ON SYSTEMS OF LINEAR DIFFERENTIO-STIELTJES-INTEGRAL EQUATIONS

BY

T. H. HILDEBRANDT

It is well known that the solution of the initial condition problem for a system of linear differential equations $dy_i/dx = \sum_{j=1}^n a_{ij}(x)y_j(x)$, $i = 1, 2, \cdots, n$, with initial conditions $y_i(a) = c_i$ is equivalent to the solution of the system of linear integral equations $y_i(x) = \int_a^x \sum_j a_{ij}(s)y_j(s)\,ds + c_i$, and that most of the properties of the solutions of such systems of differential equations are deducible from the integral equation equivalent. (See Birkhoff-Langer [2], pp. 51–60.) Generalizations of the differential system are simpler in the integral form; e.g., in the integral form, one can assume the $a_{ij}(x)$ to be Lebesgue-integrable, and the solutions are then absolutely continuous functions. Recently H. S. Wall [14] and J. S. MacNerney [5] by assuming that the functions $a_{ij}(x)$ are continuous and of bounded variation, and by using a Riemann-Stieltjes (R-S) integral, have shown that the solution of the system of linear Stieltjes integral equations

$$y_i(x) = \int_a^x \sum_j da_{ij}(s)y_j(s) + y_i(a)$$

as well as that of the corresponding nonhomogeneous system

$$y_i(x) = \int_a^x \sum_j da_{ij}(s)y_j(s) + u_i(x) - u_i(a)$$

parallels in many ways the differential case. The assumption that the $a_{ij}(x)$ be allowed discontinuities, however, introduces difficulties, since the solutions of the initial value problem would have discontinuities at the same points as $a_{ij}(x)$ so that the R-S integral as well as its properties may no longer apply.

The purpose of this paper is to indicate what changes in the theory are necessitated by dropping the continuity requirement. In this theory, the R-S integral is replaced by a form of the Lebesgue-Stieltjes (L-S) integral as applied to functions of bounded variation with respect to functions of bounded variation. While some of the properties valid in the case when $a_{ij}(x)$ are continuous do not carry over, modification of others leads to comparatively elegant results. It turns out that a form of the product integral, a modification of the product integral used effectively by L. Schlesinger ([10] and [11]) in connection with systems of linear differential equations, plays an important role in the form of the solutions for the case under consideration.

The desirability of considering the generalization treated here is an outgrowth of correspondence some time back with W. H. Ingham, who attempted

Received March 18, 1958.

to develop an integral and a theory which would handle this case. Unfortunately at the time, the attempt to effect a rigorous treatment of Ingham's approach did not succeed. The present development was carried through some time after the close of correspondence and is not related to nor dependent on the theory proposed by Ingham. However, we feel indebted to him for having insisted on calling our attention to the desirability of studying the extensions treated here.

1. Vectors and matrices

We shall assume that our basis is an n-dimensional real vector space, and for purposes of notation shall designate such vectors by capital letters at the end of the alphabet, e.g., $Y = (y_1, \cdots, y_n)$. When a norm is needed we shall assume that $\| Y \| = \max_i | y_i |$. $n \times n$ matrices will be denoted by letters at the beginning of the alphabet, e.g., $A = [a_{ij}]$, $i, j = 1, \cdots, n$. Matrices might be normed by assuming $\| A \| = \max_{ij} | a_{ij} |$, but this norm does not possess the property $\| AB \| \leqq \| A \| \cdot \| B \|$, desirable in considering matrices as constituting a normed ring. So we shall assume $\| A \| = \max_i \sum_j | a_{ij} |$, which is related to the fact that matrices provide a linear transformation on n-vectors to n-vectors normed as above. Obviously for this norm, if I is the identity matrix, then $\| I \| = 1$, and $\| AB \| \leqq \| A \| \cdot \| B \|$. Also $\lim_m \| A_m - A \| = 0$ is equivalent here to $\lim_m a_{ijm} = a_{ij}$ for each i, j. Other norms for Y and A could have been used, keeping in mind the parallels: Y belonging to a linear normed complete vector space $\mathfrak{Y}$, and A being a linear continuous transformation on $\mathfrak{Y}$ to $\mathfrak{Y}$.

2. Functions

Our considerations will be limited to real valued functions of bounded variation on the closed interval $a \leqq x \leqq b$. We denote the total variation on (a, x) of a function $\alpha(x)$ of bounded variation by $V_a^x \alpha$. However if $A(x)$ is a matrix of functions of bounded variation, we shall define

$$\begin{aligned} V_a^b A &= \text{l.u.b.}_\sigma \textstyle\sum_k \| A(x_k) - A(x_{k-1}) \| \\ &= \text{l.u.b.}_\sigma \textstyle\sum_k \max_i \sum_j | a_{ij}(x_k) - a_{ij}(x_{k-1}) | , \end{aligned}$$

where the least upper bound is taken with respect to all subdivisions $\sigma \equiv \{a = x_0 < x_1 < x_2 < \cdots < x_n = b\}$ of (a, b). It is obvious that $V_c^d a_{ij} \leqq V_c^d A$ for all subintervals (c, d) of (a, b); also that the function $V_a^x A$ is discontinuous if and only if some a_{ij} is discontinuous; and that

$$\| A(x + 0) - A(x - 0) \| \leqq V_{x-0}^{x+0} A,$$

so that $\sum_x \| A(x + 0) - A(x - 0) \| < \infty$. Here $\sum_x a(x)$ means the limit of $\sum_\sigma a(x_i)$ where σ is any finite subset of (a, b) and the σ are directed by inclusion. It is well known (E. H. Moore [8], pp. 61–67) that $\sum_x a(x)$ exists if and only if $a(x)$ is zero except for a denumerable number of x: x_n and $\sum_n | a(x_n) | < \infty$.

3. Integrals

The Stieltjes integral we shall use is a modification of the R-S integral. We write $g(x) = g_c(x) + g_b(x)$, where $g_c(x)$ is the continuous part of $g(x)$ and $g_b(x)$ is the function of the breaks, i.e.,

$$g_b(x) = g(a+0) - g(a) + \textstyle\sum_{a<y<x} (g(y+0) - g(y-0)) + g(x) - g(x-0).$$

If $\int_a^b f\, dg_c$ exists as an R-S integral, then for any closed interval (c, d) of (a, b) we define

$$\int_c^d f(x)\, dg(x) = \int_c^d f(x)\, dg_c(x) + f(c)(g(c+0) - g(c)) + \sum_{c<x<d} f(x)(g(x+0) - g(x-0)) + f(d)(g(d) - g(d-0)).$$

For convenience we shall assume that the symbol

$$\sum_{c \leqq x \leqq d} f(x)(g(x+0) - g(x-0))$$

includes the last three terms in this expression. It is possible to show that the integral in question can be obtained as the limit by successive subdivisions of a Riemann sum suggested by W. H. Young [13] (see Hildebrandt [4], p. 275), viz.,

$$\int_c^d fdg = \lim_\sigma \textstyle\sum_i \{f(x_{i-1})(g(x_{i-1}+0) - g(x_{i-1})) + f(x_i')(g(x_i - 0) - g(x_{i-1}+0)) + f(x_i)(g(x_i) - g(x_i - 0))\}$$

where $\sigma \equiv \{c = x_0 < x_1 < \cdots < x_n = d\}$ and $x_{i-1} < x_i' < x_i$. For the closed interval (a, b) it agrees with the Lebesgue-Stieltjes (L-S) integral. For any closed subinterval (c, d) of (a, b), however, we have

$$\text{L-S} \int_c^d fdg = \int_{c-0}^{d+0} fdg$$

(see (b) below), with obvious adjustments if $c = a$ or $d = b$.

We note the following properties of this integral:

(a) $\int_a^b fdg$ is a bilinear functional on the space BV $\times$ BV.

(b) If $h(x) = \int_a^x fdg$, then $h(x)$ is of bounded variation, and continuous at all points of continuity of g. If x_0 is a point of discontinuity of $g(x)$, then $h(x_0) - h(x_0 - 0) = f(x_0)(g(x_0) - g(x_0 - 0))$, with similar expressions for $h(x_0 + 0) - h(x_0)$ and $h(x_0+0) - h(x_0 - 0)$.

