

ON FINITE-ROWED SYSTEMS OF LINEAR INEQUALITIES IN INFINITELY MANY VARIABLES*

BY

I. J. SCHOENBERG

1. Introduction. The theory of systems of linear inequalities was originated by Minkowski† and Farkas‡ and extended in several directions by L. L. Dines, A. Haar§ and others. While in this theory only systems involving a finite number of variables were investigated, more recent papers of F. Hausdorff contain explicit solutions for interesting particular systems|| of linear inequalities involving infinitely many variables. In the present paper a larger class of such systems, which include the systems of Hausdorff as a special case, are solved.

This paper is divided into three parts.¶ In the first part a brief sketch of a theory of convex bodies in the space of infinitely many dimensions is given.** Such bodies occurred implicitly in the earlier work of F. Riesz†† and Carathéodory.§§ At the same time finite-rowed systems of linear inequalities in infinitely many variables of the type (4.1) are considered and their relation to convex bodies ((4.2) and Theorem 4.1) as well as to a problem of F. Riesz (§5) is investigated. In the second part a particular class of such finite-rowed systems are solved by means of Stieltjes integrals (Theorem 8.1). Essentially the same method has already been used by the author to prove Hausdorff's

* Presented to the Society, December 30, 1931; received by the editors March 14, 1932.

† H. Minkowski, *Geometrie der Zahlen*, Leipzig, 1896, §19.

‡ J. Farkas, *Theorie der einfachen Ungleichungen*, Journal für Mathematik, vol. 124 (1902), pp. 1-27.

§ A. Haar, *Über lineare Ungleichungen*, Szeged Acta, vol. 2, No. 1 (1924), pp. 1-14. For a complete bibliography concerning the work of Dines, Carver, Fujiwara and Stokes see the recent paper of R. W. Stokes, *A geometric theory of solution of linear inequalities*, these Transactions, vol. 33(1931), pp. 782-805. A geometric theory more complete than those given by Haar and Stokes is given by the author in a paper which will appear in the American Mathematical Monthly.

|| See §9 and the references given below.

¶ The theorems stated in this paper which are not new will bear in parentheses the names of their authors.

** A theory of integration in the same space was given by P. J. Daniell, *Integrals in an infinite number of dimensions*, Annals of Mathematics, (2), vol. 20 (1918-1919), pp. 281-288.

†† F. Riesz, *Sur certains systèmes singuliers d'équations intégrales*, Annales de l'Ecole Normale Supérieure, (3), vol. 28 (1911), pp. 33-62.

§§ C. Carathéodory, *Über den Variabilitätsbereich der Fourier'schen Konstanten von positiven harmonischen Funktionen*, Rendiconti di Palermo, vol. 32 (1911), pp. 193-217.

theorem on completely monotonic sequences.* The third and last part of this paper contains some applications of the general results of Part II. Hausdorff's theorems concerning his sequences of moments† appear as a special case (§9). This way of stating Hausdorff's results (Theorem 9.1) raises some new questions which are readily answered in §10. In 1929 S. Bernstein‡ proved a remarkable theorem (Theorem 11.1) on completely monotonic functions. This theorem was recently rediscovered by D. V. Widder.§ In the last section (§11) it is shown that the theorem of S. Bernstein and Widder is essentially equivalent to Hausdorff's prior results concerning his *Moment-funktionen* (loc. cit., §4); only a slight modification of Hausdorff's Theorem 9.1 (Corollary 9.1) is needed to prove it.

PART I. CONVEX BODIES AND FINITE-ROWED SYSTEMS OF LINEAR INEQUALITIES IN INFINITELY MANY VARIABLES

2. **Sets of points in S_∞ .** Let us denote by S_p the p -dimensional space of the real variables (x_1, x_2, \dots, x_p) . For $p < q$, S_p shall be called the *p th partial space of S_q* and this space may be defined within S_q by the system $x_{p+1} = x_{p+2} = \dots = x_q = 0$. Similarly, the point $X_p = (x_1, x_2, \dots, x_p)$ shall be called the *p th partial point of $X_q = (x_1, x_2, \dots, x_q)$ of S_q* , and X_p is precisely the orthogonal projection of the point X_q on the sub-space S_p . Let us consider the space $S = S_\infty$ of the infinite sequence of variables $X = X_\infty = (x_1, x_2, x_3, \dots)$. Just as above, the spaces S_1, S_2, S_3, \dots shall be called the *partial spaces* of $S = S_\infty$, and let us refer to $X_p = (x_1, x_2, \dots, x_p)$ as the *p th partial point of X* .

The same definition may be extended to any set $E = \{X\}$ of points X of S . The set $E_p = \{X_p\}$ of the p th partial points of the X shall be called the *p th partial set of E* . The set E shall be called *bounded* in case that every partial set $E_p (p = 1, 2, \dots)$ is bounded in the corresponding partial space S_p .

Let $X^{(n)} = (x_1^{(n)}, x_2^{(n)}, x_3^{(n)}, \dots)$ ($n = 1, 2, 3, \dots$) be a sequence of points of S . We shall say that $X^{(n)} \rightarrow X = (x_1, x_2, x_3, \dots)$ for $n \rightarrow \infty$, in case $x_p^{(n)} \rightarrow x_p$ for $n \rightarrow \infty$ holds for every value of $p = 1, 2, 3, \dots$. A point X of the space S is said to be an *accumulation point* of the set E , if X is the limit of a sequence of points $X^{(n)}$ of E in the sense given above. A point set E shall be called

* I. J. Schoenberg, *On finite and infinite completely monotonic sequences*, Bulletin of the American Mathematical Society, vol. 38 (1932), pp. 72-76.

† F. Hausdorff, *Summationsmethoden und Momentfolgen*, II, Mathematische Zeitschrift, vol. 9 (1921), pp. 280-299.

‡ S. Bernstein, *Sur les fonctions absolument monotones*, Acta Mathematica, vol. 52 (1929), pp. 1-66. See his Theorem E on page 20 and formula (72) on page 56.

§ D. V. Widder, *Necessary and sufficient conditions for the representation of a function as a Laplace integral*, these Transactions, vol. 33 (1931), pp. 851-892. See also J. D. Tamarkin, *On a theorem of S. Bernstein-Widder*, same volume, pp. 893-896.

closed, in case E contains all its points of accumulation. This definition does not imply that any of the partial sets E_1, E_2, E_3, \dots has to be bounded. For example the whole space S is closed. Nor can we conclude that the partial sets E_p are closed if E is closed. To show this let E be the set of points $X = (x_1, x_2, x_3, \dots)$ defined by the inequalities $x_1 > 0, x_2 > 0, x_1 x_2 \geq 1$. While E is closed, the first partial set E_1 which is defined by $x_1 > 0$ is not closed. The converse statement which says that E is closed, if E_1, E_2, E_3, \dots are closed, is also not true. Let us take for instance for the set E the sequence of points

$$(1, 0, 0, 0, \dots), (1, 1, 0, 0, 0, \dots), (1, 1, 1, 0, 0, \dots), \dots$$

The partial set $E_p (p = 1, 2, 3, \dots)$ contains precisely p different points and is therefore closed. E however is not closed, for E does not contain its accumulation point $(1, 1, 1, 1, \dots)$.

LEMMA 2.1. *If the set E is bounded and closed, then also all the partial sets E_p are bounded and closed.*

The partial set E_p is bounded by definition. To show that E_p is also closed let $X_p^{(n)} = (x_1^{(n)}, x_2^{(n)}, \dots, x_p^{(n)})$ be a sequence of points of E_p with $X_p^{(n)} \rightarrow X_p = (x_1, x_2, \dots, x_p)$ for $n \rightarrow \infty$. We have to show that $X_p \in E_p$. The point $X_p^{(n)}$ is the p th partial point of a point $X^{(n)} = (x_1^{(n)}, x_2^{(n)}, x_3^{(n)}, \dots)$ of E , and $E_q (q > p)$ being all bounded sets, applying the theorem of Weierstrass-Bolzano and the diagonal method of Cantor we get a convergent sub-sequence $X^{(n_i)}$ out of the sequence $X^{(n)}$, that is to say $X^{(n_i)} \rightarrow X = (x_1, x_2, x_3, \dots)$ for $i \rightarrow \infty$. But E is closed and hence $X \in E$. Hence $X_p \in E_p$.

LEMMA 2.2. *Let the set E and also all its partial sets E_p be closed. Necessary and sufficient conditions which insure that the point $X = (x_1, x_2, x_3, \dots)$ belong to E are that its p th partial point $X_p = (x_1, x_2, \dots, x_p)$ shall belong to E_p and this for every $p = 1, 2, 3, \dots$.*

The conditions are obviously necessary by the definition of E . To show their sufficiency let us suppose that they are satisfied. To $X_p = (x_1, x_2, \dots, x_p)$ of E_p corresponds a point $X^{(p)} = (x_1, \dots, x_p, x_{p+1}^{(p)}, x_{p+2}^{(p)}, \dots)$ of E . Obviously $X^{(p)} \rightarrow X$ for $p \rightarrow \infty$, and E being closed, we infer that $X \in E$.

