Chapter 3

Lebesgue product integration

While it is sufficient to use the Riemann integral in applications, it is rather unsatisfactory from the viewpoint of theoretical mathematics. The generalization of Riemann integral due to Henri Lebesgue is based on the notion of measure. The problem of extending Volterra's definition of product integral in a similar way has been solved by Ludwig Schlesinger.

Volterra's and Schlesinger's works differ in yet another way: Volterra did not worry about using infinitesimal quantities, and it is not always easy to translate his ideas into the language of modern mathematics. Schlesinger's proofs are rather precise and can be read without greater effort except for occasionally strange notation. The foundations of mathematical analysis in 1930's were firmer than in 1887; moreover, Schlesinger inclined towards theoretical mathematics, as opposed to Volterra, who always kept applications in mind.

Ludwig Schlesinger[1]

Schlesinger's biographies can be found in [Lex, McT]: Ludwig (Lajos in Hungarian) Schlesinger was born on the 1st November 1864 in a Hungarian town Trnava (Nagyszombat), which now belongs to Slovakia. He studied mathematics and physics at the universities of Heidelberg and Berlin, where he received a doctorate in 1887.

[1] Photo from [McT]

The advisors of his thesis (which was concerned with homogeneous linear differential equations of the fourth order) were Lazarus Fuchs (who later became his father-in-law) and Leopold Kronecker. Two years later Schlesinger became an associate professor in Berlin and in 1897 an invited professor at the University of Bonn. During the years 1897 to 1911 he served as an ordinary professor and also as the head of the department of higher mathematics at the University of Kolozsvár (now Cluj in Romania). In 1911 he moved to Giessen in Germany where he continued to teach until his retirement in 1930. Ludwig Schlesinger died on the 16th December 1933.

Schlesinger devoted himself especially to complex function theory and linear differential equations; he also made valuable contributions to the history of mathematics. He translated Descartes' *Geometrie* into German, and was one of the organizers of the centenary festivities dedicated to the hundredth anniversary of János Bolyai, one of the pioneers of non-Euclidean geometry. The most important works of Schlesinger include *Handbuch der Theorie der linearen Differentialgleichungen* (1895–98), *J. Bolyai in Memoriam* (1902), *Vorlesungen über lineare Differentialgleichungen* (1908) and *Raum, Zeit und Relativitätstheorie* (1920).

Schlesinger's paper on product integration called *Neue Grundlagen für einen Infinitesimalkalkul der Matrizen* [LS1] was published in 1931. The author links up to Volterra's theory of product integral. He starts with the Riemann-type definition and establishes the basic properties of the product integral. His proofs are nevertheless original – while Volterra proved most of his statements using the Peano series expansion, Schlesinger prefers the "$\varepsilon - \delta$" proofs. He then proceeds to define the Lebesgue product integral (as a limit of product integrals of step functions) and explores its properties.

A continuation of this paper appeared in 1932 under the title *Weitere Beiträge zum Infinitesimalkalkul der Matrizen* [LS2]. Schlesinger again studies the properties of Lebesgue product integral and is also concerned with contour product integration in $\mathbf{R}^2$ and in $\mathbf{C}$.

This chapter summarizes the most important results from both Schlesinger's papers; the final section then presents a generalization of Schlesinger's definition of the Lebesgue product integral.

3.1 Riemann integrable matrix functions

When dealing with product integral we need to work with sequences of matrices and their limits. Volterra was mainly working with the individual entries of the matrices and convergence of a sequence of matrices was for him equivalent to convergence of all entries.

Schlesinger chooses a different approach: He defines the norm of a matrix $A = \{a_{ij}\}_{i,j=1}^n$ by

$$[A] = n \cdot \max_{1 \le i,j \le n} |a_{ij}|.$$

He also mentions another norm

$$\Omega_A = \max\{|\lambda|;\ \lambda \text{ is an eigenvalue of } A\}$$

and states that

$$\Omega_A \le \sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2} \le [A].$$

The second inequality is obvious, the first is proved in [LS1][1].

Schlesinger's norm $[A]$ has the nice property that $[A \cdot B] \le [A] \cdot [B]$ for every $A, B \in \mathbf{R}^{n\times n}$, but its disadvantage is that $[I] = n$. In the following text we will use the operator norm

$$\|A\| = \sup\{\|Ax\|;\ \|x\| \le 1\},$$

where $\|Ax\|$ and $\|x\|$ denote the Euclidean norms of vectors Ax, $x \in \mathbf{R}^n$. This simplifies Schlesinger's proofs slightly, because $\|I\| = 1$ and

$$\|A \cdot B\| \le \|A\| \cdot \|B\|$$

still holds for every $A, B \in \mathbf{R}^{n\times n}$. It should be noted that the space $\mathbf{R}^{n\times n}$ is finite-dimensional, therefore it doesn't matter which norm we choose since they are all equivalent.

The convergence of a sequence of matrices and the limit of a matrix function is now defined in a standard way using the norm introduced above.

For an arbitrary matrix function $A : [a, b] \to \mathbf{R}^{n\times n}$ and a tagged partition

$$D:\ a = t_0 \le \xi_1 \le t_1 \le \xi_2 \le \cdots \le t_{m-1} \le \xi_m \le t_m = b$$

of interval $[a, b]$ with division points t_i and tags ξ_i we denote

$$P(A, D) = \prod_{k=1}^{m} (I + A(\xi_k)\Delta t_k),$$

where $\Delta t_k = t_k - t_{k-1}$.

Schlesinger is now interested in the limit value of $P(A, D)$ as the lengths of the intervals $[t_{k-1}, t_k]$ approach zero (if the limit exists independently on the choice of $\xi_k \in [t_{k-1}, t_k]$). Clearly, the limit is nothing else than Volterra's right product integral.

Definition 3.1.1. Consider function $A : [a, b] \to \mathbf{R}^{n\times n}$. In case the limit

$$\lim_{\nu(D)\to 0} P(A, D)$$

exists, it is called the product integral of function A on interval $[a, b]$ and denoted by the symbol

$$(I + A(t)\,\mathrm{d}t) \prod_{a}^{b}.$$

---

[1] [LS1], p. 34–35

Remark 3.1.2. Schlesinger in fact defines the product integral as the limit of the products

$$P(A,D) = Y^0 \prod_{k=1}^{m} (I + A(\xi_k)\Delta t_k),$$

where Y^0 is an arbitrary regular matrix (which plays the role of an "integration constant"). In the following text we assume for simplicity that $Y^0 = I$. Also, instead of Schlesinger's notation

$$\widehat{\int_a^b} (I + A(x)\,\mathrm{d}x)$$

we use the symbol $(I + A(x)\,\mathrm{d}x) \prod_a^b$ to denote the product integral.

Lemma 3.1.3.[1] Let $A_1, A_2, \dots, A_m \in \mathbf{R}^{n\times n}$ be arbitrary matrices. Then

$$\|(I + A_1)(I + A_2)\cdots(I + A_m)\| \le \exp\left(\sum_{k=1}^{m} \|A_k\|\right).$$

Proof. A simple consequence of the inequalities

$$\|I + A_k\| \le 1 + \|A_k\| \le \exp\|A_k\|.$$

□

Corollary 3.1.4.[2] If $\|A(x)\| \le M$ for every $x \in [a,b]$, then $\|P(A,D)\| \le e^{M(b-a)}$ for every tagged partition D of interval $[a,b]$.

Corollary 3.1.5. If the function $A : [a,b] \to \mathbf{R}^{n\times n}$ is product integrable and $\|A(x)\| \le M$ for every $x \in [a,b]$, then

$$\left\| (I + A(x)\,\mathrm{d}x) \prod_a^b \right\| \le e^{M(b-a)}.$$

Schlesinger's first task is to prove the existence of product integral for Riemann integrable matrix functions, i.e. functions $A : [a,b] \to \mathbf{R}^{n\times n}$ whose entries a_{ij} are Riemann integrable on $[a,b]$. The proof is substantially different from the proof given by Volterra; the technique is similar to Cauchy's proof of the existence of $\int_a^b f$ for a continuous function f (see [CE, SŠ]).

Definition 3.1.6. Consider function $A : [a,b] \to \mathbf{R}^{n\times n}$ and let $[c,d] \subseteq [a,b]$. The oscillation of A on interval $[c,d]$ is the number

$$\operatorname{osc}(A,[c,d]) = \sup\{\|A(\xi_1) - A(\xi_2)\|;\ \ \xi_1, \xi_2 \in [c,d]\}.$$

---

[1] [LS1], p. 37

[2] [LS1], p. 38

The following characterization of Riemann integrable function will be needed in subsequent proofs:

Lemma 3.1.7. If $A : [a, b] \to \mathbf{R}^{n\times n}$ is a Riemann integrable function, then

$$\lim_{\nu(D)\to 0} \sum_{k=1}^{m} \operatorname{osc}(A, [t_{k-1}, t_k])\Delta t_k = 0.$$

Proof. The statement follows easily from Darboux's definition of the Riemann integral which is based on upper and lower sums; it is in fact equivalent to Riemann integrability of the given function (see e.g. [Sch2]). □

Definition 3.1.8. We say that a tagged partition D' is a refinement of a tagged partition D (we write $D' \prec D$), if every division point of D is also a division point of D' (no condition being imposed on the tags).

Lemma 3.1.9.[1] Let the function $A : [a, b] \to \mathbf{R}^{n\times n}$ be such that $\|A(x)\| \le M$ for every $x \in [a, b]$. Then for every pair of tagged partitions D, D' of interval $[a, b]$ such that $D' \prec D$ we have

$$\|P(A, D) - P(A, D')\| \le e^{M(b-a)} \sum_{k=1}^{m} (\operatorname{osc}(A, [t_{k-1}, t_k])\Delta t_k + (M\Delta t_k)^2 e^{M\Delta t_k}),$$

where t_i, $i = 0, \dots, m$ are division points of the partition D.