(c) A *substitution theorem* is valid, viz., if $k(x) = \int_a^x gdh$, then $\int_a^b fdk(x) = \int_a^b fgdh$. This results immediately from the integral definition by using the substitution theorem for R-S integrals and item (b).

(d) The following *convergence theorem* is adequate for the purposes of this paper: If $\lim_n f_n(x) = f(x)$ for all x on (a, b), the $f_n(x)$ are uniformly bounded on (a, b), and $g_m(x)$ converges to $g(x)$ in the sense $\lim_m V_a^b(g_m - g) = 0$, then $\lim_{m,n} \int_a^b f_n\, dg_m = \int_a^b f dg$. This follows immediately from the identity $\int f_n\, dg_m - \int f dg = \int f_n\, d(g_m - g) + \int (f_n - f)\, dg$, the inequality

$$\left| \int_a^b f dg \right| \leqq \text{l.u.b.}_x | f(x) | V_a^b g,$$

and the convergence theorem for L-S integrals.

(e) The integration by parts theorem must be modified. We have (Hildebrandt [4], p. 276)

$$\int_a^b f dg + \int_a^b g df = fg \Big|_a^b - \sum_{a \leqq x \leqq b} [(f(x + 0) - f(x))(g(x + 0) - g(x)) + (f(x) - f(x - 0))(g(x) - g(x - 0))].$$

Since

$$\int_a^b f dg + \int_a^b g df - fg \Big|_a^b = \int_a^b df(x) \int_a^x dg(y) - \int_a^b \left(\int_y^b df(x) \right) dg(y),$$

we can rewrite the integration by parts formula as follows:

$$\int_a^b df(x) \int_a^x dg(y) = \int_a^b \left(\int_y^b df(x) \right) dg(y) - \sum_{a \leqq x \leqq b} [(f(x + 0) - f(x))(g(x + 0) - g(x)) - (f(x) - f(x - 0))(g(x) - g(x - 0))],$$

which makes it a special case of the Dirichlet formula considered below.

Notation. Since the differences $f(x + 0) - f(x)$ and $f(x) - f(x - 0)$ will be of frequent occurrence in the sequel, we shall abbreviate them by $\Delta^+ f(x)$ and $\Delta^- f(x)$, respectively, and set $\Delta^{\pm} f(x) = f(x + 0) - f(x - 0)$.

(f) *Dirichlet formula.* If $h(x, y)$ is bounded on $a \leqq x \leqq b$, $a \leqq y \leqq b$ and of bounded variation in y for each x and in x for each y, then

$$\int_a^b df(x) \int_a^x h(x, y)\, dg(y) = \int_a^b \left(\int_y^b df(x) h(x, y) \right) dg(x) - \sum_{a \leqq x \leqq b} [\Delta^+ f(x) h(x, x) \Delta^+ g(x) - \Delta^- f(x) h(x, x) \Delta^- g(x)].$$

This theorem can be proved by breaking up each of the functions $f(x)$ and $g(y)$ into their continuous and discontinuous parts, giving rise to four parts for each of the iterated integrals. The three parts involving continuous parts of f or g are equal in pairs; the integrals for the purely discontinuous parts of f and g give rise to the sum term. A less cumbersome procedure can be based on the following lemma on iterated integrals, the proof of which can be made by an adaptation of the reasoning used by H. J. Ettlinger [3], p. 65, for the case when $g(x) = h(x) = x$, and the integrals are Riemann integrals:

LEMMA. *If $h(x, y)$ is bounded on the rectangle $a \leqq x \leqq b$, $c \leqq y \leqq d$; if $f(x)$ is of bounded variation on $a \leqq x \leqq b$, and $g(y)$ of bounded variation on $c \leqq y \leqq d$; and if $\int_a^b h(x, y)\, df(x)$ exists for every y, and $\int_c^d h(x, y)\, dg(y)$ exists for every x; then $\int_a^b df(x) \int_c^d h(x, y)\, dg(y)$ and $\int_c^d \left(\int_a^b df(x) h(x, y)\right) dg(y)$ both exist and are equal.*

All integrals are to be taken in the sense of this section.

If $h(x, y)$ is bounded on $a \leqq x \leqq b$, $a \leqq y \leqq b$ and of bounded variation in x for each y and in y for each x, and if we set $H(x, y) = h(x, y)$ on $a \leqq y \leqq x \leqq b$ and zero on $a \leqq x < y \leqq b$, then $H(x, y)$ satisfies the hypotheses of the lemma, and its iterated integrals are equal. Now

$$\int_a^b H(x, y)\, dg(y) = \text{L-S} \int_a^x h(x, y)\, dg(y)$$

$$= \int_a^x h(x, y)\, dg(y) + h(x, x)\Delta^+ g(x),$$

and similarly

$$\int_a^b df(x) H(x, y) = \int_y^b df(x) h(x, y) + \Delta^- f(y) h(y, y).$$

Consequently

$$\int_a^b df(x) \int_a^x h(x, y)\, dg(y) + \int_a^b df(x) h(x, x) \Delta^+ g(x)$$
$$= \int_a^b \left(\int_y^b df(x) h(x, y)\right) dg(y) + \int_a^b \Delta^- f(y) h(y, y)\ dg(y).$$

Since $h(x, x)\Delta^+ g(x)$ and $\Delta^- f(y) h(y, y)$ vanish excepting at the points of discontinuity of g and f, respectively, we have

$$\int_a^b df(x) \int_a^x h(x, y)\, dg(y) + \sum_{a \leqq x \leqq b} \Delta^{\pm} f(x) h(x, x) \Delta^+ g(x)$$
$$= \int_a^b \left(\int_y^b df(x) h(x, y)\right) dg(y) + \sum_{a \leqq x \leqq b} \Delta^- f(x) h(x, x) \Delta^{\pm} g(x),$$

or

$$\int_a^b df(x) \int_a^x h(x, y)\, dg(y) = \int_a^b \left(\int_y^b df(x) h(x, y)\right) dg(y)$$
$$+ \sum_{a \leqq x \leqq b} [\Delta^- f(x) h(x, x) \Delta^- g(x) - \Delta^+ f(x) h(x, x) \Delta^+ g(x)],$$

which is the Dirichlet formula desired. Obviously the formula in its simple form holds if the sum term vanishes, which occurs, for instance, if f and g have no common discontinuities on the same side of any point of (a, b). As noted above, setting $h(x, y) \equiv 1$ on $a \leqq x \leqq b$, $a \leqq y \leqq b$ yields the integration by parts formula.

The theorems of this section have been stated for the case of single functions. Their validity can be extended to the case where matrix or vector functions are involved, if proper attention is paid to order in the products.

4. The matrix $A(x)$ continuous

For future reference and comparison, it seems desirable to collect the basic results relative to the solutions of the systems

$$\text{(I)} \qquad Y(x) = \int_a^x dA(s)Y(s) + Y(a)$$

$$\text{(II)} \qquad Y(x) = \int_a^x dA(s)Y(s) + U(x) \quad \text{with} \quad Y(a) = U(a),$$

for the case where the matrix $A(x)$ is continuous (Wall [14], pp. 160–163; MacNerney [5], pp. 354–362). We have

(a) The system (I) has a unique solution valid on (a, b) expressible in the form $Y(x) = B(a, x)Y(a)$.

(b) The matrix $B(a, x)$, the result of successive substitutions applied to (I), is expressible as the Peano [9] series:

$$B(a, x) = I + \int_a^x dA(s) + \int_a^x dA(s_1) \int_a^{s_1} dA(s_2) + \cdots + \int_a^x dA(s_1) \int_a^{s_1} dA(s_2) \cdots \int_a^{s_n} dA(s_{n+1}) + \cdots,$$

the convergence being uniform on (a, b). As a matter of fact

$$\| B(a, x) \| \leqq \exp V_a^b A,$$

the inequality being actually term by term. The proof of this uses the fact that if $v(x)$ is continuous, then $\int_a^b v^n dv = (v^{n+1}(b) - v^{n+1}(a))/(n + 1)$, which in turn depends on the integration by parts theorem for R-S integrals.

