Chapter 2

The origins of product integration

The notion of product integral has been introduced by Vito Volterra at the end of the 19th century. We start with a short biography of this eminent Italian mathematician and then proceed to discuss his work on product integration.

Vito Volterra was born at Ancona on 3rd May 1860. His father died two years later; Vito moved in with his mother Angelica to Alfonso, Angelica's brother, who supported them and was like a boy's father. Because of their financial situation, Angelica and Alfonso didn't want Vito to study his favourite subject, mathematics, at university, but eventually Edoardo Almagià, Angelica's cousin and a railroad engineer, helped to persuade them. An important role was also played by Volterra's teacher Ròiti, who secured him a place of assistant in a physical laboratory.

Vito Volterra[1]

In 1878 Volterra entered the University of Pisa; among his professors was the famous Ulisse Dini. In 1879 he passed the examination to Scuola Normale Superiore of Pisa. Under the influence of Enrico Betti, his interest shifted towards mathematical physics. In 1882 he offered a thesis on hydrodynamics, graduated doctor of physics and became Betti's assistent. Shortly after, in 1883, the young Volterra won the competition for the vacant chair of rational mechanics and was promoted

[1] Photo from [McT]

to professor of the University of Pisa. After Betti's death he took over his course in mathematical physics. In 1893 he moved to the University of Turin, but eventually settled in Rome in 1900. The same year he married Virginia Almagià (the daughter of Edoardo Almagià).

During the first quarter of the 20th century Volterra not only represented the leading figure of Italian mathematics, but also became involved in politics and was nominated a Senator of the Kingdom in 1905.

When Italy entered the world war in 1915, Volterra volunteered the Army Corps of Engineers and engaged himself in perfecting of airships and firing from them; he also promoted the collaboration with French and English scientists. After the end of the war he returned to scientific work and teaching at the university.

Volterra strongly opposed the Mussolini regime which came to power in 1922. As one of the few professors who refused to take an oath of loyalty imposed by the fascists in 1931, he was forced to leave the University of Rome and other scientific institutions. After then he spent a lot of time abroad (giving lectures e.g. in France, Spain, Czechoslovakia or Romania) and also at his country house in Ariccia. Volterra, who was of Jewish descent, was also affected by the antisemitic racial laws of 1938. Although he began to suffer from phlebitis, he still devoted himself actively to mathematics. He died in isolation on 11th October 1940 without a greater interest of Italian scientific community.

Despite the fact that Volterra is best known as a mathematician, he was a man of universal interests and devoted himself also to physics, biology and economy. His mathematical research often had origins in physical problems. Volterra was also an enthusiastic bibliophile and his collection, which reached nearly seven thousand volumes and is now deposited in the United States, included rare copies of scientific papers e.g. by Galileo, Brahe, Tartaglia, Fermat etc. The monograph [JG] contains a wealth of information about Volterra's life and times.

Volterra's name is closely connected with integral equations. He contributed the method of successive approximations for solving integral equations of the second kind, and also noted that an integral equation might be considered as a limiting case of a system of algebraic linear equations; this observation was later utilized by Ivar Fredholm (see also the introduction to Chapter 4).

His investigations in calculus of variations led him to the study of functionals (he called them "functions of lines"); in fact he built a complete calculus including the definitions of continuity, derivative and integral of a functional. Volterra's pioneering work on integral equations and functionals is often regarded as the dawn of functional analysis. An overview of his achievements in this field can be obtained from the book [VV5].

Volterra was also one of the founders of mathematical biology. The motivation came from his son-in-law Umberto D'Ancona, who was studying the statistics of Adriatic fishery. He posed to Volterra the problem of explaining the relative increase of predator fishes, as compared with their prey, during the period of First World War (see e.g. [MB]). Volterra interpreted this phenomenon with the help of mathematical models of struggle between two species; from mathematical point of view, the models were combinations of differential and integral equations. Volterra's correspondence concerning mathematical biology was published in the book [IG].

A more detailed description of Volterra's activites (his work on partial differential equations, theory of elasticity) can be found in the biographies [All, JG] and also in the books [IG, VV5]. An interesting account of Italian mathematics and its intertwining with politics in the first half of the 20th century is given in [GN].

