives the convergence of $\{D_m\}$ by (25). A determinant which is such that the set $\{a_{ik}\}$ satisfies condition (21) will be said to be in *normal form.*

In showing that the value of the limit of a convergent determinant is independent of the choice of origin, von Koch actually proved more. Let

$$D_{mn} = \det\{A_{ik};\ i, k = -n, \ldots, m\},$$

and similarly, let

$$\bar{P}_{mn} = \prod_{i=-n}^{m} (1 + |a_{ik}|).$$

Note that D_{pp} and $\bar{P}_{pp}$ are the same as D_p and $\bar{P}_p$, respectively. This led von Koch to the

Theorem. *Let A be in normal form. Then* $\lim\limits_{\substack{n\to\infty\\ m\to\infty}} D_{mn} = D$.

Proof. By the previous theorem we know that D is finite and that (22) converges. Now for any pair (m, n), let $p = \max(m, n)$. Then, as before,

$$|D_{pp} - D_{mn}| \leqq \bar{P}_{pp} - \bar{P}_{mn}.$$

The right hand side can be made arbitrarily small for sufficiently large m and n (hence, also for sufficiently large p), because of the convergence of (25). The triangle inequality can now be applied to the last inequality to give the result.

Following the theorem, certain properties of D were deduced under the assumption that D is in normal form: If any row or column of A is replaced by a bounded sequence of numbers, the new determinant will also be convergent. If two rows (or columns) are interchanged, the new determinant will have $-D$ as its numerical value. von Koch also implies, but does not state, that if a row or column of A is multiplied by a constant c, then the new determinant will have the value cD.

von Koch next showed that various techniques can be used to compute D. For example, he stated that

$$D = \sum_{\varphi} \pm \cdots A_{-m\varphi(-m)} \cdots A_{0\varphi(0)} \cdots A_{m\varphi(m)} \cdots \tag{26}$$

where the sum is to be taken over all permutations φ and the sign of each term is determined by the parity of φ. [VON KOCH does not state what he means by a permutation of an infinite set, nor what he means by its parity. Presumably, one is to permute only a finite set of numbers at one time and calculate parity by counting the number of interchanges of the permuted numbers.]

Using (26), VON KOCH developed an interesting proof of the fact that D can be expanded by minors. It is clear that each term in (26) contains as a factor exactly one entry from each row and exactly one entry from each column of A. Thus D can be considered as a linear functional of any row or any column. Suppose we are interested in an expansion by minors by the i^{th} row. To determine the coefficient of A_{ik}, one replaces A_{jk} $(j \neq i)$ by zero and A_{ik} by 1 in A, and calculates the resulting determinant which VON KOCH denoted by

$$\text{adj } A_{ik} = \binom{i}{k} = \alpha_{ik} = \frac{\partial D}{\partial A_{ik}}. \tag{27}$$

The α_{ik} will be called *minors* or *subdeterminants of order one*. From these considerations, it is immediate that

$$D = \sum_{k=-\infty}^{\infty} A_{ik}\,\alpha_{ik} \tag{28}$$

which is analogous to the usual expression for the expansion of D by minors of the i^{th} row for finite matrices. Similarly, the expansion by the k^{th} row can be given by

$$D = \sum_{i=-\infty}^{\infty} A_{ik}\,\alpha_{ik}. \tag{29}$$

Exactly as in the case of finite matrices

$$\sum_{i=-\infty}^{\infty} A_{ij}\,\alpha_{ik} = 0 \qquad (j \neq k) \tag{30}$$

and

$$\sum_{k=-\infty}^{\infty} A_{jk}\,\alpha_{ik} = 0 \qquad (j \neq k). \tag{31}$$

It is also clear that α_{ik} can be calculated by suppressing the i^{th} row and the k^{th} column of A, finding the new determinant D' and then taking $\alpha_{ik} = (-1)^{i-k} D'$.

These ideas can also be extended to expand D by two or more rows (or columns). Suppose we wish to expand by the i^{th} and m^{th} row. Then, to find the coefficient of $A_{ik}A_{mn}$, we replace A_{ik} and A_{mn} by one in A, and all other entries in rows i and m by zero. The determinant of this new matrix is called a *minor of the second order* and is designated by

$$\text{adj} \begin{vmatrix} A_{ik} & A_{in} \\ A_{mk} & A_{mn} \end{vmatrix} = \frac{\partial^2 D}{\partial A_{ik} A_{mn}} = \begin{pmatrix} i & m \\ k & n \end{pmatrix}.$$

By interchanging the k^{th} and n^{th} columns, one sees that

$$\begin{pmatrix} i & m \\ n & k \end{pmatrix} = -\begin{pmatrix} i & m \\ k & n \end{pmatrix}.$$

Consequently the expansion for D can be written as

$$D = \sum_k \sum_n \begin{vmatrix} A_{ik} & A_{in} \\ A_{mk} & A_{mn} \end{vmatrix} \begin{pmatrix} i & m \\ k & n \end{pmatrix}$$

where $-\infty < n < \infty$ and $k < n$. [Here we have used the notation

$$\begin{vmatrix} A_{ik} & A_{in} \\ A_{mk} & A_{mn} \end{vmatrix} = \det \begin{pmatrix} A_{ik} & A_{in} \\ A_{mk} & A_{mn} \end{pmatrix}.]$$

Similarly, the r^{th} *order minor* can be constructed by replacing $A_{i_1k_1}, A_{i_2k_2}, \ldots, A_{i_rk_r}$ with the number one in the distinct rows $i_1, i_2, \ldots, i_r$ and columns $k_1, k_2, \ldots, k_r$ of A, and all other elements in those rows or columns with zeros. This minor is designated by

$$\operatorname{adj} \begin{vmatrix} A_{i_1k_1} \ldots A_{i_1k_r} \\ A_{i_2k_1} \ldots A_{i_2k_r} \\ \vdots \quad \ddots \quad \vdots \\ A_{i_rk_r} \ldots A_{i_rk_r} \end{vmatrix} = \begin{pmatrix} i_1\, i_2 \ldots i_r \\ k_1\, k_2 \ldots k_r \end{pmatrix},$$

and one has

$$D = \sum_{k_1} \sum_{k_2} \cdots \sum_{k_r} \begin{vmatrix} A_{i_1k_1} \ldots A_{i_1k_r} \\ \vdots \quad \ddots \quad \vdots \\ A_{i_rk_r} \ldots A_{i_rk_r} \end{vmatrix} \begin{pmatrix} i_1\, i_2 \ldots i_r \\ k_1\, k_2 \ldots k_r \end{pmatrix}, \tag{32}$$

where $k_1 < k_2 < \cdots < k_r$ and $-\infty < k_r < \infty$. This is, of course, the generalization to infinite matrices of the Laplace expansion for D. Also

$$\begin{pmatrix} i_1\, i_2 \ldots i_r \\ k_1\, k_2 \ldots k_r \end{pmatrix}$$

can be calculated by suppressing the appropriate r rows and r columns of A, finding the resulting determinant D', and taking the minor to be $(-1)^p D'$ where $p = \sum_{q=1}^{r} (i_q - k_q)$. There is no need to restrict this type of expansion of D to a method which employs only rows or only columns. As von Koch pointed out, any combination of rows and columns can be used, and in fact they need not form a finite set.

von Koch's final expansion for D is given by the formula

$$D = 1 + \sum_{p=-\infty}^{\infty} a_{pp} + \sum_{p<q} \begin{vmatrix} a_{pp} & a_{pq} \\ a_{qp} & a_{qq} \end{vmatrix} + \sum_{p<q<r} \begin{vmatrix} a_{pp} & a_{pq} & a_{pr} \\ a_{pq} & a_{qq} & a_{qr} \\ a_{rp} & a_{rq} & a_{rr} \end{vmatrix} + \cdots. \tag{33}$$

Here, the largest summation index appearing in each term is to range over all integers, and the others are to range over all integers as indicated. [In von Koch's paper (2; p. 228) the second term on the right hand side of (33) is absent; this seems to be a typographical error.] Expression (33) is particularly important since it is the form used by Ivar Fredholm to solve the integral equation

$$\varphi(x) + \int_0^1 f(x, y)\, \varphi(y)\, dy = \psi(x)$$

(see BERNKOPF (1; p. 8)). VON KOCH did not indicate a proof of (33), satisfying himself with the remark that the proof is analogous to the finite dimensional case.

The usual product theorem was proved next. *Let* $A=\{A_{ik}\}$ *and* $B=\{B_{ik}\}$ *define* $C_{ik}=\sum_{j=-\infty}^{\infty} A_{ij}B_{jk}$, *and* $C=\{C_{ik}\}$ *for* $i, k=-\infty, \ldots, \infty$. *Then, if* det A *and* det B *are in normal form,* det C *is in normal form and* (det A) (det B) = det C. There is no need to reconstruct VON KOCH's proof here, but it depends on getting various estimates and using the triangle inequality. He also noted that the theorem can be proved in a manner similar to the proof of the finite case.

VON KOCH next observed that a determinant may converge, even though it is not in normal form. As an example, he showed that if $A=\{A_{ik}\}$ is such that 1) $\prod_{i=-\infty}^{\infty} A_{ii}$ converges absolutely and 2) there exists a sequence of numbers $\{x_k;\ k=-\infty, \ldots, \infty\}$ so that the double series $\sum_{i=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} A_{ik}x_i/x_k$ converges absolutely, then the determinant of A converges and has the same properties as if it were in normal form.

Also determinants of matrices whose elements are functions were studied. Consider $A(\varrho)=\{A_{ik}(\varrho);\ i, k=-\infty, \ldots, \infty\}$ where each $A_{ik}(\varrho)$ is an analytic function of ϱ in the same domain T, and is continuous and bounded on the boundary of T. Then, as in (19), set

$$D_m(\varrho)=\det\{A_{ik}(\varrho);\ i, k=-m, \ldots, m\}.$$

$D(\varrho)$ is said to be uniformly convergent if the sequence $\{D_m(\varrho)\}$ converges uniformly in the domain T and on its boundary. Thus $D(\varrho)$ is analytic in T. Now as in (20) set $A_{ik}(\varrho)=\delta_{ik}+a_{ik}(\varrho)$, and suppose that the double series $\sum_{i=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}|a_{ik}(\varrho)|$ converges uniformly in T. Then expansions analogous to (26), (28), (29) and (32) all are shown to be valid. Also the expression

$$\frac{dD(\varrho)}{d(\varrho)}=\sum_{i=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\frac{\partial D}{\partial A_{ik}}\cdot\frac{\partial A_{ik}}{\partial \varrho}$$

is proved to hold uniformly in the interior of T, where, recall, $\partial D/\partial A_{ik}$ is just the first order minor $\binom{i}{k}$. This, of course, corresponds exactly to the finite case.

The final investigation of interest in the infinite matrix theory of VON KOCH was the study of the solution of infinitely many equations in infinitely many unknowns. Although he claimed a certain amount of generality, actually he considers only the homogeneous case

$$\sum_{k=-\infty}^{\infty} A_{ik}x_k=0 \qquad (i=-\infty, \ldots, \infty), \tag{34}$$

where, as before, $D=\det\{A_{ik}\}$ is in normal form.

First, suppose $D\neq 0$. Then a solution, $\{x_k\}$, of (34) was sought which satisfied

$$|x_k|\leqq X<\infty \qquad (k=-\infty, \ldots, \infty). \tag{35}$$

Since D is in normal form [by use of (35)],

$$\sum_{k=-\infty}^{\infty} |A_{ik}|\,|x_k| < U \quad \text{for} \quad i=-\infty, \ldots, \infty.$$

Thus, the series

$$S = \sum_{i=-\infty}^{\infty} \sum_{\lambda=-\infty}^{\infty} \binom{i}{k} A_{i\lambda} x_\lambda \tag{36}$$

converges absolutely for each k, and so the order of summation can be interchanged, which gives

$$S = \sum_{\lambda=-\infty}^{\infty} x_\lambda \sum_{i=-\infty}^{\infty} \binom{i}{k} A_{i\lambda}.$$

But, for $\lambda \neq k$, $\sum_{i=-\infty}^{\infty} \binom{i}{k} A_{i\lambda} = 0$ (see (30)); thus

$$S = x_k \sum_{i=-\infty}^{\infty} \binom{i}{k} A_{ik} = x_k D \tag{37}$$

(see (29)). On the other hand, (36) can also be written as

$$S = \sum_{i=-\infty}^{\infty} \binom{i}{k} \sum_{\lambda=-\infty}^{\infty} A_{i\lambda} x_\lambda = 0 \tag{38}$$

since $\{x_k\}$ is supposed to be a solution of (34). Thus, from (37) and (38)

$$x_k D = 0$$

or $x_k = 0$ for $k=-\infty, \ldots, \infty$. That is, if $D \neq 0$, the only solution for (34) which satisfies (35) is the trivial solution, or in VON KOCH's words, there is no solution.

Now suppose $D=0$. VON KOCH showed that unless $A_{ik} \equiv 0$ there would always exist, for some m, a minor of order m which is not zero. Now, determine the indices $i_1, i_2, \ldots, i_r$ and $k_1, k_2, \ldots, k_r$ so that the r^{th} order minor

$$\begin{pmatrix} i_1 & i_2 & \ldots & i_r \\ k_1 & k_2 & \ldots & k_r \end{pmatrix} \neq 0, \tag{39}$$

and also so that if $r > 1$ and if

$$W = \begin{pmatrix} t_1 & t_2 & \ldots & t_\nu \\ u_1 & u_2 & \ldots & u_\nu \end{pmatrix}$$

is any ν^{th} order minor $(1 \leqq \nu < r)$ where the t_j are selected from $\{i_1, i_2, \ldots, i_r\}$ and the u_j are selected from $\{k_1, k_2, \ldots, k_r\}$, then $W=0$. [Recall the lower the order of a minor, the "larger" is the matrix from which it is calculated.] Under these conditions, *every* minor of order less than r will vanish.

Consider equation (34), and suppose the minor is selected as in (39). Then, by using some previously obtained but uncited results, VON KOCH showed that equations $i_1, i_2, \ldots, i_r$ of (34) are linearly dependent on the remaining equations, that $x_{k_1}, x_{k_2}, \ldots, x_{k_r}$ may be selected arbitrarily, and that a solution for (34) is then given by the expression

$$\begin{pmatrix} i_1 & i_2 & \ldots & i_r \\ k_1 & k_2 & \ldots & k_r \end{pmatrix} x_k = \begin{pmatrix} i_1 & i_2 & \ldots & i_r \\ k & k_2 & \ldots & k_r \end{pmatrix} x_{k_1} + \cdots + \begin{pmatrix} i_1 & i_2 & \ldots & i_{r-1} & i_r \\ k_1 & k_2 & \ldots & k_{r-1} & k \end{pmatrix} x_{k_r} \qquad (k=-\infty \cdots \infty).$$

He remarked that analogous results can be obtained for a non-homogeneous case of (34) provided that the right-hand side satisfies a boundedness condition similar to (35). This remark was, as we now know, somewhat optimistic; in fact, it is false unless further hypotheses are put on the matrix A.