(c) The matrix B can be defined for any two points x', x'' of (a, b) in either order. As a function of the upper limit it satisfies the integral equation:

$$\text{(I}'\text{)} \qquad B(x', x'') = \int_{x'}^{x''} dA(s)B(x', s) + I,$$

so that $B(a, x)$ is a fundamental system of solutions of (I).

(d) For any x, $B(x, x) = I$, and for any three points x', x'', x''' of (a, b) we have $B(x'', x''')B(x', x'') = B(x', x''')$. Then $B(x', x'')$ and $B(x'', x')$ are reciprocals of each other. Matrix functions possessing these properties are called *harmonic functions* by H. S. Wall ([14], p. 160).

(e) The matrix A is expressible in terms of such a $B(x', x'')$ via the relation $A(x) - A(a) = \int_a^x d_s B(x', s)B(s, x')$, the latter expression being independent of x'.

(f) The matrix $B(x', x'')$ is also expressible as a product integral (Schlesinger [11]; MacNerney [5], p. 362), viz., if

$$\sigma \equiv \{x' < x_1 < x_2 < \cdots < x_n < x''\},$$

then

$$B(x', x'') = \widehat{\int_{x'}^{x''}} (I + dA) = \lim (I + A(x'') - A(x_n))$$
$$\cdot (I + A(x_n) - A(x_{n-1})) \cdots (I + A(x_1) - A(x')),$$

the limit being taken as $|\sigma| = \max |x_i - x_{i-1}|$ approaches zero. Consequently the limit exists also as a directed limit in σ ordered by inclusion.

(g) As a function of the lower limit x', $B(x', x'')$ satisfies the system of integral equations

$$B(x', x'') = \int_{x'}^{x''} B(s, x'')\, dA(s) + I,$$

which in differential equations is equivalent to the statement that $B(x', x'')$ in x' is a solution of the adjoint differential equation $dZ/dx = -ZA$. This property can be proved by applying the Dirichlet formula for interchange of order of integration to the Peano series.

(h) The solution of the nonhomogeneous system (II) is expressible in the form

$$Y(x) = B(a, x)U(a) + \int_a^x B(s, x)\, dU(s),$$

$U(x)$ being assumed to be of bounded variation (discontinuities allowed).

(i) The determinant of $B(a, x)$ satisfies the integral equation

$$\det B(a, x) = \int_a^x d\left(\sum_{i=1}^n (a_{ii}(s) - a_{ii}(a))\right) \det B(a, s) + 1.$$

This can be shown by applying a multiple integration by parts theorem and the substitution theorem for R-S integrals.

5. The matrix $A(x)$ has a finite number of discontinuities on (a, b)

In order to see what effect discontinuities of $A(x)$ have on the solution of equation (I), we consider first the case where $A(x)$ has a single discontinuity at x_0. Since for $a \leqq x < x_0$, $A(x)$ is continuous, it follows that for this interval we have

$$Y(x) = \int_a^x dA(s)Y(s) + Y(a) \quad \text{is equivalent to} \quad Y(x) = B(a, x)Y(a).$$

At x_0, we have

$$Y(x_0) = \int_a^{x_0} dA(s)Y(s) + Y(a) = \int_a^{x_0-0} dA \cdot Y + Y(a) + \int_{x_0-0}^{x_0} dA \cdot Y$$
$$= Y(x_0 - 0) + \Delta^- A(x_0)Y(x_0) = B(a, x_0 - 0)Y(a) + \Delta^- A(x_0)Y(x_0).$$

Consequently $Y(x_0)$ is determined uniquely, if and only if the matrix $I - \Delta^{-}A(x_0)$ has a reciprocal. Assuming this to be the case, we have

$$Y(x_0) = (I - \Delta^{-}A(x_0))^{-1}B(a, x_0 - 0)Y(a).$$

Similarly we show that

$$\begin{aligned} Y(x_0 + 0) &= Y(x_0) + \Delta^{+}A(x_0)Y(x_0) \\ &= (I + \Delta^{+}A(x_0))(I - \Delta^{-}A(x_0))^{-1}B(a, x_0 - 0)Y(a). \end{aligned}$$

For $x > x_0 + 0$

$$Y(x) = \int_{x_0+0}^{x} dA \cdot Y + Y(x_0 + 0),$$

so that

$$Y(x) = B(x_0 + 0, x)(I + \Delta^{+}A(x_0))(I - \Delta^{-}A(x_0))^{-1}B(a, x_0 - 0)Y(a).$$

This suggests setting

$$\begin{aligned} \bar{B}(a, x) &= B(a, x) && \text{for } x < x_0 \\ &= (I - \Delta^{-}A(x_0))^{-1}B(a, x_0 - 0) && \text{for } x = x_0 \\ &= B(x_0 + 0, x)(I + \Delta^{+}A(x_0))(I - \Delta^{-}A(x_0))^{-1}B(a, x_0 - 0) && \\ & && \text{for } x > x_0 . \end{aligned}$$

If we examine this expression carefully, we find that the definition of the matrix $B(a, x)$ for the case when $A(x)$ has a finite number of discontinuities centers in the interval function

$$C(x', x'') = (I - \Delta^{-}A(x''))^{-1}B(x' + 0, x'' - 0)(I + \Delta^{+}A(x'))$$

where $A(x)$ is continuous for $x' < x < x''$. Then

$$\bar{B}(a, x) = C(x_k , x)C(x_{x-1} , x_k) \cdots C(a, x_1)$$

where $a \leqq x_1 < x_2 < \cdots < x_k \leqq x$, are points of discontinuity of $A(x)$ on (a, x). In general

$$\bar{B}(x', x'') = C(x_m , x'')C(x_{m-1} , x_m) \cdots C(x', x_k)$$

where $x_k , \cdots , x_m$ are the ordered points of discontinuity of A in the closed interval $x' \leqq x \leqq x''$. We can then state the following theorem:

THEOREM. *If $A(x)$ has a finite number of discontinuities on (a, b), then a unique solution of the system* (I) *on $a \leqq x \leqq b$ exists if and only if the matrices $I - \Delta^{-}A(x)$ have reciprocals for all points of discontinuity of A. The solution can then be expressed $Y(x) = \bar{B}(a, x)Y(a)$.*

It must be pointed out that the order $a < b$ is important here. If $b < a$, or in $C(x', x'')$, $x' > x''$, then the differences are reversed; $x'' - 0$ is replaced by $x'' + 0$ and $x' - 0$ by $x' + 0$, so that for $x' > x''$ we define

$$C(x', x'') = (I + \Delta^{+}A(x''))^{-1}B(x' - 0, x'' + 0)(I - \Delta^{-}A(x'))$$

which exists if and only if $I + \Delta^+A(x)$ has a reciprocal for all discontinuities of A. It is obvious that $C(x', x'')C(x'', x') = I$ for all (x', x''), so that $\bar{B}(x', x'')\bar{B}(x'', x') = I$. As a matter of fact if x', x'', x''' are any three points of (a, b), then $\bar{B}(x'', x''')\bar{B}(x', x'') = \bar{B}(x', x''')$. Further the fact that $\bar{B}(x', x'') = \int_{x'}^{x''} dA(s)\bar{B}(x', s) + I$ can be obtained from the uniqueness of the solution of equations (I) by setting $Y(a)$ successively equal to the rows of the identity matrix. For the unique existence of $\bar{B}(x', x'')$ for all pairs x', x'', it must be assumed that both $I - \Delta^-A(x)$ and $I + \Delta^+A(x)$ have reciprocals for all points of discontinuity of A (and so for all x).