Connected with Lemma 2.2 is also the following

LEMMA 2.3. *Let E_1, E_2, E_3, \dots be a sequence of closed point sets in S_1, S_2, S_3, \dots respectively, and such that for every pair of integers p and q ($0 < p < q$) E_p is the p th partial set of E_q . Denote by E the set of points $X = (x_1, x_2, x_3, \dots)$ of S , with the property that for every $p = 1, 2, 3, \dots$ the partial point $X_p = (x_1, x_2, \dots, x_p)$ belongs to E_p . The set E of S is then closed and E_p is its p th partial*

set and E is thus completely defined by the given sequence, which dependence shall be written thus:

$$E = (E_1, E_2, E_3, \dots).$$

We have to prove two things: first that E_p is the p th partial set of E , second that E is closed. If $X \subset E$ then certainly $X_p \subset E_p$, according to the definition of E . Obviously also every $X_p \subset E_p$ is the p th partial point of a point X in E . To show that E is closed let $X^{(1)}, X^{(2)}, X^{(3)}, \dots$ be a sequence of points of E with $X^{(n)} \rightarrow X$. We have to show that $X \subset E$. But $X_p^{(n)} \rightarrow X_p$, hence $X_p \subset E_p$ and therefore also $X \subset E$.

3. Convex bodies in S_∞ . The point set E in $S = S_\infty$ is said to be *convex* if together with the points

$$X = (x_1, x_2, x_3, \dots), \quad X' = (x'_1, x'_2, x'_3, \dots),$$

also the point

$$X'' = \alpha X + \alpha' X' = (\alpha x_1 + \alpha' x'_1, \alpha x_2 + \alpha' x'_2, \dots)$$

with $\alpha > 0, \alpha' > 0, \alpha + \alpha' = 1$, belongs to E . A closed and convex set of points in S shall be called a *convex body* and denoted by K , and its partial sets by K_p . It is obvious that also $K_p (p = 1, 2, 3, \dots)$ is convex in S_p , but not necessarily closed. However, from Lemmas 2.1 and 2.2 we derive the following

COROLLARY 3.1. *If K is bounded, then all its partial sets $K_p (p = 1, 2, 3, \dots)$ are closed, bounded and convex bodies. Moreover*

$$K = (K_1, K_2, K_3, \dots)$$

in the sense of Lemma 2.3.

An example of a bounded convex body K is the set of points $X = (x)$ with

$$0 \leq x_n \leq 1 \quad (n = 1, 2, 3, \dots).$$

The partial body K_p is defined by

$$0 \leq x_n \leq 1 \quad (n = 1, 2, \dots, p)$$

and $K = (K_1, K_2, K_3, \dots)$. This K may be called the *unit-cube* in S . Another example is the set of points with

$$x_1^2 + x_2^2 + x_3^2 + \dots \leq 1.$$

The partial body K_p is the unit-hypersphere $x_1^2 + x_2^2 + \dots + x_p^2 \leq 1$ in S_p and $K = (K_1, K_2, K_3, \dots)$. It may be called the *unit-sphere* in S . The set of points $X = (x_1, x_2, x_3, \dots)$ such that $x_1^2 + x_2^2 + x_3^2 + \dots$ is convergent (Hilbert space) is convex but neither bounded nor closed.

4. **Finite-rowed systems of linear inequalities in infinitely many variables.** Let us consider the finite-rowed system of linear inequalities

$$(4.1) \quad f_i \equiv a_{i0} + a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in_i}x_{n_i} \geq 0 \quad (i = 1, 2, 3, \cdots).$$

This system will involve all the variables x_1, x_2, x_3, \cdots , or only a finite number of them, according as the sequence $n_i (i=1, 2, 3, \cdots)$ is unbounded or not. As usual we shall say that $X = (x_1, x_2, x_3, \cdots)$ is a solution of (4.1), in case all the inequalities of (4.1) are satisfied, and (4.1) is *consistent* if there is at least one such solution (different from the origin $0 = (0, 0, 0, \cdots)$) in the homogeneous case $a_{i0} = a_{20} = \cdots = 0$). It is obvious that

(4.2) *If (4.1) is consistent, then the set of points X of S which are solutions of (4.1), represent a closed and convex body K in S .*

The system of inequalities $rx_1 + x_2 - 2r^{1/2} \geq 0$, where r takes all positive rational values, is of the type (4.1), and because $rx_1 + x_2 - 2r^{1/2} = 0$ is a tangent to the branch of a hyperbola defined by $x_1x_2 = 1$, $x_1 > 0$, $x_2 > 0$, the partial set K_1 of the set of its solutions is identical with the set of S_1 defined by $x_1 > 0$, hence not closed. We see from this example that while the set K of solutions of (4.1) is convex and closed, its partial sets K_p need not to be closed also. But if K is bounded, then also all the K_p are closed according to Corollary 3.1. For this case we shall prove the following converse:

THEOREM 4.1. *Let $K = (K_1, K_2, K_3, \cdots)$ be a closed, convex and bounded set of points in S (a bounded convex body in S). There is then a system of the type (4.1) such that K is identical with the set of solutions of this system.*

The partial set $K_p (p=1, 2, 3, \cdots)$ is bounded and therefore closed in S_p (Corollary 3.1). Let $P_1^{(p)}, P_2^{(p)}, P_3^{(p)}, \cdots$ be all the points of S_p exterior to K_p and which have only rational coordinates. To every such point $P_i^{(p)}$ there corresponds a hyperplane

$$\pi_i^{(p)} \equiv c_{i0}^{(p)} + c_{i1}^{(p)}x_1 + \cdots + c_{ip}^{(p)}x_p = 0$$

passing through this point and which is a bound for K_p . Every point $X_p = (x_1, x_2, \cdots, x_p)$ in K_p is a solution of the system $\pi_i^{(p)} \geq 0$ ($i=1, 2, 3, \cdots$), and conversely, every solution of it is a point of K_p . The combined system

$$\pi_i^{(p)} = c_{i0}^{(p)} + c_{i1}^{(p)}x_1 + \cdots + c_{ip}^{(p)}x_p \geq 0 \quad (i, p = 1, 2, 3, \cdots)$$

is of the type (4.1) and from Corollary 3.1 we infer that K is identical with the set of its solutions in S .

5. A problem of F. Riesz. F. Riesz (loc. cit., §IX) has investigated the following problem.

(5.1) Let $\phi_1(t), \phi_2(t), \phi_3(t), \dots$ be a given sequence of real continuous functions for $0 \leq t \leq 1$. There is to be determined the domain of variability of the sequence

$$(5.1') \quad x_n = \int_0^1 \phi_n(t) d\chi(t) \quad (n = 1, 2, 3, \dots)$$

where $\chi(t)$ is monotonic for $0 \leq t \leq 1$ with

$$(5.1'') \quad \int_0^1 d\chi(t) = 1.$$

Riesz' solution is contained in the following

THEOREM 5.1 (F. Riesz). Consider in S the continuous arc

$$C_p : x_n = \phi_n(t) \quad (0 \leq t \leq 1; n = 1, 2, \dots, p),$$

and let $K_p = K(C_p)$. The domain of variability of the sequence (5.1') is the bounded convex body in S defined by

$$K = (K_1, K_2, K_3, \dots).$$

It is obvious that $X = (x_1, x_2, x_3, \dots)$ as given by (5.1') and (5.1'') is a point of K , because, for every value of p , $X_p = (x_1, x_2, \dots, x_p)$ belongs to K_p (X_p is the limiting point of a sequence of points of K_p by the definition of the Stieltjes integral and K_p is closed). Let us prove now that any point X of K may be written in the form (5.1') with (5.1''). If X belongs to K , then X_p ($p = 1, 2, 3, \dots$) belongs to $K_p = K(C_p)$ and a theorem of Carathéodory (Carathéodory, loc. cit., §9) permits us to write

$$(5.2) \quad x_n = \sum_{m=0}^p \phi_n(t_{p,m}) \lambda_{p,m} \quad (n = 1, 2, \dots, p),$$

with $0 \leq t_{p,0} < t_{p,1} < \dots < t_{p,p} \leq 1$, $\lambda_{p,m} \geq 0$ and $\sum_m \lambda_{p,m} = 1$. Let $\chi_p(t)$ be the monotonic step-function defined on $0 \leq t \leq 1$ by $\chi_p(0) = 0$, $2\chi_p(t) = \chi_p(t+0) + \chi_p(t-0)$ for $0 < t < 1$, and whose jump at the point $t = t_{p,m}$ ($m = 0, 1, \dots, p$) is $\lambda_{p,m}$. The system (5.2) may be written as

$$(5.3) \quad x_n = \int_0^1 \phi_n(t) d\chi_p(t) \quad (n = 1, 2, \dots, p).$$

The sequence $\chi_p(t)$ is uniformly bounded and a theorem of Helly* insures the

* E. Helly, *Über lineare Funktionaloperationen*, Sitzungsberichte der Wiener Akademie, vol. 121, IIa (1912), p. 286.

existence of a sub-sequence $\chi_q(t)$ such that $\chi_q(t) \rightarrow \chi(t)$ for $q \rightarrow \infty$ and $0 \leq t \leq 1$. When $p = q \rightarrow \infty$, another theorem of Helly (loc. cit., pp. 288–289) shows that (5.3) becomes (5.1') with (5.1''). Theorem 5.1 is therefore proved.