2.1 Product integration in the work of Vito Volterra

Volterra's first work devoted to product integration [VV1] was published in 1887 and was written in Italian. It introduces the two basic concepts of the multiplicative calculus, namely the derivative of a matrix function and the product integral. The topics discussed in [VV1] are essentially the same as in Sections 2.3 to 2.6 of the present chapter. The publication [VV1] was followed by a second part [VV2] printed in 1902, which is concerned mainly with matrix functions of a complex variable. It includes results which are recapitulated in Sections 2.7 and 2.8, and also a treatment of product integration on Riemann surfaces. Volterra also published two short Italian notes, [VV3] from 1887 and [VV4] from 1888, which summarize the results of [VV1, VV2] but don't include proofs.

Volterra's final treatment of product integration is represented by the book *Opérations infinitésimales linéaires* [VH] written together with a Czech mathematician Bohuslav Hostinský. The publication appeared in the series *Collection de monographies sur la théorie des fonctions* directed by Émile Borel in 1938.

More than two hundred pages of [VH] are divided into eighteen chapters. The first fifteen chapters represent a French translation of [VV1, VV2] with only small changes and complements. The remaining three chapters, whose author is Bohuslav Hostinský, will be discussed in Chapter 4.

As Volterra notes in the book's preface, the publication of [VH] was motivated by the results obtained by Bohuslav Hostinský, as well as by an increased interest in matrix theory among mathematicians and physicists. As the bibliography of [VH] suggests, Volterra was already aware of the papers [LS1, LS2] by Ludwig Schlesinger, who linked up to Volterra's first works (see Chapter 3).

The book [VH] is rather difficult to read for contemporary mathematicians. One of the reasons is a somewhat cumbersome notation. For example, Volterra uses the same symbol to denote additive as well as multiplicative integration: The sign $\int$ applied to a matrix function denotes the product integral, while the same sign applied to a scalar function stands for the ordinary (additive) integral. Calculations with matrices are usually written out for individual entries, whereas using the matrix notation would have greatly simplified the proofs. Moreover, Volterra didn't hesitate to calculate with infinitesimal quantities, he interchanges the order of summation and integration or the order of partial derivatives without any justification etc. The conditions under which individual theorems hold (e.g. continuity or differentiability of the given functions) are often omitted and must be deduced from the proof. This is certainly surprising, since the rigorous foundations of mathematical analysis were already laid out at the end of the 19th century, and even Volterra contributed to them during his studies by providing an example of a function which is not Riemann integrable but has an antiderivative.

The following sections summarize Volterra's achievements in the field of product integration. Our discussion is based on the text from [VH], but the results are stated in the language of contemporary mathematics (with occasional comments on Volterra's notation). Proofs of most theorems are also included; they are generally based on Volterra's original proofs except for a few cases where his calculations with infinitesimal quantities were replaced by a different, rigorous argument.

2.2 Basic results of matrix theory

The first four chapters of the book [VH] recapitulate some basic results of matrix theory. Most of them are now taught in linear algebra courses and we repeat only some of them for reader's convenience, as we will refer to them in subsequent chapters.

Volterra refers to matrices as to substitutions, because they can be used to represent a linear change of variables. A composition of two substitutions then corresponds to multiplication of matrices: If

$$x_i' = \sum_{j=1}^{n} a_{ij} x_j \quad \text{and} \quad x_i'' = \sum_{j=1}^{n} b_{ij} x_j',$$

then

$$x_i'' = \sum_{j=1}^{n} c_{ij} x_j,$$

where

$$c_{ij} = \sum_{k=1}^{n} b_{ik} a_{kj}, \qquad (2.2.1).$$

We will use the symbol $\mathbf{R}^{n\times n}$ to denote the set of all square matrices with n rows and n columns. If

$$A = \{a_{ij}\}_{i,j=1}^{n}, \quad B = \{b_{ij}\}_{i,j=1}^{n}, \quad C = \{c_{ij}\}_{i,j=1}^{n},$$

we can write Equation (2.2.1) in the form $C = B \cdot A$.