Proof. Let the partition D consist of division points and tags

$$D:\ a = t_0 \le \xi_1 \le t_1 \le \xi_2 \le \dots \le t_{m-1} \le \xi_m \le t_m = b.$$

First, we refine it only on the subinterval $[t_{k-1}, t_k]$, i.e. we consider a partition D^* which contains division points and tags

$$t_{k-1} = u_0 \le \eta_1 \le u_1 \dots \le u_{l-1} \le \eta_l \le u_l = t_k$$

and coincides with the partition D on the rest of interval $[a, b]$. Then

$$\|P(A, D^*) - P(A, D)\| \le \left\| \prod_{i=1}^{k-1} (I + A(\xi_i)\Delta t_i) \right\| \cdot$$

$$\cdot \left\| \prod_{j=1}^{l} (I + A(\eta_j)\Delta u_j) - I - A(\xi_k)\Delta t_k \right\| \cdot \left\| \prod_{i=k+1}^{m} (I + A(\xi_i)\Delta t_i) \right\|.$$

We estimate

$$\left\| \prod_{i=1}^{k-1} (I + A(\xi_i)\Delta t_i) \right\| \cdot \left\| \prod_{i=k+1}^{m} (I + A(\xi_i)\Delta t_i) \right\| \le e^{M(b-a)}$$

[1] [LS1], p. 39–41

and

$$\left\|\prod_{j=1}^{l}(I+A(\eta_j)\Delta u_j)-I-A(\xi_k)\Delta t_k\right\| \le \left\|\sum_{j=1}^{l}(A(\eta_j)-A(\xi_k))\Delta u_j\right\|+$$

$$+\left\|\sum_{p=2}^{l}\sum_{1\le r_1<\cdots<r_p\le l}A(\eta_{r_1})\cdots A(\eta_{r_p})\Delta u_{r_1}\cdots\Delta u_{r_p}\right\|\le$$

$$\le \operatorname{osc}(A,[t_{k-1},t_k])\Delta t_k+\sum_{p=2}^{l}\sum_{1\le r_1<\cdots<r_p\le l}M^p\,\Delta u_{r_1}\cdots\Delta u_{r_p}=$$

$$=\operatorname{osc}(A,[t_{k-1},t_k])\Delta t_k+\prod_{j=1}^{l}(1+M\Delta u_j)-1-\sum_{j=1}^{l}M\Delta u_j\le$$

$$\le \operatorname{osc}(A,[t_{k-1},t_k])\Delta t_k+e^{M\Delta t_k}-1-M\Delta t_k\le \operatorname{osc}(A,[t_{k-1},t_k])\Delta t_k+(M\Delta t_k)^2e^{M\Delta t_k}.$$

Therefore we conclude that

$$\|P(A,D)-P(A,D^*)\|\le e^{M(b-a)}(\operatorname{osc}(A,[t_{k-1},t_k])\Delta t_k+(M\Delta t_k)^2e^{M\Delta t_k}).$$

Now, since the given partition D' can be obtained from D by successively refining the subintervals $[t_0,t_1],\ldots,[t_{m-1},t_m]$, we obtain

$$\|P(A,D)-P(A,D')\|\le e^{M(b-a)}\sum_{k=1}^{m}(\operatorname{osc}(A,[t_{k-1},t_k])\Delta t_k+(M\Delta t_k)^2e^{M\Delta t_k}).$$

□

Corollary 3.1.10.[1] Consider a Riemann integrable function $A:[a,b]\to\mathbf{R}^{n\times n}$. Then for every $\varepsilon>0$ there exists $\delta>0$ such that

$$\|P(A,D)-P(A,D')\|<\varepsilon$$

whenever $\nu(D)<\delta$ and $D'\prec D$.

Proof. The statement follows from the previous lemma, Lemma 3.1.7 and the estimate

$$\sum_{k=1}^{m}(M\Delta t_k)^2e^{M\Delta t_k}\le\nu(D)M^2e^{M\nu(D)}\sum_{k=1}^{m}\Delta t_k=(b-a)\nu(D)M^2e^{M\nu(D)}.$$

□

[1] [LS1], p. 39–41

Theorem 3.1.11.[1] The product integral $(I+A(x)\,\mathrm{d}x)\prod_a^b$ exists for every Riemann integrable function $A:[a,b]\to\mathbf{R}^{n\times n}$.

Proof. Take $\varepsilon>0$. Corollary 3.1.10 guarantees the existence of a $\delta>0$ such that

$$\|P(A,D)-P(A,D')\|<\varepsilon/2$$

whenever $\nu(D)<\delta$ and $D'\prec D$. Consider a pair of tagged partitions D_1, D_2 of interval $[a,b]$ satisfying $\nu(D_1)<\delta$ and $\nu(D_2)<\delta$. These partitions have a common refinement, i.e. a partition D such that $D\prec D_1$, $D\prec D_2$ (the tags in D can be chosen arbitrarily). Then

$$\|P(A,D_1)-P(A,D_2)\|\le\|P(A,D_1)-P(A,D)\|+\|P(A,D)-P(A,D_2)\|<\varepsilon.$$

We have proved that every Riemann integrable function $A:[a,b]\to\mathbf{R}^{n\times n}$ satisfies a certain Cauchy condition and this is also the end of Schlesinger's proof; the existence of product integral follows from the Cauchy condition in the same way as in the analoguous theorem for the ordinary Riemann integral (see e.g. [Sch2]). □

Theorem 3.1.12.[2] Consider a Riemann integrable function $A:[a,b]\to\mathbf{R}^{n\times n}$. If $c\in[a,b]$, then

$$(I+A(x)\,\mathrm{d}x)\prod_a^b=(I+A(x)\,\mathrm{d}x)\prod_a^c\cdot(I+A(x)\,\mathrm{d}x)\prod_c^b.$$

Proof. As Schlesinger remarks, the proof follows directly from the definition of product integral (see the proof in Chapter 2). □

3.2 Matrix exponential function

Let $A:[a,b]\to\mathbf{R}^{n\times n}$ be a constant function. If D_m is a partition of $[a,b]$ to m subintervals of length $(b-a)/m$, then

$$P(A,D_m)=\left(I+\frac{b-a}{m}A\right)^m.$$

Since $\nu(D_m)\to 0$ as $m\to\infty$, we have

$$(I+A(x)\,\mathrm{d}x)\prod_a^b=\lim_{m\to\infty}\left(I+\frac{b-a}{m}A\right)^m=e^{(b-a)A}.$$

The last equality follows from the fact that $e^A=\lim_{m\to\infty}(I+A/m)^m$ for every $A\in\mathbf{R}^{n\times n}$; recall that the matrix exponential was defined in Chapter 2 using the series

$$e^A=\sum_{m=0}^{\infty}\frac{A^m}{m!}. \tag{3.2.1}$$

1 [LS1], p. 41
2 [LS1], p. 41

Lemma 3.2.1. If $A_1, \dots, A_m \in \mathbf{R}^{n\times n}$ and $B_1, \dots, B_m \in \mathbf{R}^{n\times n}$, then

$$\prod_{i=1}^{m} A_i - \prod_{i=1}^{m} B_i = \sum_{i=1}^{m} \left(\prod_{j=1}^{i-1} B_j \cdot (A_i - B_i) \cdot \prod_{j=i+1}^{m} A_j \right).$$

Proof.

$$\prod_{i=1}^{m} A_i - \prod_{i=1}^{m} B_i = \sum_{i=1}^{m} (B_1 \cdots B_{i-1} A_i \cdots A_m - B_1 \cdots B_i A_{i+1} \cdots A_m) =$$

$$= \sum_{i=1}^{m} \left(\prod_{j=1}^{i-1} B_j \cdot (A_i - B_i) \cdot \prod_{j=i+1}^{m} A_j \right).$$

□

Theorem 3.2.2.[1] Consider a Riemann integrable function $A : [a,b] \to \mathbf{R}^{n\times n}$. Then

$$\lim_{\nu(D)\to 0} \prod_{k=1}^{m} e^{A(\xi_k)\Delta t_k} = \lim_{\nu(D)\to 0} \prod_{k=1}^{m} (I + A(\xi_k)\Delta t_k) = (I + A(t)\,\mathrm{d}t) \prod_{a}^{b}.$$

Proof. Since every Riemann integrable function is bounded, we have $\|A(x)\| \le M$ for some $M \in \mathbf{R}$ and for every $x \in [a,b]$. The definition of matrix exponential (3.2.1) implies

$$\left\| e^{A(\xi_k)\Delta t_k} - (I + A(\xi_k)\Delta t_k) \right\| \le (\|A(\xi_k)\|\Delta t_k)^2 e^{\|A(\xi_k)\|\Delta t_k} \le (M\Delta t_k)^2 e^{M\Delta t_k}$$

for $k = 1, \dots, m$. According to Lemma 3.2.1,

$$\left\| \prod_{k=1}^{m} e^{A(\xi_k)\Delta t_k} - \prod_{k=1}^{m} (I + A(\xi_k)\Delta t_k) \right\| =$$

$$= \left\| \sum_{j=1}^{m} \left(\prod_{k=1}^{j-1} (I + A(\xi_k)\Delta t_k) \cdot (e^{A(\xi_j)\Delta t_j} - I - A(\xi_j)\Delta t_j) \cdot \prod_{k=j+1}^{m} e^{A(\xi_k)\Delta t_k} \right) \right\| \le$$

$$\le e^{M(b-a)} \sum_{j=1}^{m} \left\| e^{A(\xi_j)\Delta t_j} - I - A(\xi_j)\Delta t_j \right\| \le e^{M(b-a)} M^2 \sum_{j=1}^{m} (\Delta t_j)^2 e^{M\Delta t_j} \le$$

$$\le e^{M(b-a)} M^2 \nu(D) e^{M\nu(D)} \sum_{j=1}^{m} \Delta t_j = (b-a) e^{M(b-a)} M^2 \nu(D) e^{M\nu(D)}.$$

[1] [LS1], p. 42

By choosing a sufficiently fine partition D of $[a,b]$, the last expression can be made arbitrarily small. □

Definition 3.2.3. The trace of a matrix $A=\{a_{ij}\}_{i,j=1}^n$ is the number

$$\operatorname{Tr} A=\sum_{i=1}^n a_{ii}.$$

Theorem 3.2.4.[1] If $A:[a,b]\to\mathbf{R}^{n\times n}$ is a Riemann integrable function, then

$$\det\left((I+A(x)\,\mathrm{d}x)\prod_a^b\right)=\exp\left(\int_a^b \operatorname{Tr} A(x)\,\mathrm{d}x\right).$$

Proof.

$$\det\left((I+A(x)\,\mathrm{d}x)\prod_a^b\right)=\det\left(\lim_{\nu(D)\to 0}\prod_{k=1}^m e^{A(\xi_k)\Delta t_k}\right)=\lim_{\nu(D)\to 0}\prod_{k=1}^m \det\, e^{A(\xi_k)\Delta t_k}=$$

$$=\lim_{\nu(D)\to 0}\prod_{k=1}^m e^{\operatorname{Tr} A(\xi_k)\Delta t_k}=\lim_{\nu(D)\to 0}\exp\left(\sum_{k=1}^m \operatorname{Tr} A(\xi_k)\Delta t_k\right)=\exp\left(\int_a^b \operatorname{Tr} A(x)\,\mathrm{d}x\right)$$

(we have used a theorem from linear algebra: $\det\exp A=\exp\operatorname{Tr} A$). □

Remark 3.2.5. This formula (sometimes called the Jacobi formula) appeared already in Volterra's work. Schlesinger employs a different proof and his statement is also more general – it requires only the Riemann integrability of A, in contrast to Volterra's assumption that A is continuous.

Corollary 3.2.6. If $A:[a,b]\to\mathbf{R}^{n\times n}$ is a Riemann integrable function, then the product integral $(I+A(x)\,\mathrm{d}x)\prod_a^b$ is a regular matrix.