This work of VON KOCH is disappointing. It explored only a single aspect of infinite matrices, and it raised more questions than it answered, questions which begged for answers. For example, there is no discussion of eigenvalues, nor is the matter of an *alternative theorem** investigated. It is true that it would be unfair to expect a complete theory to be developed by VON KOCH, since the tools for such a theory simply were not available at the time, yet surely more could have been accomplished to open up this new field. This is particularly true when one considers that it was already apparent that the subject would have far reaching applications in analysis and algebra.

6. Consequences of Integral Equation Theory in the Study of Infinite Systems

After VON KOCH's paper of 1893, the first significant work on the theory of infinite systems was done by DAVID HILBERT (1862—1943) beginning in 1904 (HILBERT (1)). It is true that an important application of VON KOCH's work had appeared in FREDHOLM's solution of the integral equation of the second kind (see (33) above), but this did not advance the infinite matrix theory.

HILBERT, after hearing of FREDHOLM's results, also took up the study of integral equation theory. Initially, HILBERT had no interest in infinite matrices *per se*; he was exclusively concerned with solving the integral equation

$$f(s) = \varphi(s) - \lambda \int_0^1 K(s,t)\,\varphi(t)\,dt. \tag{40}$$

Here f and K are assumed to be known, and φ is a function to be determined. The function $K(s,t)$ is called the *kernel* of equation (40). More precisely, HILBERT wished to extend the previous work of FREDHOLM and to develop an eigenvalue theory for this equation. Consequently, HILBERT's early approach to infinite matrix theory is his own. VON KOCH, as we have seen, had started with a given infinite matrix, and then considered it as the limit of a sequence of its square finite truncations. HILBERT, in his first three papers, never actually had any specific infinite matrix under consideration; instead, he looked at the limit of a sequence of finite matrices which increase monotonically in dimension, but which are not truncations of any single infinite matrix. (See BERNKOPF (1) for details, especially II-1 and II-2.)

HILBERT's first specific approach to the countably infinite problem occurs in his fourth paper, published in 1906. In this work HILBERT observes that the theory of infinite quadratic forms, on the one hand, is an essential extension of the theory of finite forms, and on the other hand, has wide applications in integral equations, in continued fractions, and, of course in the solution of infinite linear

* An *alternative theorem* is of the following type: the system of equations $Bx=y$ either has a unique solution for all y or the associated homogeneous equation $Bx=0$ has a non-trivial solution. In this case $Bx=y$ has a solution only if y satisfies certain orthogonality conditions.

systems. Therefore, he finds it more convenient to tackle the problem from the point of view of infinite quadratic forms rather than considering an infinite system of linear equations with infinitely many unknowns.

HILBERT then undertook the study of the *infinite quadratic form*

$$\sum_{p,q=1}^{\infty} k_{pq} x_p x_q \tag{41}$$

of the infinitely many variables $x_1, x_2, x_3, \ldots$, with the constant coefficients $\{k_{pq}\}$. [We shall sometimes write (41) as $K(x, x)$ where $x = (x_1, x_2, \ldots)$.] Associated with (41) is the *bilinear form*,

$$K(x, y) = \sum_{p,q=1}^{\infty} k_{pq} x_p y_q; \tag{42}$$

HILBERT also introduced the n^{th} *section of* $K(x, y)$, as

$$\sum_{p,q=1}^{n} k_{pq} x_p y_q .$$

In addition the *product form* of the forms $A(x, y)$ and $B(x, y)$ was defined to be the form

$$A\big(B(x, y)\big) = \sum_{p,q,r=1}^{\infty} a_{pq} b_{qr} x_p y_r .$$

This is nothing but the bilinear form associated with the (infinite) product matrix AB. Finally the special form $(x, y) = \sum_{p=1}^{\infty} x_p y_p$ was defined which is (42) with $k_{pq} = \delta_{pq}$; its n^{th} section is denoted by $(x, y)_n$.

HILBERT'S problem can now be stated as this: He wishes to find a *resolvent* (or inverse) form for the expression $(x, y) - \lambda K(x, y)$ where λ represents a real or complex parameter; in other words, a form, $\overline{K}(\lambda; x, y)$, is sought which will satisfy

$$\overline{K}(\lambda; x, y) - \lambda K\big(\overline{K}(\lambda; x, y)\big) = (x, y) .$$

This problem was solved by HILBERT under quite general hypotheses. Specifically, for a *bounded form* $K(x, y)$, namely, a form $K(x, y)$ which satisfies $|K(x, y)| \leqq M$ for all x and y with $(x, x) \leqq 1$ and $(y, y) \leqq 1$, an explicit representation for the resolvent $\overline{K}(x, y)$ was obtained. Then a theorem analogous to the principal axis (diagonalization) theorem for finite dimensional forms was shown to be valid for bounded infinite dimensional quadratic forms.

But it was HILBERT'S introduction of the concept of complete continuity* which proved to be the fundamental tool in showing that there are still more properties of finite dimensional spaces which have their analogues in infinite dimensional sequence spaces. For example, a completely continuous quadratic

* A bounded linear operator T is said to be *completely continuous* (or, more recently, *compact*) if it maps bounded sets into compact ones. This condition, obviously stronger than ordinary continuity, insures that the algebraic kernel of $T - I$ (I is the identity) is finite dimensional. The modern definition, referring to operators, is equivalent to HILBERT'S original definition referring to forms if we observe that the bilinear form $K(x, y)$ is completely continuous if and only if its matrix $\{k_{pq}\}$ defines a completely continuous operator.

form has an eigenvalue representation; that is, it satisfies a simple principal axis theorem. In addition, an alternative theorem holds for an infinite system of linear equations in infinitely many unknowns if the equation has the form $(I+A)x=y$ where A is the matrix of coefficients of a completely continuous form.

This alternative theorem gave HILBERT the means for a fresh attack on the integral equation

$$f(s)=\varphi(s)-\int_0^1 K(s,t)\,\varphi(t)\,dt. \tag{42a}$$

Specifically, he transformed (42a) into

$$x_p-\sum_{q=1}^{\infty} a_{pq}\,x_q=a_p \tag{42b}$$

by taking $\{a_p\}$ and $\{x_p\}$ to be the Fourier coefficients of $f(s)$ and $\varphi(s)$ respectively, and $\{a_{pq}\}$ to be the double Fourier coefficients of $K(s,t)$. Since $K(s,t)$ is supposed to be continuous, $\{a_{pq}\}$ defines a completely continuous form, and thus (42b) is a system satisfying the hypothesis of the alternative theorem. Now, the function $\varphi(s)$ is determined from the known solution $\{x_p\}$ of (42b) when any solution exists, and then this φ is shown to be a solution of (42a). More generally, the alternative theorem for (42b) gives HILBERT an alternative for (42a). The substitution of $\lambda K(s,t)$ (where $K(s,t)$ is a symmetric kernel and λ is a real parameter) for K in the above discussion gave HILBERT his eigenvalue theory. [The preceding paragraphs have been summarized from BERNKOPF (1; II-3 and II-4).]

It would be hard to overestimate the significance of HILBERT'S work in the budding field of functional analysis. His success in opening up the hitherto stubborn subject of integral equations would have, in itself, insured that active research would continue beyond the relatively limited areas HILBERT himself had considered. Also, he was able to define and utilize two concepts which have turned out to be fundamental for the study of linear operators: boundedness and complete continuity.

HILBERT'S chief contribution is that he showed that the techniques of algebra are appropriate to apply to the problems of analysis. He was not the first to use algebra; FREDHOLM'S earlier work on integral equations (FREDHOLM (1)) is but one example. But HILBERT did confirm that the introduction of algebra into analysis was not accidental, as might have been inferred from earlier scattered successes, but that it was a natural tool which would prove to be extremely valuable when fully developed.

The impetus given to the work on infinite systems and integral equations — the two topics tended to merge, at least in Germany — by HILBERT'S work of 1906 was enormous. To young research mathematicians, the theory of infinite systems coupled with its apparent wealth of applications must have seemed like the promised land, and many, particularly in Germany, devoted their energies to the study of infinite matrices while ignoring the abstract theory.* This history can cover only a few of the papers published in this period [but see HELLINGER & TOEPLITZ (1), particularly the footnotes, for a good bibliography].

* FRIEDRICHS, conversation previously cited.

Typical of the work of this time is a pair of papers by OTTO TOEPLITZ (1881—1940) which appeared in 1907. In the first (TOEPLITZ (1)), the so-called Jacobi Transformation for finite matrices was considered and utilized. This transformation gives the result that if

$$S_n = \sum_{i,k=1}^{n} a_{ik} x_i x_k \qquad (a_{ik} = a_{ki})$$

is a quadratic form, and if all the first order minors of S_α ($\alpha = 1, 2, \ldots, n$) do not vanish, then there exists a matrix U_n with $U_n U_n' = S_n^{-1}$. [A' is the transpose of A.] This was used to simplify HILBERT's construction of the resolvent for the bounded infinite real quadratic form

$$S = \sum_{i,k=1}^{\infty} a_{ik} x_i x_k$$

under the hypotheses that S is positive definite and that the zeros (in λ) of $I_n - \lambda S_n$ do not have infinity as a point of accumulation. For each n^{th} section S_n of S the corresponding U_n was constructed, and it was shown that $\lim_{n\to\infty} U_n$ exists and is a real bounded bilinear form. Then U was defined to be this limit, and S^{-1} was taken to be UU'. Finally, the notation S^{-1} was justified by showing that $SS^{-1} = S^{-1}S = I$. It may be that this is the first appearance of the "S^{-1}" notation. Also, in this paper is the probable first use of the term "reciprocal" *(Reziproke)* in connection with infinite matrices.

TOEPLITZ pointed out that if S_n^{-1} is the inverse of the n^{th} section of S — where S is now an arbitrary bilinear form — the sequence $\{S_n^{-1}\}$ may not converge, even if S has a bounded inverse. He also showed that S may have a left inverse but not a right, or *vice versa*. He summed up his results in the following

Theorem. *A real bounded bilinear form S has a bounded right inverse if and only if SS' does not have infinity as a point of accumulation; that is, if the numerical values of λ for which* $\det\left((SS')_n - \lambda I_n\right) = 0$ *are bounded. Similarly, S has a bounded left inverse if and only if $S'S$ satisfies the same condition. SS' and $S'S$ both satisfy the condition when and only when S has a unique (two sided) inverse.*

TOEPLITZ's next paper (TOEPLITZ (2)) presents an application of the theory of infinite matrices to Laurent expansions. In a slight shift in notation he calls a bounded bilinear form A complete *(abgeschlossen)* if A^{-1} exists; *i.e.*, if

$$AA^{-1} = A^{-1}A = I = \sum_{i=-\infty}^{\infty} x_i y_i .$$

[We shall not use TOEPLITZ's terminology in what follows.] Note also the change in the domain of the summation index. Two forms, A and B, are defined to be similar if there exists a complete form P so that $P^{-1}AP = B$. Then he points out that two forms have the same spectrum if they are similar.

He considers next a Laurent expansion

$$f(z) = \sum_{-\infty}^{\infty} a_n z^n. \tag{43}$$

From this expansion a *Laurent form* is constructed which TOEPLITZ writes as

$$\sum_{i,k=-\infty}^{\infty} a_{k-i} x_i y_k, \tag{44}$$

where the coefficients a_{k-i} are taken from (43). One sees that (44) can be written, matrix fashion, as

$$\begin{pmatrix} \cdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots \\ \cdots & a_0 & a_1 & a_2 & a_3 & \cdots \\ \cdots & a_{-1} & a_0 & a_1 & a_2 & \cdots \\ \cdots & a_{-2} & a_{-1} & a_0 & a_2 & \cdots \\ \cdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots \end{pmatrix}.$$

The following connections between Laurent forms and Laurent expansions are given: A form is bounded if and only if $\sum_{-\infty}^{\infty} a_n$ converges absolutely, and in this case the expansion is a single valued analytic function in a neighborhood of the unit circle. The sum (or product) of two forms is a form which corresponds to the sum (or product) of their associated expansions. If a bounded Laurent form has a bounded inverse, then this inverse is also a bounded Laurent form. If a form has no (two sided) inverse, then it has neither a left inverse nor a right inverse. He goes on to note that the unit form I is a Laurent form associated with the constant function one.

Now let $A=[f(z)]$ represent the Laurent form (44) associated with (43), under the further assumption that f is analytic in a neighborhood of the unit circle. Then A is a bounded form. Consider the spectrum of A, *i.e.*, the values of λ for which $A-\lambda I$ has no inverse. But, by the preceding paragraph, $A-\lambda I=[f(z)]-\lambda I=[f(z)-\lambda]$. Thus the spectrum of A will include all values of λ for which $f(z)=\lambda$, with $|z|=1$. Hence, the spectrum of A is the range of f restricted to the unit circle, and so will, in general, include an entire arc. This yields a sufficient condition for the similarity of two Laurent forms, namely that their spectral values be the same with the same multiplicity.

The work of HILBERT on solutions of infinitely many equations in infinitely many unknowns was taken up by, among others, ERHARD SCHMIDT (1876–1959). To him belongs the honor of being the first to employ properties of the underlying Hilbert (sequence) space to determine necessary and sufficient conditions for the solvability of such equations, and his paper (SCHMIDT (1)) represents an application of some earlier work on integral equations (see SCHMIDT (2)). Before SCHMIDT's work only necessary *or* sufficient conditions had been established; SCHMIDT found conditions which are both necessary *and* sufficient.

We summarize the first part of SCHMIDT's paper, in which he introduced geometric concepts into Hilbert space theory. An element of such a space H [SCHMIDT calls his elements functions] is a square summable sequence $z=\{z_n\}$ [we do not follow SCHMIDT's notation] of complex numbers; *i.e.*, a sequence which has the property that $\sum_{p=1}^{\infty} |z_p|^2 < \infty$. A *norm* for z (denoted by $\|z\|$) is defined by taking $\|z\|^2 = \sum_{p=1}^{\infty} z_p \bar{z}_p$;* and the *inner product* of z and w (denoted by (z, w))

* $\bar{z}$ denotes complex conjugate.

is defined as $\sum_{p=1}^{\infty} z_p w_p$; z and w are said to be *orthogonal* if $(z, w)=0$. The sequence of elements $\{z^n\}$ is said to *converge strongly* if $\lim_{\substack{n\to\infty \\ m\to\infty}} \|z^n - z^m\| = 0$, and it is shown that the space H is complete in the norm; that is, if $\{z^n\}$ converges strongly, then there is an element z of H with $\lim_{n\to\infty} z^n = z$.