A matrix $A = \{a_{ij}\}_{i,j=1}^{n}$ is called regular if it has a nonzero determinant. If $i \in \{1, \ldots, n\}$, the theorem on expansion by minors gives

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{k=1}^{n} a_{ik} A_{ik}, \qquad (2.2.2)$$

where

$$A_{ik} = (-1)^{i+k} M_{ik}$$

is the so-called cofactor corresponding to minor M_{ik}; the minor is defined as the determinant of a matrix obtained from A by deleting the i-th row and k-th column. Since the determinant of a matrix with two or more identical rows is zero, it follows that

$$\sum_{k=1}^{n} a_{jk} A_{ik} = \delta_{ij} \det A. \tag{2.2.3}$$

for each pair of numbers $i, j \in \{1, \ldots, n\}$; recall that δ_{ij} is the Kronecker symbol

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

If we thus define the matrix

$$A^{-1} = \left\{ \frac{A_{ji}}{\det A} \right\}_{i,j=1}^{n},$$

then Equation (2.2.3) yields

$$AA^{-1} = I = A^{-1}A,$$

where

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

is the identity matrix. The matrix A^{-1} is called the inverse of A.

The definition of matrix multiplication gives the following rule for multiplication of block matrices:

Theorem 2.2.1.[1] Consider a matrix that is partitioned into m^2 square blocks A_{ij} and a matrix partitioned into m^2 square blocks B_{ij} such that A_{ij} has the same dimensions as B_{ij} for every $i, j \in \{1, \ldots, m\}$. Then

$$\begin{pmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mm} \end{pmatrix} \begin{pmatrix} B_{11} & \cdots & B_{1m} \\ \vdots & \ddots & \vdots \\ B_{m1} & \cdots & B_{mm} \end{pmatrix} = \begin{pmatrix} C_{11} & \cdots & C_{1m} \\ \vdots & \ddots & \vdots \\ C_{m1} & \cdots & C_{mm} \end{pmatrix},$$

where $C_{ik} = \sum_j A_{ij} B_{jk}$.

We will be often dealing with block diagonal matrices, i.e. with matrices of the form

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_m \end{pmatrix}$$

[1] [VH], p. 27

composed of smaller square matrices $A_1, A_2, \dots, A_m$; Volterra denotes such a matrix by the symbols

$$\{\overline{\underline{A_1 \cdot A_2 \cdots A_m}}\} \quad \text{or} \quad \left\{\overline{\underline{\prod_{i=1}^{m} A_i}}\right\},$$

but we don't follow his notation.

The following theorem expresses the fact that every square matrix can be transformed to a certain canonical form called the Jordan normal form. Volterra proves the theorem by induction on the dimension of the matrix; we refer the reader to any good linear algebra textbook.

Theorem 2.2.2.[1] To every matrix $A \in \mathbf{R}^{n\times n}$ there exist matrices $C, J \in \mathbf{R}^{n\times n}$ such that

$$A = C^{-1}JC$$

and J has the form

$$J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_m \end{pmatrix}, \quad \text{where } J_i = \begin{pmatrix} \lambda_i & 0 & \cdots & 0 & 0 \\ 1 & \lambda_i & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \lambda_i \end{pmatrix}$$

and $\{\lambda_1, \dots, \lambda_m\}$ are (not necessarily distinct) eigenvalues of A.

Recall that if $A = C^{-1}BC$ for some regular matrix C, then the matrices A, B are called similar. Thus the previous theorem says that every square matrix is similar to a certain Jordan matrix.

The next two theorems concern the properties of block diagonal matrices and are simple consequences of Theorem 2.2.1.

Theorem 2.2.3. If A_i is a square matrix of the same dimensions as B_i for every $i \in \{1, \dots, m\}$, then

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_m \end{pmatrix} \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_m \end{pmatrix} = \begin{pmatrix} A_1B_1 & 0 & \cdots & 0 \\ 0 & A_2B_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_mB_m \end{pmatrix}.$$

Theorem 2.2.4. The inverse of a block diagonal matrix composed of invertible matrices $A_1, \dots, A_m$ is equal to

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_m \end{pmatrix}^{-1} = \begin{pmatrix} A_1^{-1} & 0 & \cdots & 0 \\ 0 & A_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_m^{-1} \end{pmatrix}.$$

[1] [VH], p. 20–24

2.3 Derivative of a matrix function

In this section we focus on first of the basic operations of Volterra's matrix calculus, which is the derivative of a matrix function.