Recall that Volterra has also assigned meaning to product integrals whose lower limit is greater than the upper limit; his definition for the right integral was

$$(I+A(t)\,\mathrm{d}t)\prod_b^a=\lim_{\nu(D)\to 0}\prod_{k=m}^1 (I-A(\xi_k)\Delta t_k).$$

If A is Riemann integrable, we know that this is equivalent to

$$(I+A(t)\,\mathrm{d}t)\prod_b^a=\lim_{\nu(D)\to 0}\prod_{k=m}^1 e^{-A(\xi_k)\Delta t_k}.$$

[1] [LS1], p. 43–44

Thus

$$I = \lim_{\nu(D)\to 0} \left(\prod_{k=1}^{m} e^{A(\xi_k)\Delta t_k} \cdot \prod_{k=m}^{1} e^{-A(\xi_k)\Delta t_k} \right) =$$

$$= \lim_{\nu(D)\to 0} \prod_{k=1}^{m} e^{A(\xi_k)\Delta t_k} \cdot \lim_{\nu(D)\to 0} \prod_{k=m}^{1} e^{-A(\xi_k)\Delta t_k} =$$

$$(I + A(t)\,\mathrm{d}t) \prod_{a}^{b} \cdot (I + A(t)\,\mathrm{d}t) \prod_{b}^{a},$$

which proves that $(I+A(t)\,\mathrm{d}t)\prod_b^a$ is the inverse matrix of $(I+A(t)\,\mathrm{d}t)\prod_a^b$; compare with Volterra's proof of Theorem 2.4.10.

3.3 The indefinite product integral

Schlesinger now proceeds to study the properties of the indefinite product integral, i.e. of the function $Y(x) = (I + A(t)\,\mathrm{d}t)\prod_a^x$.

Theorem 3.3.1.[1] If $A : [a, b] \to \mathbf{R}^{n\times n}$ is Riemann integrable, then the function $Y(x) = (I + A(t)\,\mathrm{d}t)\prod_a^x$ is continuous on $[a, b]$.

Proof. We prove the right-continuity of Y at $x_0 \in [a, b)$; continuity from left is proved similarly. Let $x_0 \le x_0 + h \le b$. The function A is bounded: $\|A(x)\| \le M$ for some $M \in \mathbf{R}$. We now employ the inequality from Lemma 3.1.9. Let D' be a partition of interval $[x_0, x_0 + h]$. Then

$$\|I + A(x_0)h - P(A, D')\| \le e^{Mh}(\mathrm{osc}(A, [x_0, x_0 + h])h + (Mh)^2 e^{Mh}).$$

Passing to the limit $\nu(D') \to 0$ we obtain

$$\left\| I + A(x_0)h - (I + A(t)\,\mathrm{d}t) \prod_{x_0}^{x_0+h} \right\| \le e^{Mh}(\mathrm{osc}(A, [x_0, x_0 + h])h + (Mh)^2 e^{Mh}),$$

which implies

$$\lim_{h\to 0+} (I + A(t)\,\mathrm{d}t) \prod_{x_0}^{x_0+h} = I.$$

Therefore

$$\lim_{h\to 0+} (Y(x_0 + h) - Y(x_0)) = Y(x_0) \left((I + A(t)\,\mathrm{d}t) \prod_{x_0}^{x_0+h} - I \right) = 0.$$

□

[1] [LS1], p. 44–46

Theorem 3.3.2.[1] If $A : [a,b] \to \mathbf{R}^{n\times n}$ is Riemann integrable, then the function

$$Y(x) = (I + A(t)\,\mathrm{d}t)\prod_a^x,$$

satisfies the integral equation

$$Y(x) = I + \int_a^x Y(t)A(t)\,\mathrm{d}t, \quad x \in [a,b].$$

Proof. It is sufficient to prove the statement for $x = b$. Let

$$D: \ a = t_0 \le \xi_1 \le t_1 \le \xi_2 \le \cdots \le t_{m-1} \le \xi_m \le t_m = b$$

be a tagged partition of interval $[a,b]$. We define

$$Y^k = \prod_{i=1}^k (I + A(\xi_i)\Delta t_i), \quad k = 0,\ldots,m.$$

Then

$$Y^k - Y^{k-1} = Y^{k-1}A(\xi_k)\Delta t_k, \quad k = 1,\ldots,m. \tag{3.3.1}$$

Since $Y^0 = I$ and $Y^m = P(A,D)$, adding the equalities (3.3.1) for $k = 1,\ldots,m$ yields

$$P(A,D) - I = \sum_{k=1}^m Y^{k-1}A(\xi_k)\Delta t_k.$$

The function A is bounded: $\|A(x)\| \le M$ for some $M \in \mathbf{R}$. We estimate

$$\left\| Y(b) - I - \int_a^b Y(t)A(t)\,\mathrm{d}t \right\| \le \|Y(b) - P(A,D)\| +$$

$$+ \left\| P(A,D) - I - \int_a^b Y(t)A(t)\,\mathrm{d}t \right\| \le \|Y(b) - P(A,D)\| +$$

$$+ \left\| \sum_{k=1}^m (Y^{k-1} - Y(t_{k-1}))A(\xi_k)\Delta t_k \right\| + \left\| \sum_{k=1}^m (Y(t_{k-1}) - Y(\xi_k))A(\xi_k)\Delta t_k \right\| +$$

$$+ \left\| \sum_{k=1}^m Y(\xi_k)A(\xi_k)\Delta t_k - \int_a^b Y(t)A(t)\,\mathrm{d}t \right\|. \tag{3.3.2}$$

[1] [LS1], p. 46–47

Using the inequalities

$$\left\|\sum_{k=1}^{m}(Y^{k-1}-Y(t_{k-1}))A(\xi_k)\Delta t_k\right\| \leq M\sum_{k=1}^{m}\|Y^{k-1}-Y(t_{k-1})\|\Delta t_k \leq$$

$$\leq M\sum_{k=1}^{m}e^{M(b-a)}\Delta t_k\left(\sum_{j=1}^{m}(\operatorname{osc}(A,[t_{j-1},t_j])\Delta t_j+(M\Delta t_j)^2e^{M\Delta t_j})\right) \leq$$

$$\leq Me^{M(b-a)}(b-a)\left(\sum_{j=1}^{m}\operatorname{osc}(A,[t_{j-1},t_j])\Delta t_j+M^2\nu(D)e^{M\nu(D)}\right)$$

(we have used Lemma 3.1.9) and

$$\left\|\sum_{k=1}^{m}(Y(t_{k-1})-Y(\xi_k))A(\xi_k)\Delta t_k\right\| \leq M\sum_{k=1}^{m}\operatorname{osc}(Y,[t_{k-1},t_k])\Delta t_k,$$

we see that all terms on the right-hand side of (3.3.2) can be made arbitrarily small if the partition D is sufficiently fine. □

Corollary 3.3.3.[1] If $A:[a,b]\to\mathbf{R}^{n\times n}$ is continuous, then the function

$$Y(x)=(I+A(t)\,\mathrm{d}t)\prod_a^x$$

provides a solution of the differential equation

$$Y'(x)=Y(x)A(x),\quad x\in[a,b]$$

and satisfies the initial condition $Y(a)=I$.

Remark 3.3.4. The function Y is therefore the fundamental matrix of the system

$$y_i'(x)=\sum_{j=1}^{n}a_{ji}(x)y_j(x),\quad i=1,\ldots,n.$$

Schlesinger uses the notation $D_xY(x)=A(x)$, where

$$D_xY=Y^{-1}Y',$$

i.e. D_x is exactly Volterra's right derivative of a matrix function.

3.4 Product integral inequalities

In this section we summarize various inequalities that will be useful later.

[1] [LS1], p. 47–48

Lemma 3.4.1.[1] If $A : [a, b] \to \mathbf{R}^{n\times n}$ is a Riemann integrable function, then

$$\left\| (I + A(x)\,\mathrm{d}x) \prod_{a}^{b} \right\| \leq \exp\left(\int_a^b \|A(x)\|\,\mathrm{d}x \right).$$

Proof. Lemma 3.1.3 implies that

$$\left\| \prod_{i=1}^{m} (I + A(\xi_i)\Delta t_i) \right\| \leq \exp\left(\sum_{i=1}^{m} \|A(\xi_i)\|\Delta t_i \right)$$

for every tagged partition D of interval $[a, b]$; the proof is completed by passing to the limit $\nu(D) \to 0$. □

Lemma 3.4.2.[2] Let $m \in \mathbf{N}$, $A_k, B_k \in \mathbf{R}^{n\times n}$ for every $k = 1, \dots, m$. Then

$$\left\| \prod_{k=1}^{m} (I + B_k) - \prod_{k=1}^{m} (I + A_k) \right\| \leq \exp\left(\sum_{k=1}^{m} \|A_k\| \right) \left(\exp\left(\sum_{k=1}^{m} \|B_k - A_k\| \right) - 1 \right).$$

Proof. Define

$$Y^k = \prod_{i=1}^{k} (I + A_i), \quad Z^k = \prod_{i=1}^{k} (I + B_i), \quad k = 0, \dots, m$$

(where the empty product for $k = 0$ equals the identity matrix). Then

$$Y^k - Y^{k-1} = Y^{k-1} A_k,$$

$$Z^k - Z^{k-1} = Z^{k-1} B_k,$$

for $k = 1, \dots, m$. This implies

$$Z^k - Y^k = (Z^{k-1} - Y^{k-1})(I + B_k) + E_k, \tag{3.4.1}$$

where

$$E_k = Y^{k-1}(B_k - A_k).$$

Applying the equality (3.4.1) m times on the difference $Z^m - Y^m$ we obtain

$$Z^m - Y^m = \sum_{k=1}^{m-1} E_k (I + B_{k+1}) \cdots (I + B_m) + E_m.$$

---

[1] [LS1], p. 51

[2] [LS1], p. 52–53

We also estimate

$$\|E_k\| \le \exp\left(\sum_{i=1}^{k-1} \|A_i\|\right) \|B_k - A_k\|$$

(the empty sum for $k = 0$ equals zero),

$$\|Z^m - Y^m\| \le \sum_{k=1}^{m-1} \exp\left(\sum_{i=1}^{k-1} \|A_i\|\right) \|B_k - A_k\| \cdot \exp\left(\sum_{i=k+1}^{m} (\|B_i - A_i\| + \|A_i\|)\right) +$$

$$+ \exp\left(\sum_{i=1}^{m-1} \|A_i\|\right) \|B_m - A_m\| =$$

$$= \sum_{k=1}^{m-1} \exp\left(\sum_{i \neq k} \|A_i\|\right) \|B_k - A_k\| \cdot \exp\left(\sum_{i=k+1}^{m} \|B_i - A_i\|\right) +$$

$$+ \exp\left(\sum_{i=1}^{m-1} \|A_i\|\right) \|B_m - A_m\| \le$$

$$\le \sum_{k=1}^{m} \exp\left(\sum_{i=1}^{m} \|A_i\|\right) \|B_k - A_k\| \cdot \exp\left(\sum_{i=k+1}^{m} \|B_i - A_i\|\right).$$

Since

$$\|B_k - A_k\| \le \exp(\|B_k - A_k\|) - 1,$$

we conclude that

$$\left\|\prod_{k=1}^{m}(I + B_k) - \prod_{k=1}^{m}(I + A_k)\right\| = \|Z^m - Y^m\| \le$$

$$\le \exp\left(\sum_{i=1}^{m} \|A_i\|\right) \sum_{k=1}^{m} \left((\exp(\|B_k - A_k\|) - 1) \exp\left(\sum_{i=k+1}^{m} \|B_i - A_i\|\right)\right) =$$

$$= \exp\left(\sum_{i=1}^{m} \|A_i\|\right) \sum_{k=1}^{m} \left(\exp\left(\sum_{i=k}^{m} \|B_i - A_i\|\right) - \exp\left(\sum_{i=k+1}^{m} \|B_i - A_i\|\right)\right) =$$

$$= \exp\left(\sum_{i=1}^{m} \|A_i\|\right) \left(\exp\left(\sum_{i=1}^{m} \|B_i - A_i\|\right) - 1\right).$$

□

Corollary 3.4.3.[1] If A, $B : [a, b] \to \mathbf{R}^{n \times n}$ are Riemann integrable functions, then

$$\left\|(I + B(x)\,\mathrm{d}x) \prod_{a}^{b} - (I + A(x)\,\mathrm{d}x) \prod_{a}^{b}\right\| \le$$

[1] [LS1], p. 53

$$\leq \exp\left(\int_a^b \|A(x)\|\,\mathrm{d}x\right)\left(\exp\left(\int_a^b \|B(x)-A(x)\|\,\mathrm{d}x\right)-1\right).$$

Proof. The previous lemma ensures that for every tagged partition D of interval $[a,b]$ we have

$$\|P(B,D)-P(A,D)\| = \left\|\prod_{k=1}^m (I+B(\xi_k)\Delta t_k) - \prod_{k=1}^m (I+A(\xi_k)\Delta t_k)\right\| \leq$$

$$\leq \exp\left(\sum_{k=1}^m \|A(\xi_k)\|\Delta t_k\right)\left(\exp\left(\sum_{k=1}^m \|B(\xi_k)-A(\xi_k)\|\Delta t_k\right)-1\right).$$

The proof is completed by passing to the limit $\nu(D)\to 0$. □

Remark 3.4.4. Lemma 3.4.2 is not present in Schlesinger's work, he proves directly the Corollary 3.4.3; our presentation is perhaps more readable.