6. The nonhomogeneous system if $A(x)$ has a finite number of discontinuities

In the case when $A(x)$ is continuous the solution of

$$\text{(II)} \qquad Y(x) = \int_a^x dA(s)Y(s) + U(x) \quad \text{with} \quad Y(a) = U(a)$$

is expressible in the form $Y(a) = B(a, x)U(a) + \int_a^x B(t, x)\, dU(t)$. If we test this form for the case when $A(x)$ has a finite number of discontinuities with $\bar{B}(a, x)$ defined as above, we find by using Dirichlet's formula

$$\begin{aligned}\int_a^x dA(s)Y(s) &= \int_a^x dA(s)B(a, s)U(a) + \int_a^x dA(s) \int_a^s B(t, s)\, dU(t) \\ &= (B(a, x) - I)U(a) + \int_a^x \left(\int_t^x dA(s)B(t, s)\right) dU(t) \\ &\quad - \sum_{a \leqq y \leqq x} [\Delta^+A(y)B(y, y)\Delta^+U(y) - \Delta^-A(y)B(y, y)\Delta^-U(y)] \\ &= B(a, x)U(a) - U(a) + \int_a^x (B(t, x) - I)\, dU(t) \\ &\qquad - \sum_{a \leqq y \leqq x} [\Delta^+A(y)\Delta^+U(y) - \Delta^-A(y)\Delta^-U(y)] \\ &= B(a, x)U(a) + \int_a^x B(t, x)\, dU(t) - U(x) \\ &\qquad - \sum_{a \leqq y \leqq x} [\Delta^+A(y)\Delta^+U(y) - \Delta^-A(y)\Delta^-U(y)].\end{aligned}$$

It follows that $Y(x) = B(a, x)U(a) + \int_a^x B(t, x)\, dU(t)$ is a solution of the nonhomogeneous equation (II) if the term $\sum_y$ vanishes, in particular, if A and U have no common discontinuities. This solution will be unique since $Y(x) = \int_a^x dA(s)Y(s)$ is equivalent to $Y(x) = B(a, x)Y(a) = 0$, where $B(a, x)$ is uniquely determined.

To determine the form of the corrective term in our assumed solution, we again assume that $A(x)$ has a single discontinuity at x_0 and proceed as in system (I). As a result of rather tedious but fairly obvious manipulation we find that if $I - \Delta^{-}A(x)$ has a reciprocal at x_0 then the solution of (II) is as follows

$$\begin{aligned} Y(x) &= B(a, x)U(a) + \int_a^x B(t, x)\, dU(t) && \text{for} \quad x < x_0, \\ &= \bar{B}(a, x_0)U(a) + \int_a^x \bar{B}(t, x_0)\, dU(t) \\ &\qquad + (I - \Delta^{-}A(x_0))^{-1}\Delta^{-}A(x_0)\Delta^{-}U(x_0) && \text{for} \quad x = x_0, \\ &= \bar{B}(a, x)U(a) + \int_a^x \bar{B}(t, x)\, dU(t) + \bar{B}(x_0 - 0, x)\Delta^{-}A(x_0)\Delta^{-}U(x_0) \\ &\qquad - \bar{B}(x_0 + 0, x)\Delta^{+}A(x_0)\Delta^{+}U(x_0) && \text{for} \quad x > x_0. \end{aligned}$$

This suggests, and it is easily verifiable, that if $A(x)$ has a finite number of discontinuities and $I - \Delta^{-}A(x)$ has a reciprocal for each of these, then the unique solution of the system (II) with $a < b$ is expressible in the form

$$\begin{aligned} Y(x) &= \bar{B}(a, x)U(a) + \int_a^x \bar{B}(t, x)\, dU(t) \\ &\quad + \sum_{a \leqq y \leqq x} [\bar{B}(y - 0, x)\Delta^{-}A(y)\Delta^{-}U(y) - \bar{B}(y + 0, x)\Delta^{+}A(y)\Delta^{+}U(y)]. \end{aligned}$$

Obviously the corrective term vanishes if A and U have no common discontinuities.

7. The matrix $A(x)$ has an infinite number of discontinuities

In case there are no restrictions on the number of discontinuities of $A(x)$, we set $A(x) = A_c(x) + A_b(x)$, where $A_c(x)$ is the continuous part of A, and $A_b(x) = \sum_k A_k(x)$, where $A_k(x)$ is the simple break function corresponding to the discontinuity x_k. The convergence in this series is in the terms of total variation, i.e., $\lim_m V_a^b(A_b(x) - \sum_{k=1}^m A_k(x)) = 0$. If we denote by σ any finite number of the discontinuities and set

$$A_\sigma(x) = A_c(x) + \textstyle\sum_\sigma A_k(x),$$

then on any interval containing no points of σ, $\Delta A_\sigma(x) = \Delta A_c(x)$. As a consequence the fundamental matrix $B_\sigma(x', x'')$ corresponding to $A_\sigma(x)$ can be written

$$B_\sigma(x', x'') = C(x_m, x'')C(x_{m-1}, x_m) \cdots C(x', x_k),$$

where $x_k < \cdots < x_m$, are the points of σ in the interval (x', x''),

$$C(s, t) = (I - \Delta^{-}A(t))^{-1}B_c(s, t)(I + \Delta^{+}A(s)),$$

while $B_c(s, t)$ corresponds to $A_c(x)$. This suggests that $\lim_\sigma B_\sigma(x', x'')$ may exist, and that this can serve as the fundamental matrix for $A(x)$.

We note in the first place that $B_\sigma(x', x'')$ is uniformly bounded as to σ, x', and x'', i.e., there exists an M such that $\|B_\sigma(x', x'')\| \leqq M$. Because of the properties of norms we have

$$\| C(s, t) \| \leqq \| (I - \Delta^- A(t))^{-1} \| \cdot \| B_c(s, t) \| \cdot \| I + \Delta^+ A(s) \| .$$

Now

$$\| B_c(s, t) \| \leqq \exp V_s^t A_c \quad \text{and} \quad \| I + \Delta^+ A(a) \| \leqq 1 + \| \Delta^+ A(s) \| .$$

Further if $\| \Delta^- A(t) \| < 1$, then

$$\| (I - \Delta^- A(t))^{-1} \| \leqq 1 + \textstyle\sum_m \| \Delta^- A(t) \|^m = (1 - \| \Delta^- A(t) \|)^{-1}.$$

Using the fact that $\sum_s \| \Delta^+ A(s) \| < \infty$ and $\sum_s \| \Delta^- A(s) \| < \infty$ so that $\prod_s (1 + \| \Delta^+ A(s) \|)$ and $\prod_s (1 - \| \Delta^- A(s) \|)$ converge, we can show that

$$\| B(x', x'') \| \\ \leqq N \textstyle\prod_x (1 + \| \Delta^+ A(x) \|) \prod_s (1 - \| \Delta^- A(s) \|)^{-1} \exp V_a^b A_c = M,$$

where s is limited to the discontinuities for A for which $\| \Delta^- A(s) \| < 1$, and $N = \prod_t \| (I - \Delta^- A(t))^{-1} \|$, where t ranges over the points, finite in number for which $\| \Delta^- A(t) \| \geqq 1$ and $\| (I - \Delta^- A(t))^{-1} \| > 1$. It is obvious that if any terms in the product definition of $B_\sigma(x', x'')$ are omitted, the same inequality holds, i.e., the same upper bound M serves.

In order to show that $\lim_\sigma B_\sigma(x', x'')$ exists, we follow the usual procedure involved in proving convergence of infinite products and observe that if $A_1, \cdots, A_m$ and $B_1, \cdots B_m$ are $2m$ matrices, then

$$\| A_1 A_2 \cdots A_m - B_1 B_2 \cdots B_m \| \leqq K(\textstyle\sum_{i=1}^m \| A_i - B_i \|),$$

where K is the maximum of $\| A_1 \cdots A_j B_{j+1} \cdots B_m \|$, $j = 1, \cdots, m$. Suppose $\sigma_1 \geqq \sigma$. Let σ consist of $a = x_0 < x_1 < x_2 < \cdots < x_k = b$, and let σ_1 consist of $x_{i-1} = x_{i0} < x_{i1} < \cdots < x_{ik_i} = x_i$, $i = 1, \cdots, k$. Then if we remember that $B_c(x_{i-1}, x_i) = \prod_{j=1}^{k_i} B_c(x_{ij-1}, x_{ij})$, we obtain

$$\begin{aligned} \| B_{\sigma_1}(a, b) - B_\sigma(a, b) \| &= \left\| \prod_{i=1}^k C(x_{i-1}, x_i) - \prod_{i=1}^k \prod_{j=1}^{k_i} C(x_{i,j-1}, x_{ij}) \right\| \\ &\leqq M \sum_{i=1}^k \left\{ \sum_{j=1}^{k_i-1} (\| I - (I - \Delta^+ A(x_{ij})) \| \right. \\ &\qquad \left. + \| I - (I - \Delta^- A(x_{ij}))^{-1} \|) \right\} \\ &\leqq M \sum_{i=1}^k \left\{ \sum_{j=1}^{k_i-1} (\| \Delta^+ A(x_{ij}) \| \right. \\ &\qquad \left. + \| (I - \Delta^- A(x_{ij}))^{-1} \| \, \| \Delta^- A(x_{ij}) \|) \right\} . \end{aligned}$$