A matrix function $A : [a, b] \to \mathbf{R}^{n\times n}$ will be called differentiable at a point $x \in (a, b)$ if all the entries a_{ij}, $i, j \in \{1, \dots, n\}$ of A are differentiable at x; in this case we denote

$$A'(x) = \{a'_{ij}(x)\}_{i,j=1}^{n}.$$

We also define $A'(x)$ for the endpoints $x = a$ and $x = b$ as the matrix of the corresponding one-sided derivatives (provided they exist).

Definition 2.3.1. Let $A : [a, b] \to \mathbf{R}^{n\times n}$ be a matrix function that is differentiable and regular at a point $x \in [a, b]$. We define the left derivative of A at x as

$$\frac{\mathrm{d}}{\mathrm{d}x}A(x) = A'(x)A^{-1}(x) = \lim_{\Delta x\to 0}\frac{A(x+\Delta x)A^{-1}(x) - I}{\Delta x}$$

and the right derivative of A at x as

$$A(x)\frac{\mathrm{d}}{\mathrm{d}x} = A^{-1}(x)A'(x) = \lim_{\Delta x\to 0}\frac{A^{-1}(x)A(x+\Delta x) - I}{\Delta x}.$$

Volterra doesn't use the matrix notation and instead writes out the individual entries:

$$\frac{\mathrm{d}}{\mathrm{d}x}\{a_{ij}\} = \left\{\lim_{\Delta x\to 0}\sum_{k=1}^{n}\frac{a_{ik}(x+\Delta x) - a_{ik}(x)}{\Delta x}A_{jk}(x)\right\},$$

$$\{a_{ik}\}\frac{\mathrm{d}}{\mathrm{d}x} = \left\{\lim_{\Delta x\to 0}\sum_{k=1}^{n}A_{ki}(x)\frac{a_{kj}(x+\Delta x) - a_{kj}(x)}{\Delta x}\right\},$$

where $\{A_{ji}\}_{i,j=1}^{n}$ denote the entries of A^{-1}. He also defines the left differential as

$$\mathrm{d}\{a_{ij}\} = A(x+\mathrm{d}x)A(x)^{-1} = I + A'(x)A(x)^{-1}\,\mathrm{d}x = \left\{\delta_{ij} + \sum_{k=1}^{n}a'_{ik}(x)A_{jk}(x)\,\mathrm{d}x\right\}$$

and the right differential as

$$\{a_{ij}\}\mathrm{d} = A(x)^{-1}A(x+\mathrm{d}x) = I + A(x)^{-1}A'(x)\,\mathrm{d}x = \left\{\delta_{ij} + \sum_{k=1}^{n}A_{ki}(x)a'_{kj}(x)\,\mathrm{d}x\right\},$$

where $\mathrm{d}x$ is an infinitesimal quantity. Both differentials are considered as matrices that differ infinitesimally from the identity matrix.

Volterra uses infinitesimal quantities without any scruples, which sometimes leads to very unreadable proofs. This is also the case of the following theorem; Volterra's justification has been replaced by a rigorous proof.

Theorem 2.3.2.[1] If $A, B : [a,b] \to \mathbf{R}^{n\times n}$ are differentiable and regular matrix functions at $x \in [a,b]$, then

$$\frac{\mathrm{d}}{\mathrm{d}x}(AB) = \frac{\mathrm{d}}{\mathrm{d}x}A + A\left(\frac{\mathrm{d}}{\mathrm{d}x}B\right)A^{-1} = A\left(A\frac{\mathrm{d}}{\mathrm{d}x} + \frac{\mathrm{d}}{\mathrm{d}x}B\right)A^{-1},$$

$$(AB)\frac{\mathrm{d}}{\mathrm{d}x} = B\frac{\mathrm{d}}{\mathrm{d}x} + B^{-1}\left(A\frac{\mathrm{d}}{\mathrm{d}x}\right)B = B^{-1}\left(A\frac{\mathrm{d}}{\mathrm{d}x} + \frac{\mathrm{d}}{\mathrm{d}x}B\right)B,$$

where all derivatives are taken at the given point x.