SCHMIDT also introduced the concept of a closed subspace A of H. A is a closed subspace of H if A is topologically closed under strong convergence, and if it is also algebraically closed under the operations of scalar multiplication and addition. The idea of a basis is also defined, and it is shown by the Gram-Schmidt orthogonalization process that given any basis for A, there is an equivalent orthonormal basis; that is, there exists a set of elements which span A, which are linearly independent, which are pairwise orthogonal, and all of which have norm one. Finally, and most important, given any element z of H and any closed subspace A, there exist unique elements w_1 and w_2 with $z = w_1 + w_2$, where $(w_1, w_2) = 0$, and w_1 is an element of A. w_1 is called the *projection of z on A*; and SCHMIDT calls w_2 the perpendicular function (of z) to A. We shall call w_2 the *part of z perpendicular to A*. It is easy to see that $w_2 = 0$ if and only if $z \in A$.

Consider now the infinite set of homogeneous equations

$$\sum_{p=1}^{\infty} a_{np} z_p = 0 \qquad (n = 1, 2, \ldots), \tag{45}$$

and suppose that for each n

$$\sum_{p=1}^{\infty} |a_{np}|^2 < \infty. \tag{46}$$

A solution for (45) is called *regular* if it is square summable, and SCHMIDT was concerned only with regular solutions. If the element a^n of H is defined by

$$a^n = \{\bar{a}_{np}\} \qquad (n = 1, 2, \ldots), \tag{47}$$

then the system (45) can be written in the inner product notation

$$(\bar{a}^n, z) = 0 \qquad (n = 1, 2, \ldots). \tag{48}$$

Let A be the closed linear subspace of H spanned by the sequence $\{a^n\}$, and let $e^\nu = \{e_{\nu p} = \delta_{\nu p}\}$ $(\nu = 1, 2, \ldots)$. Let φ^ν be the part of e^ν perpendicular to A. Let R be the closed linear subspace spanned by the sequence $\{\varphi^\nu\}$. It is easy to see that R is the orthogonal complement of A in H. Thus from (48), *z is a solution of* (45) *if and only if z is an element of R.*

However, SCHMIDT was not satisfied with this abstract result. He wished to obtain a more specific representation for the solutions of (45). These results are included in Appendix A for the reader who may wish to see these solutions and an indication of SCHMIDT's proofs.

We observe that SCHMIDT finally settled many questions concerning infinite systems of linear equations. Specifically, he determined the solvability of such systems under the hypothesis that the rows of the coefficient matrix are square summable and only regular (square summable) solutions are sought. [It should be noted, however, that many problems remain. For example, what are the most general conditions under which some form of an alternative theorem holds? As far as I have been able to determine, this is still an open question.]

Nevertheless, this work of SCHMIDT has significance beyond the solution of infinite linear equations. As we noted earlier, he was the first to introduce geometric language into Hilbert space theory, and the results obtained by SCHMIDT show that these geometric notions are not mere pedantry. Rather, the concepts of subspace, orthogonality, *etc.*, form an integral part of the circle of ideas centered about the term "function spaces."

After SCHMIDT, the next important work was a 1910 paper of HELLINGER & TOEPLITZ (HELLINGER & TOEPLITZ (2)). It presented amplifications and extensions, as well as proofs of results which were first announced in 1906 (HELLINGER & TOEPLITZ (3)). Their aim was to present an "axiomatic" treatment for a "Calculus of Infinite Matrices." Their use of "axiomatic" is not the same as the current usage. They used the term to mean that their presentation was independent of any specific problem; *i.e.*, infinite matrix theory was to be considered independently of any integral equation or algebraic theories, *etc.* In addition, they also included a foundation for the theory, since the work was not intended to depend on any prior knowledge of infinite matrices or integral equations. Thus the first chapter of the article represents a good summary of the state of the theory up to the time it was written, probably in 1909.

It is interesting to note that although the inspiration for this work on the general theory of bounded infinite matrices was the integral equation

$$f(s) = \varphi(s) + \int_a^b k(s, t)\, \varphi(t)\, dt, \tag{93}$$

nevertheless HELLINGER & TOEPLITZ were interested also in generalizing various problems of algebra. In fact, they go to some lengths to state some of the already well-known classifications of problems from finite linear algebra-matrix theory, and point out that these can be extended to problems involving infinite matrices. However, they do not undertake to solve any of these problems. As we shall see later, in the discussion of VON NEUMANN's work, at least one of these problems, that of unitary equivalence, has no solution in the expected sense.

We cite a few results and definitions from the first chapter of HELLINGER & TOEPLITZ's paper of 1910 before examining some of its novel aspects (HELLINGER & TOEPLITZ (2)). *Schwarz's inequality*

$$\left|\sum_p u_p v_p\right| \leqq \left(\sum_p u_p^2\right)^{\frac{1}{2}} \left(\sum_p v_p\right)^{\frac{1}{2}} \tag{94}$$

is proved, first for finite sums, then for infinite sums. Then, following HILBERT, an infinite sequence of real numbers $\{a_n\}$ is said to define a *linear form* of infinitely many variables $\{x_n\}$ if

$$\left|\sum_{n=1}^{\infty} a_n x_n\right| < \infty. \tag{95}$$

The form is said to be *bounded* if, for all $\{x_n\}$ satisfying

$$\sum_{n=1}^{\infty} x_n^2 \leqq 1 \tag{96}$$

the form satisfies

$$\left|\sum_{n=1}^{\infty} a_n x_n\right| \leqq M < \infty. \tag{97}$$

Using (94) and (97), they then prove that *a necessary and sufficient condition that a linear form defined by $\{a_n\}$ be bounded is that $\{a_n\}$ be square summable;* that is, $\sum_{n=1}^{\infty} a_n^2 < \infty$. In this case the least upperbound of all M's which satisfy (97) is $\sum_{n=1}^{\infty} a_n^2$. This is, of course, the Riesz representation theorem for Hilbert sequence spaces.

Similar concepts are then defined and proved for infinite matrices. The infinite matrix $A=\{a_{pq}\}$ is said to define *bounded bilinear form* when for all $\{x_n\}$ and $\{y_n\}$ satisfying

$$\sum_{n=1}^{\infty} x_n^2 \leqq 1 \quad \text{and} \quad \sum_{n=1}^{\infty} y_n^2 \leqq 1 \tag{98}$$

one has

$$\left|\sum_{p,q=1}^{\infty} a_{pq} x_p y_q\right| \leqq M. \tag{99}$$

Equivalently, (98) and (99) can be written jointly as

$$\left|\sum_{p,q=1}^{\infty} a_{pq} x_p y_q\right| \leqq M \left(\sum_{n=1}^{\infty} x_n^2\right)^{\frac{1}{2}} \left(\sum_{n=1}^{\infty} y_n^2\right)^{\frac{1}{2}} \tag{100}$$

for all square summable $\{x_n\}$ and $\{y_n\}$. In what follows, a bilinear form will be denoted by $A(x, y)$. It is then shown that *every row and every column of A is square summable if A defines a bounded bilinear form.* Furthermore, if $A_n(x, y) = \sum_{p,q=1}^{n} a_{pq} x_p y_q$, then, again for bounded bilinear forms,

$$\lim_{n\to\infty} A_n(x, y) = A(x, y).$$

Next the *product form* $C(x, y) = A B(x, y)$ is defined to be the form associated with the product matrix $C=\{c_{pq}\}$, where

$$c_{pq} = \sum_{r=1}^{\infty} a_{pr} b_{rq},$$

and it is shown that $C(x, y)$ is bounded whenever $A(x, y)$ and $B(x, y)$ are. Furthermore, multiplication for bounded forms is proved to be associative. Notice is also given to the special case of bilinear forms, namely quadratic forms. These are the bilinear forms, $A(x, x)$ where $A=\{a_{pq}\}$ with $a_{pq}=a_{qp}$, $p, q = 1, 2, \ldots$.

Perhaps the most interesting part of the first chapter of this paper of HELLINGER & TOEPLITZ is a section dealing with counter examples. As is, and was, well known in the finite case, if a square matrix has a one-sided inverse, then this inverse is, in fact, two-sided and unique. For infinite matrices, however, things are different. Consider

$$A=\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & . & . & . \\ 0 & 0 & 1 & 0 & 0 & . & . & . \\ 0 & 0 & 0 & 1 & 0 & . & . & . \\ 0 & 0 & 0 & 0 & 1 & . & . & . \\ . & . & . & . & . & . & . & . \end{pmatrix}$$

and

$$B=\begin{pmatrix} b_1 & b_2 & b_3 & b_4 & . & . & . \\ 1 & 0 & 0 & 0 & . & . & . \\ 0 & 1 & 0 & 0 & . & . & . \\ 0 & 0 & 1 & 0 & . & . & . \\ . & . & . & . & . & . & . \end{pmatrix}$$

where $\{b_n\}$ is an arbitrary sequence of numbers. Then $AB=I$, and A has a non-countable set of right inverses. On the other hand, the matrix equation $XA=I$ has no solution, since the entry in the first row and first column will always be zero. However, if a bounded matrix has both a left and right inverse, they are equal to each other; *i.e.*, such a matrix has a unique two-sided inverse.

Several other matrices were also constructed to show that certain conditions may be necessary but not sufficient, and conversely. For example, a bounded matrix has the property that each of its rows (or columns) is square summable. However, this condition is not sufficient for boundedness. Consider the matrix $A=\left\{a_{pq}=\frac{1}{(p+q)^{\frac{3}{4}}}\right\}$. Then every row or column of A is square summable, as is easily shown by the integral test, but this matrix does not define a bounded bilinear form. To see this, take $x=y=\{x_p=y_p=p^{-\frac{5}{8}}\}$. Then x is square summable, but

$$A(x,y)=\sum_{p,q=1}^{\infty} a_{pq}\,x_p\,y_q=\sum_{p,q=1}^{\infty}\frac{1}{(p+q)^{\frac{3}{4}}}\cdot\frac{1}{p^{\frac{5}{8}}}\cdot\frac{1}{q^{\frac{5}{8}}}. \tag{101}$$

Now by suppressing some terms we see, since everything is positive, that (101) is greater than

$$\sum_{p,q=1}^{\infty}\frac{1}{(p+q)^{\frac{3}{4}}}\,\frac{1}{(p+q)^{\frac{5}{8}}}\,\frac{1}{(p+q)^{\frac{5}{8}}}=\sum_{p,q=1}^{\infty}\frac{1}{(p+q)^{2}}. \tag{102}$$

But

$$\sum_{p,q=1}^{\infty}\frac{1}{(p+q)^{2}}=\sum_{n=1}^{\infty}\frac{n}{(n+1)^{2}}$$

which diverges; hence (101) diverges; that is, A does not define a bounded bilinear form.

In the same vein is the consideration of square summability and matrices. The precise statement is that *if* $A=\{a_{pq}\}$ *satisfies*

$$\sum_{p,q=1}^{\infty} a_{pq}^{2}\leqq M<\infty, \tag{103}$$

then the form $A(x,y)$ *is bounded*. However, this condition is only sufficient. A double application of SCHWARZ's inequality (94) proves the theorem, while to show that (103) is not necessary, take $A = I$. I does not satisfy (103) but certainly defines a bounded bilinear form.

Next, HELLINGER & TOEPLITZ examined continuity questions. Every bounded bilinear form is shown to be continuous. Complete continuity is defined, and it is shown that not every bounded (continuous) form is completely continuous. Also, HELLINGER & TOEPLITZ show that (in modern terminology) the unit sphere is not compact. They do this by considering the form $A = \left\{a_{pq} = \delta_{pq}\left(\frac{p}{p+1}\right)\right\}$. Then A defines a bounded, hence continuous, bilinear form. But this form does not assume its least upper bound (equal to one) anywhere in the unit ball since

$$(104) \qquad |A(x,y)| = \left|\sum_{p=1}^{\infty} \frac{p}{p+1} x_p y_p\right| \leqq \left(\sum_{p=1}^{\infty} \frac{p}{p+1} x_p^2\right)^{\frac{1}{2}} \left(\sum_{p=1}^{\infty} \frac{p}{p+1} y_p^2\right)^{\frac{1}{2}}$$

$$(105) \qquad < \left(\sum_{p=1}^{\infty} x_p^2\right)^{\frac{1}{2}} \left(\sum_{p=1}^{\infty} y_p^2\right)^{\frac{1}{2}}.$$

Under the side conditions that $\sum_{p=1}^{\infty} x_p^2 \leqq 1$ and $\sum_{p=1}^{\infty} y_p^2 \leqq 1$ we see that (104) will come arbitrarily close to the value 1 by taking $x_p = y_p = 0$ for $p \neq n$, $x_n = y_n = 1$ and by letting $n \to \infty$, while (105) shows that the value 1 can never be assumed.

The second chapter of the paper by HELLINGER & TOEPLITZ contains new results. Of particular importance is the theorem which they refer to as the theorem on *uniform finiteness of bilinear forms* of infinitely many variables (HELLINGER & TOEPLITZ (2; p. 321)), which leads to what today is called TOEPLITZ's Theorem. Their statement can be paraphrased as follows: Let $A = \{a_{pq}\}$ be a matrix having the property that $\sum_{p,q=1}^{\infty} a_{pq} x_p y_q$ converges for every $x = \{x_p\}$ and $y = \{y_p\}$ in the unit ball. Then $\sum_{p,q=1}^{\infty} a_{pq} x_p y_q$ is uniformly bounded in the unit ball; that is, A defines a bounded bilinear form. This has the immediate consequence (TOEPLITZ's Theorem) that an operator is bounded if it is defined for every x in H, and is an early form of the Closed Graph Theorem.*

To prove TOEPLITZ's Theorem, let A be a linear operator defined for all x in H. If $x = 0$, $Ax = 0$, and A is bounded at x. If x is not zero, there is a scalar α such that αx is in the unit sphere. Hence, by the uniform finiteness theorem restated for operators, $\|A(\alpha x)\| \leqq M$ or $\|A(x)\| \leqq M/|\alpha|$. But $\|\alpha x\| = 1$, or $\|x\| = 1/|\alpha|$; hence $\|A(x)\| \leqq M\|x\|$, which is the statement that A is bounded. [Here we have used modern notation for clarity and $\|x\|$ to denote $(x,x)^{\frac{1}{2}} = \left(\sum_{p=1}^{\infty} x_p^2\right)^{\frac{1}{2}}$.]