If we assume that σ contains all points for which $\| \Delta^- A(x_i) \| \geqq \frac{1}{2}$, we have

$$\| B_{\sigma_1}(a, b) - B_\sigma(a, b) \| \leqq M \sum_{i=1}^k \left\{ \sum_{j=1}^{k_i-1} (\| \Delta^+ A(x_{ij}) \| + 2 \| \Delta^- A(x_{ij}) \|) \right\} ,$$

where now the points x_{ij} involved on the right-hand side of this inequality belong to $\sigma_1 - \sigma$. Since $\sum_x \| \Delta^+ A(x) \| < \infty$ and $\sum_x \| \Delta^- A(x) \| < \infty$, we conclude that there exists a σ_e such that if $\sigma \geqq \sigma_e$ and $\sigma_1 \geqq \sigma$, then

$$\| B_{\sigma_1}(a, b) - B_\sigma(a, b) \| \leqq e,$$

from which one deduces in the usual way that $\lim_\sigma B_\sigma(a, b)$ exists. The same reasoning proves that if $a \leqq x' < x'' \leqq b$, then $\lim_\sigma B_\sigma(x', x'')$ exists (denoted by $B(x', x'')$), as a matter of fact uniformly in x', x''. A similar procedure shows that if $x' > x''$, then $\lim_\sigma B_\sigma(x', x'')$ also exists, and that since $B_\sigma(x'', x''')B_\sigma(x', x'') = B_\sigma(x', x''')$ for all x', x'', x''', we have also $B(x'', x''')B(x', x'') = B(x', x''')$. Further $B_\sigma(x, x) = I$ for all σ implies $B(x, x) = I$. We then have fundamental matrices having the same (harmonic) properties as the matrices B for the continuous case, subject, of course, to the proviso that $I + \Delta^+ A(x)$ and $I - \Delta^- A(x)$ have reciprocals for all discontinuities of $A(x)$.

Since $B_\sigma(x', x'') = \int_{x'}^{x''} dA_\sigma(s)B_\sigma(x', s) + I$ for all σ, $\lim_\sigma V(A - A_\sigma) = 0$, and $\lim_\sigma B_\sigma(x', x'') = B(x', x'')$ uniformly in x', x'', the convergence theorem on integrals assures us that

$$B(x', x'') = \int_{x'}^{x''} dA(s)B(x', s) + I,$$

and that a solution of $Y(x) = \int_a^x dA(s)Y(s) + Y(a)$ can be expressed in the form $Y(x) = B(a, x)Y(a)$. To show that this solution is unique, we show that $Y(x) = \int_a^x dA(s)Y(s)$ implies $Y(x) = 0$, $Y(x)$ being of bounded variation and so bounded in x. Now

$$Y(s) = \int_a^x dA(s)Y(s) = \int_a^x dA_\sigma(s)Y(s) + \int_a^x d(A(s) - A_\sigma(s))Y(s).$$

Since $A_\sigma(s)$ has only a finite number of discontinuities and has no discontinuities in common with $\int_a^x d(A(s) - A_\sigma(s))Y(s) = U(x)$, it follows from §6 that $Y(x) = \int_a^x B_\sigma(s, x)\, d(A(s) - A_\sigma(s))Y(s)$. Then

$$\| Y(x) \| \leqq MV_a^b(A - A_\sigma)\text{l.u.b.}_x \| Y(x) \| .$$

From this it follows that $\| Y(x) \| \leqq e$ for any e, i.e., $Y(x) = 0$ in x. As a consequence, $B(x', x'')$ is the unique solution of the system

$$B(x', x'') = \int_{x'}^{x''} dA(s)B(x', s) + I,$$

and gives a fundamental system of solutions for system (I).

8. Properties of the fundamental matrix $B(x', x'')$

We shall assume that $x' < x''$ except when stated otherwise. The changes to be made when $x' > x''$ are easily apparent from the change in the definition of the fundamental interval function $C(x', x'')$.

(a) The matrix $B(x', x'')$ is of bounded variation in x'' uniformly for

$a \leqq x'' \leqq b$, and in x' for $a \leqq x' \leqq b$. The only points of discontinuity in x' or x'' are those of $A(x)$.

From the integral equation satisfied by $B(x', x'')$ it follows that

$$\| B(x', x_2) - B(x', x_1) \| = \left\| \int_{x_1}^{x_2} dA(s)B(x', s) \right\| \leqq M \mid VA(x_2) - VA(x_1) \mid .$$

The continuity and bounded variation properties in x'' follow from this inequality. The identity

$$B(x', x'' + 0) - B(x', x'' - 0) = (A(x'' + 0) - A(x'' - 0))B(x', x'')$$

follows from the integral equation.

For $B(x', x'')$ as a function of x' we have

$$B(x_2, x'') - B(x_1, x'') = \int_{x_2}^{x''} dA(s)B(x_2, s) - \int_{x_1}^{x''} dA(s)B(x_1, s)$$
$$= \int_{x_2}^{x''} dA(s)(B(x_2, s) - B(x_1, s)) - \int_{x_1}^{x_2} dA(s)B(x_1, s).$$

Considering this as an integral equation to be solved for

$$B(x_2, x'') - B(x_1, x''),$$

we find

$$B(x_2, x'') - B(x_1, x'') = B(x_2, x'') \int_{x_1}^{x_2} dA(s)B(x_1, s).$$

It follows that

$$\| B(x_2, x'') - B(x_1, x'') \| \leqq M^2 \mid VA(x_2) - VA(x_1) \mid ,$$

from which the continuity and bounded variation properties in x' follow.

(b) The matrix $B(x', x'')$ can be expressed as a product integral. We shall set $x' = a$, $x'' = b$ with $a < b$; the alterations for the general case will be obvious. If $A(x)$ is continuous, then

$$B(a, b) = \lim (I + A(b) - A(x_n)) \cdots (I + A(x_1) - A(a)),$$

where the limit exists either as $\mid \sigma \mid = \max \mid x_i - x_{i-1} \mid$ approaches zero or by successive refinements of σ. The corrective terms involved in the definition of $B_\sigma(a, b)$ in §7, suggest that we consider the product

$$(I - \Delta^- A(b))^{-1}(I + A(b - 0) - A(x_n + 0))(I + \Delta^+ A(x_n)) \cdots$$
$$(I - \Delta^- A(x_1))^{-1}(I + A(x_1 - 0) - A(a + 0))(I + \Delta^+ A(a))$$
$$= D(x_n, b)D(x_{n-1}, x_n) \cdots D(a, x_1),$$

where we define

$$D(x', x'')$$
$$= (I - \Delta^- A(x''))^{-1}(I + A(x'' - 0) - A(x' + 0))(I + \Delta^+ A(x')).$$

We prove the following theorem:

THEOREM. *If $A(x)$ is of bounded variation on (a, b), then*

$$\lim_\sigma \prod_\sigma D(x_{i-1}, x_i)$$

exists in the sense of successive subdivisions and is equal to $B(a, b)$.