Proof. The definition of the left derivative gives

$$\frac{\mathrm{d}}{\mathrm{d}x}(AB) = (AB)'(AB)^{-1} = (A'B + AB')B^{-1}A^{-1} = A'A^{-1} + AB'B^{-1}A^{-1},$$

where the expression on the right hand side is equal to

$$\frac{\mathrm{d}}{\mathrm{d}x}A + A\left(\frac{\mathrm{d}}{\mathrm{d}x}B\right)A^{-1},$$

but can be also transformed to the form

$$AA^{-1}A'A^{-1} + AB'B^{-1}A^{-1} = A\left(A\frac{\mathrm{d}}{\mathrm{d}x} + \frac{\mathrm{d}}{\mathrm{d}x}B\right)A^{-1}.$$

The second part is proved in a similar way. □

Corollary 2.3.3.[2] Consider a matrix function $A : [a,b] \to \mathbf{R}^{n\times n}$ that is differentiable and regular on $[a,b]$. Then for an arbitrary regular matrix $C \in \mathbf{R}^{n\times n}$ we have

$$\frac{\mathrm{d}}{\mathrm{d}x}(AC) = \frac{\mathrm{d}}{\mathrm{d}x}A.$$

The corollary can be expressed like this: *The left derivative of a matrix function doesn't change, if the function is multiplied by a constant matrix from right.* It is also easy to prove a dual statement: *The right derivative of a matrix function doesn't change, if the function is multiplied by a constant matrix from left.* Symbolically written,

$$(CA)\frac{\mathrm{d}}{\mathrm{d}x} = A\frac{\mathrm{d}}{\mathrm{d}x}.$$

As Volterra notes, this is a general principle: Each statement concerning matrix functions remains true, if we replace all occurences of the word "left" by the word

[1] [VH], p. 43
[2] [VH], p. 39

“right” and vice versa. A precise formulation and justification of this duality principle is due to P. R. Masani and will be given in Chapter 5, Remark 5.2.2.

Theorem 2.3.4.[1] If $A : [a,b] \to \mathbf{R}^{n\times n}$ is differentiable and regular at $x \in [a,b]$, then

$$\frac{\mathrm{d}}{\mathrm{d}x}(A^{-1}) = -A\frac{\mathrm{d}}{\mathrm{d}x}, \quad (A^{-1})\frac{\mathrm{d}}{\mathrm{d}x} = -\frac{\mathrm{d}}{\mathrm{d}x}A,$$

where all derivatives are taken at the given point x.

Proof. Differentiating the equation $AA^{-1} = I$ yields $A'A^{-1} + A(A^{-1})' = 0$, and consequently $(A^{-1})' = -A^{-1}A'A^{-1}$. The statement follows easily. □

Corollary 2.3.5.[2] If $A, B : [a,b] \to \mathbf{R}^{n\times n}$ are differentiable and regular matrix functions at $x \in [a,b]$, then

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

$$(AB^{-1})\frac{\mathrm{d}}{\mathrm{d}x} = B\left(A\frac{\mathrm{d}}{\mathrm{d}x} - B\frac{\mathrm{d}}{\mathrm{d}x}\right)B^{-1},$$

where all derivatives are taken at the given point x.

Proof. A simple consequence of Theorems 2.3.2 and 2.3.4. □

Theorem 2.3.6.[3] Consider functions $A, B : [a,b] \to \mathbf{R}^{n\times n}$ that are differentiable and regular on $[a,b]$. If

$$\frac{\mathrm{d}}{\mathrm{d}x}A = \frac{\mathrm{d}}{\mathrm{d}x}B$$

on $[a,b]$, then there exists a matrix $C \in \mathbf{R}^{n\times n}$ such that $B(x) = A(x)C$ for every $x \in [a,b]$.