Consider then the *Uniform Finiteness*

Theorem. *Let $A = \{a_{pq}\}$ be an infinite matrix for which the double series*

$$(106) \qquad \sum_{p=1}^{\infty}\left(\sum_{q=1}^{\infty} a_{pq} x_p y_q\right)$$

* One version of the Closed Graph Theorem is (DUNFORD & SCHWARTZ (1)): A closed linear operator defined on all of an F-space with values in an F-space is continuous. [An F-space is a generalization of a Hilbert space; for the definition of a closed linear operator, see footnote p. 338 below.]

converges whenever $x=\{x_p\}$ *and* $y=\{y_p\}$ *satisfy*

$$\sum_{p=1}^{\infty} x_p^2 \leqq 1 \quad \textit{and} \quad \sum_{p=1}^{\infty} y_p^2 \leqq 1. \tag{107}$$

Then there is an $M>0$ *so that*

$$\left|\sum_{p=1}^{\infty}\left(\sum_{q=1}^{\infty} a_{pq}\, x_p\, y_q\right)\right| \leqq M \tag{108}$$

for all x *and* y *satisfying* (107).

The proof uses a previous (but uncited) theorem on the uniform finiteness of linear forms which is similar to the above. The idea is to suppose that the theorem is false and from this assumption construct a pair of elements $\xi=\{\xi_p\}$ and $\eta=\{\eta_p\}$ for which $\sum_{p=1}^{\infty}\sum_{q=1}^{\infty} a_{pq}\xi_p\eta_p$ diverges, thus contradicting (106). See Appendix B for details of the proof.

After the paper of 1910 by HELLINGER & TOEPLITZ there appeared many works which extended their results and simplified their proofs; see, for example, SCHUR (1), where the term *norm* is used, presumably for the first time in connection with infinite matrices, but not in the modern sense. In this period we also see early investigations into the relation between infinite matrices and sequences. SCHUR (2), in 1918, proves some theorems concerning what he calls "convergence-preserving" and "convergence-producing" matrices. This is a topic of current interest (see COOKE (1)) but is outside of the scope of this paper. A good exposition on the state of the theory up to 1929 can be found in WINTNER (1).

7. The Work of John von Neumann: Limitations of the General Theory

In the late 1920's JOHN VON NEUMANN (1903—1957) took up the study of Hilbert spaces and operators on these spaces. His particular concern was the study of Hermitian operators (see below), and an examination of his later work shows that it is likely that this interest was, in turn, generated by his studies in quantum mechanics. VON NEUMANN'S contribution to the theory of infinite matrices remains unique. He showed, as we shall see, that the subject of infinite matrices was the wrong road to the study of linear operators defined on Hilbert spaces, even Hilbert sequence spaces. This is, of course, in sharp distinction to the case of finite dimensional spaces, where the theories of matrices and continuous linear operators are equivalent.

To VON NEUMANN belongs the honor of developing abstract Hilbert space theory. In his first paper on the subject in 1929 (VON NEUMANN (1)) he was the first to publish an axiom system for such spaces.* In this work he also studied the relation between subspaces and projection operators (see BERNKOPF (1; IV-5)).

VON NEUMANN begins his paper by considering Hermitian operators; that is, R is an *Hermitian Operator* [abbreviated HO by VON NEUMANN] if it is a linear operator in H** which satisfies $(f, Rg)=(Rf, g)$ for all f and g in H. Then he

* Specifically, a Hilbert space is an abstract infinite dimensional vector space with an inner product, (x, y), defined on it; see DUNFORD & SCHWARTZ (1).

** A is a linear operator defined on H if $A(\lambda a+\gamma b)=\lambda A\,a+\gamma A\,b$ for all complex numbers λ and γ and all vector a and b of H. It is defined *in* H if its range is also a subset of H.

specializes to Hermitian matrices defined on the space of square summable complex sequences. A matrix $A=\{a_{ij}\}$ is said to be *Hermitian square summable if*

(1) $a_{ij}=\bar{a}_{ji}, \quad i, j=1, 2, \ldots$ (*Hermitian property*)

and

(2) $\sum_{j=1}^{\infty}|a_{ij}|^2=M_i<\infty, \quad i=1, 2, \ldots$ (condition of *square summability*, but the set $\{M_i\}$ is not required to be bounded).

Now, let $\{\varphi_j\}$ be a complete orthonormal set*. Then an operator R can be defined by taking $R\varphi_j=\sum_{j=1}^{\infty}a_{ij}\varphi_j$. This converges because $\{a_{ij}\}$ satisfies condition (2) just above. R will be called the *elementary operator associated with the pair* $(A, \{\varphi_j\})$. Furthermore, $(\varphi_j, R\varphi_i)=a_{ij}=\bar{a}_{ji}=(R\varphi_j, \varphi_i)$. Thus R is a Hermitian operator on the orthonormal set $\{\varphi_j\}$. Next, extend R by linearity to all finite linear combinations of the φ_j; that is, define $R\left(\sum_{j=1}^{n}\lambda_j\varphi_j\right)=\sum_{j=1}^{n}\lambda_j R\varphi_j$. Designate the extended operator by $\hat{R}$; then $\hat{R}$ is defined on a dense subset of H.

VON NEUMANN next wished to characterize the closed linear operator** $\tilde{R}$, associated with the pair $(A, \{\varphi_j\})$. If we knew that $\hat{R}$ was bounded, and hence continuous, it could be extended by continuity; that is, $\tilde{R}(\lim_{n\to\infty} f_n)$ is defined to be $\lim_{n\to\infty}\hat{R}(f_n)$, where $\{f_n\}$ is any Cauchy sequence of elements in the domain of $\hat{R}$. However, this is not, in general, the case. In order to extend an unbounded $\hat{R}$, VON NEUMANN introduces the concept of an extension element.

Suppose, for all $g\in D_R$ and some $f\in H$, there exists $f^*\in H$ with $(f, Rg)=(f^*, g)$; this f will be called an *extension element of* R. It is easy to see that the assignment $f\to f^*$ is single valued, but the question now arises as to whether R can be extended to be meaningful for f. (Obviously, if the extension is possible, $Rf=f^*$.) But, by the Hermitian property of R,

$$(f, f^*)=(f, Rf)=(Rf, f)=(f^*, f),$$

or

$$(f, f^*)=\overline{(f, f^*)}.$$

[The property that $(x, y)=\overline{(y, x)}$ has been used.] Hence a necessary condition for R to be extendable to f is that the imaginary part of (f, f^*) is equal to zero. (Such extension elements are said to be in the *zero class*.) It turns out that this is also sufficient.

Now the closed operator associated with R is clearly a (possibly improper) extension of $\hat{R}$. Thus the question of finding this operator is reduced to finding those extension elements which are the zero class. Let f be an extension element,

* A set $\{\varphi_j\}$ is orthonormal if $(\varphi_i, \varphi_j)=\delta_{ij}$. It is complete if it spans H, *or* if $f=\sum_{p=1}^{\infty}(f, \varphi_p)\varphi_p$ for all f in H, *or* if $(f, g)=\sum_{p=1}^{\infty}(f, \varphi_p)\cdot\overline{(g, \varphi_p)}$ for all f and g in H; the three conditions are equivalent.

** The operator, S, is said to be closed if wherever $\{f_n\}\subset D_s$ [D_s is the domain of S], and if $f_n\to f$, and also if $Sf_n\to\tilde{f}$, then $f\in D_s$ and $Sf=\tilde{f}$. According to HILLE & PHILLIPS (1; p. 45) "the closed linear transformations include all of the linear operators which the analyst is likely to use."

and let f^* be as in the preceding paragraph. Then

$$(f^*, \varphi_i) = (f, R\varphi_i) = \left(f, \sum_{j=1}^{\infty} a_{ij}\varphi_j\right)$$

from which

(109) $$(f^*, \varphi_i) = \sum_{j=1}^{\infty} a_{ji}(f, \varphi_j).$$

Now write $f = \sum_{j=1}^{\infty} x_j \varphi_j$. This is possible because $\{\varphi_j\}$ is an orthonormal set, and furthermore,

(110) $$x_j = (f, \varphi_j), \quad j = 1, 2, \ldots.$$

From (109) and (110) we then have that since $\{\varphi_i\}$ is an orthonormal set, f^* will exist if and only if

(111) $$\sum_{j=1}^{\infty} |y_i|^2 < \infty,$$

where we have written

(112) $$y_i = \sum_{j=1}^{\infty} a_{ji} x_j.$$

In this case, from (109), (110), (112) and the fact that $\{\varphi_i\}$ is also complete,

(113) $$f^* = \sum_{i=1}^{\infty} y_i \varphi_i.$$

To determine which of the extensions are of the zero class, consider

(114) $$(f^*, f) = \sum_{j=1}^{\infty} y_j \bar{x}_j = \sum_{j=1}^{\infty} \left(\sum_{i=1}^{\infty} a_{ij} x_i\right) \bar{x}_j = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} (a_{ij} x_i \bar{x}_j)$$

and

(115) $$(f, f^*) = \sum_{i=1}^{\infty} x_i \bar{y}_i = \sum_{i=1}^{\infty} x_i \left(\sum_{j=1}^{\infty} \bar{a}_{ji} \bar{x}_j\right) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (a_{ij} x_i \bar{x}_j).$$

If f is of the zero class, then $(f, f^*) = (f^*, f)$. Hence the zero class consists of those elements for which the order of summation in the last element of (114) (or (115)) can be reversed.

von Neumann next showed that given any closed Hermitian operator R, there exists a matrix A and an orthonormal sequence $\{\varphi_i\}$ so that R is the closed operator associated with the pair $(A, \{\varphi_i\})$. This was demonstrated by constructing the set $\{\varphi_i\}$ in a way which depended on R, and then taking $A = \{a_{ij}\}$ with $a_{ij} = (\varphi_i, R\varphi_j)$. However, A and $\{\varphi_i\}$ are not unique. For example, it is possible to construct a second sequence $\{\psi_i\}$ which gives rise to a different matrix $B = \{b_{ij} = (\psi_i, R\psi_j)\}$. As in the finite dimensional case, the connection between the matrices A and B is as follows: Consider the unitary matrix* $U = \{u_{ij} = (\varphi_i, \psi_j)\}$. Then $A = UBU^{-1}$, in direct analogy to the finite case. von Neumann inserts a word of caution at this point that this is as far as the finite theory of unitary transformations of Hermitian matrices can be generalized to Hilbert spaces.

* A matrix, U, is called *unitary* if U is defined in H, and if $|Uf| = |f|$ for all f in H. U is then continuous, one-one, and satisfies $(Uf, Ug) = (f, g)$. Here $|f| = \sqrt{(f, f)}$.

In another paper, *On the Theory of Unbounded Matrices*, VON NEUMANN (2) continues his investigation of infinite matrices. Here he is interested in developing the relation between bounded and unbounded Hermitian matrices; that is, between matrices which define bounded and unbounded Hermitian operators.* Since this paper is essentially a continuation of his first, the definitions and terminology of the prior work are used, and the results mentioned are assumed.

Thus, we already know that given a closed Hermitian operator R, there exists a complete orthonormal sequence $\{\varphi_j\}$ and a square summable Hermitian matrix $A=\{a_{ij}=(R\varphi_i, \varphi_j)\}$ for which R is the associated closed Hermitian operator. But suppose we are given R and $\{\varphi_j\}$, with $R\varphi_j$ defined for all $j=1, 2, \ldots$; form A, its associated elementary operator S, and (in the notation of VON NEUMANN's previous paper) the extensions $\hat{S}$ and $\tilde{S}$, of S. In particular, $\tilde{S}$ is a closed Hermitian operator, and it is easy to see that R is an extension of $\tilde{S}$, but whether $R=\tilde{S}$ still must be settled, *i.e.*, is R the closed linear operator associated with the pair $(A, \{\varphi_j\})$?

VON NEUMANN answered this with the following

Theorem. *Let R be a closed Hermitian operator, and let $\{\varphi_i\}$ be a given complete orthonormal set. Then there exists either no matrix A (as above) or exactly one such A. Furthermore, if R is bounded, such an A will always exist for every set $\{\varphi_i\}$; if R is unbounded, there exist sets $\{\varphi_i\}$ for which an A exists and sets for which no A exists.*

Now let A be a square summable Hermitian matrix, and let $U=\{u_{ij}\}$ be a unitary matrix. The u_{ij} satisfy

$$\sum_{j=1}^{\infty} u_{ij}\bar{u}_{kj} = \sum_{j=1}^{\infty} u_{ji}\bar{u}_{jk} = \delta_{ik}. \tag{116}$$

Next, VON NEUMANN defined the concept of convergent applicability.

Definition 1. U is *convergently applicable* to A if

1.1) the series $\sum_{j=1}^{\infty} a_{ij}u_{jk}$ and $\sum_{i=1}^{\infty} a_{ij}\bar{u}_{ik}$ converge absolutely and, furthermore, the sums

$$\sum_{i=1}^{\infty}\left|\sum_{j=1}^{\infty} a_{ij}u_{jk}\right|^2 \quad \text{and} \quad \sum_{j=1}^{\infty}\left|\sum_{i=1}^{\infty} a_{ij}\bar{u}_{ik}\right|^2$$

are finite.

1.2) Also, we must have

$$\sum_{k=1}^{\infty}\left(\sum_{p=1}^{\infty} a_{kp}\bar{u}_{ki}u_{pj}\right) = \sum_{p=1}^{\infty}\left(\sum_{k=1}^{\infty} a_{kp}\bar{u}_{ki}u_{pj}\right).$$

[The convergence of the series in (1.2) is assured by (1.1).] The common value of the sums in (1.2) will be denoted by b_{ij}, and it can be seen that the matrix $B=\{b_{ij}\}$ is a square summable Hermitian matrix. Note that if $\psi_j = \sum_{i=1}^{\infty} \overline{u_{ij}}\varphi_i$, then $\{\psi_i\}$ will be a complete orthonormal sequence whenever $\{\varphi_i\}$ is. These concepts lead to

* Recall a linear operator R is *bounded* if $|Rf| \leqq c|f|$ for all $f\in H$ and some constant $c>0$, independent of f; or, equivalently, if $|Rf| \leqq c$ for all f with $|f| \leqq 1$. An operator is *unbounded* if it is not bounded.