If we set

$$E(x', x'') = (I - \Delta^- A(x''))^{-1} (I + A_c(x'') - A_c(x'))(I + \Delta^+ A(x')),$$

then we obtain our theorem by comparing

$$\begin{aligned} B(a, b) &\quad \text{with} \quad B_\sigma(a, b) = \textstyle\prod_\sigma C(x_{i-1}, x_i), \\ \textstyle\prod_\sigma C(x_{i-1}, x_i) &\quad \text{with} \quad \textstyle\prod_\sigma E(x_{i-1}, x_i), \quad \text{and} \\ \textstyle\prod_\sigma E(x_{i-1}, x_i) &\quad \text{with} \quad \textstyle\prod_\sigma D(x_{i-1}, x_i). \end{aligned}$$

Since $\lim_\sigma B_\sigma(a, b) = B(a, b)$, we have that, for every $e > 0$, there exists a σ_e such that if $\sigma \geqq \sigma_e$, then $\| B(a, b) - B_\sigma(a, b) \| \leqq e$. Here the σ is limited to the points of discontinuity of $A(x)$. However since

$$B_c(x'', x''')B_c(x', x'') = B_c(x', x''')$$

and the "corrective" multipliers on B_c drop out if x_{i-1} or x_i are points of continuity of A, we see that limiting the points of σ to discontinuity points of A is unnecessary.

In order to compare $C(x', x'')$ with $E(x', x'')$ we note that

$$\begin{aligned} &\| B_c(x', x'') - (I + A_c(x'') - A_c(x')) \| \\ &\qquad \leqq | VA_c(x'') - VA_c(x') |^2 \exp | VA_c(x'') - VA_c(x') |, \end{aligned}$$

a consequence of the Peano series for $B_c(x', x'')$. Consequently

$$\begin{aligned} &\textstyle\sum_\sigma \| B_c(x_{i-1}, x_i) - (I + A_c(x_i) - A_c(x_{i-1})) \| \\ &\qquad \leqq \max_\sigma | VA_c(x_i) - VA_c(x_{i-1}) | \, V_a^b A \exp V_a^b A, \end{aligned}$$

so that

$$\begin{aligned} &\| \textstyle\prod_\sigma C(x_{i-1}, x_i) - \prod_\sigma E(x_{i-1}, x_i) \| \\ &\qquad \leqq M \textstyle\sum_\sigma \| B_c(x_{i-1}, x_i) - (I + A_c(x_i) - A_c(x_{i-1})) \| \\ &\qquad \leqq M \max_\sigma | VA_c(x_i) - VA_c(x_{i-1}) | \, V_a^b A \exp V_a^b A. \end{aligned}$$

Because of the continuity of $A_c(x)$, it follows that if $| \sigma |$ is made small enough, this expression can be made smaller than e.

Finally since

$$\begin{aligned} &\| E(x', x'') - D(x', x'') \| \leqq \| (I - \Delta^- A(x''))^{-1} \| \\ &\quad \cdot \| A(x'' - 0) - A(x' + 0) - (A_c(x'') - A_c(x')) \| \cdot \| I + \Delta^+ A(x') \|, \end{aligned}$$

and

$$\| A(x'' - 0) - A(x' + 0) - (A_c(x'') - A_c(x')) \| \leqq \textstyle\sum_{x' \leqq x \leqq x''} \| A(x + 0) - A(x - 0) \|,$$

we have

$$\| \textstyle\prod_\sigma E(x_{i-1}, x_i) - \prod_\sigma D(x_{i-1}, x_i) \| \leqq M \sum_{x \in C\sigma} \| A(x + 0) - A(x - 0) \|,$$

where $C\sigma$ are the points *not* in σ. The right-hand side of this inequality can be made to be less than e by choosing σ so that

$$\textstyle\sum_{x \in C\sigma} \| A(x + 0) - A(x - 0) \| < e/M.$$

Combining these various considerations, we find that there exists a σ_e' such that if $\sigma \geqq \sigma_e'$, then $\| B(a, b) - \prod_\sigma D(x_{i-1}, x_i) \| \leqq 3e$, the desired result.

If $x' > x''$, we define

$$D(x', x'') = (I + \Delta^+ A(x''))^{-1}(I - (A(x' - 0) - A(x'' + 0)))(I - \Delta^- A(x')).$$

From the expression for $B(x', x'')$ with $x' < x''$ we can check that

$$B(x', x'' + 0) - B(x', x'' - 0) = (A(x'' + 0) - A(x'' - 0))B(x', x'').$$

On the other hand,

$$B(x' + 0, x'') = B(x', x'')(I + \Delta^+ A(x'))^{-1},$$

and

$$B(x' - 0, x'') = B(x', x'')(I - \Delta^- A(x'))^{-1},$$

so that

$$B(x' + 0, x'') - B(x' - 0, x'') = B(x', x'')[(I + \Delta^+ A(x'))^{-1} - (I - \Delta^- A(x'))^{-1}].$$

It follows that $B(x', x'')$ considered as a function of x' does not in general satisfy the adjoint system $B(x', x'') = \int_{x'}^{x''} B(s, x'')\, dA(s) + I$, since this would imply that

$$B(x' + 0, x'') - B(x' - 0, x'') = B(x', x'')(A(x' - 0) - A(x' + 0)).$$

(c) The matrix $A(x)$ in terms of $B(x', x'')$. We have (Wall [14], p. 161)

THEOREM. *If $B(x', x'')$ is the fundamental matrix for $A(x)$, then*

$$A(x) - A(a) = \int_a^x d_s\, B(x', s)B(s, x'),$$

so that the integral on the right is independent of x'.

If we replace $B(x', s)$ by $\int_{x'}^{s} dA(t)B(x', t) + I$ and use the substitution theorem for integrals, we obtain

$$\int_a^x d_s\, B(x', s)B(s, x') = \int_a^x dA(s)B(x', s)B(s, x')$$
$$= \int_a^x dA(s) = A(x) - A(a).$$

Conversely,

THEOREM. *If $B(x, x) = I$ for all x and $B(x'', x''')B(x', x'') = B(x', x''')$ for all x', x'', x''', $B(x', x'')$ being of bounded variation in x' for each x'', and in x'' for each x', then $A(x, t) = \int_a^x d_s\, B(t, s)B(s, t)$ is independent of t, and $B(x', x'') = \int_{x'}^{x''} dA(s)B(x', s) + I$.*

The fact that

$$\int_a^x d_s\, B(t, s)B(s, t) = \int_a^x d_s\, B(t, s)B(t', t)B(t, t')B(s, t)$$
$$= \int_a^x d_s\, B(t', s)B(s, t')$$

shows that $A(x, t)$ is independent of t. By the substitution theorem, we have

$$\int_{x'}^{x''} d_s\, A(s, t)B(x', s) = \int_{x'}^{x''} d_s\, B(t, s)B(s, t)B(x', s)$$
$$= \int_{x'}^{x''} d_s\, B(t, s)B(x', t) = (B(t, x'') - B(t, x'))B(x', t)$$
$$= B(x', x'') - I,$$

i.e., $B(x', x'')$ satisfies the integral equation (I) for $A(s, t)$ for all t.

As already noted in the case when $A(x)$ is continuous, $B(x', x'')$ is expressible as the product integral $\int_{x'}^{x''} (I + dA) = \lim_\sigma \prod (I + \Delta_i A)$. It follows from the above theorem that if A is discontinuous, the same product integral which can readily be shown to exist as a limit in the sense of successive subdivisions σ (see MacNerney [6], pp. 186–187) does not correspond to the matrix of a differentio-integral equation. For it is not in general true that $\int_a^b (I + dA) \int_b^a (I + dA) = I$. Take for instance $A(x) = 0$ for $0 \leqq x < \frac{1}{2}$; c_1 for $x = \frac{1}{2}$; and c_2 for $\frac{1}{2} < x \leqq 1$. Then

$$\int_0^1 (1 + dA(x)) = (1 + c_1)(1 + c_2 - c_1),$$

while $\int_1^0 (1 + dA(x)) = (1 - (c_2 - c_1))(1 - c_1)$, so that these two integrals are in general not reciprocals of each other.

Incidentally this example also shows that if A is discontinuous,

$$\lim \prod (I + \Delta_i A)$$

need not exist as $|\sigma|$ approaches zero, contrary to a theorem stated by F. M. Stewart ([12], pp. 101). For if σ does not contain the point $\frac{1}{2}$, then $\prod (1 + \Delta_i A) = 1 + c_2$, while if σ contains the point $\frac{1}{2}$, then

$$\prod (1 + \Delta_i A) = (1 + c_1)(1 + c_2 - c_1),$$

so that the limit exists in the norm sense if and only if $c_1(c_1 - c_2) = 0$, i.e., $A(x)$ is continuous on the right or on the left.