Combining Theorem 5.1 with Theorem 4.1 we get at once the following

COROLLARY 5.1. *There is a system of linear inequalities of the type (4.1), such that the domain of variability asked for in (5.1) is identical with the set of solutions of this system.*

A typical example for such a solution of Riesz' problem has been given by Hausdorff* for the special sequence $\phi_n(t) = t^n$ ($n = 1, 2, 3, \dots$). He found that the domain of variability of the sequence

$$(5.4) \quad x_n = \int_0^1 t^n d\chi(t) \quad (n = 1, 2, 3, \dots),$$

where $\chi(t)$ is monotonic and $\chi(1) - \chi(0) = 1$, is identical with the set of sequences x_1, x_2, x_3, \dots which are solutions of the system

$$(5.5) \quad \Delta^k x_l \equiv x_l - \binom{k}{1} x_{l+1} + \binom{k}{2} x_{l+2} - \dots + (-1)^k x_{l+k} \geq 0$$

$$(k, l = 0, 1, 2, \dots, x_0 = 1).$$

PART II. CONCERNING A CERTAIN CLASS OF FINITE-ROWED SYSTEMS OF LINEAR INEQUALITIES IN INFINITELY MANY VARIABLES

6. Statement of the problem. Let

$$(6.1) \quad A = \left\| \begin{array}{cccc} a_{01} & a_{02} & a_{03} & \dots \\ a_{11} & a_{12} & a_{13} & \dots \\ a_{21} & a_{22} & a_{23} & \dots \\ \vdots & \vdots & \vdots & \vdots \end{array} \right\|$$

be an infinite matrix of real numbers whose minors satisfy the conditions

$$(6.2) \quad (i_1, i_2, \dots, i_m) = \left| \begin{array}{cccc} a_{i_1,1} & a_{i_1,2} & \dots & a_{i_1,m} \\ a_{i_2,1} & a_{i_2,2} & \dots & a_{i_2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_m,1} & a_{i_m,2} & \dots & a_{i_m,m} \end{array} \right| > 0,$$

for $0 \leq i_1 < i_2 < \dots < i_m$ ($m = 1, 2, 3, \dots$). Let us also consider the matrix

* F. Hausdorff, *Über das Momentenproblem für ein endliches Intervall*, Mathematische Zeitschrift, vol. 16 (1923), §1. See also I. J. Schoenberg (loc. cit.).

$$(6.3) \quad \|x, A\| = \left\| \begin{array}{cccc} x_0 & a_{01} & a_{02} & a_{03} \cdots \\ x_1 & a_{11} & a_{12} & a_{13} \cdots \\ x_2 & a_{21} & a_{22} & a_{23} \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{array} \right\|$$

obtained by bordering the matrix A with a column of variables, and let us define

$$(6.4) \quad D^k x_l \equiv \left| \begin{array}{cccc} x_l & a_{l,1} & a_{l,2} & \cdots a_{l,k} \\ x_{l+1} & a_{l+1,1} & a_{l+1,2} & \cdots a_{l+1,k} \\ \vdots & \vdots & \vdots & \vdots \\ x_{l+k} & a_{l+k,1} & a_{l+k,2} & \cdots a_{l+k,k} \end{array} \right| \quad (k, l = 0, 1, 2, \cdots),$$

where $D^0 x_l$ means x_l .

We shall be concerned with the problem of solving the system

$$(6.5) \quad D^k x_l \geq 0 \quad (k, l = 0, 1, 2, \cdots),$$

which is a homogeneous, finite-rowed system of linear inequalities in the variables x_0, x_1, x_2, \cdots .

An interesting special case of (6.5) is obtained for

$$(6.6) \quad A = \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \cdots \\ 1 & 1 & 0 & 0 \cdots \\ 1 & 2 & 1 & 0 \cdots \\ 1 & 3 & 3 & 1 \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{array} \right\|,$$

which is the matrix of the binomial coefficients. It is readily shown that the conditions (6.2) are satisfied (hence those conditions are consistent) and that

$$D^k x_l \equiv \Delta^k x_l \equiv x_l - \binom{k}{1} x_{l+1} + \cdots + (-1)^k x_{l+k}.$$

Hence (6.5) is identical in this particular case with Hausdorff's system (5.5), whose most general solution is given by (5.4), as shown by Hausdorff. A similar solution will be given for the general system (6.5).

7. The solution of the finite partial system. In this section we shall solve the finite partial system

$$(7.1) \quad D^k x_l \geq 0 \quad (k, l = 0, 1, 2, \cdots, p; k + l \leq p),$$

which involves only the variables x_0, x_1, \cdots, x_p .

Let us first prove the following recursion formula:

$$(7.2) \quad D^k x_l = \frac{(l+1, l+2, \dots, l+k)}{(l+1, l+2, \dots, l+k+1)} D^{k+1} x_l + \frac{(l, l+1, \dots, l+k)}{(l+1, l+2, \dots, l+k+1)} D^k x_{l+1}.$$

The algebraic complements of the four corner elements of the determinant

$$D^{k+1} x_l = \begin{vmatrix} x_l & a_{l,1} & \cdots & a_{l,k} & a_{l,k+1} \\ x_{l+1} & a_{l+1,1} & \cdots & a_{l+1,k} & a_{l+1,k+1} \\ \vdots & \vdots & & \vdots & \vdots \\ x_{l+k} & a_{l+k,1} & \cdots & a_{l+k,k} & a_{l+k,k+1} \\ x_{l+k+1} & a_{l+k+1,1} & \cdots & a_{l+k+1,k} & a_{l+k+1,k+1} \end{vmatrix}$$

are respectively equal to

$$(l+1, l+2, \dots, l+k+1), \pm (l, l+1, \dots, l+k), D^k x_l \text{ and } \pm D^k x_{l+1}.$$

Their determinant is

$$(l+1, l+2, \dots, l+k+1) D^k x_l - (l, l+1, \dots, l+k) D^k x_{l+1}$$

and is equal, according to a theorem of Sylvester,* to the original determinant $D^{k+1} x_l$ times its central minor $(l+1, \dots, l+k)$. This proves the identity (7.2).

From (7.2) and (6.2) we infer that the system (7.1) is equivalent to its partial system

$$(7.3) \quad D^{p-n} x_n \geq 0 \quad (n = 0, 1, 2, \dots, p).$$

Indeed, a repeated application of (7.2) shows that every $D^k x_l$ with $k+l < p$ appears as a linear combination with positive coefficients of some of the left hand members of (7.3). It suffices therefore to solve the system (7.3). All we have to do is to find the linear transformation which is inverse to the transformation

$$(7.4) \quad y_{p,n} = D^{p-n} x_n \quad (n = 0, 1, 2, \dots, p),$$

which in more explicit form is

$$(7.5) \quad \begin{aligned} y_{p,n} = & (n+1, n+2, \dots, p) x_n - (n, n+2, \dots, p) x_{n+1} \\ & + (n, n+1, n+3, \dots, p) x_{n+2} + \cdots \\ & + (-1)^{p-n} (n, n+1, \dots, p-1) x_p \quad (n = 0, 1, \dots, p). \end{aligned}$$

* Maxime Bôcher, *Introduction to Higher Algebra*, New York, 1924, p. 33, Corollary 3.

Let us verify directly that the inverse transformation is

$$(7.6) \quad x_n = \sum_{m=0}^p \frac{(n, m+1, m+2, \dots, p)}{(m+1, \dots, p)(m, m+1, \dots, p)} y_{p,m} \quad (n = 0, 1, \dots, p),$$

where the coefficient of $y_{p,p}$ is $(n)/(p) = a_{n1}/a_{p1}$. Let c_{nm} ($n, m = 0, 1, \dots, p$) be the matrix product of the matrices of the transformations (7.5) and (7.6). In both transformations the coefficients below the principal diagonal are zero and therefore also $c_{nm} = 0$ for $n > m$. On the other hand $c_{nn} = 1$ for $n = 0, 1, \dots, p$. Finally for $n < m$ we get

$$c_{nm} = \frac{1}{(m+1, \dots, p)(m, m+1, \dots, p)} \times \begin{vmatrix} (n, m+1, \dots, p) & a_{n,1} & a_{n,2} & \dots & a_{n,p-n} \\ (n+1, m+1, \dots, p) & a_{n+1,1} & a_{n+1,2} & \dots & a_{n+1,p-n} \\ (n+2, m+1, \dots, p) & a_{n+2,1} & a_{n+2,2} & \dots & a_{n+2,p-n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (p, m+1, \dots, p) & a_{p,1} & a_{p,2} & \dots & a_{p,p-n} \end{vmatrix}$$

and this determinant vanishes because its first column is a linear combination of the next $p-m+1$ columns ($p-m+1 \leq p-n$). Hence $\|c_{nm}\|$ is the unit matrix and (7.6) is indeed the inverse transformation of (7.4).

Let us use for convenience instead of the $y_{p,m} = D^{p-m}x_m$, the quantities

$$(7.7) \quad \lambda_{p,m} = \frac{(0, m+1, m+2, \dots, p)}{(m+1, \dots, p)(m, m+1, \dots, p)} D^{p-m}x_m,$$

and the system (7.6) becomes

$$(7.8) \quad x_n = \sum_{m=0}^p \frac{(n, m+1, m+2, \dots, p)}{(0, m+1, m+2, \dots, p)} \lambda_{p,m} \quad (n = 0, 1, \dots, p),$$

where the coefficient of $\lambda_{p,p}$ is $(n)/(0)$.