Definition 2. Let A be a summable Hermitian matrix, and let $\{\varphi_i\}$ be a complete orthonormal sequence. If there exists a unitary matrix U, convergently applicable to A, then the pair $(A, \{\varphi_i\})$ will be said to be *convergently-unitarily equivalent* [following VON NEUMANN, this will be abbreviated *c.-u. equiv.*] to the pair $(B, \{\psi_i\})$ constructed immediately above.

In what follows, it will be said that a pair $(A, \{\varphi_i\})$ has a certain property if the closed linear operator associated with it has that property. For example $(A, \{\varphi_i\})$ will be called maximal if $\tilde{R}$ is maximal.*

From these ideas VON NEUMANN gets an important characterization of bounded matrices. A summable Hermitian matrix A is bounded if and only if U is convergently applicable to A for every unitary matrix U. A second theorem states that two pairs $(A, \{\varphi_i\})$ and $(B, \{\psi_i\})$ are c.-u. equiv. if and only if they have a common Hermitian extension. Thus the question of c.-u. equiv. can be reduced to the study of common extensions. It can be seen, in particular, that a maximal Hermitian operator is only c.-u. equiv. to a restriction of itself; that a bounded Hermitian operator is c.-u. equiv. only to itself (since a bounded Hermitian operator has no proper closed restrictions); and that *two c.-u. equiv. maximal Hermitian operators must be identical.*

The question now arises as to whether c.-u. equiv. is actually an equivalence relation; *i.e.*, is it symmetric, reflexive and transitive? As is well known for finite matrices, where c.-u. equiv. reduces to unitary equivalence, there is a true equivalence relation. However this is in general not true for the infinite case; it is the transitivity which breaks down. This was more precisely formulated by VON NEUMANN in the following

Theorem. *Let $R \sim S$ denote that R is c.-u. equiv. to S. Then $R \sim R$, and if $R \sim S$, then $S \sim R$. However, if S is fixed, then $R \sim S$ and $S \sim T$ implies $R \sim T$ for all such R and T if and only if S is maximal.*

We outline VON NEUMANN'S proof of this last assertion. By a previous theorem the pairs, R, S and S, T each have common extensions. But if S is maximal, these common extensions are S in each case. Thus S is an extension of both R and T; *i.e.*, $R \sim T$. Conversely, if S is not maximal, VON NEUMANN had shown that there are several distinct maximal extensions of S. Let two of them be R and T. Then by a previous theorem, $R \sim S$ and $S \sim T$, but since R (say) is maximal, the only *maximal* operator c.-u. equiv. to R is R itself. Since $R \neq T$, it necessarily holds that R is not c.-u. equiv. to T.

Thus, putting together the above theorem and the remarks preceding it, we see that the analogy of unitary equivalence for finite matrices carries over to infinite matrices completely only in the case of bounded Hermitian operators, and for these operators the theory is entirely trivial since every bounded Hermitian operator is maximal.

Next, the relation between eigenvalues and c.-u. equiv. was explored by VON NEUMANN. As is well known, in the finite case the eigenvalues are invarient

* Let D_R mean the domain of R, *etc.* Then an operator S is called an *extension* of an operator R if $D_R \subset D_S$ and for all $f \in D_R$, $Rf = Sf$. Also R will be called a *restriction* of S. S is a proper extension of R if $D_R \neq D_S$. S is *maximal* if S has no proper extensions.

under unitary equivalence. To discuss this question for unbounded matrices, the concept of semiboundedness *(halbbeschränkte)* must be introduced from the earlier paper (von Neumann (1)).

Definition 3. The operator R is *semibounded* from above (resp. below) if, for all $f \in D_R$ $(Rf, f) \leqq c|f|^2$ (resp. $(Rf, f) \geqq -c|f|^2$). Note that the pair $(A, \{\varphi_i\})$ is semibounded if the closed linear operator associated with A is semibounded. Furthermore, if $(Rf, f) \geqq 0$, the operator R is said to be *definite,* as is the associated Hermitian pair. Also, a bounded operator is semibounded both from above and from below.

Let $(A, \{\varphi_i\})$ be a definite Hermitian pair; then not only is R a definite operator; it is clear that $\hat{R}$, the linear extension of R, is also definite. Thus, for $f = \sum_{i=1}^{N} x_i \varphi_i$ we have

$$(117) \qquad (Rf, f) = \sum_{j=1}^{N} \left(\sum_{i=1}^{N} a_{ij} x_i \right) \bar{x}_j = \sum_{i,j=1}^{N} a_{ij} x_i \bar{x}_j \geqq 0.$$

In other words, the N^{th} section* of A is positive-semidefinite for every N. In particular, we observe that the eigenvalues of every N^{th} section are all non-negative.

Now let $(A, \{\varphi_i\})$ be a pair which is *not* semibounded from above. Then von Neumann had shown (von Neumann (1); Theorem 46) that the closed operator R associated with $(A, \{\varphi_i\})$ can be extended to a definite Hermitian operator S, which may be assumed to be also closed. Let $(B, \{\psi_i\})$ be the pair associated with S; then B is definite and $(B, \{\psi_i\})$ is c.-u. equiv. to $(A, \{\varphi_i\})$.

In fact, this result can be sharpened and extended to the following

Theorem. *Let A be a Hermitian matrix which is not semibounded from above or below, respectively, and let c be an arbitrary number. Then A is c.-u. equiv. to a matrix B, where B has the property that for every N, its N^{th} section has all its eigenvalues $\geqq c$ or $\leqq c$, respectively.*

Proof. For non-semiboundedness from above and $c = 0$, this is the assertion of the preceding paragraph. For the former alternative and for arbitrary c, replace A by $A - cI$ and B by $B - cI$, where I is the identity matrix; for the latter alternative by $-A + cI$ and $-B + cI$.

This theorem shows that the spectra of the N^{th} sections of Hermitian matrix which defines an unbounded operator have little relationship with the spectrum of the matrix itself. This is in marked distinction to Hilbert's theory of bounded operators. [See Bernkopf (1; pp. 15—16). Because of the continuity of the kernel K, the Hermitian operator appearing there is bounded.] Furthermore, there exists a maximal matrix A which is semibounded neither from above nor from below; thus there is a matrix B whose N^{th} sections have eigenvalues all $\geqq 1$, and a matrix C whose N^{th} sections all have eigenvalues $\leqq -1$, and such that B and C are both c.-u. equiv. to A. But since A is maximal, a previous theorem shows that B is also c.-u. equiv. to C! This pathology is clearly not just limited to a peculiar special case but holds generally for all non-semibounded matrices A.

* The N^{th} section of $A = \{a_{ij} | i, j = 1, 2, \ldots, N\}$.

To explore further the pathology of unbounded Hermitian matrices, VON NEUMANN introduced yet two more concepts.

Definition 4. Two pairs $(A, \{\varphi_i\})$ and $(B, \{\psi_i\})$ are said to be *c.-u. equiv. in n steps* (denoted by $\underset{n}{\sim}$) if there exists $n+1$ pairs $(A_k, \{\psi_i^k\})$ with $(A_0, \{\psi_i^0\}) = (A, \{\varphi_i\})$ and $(A_n, \{\psi_i^n\}) = (B, \{\psi_i\})$ such that for each $k = 1, 2, \ldots, n$, $(A_{k-1}, \{\psi_i^{k-1}\})$ is c.-u. equiv. to $(A_k, \{\psi_i^k\})$. *A is c.-u. equiv. in n steps to B* if for every orthonormal set $\{\varphi_i\}$ there is a corresponding set $\{\psi_i\}$ so that $(A, \{\varphi_i\})$ is c.-u. equiv. to $(B, \{\psi_i\})$ in n steps. As before, this definition can be extended to the Hermitian operators or the closed Hermitian operators associated with $(A, \{\varphi_i\})$ and $(B, \{\psi_i\})$.

Suppose that a partial ordering has been assigned to the set of Hermitian operators by writing $S \leqq R$ (resp. $S < R$) if R is an extension (resp. proper extension) of S. Then if $R \underset{n}{\sim} S$, there exists $2n-1$ closed linear operators which satisfy

$$R \leqq T_1 \geqq S_1 \leqq T_2 \geqq S_2 \leqq \cdots \geqq S_{n-1} \leqq T_n \geqq S \tag{118}$$

where the T_i $(i = 1, 2, \ldots, n)$ can be assumed to be maximal, since by a previous theorem two operators are c.-u. equiv. if and only if they have a common maximal extension.

The second concept occurs in

Definition 5. Two operators R and S are called *adjacent (benachbart)* if they are both extensions of the same Hermitian operator. They are said to be *adjacent in n steps* if there are $n+1$ closed linear operators T_k so that $R = T_0$ and $S = T_n$ and for all $k = 1, 2, \ldots, n$, T_{k-1} is adjacent to T_k.

We see that if R and S are adjacent, then $R \geqq T \leqq S$, or if R and S are adjacent in n steps, then

$$R \geqq S_1 \leqq T_1 \geqq S_2 \leqq \cdots \leqq T_{n-1} \geqq S_n \leqq S \tag{119}$$

where as in (118) it may be assumed that the T_i are all maximal. The similarity between (118) and (119) is clear, and VON NEUMANN describes them as "*hill and valley roads*", the difference being that (118) begins and ends with hills and (119) with valleys. From this it follows that if R is c.-u. equiv. to S in n steps, then R and S are adjacent in $n+1$ steps and conversely. This is shown by replacing T_1 and T_n by R and S respectively in (118) and by replacing S_1 and S_n by R and S in (119).

What VON NEUMANN then showed, as we shall see, is this: *Given two arbitrary unbounded square summable Hermitian matrices A and B, then for every orthonormal sequence $\{\varphi_i\}$ there is an orthonormal sequence $\{\psi_i\}$ so that pair $(A, \{\varphi_i\})$ is c.-u. equiv. to the pair $(B, \{\psi_i\})$ in no more than three steps.* He further showed that if $\{\psi_i\}$ is prescribed, then $(A, \{\varphi_i\})$ is c.-u. equiv. to $(B, \{\psi_i\})$ in no more than nine steps; that is, every pair of unbounded Hermitian operators is c.-u. equiv. in less than ten steps.

To illustrate the significance of this result, suppose we say that the Hermitian matrix A is *weakly unitarily equivalent* to the Hermitian matrix B if A is c.-u. equiv. to B in finitely many steps as just described. Clearly weak unitary equivalence is an equivalence relation [transitivity is trivial], and in case A and B are both finite matrices, weak unitary equivalence reduces to ordinary unitary equivalence. As is well known, in the finite case there are infinitely many equivalence classes. Moreover, for fixed k, the classes of $k \times k$ Hermitian matrices can be characterized looking at eigenvalues; that is, two matrices belong to the same equivalence class if and only if they have exactly the same k eigenvalues (see, for example, Perlis (1; p. 191)). Consider now the situation for unbounded matrices. In effect what von Neumann's result means is that *every unbounded matrix lies in the same weak unitary equivalence class* without regard to eigenvalue considerations.

The situation for bounded infinite Hermitian matrices is almost as bad. We have already noted that a bounded Hermitian matrix is c.-u. equiv. only to itself. Thus, for these matrices each weak unitary equivalence class contains only a single element. To summarize then, we see that weak unitary equivalence produces, on the sets of infinite bounded or unbounded Hermitian matrices, only a trivial equivalence relation, in sharp contrast to the state of affairs in the finite case.

We outline, then, von Neumann's proof that weak unitary equivalence is trivial for unbounded matrices. First, let $A = \{a_i \delta_{ik}: \delta_{ik} = \text{Kronecker delta}\}$ be a diagonal Hermitian matrix. Then the a_i $(i = 1, 2, \ldots)$, are all real. The closed linear operator associated with the pair $(A, \{\varphi_i\})$ is called *a diagonal operator*. The key to von Neumann's assertion is the following

Theorem. *Let R be an unbounded closed Hermitian operator. Then there is a diagonal operator S associated with a matrix $A = \{a_i \delta_{ik}\}$ and an orthonormal set $\{\psi_i\}$ such that R and S are adjacent. Furthermore, a_1 can be arbitrarily prescribed.*

In the proof von Neumann uses his abstract theory of operators [which will be discussed in a subsequent paper] as generated in von Neumann (1). However, the general ideas used are these: Subspaces M_q of H are found on each of which R has an inverse. Denote the restrictions of R to M_q by R_q, and their (operator) closures by $\widetilde{R}_q$. Now let M_p be such that the domain of $\widetilde{R}_p^{-1}$ $(\equiv N_p)$ is not all of H. N_p is closed; denote its orthogonal complement in H by L. Then it was shown that the subspace L is also orthogonal to the domain of $\widetilde{R}_p$. Thus if $f \in H$, f can be uniquely decomposed into $g+h$ where g is in the domain of $\widetilde{R}_p$ and h is in L. Now define $Sf = \widetilde{R}_p(g)$. S is a linear Hermitian operator which is an extension of $\widetilde{R}_p$. Next von Neumann showed that N_p (the range of $\widetilde{R}_p$) reduces S; that is, S maps N_p into N_p. If the restriction of S to N_p is written S', then S' (considered as an operator *in* N_p) has an inverse which is completely continuous. Also $Sh = 0$ for $h \in L$. Thus, by a principal axis theorem proved by Hilbert, there exists a sequence of real numbers $\{\alpha_i\}$ and an orthonormal set $\{\theta_i\}$ with $(S')^{-1}(\theta_i) = \alpha_i \theta_i$ $(i = 1, 2, \ldots)$. Since $(S')^{-1}$ has an inverse, $\alpha_i \neq 0$ $(i = 1, 2, \ldots)$, and also $S\theta_i = \frac{1}{\alpha_i}\theta_i$. Now if $\{\omega_i\}$ is an orthonormal set spanning L, $S\omega_i = 0$. If the set $\{\theta_1, \theta_2, \ldots, \omega_1, \omega_2, \ldots\}$ is denoted by $\psi_1, \psi_2, \ldots$, then $\{\psi_i\}$ is the required set, and S is the required operator. This proves the theorem whenever R is not semibounded from above and a_1 is chosen to be zero, provided the numbering of the ψ_i is selected so that ψ_1 is an ω_i. For other choices of a_1, use $R - a_1 I$ and $S - a_1 I$, and if R is not semibounded from below (but is semibounded from above) replace R and S by $-R + a_1 I$ and $-S + a_1 I$, respectively.

von Neumann next proved that if R and S are maximal Hermitian operators, there is a unitary operator U so that R and USU^{-1} are adjacent in two steps. The proof used the previous theorem together with further ideas from his earlier paper on operator theory. He next showed the

Theorem. *Let A and B be any unbounded Hermitian square summable operators. Then A and B are c.-u. equiv. in three steps. In general the number three cannot be decreased.*

Proof. Let R and S be the closed linear Hermitian operators associated with $(A, \{\varphi_i\})$ and $(B, \{\psi_i\})$ respectively; let R' and S' be the maximal extensions of R and S. Then the theorem of the previous paragraph can be applied; that is, R' and $US'U^{-1}$ are adjacent in two steps. Also $US'U^{-1}$ is a maximal extension of USU^{-1}. This shows, according to the remark about hill and valley roads, that R and USU^{-1} are c.-u. equiv. in three steps, or, since R is the operator associated with $(A, \{\varphi_i\})$ and USU^{-1} with $(B, \{U\psi_i\})$, that A and B are c.-u. equiv. in three steps.