9. The determinant of the matrix $B(a, b)$

If $A(x)$ has a finite number of discontinuities at $a \leqq x_1 < \cdots < x_m \leqq b$, then (with $x_0 = a$, $x_{m+1} = b$) $B(a, b) = \prod_{i=1}^{m+1} C(x_{i-1}, x_i)$. Now $\det C(x', x'') = \det [(I - \Delta^- A(x''))^{-1}] \cdot \det B_c(x', x'') \cdot \det (I + \Delta^+ A(x'))$, and $\det B_c(x', x'') = \exp \left(\sum_{j=1}^n (a_{jj}(x'' - 0) - a_{jj}(x' + 0))\right)$. Hence

$$\det B(a, b) = \exp \left(\textstyle\sum_{j=1}^n (a_{cjj}(b) - a_{cjj}(a))\right) \cdot \prod \det (I + \Delta^+ A(x_i)) / \prod \det (I - \Delta^- A(x_i)),$$

where $a_{cjj}(x)$ is the continuous part of $a_{jj}(x)$. Since the determinant is a continuous function of its elements, it follows that in any case

$$\det B(a, b) = (\exp \textstyle\sum_j (a_{cjj}(b) - a_{cjj}(a))) \cdot \prod_x \det (I + \Delta^+ A(x) / \prod_x \det (I - \Delta^- A(x)).$$

By the use of Hadamard's theorem on the maximum of a determinant it can be shown that the infinite products involved in this expression converge absolutely. Since $\det B(a, b)$ and $\det B(b, a)$ are reciprocals of each other, this formula points up the fact that for $B(b, a)$ to exist it is necessary that $I + \Delta^+ A(x)$ have a reciprocal for every x. (See also MacNerney [7].)

10. The nonhomogeneous equation

From the preceding developments and the treatment of equation (II) for the case of a finite number of discontinuities of $A(x)$, the following theorem is suggested:

THEOREM. *If $a \leqq x \leqq b$, $I - \Delta^- A(x)$ has a reciprocal for all x, and $U(x)$ is of bounded variation, then the system $Y(x) = \int_a^x dA(s) Y(s) + U(x)$ with $Y(a) = U(a)$ has as unique solution*

$$Y(x) = B(a, x) U(a) + \int_a^x B(s, x)\, dU(s) + \sum_{a \leqq y \leqq x} [B(y - 0, x) \Delta^- A(y) \Delta^- U(y) - B(y + 0, x) \Delta^+ A(y) \Delta^+ U(y)].$$

The fact that this is a solution can be easily verified. Since the infinite series on the right is uniformly convergent, term by term integration is permitted, and we have

$$\int_a^x dA(s)Y(s) = \int_a^x dA(s)B(a, s)U(a) + \int_a^x dA(s) \int_a^s B(t, s)\, dU(t)$$
$$+ \sum_{a \leqq y \leqq x} \left[\int_{y-0}^x dA(s)B(y - 0, s)\Delta^- A(y)\Delta^- U(y) \right.$$
$$\left. - \int_{y+0}^x dA(s)B(y + 0, s)\Delta^+ A(y)\Delta^+ U(y) \right].$$

The lower limits on the last two integrals are as indicated, because the term does not enter into the expression for $Y(x)$ until x or s is $\geqq y - 0$ or $y + 0$, respectively. Applying the Dirichlet formula of §3 and the fact that

$$\int_{y-0}^x dA(s)B(y - 0, s) = B(y - 0, x) - I,$$

(similarly for $y + 0$), we find that $\int_a^x dA(s)Y(s) = Y(x) - F(x)$. The uniqueness of the solution has already been considered.

Another form for this solution of the nonhomogeneous equation can be obtained, if we note that by §8(b)

$$B(y - 0, x)\Delta^- A(y) = B(y - 0, x) - B(y, x) = -\Delta_y^- B(y, x),$$

and

$$B(y + 0, x)\Delta^+ A(y) = B(y, x) - B(y + 0, x) = -\Delta_y^+ B(y, x).$$

Then

$$Y(x) = B(a, x)U(a) + \int_a^x B(s, x)\, dU(s)$$
$$- \sum_{a \leqq y \leqq x} [\Delta_y^- B(y, x)\Delta^- U(y) - \Delta_y^+ B(y, x)\Delta^+ U(y)].$$

If we apply the integration by parts theorem of §3(e) to the last two terms, we obtain

$$Y(x) = B(a, x)U(a) + B(s, x)U(s) \Big|_{s=a}^{s=x} - \int_a^x d_s B(s, x)U(s)$$
$$= U(x) - \int_a^x d_s B(s, x)U(s).$$

If $U(a) \neq Y(a)$, i.e., if $(Y(x) = \int_a^x dA(s)Y(s) + U(x) - U(a) + Y(a)$, then

$$Y(x) = B(a, x)Y(a) + U(x) - U(a) - \int_a^x d_s B(s, x)(U(s) - U(a)).$$

11. The matrix $B(x', x'')$ as a function of the lower limit x'

The fact that the solution of the nonhomogeneous equation discussed in the preceding paragraph involves the integration of the matrix $B(x', x'')$ with respect to the lower limit x', suggests that this matrix may perhaps satisfy such an equation with the function $U(x)$ replaced by $A(x)$. This is actually the case. We rewrite the equation

$$B(x', x'') = \int_{x'}^{x''} dA(s)B(x', s) + I$$

in the form

$$B(x', x'') - I = \int_{x'}^{x''} dA(s)(B(x', s) - I) + A(x'') - A(x'),$$

and then apply the formula of §10 to obtain the relation:

$$\begin{aligned} B(x', x'') &= I + \int_{x'}^{x''} B(s, x'')\, dA(s) \\ &\quad + \sum_{x' \leqq y \leqq x''} [B(y - 0, x'')(\Delta^- A(y))^2 - B(y + 0, x'')(\Delta^+ A(Y))^2]. \end{aligned}$$

This reduces to the usual adjoint equation if the sum term vanishes, which is true in particular if $A(x)$ is continuous. An alternative form can be obtained by observing that $B(y - 0, x) = B(y, x)(I - \Delta^- A(y))^{-1}$ and $B(y + 0, x) = B(y, x)(I + \Delta^+ A(y))^{-1}$. Since for any matrix C,

$$(I - C)^{-1}C^2 = -I - C + (I - C)^{-1},$$

we have

$$\begin{aligned} &\textstyle\sum_{x' \leqq y \leqq x''} B(y, x'')[(I - \Delta^- A(y))^{-1}(\Delta^- A(y))^2 - (I + \Delta^+ A(y))^{-1}(\Delta^+ A(y))^2] \\ &\quad = \textstyle\sum_{x' \leqq y \leqq x''} B(y, x'') \\ &\qquad \cdot[-\Delta^- A(y) + (I - \Delta^- A(y))^{-1} - \Delta^+ A(y) - (I + \Delta^+ A(y))^{-1}] \\ &\quad = \textstyle\sum_{x' \leqq y \leqq x''} B(y, x'')[-\Delta^{\pm} A(y) + (I - \Delta^- A(y))^{-1} - (I + \Delta^+ A(y))^{-1}]. \end{aligned}$$

As a consequence if we set

$$\begin{aligned} \bar{A}(x) &= A(x) - \textstyle\sum_{a \leqq y \leqq x} \Delta^{\pm} A(y) \\ &\qquad + \textstyle\sum_{a \leqq y \leqq x} [(I - \Delta^- A(y))^{-1} - (I + \Delta^+ A(y))^{-1}] \\ &= A_c(x) + \textstyle\sum_{a \leqq y \leqq x} [(I - \Delta^- A(y))^{-1} - (I + \Delta^+ A(y))^{-1}], \end{aligned}$$

then

$$B(x', x'') = \int_{x'}^{x''} B(s, x'')\, d\bar{A}(s) + I,$$

i.e., $B(x, a)$ as a function of x satisfies an equation of the form

$$Z(x) = \int_x^a Z(s)\, d\bar{A}(s) + Z(a),$$

where the continuous parts of $A(s)$ and $\bar{A}(s)$ are the same.