We have proved so far the following

LEMMA 7.1. *The most general solution of the finite system (7.1) is given by the set (7.8) for arbitrary values of the $\lambda_{p,m}$ with*

$$(7.9) \quad \lambda_{p,0} \geq 0, \lambda_{p,1} \geq 0, \dots, \lambda_{p,p} \geq 0.$$

8. The solution of the infinite system (6.5). The solution of the infinite system (6.5) will be prepared by a closer study of the structure of the system (7.8), for various values of p . The fundamental conditions (6.2) require in particular that $(i) = a_{i1} > 0$ for $i = 0, 1, 2, \dots$. Without any loss of generality

we may and shall suppose that

$$(i) = a_{i1} = 1 \quad (i = 0, 1, 2, \dots).$$

For convenience we write

$$(8.1) \quad c_{n,m,p} = \frac{(n, m+1, m+2, \dots, p)}{(0, m+1, m+2, \dots, p)} \left(m = 0, 1, \dots, p; c_{n,p,p} = \frac{(n)}{(0)} = 1 \right),$$

and (7.8) becomes

$$(8.2) \quad x_n = \sum_{m=0}^p c_{n,m,p} \lambda_{p,m} \quad (n = 0, 1, \dots, p).$$

From (7.7) and (7.2) we readily get the recursion formula

$$\begin{aligned} \lambda_{p,m} = & \frac{(m, m+1, \dots, p+1)}{(m, m+1, \dots, p)} \frac{(0, m+1, m+2, \dots, p)}{(0, m+1, m+2, \dots, p+1)} \lambda_{p+1,m} \\ & + \frac{(m+2, m+3, \dots, p+1)}{(m+1, m+2, \dots, p)} \frac{(0, m+1, m+2, \dots, p)}{(0, m+2, m+3, \dots, p+1)} \lambda_{p+1,m+1}, \end{aligned}$$

or

$$(8.3) \quad \lambda_{p,m} = g_{p,m} \lambda_{p+1,m} + h_{p,m+1} \lambda_{p+1,m+1} \quad (m = 0, 1, \dots, p),$$

if we define

$$(8.4) \quad \begin{aligned} g_{p,m} &= \frac{(m, \dots, p+1)}{(m, \dots, p)} \frac{(0, m+1, \dots, p)}{(0, m+1, \dots, p+1)}, \\ h_{p,m} &= \frac{(m+1, \dots, p+1)}{(m, \dots, p)} \frac{(0, m, \dots, p)}{(0, m+1, \dots, p+1)}, \end{aligned}$$

while in (8.3) $h_{p,p+1} = 1$.

For a particular value of $m (m = 1, 2, \dots, p)$ we put

$$(8.5) \quad \lambda_{p+1,m} = 1, \text{ and } \lambda_{p+1,r} = 0 \quad (0 \leq r \leq p+1; r \neq m).$$

For these particular values of the λ , the set (8.2), taken for $p+1$ instead of p , becomes

$$(8.6) \quad x_n = c_{n,m,p+1} \quad (n = 0, 1, \dots, p+1).$$

On the other hand we get from (8.5) and (8.3)

$$\lambda_{p,m-1} = h_{p,m}, \lambda_{p,m} = g_{p,m} \text{ and } \lambda_{p,s} = 0 \quad (0 \leq s \leq p; s \neq m-1, m).$$

With these particular values we get from (8.2)

$$(8.7) \quad x_n = c_{n,m-1,p} h_{p,m} + c_{n,m,p} g_{p,m} \quad (n = 0, 1, \dots, p).$$

Equating the two results (8.6) and (8.7) we get the recursion formulas

$$(8.8) \quad c_{n,m,p+1} = h_{p,m}c_{n,m-1,p} + g_{p,m}c_{n,m,p} \quad \begin{pmatrix} m = 1, 2, \dots, p \\ n = 0, 1, \dots, p \end{pmatrix},$$

while $c_{n,p+1,p+1} = c_{n,p,p} = (n)/(0) = 1$, which relation will also be included in (8.8) for $m = p+1$, if we define $h_{p,p+1} = 1$ and $g_{p,p+1} = 0$. On the other hand, from (8.8) and (8.1) we get for $n=0$ the relations

$$(8.9) \quad h_{p,m} + g_{p,m} = 1, \quad h_{p,m} > 0, \quad g_{p,m} > 0 \quad (m = 1, 2, \dots, p).$$

Let us define

$$(8.10) \quad t_{p,m} = c_{1,m,p} \quad (m = 0, 1, \dots, p).$$

For a particular value of n ($n=0, 1, 2, 3, \dots$), let us consider the plane of the variables (t, u_n) and let $P_n^{(p)}$ ($p \geq n$) denote the polygonal line joining successively the $p+1$ vertices

$$(t_{p,m}, c_{n,m,p}) \quad (m = 0, 1, \dots, p).$$

We shall first prove that

$$(8.11) \quad 0 = t_{p,0} < t_{p,1} < \dots < t_{p,p-1} < t_{p,p} = 1 \quad (p = 1, 2, 3, \dots).$$

For $p=1$ we have indeed $t_{1,0} = c_{1,0,1} = 0$, $t_{1,1} = c_{1,1,1} = 1$. Suppose (8.11) to be true for a particular value of p . From (8.8) and (8.10) we get for $n=1$

$$(8.12) \quad t_{p+1,m} = h_{p,m}t_{p,m-1} + g_{p,m}t_{p,m} \quad (m = 1, 2, \dots, p),$$

while $t_{p+1,0} = 0$, $t_{p+1,p+1} = 1$. From (8.11), (8.12) and (8.9) we infer that (8.11) is also true for $p+1$ instead of p . Hence (8.11) holds for every value of $p=1, 2, 3, \dots$. In particular it follows that $P_0^{(p)}$ is the straight segment joining the points $(0, 1)$ and $(1, 1)$, while $P_1^{(p)}$ is the segment joining $(0, 0)$ with $(1, 1)$.

Let us now consider the polygonal line $P_n^{(p)}$ for $p=n, n+1, n+2, \dots$. The polygonal line $P_n^{(n)}$ has n sides and its consecutive vertices are $(0, 0)$, $(t_{n,1}, 0)$, \dots , $(t_{n,n-1}, 0)$, $(1, 1)$. The formulas (8.12), (8.8) and (8.9), all taken for $p=n$, show that $P_n^{(n+1)}$ is *inscribed* in $P_n^{(n)}$ in the sense that they have the same end points, while the other vertices of $P_n^{(n+1)}$ are consecutively situated on the sides of $P_n^{(n)}$ (limits excluded). A repeated application of the same arguments will show that if

$$(8.13) \quad u_n = P_n^{(p)}(t) \quad (0 \leq t \leq 1, p \geq n)$$

denotes the analytical representation of the polygonal line $P_n^{(p)}$, then

$$(8.14) \quad P_n^{(p)}(t) \text{ is a non-negative, continuous, non-decreasing and convex function of } t \text{ for } 0 \leq t \leq 1,$$

and

$$(8.15) \quad P_n^{(p)}(t) \leq P_n^{(p')}(t) \leq t \text{ for } 1 \leq n \leq p < p', \quad 0 \leq t \leq 1.$$

From the convexity of the function $P_n^{(p)}(t)$ and the fact that the last side of $P_n^{(p)}$ is the steepest side and has the slope

$$k_n = (1 - t_{n,n-1})^{-1},$$

we infer that

$$(8.16) \quad 0 \leq P_n^{(p)}(t'') - P_n^{(p)}(t') \leq k_n(t'' - t') \text{ for } 0 \leq t' < t'' \leq 1 \quad (p \geq n).$$

Now we can readily prove that

$$(8.17) \quad \lim_{p \rightarrow \infty} P_n^{(p)}(t) = \phi_n(t),$$

where $\phi_n(t)$ is a non-negative, continuous, non-decreasing and convex function of t for $0 \leq t \leq 1$, while the relation (8.17) holds uniformly with respect to t for $0 \leq t \leq 1$.

We have previously seen that $P_0^{(p)}(t) \equiv 1$, hence $\phi_0(t) \equiv 1$, while $P_1^{(p)}(t) \equiv t$ and hence $\phi_1(t) \equiv t$. The existence of the limit (8.17) for $n > 1$ follows from (8.15). The limiting function $\phi_n(t)$ is obviously non-negative, non-decreasing and convex from (8.14). We still have to show first, that $\phi_n(t)$ is continuous, and second, that (8.17) holds uniformly with respect to t . From (8.17) and (8.16) we get the Lipschitz condition

$$(8.18) \quad 0 \leq \phi_n(t'') - \phi_n(t') \leq k_n(t'' - t') \text{ for } 0 \leq t' < t'' \leq 1,$$

which insures the continuity of $\phi_n(t)$. The uniformity of convergence in (8.17) follows immediately from the fact that the sequence $P_n^{(n)}(t)$, $P_n^{(n+1)}(t)$, \dots is monotone, while its limiting function $\phi_n(t)$ is continuous.*

Now we can readily prove the following

THEOREM 8.1. (1) *Every solution of the finite-rowed system*

$$(8.19) \quad D^k x_l \geq 0 \quad (k, l = 0, 1, 2, 3, \dots), \quad a_{01} = a_{11} = a_{21} = \dots = 1,$$

may be expressed in the form

$$(8.20) \quad x_n = \int_0^1 \phi_n(t) d\chi(t) \quad (n = 0, 1, 2, \dots),$$

where the $\phi_n(t)$ are the functions defined by (8.17), while $\chi(t)$ is a non-decreasing function for $0 \leq t \leq 1$, and conversely, every sequence x_n defined by (8.20) with $\chi(t)$ non-decreasing, is a solution of the system (8.19).