To show that, in general, three is the least possible number, consider the diagonal matrices $A=\{a_k\delta_{jk}\}$ and $B=\{b_k\delta_{jk}\}$. Suppose that $a_k \to \infty$ and $b_k \to -\infty$. Then A is not semibounded from above but is semibounded from below, and B is not semibounded from below but is from above. Suppose $(A, \{\varphi_i\})$ and $(B, \{\psi_i\})$ were c.-u. equiv. in two steps. Let R and S be the maximal operators associated with the above pairs; then R and S are c.-u. equiv. in two steps, and since they are maximal, they are also adjacent. Let T be such that

$$R \geqq T \leqq S. \tag{120}$$

Since T is a restriction of R, it is bounded from below, and since it is a restriction of S, T is bounded from above; that is, T is bounded. Hence it is maximal, hence $T=R=S$. But this contradicts the non-boundedness of R and S. It is easy to see that this state of affairs will hold for any pair of matrices satisfying only the boundedness properties of A and B above.

von Neumann then completed his assertion that no more than nine steps were needed if the orthonormal sequence $\{\psi_i\}$ is prescribed. We only indicate the results. He first showed that operators R and S associated with special types of diagonal matrices are adjacent in six steps. This fact was then used to show the nine-step c.-u. equiv. of two arbitrary unbounded operators. In the balance of his paper, von Neumann explored the further pathology of unbounded abstract operators, with particular attention being paid to their domains of definition. This study is primarily concerned with abstract operator theory, and as such is outside the scope of this paper; it will be taken up in a subsequent history.

8. Postface

The publication of von Neumann's paper essentially meant an end to the use of matrix theory as an effective tool for the study of operators defined on Hilbert spaces. Not that there were any logical difficulties; it is just that matrices were unsuitable because of their clumsiness. We let von Neumann himself explain some of the difficulties (von Neumann (2; pp. 208—209)).

The operator-matrix relation — which, in finite dimensional (Euclidean) spaces and also in bounded operators in Hilbert Space, is simple and even one to one — displays, of course, for unbounded operators in Hilbert Space, new features. It is essentially more complicated, and the theory of matrices is in no way, as in the above mentioned cases, equivalent and corresponding to that of operators. ... [This is because] in finite dimensions as well as for bounded operators in a Hilbert Space ... a matrix can be assigned to a Hermitian operator R and an arbitrary complete orthonormal sequence $\{\varphi_k\}$ [by the formula].

$$a_{jk}=(\varphi_j, R\,\varphi_k).$$

But for an unbounded operator R, which according to TOEPLITZ's theorem is not defined on the whole space, this assignment is only possible when all of [the terms] $R\,\varphi_1, R\,\varphi_2, \ldots$ are defined.

And also, if this is the case, a series of convergence difficulties and limitations appear which are completely absent in the "bounded" theory.

It so happens that, for example, the usual ordinary operator-matrix assignment is not even unitary-invariant.* Because of this, although the operator is entirely reasonable, a peculiar pathology can exist in the matrix. ...

Thus, it becomes evident that from the simple situation which exists in finite dimensional as well as in infinite dimensional but bounded matrices, ... only very little is carried over. For operators [the theory is], indeed, most in order: the eigenvalue problem has [either] no or exactly one solution, and if it has none then other simple normal forms, *etc.*, take the place of HILBERT's "Spectral form." But for matrices such an unexpected pathology prevails that a construction on this [operator-matrix] foundation appears to be very difficult. The most characteristic aspect of this is that the Hermitian matrices behave in a relatively rational fashion; the real sources of the pathology are the unitary matrices (in which case, because of their boundedness, was really not to be expected).

There were other difficulties. Some of the techniques employed to solve problems concerning matrices and matrix theory appeared to be special, even "accidental". Furthermore, certain properties, such as self-adjointness, were often not readily visible in infinite matrices. It could happen that a matrix representation of an unbounded operator was not effective; generally, care had to be exercised in the selection of the proper orthonormal sequence. Thus, "in this one case at least there was little to be gained from the concrete [matrix] representation. The abstract formulation was of genuine help in describing real problems and making them more transparent."**

Not unnaturally there was an initial hesitancy by some mathematicians to accept VON NEUMANN's new abstract formulation of operator theory. Even ERHARD SCHMIDT, who had been a helpful referee for the earlier work of VON NEUMANN (cited in the previous section) which introduced the abstract concept to Hilbert space theory, is reported to have remarked in this connection, "No, no! Don't say operator, say matrix." Yet there was no denying that infinite matrix theory was not the effective setting for operators, and so the abstract approach soon prevailed.***

* That is, two pairs $(A, \{\varphi_i\})$ and $(B, \{\psi_i\})$ are not c.-u. equiv. if they do not have a common extension.

** FRIEDRICHS, K. O., Conversation with the author, April 1967.

*** FRIEDRICHS, K. O., Conversation, previously cited, of February 1967. Professor FRIEDRICHS tells me that he had a paper ready for publication at the time when VON NEUMANN's work appeared, and that he found it necessary to rewrite it using the new abstract concepts.

It should not be assumed, however, that research in infinite matrices has ceased. Today the chief interest is centered about their application in the theory of divergent sequences and series. The theory lies outside of the stream of the work in operator theory, but the interested reader should consult COOKE (1) for details.

Appendix A

We present SCHMIDT's theorems and proofs concerning the solutions of

$$\sum_{\mu=1}^{\infty} a_{np} z_p = 0, \qquad (n = 1, 2, \ldots). \tag{45}$$

In order to get his representations, SCHMIDT first added the mild restriction that no finite set of the a^n are linearly dependent. [The notation is as before, and the equation numbering is a continuation of that in the main text, Chapter 6.]

Let

$$\alpha_{ik} = (a^i, \bar{a}^k); \tag{49}$$

then $\alpha_{ik} = \alpha_{ki}$. Observe the independence condition implies that $\det\{\alpha_{ik};\ i, k = 1, 2, \ldots, m\}$ will be different from zero for any n, and it will, in fact, be real and positive. Next, take $w = \{w_j\}$ to be any element of H, and construct

$$p_j^m = \frac{\det\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1m} & \bar{a}_{1j} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{2m} & \bar{a}_{2j} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_{1m} & \alpha_{2m} & \cdots & \alpha_{mm} & \bar{a}_{mj} \\ (\bar{a}^1, w) & (\bar{a}^2, w) & \cdots & (\bar{a}^m, w) & w_j \end{pmatrix}}{\det\{\alpha_{ik};\ i, k = 1, 2, \ldots, m\}} \tag{50}$$

(see (47) and (49)). Then, if $p^m = \{p_j^m;\ j = 1, 2, \ldots\}$, since the j^{th} and last columns are the same, one has

$$(p^m, a^k) = 0$$

for $k = 1, 2, \ldots, m$. Now let p be the part of w which is perpendicular to the space spanned by all of the a^n. Then p is the strong limit of the sequence $\{p^m;\ m = 1, 2, \ldots\}$. We omit the proof.

In view of the results just obtained for the homogeneous equation (45), SCHMIDT then utilized (48) to obtain an explicit representation for the φ^ν; see the paragraph following (48). From the definition of φ^ν, using the fact that φ^ν is the part of e^ν perpendicular to A one sees that

$$\varphi_j^\nu = \lim_{m\to\infty} \left[\frac{\det\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1m} & \bar{a}_{1j} \\ \vdots & \ddots & \vdots & \vdots \\ \alpha_{1m} & \cdots & \alpha_{mm} & \bar{a}_{mj} \\ a_{1\nu}\ a_{2\nu} & \cdots & a_{m\nu} & e_{\nu j} \end{pmatrix}}{\det\{\alpha_{ik};\ i, k = 1, 2, \ldots, m\}} \right]. \tag{51}$$

Also

$$\|\varphi\|^2 = \lim_{m\to\infty}\left[\frac{\begin{pmatrix}\alpha_{11} & \cdots & \alpha_{1m} & \bar{a}_{1\nu}\\ \vdots & & \vdots & \vdots\\ \alpha_{1m} & \cdots & \alpha_{mm} & \bar{a}_{m\nu}\\ a_{1\nu} & \cdots & a_{m\nu} & 1\end{pmatrix}}{\det\{\alpha_{ik}\}}\right] \tag{52}$$

by the last remark of the preceding paragraph. Thus, in consideration of the previous theorem concerning equation (45) and the fact that the φ^ν generate R, we have the following

Theorem. *A necessary and sufficient condition that equation* (45) *has no regular solution except for the trivial one is that* (51) *vanish for all* ν, *or, equivalently, that* (52) *vanish for all* ν.

In the special case that $\{a^n\}$ is an orthonormal set, condition (52) becomes

$$1 - \sum_{m=1}^{\infty} |a_{m\nu}| = 0, \qquad \nu = 1, 2, \ldots. \tag{53}$$

Next SCHMIDT considered the inhomogeneous system

$$\sum_{p=1}^{\infty} a_{np} z_p = c_n, \qquad n = 1, 2, \ldots, \tag{54}$$

for which he demonstrated three methods of solution. For the first method, define $g^n = \{g_{np}: p = 1, 2, \ldots\}$ by setting

$$g_{n1} = \bar{c}_n \quad \text{and} \quad g_{nj} = \bar{a}_{n,j-1} \tag{55}$$

for $n = (1, 2, \ldots)$ and $j = (2, 3, \ldots)$. Also define $x = \{x_p: p = 1, 2, \ldots\}$ by placing

$$x_1 = 1 \quad \text{and} \quad x_p = z_{p-1} \quad \text{for} \quad p = 2, 3, \ldots. \tag{56}$$

Then (54) can be replaced by

$$(\bar{g}^n, x) = 0. \tag{57}$$

Also, $x_1 = 1$ can be replaced by

$$(\bar{e}^1, x) = 1 \tag{58}$$

where $e^1 = \{1, 0, 0, \ldots\}$. Now let $h = \{h_p\}$ be the part of e^1 which is perpendicular to the space G generated by the g^n. If $\|h\| = 0$, then e^1 is an element of G, and hence (by (57))

$$(e^1, x) = 0. \tag{59}$$

But this contradicts (58), and so, in such a case, there is no solution to (54).

Consequently, we suppose now that $\|h\| \neq 0$. By a previous result, $\|h\|^2 = (\bar{e}^1, h)$. Then a solution of (54) can be gotten by taking $x^1 = \{x_{1j}\}$, where

$$x_{1j} = \frac{h_j}{\|h\|^2} \qquad (j = 1, 2, \ldots), \tag{60}$$

and then using (56) to get a solution, we have $z^1=\{z_{1j}\}$ with

(61) $$z_{1j}=\frac{h_{j+1}}{\|h\|^2}\qquad (j=1,2,\ldots).$$

[To see that (61) is a solution, we observe that the theory of homogeneous equations shows that (60) is a solution of (57).] Thus we have, rather quaintly stated, the following

Theorem. *The vanishing of $\|h\|$ is a necessary and sufficient condition for the non-existence of a regular solution of* (54).

Schmidt pointed out that if (54) has a solution then, as in the finite case, all other solutions can be derived by adding solutions of the homogeneous system associated with (54) to z^1.

To get an explicit formulation for the solution z^1, Schmidt again assumed that no finite set of the a^n are linearly dependent and set

(62) $$\gamma_{ik}=(g^i,\bar{g}^k)=\bar{c}_i c_k+\alpha_{ik}$$

(see (49)). Then, as before,

$$h_{j+1}=\lim_{m\to\infty}\left[\frac{\begin{pmatrix}\gamma_{11} & \cdots & \gamma_{1m} & \bar{a}_{1j}\\ \vdots & \ddots & \vdots & \vdots\\ \gamma_{1m} & \cdots & \gamma_{mm} & \bar{a}_{mj}\\ -c_1 & \ldots & -c_m & 0\end{pmatrix}}{\det\{\gamma_{ik};\ i,k=1,2,\ldots,m\}}\right]$$

and

$$\|h\|^2=\lim_{n\to\infty}\left[\frac{\begin{pmatrix}\gamma_{11} & \cdots & \gamma_{1m} & -\bar{c}_1\\ \vdots & \ddots & \vdots & \vdots\\ \gamma_{1m} & \cdots & \gamma_{mm} & -\bar{c}_m\\ -c_1 & \ldots & -c_m & 1\end{pmatrix}}{\det\{\gamma_{ik}\}}\right].$$

Now, suppose that x^2 is another solution of (57) and (58). Then x^2-x^1 is orthogonal to G and also to e^1. Thus x^2-x^1 is orthogonal to h, and hence to x^1. Thus, the Pythagorean theorem [which holds in a Hilbert space] can be applied, and we have

(63) $$\|x^2\|^2=\|x^2-x^1\|^2+\|x^1\|^2.$$

As a consequence, from (56),

$$\|z^1\|\leqq\|z^2\|.$$

In other words, Schmidt has proved the following

Theorem. *If* $\|h\|\neq 0$, *formula* (61) *gives the solution of* (54) *with the smallest norm.*

Schmidt's second method of dealing with (54) is closely analogous to, and gives a result similar to the case of finite systems. Consider (54) as a system of equations involving inner products; that is, rewrite (54) as

(64) $$(\bar{a}^n,z)=c_n,\qquad n=1,2,\ldots.$$

Let A be the closed subspace generated by the set $\{a^n\}$, and let A_n be the space generated by $\{a^n\}$ after deleting the element a^n. Let $p_n = \{p_{nj}\}$ be the part of a^n which is perpendicular to A_n. Then SCHMIDT's proposition is the

Theorem. *If, for every value of n, $\|p_n\| \neq 0$; that is, if $a^n \notin A_n$ for every n, then* (64) *has a solution whenever*

$$\sum_{n=1}^{\infty} \frac{|c|}{\|p\|} < \infty. \tag{65}$$

In this case a solution is given by $z^1 = \{z_{1j}\}$ where z_{1j} is given by the formula

$$z_{1j} = \sum_{n=1}^{\infty} \frac{c_n}{\|p_n\|^2} p_{nj}, \qquad j = 1, 2, \ldots, \tag{66}$$

which converges strongly. Furthermore z^1, as given by (66) *represents the solution to* (64) *which has the smallest norm.*

Proof. To establish strong convergence of (66), by the triangle inequality we have that

$$\left\| \sum_{n=p}^{p+q} \frac{c_n}{\|p_n\|^2} p_{nj} \right\| \leqq \sum_{n=p}^{p+q} \left\| \frac{c_n}{\|p_n\|^2} p_{nj} \right\| \leqq \sum_{n=p}^{p+q} \frac{|c_n|}{\|p_n\|^2}$$

which goes to zero as p goes to infinity because of (65). Now, from the definition of p_n and the fact that $p_n \in A_m$ if $n \neq m$ gives

$$(\bar{a}^n, p_m) = 0 \tag{67}$$

if $n \neq m$. Also

$$(\bar{a}^n, p_n) = \|p_n\|^2. \tag{68}$$

Thus, (67) and (68) give that

$$(\bar{a}^m, z^1) = \sum_{n=1}^{\infty} \frac{c_n}{|p_n|^2} (\bar{a}^m, p_n) = c_m,$$

that is z^1 is a solution of (64), hence of (54). The last part of the theorem is proved as before.