3.5 Lebesgue product integral

The most valuable contribution of Schlesinger's paper is his generalized definition of product integral which is applicable to all matrix functions with bounded and measurable (i.e. bounded Lebesgue integrable) entries.

From a historical point of view, such a generalization certainly wasn't a straightforward one. Recall the original Lebesgue's definition: To compute the integral $\int_a^b f$ of a bounded measurable function $f:[a,b]\to[m,M]$, we choose a partition

$$D:\ m = m_0 < m_1 < \cdots < m_p = M,$$

then form the sets

$$E_0 = \{x\in[a,b];\ f(x)=m\},$$
$$E_j = \{x\in[a,b];\ m_{j-1} < f(x) \leq m_j\},\quad j=1,\ldots,p,$$

and compute the lower and upper sums

$$s(f,D) = m_0\mu_0 + \sum_{j=1}^p m_{j-1}\mu_j,\quad S(f,D) = m_0\mu_0 + \sum_{j=1}^p m_j\mu_j, \tag{3.5.1}$$

where $\mu_j = \mu(E_j)$ is the Lebesgue measure of the set E_j. Since

$$S(f,D)-s(f,D) = \sum_{j=1}^p (m_j - m_{j-1})\mu_j \leq \nu(D)(b-a),$$

the sums in (3.5.1) approach a common limit as $\nu(D) \to 0$ and we define

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{\nu(D)\to 0} s(f,D) = \lim_{\nu(D)\to 0} S(f,D).$$

Similar procedure cannot be used to define product integral of a matrix function $A : [a,b] \to \mathbf{R}^{n\times n}$, because $\mathbf{R}^{n\times n}$ is not an ordered set. Schlesinger was instead inspired by an equivalent definition of Lebesgue integral which is due to Friedrich Riesz (see [FR, KZ]): A bounded function $f : [a,b] \to \mathbf{R}$ is integrable, if and only if there exists a uniformly bounded sequence of step (i.e. piecewise-constant) functions $\{f_n\}_{n=1}^\infty$ such that $f_n \to f$ almost everywhere on $[a,b]$; in this case,

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x.$$

To proceed to the definition of product integral we first recall that (see Theorem 3.2.2)

$$(I + A(x)\,\mathrm{d}x)\prod_a^b = \lim_{\nu(D)\to 0} \prod_{k=1}^m e^{A(\xi_k)\Delta t_k}$$

for every Riemann integrable function $A : [a,b] \to \mathbf{R}^{n\times n}$. The product on the right side might be interpreted as

$$\prod_{k=1}^m e^{A(\xi_k)\Delta t_k} = (I + A_D(t)\,\mathrm{d}t)\prod_a^b,$$

where A_D is a step function defined by

$$A_D(t) = A(\xi_k), \quad t \in (t_{k-1}, t_k)$$

(the values $A(t_k)$, $k = 0,\ldots,m$, might be chosen arbitrarily). If $\{D_k\}_{k=1}^\infty$ is a sequence of tagged partitions of $[a,b]$ such that $\lim_{k\to\infty} \nu(D_k) = 0$, it is easily proved that

$$\lim_{k\to\infty} A_{D_k}(t) = A(t) \tag{3.5.2}$$

at every point $t \in [a,b]$ at which A is continuous. Since Riemann integrable functions are continuous almost everywhere, the Equation (3.5.2) holds a.e. on $[a,b]$. We are therefore led to the following generalized definition of product integral:

$$(I + A(x)\,\mathrm{d}x)\prod_a^b = \lim_{k\to\infty} (I + A_k(x)\,\mathrm{d}x)\prod_a^b,$$

where $\{A_k\}_{k=1}^\infty$ is a suitably chosen sequence of matrix step functions that converge to A almost everywhere.

Definition 3.5.1. A function $A : [a,b] \to \mathbf{R}^{n\times n}$ is called a step function if there exist numbers

$$a = t_0 < t_1 < \cdots < t_m = b$$

such that A is a constant function on every interval (t_{k-1}, t_k), $k = 1, \dots, m$.

Clearly, a matrix function $A = \{a_{ij}\}_{i,j=1}^n$ is a step function if and only if all the entries a_{ij} are step functions.

Definition 3.5.2. A sequence of functions $A_k : [a, b] \to \mathbf{R}^{n\times n}$, $k \in \mathbf{N}$, is called uniformly bounded if there exists a number $M \in \mathbf{R}$ such that $\|A_k(x)\| \le M$ for every $k \in \mathbf{N}$ and every $x \in [a, b]$.

Definition 3.5.3. A function $A : [a, b] \to \mathbf{R}^{n\times n}$ is called measurable if all the entries a_{ij} are measurable functions.

Lemma 3.5.4. Let $A_k : [a, b] \to \mathbf{R}^{n\times n}$, $k \in \mathbf{N}$, be a uniformly bounded sequence of measurable functions such that

$$\lim_{k\to\infty} A_k(x) = A(x)$$

a.e. on $[a, b]$. Then $A_k \to A$ in the norm of the space L^1, i.e.

$$\lim_{k\to\infty} \int_a^b \|A_k(x) - A(x)\| \,\mathrm{d}x = 0.$$

Proof. Choose $\varepsilon > 0$. As $\|A_k(x)\| \le M$ for every $k \in \mathbf{N}$ and every $x \in [a, b]$, we can estimate

$$\int_a^b \|A_k(x) - A(x)\| \,\mathrm{d}x \le \varepsilon(b - a) + 2M\mu(\{x; \|A_k(x) - A(x)\| \ge \varepsilon\}).$$

The convergence $A_k \to A$ a.e. implies convergence in measure[1], i.e. for every $\varepsilon > 0$ we have

$$\lim_{k\to\infty} \mu(\{x; \|A(x) - A_k(x)\| \ge \varepsilon\}) = 0.$$

Therefore

$$\lim_{k\to\infty} \int_a^b \|A_k(x) - A(x)\| \,\mathrm{d}x \le \varepsilon(b - a)$$

for every $\varepsilon > 0$. □

Theorem 3.5.5.[2] Let $A_k : [a, b] \to \mathbf{R}^{n\times n}$, $k \in \mathbf{N}$, be a sequence of step functions such that

$$\lim_{k\to\infty} \int_a^b \|A_k(x) - A(x)\| \,\mathrm{d}x = 0.$$

Then the limit

$$\lim_{k\to\infty} (I + A_k(x)\,\mathrm{d}x)\prod_a^b$$

---

1 [IR], Proposition 8.3.3, p. 256

2 [LS1], p. 55–56

exists and is independent on the choice of the sequence $\{A_k\}_{k=1}^{\infty}$.

Proof. We verify that $(I + A_k(x)\,\mathrm{d}x)\prod_a^b$ is a Cauchy sequence. According to Corollary 3.4.3 we have

$$\left\| (I + A_l(x)\,\mathrm{d}x)\prod_a^b - (I + A_m(x)\,\mathrm{d}x)\prod_a^b \right\| \leq$$

$$\leq \exp\left(\int_a^b \|A_m(x)\|\,\mathrm{d}x\right)\left(\exp\left(\int_a^b \|A_l(x) - A_m(x)\|\,\mathrm{d}x\right) - 1\right).$$

The assumption of our theorem implies that the sequence of numbers $\int_a^b \|A_m(x)\|\,\mathrm{d}x$ is bounded and that

$$\lim_{l,m\to\infty} \int_a^b \|A_l(x) - A_m(x)\|\,\mathrm{d}x = 0,$$

which proves the existence of the limit. To verify the uniqueness consider two sequences of step functions $\{A_k\}$, $\{B_k\}$ that satisfy the assumption of the theorem. We construct a sequence $\{C_k\}$, where $C_{2k-1} = A_k$ and $C_{2k} = B_k$. Then $C_k \to A$ a.e. and

$$\lim_{k\to\infty} \int_a^b \|C_k(x) - A(x)\|\,\mathrm{d}x = 0,$$

which means that $\lim_{k\to\infty}(I + C_k(x)\,\mathrm{d}x)\prod_a^b$ exists. Every subsequence of $\{C_k\}$ must have the same limit, therefore

$$\lim_{k\to\infty}(I + A_k(x)\,\mathrm{d}x)\prod_a^b = \lim_{k\to\infty}(I + B_k(x)\,\mathrm{d}x)\prod_a^b.$$

□

Definition 3.5.6. Consider function $A : [a,b] \to \mathbf{R}^{n\times n}$. Assume there exists a uniformly bounded sequence of step functions $A_k : [a,b] \to \mathbf{R}^{n\times n}$ such that

$$\lim_{k\to\infty} A_k(x) = A(x)$$

a.e. on $[a,b]$. Then the function A is called product integrable and we define

$$(I + A(x)\,\mathrm{d}x)\prod_a^b = \lim_{k\to\infty}(I + A_k(x)\,\mathrm{d}x)\prod_a^b.$$

We use the symbol $L^*([a,b],\mathbf{R}^{n\times n})$ to denote the set of all product integrable functions.