* On the basis of the classical theorem of Dini. Cf. Carathéodory, *Vorlesungen über reelle Funktionen*, 2d edition, 1927, p. 276, Satz 4.

(2) A necessary and sufficient condition that the function $\chi(t)$ be uniquely defined by the set (8.20) and the additional conditions

$$(8.21) \quad \chi(0) = 0, \quad 2\chi(t) = \chi(t+0) + \chi(t-0) \text{ for } 0 < t < 1,$$

is that every function $f(t)$, continuous on $0 \leq t \leq 1$, shall be uniformly approximable as close as we want by a suitable linear combination of functions of the sequence

$$(8.22) \quad \phi_0(t) \equiv 1, \quad \phi_1(t) \equiv t, \quad \phi_2(t), \quad \phi_3(t), \quad \dots$$

Let x_0, x_1, x_2, \dots be a solution of (8.19). We have to express this sequence in the form (8.20). It is obvious that the set x_0, x_1, \dots, x_p is a solution of the partial system (7.1) for any value of p , and hence (7.7), (7.8) or (8.2), and (7.9) will hold for every value of p . Let $\chi_p(t)$ be a step-function on $0 \leq t \leq 1$, whose jump at the point $t = t_{p,m}$ ($m = 0, 1, \dots, p$) is $\lambda_{p,m}$, and which is uniquely defined by the additional conditions $\chi_p(0) = 0$, $2\chi_p(t) = \chi_p(t+0) + \chi_p(t-0)$ for $0 < t < 1$. We infer from (7.9) that $\chi_p(t)$ is non-decreasing and the first equation of the set (8.2) gives

$$(8.23) \quad \chi_p(1) = \sum_{m=0}^p \lambda_{p,m} = x_0.$$

Let n have a particular fixed value. By the definition of the function $P_n^{(p)}(t)$ ($p \geq n$), the n th equation of the set (8.2) may be written thus:

$$(8.24) \quad x_n = \sum_{m=0}^p P_n^{(p)}(t_{p,m}) \lambda_{p,m}.$$

The uniform convergence of (8.17) and (8.23) permit us to write (8.24) thus:

$$(8.25) \quad x_n = \sum_{m=0}^p \phi_n(t_{p,m}) \lambda_{p,m} + \epsilon_n^{(p)} = \int_0^1 \phi_n(t) d\chi_p(t) + \epsilon_n^{(p)}$$

with $\epsilon_n^{(p)} \rightarrow 0$ for $p \rightarrow \infty$.

On the other hand, (8.23) and $\chi_p(0) = 0$ show that the sequence of non-decreasing functions $\chi_p(t)$ is uniformly bounded. From Helly's first theorem (loc. cit.) we can get a subsequence $\chi_q(t)$ with $\chi_q(t) \rightarrow \chi(t)$ for $q \rightarrow \infty$ and $0 \leq t \leq 1$. For $p = q \rightarrow \infty$, the second theorem of Helly shows that (8.25) becomes (8.20).

In order to complete the proof of the first part of Theorem 8.1, we have to prove that the sequence given by 8.20 is a solution of (8.19). It will suffice to show that the partial sequence x_0, x_1, \dots, x_p is a solution of the partial set

$$(8.26) \quad D^k x_l \geq 0 \quad (k + l \leq p),$$

for every particular value of p . From Lemma 7.1 we know that

$$x_n^{(m)} = c_{n,m,p} \quad (n = 0, 1, \dots, p)$$

is a solution of (8.26), for every $m = 0, 1, \dots, p$. Hence also

$$x_n^{(p)}(t) = P_n^{(p)}(t) \quad (n = 0, 1, \dots, p)$$

is a solution of (8.26) for every t with $0 \leq t \leq 1$. The relation (8.17) shows that also

$$x_n(t) = \phi_n(t) \quad (n = 0, 1, \dots, p)$$

is a solution of (8.26) for every t with $0 \leq t \leq 1$. Therefore also

$$x_n = \int_0^1 \phi_n(t) d\chi(t) \quad (n = 0, 1, \dots, p)$$

is a solution of (8.26) for every non-decreasing function $\chi(t)$. This last result holds for every value of p and shows that the sequence defined by (8.20) is a solution of (8.19).

The second part of Theorem 8.1 is a direct consequence of a theorem of F. Riesz.* We may add the following remark.

(8.27) *A necessary condition that the function $\chi(t)$ be uniquely defined by (8.20) and (8.21) is that the set of points*

$$t_{p,0}, t_{p,1}, t_{p,2}, \dots, t_{p,p} \quad (p = 1, 2, 3, \dots)$$

shall be everywhere dense on the interval $0 \leq t \leq 1$.†

In order to prove (8.27), let us suppose our set to be not everywhere dense and let the interval $\alpha\beta$ ($0 \leq \alpha < \beta \leq 1$) contain no point of the set. The polygonal lines $P_n^{(p)}$ ($p \geq n$) have no vertex for values of t on $\alpha\beta$ and $P_n^{(p)}(t)$ and therefore also $\phi_n(t)$ are linear functions for $\alpha \leq t \leq \beta$, say $\phi_n(t) = a_n + b_n t$. Let us suppose $\chi(t)$ to be non-decreasing with $\chi(\alpha) < \chi(\beta)$. We may write

$$x_n = \int_0^1 \phi_n(t) d\chi(t) = \int_0^\alpha + \int_\beta^1 + \int_\alpha^\beta \phi_n(t) d\chi(t)$$

and

* F. Riesz, loc. cit., p. 53. The theorem says that if every continuous function on $0 \leq t \leq 1$ is uniformly approximable as close as we want by a suitable linear combination of functions of the sequence $\phi_0(t), \phi_1(t), \phi_2(t), \dots$ then the function $\chi(t)$ of bounded variation is essentially uniquely defined by the system (8.20). For a geometric proof of this theorem see the recent paper of W. Seidel, *On the approximation of continuous functions by linear combinations of continuous functions*, Annals of Mathematics, (2), vol. 32 (1931), pp. 777-784.

† It would be of interest to prove that this condition is also sufficient for the uniqueness of $\chi(t)$.

$$\begin{aligned}\int_{\alpha}^{\beta} \phi_n(t) d\chi(t) &= \phi_n(t)\chi(t) \Big|_{\alpha}^{\beta} - \int_{\alpha}^{\beta} \chi(t) d\phi_n(t) \\ &= \phi_n(\beta)\chi(\beta) - \phi_n(\alpha)\chi(\alpha) - b_n \int_{\alpha}^{\beta} \chi(t) dt.\end{aligned}$$

These formulas show that the sequence x_n defined by (8.20) is not changed by altering the function $\chi(t)$ within the interval $\alpha\beta$, so as to leave the three quantities $\chi(\alpha)$, $\chi(\beta)$ and $\int_{\alpha}^{\beta} \chi(t) dt$ unchanged. This proves the statement (8.27).

The first part of Theorem 8.1 can also be stated as follows. *The sequence x_0, x_1, x_2, \dots being given, a necessary and sufficient condition that there be a non-decreasing function $\chi(t)$, solution of the set (8.20), is that all the inequalities (8.19) shall hold.* Following a method devised by F. Hausdorff in a special case,* we may readily answer the similar question concerning functions $\chi(t)$ of bounded variation. I may omit the proof of the result which I shall state as

COROLLARY 8.1. *The sequence x_0, x_1, x_2, \dots being given, a necessary and sufficient condition that there be a function $\chi(t)$ of bounded variation satisfying the set of equations (8.20), is that the sum*

$$\sum_{m=0}^p |\lambda_{p,m}| = \sum_{m=0}^p \frac{(0, m+1, m+2, \dots, p)}{(m+1, \dots, p)(m, m+1, \dots, p)} |D^{p-m} x_m|$$

shall be bounded for $p \rightarrow \infty$. The uniqueness of $\chi(t)$ is insured under the same conditions as in Theorem 8.1.

PART III. SOME APPLICATIONS

9. The systems of Hausdorff. In this section we shall be concerned with a class of finite-rowed systems of linear inequalities which has been first considered by Hausdorff in his paper *Summationsmethoden und Momentfolgen*, II. His system is essentially equivalent to the system†

* F. Hausdorff, *Über das Momentenproblem für ein endliches Intervall*, Mathematische Zeitschrift, vol. 16 (1923), pp. 220–248. See also I. J. Schoenberg, loc. cit.