Next, suppose that there exists m so that $\|p_m\| = 0$; that is, $a^m \in A_m$. Let $L_1, L_2, \ldots,$ be an orthonormal basis of A_m obtained by the Gram-Schmidt process. Then, if $L_k = \{L_{kj}\}$,

$$L_{kj} = \sum_p \gamma_{kp} a_{pj}, \tag{69}$$

where, for each k, the domain of p is some finite set of natural numbers which does not include m, and the γ_{kp} are uniquely determined by the Gram-Schmidt calculation. Also, since $a^m \in A_m$, and since $\{L_k\}$ is a complete orthonormal set, we have that

$$a_{mj} = \sum_{k=1}^{\infty} (a^m, \bar{L}_k) L_{kj}. \tag{70}$$

Hence, using (69) and (70), we have

$$\begin{aligned} (\bar{a}^m, z) &= \sum_{k=1}^{\infty} (\bar{a}^m, L_k)(\bar{L}_k, z) \\ &= \sum_{k=1}^{\infty} (\bar{a}^m, L_k) \sum_p \gamma_{kp} (\bar{a}^p, z). \end{aligned} \tag{71}$$

Thus, if z is a solution of (64), this last equation can be written as

$$c_m = \sum_{k=1}^{\infty} (\bar{a}^m, L_k) \sum_p \gamma_{kp} c_p. \tag{72}$$

Consequently,

Theorem. *A necessary condition that* (64) *have a solution is that the right-hand side of* (72) *converge to* c_m. Furthermore, (71) shows that the m^{th} equation of (64) (or of (54)) is a linear combination of the remaining ones.

For SCHMIDT's third method, suppose once more that no finite subset of $\{a^n\}$ is linearly dependent, and let A be the closed subspace spanned by $\{a^n\}$. Let $\{b^n\}$ be the orthonormal basis of A obtained by the Gram-Schmidt process. Then if $b^n = \{b_{nj}\}$, it can be arranged that

$$b_{nj} = \sum_{k=1}^{n} \beta_{nk} a_{kj}. \tag{73}$$

Thus (64) is transformed into the system

$$(\bar{b}^n, z) = g_n \qquad (n = 1, 2, \ldots), \tag{74}$$

where the g_n are calculated by the formula,

$$g_n = \sum_{k=1}^{n} \bar{p}_{nk} c_k. \tag{75}$$

This leads to SCHMIDT's third theorem: *A necessary and sufficient condition that* (54) *have a solution is that*

$$\sum_{n=1}^{\infty} |g_n|^2 = \sum_{n=1}^{\infty} \left| \sum_{k=1}^{n} \bar{\beta}_{nk} c_k \right|^2 < \infty. \tag{76}$$

Proof. The necessity of (76) follows from BESSEL's inequality* and (74), for if (76) were not satisfied, there could be no regular solutions. Conversely, if (76) is satisfied, then, by a previous but uncited result,

$$z^1 = \sum_{n=1}^{\infty} g_n b^n \tag{77}$$

converges strongly. Also, z^1 is a solution of (74) and thus of (54), since

$$(\bar{b}^m, z) = \sum_{n=1}^{\infty} g_n (\bar{b}^m, b^n) = g_n (\bar{b}^n, b^n) = g_n,$$

where we have used the fact that $\{b^m\}$ is an orthonormal sequence.

As an application of his theories, SCHMIDT proves a result previously obtained by HILBERT. Suppose the same condition of independence is placed upon the set $\{a^n\}$ as in the previous theorem, and that α_{ik} is defined by (49). Then, since $\alpha_{ik} = \bar{\alpha}_{ki}$, the zeros of

$$\det\{\alpha_{ik} - \lambda \delta_{ik};\ i, k = 1, 2, \ldots, n\} \tag{78}$$

* BESSEL's inequality states that if b^n is an orthonormal sequence, then for any $z \in H$,

$$\left| \sum_{n=1}^{\infty} \right| (\bar{b}^n, z)|^2 \leqq \|z\|^2.$$

are all real and positive. Also the associated bilinear Hermitian form

$$H_n = \sum_{i=1}^{n} \sum_{k=1}^{n} \alpha_{ik} y_i \bar{y}_k \tag{79}$$

is positive definite. In fact, if the zeros of (78) are denoted by

$$\lambda_{n1}, \lambda_{n2}, \ldots, \lambda_{nn} \tag{80}$$

with $\lambda_{n1} \leqq \lambda_{n2} \leqq \cdots \leqq \lambda_{nn}$, then λ_{n1} is the minimum value and λ_{nn} the maximum value assumed by (79) for $\sum_{i=1}^{n} |y_i|^2 = 1$. Thus, since H_n can be considered as a "subform" obtained from H_{n+1} by setting $y_{n+1} \equiv 0$, we see that λ_{n1} (considered as a function of n) is non-increasing. Furthermore, as $\lambda_{n1} > 0$, we have that $\lim_{n\to\infty} \lambda_{n1}$ will exist and will be non-negative. Set

$$\lim_{n\to\infty} \lambda_1 = l. \tag{81}$$

This leads to the following

Theorem. *If* $l > 0$, *and if* $\sum_{n=1}^{\infty} |c_n|^2 < \infty$, *then system* (54) *always has a solution.*

(*Note.* If the existence of a solution is established, explicit representations can be given by (61) or (77).)

Proof. Let the numbers $\{g_n\}$ be defined as in (74) and (75), and let $\{b^n\}$ be the orthonormal basis of A given by (73). Then, since $\{a^k\}$ and $\{b^k\}$ generate the same space for $k = 1, 2, \ldots, n$, let the numbers h_k $(k = 1, 2, \ldots, n)$ be determined by setting

$$\sum_{k=1}^{n} g_k b^k = \sum_{k=1}^{n} h_k a^k. \tag{82}$$

Then, by the generalized Pythagorean theorem* and the fact that $\{b^k\}$ is an orthonormal set, we have

$$\sum_{k=1}^{n} |g_k|^2 = \left\| \sum_{k=1}^{n} g_k b^k \right\|^2 = \left\| \sum_{k=1}^{n} h_k a^k \right\|^2 = \sum_{i=1}^{n} \sum_{k=1}^{n} \alpha_{ik} h_i \bar{h}_k = U. \tag{83}$$

Again, since $\{b^k\}$ is an orthogonal set, we have

$$\left(\bar{b}^\varrho, \sum_{k=1}^{n} g_k b^k \right) = g_\varrho, \quad \varrho = 1, 2, \ldots, n, \tag{84}$$

and from (73) and (74), and the fact that (73) is invertable,

$$\left(\bar{a}^\varrho, \sum_{k=1}^{n} g_k b^k \right) = c_\varrho; \tag{85}$$

we also have that

$$\sum_{k=1}^{n} \alpha_{k\varrho} h_k = \sum_{k=1}^{n} (a^k, \bar{a}^\varrho) h_k = \left(\bar{a}^\varrho, \sum_{k=1}^{n} h_k a^k \right) = c_\varrho, \tag{86}$$

and so

$$\sum_{\varrho=1}^{n} |c_\varrho|^2 = \sum_{\varrho=1}^{n} c_\varrho \bar{c}_\varrho = \sum_{i=1}^{n} \sum_{k=1}^{n} \xi_{ik} h_i \bar{h}_k = V \tag{87}$$

* The generalized Pythagorean theorem states that if $z^1, z^2, \ldots, z^n$ are n mutually orthogonal elements, then

$$\left\| \sum_{k=1}^{n} z^k \right\|^2 = \sum_{k=1}^{n} \|z^k\|^2.$$

where

(88) $$\xi_{ik} = \sum_{\varrho=1}^{n} \alpha_{i\varrho} \bar{\alpha}_{k\varrho} = \sum_{\varrho=1}^{n} \alpha_{i\varrho} \alpha_{\varrho k}.$$

Thus if the matrix $F = \{\alpha_{ik};\ i, k = 1, 2, \ldots, n\}$, then F^2 is the matrix

(89) $$F^2 = \{\xi_{ik};\ i, k = 1, 2, \ldots, n\}.$$

Also, by the principal axis theorem there is a matrix D, with $D'\overline{D} = I$ [D' is the transpose of D] so that $D'F\overline{D} = L$ where L is the diagonal matrix $\{\lambda_{n1}, \lambda_{n2}, \ldots, \lambda_{nn}\}$ (see (80)). Furthermore $D'F^2\overline{D} = L^2 = \{\lambda_{n1}^2, \lambda_{n2}^2, \ldots, \lambda_{nn}^2\}$. Thus, if $t = Dh$, where $h = (h_1, h_2, \ldots, h_n)$ and $t = (t_1, t_2, \ldots, t_n)$, then (see (83))

(90) $$U = \sum_{k=1}^{n} \lambda_{nk} t_k \bar{t}_k$$

and (see (87) and (88))

(91) $$V = \sum_{k=1}^{n} \lambda_{nk}^2 t_k \bar{t}_k.$$

Then (80) and (90) and (91) imply that

(92) $$\lambda_{nn} U \geqq V \geqq \lambda_{n1} U,$$

or, in view of (83) and (87),

$$\lambda_{nn} \sum_{\varrho=1}^{n} |g_\varrho|^2 \geqq \sum_{\varrho=1}^{n} |c_\varrho|^2 \geqq \lambda_{n1} \sum_{\varrho=1}^{n} |g_\varrho|^2.$$

In particular, since $\lambda_{n1} \geqq l > 0$,

$$\sum_{\varrho=1}^{n} |g_\varrho|^2 \leqq \frac{1}{l} \sum_{\varrho=1}^{n} |c_\varrho|^2,$$

and since this is true for every n, the convergence of $\sum_{\varrho=1}^{\infty} |c_\varrho|^2$ implies the convergence $\sum_{j=1}^{\infty} |g_\varrho|^2$ which by a previous theorem (see (76)) insures that (54) has a solution.

Finally, SCHMIDT obtained another of HILBERT's results. He considered the resolvent of equation (54); that is, an array $\{b_{kj}\}$ which has the property that $\sum_{j=1}^{\infty} |b_{kj}|^2 < \infty$, and satisfies

$$\sum_{j=1}^{\infty} a_{ij} b_{kj} = \delta_{ik}, \quad i, k = 1, 2, \ldots.$$

Then, by the theorem associated with (65) and (66), *a necessary and sufficient condition for the existence of a resolvent is that* $\|p_n\| \neq 0$ *for all* n. *In this case* b_{kj} *is given by*

$$b_{kj} = \frac{p_{kj}}{\|p_k\|^2}.$$

Appendix B

We present HELLINGER & TOEPLITZ's proof of the uniform finiteness of convergent bilinear forms (HELLINGER & TOEPLITZ (2; pp. 321 ff.)). The notation and numbering are that of the theorem. It should be noted that the proof is accomplished without the use of high-powered abstract techniques usually associated with the modern proofs of similar theorems.

Suppose the theorem is false. Then (108) is false; that is, there exists a pair of sequences $\{x^k=\{x_p^k\}\}$ and $\{y^k=\{y_p^k\}\}$ such that x^k and y^k all satisfy (107) (lie in the unit ball) for $k=1, 2, \ldots$, and such that

$$\lim_{k\to\infty} \sum_{p=1}^{\infty}\left(\sum_{q=1}^{\infty} a_{pq}\, x_p^k y_p^k\right)=\infty. \tag{A 1}$$

We may assume without loss of generality that the divergence is monotonic in k let $\{c_k\}$ be any sequence of positive numbers such that

$$\sum_{k=1}^{\infty} c_k^2=1. \tag{A 2}$$

Now by (A 1), we can find a sequence x^{k_1} from $\{x^k\}$ and a sequence y^{k_1} from $\{y^k\}$, and a number n_1 such that

$$A_{n_1}(x^{k_1}, y^{k_1})=\sum_{p=1}^{n_1}\left(\sum_{q=1}^{\infty} a_{pq}\, x_p^{k_1} y_q^{k_1}\right)>\frac{2}{c_1^2}. \tag{A 3}$$

Set

$$\begin{aligned} &\xi=\{\xi_p=c_1 x_p^{k_1}\} \quad \text{for} \quad p=1, 2, \ldots, n, \qquad \text{and} \\ &\eta^{(1)}=\{\eta_p^{(1)}=c_1 y_p^{k_1}\} \quad \text{for} \quad p=1, 2, \ldots. \end{aligned} \tag{A 4}$$

Then, by (A 2) and (107)

$$\sum_{p=1}^{n_1}\xi_p^2=\sum_{p=1}^{n_1} c_1^2 (x_p^{k_1})^2 \leqq c_1^2 \sum_{p=1}^{\infty}(x_p^{k_1})^2 \leqq 1 \quad \text{and also} \quad \sum_{p=1}^{\infty}(\eta_p^{(1)})^2 \leqq 1. \tag{A 5}$$

Thus, from (A 3)

$$A_{n_1}(\xi, \eta_1)=\sum_{p=1}^{n_1}\left(\sum_{q=1}^{\infty} a_{pq}\,\xi_p \eta_q^1\right)=c_1^2 A_{n_1}(x^{k_1}, y^{k_1})>2. \tag{A 6}$$

The hypothesis of the theorem shows that $\sum_{q=1}^{\infty} a_{pq} y_q$ converges for each p and for all $y=\{y_q\}$ in the unit ball; hence, by a previous theorem, the sequence $\{a_{pq}\}$ defines a bounded linear form for each p. Thus, $\sum_{q=1}^{\infty} a_{pq}^2$ converges. Let m_1 be any number greater than n_1, to be determined later, and split $A_{n_1}(x, y)$ by placing

$$A_{n_1}(x, y)=\sum_{p=1}^{n_1}\sum_{q=1}^{m_1} a_{pq}\, x_p y_q+\sum_{p=1}^{n_1}\left(\sum_{q=m_1+1}^{\infty} a_{pq}\, x_p y_q\right). \tag{A 7}$$

Consider the second term on the right hand side of (A 7) (the remainder). By the Schwarz inequality (94) applied twice, we have

$$\left(\sum_{p=1}^{n_1}\left(\sum_{q=m_1+1}^{\infty} a_{pq}\, x_p y_q\right)\right)^2 \leqq \left(\sum_{p=1}^{n_1}\sum_{q=m_1+1}^{\infty} a_{pq}^2\right)\left(\sum_{p=1}^{n_1} x_p^2\right)\left(\sum_{q=m_1+1}^{\infty} y_q^2\right). \tag{A 8}$$

Since $\sum_{q=1}^{\infty} a_{pq}^2$ converges independently of n_1, the first factor on the right hand side of (A 8) becomes arbitrarily small for sufficiently large m_1, and in fact m_1 can be chosen so that (A 8) will be small for every x and y which are in the unit ball. In particular, if m_1 is chosen so that

$$\sum_{p=1}^{n_1}\sum_{q=m_1+1}^{\infty} a_{pq}^2<1, \tag{A 9}$$

we will have, for $\eta=\{\eta_p=\eta_p^{(1)};\ p=1,2,\ldots,m_1\}$ (see (A4)), using (A6), (A7) and (A9), that

(A10) $$A_{n_1}(\xi,\eta)=\sum_{p=1}^{n_1}\left(\sum_{q=1}^{\infty}a_{pq}\xi_p\eta_q\right)>1$$

as long as the undetermined η_p $(p=m_1+1, m_1+2, \ldots)$ are selected so that $\sum_{p=1}^{\infty}\eta_p^2\leqq 1$.