Remark 3.5.7. The correctness of the previous definition is guaranteed by Lemma 3.5.4 and Theorem 3.5.5. Every function $A \in L^*([a,b],\mathbf{R}^{n\times n})$ is clearly bounded and measurable (step functions are measurable and the limit of measurable functions is again measurable). Assume on the contrary that $A : [a,b] \to \mathbf{R}^{n\times n}$ is a measurable function on $[a,b]$ such that

$$\|a_{ij}(x)\| \le M, \quad x \in [a,b], \;\; i,j = 1,\ldots,n.$$

There exists[1] a sequence of step functions $\{A_k\}_{k=1}^{\infty}$ which converge to A in the L^1 norm. This sequence contains[2] a subsequence $\{B_k\}_{k=1}^{\infty}$ of matrix functions $B_k = \{b_{ij}^k\}_{i,j=1}^n$ such that $B_k \to A$ a.e. on $[a,b]$. Without loss of generality we can assume that the sequence $\{B_k\}_{k=1}^{\infty}$ is uniformly bounded (otherwise consider the functions $\min(\max(-M, b_{ij}^k), M)$). We have thus found a uniformly bounded sequence of step functions which converge to A a.e. on $[a,b]$. This means that

$$L^*([a,b],\mathbf{R}^{n\times n}) = \left\{A : [a,b] \to \mathbf{R}^{n\times n};\ A \text{ is measurable and bounded}\right\}.$$

Schlesinger remarks that it is possible to further extend the definition of product integral to encompass all matrix functions with Lebesgue integrable (not necessarily bounded) entries, but he doesn't give any details. We return to this question at the end of the chapter.

3.6 Properties of Lebesgue product integral

After having defined the Lebesgue product integral in [LS1], Schlesinger carefully studies its properties. Interesting results may be found also in [LS2].

Lemma 3.6.1. Assume that $\{A_k\}_{k=1}^{\infty}$ is a uniformly bounded sequence of functions from $L^*([a,b],\mathbf{R}^{n\times n})$, and that $A_k \to A$ a. e. on $[a,b]$. Then

$$\int_a^b \|A(x)\|\,\mathrm{d}x = \lim_{k\to\infty} \int_a^b \|A_k(x)\|\,\mathrm{d}x.$$

Proof. According to the Lebesgue's dominated convergence theorem,

$$\lim_{k\to\infty} \int_a^b \|A_k(x)\|\,\mathrm{d}x = \int_a^b \lim_{k\to\infty} \|A_k(x)\|\,\mathrm{d}x = \int_a^b \|A(x)\|\,\mathrm{d}x$$

(we have used continuity of the norm). □

Corollary 3.6.2. Inequalities 3.4.1 and 3.4.3 are satisfied for all step functions. As a consequence of the previous lemma we see they are valid even for functions from $L^*([a,b],\mathbf{R}^{n\times n})$.

The next statement represents a dominated convergence theorem for the Lebesgue product integral.

---

[1] [RG], Corollary 3.29, p. 47

[2] [IR], Theorem 8.4.14, p. 267, and Theorem 8.3.6, p. 257

Theorem 3.6.3.[1] Assume that $\{A_k\}_{k=1}^{\infty}$ is a uniformly bounded sequence of functions from $L^*([a,b],\mathbf{R}^{n\times n})$ such that $A_k \to A$ a. e. on $[a,b]$. Then

$$(I + A(x)\,\mathrm{d}x)\prod_a^b = \lim_{k\to\infty}(I + A_k(x)\,\mathrm{d}x)\prod_a^b.$$

Proof. The function A is measurable and bounded, therefore $A \in L^*([a,b],\mathbf{R}^{n\times n})$. To complete the proof we use Corollary 3.4.3 in the form

$$\left\|(I + A(x)\,\mathrm{d}x)\prod_a^b - (I + A_k(x)\,\mathrm{d}x)\prod_a^b\right\| \le$$

$$\exp\left(\int_a^b \|A(x)\|\,\mathrm{d}x\right)\left(\exp\left(\int_a^b \|A_k(x) - A(x)\|\,\mathrm{d}x\right) - 1\right)$$

and Lemma 3.5.4. □

Remark 3.6.4. The previous theorem holds also for Riemann product integral in case we add an extra assumption that the limit function A is Riemann product integrable.

Definition 3.6.5. If M is a measurable subset of $[a,b]$ and $A \in L^*([a,b],\mathbf{R}^{n\times n})$, we define

$$(I + A(x)\,\mathrm{d}x)\prod_M = (I + \chi_M(x)A(x)\,\mathrm{d}x)\prod_a^b$$

(where χ_M is the characteristic function of the set M).

The previous definition is correct, because the product $\chi_M A$ is obviously a measurable bounded function.

Remark 3.6.6. The following theorem is proved in the theory of Lebesgue integral[2]: For every $f \in L^1([a,b])$ and every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left\|\int_M f(x)\,\mathrm{d}x\right\| < \varepsilon$$

whenever M is a measurable subset of $[a,b]$ and $\mu(M) < \delta$. Schlesinger proceeds to prove an analoguous theorem for the product integral (he speaks about "total continuity").

Theorem 3.6.7.[3] For every $A \in L^*([a,b],\mathbf{R}^{n\times n})$ and every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left\|(I + A(x)\,\mathrm{d}x)\prod_M - I\right\| < \varepsilon$$

---

[1] [LS1], p. 57–58
[2] [RG], theorem 3.26, p. 46
[3] [LS1], p. 59

whenever M is a measurable subset of $[a,b]$ and $\mu(M) < \delta$.

Proof. Substituting $B = 0$ to Corollary 3.4.3 we obtain

$$\left\| (I + A(x)\,\mathrm{d}x) \prod_M - I \right\| \le \exp\left(\int_M \|A(x)\|\,\mathrm{d}x \right) \left(\exp\left(\int_M \|A(x)\|\,\mathrm{d}x \right) - 1 \right),$$

which completes the proof (see Remark 3.6.6). □

Schlesinger now turns his attention to the indefinite product integral. Recall that if $f \in L^1([a,b])$, then the indefinite integral

$$F(x) = \int_a^x f(t)\,\mathrm{d}t, \quad x \in [a,b]$$

is an absolutely continuous function and $F'(x) = f(x)$ a. e. on $[a,b]$. Before looking at a product analogy of this theorem we state the following lemma.

Lemma 3.6.8. If $A \in L^*([a,b], \mathbf{R}^{n\times n})$, then

$$\lim_{h\to 0} \frac{1}{h} \int_x^{x+h} \|A(t) - A(x)\|\,\mathrm{d}t = 0$$

for almost all $x \in (a,b)$.

Proof. If $f \in L^1([a,b])$, then[1]

$$\lim_{h\to 0} \frac{1}{h} \int_0^h |f(x+t) - f(x)|\,\mathrm{d}t = 0$$

for almost all $x \in (a,b)$ (every such x is called the Lebesgue point of f). Applying this equality to the entries of A we obtain

$$\lim_{h\to 0} \frac{1}{h} \int_x^{x+h} \|A(t) - A(x)\|\,\mathrm{d}t = \lim_{h\to 0} \frac{1}{h} \int_0^h \|A(x+t) - A(x)\|\,\mathrm{d}t = 0$$

for almost all $x \in (a,b)$. □

Theorem 3.6.9.[2] If $A \in L^*([a,b], \mathbf{R}^{n\times n})$, then the indefinite integral

$$Y(x) = (I + A(t)\,\mathrm{d}t) \prod_a^x.$$

satisfies $Y^{-1}(x)Y'(x) = A(x)$ for almost all $x \in [a,b]$.

Proof. According to the definition of derivative,

$$Y^{-1}(x)Y'(x) = \lim_{h\to 0} \frac{Y^{-1}(x)Y(x+h) - I}{h}.$$

[1] [IR], Theorem 6.3.2, p. 194
[2] [LS1], p. 60–61

We now prove that

$$\lim_{h\to 0+}\frac{Y^{-1}(x)Y(x+h)-I}{h}=\lim_{h\to 0+}\frac{1}{h}\left((I+A(t)\,\mathrm{d}t)\prod_x^{x+h}-I\right)=A(x)\qquad(3.6.1)$$

for almost all $x\in[a,b]$; the procedure is similar for the limit from left. We estimate

$$\left\|\frac{1}{h}\left((I+A(t)\,\mathrm{d}t)\prod_x^{x+h}-I\right)-A(x)\right\|\leq\left\|\frac{1}{h}\left((I+A(t)\,\mathrm{d}t)\prod_x^{x+h}-e^{A(x)h}\right)\right\|+$$

$$+\left\|\frac{1}{h}\sum_{k=2}^{\infty}\frac{A^k(x)h^k}{k!}\right\|\leq\left\|\frac{1}{h}\left((I+A(t)\,\mathrm{d}t)\prod_x^{x+h}-e^{A(x)h}\right)\right\|+\|A(x)\|^2|h|e^{\|A(x)\|h}\qquad(3.6.2)$$

Since $\|A(x)\|\leq M$ for some $M\in\mathbf{R}$, the Corollary 3.4.3 yields

$$\left\|\frac{1}{h}\left((I+A(t)\,\mathrm{d}t)\prod_x^{x+h}-e^{A(x)h}\right)\right\|=\left\|\frac{1}{h}\left((I+A(t)\,\mathrm{d}t)\prod_x^{x+h}-(I+A(x)\,\mathrm{d}t)\prod_x^{x+h}\right)\right\|\leq$$

$$\leq\frac{1}{|h|}\exp\left(\int_x^{x+h}\|A(x)\|\,\mathrm{d}t\right)\left(\exp\left(\int_x^{x+h}\|A(t)-A(x)\|\,\mathrm{d}t\right)-1\right)=$$

$$=\exp(\|A(x)\|h)\,\frac{1}{|h|}\sum_{k=1}^{\infty}\frac{1}{k!}\left(\int_x^{x+h}\|A(t)-A(x)\|\,\mathrm{d}t\right)^k\leq$$

$$\leq\exp(Mh)\left(\frac{1}{|h|}\int_x^{x+h}\|A(t)-A(x)\|\,\mathrm{d}t+(2M)^2h\exp(2Mh)\right).\qquad(3.6.3)$$

Equations (3.6.2), (3.6.3), and Lemma 3.6.8 imply Equation (3.6.1). □

Remark 3.6.10. In the previous theorem we have tacitly assumed that the matrix

$$Y(x)=(I+A(t)\,\mathrm{d}t)\prod_a^x$$

is regular for every $x\in[a,b]$. Schlesinger proved it only for $A\in R([a,b],\mathbf{R}^{n\times n})$ (see Corollary 3.2.6), but the proof is easily adjusted to $A\in L^*([a,b],\mathbf{R}^{n\times n})$:
If $\{A_k\}_{k=1}^{\infty}$ is a uniformly bounded sequence of step functions such that $A_k\to A$ a. e. on $[a,b]$, then (using 3.2.4 and Lebesgue's dominated convergence theorem)

$$\det(I+A(t)\,\mathrm{d}t)\prod_a^b=\det\lim_{k\to\infty}(I+A_k(t)\,\mathrm{d}t)\prod_a^b=\lim_{k\to\infty}\det(I+A_k(t)\,\mathrm{d}t)\prod_a^b=$$

$$=\lim_{k\to\infty}\exp\left(\int_a^b\operatorname{Tr}A_k(t)\,\mathrm{d}t\right)=\exp\left(\int_a^b\operatorname{Tr}A(t)\,\mathrm{d}t\right)>0.$$

Theorem 3.6.11.[1] If $A \in L^*([a,b], \mathbf{R}^{n\times n})$, then

$$(I + A(x)\,\mathrm{d}x)\prod_a^b = I + \sum_{k=1}^{\infty} \int_a^b \int_a^{x_k} \cdots \int_a^{x_2} A(x_1)\cdots A(x_k)\,\mathrm{d}x_1 \cdots \mathrm{d}x_k$$

(where the integrals on the right side are taken in the sense of Lebesgue).