† Hausdorff's terminology and notations are quite different from those used here. He calls the sequence x_n *completely monotonic with respect to the sequence a_n* , whenever the sequence x_n is a solution of the system (9.1), while in his notation

$$D^k x_l = (-1)^k \begin{pmatrix} x_l & x_{l+1} & \dots & x_{l+k} \\ a_l & a_{l+1} & \dots & a_{l+k} \end{pmatrix} \begin{vmatrix} 1 & a_l & a_l^2 & \dots & a_l^k \\ 1 & a_{l+1} & a_{l+1}^2 & \dots & a_{l+1}^k \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_{l+k} & a_{l+k}^2 & \dots & a_{l+k}^k \end{vmatrix}.$$

Our additional assumption (9.3), which is essential for our point of view, does not restrict the problem.

$$(9.1) \quad D^k x_l \geq 0 \quad (k, l = 0, 1, 2, \dots),$$

derived as stated in §6 from the infinite Vandermondean matrix

$$(9.2) \quad A = \begin{vmatrix} 1 & a_0 & a_0^2 & a_0^3 & \cdots \\ 1 & a_1 & a_1^2 & a_1^3 & \cdots \\ 1 & a_2 & a_2^2 & a_2^3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{vmatrix},$$

whose elements a_n satisfy the following conditions:

$$(9.3) \quad 0 \leq a_0 < a_1 < a_2 < \cdots,$$

$$(9.4) \quad \lim_{r \rightarrow \infty} a_r = \infty,$$

$$(9.5) \quad \sum_{r=1}^{\infty} 1/a_r = \infty.$$

Hausdorff's results concerning the system (9.1) are essentially contained in the following

THEOREM 9.1 (Hausdorff). *The first part of Theorem 8.1 holds for the system (9.1), derived from the Vandermondean matrix (9.2), whose elements a_n satisfy the conditions (9.3), (9.4) and (9.5), with the following expressions for the functions $\phi_n(t)$:*

$$(9.6) \quad \phi_n(t) = t^{(a_n - a_0)/(a_1 - a_0)} \quad (n = 0, 1, 2, \dots; 0 \leq t \leq 1).$$

The condition for the uniqueness of the function $\chi(t)$ in (8.20) is always satisfied in this particular case.

As a consequence of (9.3), the fundamental conditions (6.2) are satisfied. Hence Theorem 8.1 may be applied. The coefficients of the variables x_n in the linear form $D^k x_l$ are either numerical constants or else Vandermondean determinants of some of the elements a_n . All these coefficients and therefore also the entire system (9.1) depend only on the various differences $a_k - a_l$ and do not change when the elements a_n are replaced by $a_n - a_0$. This shows that *without any loss of generality we may assume*

$$(9.7) \quad a_0 = 0.$$

From (8.1), (9.2) and (9.7) we readily get

$$(9.8) \quad c_{n,m,p} = \prod_{r=m+1}^p \left(1 - \frac{a_n}{a_r}\right) \quad (c_{n,p,p} = 1),$$

and (9.8) and (8.10) give

$$(9.9) \quad t_{p,m} = c_{1,m,p} = \prod_{r=m+1}^p \left(1 - \frac{a_1}{a_r}\right) \quad (t_{p,p} = 1).$$

From (9.9), (9.4) and (9.5) we may readily prove that *the numbers $t_{p,m}$ are everywhere dense on the interval $0 \leq t \leq 1$* . This is best shown by the consideration of

$$\log t_{p,m} = \sum_{r=m+1}^p \log \left(1 - \frac{a_1}{a_r}\right).$$

For m sufficiently large, the sequence $\log t_{p,m}$ ($p = m+1, m+2, \dots$) will cover the entire half-axis $-\infty \dots 0$ with gaps as small as we please, a fact which is due to the divergence of $\sum_{m+1}^{\infty} \log(1 - a_1/a_r)$ (a consequence of (9.5)) and (9.4).

Taking into account the results of §8 we see that in order to prove that $\phi_n(t) = t^{a_n/a_1}$, it suffices to prove that for every particular value of t with $0 \leq t \leq 1$, from $t_{p,m} \rightarrow t$ (for $p \rightarrow \infty$), it follows that $c_{n,m,p} \rightarrow t^{a_n/a_1}$. We have therefore to prove that

(9.10) *If $m = m(p)$ is such a function of p that*

$$(9.10') \quad \lim_{p \rightarrow \infty} \prod_{r=m+1}^p \left(1 - \frac{a_1}{a_r}\right) = t \quad (t \text{ is a particular value with } 0 \leq t \leq 1),$$

then also

$$(9.10'') \quad \lim_{p \rightarrow \infty} \prod_{r=m+1}^p \left(1 - \frac{a_n}{a_r}\right) = t^{a_n/a_1} \quad (n \geq 2)$$

will hold.

This assertion is obviously true for $t=0$, because of

$$0 \leq c_{n,m,p} \leq t_{p,m} \quad (n \geq 2).$$

We may hence suppose that

$$(9.11) \quad 0 < t \leq 1.$$

The inequality $1 - x \leq e^{-x}$ ($x \geq 0$) insures the inequality

$$\prod_{m+1}^p \left(1 - \frac{a_1}{a_r}\right) \leq \exp \left\{ -a_1 \sum_{m+1}^p \frac{1}{a_r} \right\}^*,$$

which shows that (9.10') and (9.11) imply that

* We write $e^y = \exp\{y\}$.

$$(9.12) \quad \sum_{m+1}^p \frac{1}{a_r} \text{ is bounded for } p \rightarrow \infty.$$

A consequence of (9.5) and (9.12) is

$$(9.13) \quad \lim_{p \rightarrow \infty} m(p) = \infty.$$

Applying Taylor's formula

$$\log(1-x) = -x - \frac{x^2}{2}(1-\theta x)^{-2} \quad (0 \leq x < 1, \theta = \theta(x) \text{ with } 0 < \theta < 1),$$

we get for $m(p) \geq 1$,

$$(9.14) \quad \prod_{m+1}^p \left(1 - \frac{a_1}{a_r}\right) = \exp \left\{ -a_1 \sum_{m+1}^p \frac{1}{a_r} \right\} \cdot \exp \left\{ -\frac{a_1^2}{2} \sum_{m+1}^p \frac{1}{a_r^2} \left(1 - \theta_r \frac{a_1}{a_r}\right)^{-2} \right\}.$$

For the exponent of the second factor on the right side of (9.14) we have

$$0 \leq \frac{a_1^2}{2} \sum_{m+1}^p \frac{1}{a_r^2} \left(1 - \theta_r \frac{a_1}{a_r}\right)^{-2} < \frac{a_1^2}{2} \left(1 - \frac{a_1}{a_2}\right)^{-2} \frac{1}{a_{m+1}} \sum_{m+1}^p \frac{1}{a_r} \\ (m = m(p) \geq 1).$$

These inequalities, in connection with (9.13), (9.4), (9.12) and (9.14), show that the second factor on the right side of (9.14) tends to 1 for $p \rightarrow \infty$. Hence from (9.10')

$$(9.15) \quad \lim_{p \rightarrow \infty} \exp \left\{ -a_1 \sum_{m+1}^p \frac{1}{a_r} \right\} = i.$$

For p sufficiently large we have $m(p) \geq n$ and for such values of p

$$\prod_{m+1}^p \left(1 - \frac{a_n}{a_r}\right) = \exp \left\{ -a_n \sum_{m+1}^p \frac{1}{a_r} \right\} \cdot \exp \left\{ -\frac{a_n^2}{2} \sum_{m+1}^p \frac{1}{a_r^2} \left(1 - \theta'_r \frac{a_n}{a_r}\right)^{-2} \right\}.$$

A similar argument as above and (9.15) show that

$$\lim_{p \rightarrow \infty} \prod_{m+1}^p \left(1 - \frac{a_n}{a_r}\right) = \lim_{p \rightarrow \infty} \exp \left\{ -a_n \sum_{m+1}^p \frac{1}{a_r} \right\} \\ = \lim_{p \rightarrow \infty} \left[\exp \left\{ -a_1 \sum_{m+1}^p \frac{1}{a_r} \right\} \right]^{a_n/a_1} = i^{a_n/a_1},$$

and (9.10) is proved.

The condition for the uniqueness of the non-decreasing function $\chi(t)$ in (8.22) is insured by a theorem of Müntz* and by our assumption (9.5).

We get again the special system (5.5) and its solution (5.4) by taking $a_n = n$ ($n = 0, 1, 2, \dots$). For these values the following identities hold:

$$D^k x_l = 1!2!3! \cdots (k-1)! \Delta^k x_l \quad (k, l = 0, 1, 2, \dots),$$

which show that the systems (5.5) and (9.1) are equivalent. The result (5.4) is then a special case of (8.20) for $\phi_n(t) = t^n$. The system (5.5) appeared twice as a special case of (6.5), first, by replacing the general matrix (6.1) by the matrix of the binomial coefficients (6.6), and second, as a Hausdorff system for $a_n = n$ ($n = 0, 1, 2, \dots$).