12. The adjoint equation

It is obvious that the adjoint equation $Z(x) = -\int_a^x Z(s)\, dA(s) + Z(a)$ can be subjected to the same treatment as equation (I) and yields a matrix

$C(a, x)$ which satisfies the equation $C(a, x) = -\int_a^x C(a, s)\, dA(s) + I$. The question of the relation between $B(a, x)$ and $C(a, x)$ is pertinent. We note that

$$\int_a^x C(a, s)\, dB(a, s) + \int_a^x dC(a, s)B(a, s) = 0.$$

Applying the integration by parts theorem we obtain

$$0 = C(a, s)B(a, s) \big|_a^x - \sum_{a \leqq y \leqq x} [\Delta^+ C(a, y)\Delta^+ B(a, y) - \Delta^- C(a, y)\Delta^- B(a, y)].$$

Solving this equation for $C(a, x)$ and remembering that the reciprocal of $B(a, x)$ is $B(x, a)$, we have

$$C(a, x) = B(x, a) - \sum_{a \leqq y \leqq x} [\Delta^+ C(a, y)\Delta^+ B(x, y) - \Delta^- C(a, y)\Delta^- B(x, y)].$$

Now $\Delta^+ C(a, y) = -C(a, y)\Delta^+ A(y)$, $\Delta^+ B(x, y) = \Delta^+ A(y)B(x, y)$, and similarly for Δ^-. Then

$$C(a, x) = B(x, a) + \sum_{a \leqq y \leqq x} [C(a, y)(\Delta^+ A(y))^2 B(x, y) - C(a, y)(\Delta^- A(y))^2 B(x, y)].$$

Obviously $C(a, x) = B(x, a)$ if $A(x)$ is continuous.

13. Singular solutions

The preceding developments are valid provided that $(I + \Delta^+ A(x))$ and $(I - \Delta^- A(x))$ have reciprocals for every x. The contrary can occur at only a finite number of points, since $\sum(\| \Delta^+ A(x) \| + \| \Delta^- A(x) \|) \leqq V_a^b A$, and the matrix $I + A$ has a reciprocal if $\| A \| < 1$. Suppose that x_1 is the first point on (a, b) for which $I - \Delta^- A(x)$ fails to have a reciprocal. Then since at x_1 we would have

$$\begin{aligned} Y(x_1) &= \int_{x_1 - 0}^{x_1} dA(s)\, Y(s) + \int_a^{x_1 - 0} dA(s) Y(s) + U(x_1) \\ &= (\Delta^- A(x_1))Y(x_1) + Y(x_1 - 0) + U(x_1) - U(x_1 - 0), \end{aligned}$$

the value of $Y(x_1)$ is determined if and only if this system of equations has a solution. If r_1 is the rank of the matrix $I - \Delta^- A(x_1)$, then such a solution would contain $n - r_1$ arbitrary constants. These constants would carry over into the interval x_1, x_2, where x_2 is the next point of (a, b) for which $I - \Delta^- A(x)$ does not have a reciprocal. Proceeding in this way, we find that if a solution exists between a and x_k, then in the interval x_{k-1}, x_k such a solution may depend on at most $(n - r_1) + \cdots + (n - r_{k-1})$ arbitrary constants, since it is conceivable that the arbitrary constants may be reduced by the conditions for the existence of a solution at some point $x = x_i$. Similar statements can be made concerning the existence of nonvanishing solutions of the equation $Y(x) = \int_a^x dA(s)Y(s)$.

14. Generalization

The case when the system (II) reads

$$Y(x) = \int_a^x F(s)\, dA(s) Y(s) + U(x)$$

or

$$Y(x) = \int_a^x dA(s) F(s) Y(s) + U(x), \quad \text{with} \quad Y(a) = U(a),$$

where $F(x)$ is a matrix of functions L-S integrable with respect to A, does not introduce any new difficulties. Such a system is reduced to the system (II) by the substitution

$$A^*(x) = \int_a^x F(s)\, dA(s) \qquad \left(A^*(x) = \int_a^x dA(s) F(s)\right),$$

where now $A^*(x)$ will be a matrix of functions of bounded variation. The right to make this substitution depends on the substitution theorem for L-S integrals, viz., if $Y(s)$ is of bounded variation and $F(s)$ is L-S integrable with respect to $A(x)$, then

$$\int_a^x F(s)\, dA(s) Y(s) = \int_a^x dA^*(s) Y(s).$$

This suggests that an effective way to define a product integral for the expression $(I + F(x)\, dA(x))$, $F(x)$ L-S integrable with respect to $A(x)$, is to apply the definition given in §8 to $A^*(x) = \int_a^x F(s)\, dA(s)$. This avoids the complication inherent in the fact that measurable sets on the linear interval are difficult to arrange in a linear order.

15. Linear normed spaces

The developments of this paper have been made in such a way that they indicate the generalization involved by assuming that $Y(x)$ is on (a, b) to a linear normed complete space $\mathfrak{Y}$, and $A(x)$ is on (a, b) to the class of linear continuous transformations on $\mathfrak{Y}$ to $\mathfrak{Y}$, the $A(x)$ being strongly of bounded variation on (a, b). This would extend the situation developed by MacNerney [5] by dropping the condition that $A(x)$ be continuous in the strong sense on (a, b) (see also MacNerney [6]). So far as we can see, the basic results of this paper can be extended to this more general setting without any difficulty. The exception is, of course, that there is no parallel to the determinant of the matrix $B(a, x)$ in the general case. Also, some of the statements made for the singular case where for some points of discontinuity of $A(x)$, $I + \Delta^+ A$ and $I - \Delta^- A$ do not have reciprocals need alterations. Although there will still be only a finite number of such singular points, the question of solutions is tied up with the inversion of a continuous linear transformation T which does not have a reciprocal.

REFERENCES

1. GARRETT BIRKHOFF, *On product integration*, J. Math. Phys., vol. 16 (1937), pp. 104–132.
2. G. D. BIRKHOFF AND R. E. LANGER, *The boundary problems and developments associated with a system of ordinary linear differential equations of the first order*, Proc. Amer. Acad. Arts Sci., vol. 58 (1923), pp. 49–128; also George David Birkhoff: Collected mathematical papers, vol. I, pp. 345–424.
3. H. J. ETTLINGER, *On the inversion of the order of integration of a two fold iterated integral*, Ann. of Math. (2), vol. 28 (1926), pp. 65–68.
4. T. H. HILDEBRANDT, *Definitions of Stieltjes integrals of the Riemann type*, Amer. Math. Monthly, vol. 45 (1938), pp. 265–278.
5. J. S. MACNERNEY, *Stieltjes integrals in linear spaces*, Ann. of Math. (2), vol. 61 (1955), pp. 354–367.
6. ———, *Continuous products in linear spaces*, J. Elisha Mitchell Sci. Soc., vol. 71 (1955), pp. 185–200.
7. ———, *Determinants of harmonic matrices*, Proc. Amer. Math. Soc., vol. 7 (1956), pp. 1044–1046.
8. E. H. MOORE, *General analysis*, *Part II*, Memoirs of the American Philosophical Society, Philadelphia, 1939.
9. G. PEANO, *Intégration par séries des équations différentielles linéaires*, Math. Ann., vol. 32 (1888), pp. 450–456.
10. L. SCHLESINGER, *Vorlesungen ueber lineare Differentialgleichungen*, Leipzig, Teubner, 1908.
11. ———, *Neue Grundlagen für einen Infinitesimalkalkul der Matrizen*, Math. Zeit., vol. 33 (1931), pp. 33–61.
12. F. M. STEWART, *Integration in noncommutative systems*, Trans. Amer. Math. Soc., vol. 68 (1950), pp. 76–104.
13. W. H. YOUNG, *On integration with respect to a function of bounded variation*, Proc. London Math. Soc. (2), vol. 13 (1914), pp. 109–150.
14. H. S. WALL, *Concerning harmonic matrices*, Arch. Math., vol. 5 (1954), pp. 160–167

UNIVERSITY OF MICHIGAN
ANN ARBOR, MICHIGAN