Proof. Let $\{A_k\}_{k=1}^{\infty}$ be a uniformly bounded sequence of step functions such that $A_k \to A$ a. e. on $[a,b]$. Every function A_k is associated with a partition

$$D_k : a = t_0^k < t_1^k < \cdots < t_{m(k)}^k = b$$

such that

$$A_k(x) = A_j^k, \quad x \in (t_{j-1}^k, t_j^k).$$

According to the definition of Lebesgue product integral,

$$(I + A(x)\,\mathrm{d}x)\prod_a^b = \lim_{k\to\infty} (I + A_k(x)\,\mathrm{d}x)\prod_a^b = \lim_{k\to\infty} \prod_{j=1}^{m(k)} \exp(A_j^k \Delta t_j^k).$$

Schlesinger proves[2] first that the product integral might be also calculated as

$$(I + A(x)\,\mathrm{d}x)\prod_a^b = \lim_{k\to\infty} \prod_{i=1}^{m(k)} (I + A_i^k \Delta t_i^k), \tag{3.6.4}$$

provided that

$$\lim_{k\to\infty} \nu(D_k) = 0 \tag{3.6.5}$$

(which can be assumed without loss of generality); note that if (3.6.5) is not satisfied, (3.6.4) need not hold (consider $A = A_k = I$ and the partitions $a = t_0^k < t_1^k = b$ for every $k \in \mathbf{N}$). Schlesinger's proof of (3.6.4) seems too complicated and even faulty; we instead argue similarly as in the proof of Theorem 3.2.2: Take a positive number M such $\|A_k(x)\| \le M$ for every $k \in \mathbf{N}$ and $x \in [a,b]$. Then

$$\left\|\exp(A_j^k \Delta t_j^k) - I - A_j^k \Delta t_j^k\right\| \le (M\Delta t_j^k)^2 e^{M\Delta t_j^k}$$

for every $k \in \mathbf{N}$ and $j = 1, \ldots, m(k)$. According to Lemma 3.2.1,

$$\left\|\prod_{j=1}^{m(k)} \exp(A_j^k \Delta t_j^k) - \prod_{j=1}^{m(k)} (I + A_j^k \Delta t_j^k)\right\| =$$

$$= \left\|\sum_{j=1}^{m(k)} \left(\prod_{l=1}^{j-1} (I + A_l^k \Delta t_l^k) \cdot (\exp(A_j^k \Delta t_j^k) - I - A_j^k \Delta t_j^k) \cdot \prod_{l=j+1}^{m(k)} \exp(A_l^k \Delta t_l^k)\right)\right\| \le$$

[1] [LS2], p. 487
[2] [LS2], p. 485–486

$$\leq e^{M(b-a)}\sum_{j=1}^{m(k)}\left\|\exp(A_j^k\Delta t_j^k)-I-A_j^k\Delta t_j^k\right\|\leq e^{M(b-a)}M^2\sum_{j=1}^{m(k)}(\Delta t_j^k)^2 e^{M\Delta t_j^k}\leq$$

$$\leq e^{M(b-a)}M^2\nu(D_k)e^{M\nu(D_k)}\sum_{j=1}^{m(k)}\Delta t_j^k = e^{M(b-a)}M^2\nu(D_k)e^{M\nu(D_k)}(b-a).$$

This completes the proof of (3.6.4). Schlesinger now states that

$$\prod_{i=1}^{m(k)}(I+A_i^k\Delta t_i^k)=I+\sum_{s=1}^{m(k)}\sum_{1\leq i_1<\cdots<i_s\leq m(k)}A_{i_1}^k\cdots A_{i_s}^k\Delta t_{i_1}^k\cdots\Delta t_{i_s}^k$$

and concludes the proof saying that

$$\lim_{k\to\infty}\sum_{1\leq i_1<\cdots<i_s\leq m(k)}A_{i_1}^k\cdots A_{i_s}^k\Delta t_{i_1}^k\cdots\Delta t_{i_s}^k =$$

$$=\int_a^b\int_a^{x_s}\cdots\int_a^{x_2}A(x_1)\cdots A(x_s)\,\mathrm{d}x_1\cdots\mathrm{d}x_s$$

The last step perhaps deserves a better explanation: Denote

$$X^s=\{(x_1,\ldots,x_s)\in\mathbf{R}^s;a\leq x_1<x_2<\cdots<x_s\leq b\},$$

and

$$X_k^s = \bigcup_{1\leq i_1<\cdots<i_s\leq m(k)}[t_{i_1-1},t_{i_1}]\times[t_{i_2-1},t_{i_2}]\times\cdots\times[t_{i_s-1},t_{i_s}],$$

where s and k are arbitrary positive integers. If χ^s and χ_k^s denote the characteristic functions of X^s and X_k^s, then $\chi_k^s\to\chi^s$ for $k\to\infty$. Consequently

$$\lim_{k\to\infty}\sum_{1\leq i_1<\cdots<i_s\leq m(k)}A_{i_1}^k\cdots A_{i_s}^k\Delta t_{i_1}^k\cdots\Delta t_{i_s}^k =$$

$$=\lim_{k\to\infty}\int_a^b\int_a^b\cdots\int_a^b A_k(x_1)\cdots A_k(x_s)\chi_k^s(x_1,\ldots,x_s)\,\mathrm{d}x_1\cdots\mathrm{d}x_s=$$

$$=\int_a^b\int_a^b\cdots\int_a^b A(x_1)\cdots A(x_s)\chi^s(x_1,\ldots,x_s)\,\mathrm{d}x_1\cdots\mathrm{d}x_s=$$

$$=\int_a^b\int_a^{x_s}\cdots\int_a^{x_2}A(x_1)\cdots A(x_s)\,\mathrm{d}x_1\cdots\mathrm{d}x_s$$

(we have used the dominated convergence theorem). □

Remark 3.6.12. The deficiency in the previous proof is that Schlesinger didn't justify the equality

$$\lim_{k\to\infty}\left(I+\sum_{s=1}^{m(k)}\sum_{1\leq i_1<\cdots<i_s\leq m(k)}A_{i_1}^k\cdots A_{i_s}^k\Delta t_{i_1}^k\cdots\Delta t_{i_s}^k\right)=$$

$$I + \sum_{s=1}^{\infty} \lim_{k\to\infty} \left(\sum_{1\le i_1<\cdots<i_s\le m(k)} A_{i_1}^k \cdots A_{i_s}^k \Delta t_{i_1}^k \cdots \Delta t_{i_s}^k \right).$$

We have already encountered a similar inaccuracy when discussing Volterra's proof of the Peano series expansion theorem for product integral; see also Masani's proof of Theorem 5.5.10.

Remark 3.6.13. Recall that, according to Theorem 2.3.5, the right derivative of a matrix function satisfies

$$(CD^{-1})\frac{\mathrm{d}}{\mathrm{d}x} = D\left(C\frac{\mathrm{d}}{\mathrm{d}x} - D\frac{\mathrm{d}}{\mathrm{d}x}\right)D^{-1}.$$

Consider two continuous matrix functions A, B defined on $[a,b]$. Using the previous formula and also the convention that

$$(I + A(t)\,\mathrm{d}t)\prod_y^x = \left((I + A(t)\,\mathrm{d}t)\prod_x^y\right)^{-1}$$

for $y > x$, we infer the equality

$$\left((I + B(t)\,\mathrm{d}t)\prod_b^x (I + A(t)\,\mathrm{d}t)\prod_x^b\right)\frac{\mathrm{d}}{\mathrm{d}x} =$$

$$= (I + A(t)\,\mathrm{d}t)\prod_b^x (B(x) - A(x))(I + A(t)\,\mathrm{d}t)\prod_x^b$$

for every $x \in [a,b]$. Denoting $S(x) = (I + A(t)\,\mathrm{d}t)\prod_b^x$ we obtain

$$\left((I + B(t)\,\mathrm{d}t)\prod_b^x (I + A(t)\,\mathrm{d}t)\prod_x^b\right)\frac{\mathrm{d}}{\mathrm{d}x} = S(x)(B(x) - A(x))S^{-1}(x),$$

and consequently (since the left hand side is equal to I for $x = b$)

$$(I + B(t)\,\mathrm{d}t)\prod_b^x (I + A(t)\,\mathrm{d}t)\prod_x^b = (I + S(t)(B(t) - A(t))S^{-1}(t)\,\mathrm{d}t)\prod_b^x.$$

Substituting $x = a$ and inverting both sides of the equation yields

$$(I + A(t)\,\mathrm{d}t)\prod_b^a (I + B(t)\,\mathrm{d}t)\prod_a^b = (I + S(t)(B(t) - A(t))S^{-1}(t)\,\mathrm{d}t)\prod_a^b. \quad (3.6.6)$$

A similar theorem (concerning the left product integral) was already present in Volterra's work[1]. Schlesinger proves[2] that the statement remains true even if A, $B \in L^*([a,b], \mathbf{R}^{n\times n})$. The proof is rather technical and we don't reproduce it here.

[1] [VH], p. 85–86
[2] [LS2], p. 488–489

Theorem 3.6.14.[1] Let $A : [a,b] \times [c,d] \to \mathbf{R}^{n\times n}$ be such that the integral

$$P(t) = (I + A(x,t)\,\mathrm{d}x)\prod_a^b$$

exists for every $t \in [c,d]$ and that

$$\left\| \frac{\partial A}{\partial t}(x,t) \right\| \le M, \quad x \in [a,b],\ t \in [c,d],$$

for some $M \in \mathbf{R}$. Then

$$P\frac{\mathrm{d}}{\mathrm{d}t} = P^{-1}(t)P'(t) = \int_a^b S(x,t)\frac{\partial A}{\partial t}(x,t)S^{-1}(x,t)\,\mathrm{d}x,$$

where $S(x,t) = (I + A(u,t)\,\mathrm{d}u)\prod_b^x$.

Proof. The definition of derivative gives

$$P^{-1}(t)P'(t) = \lim_{h\to 0}\frac{1}{h}\left((I + A(x,t)\,\mathrm{d}x)\prod_b^a (I + A(x,t+h)\,\mathrm{d}x)\prod_a^b - I \right).$$

Using Equation (3.6.6) we convert the above limit to

$$\lim_{h\to 0}\frac{1}{h}\left((I + S(x,t)(A(x,t+h) - A(x,t))S^{-1}(x,t)\,\mathrm{d}x)\prod_a^b - I \right).$$

Expanding the product integral to Peano series (see Theorem 3.6.11) we obtain

$$\lim_{h\to 0}\frac{1}{h}\sum_{k=1}^{\infty}\int_a^b\int_a^{x_k}\cdots\int_a^{x_2}\Delta(x_1,t,h)\cdots\Delta(x_k,t,h)\,\mathrm{d}x_1\cdots dx_k, \tag{3.6.7}$$

where

$$\Delta(x,t,h) = S(x,t)(A(x,t+h) - A(x,t))S^{-1}(x,t).$$

As the Peano series converges uniformly (the Weierstrass M-test, see Theorem 2.4.5), we can interchange the order of limit and summation. According to the mean value theorem there is a $\xi(h) \in [t,t+h]$ such that

$$\left\| \frac{A(x,t+h) - A(x,t)}{h} \right\| = \left\| \frac{\partial A}{\partial t}(x,\xi(h)) \right\| \le M.$$

The dominated convergence theorem therefore implies

$$\lim_{h\to 0}\frac{1}{h}\int_a^b \Delta(x_1,t,h)\,\mathrm{d}x_1 = \int_a^b \lim_{h\to 0}\frac{\Delta(x_1,t,h)}{h}\,\mathrm{d}x_1 =$$

[1] [LS2], p. 490–491

$$= \int_a^b S(x_1,t)\frac{\partial A}{\partial t}(x_1,t)S^{-1}(x_1,t)\,\mathrm{d}x_1,$$

and for $k \geq 2$

$$\lim_{h\to 0}\frac{1}{h}\int_a^b\int_a^{x_k}\cdots\int_a^{x_2}\Delta(x_1,t,h)\cdots\Delta(x_k,t,h)\,\mathrm{d}x_1\cdots dx_k =$$

$$= \int_a^b\int_a^{x_k}\cdots\int_a^{x_2}\lim_{h\to 0}\left(h^{k-1}\frac{\Delta(x_1,t,h)}{h}\cdots\frac{\Delta(x_k,t,h)}{h}\right)\mathrm{d}x_1\cdots dx_k = 0,$$

which completes the proof. □

The following statement generalizes Theorem 2.5.12; Schlesinger replaces Volterra's assumption $A \in \mathcal{C}([a,b],\mathbf{R}^{n\times n})$ by a weaker condition $A \in L^*([a,b],\mathbf{R}^{n\times n})$.