Returning to the more general case $a_0 \geq 0$, (8.20) and (9.6) give

$$x_n = \int_0^1 t^{(a_n - a_0) / (a_1 - a_0)} d\chi(t) \quad (n = 0, 1, 2, \dots).$$

Replacing in this integral the variable t by $t^{a_1 - a_0}$ and writing $\chi(t^{a_1 - a_0}) = \phi(t)$, we get

$$(9.16) \quad x_n = \int_0^1 t^{a_n - a_0} d\phi(t) \quad (n = 0, 1, 2, \dots),$$

the non-decreasing function $\phi(t)$, with $\phi(0) = 0$, being uniquely defined by (9.16) in all its points of continuity.

Let us now consider the function $\psi(t)$ defined by

$$(9.17) \quad \psi(t) = - \int_t^1 t^{-a_0} d\phi(t) \quad \text{for } 0 < t \leq 1.$$

This function $\psi(t)$ is non-decreasing on $0 < t \leq 1$, $\psi(1) = 0$, but not necessarily bounded for $t \rightarrow 0$. Moreover, $\psi(t)$ is uniquely defined by (9.17) and (9.16) in all the points of continuity of $\phi(t)$. Let us consider the following convergent improper Stieltjes integral:

$$(9.18) \quad \begin{aligned} \int_0^1 t^{a_n} d\psi(t) &= \lim_{\epsilon \rightarrow 0} \int_{\epsilon}^1 t^{a_n} d\psi(t) = \lim_{\epsilon \rightarrow 0} \int_{\epsilon}^1 t^{a_n - a_0} d\phi(t) \\ &= - [t^{a_n - a_0}]_{t=0} \cdot [\phi(+0) - \phi(0)] + \int_0^1 t^{a_n - a_0} d\phi(t). \end{aligned}$$

The first term of the last member of (9.18) is 0 for $n > 0$, and ≤ 0 for $n = 0$.

* Ch. H. Müntz, *Über den Approximationssatz von Weierstrass*, Mathematische Abhandlungen H. A. Schwarz gewidmet, Berlin, 1914, pp. 303-312. The theorem says that if $0 < p_1 < p_2 < p_3 < \dots$ and $\sum_{n=1}^{\infty} 1/p_n = \infty$, then every continuous function on $0 \leq t \leq 1$ is uniformly approximable as close as we want by linear combinations of functions of the sequence $1, t^{p_1}, t^{p_2}, t^{p_3}, \dots$.

From (9.16) and (9.18) we therefore get

$$(9.19) \quad \begin{aligned} x_0 &= C + \int_0^1 t^{a_0} d\psi(t), \\ x_n &= \int_0^1 t^{a_n} d\psi(t) \end{aligned} \quad (n = 1, 2, 3, \dots; C \geq 0),$$

where the integrals have to be taken as $\lim_{\epsilon \rightarrow 0} \int_{\epsilon}^1$.

Conversely, let $\psi(t)$ be a function which is non-decreasing on $0 < t \leq 1$, with $\psi(1) = 0$, not necessarily bounded, but such that the first integral in (9.19) is convergent and $C \geq 0$. Let us define the function $\phi(t)$ by

$$(9.20) \quad \phi(0) = 0, \quad \phi(t) = C + \int_0^t t^{a_0} d\psi(t) \quad \text{for } 0 < t \leq 1.$$

This function $\phi(t)$ is non-decreasing on $0 \leq t \leq 1$. From (9.19), (9.20) and $\psi(1) = 0$, we may derive back (9.17) and (9.16). The function $\phi(t)$ is uniquely defined by (9.16) and $\phi(0) = 0$ in all its points of continuity, therefore also $\psi(t)$ is uniquely defined by (9.19) and $\psi(1) = 0$ in all its points of continuity. This proves the following

COROLLARY 9.1. *Every solution x_0, x_1, x_2, \dots of the system (9.1) with (9.2), (9.3), (9.4) and (9.5), may be expressed in the form (9.19), where $C \geq 0$, while $\psi(t)$ is non-decreasing on $0 < t \leq 1$ with $\psi(1) = 0$, and such that the first and hence all the integrals in (9.19) are convergent.*

Conversely, for every such C and $\psi(t)$, the sequence x_0, x_1, x_2, \dots given by (9.19) is a solution of the system (9.1). Both C and $\psi(t)$ are uniquely defined by the set (9.19) and the additional conditions $\psi(1) = 0$, $2\psi(t) = \psi(t+0) + \psi(t-0)$ for $0 < t < 1$.

This form of Hausdorff's Theorem 9.1 will be used in §11.

10. The solution of Hausdorff's system (9.1) when either (9.4) or (9.5) does not hold. Theorem 9.1 is based on the assumptions (9.3), (9.4) and (9.5). We shall now replace first the assumption (9.4) by

$$(10.1) \quad \lim_{r \rightarrow \infty} a_r = \alpha (< \infty),$$

and solve the system (9.1) in this new situation. For $a_0 > 0$, formula (9.8) becomes

$$(10.2) \quad c_{n,m,p} = \prod_{r=m+1}^p \left(1 - \frac{a_n}{a_r}\right) \left(1 - \frac{a_0}{a_r}\right)^{-1} \quad (c_{n,p,p} = 1).$$

From (8.10), (10.2) and (10.1) we get

$$\begin{aligned} \lim_{p \rightarrow \infty} t_{p, p-q} &= \left(1 - \frac{a_1}{\alpha}\right)^q \left(1 - \frac{a_0}{\alpha}\right)^{-q} \\ (10.3) \qquad &= \left(\frac{\alpha - a_1}{\alpha - a_0}\right)^q, \end{aligned}$$

$$\lim_{p \rightarrow \infty} c_{n, p-q, p} = \left(\frac{\alpha - a_n}{\alpha - a_0}\right)^q \quad (q = 0, 1, 2, \dots; n = 1, 2, 3, \dots).$$

These relations show that the polygonal line $P_n^{(p)}$ ($n \geq 1$) tends, for $p \rightarrow \infty$, to a polygonal line whose vertices are the origin $(0, 0)$ and the sequence of points

$$(t, u_n) = \left(\left(\frac{\alpha - a_1}{\alpha - a_0}\right)^q, \left(\frac{\alpha - a_n}{\alpha - a_0}\right)^q \right) \quad (q = 0, 1, 2, \dots),$$

which tend to the origin, while $P_0^{(p)}$ is as usual the segment joining the points $(0, 1)$ and $(1, 1)$. These limiting polygonal lines are the graphs of the functions $\phi_n(t)$ of Theorem 8.1. The arguments which proved the statement (8.27) show that the non-decreasing function $\chi(t)$ may always be replaced in this special case by a non-decreasing step-function which jumps only at the points

$$t = 0 \text{ and } t = \left(\frac{\alpha - a_1}{\alpha - a_0}\right)^q \quad (q = 0, 1, 2, \dots);$$

let γ and λ_q denote the amounts of the respective jumps. The integrals (8.20) become

$$\begin{aligned} x_0 &= \gamma + \lambda_0 + \lambda_1 + \lambda_2 + \dots & (\gamma \geq 0, \lambda_q \geq 0), \\ (10.4) \quad x_n &= \lambda_0 + \lambda_1 \left(\frac{\alpha - a_n}{\alpha - a_0}\right) + \lambda_2 \left(\frac{\alpha - a_n}{\alpha - a_0}\right)^2 + \dots & (n = 1, 2, 3, \dots). \end{aligned}$$

The substitution $\lambda_q = (\alpha - a_0)^q \mu_q$ completes the proof of the following

THEOREM 10.1. *Every solution of the system (9.1) derived from the Vandermondean matrix (9.2), whose elements satisfy the conditions (9.3) and (10.1), may be written in the form*

$$\begin{aligned} (10.5) \quad x_0 &= \gamma + \mu_0 + \mu_1(\alpha - a_0) + \mu_2(\alpha - a_0)^2 + \dots & (\gamma, \mu_0, \mu_1, \mu_2, \dots \geq 0), \\ x_n &= \mu_0 + \mu_1(\alpha - a_n) + \mu_2(\alpha - a_n)^2 + \dots & (n = 1, 2, 3, \dots). \end{aligned}$$

Conversely, every sequence of non-negative numbers $\gamma, \mu_0, \mu_1, \mu_2, \dots$, such that the first series (10.5) is convergent, defines a sequence x_0, x_1, x_2, \dots by means of the set (10.5), which is a solution of the system of linear inequalities (9.1). The coefficients $\gamma, \mu_0, \mu_1, \mu_2, \dots$ are uniquely defined by the set (10.5).

The last statement concerning the uniqueness of the coefficients follows readily from a fundamental property of power series. The connection of Theorem 10.1 with the theory of completely monotonic functions will be discussed in the next section.

Let us now replace the assumption (9.5) by the assumption

$$(10.6) \quad \sum_{r=1}^{\infty} \frac{1}{a_r} \text{ is convergent.}$$

This of course implies (9.4). From (10.2) and (10.6) we get

$$(10.7) \quad \lim_{p \rightarrow \infty} c_{n,m,p} = b_{n,m} = \prod_{r=m+1}^{\infty} \left(1 - \frac{a_n}{a_r}\right) \bigg/ \prod_{r=m+1}^{\infty} \left(1 - \frac{a_0}{a_r}\right).$$

In particular, for $n=1$,

$$(10.8) \quad \lim_{p \rightarrow \infty} t_{p,m} = b_{1,m} = \prod_{r=m+1}^{\infty} \left(1 - \frac{a_1}{a_r}\right) \bigg/ \prod_{r=m+1}^{\infty} \left(1 - \frac{a_0}{a_r}\right).$$

Formula (10.7) shows that

$$(10.9) \quad b_{0,m} = 1, \quad b_{n,m} = 0 \text{ for } n > m.$$

The limiting function $\phi_n(t)$ of Theorem 8.1 is therefore represented graphically by the polygonal line whose vertices have the coordinates

$$(b_{1,m}, b_{n,m}) \quad (m = 0, 1, 2, \dots),$$

to which sequence converging towards $(1, 1)$ we have to adjoin this limiting point itself.