Before proceeding to the next step, observe that

$$A_{n_1}(x,y)=\sum_{p=1}^{n_1}\left(\sum_{q=1}^{\infty}a_{pq}x_py_q\right)\quad\text{and}\quad\sum_{p=n_1+1}^{\infty}\left(\sum_{q=1}^{m_1}a_{pq}x_py_q\right)$$

are bounded forms for x and y in the unit ball since they are sums of finitely many bounded linear forms. That is, they are finite sums of the type $x_p\sum_{q=1}^{\infty}a_{pq}y_q$ and $y_q\sum_{p=n_1+1}^{\infty}a_{pq}x_p$. Let M_1 and μ_1, respectively, be their least upper bounds.

Now select $n_2>n_1$ and x^{k_2} and y^{k_2} so that

(A11) $$A_{n_2}(x^{k_2},y^{k_2})=\sum_{p=1}^{n_2}\left(\sum_{q=1}^{\infty}a_{pq}x_p^{k_2}x_p^{k_2}\right)>\frac{2+M_1+2\mu_1}{c_2^2}$$

where c_2 is chosen from (A2). To extend ξ, set

(A12) $$\xi_p=c_2x_p^{k_2}\quad\text{for}\quad p=n_1+1,n_1+2,\ldots;n_2$$

and set $\eta^{(2)}=\{\eta_p^{(2)}\}$ where

(A13) $$\eta_p^{(2)}=\eta_p^{(1)}\quad\text{for}\quad p=1,2,\ldots,m_1;\qquad\eta_p^{(2)}=c_2y_p\quad\text{for}\quad p=m_1+1,m_1+2,\ldots.$$

Then, as in (A5)

(A14) $$\sum_{p=1}^{n_2}\xi_p^2=c_1^2+c_2^2\sum_{p=n_1+1}^{n_2}(x_p^{k_2})^2\leqq c_1^2+c_2^2\leqq 1,\quad\text{and}\quad\sum_{p=1}^{\infty}(\eta_p^{(2)})^2\leqq 1.$$

Next, write

(A15) $$A_{n_2}(\xi,\eta^{(2)})=A_{n_1}(\xi,\eta^{(2)})+\sum_{p=n_1+1}^{n_2}\left(\sum_{q=1}^{\infty}a_{pq}\xi_p\eta_q^{(2)}\right)$$

and consider the remainder. By adding and subtracting $c_2^2\sum_{p=n_1+1}^{n_2}\left(\sum_{q=1}^{\infty}a_{pq}x_p^{k_2}y_q^{k_2}\right)$ one obtains

(A16) $$\sum_{p=n_1+1}^{n_2}\left(\sum_{q=1}^{\infty}a_{pq}\xi_p\eta_q^{(2)}\right)$$

(A17) $$=c_2^2\sum_{p=n_1+1}^{n_2}\left(\sum_{q=1}^{\infty}a_{pq}x_p^{k_2}y_p^{k_2}\right)+\sum_{p=n_1+1}^{n_2}\sum_{q=1}^{m_1}a_{pq}\xi_p(\eta_q^{(2)}-c_2y_q^{k_2})$$

(A18) $$=c_2^2A_{n_2}(x^{k_2},y^{k_2})-A_{n_1}(c_2x^{k_2},c_2y^{k_2})-\sum_{p=n_1+1}^{n_2}\sum_{q=1}^{m_1}a_{pq}\xi_p(\eta_q^{(2)}-c_2y_q^{k_2}).$$

In deriving (A17) from (A18) we have used (A4), (A12), and (A13) as well as the first parts of (A3) and (A11).

We now employ (A 18), (A 11), (A 12), (A 13), (A 10) and the definitions of M_1 and μ_1 to get that

$$A_{n_2}(\xi, \eta^{(2)}) > 1 + (2 + M_1 + 2\mu_1) - M_1 - 2\mu_1 = 3.$$

In exact analogy with the selection of n_1, m_2 is now to be determined so that η can be extended by setting

$$\eta_p = \eta_p^{(2)}, \quad p = 1, 2, \ldots, m_2$$

while $A_{n_2}(\xi, \eta) > 2$.

Inductively then, it can be seen that for every $t = 1, 2, \ldots$, one can extend ξ and η so that at each step ξ and η are within the unit ball, and furthermore

$$A_{n_t}(\xi, \eta) = \sum_{p=1}^{n_t}\left(\sum_{q=1}^{\infty} a_{pq}\,\xi_p\,\eta_q\right) > t.$$

Thus $A(\xi, \eta)$ does not satisfy (108), even though ξ and η satisfy (107), which contradicts the hypothesis.

Bibliography

APPELL, PAUL (1855—1930)

(1) Sur une méthode élémentaire pour obtenir les developpements en série trigonométrique des fonctions elliptiques. Bul. S.M.F. **13**, 13—18 (1885).

BANACH, STEFAN (1892—1945)

(1) Sur les opérations dans les ensembles abstraites et leur application aux équations intégrales. Fund. Math. **3**, 133—181 (1922).

(2) Sur les fonctionnelles linéaires I. Studia Math. **1**, 211—216 (1929); II. *Ibid.* 223—239 (1929).

BERNKOPF, MICHAEL

(1) The development of function spaces with particular reference to their origins in integral equation theory. Archive for History of Exact Sciences **3**, 1—96 (1966).

COOKE, RICHARD

(1) Infinite Matrices and Sequence Spaces. London: Macmillan, 1950. Also reprinted New York: Dover 1960.

DUNFORD, NELSON, & SCHWARTZ, JACOB T.

(1) Linear Operators: Part I. New York: Interscience 1958.

FRÉCHET, MAURICE (1878—)

(1) Sur les opérations linéaires I. Trans. Amer. Math. Soc. **5**, 493—499; II. *Ibid.* **6**, 134—140 (1905).

(2) Sur quelques-points du calcul fonctionnel. Rendiconti del Circ. mat. di Palermo **22**, 1—74 (1906).

FREDHOLM, IVAR (1866—1927)

(1) Sur une classe d'équations fonctionnelles. Acta Math. **27**, 365—390 (1903).

FOURIER, JOSEPH (1768—1830)

(1) Théorie analytique de la chaleur. Oeuvres, 1, edited with corrections and elimination of typographical errors by GASTON DARBOUX. Paris: Gauthier-Villars 1888.

(2) The Analytic Theory of Heat. Translated with notes from the French by ALEXANDER FREEMAN. New York: Dover 1955. Reprinted with corrections from the 1878 edition.

GRAVES, LAURENCE M.

(1) Implicit functions and differential equations in general analysis. Trans. Amer. Math. Soc. **29**, 514—552 (1927).

HAHN, HANS (1879—1934)

(1) Über Folgen Linearer Operationen. Monatshefte f. Math. u. Physik **32**, 3—88 (1922).

(2) Über lineare Gleichungssysteme in linearen Räumen. J. f. reine u. angewandte Math. **157**, 214—229 (1927).

HELLINGER, ERNST, & OTTO TOEPLITZ

(1) Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten. Encyk. der Math. Wissen. IIc, Leipzig: Teubner, 1335—1602 (1923—1927). Reprinted in book form. Leipzig and Berlin: Teubner 1928.

(2) Grundlagen für eine Theorie der unendlichen Matrizen. Math. Ann. **69**, 289—330 (1910).

(3) Grundlagen für eine Theorie der unendlichen Matrizen. Nachr. Akad. Wiss. Göttingen 1906, 351—355.

HELLY, EDUARD

(1) Über Systeme linearer Gleichungen mit unendlich vielen Unbekannten. Monatshefte f. Math. u. Physik **31**, 60—91 (1921).

HILBERT, DAVID (1862—1943)

(1) Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen. Leipzig and Berlin: Teubner 1912. This is a collection in book form of six papers reprinted with minor changes first published as follows: I. Nachr. Akad. Wiss. Göttingen 1904, 49—91; II. *Ibid.* 1904, 213—259; III. *Ibid.* 1904, 307—338; IV. *Ibid.* 1906, 157—227; V. *Ibid.* 1906, 439—480; VI. *Ibid.* 1910, 355—417. Book reprinted New York: Chelsea 1952.

HILL, GEORGE WILLIAM (1838—1914)

(1) On the part of the motion of the linear perigee which is a function of the motions of the sun and the moon. Acta Math. **8**, 1—36 (1886). Reprinted, with "some additions" from a privately printed monograph of the same title, Cambridge, Mass.: Wilson 1887.

HILLE, EINAR, & RALPH S. PHILLIPS

(1) Functional Analysis and Semi-Groups. Providence, R.I.: Amer. Math. Soc. (1957).

KOCH, HELGE VON (1870—1924)

(1) Sur une application des déterminants infinis à la théorie des équations différentielles linéaires. Acta Math. **15**, 53—63 (1891).

(2) Sur les déterminants infinis et les équations différentielles linéaires. Acta Math. **16**, 217—295 (1892—93).

KÖTTERITZSCH, TH.

(1) Über die Auflösung eines Systems von unendlich vielen linearen Gleichungen. Zeitschrift f. Math. u. Phys. **15**, 1—15 and 229—268 (1870).

LÉVY, PAUL

(1) Leçons d'analyse fonctionnelle. Paris: Gauthier-Villars 1922.

MOORE, ELIAKIM HASTINGS (1862—1924)

(1) On a form of general analysis with application to linear differential and integral equations. Atti del IV [1908] Congresso Internazionale, vol. **2**, Rome, 1909.

NEUMANN, JOHN VON (1903—1957)

(1) Allgemeine Eigenwerttheorie Hermitescher Funktionaloperationen. Math. Ann. **102**, part 1, 49—131 (1929).

(2) Zur Theorie der unbeschränkten Matrizen. J. f. reine u. angewandte Math. **161**, 208—236 (1929).

PERLIS, SAM

(1) Theory of Matrices. Cambridge, Mass.: Addison-Wesley 1952.

PINCHERLE, SALVATORE (1853—1936)
(1) Funktionale Operationen und Gleichungen. Encyk. der Math. Wissen. IIa, 11, Leipzig: Teubner, 761—824 (1903—1921).
(2) Opere Scelte, II, Rome: Edizioni Cremonese, 1—70, 1954. Reprinted from Math. Ann. **49**, 325—382 (1897).

PIOLA, G.
(1) Sulla teoria delle fonzioni discontinue. Memorie Soc. Italian. Scienze 20, 573—639. Cited by RIESZ (1; p. 9 fn.).

POINCARÉ, HENRI (1854—1912)
(1) Remarques sur l'emploi de la méthode précédente. Bul. S.M.F. **13**, 19—27 (1884—85).
(2) Sur les detérminants d'ordre infini. *Ibid.* **14**, 77—90 (1885—86).
(3) Introduction to "The Collected Mathematical Works of George William Hill," Vol. I, pp. vii—xvii. Washington: Carnegie Institution (1905). Reprinted by Johnson Reprint Corporation, New York, 1965.

RIESZ, FRÉDÉRIC (1880—1956)
(1) Les systemes d'équations linéaires à une infinité d'inconnues. Paris: Gauthier-Villars 1913.
(2) Untersuchungen über Systeme integrierbarer Funktionen. Math. Ann. **69**, 449—497 (1910).
(3) Über lineare Funktionalgleichungen. Acta Math. **41**, 71—98 (1918).

SANGER, RALPH G.
(1) Functions of lines and the calculus of variations. Contributions to the Calculus of Variations, pp. 193—293. Chicago: Chicago University Press 1933.

SCHMIDT, ERHARD (1876—1959)
(1) Über die Auflösung linearer Gleichungen mit unendlich vielen Unbekannten. Rendiconti del Circ. mat. di Palermo **25**, 53—77 (1908).
(2) Zur Theorie der linearen und nichtlinearen Integralgleichungen. I. Math. Ann. **63**, 433—476 (1907); II. *Ibid.* **64**, 161—174 (1907).

SCHUR, J.
(1) Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen. J. f. reine u. angewandte Math. **140**, 1—28 (1911).
(2) Über lineare Transformationen in der Theorie der unendlichen Reihen. *Ibid.* **151**, 79—111 (1921).

TOEPLITZ, OTTO (1881—1940)
(1) Die Jacobische Transformation der quadratischen Formen von unendlich vielen Veränderlichen. Nachr. Akad. Wiss. Göttingen 1907, 101—109.
(2) Zur Transformation der Scharen bilinearer Formen von unendlichen Veränderlichen. *Ibid.* 110—115.

WINTNER, AUREL
(1) Spektraltheorie der unendlichen Matrizen. Leipzig: Herzel 1929.

Pace College
New York City

(Received June 7, 1967)

Reprinted by Proff GmbH & Co. KG Starnberg