Theorem 3.6.15.[1] If $A \in L^*([a,b],\mathbf{R}^{n\times n})$ and $C \in \mathbf{R}^{n\times n}$ is a regular matrix, then

$$(I + C^{-1}A(x)C\,\mathrm{d}x)\prod_a^b = C^{-1}(I + A(x)\,\mathrm{d}x)\prod_a^b C.$$

Proof. Since $(C^{-1}AC)^k = C^{-1}A^kC$ for every $k \in \mathbf{N}$, we have

$$\exp(C^{-1}AC) = C^{-1}\exp(A)C.$$

If A is a step function, then

$$(I + C^{-1}A(x)C\,\mathrm{d}x)\prod_a^b = \prod_{i=1}^m e^{C^{-1}A(\xi_i)C\Delta t_i} =$$

$$= C^{-1}\prod_{i=1}^m e^{A(\xi_i)\Delta t_i}C = C^{-1}(I + A(x)\,\mathrm{d}x)\prod_a^b C.$$

In the general case when $A \in L^*([a,b],\mathbf{R}^{n\times n})$, we rewrite the above equation with simple functions A_k in place of A, and then pass to the limit $k \to \infty$. □

3.7 Double and contour product integrals

A considerable part of the paper [LS2] is devoted to double and contour product integrals (in $\mathbf{R}^2$ as well as in $\mathbf{C}$). Probably the most remarkable achievement is Schlesinger's proof of the "Green's theorem" for product integral, which is reproduced in the following text.

Definition 3.7.1. Let G be the rectangle $[a,b]\times[c,d]$ in $\mathbf{R}^2$ and $A : G \to \mathbf{R}^{n\times n}$ a matrix function on G. The double product integral of A over G is defined as

$$(I + A(x,y)\,\mathrm{d}x\,\mathrm{d}y)\prod_G = \left(I + \left(\int_a^b A(x,y)\,\mathrm{d}x\right)\mathrm{d}y\right)\prod_c^d,$$

[1] [LS2], p. 489

provided both integrals on the right hand side exist (in the sense of Lebesgue).

Definition 3.7.2. Let G be the rectangle $[a,b]\times[c,d]$ in $\mathbf{R}^2$ and $P,Q:G\to\mathbf{R}^{n\times n}$ continuous functions on G. We denote

$$U(x,y)=(I+P(t,c)\,\mathrm{d}t)\prod_a^x (I+Q(x,t)\,\mathrm{d}t)\prod_c^y,$$

$$T(x,y)=(I+Q(a,t)\,\mathrm{d}t)\prod_c^y (I+P(t,y)\,\mathrm{d}t)\prod_a^x$$

for every $x\in[a,b]$, $y\in[c,d]$. The contour product integral over the boundary of rectangle G is defined as the matrix

$$(I+P(x,y)\,\mathrm{d}x+Q(x,y)\,\mathrm{d}y)\prod_{\partial G}=U(b,d)T(b,d)^{-1}. \tag{3.7.1}$$

Remark 3.7.3. Schlesinger refers to the matrices $U(b,d)$ and $T(b,d)$ as to the "integral over the lower step" and "integral over the upper step" of the rectangle G. They are clearly a special case of the contour product integral as defined by Volterra (see definition 2.6.8); the matrix (3.7.1) corresponds to the value of contour product integral along the (anticlockwise oriented) boundary of G.

Theorem 3.7.4.[1] Let G be the rectangle $[a,b]\times[c,d]$ in $\mathbf{R}^2$ and $P,Q:G\to\mathbf{R}^{n\times n}$ continuous matrix functions on G. Assume that the derivatives

$$\frac{\partial P}{\partial y},\quad \frac{\partial Q}{\partial x}$$

exist and are continuous on G. Then

$$(I+P(x,y)\,\mathrm{d}x+Q(x,y)\,\mathrm{d}y)\prod_{\partial G}=(I+T\cdot\Delta^*(P,Q)\cdot T^{-1}\,\mathrm{d}x\,\mathrm{d}y)\prod_G,$$

where

$$\Delta^*(P,Q)=\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}+PQ-QP,$$

$$T(x,y)=(I+Q(a,t)\,\mathrm{d}t)\prod_c^y (I+P(t,y)\,\mathrm{d}t)\prod_a^x.$$

Proof. A simple calculation reveals that (compare to Lemma 2.6.4)

$$T\cdot\Delta^*(P,Q)\cdot T^{-1}=\frac{\partial}{\partial x}\left(T\left(Q-T\frac{\mathrm{d}}{\mathrm{d}y}\right)T^{-1}\right).$$

[1] [LS2], p. 496–497

Taking the product integral over G we obtain

$$(I + T \cdot \Delta^*(P,Q) \cdot T^{-1} \, \mathrm{d}x \, \mathrm{d}y) \prod_G =$$

$$= \left(I + \left[T(x,y)\left(Q(x,y) - T(x,y)\frac{\mathrm{d}}{\mathrm{d}y}\right)T(x,y)^{-1}\right]_a^b \mathrm{d}y\right) \prod_c^d . \tag{3.7.2}$$

According to the rules for differentiating a product of functions (see Theorem 2.3.2),

$$T\frac{\mathrm{d}}{\mathrm{d}y} = (I + P(t,y)\,\mathrm{d}t) \prod_x^a Q(a,y)(I + P(t,y)\,\mathrm{d}t) \prod_a^x + \left((I + P(t,y)\,\mathrm{d}t) \prod_a^x\right) \frac{\mathrm{d}}{\mathrm{d}y}.$$

Theorem 3.6.14 on differentiating the product integral with respect to a parameter yields

$$\lim_{x \to a} \left((I + P(t,y)\,\mathrm{d}t) \prod_a^x\right) \frac{\mathrm{d}}{\mathrm{d}y} = 0,$$

and consequently

$$\lim_{x \to a} T\frac{\mathrm{d}}{\mathrm{d}y} = Q(a,y). \tag{3.7.3}$$

The equalities (3.7.2) and (3.7.3) imply

$$(I + T \cdot \Delta^*(P,Q) \cdot T^{-1} \, \mathrm{d}x \, \mathrm{d}y) \prod_G =$$

$$= \left(I + \lim_{x \to b}\left(T(x,y)\left(Q(x,y) - T(x,y)\frac{\mathrm{d}}{\mathrm{d}y}\right)T(x,y)^{-1}\right)\mathrm{d}y\right) \prod_c^d =$$

$$= \lim_{x \to b}\left(I + T(x,y)\left(Q(x,y) - T(x,y)\frac{\mathrm{d}}{\mathrm{d}y}\right)T(x,y)^{-1}\mathrm{d}y\right) \prod_c^d \tag{3.7.4}$$

(we have used Theorem 3.6.3 on interchanging the order of limit and integral). For every $x \in [a,b]$ we have

$$T(x,y)\frac{\mathrm{d}}{\mathrm{d}y} = \left(T(x,d)^{-1}T(x,y)\right)\frac{\mathrm{d}}{\mathrm{d}y}$$

and also

$$T(x,d)^{-1}T(x,y) = \left(I + T(x,u)\frac{\mathrm{d}}{\mathrm{d}u}\,\mathrm{d}u\right) \prod_d^y = \left(I + \left(T(x,d)^{-1}T(x,u)\right)\frac{\mathrm{d}}{\mathrm{d}u}\,\mathrm{d}u\right) \prod_d^y .$$

Using Theorem 3.6.15 and Equation (3.6.6) we arrive at

$$\left(I + T(x,y)\left(Q(x,y) - T(x,y)\frac{\mathrm{d}}{\mathrm{d}y}\right)T(x,y)^{-1}\mathrm{d}y\right) \prod_c^d = T(x,d)\cdot$$

$$\cdot\left(I+T(x,d)^{-1}T(x,y)\left(Q(x,y)-\left(T(x,d)^{-1}T(x,y)\right)\frac{\mathrm{d}}{\mathrm{d}y}\right)T(x,y)^{-1}T(x,d)\,\mathrm{d}y\right)\prod_c^d\cdot$$

$$\cdot\, T(x,d)^{-1}=T(x,d)\left(I+\left(T(x,d)^{-1}T(x,y)\right)\frac{\mathrm{d}}{\mathrm{d}y}\,\mathrm{d}y\right)\prod_d^c\cdot$$

$$\cdot(I+Q(x,y)\,\mathrm{d}y)\prod_c^d T(x,d)^{-1}=T(x,d)\,T(x,d)^{-1}T(x,c)(I+Q(x,y)\,\mathrm{d}y)\prod_c^d\cdot$$

$$\cdot\, T(x,d)^{-1}=T(x,c)(I+Q(x,y)\,\mathrm{d}y)\prod_c^d T(x,d)^{-1}.$$

Finally, Equation (3.7.4) gives

$$(I+T\cdot\Delta^*(P,Q)\cdot T^{-1}\,\mathrm{d}x\,\mathrm{d}y)\prod_G=\lim_{x\to b}\left(T(x,c)(I+Q(x,y)\,\mathrm{d}y)\prod_c^d T(x,d)^{-1}\right)=$$

$$=T(b,c)(I+Q(b,y)\,\mathrm{d}y)\prod_c^d T(b,d)^{-1}=(I+P(x,y)\,\mathrm{d}x+Q(x,y)\,\mathrm{d}y)\prod_{\partial G}.$$

□

Remark 3.7.5. The previous theorem represents an analogy of Green's theorem for the product integral; we have already encountered a similar statement when discussing Volterra's work. Volterra's analogy of the curl operator was

$$\Delta(P,Q)=\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}+QP-PQ,$$

while Schlesinger's curl has the form

$$\Delta^*(P,Q)=\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}+PQ-QP.$$

The reason is that Volterra stated his theorem for the left product integral, while Schlesinger was concerned with the right product integral (see Theorem 2.6.15 and Remark 2.6.7). Whereas Volterra worked with a simply connected domain G (see definition 2.6.12), Schlesinger considers only rectangles.