Just as in the previous case, we may replace in the system (8.20) the non-decreasing function $\chi(t)$ by a non-decreasing step-function which jumps only at the points $t=b_{1,m}$ ($m=0, 1, 2, \dots$) and $t=1$; let λ_m and γ denote the amounts of the respective jumps. The integrals (8.20) become

$$(10.10) \quad x_n = b_{n,0}\lambda_0 + b_{n,1}\lambda_1 + b_{n,2}\lambda_2 + \dots + \gamma \quad (n = 0, 1, 2, \dots) \\ (\gamma, \lambda_0, \lambda_1, \lambda_2, \dots \geq 0).$$

Let

$$(10.11) \quad c_{n,m} = \prod_{r=m+1}^{\infty} \left(1 - \frac{a_n}{a_r}\right), \quad \lambda_m = c_{0,m}\mu_m.$$

From (10.10) and (10.11) we get the following

THEOREM 10.2. *Every solution of the system (9.1) derived from the Vandermondean matrix (9.2), whose elements satisfy the conditions (9.3), and (10.6), may be written in the form*

$$\begin{aligned}
 x_0 &= \gamma + c_{0,0}\mu_0 + c_{0,1}\mu_1 + c_{0,2}\mu_2 + \cdots, \\
 x_1 &= \gamma + c_{1,1}\mu_1 + c_{1,2}\mu_2 + \cdots, \\
 x_2 &= \gamma + c_{2,2}\mu_2 + \cdots, \\
 &\vdots \\
 &\vdots
 \end{aligned}
 \tag{10.12}$$

$$(\gamma, \mu_0, \mu_1, \cdots \geq 0),$$

the coefficients being defined by (10.11). Conversely, every sequence x_0, x_1, x_2, \cdots given by (10.12), where all the infinite series converge if the first one converges, is a solution of the system (9.1).

I mention without proof that the set (10.12) defines uniquely the coefficients γ and μ_m there involved. In particular $\gamma = \lim x_n$ for $n \rightarrow \infty$.

11. Applications to the theory of completely monotonic functions. A function $f(x)$ is called *completely monotonic* on an open or closed, finite or infinite, interval I , if

$$(11.1) \quad (-1)^n f^{(n)}(x) \geq 0 \quad (n = 0, 1, 2, \cdots), \text{ for every } x \text{ interior to } I,$$

which inequalities imply the existence of all derivatives involved, while only the continuity of $f(x)$ is required at the end points of a closed interval. We shall prove the following fundamental

THEOREM 11.1 (F. Hausdorff, S. Bernstein, D. V. Widder). *A necessary and sufficient condition that $f(x)$ should be completely monotonic in the interval $c < x < \infty$ is that*

$$(11.2) \quad f(x) = \int_0^\infty e^{-xt} d\alpha(t),$$

where $\alpha(t)$ is a non-decreasing function of such a nature that the integral converges for $x > c$. The non-decreasing function $\alpha(t)$ is uniquely defined in all its points of continuity by (11.2) and the additional condition $\alpha(0) = 0$.

The sufficiency of the condition is obvious since

$$f^n(x) = (-1)^n \int_0^\infty e^{-xt} t^n d\alpha(t) \quad (x > c; n = 0, 1, 2, \cdots).$$

To prove the necessity, without any loss of generality we may suppose $c=0$, that is to say, the function $f(x)$ completely monotonic for $x > 0$. For $0 < x_0 < x_1 < x_2 < \cdots < x_n$ ($n=0, 1, 2, \cdots$), a mean-value theorem of Schwarz and Stieltjes* gives

* H. A. Schwarz, *Gesammelte Mathematische Abhandlungen*, vol. 2, p. 296, Berlin, 1890. T. J. Stieltjes, *Oeuvres*, vol. 2, p. 105 and p. 110.

$$\begin{aligned}
 (11.3) \quad & \begin{vmatrix} f(x_0) & 1 & x_0 & x_0^2 & \cdots & x_0^{n-1} \\ f(x_1) & 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ f(x_n) & 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix} \\
 & = (-1)^n \frac{f^{(n)}(\theta)}{n!} \begin{vmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix} \geq 0 \quad (\theta > 0).
 \end{aligned}$$

This shows that for $0 < a_0 < a_1 < a_2 < \cdots$ the sequence $x_n = f(a_n)$ ($n = 0, 1, 2, \cdots$) is a solution of the Hausdorff system (9.1). Hence Corollary 9.1 gives

$$\begin{aligned}
 (11.4) \quad & f(a_0) = C + \int_0^1 t^{a_0} d\psi(t), \quad f(a_n) = \int_0^1 t^{a_n} d\psi(t) \\
 & (n = 1, 2, 3, \cdots; C \geq 0, \psi(1) = 0),
 \end{aligned}$$

where the integrals have to be taken as $\lim_{\epsilon \rightarrow 0} \int_{\epsilon}^1$. On the other hand it is obvious from the results of §9 that this function $\psi(t)$ is also uniquely defined in all its points of continuity by the set of equations derived from (11.4) by leaving out one of its equations. We may therefore vary continuously any one of the elements a_0, a_1, a_2, \cdots without changing $\psi(t)$. This proves that

$$(11.5) \quad f(x) = \int_0^1 t^x d\psi(t) = \lim_{\epsilon \rightarrow 0} \int_{\epsilon}^1 t^x d\psi(t)$$

holds for every positive value of x . The substitutions

$$t = e^{-\tau}, \quad \alpha(t) = -\psi(e^{-\tau})$$

transform (11.5) into

$$f(x) = \int_0^{\infty} e^{-x\tau} d\alpha(\tau) \quad \text{for } x > 0,$$

and Theorem 11.1 is proved.

Just as Corollary 9.1 has led to a proof of Theorem 11.1, similarly Theorem 10.1 will readily prove the following

THEOREM 11.2 (S. Bernstein).* *A function $f(x)$ which is completely monotonic on the finite interval $c < x < \alpha$, is necessarily regular and analytic on $c < x < \alpha + (\alpha - c)$, and hence*

* S. Bernstein, *Leçons sur les Propriétés Extrêmes des Fonctions Analytiques*, Paris, 1926, p. 190. The definitions there used are somewhat different from those given in this paper but are essentially equivalent. Our definition of a completely monotonic function has been taken from Widder's paper.

$$(11.6) \quad f(x) = \mu_0 + \mu_1(\alpha - x) + \mu_2(\alpha - x)^2 + \cdots \text{ for } c < x < 2\alpha - c$$

$$(\mu_0, \mu_1, \mu_2, \cdots \geq 0).$$

Without any loss of generality we may suppose $c=0$. For any sequence a_n with

$$0 < a_0 < a_1 < a_2 < \cdots < a_n \rightarrow \alpha,$$

the corresponding sequence $x_n = f(a_n)$ is a solution of the Hausdorff system (9.1) (a consequence of the general inequality (11.3)). Hence this sequence x_n may be represented in the form (10.5). From fundamental properties of power series we infer that the coefficients μ_n do not change if the elements a_n are varied continuously. An immediate consequence is the relation (11.6) for $0 < x < \alpha$. The analytic function $f(x)$ is therefore obviously regular for $0 < x < 2\alpha$.

By means of Theorems 11.1 and 11.2 we may combine the results of Theorems 9.1 and 10.1 in one single statement as

COROLLARY 11.1. *Let (9.1) be a Hausdorff system defined by the matrix (9.2) whose elements a_n satisfy the conditions (9.3) and (9.5). Let $\lim_{n \rightarrow \infty} a_n = \alpha$ be finite or infinite. A necessary and sufficient condition that the sequence x_n should be a solution of the Hausdorff system (9.1) is*

$$(11.7) \quad x_0 = C + f(a_0), \quad x_n = f(a_n) \quad (n = 1, 2, 3, \cdots; C \geq 0),$$

where the function $f(x)$ is completely monotonic on the interval $a_0 \leq x < \alpha$. The function $f(x)$ and the constant C are uniquely defined by the set (11.7).

For the case when (9.5) is replaced by (10.6), the general solution of (9.1) is given by (10.12), as stated in Theorem 10.2. Any sequence x_0, x_1, x_2, \cdots of the type (11.7) of course still represents a solution of (9.1), but the converse is not true that every such solution is of the type (11.7). This is readily shown by taking for example the sequence $x_0 = c_{0,1}, x_1 = c_{1,1}, x_2 = x_3 = \cdots = 0$, which according to (10.12) is a solution of (9.1) (for $\gamma = \mu_0 = \mu_2 = \mu_3 = \cdots = 0, \mu_1 = 1$). This solution admits no representation of the type (11.7).

UNIVERSITY OF CHICAGO,
CHICAGO, ILL.