Consider functions P, Q that satisfy assumptions of Theorem 3.7.4 and such that

$$\Delta^*(P,Q)=0 \tag{3.7.5}$$

everywhere in G. Then

$$(I+P(x,y)\,\mathrm{d}x+Q(x,y)\,\mathrm{d}y)\prod_{\partial G}=I,$$

which in consequence means that the values of contour product integral over the lower step and over the upper step are the same. Schlesinger then denotes the common value of the matrices $U(x,y)$ and $T(x,y)$ (see definition 3.7.2) by the symbol

$$(I + P(x,y)\,\mathrm{d}x + Q(x,y)\,\mathrm{d}y)\prod_{(a,c)}^{(b,d)}.$$

Clearly

$$\left((I + P(u,v)\,\mathrm{d}u + Q(u,v)\,\mathrm{d}v)\prod_{(a,c)}^{(x,y)}\right)\frac{\mathrm{d}}{\mathrm{d}x} = T(x,y)\frac{\mathrm{d}}{\mathrm{d}x} = P(x,y),$$

$$\left((I + P(u,v)\,\mathrm{d}u + Q(u,v)\,\mathrm{d}v)\prod_{(a,c)}^{(x,y)}\right)\frac{\mathrm{d}}{\mathrm{d}y} = U(x,y)\frac{\mathrm{d}}{\mathrm{d}y} = Q(x,y).$$

Schlesinger now proceeds to define product integral along a contour and shows that (in a simply connected domain) the condition (3.7.5) implies that the value of product integral depends only on the endpoints of the contour. His method is almost the same as Volterra's and we don't repeat it here.

At the end of paper [LS2] Schlesinger treats matrix functions of a complex variable. He defines the contour product integral in complex domain and recapitulates the results proved earlier by Volterra (theorems 2.7.4, 2.7.7, and 2.7.6).

3.8 Generalization of Schlesinger's definition

Thanks to the definition proposed by Ludwig Schlesinger it is possible to extend the class of product integrable functions and to work with bounded measurable functions instead of Riemann integrable functions. At this place we remind the notation

$$L^*([a,b],\mathbf{R}^{n\times n}) = \left\{A : [a,b] \to \mathbf{R}^{n\times n};\ A \text{ is measurable and bounded}\right\}.$$

Schlesinger was aware that his definition might be extended to all matrix functions with Lebesgue integrable (not necessarily bounded) entries, i. e. to the class

$$L([a,b],\mathbf{R}^{n\times n}) = \left\{A : [a,b] \to \mathbf{R}^{n\times n};\ (L)\int_a^b \|A(t)\|\,\mathrm{d}t < \infty\right\},$$

where the symbol (L) emphasizes that we are dealing with the Lebesgue integral. Clearly $L^* \subset L$. If $\{A_k\}_{k=1}^{\infty}$ is a uniformly bounded sequence of functions which converge to A almost everywhere, then according to lemma 3.5.4

$$\lim_{k\to\infty} \|A_k - A\|_1 = \lim_{k\to\infty} \int_a^b \|A_k(x) - A(x)\|\,\mathrm{d}x = 0,$$

i.e. A_k converge to A also in the norm of space $L([a,b],\mathbf{R}^{n\times n})$. Taking account of Theorem 3.5.5 it is natural to state the following definition.

Definition 3.8.1. A function $A:[a,b]\to\mathbf{R}^{n\times n}$ is called product integrable if there exists a sequence of step functions $\{A_k\}_{k=1}^{\infty}$ such that

$$\lim_{k\to\infty}\|A_k-A\|_1=0.$$

We define

$$(I+A(t)\,\mathrm{d}t)\prod_a^b=\lim_{k\to\infty}(I+A_k(t)\,\mathrm{d}t)\prod_a^b.$$

Remark 3.8.2. The correctness of the previous definition is ensured by theorem 3.5.5. Since step functions belong to the space $L([a,b],\mathbf{R}^{n\times n})$, which is complete, every product integrable function also belongs to this space. Moreover, step functions form a dense subset in this space[1], and therefore $(I+A(t)\,\mathrm{d}t)\prod_a^b$ exists iff $A\in L([a,b],\mathbf{R}^{n\times n})$, i. e. iff the integral $(L)\int_a^b A(t)\,\mathrm{d}t$ exists.

Interested readers are referred to the book [DF] for more details about the theory of product integral based on definition 3.8.1. As an interesting example we present the proof of theorem on differentiating the product integral with respect to the upper bound of integration. We start with a preliminary lemma (which follows also from Theorem 3.3.2, but we don't want to use it as we are seeking another way to prove it).

Lemma 3.8.3. If $A:[a,b]\to\mathbf{R}^{n\times n}$ is a step function, then

$$Y(x)=I+\int_a^x Y(t)A(t)\,\mathrm{d}t,\quad x\in[a,b].$$

Proof. There exist a partition $a=t_0<t_1<\cdots<t_m=b$ and matrices $A_1,\ldots,A_m\in\mathbf{R}^{n\times n}$ such that

$$A(x)=A_k,\quad x\in(t_{k-1},t_k).$$

Then

$$Y(x)=(I+A(t)\,\mathrm{d}t)\prod_a^x=e^{A_1(t_2-t_1)}\cdots e^{A_{k-1}(t_{k-1}-t_{k-2})}e^{A_k(x-t_{k-1})}$$

for every $x\in[t_{k-1},t_k]$. The function Y is continuous on $[a,b]$ and differentiable except a finite number of points; we have

$$Y'(x)=Y(x)A(x)\tag{3.8.1}$$

[1] [RG], Corollary 3.29, p. 47

for $x \in [a,b]\backslash\{t_0, t_1, \dots, t_m\}$. This implies

$$Y(x) = I + \int_a^x Y(t)A(t)\,\mathrm{d}t, \quad x \in [a,b].$$

□

Theorem 3.8.4.[1] Consider function $A \in L([a,b], \mathbf{R}^{n\times n})$. For every $x \in [a,b]$ the integral

$$Y(x) = (I + A(t)\,\mathrm{d}t)\prod_a^x \tag{3.8.2}$$

exists and the function Y satisfies the equation

$$Y(x) = I + \int_a^x Y(t)A(t)\,\mathrm{d}t, \quad x \in [a,b]. \tag{3.8.3}$$

Proof. Let $A_k : [a,b] \to \mathbf{R}^{n\times n}$, $k \in \mathbf{N}$ be a sequence of step functions such that

$$\lim_{k\to\infty} \|A_k - A\|_1 = \lim_{k\to\infty} \int_a^b \|A_k(t) - A(t)\|\,\mathrm{d}t = 0. \tag{3.8.4}$$

Then clearly

$$\lim_{k\to\infty} \int_a^x \|A_k(t) - A(t)\|\,\mathrm{d}t = 0, \quad x \in [a,b],$$

i. e. the definition (3.8.2) is correct. Denote

$$Y_k(x) = (I + A_k(t)\,\mathrm{d}t)\prod_a^x.$$

Because A_k are step functions, Lemma 3.8.3 implies

$$Y_k(x) = I + \int_a^x Y_k(t)A_k(t)\,\mathrm{d}t, \quad x \in [a,b]. \tag{3.8.5}$$

According to Corollary 3.4.3,

$$\|Y_l(x) - Y_m(x)\| \le \exp\left(\int_a^b \|A_l(t)\|\,\mathrm{d}t\right)\left(\exp\left(\int_a^b \|A_l(t) - A_m(t)\|\,\mathrm{d}t\right) - 1\right) =$$

$$= \exp\|A_l\|_1 \left(\exp\|A_l(t) - A_m(t)\|_1 - 1\right).$$

From Equation (3.8.4) we see that $\{A_k\}_{k=1}^{\infty}$ is a bounded and Cauchy sequence with respect to the norm $\|\cdot\|_1$. The previous inequality therefore implies that Y_k converge uniformly to Y, i. e.

$$\|Y_k - Y\|_\infty = \sup_{x\in[a,b]} \|Y_k(x) - Y(x)\| \to 0 \quad \text{pro } k \to \infty.$$

[1] [DF], p. 54–55

We now estimate

$$\left\|\int_a^x Y_k(t)A_k(t)\,\mathrm{d}t - \int_a^x Y(t)A(t)\,\mathrm{d}t\right\| \le \|Y_kA_k - YA\|_1 \le$$

$$\le \|(Y_k - Y)A_k\|_1 + \|Y(A_k - A)\|_1 \le \|A_k\|_1\|Y_k - Y\|_\infty + \|(A_k - A)\|_1\|Y\|_\infty,$$

and consequently

$$\lim_{k\to\infty}\int_a^x Y_k(t)A_k(t)\,\mathrm{d}t = \int_a^x Y(t)A(t)\,\mathrm{d}t.$$

The equality (3.8.3) is obtained by passing to the limit in equation (3.8.5). □

Corollary 3.8.5. If $A \in L([a,b],\mathbf{R}^{n\times n})$, then the function

$$Y(x) = (I + A(t)\,\mathrm{d}t)\prod_a^x, \quad x \in [a,b],$$

is absolutely continuous on $[a,b]$ and

$$Y(x)^{-1}\cdot Y'(x) = A(x)$$

almost everywhere on $[a,b]$.

Remark 3.8.6. In our proof of the previous theorem we have employed Schlesinger's estimate from Corollary 3.4.3, whose proof is somewhat laborious. The authors of [DF] instead make use of a different inequality, which is easier to demonstrate. Let A, $B : [a,b] \to \mathbf{R}^{n\times n}$ be two step functions. Denoting

$$Y(x) = (I + A(t)\,\mathrm{d}t)\prod_a^x, \quad Z(x) = (I + B(t)\,\mathrm{d}t)\prod_a^x,$$

we see that the function YZ^{-1} is continuous on $[a,b]$ and differentiable except a finite number of points. Using Equation (3.8.1) we calculate

$$(YZ^{-1})' = Y'Z^{-1} + Y(Z^{-1})' = YY^{-1}Y'Z^{-1} - YZ^{-1}Z'Z^{-1} = Y(A-B)Z^{-1},$$

and consequently

$$Y(x)Z^{-1}(x) = I + \int_a^x (YZ^{-1})'(t)\,\mathrm{d}t = I + \int_a^x Y(t)(A(t)-B(t))Z^{-1}(t)\,\mathrm{d}t.$$

Multiplying this equation by Z from right and substituting $x = b$ we obtain

$$\left\|(I + A(t)\,\mathrm{d}t)\prod_a^b - (I + B(t)\,\mathrm{d}t)\prod_a^b\right\| = \|Y(b) - Z(b)\| \le$$

$$\le \int_a^b \|Y(t)\|\cdot\|A(t)-B(t)\|\cdot\|Z^{-1}(t)\|\,\mathrm{d}t\cdot\|Z(b)\| \le e^{2\|B\|_1}e^{\|A\|_1}\|A-B\|_1$$

(we have used Lemma 3.4.1 to estimate $\|Y(t)\|$, $\|Z^{-1}(t)\|$ and $\|Z(b)\|$). The meaning of the last inequality is similar to the meaning of inequality from Corollary 3.4.3: "If two step functions A, B are close with respect to the norm $\|\cdot\|_1$, then the values of their product integrals are also close to each other."