Proof. Define $C(x) = A^{-1}(x)B(x)$ for $x \in [a,b]$. Corollary 2.3.5 gives

$$\frac{\mathrm{d}}{\mathrm{d}x}C = A^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}x}B - \frac{\mathrm{d}}{\mathrm{d}x}A\right)A = 0,$$

which implies that $0 = C'(x)$ for every $x \in [a,b]$. This means that C is a constant function. □

A combination of Theorem 2.3.6 and Corollary 2.3.3 leads to the following statement: *Two matrix functions have the same left derivative on a given interval, if and only if one of the functions is obtained by multiplying the other by a constant matrix from right.* This is the fundamental theorem of Volterra’s differential calculus; a dual statement is again obtained by interchanging the words “left” and

[1] [VH], p. 41
[2] [VH], p. 44
[3] [VH], p. 46

“right”. Both statements represent an analogy of the well-known theorem: *Two functions have the same derivative if and only if they differ by a constant.*

Theorem 2.3.7.[1] Consider a matrix function $A : [a, b] \to \mathbf{R}^{n\times n}$ that is differentiable and regular on $[a, b]$. Then for an arbitrary regular matrix $C \in \mathbf{R}^{n\times n}$ we have

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

Proof. A simple consequence of Theorem 2.3.2. □

Theorem 2.3.8. Let $A : [a, b] \to \mathbf{R}^{n\times n}$ be a matrix function of the form

$$A(x) = \begin{pmatrix} A_1(x) & 0 & \cdots & 0 \\ 0 & A_2(x) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_k(x) \end{pmatrix},$$

where $A_1, \ldots, A_k$ are square matrix functions. If

$$\frac{\mathrm{d}}{\mathrm{d}x}A_i(x) = B_i(x), \qquad i = 1, \ldots, k,$$

then

$$\frac{\mathrm{d}}{\mathrm{d}x}A(x) = \begin{pmatrix} B_1(x) & 0 & \cdots & 0 \\ 0 & B_2(x) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_k(x) \end{pmatrix}.$$

Proof. The statement follows from the definition of left derivative and from Theorems 2.2.3 and 2.2.4. □

2.4 Product integral of a matrix function

Consider a matrix function $A : [a, b] \to \mathbf{R}^{n\times n}$ with entries $\{a_{ij}\}_{i,j=1}^{n}$. For every 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

$$\Delta t_i = t_i - t_{i-1}, \;\; i = 1, \ldots, m,$$

$$\nu(D) = \max_{1\le i\le m} \Delta t_i.$$

[1] [VH], p. 41

We also put

$$P(A,D) = \prod_{i=m}^{1} (I + A(\xi_i)\Delta t_i) = (I + A(\xi_m)\Delta t_m)\cdots(I + A(\xi_1)\Delta t_1),$$

$$P^*(A,D) = \prod_{i=1}^{m} (I + A(\xi_i)\Delta t_i) = (I + A(\xi_1)\Delta t_1)\cdots(I + A(\xi_m)\Delta t_m).$$

Volterra now defines the left integral of A as the matrix

$$\int_a^b \{a_{ij}\} = \lim_{\nu(D)\to 0} P(A,D)$$

(in case the limit exists) and the right integral as

$$\{a_{ij}\} \int_a^b = \lim_{\nu(D)\to 0} P^*(A,D)$$

(again if the limit exists). Volterra isn't very precise about the meaning of the limit taken with respect to partitions; we make the following agreement:

If $M(D)$ is a matrix which is dependent on the choice of a tagged partition D of interval $[a,b]$, then the equality

$$\lim_{\nu(D)\to 0} M(D) = M$$

means that for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|M(D)_{ij} - M_{ij}| < \varepsilon$ for every tagged partition D of interval $[a,b]$ satisfying $\nu(D) < \delta$ and for $i,j = 1,\ldots,n$.

The following definition also introduces a different notation to better distinguish between ordinary and product integrals.

Definition 2.4.1. Consider function $A : [a,b] \to \mathbf{R}^{n\times n}$. If the limit

$$\lim_{\nu(D)\to 0} P(A,D) \quad \text{or} \quad \lim_{\nu(D)\to 0} P^*(A,D)$$

exists, it is called the left (or right) product integral of A over the interval $[a,b]$. We use the notation

$$\prod_a^b (I + A(t)\,\mathrm{d}t) = \lim_{\nu(D)\to 0} P(A,D)$$

for the left product integral and

$$(I + A(t)\,\mathrm{d}t)\prod_a^b = \lim_{\nu(D)\to 0} P^*(A,D)$$

for the right product integral.

We note that in the case when the upper limit of integration coincides with the lower limit, then

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

In the subsequent text we use the following convention: A function $A : [a, b] \to \mathbf{R}^{n\times n}$ is called Riemann integrable, if its entries a_{ij} are Riemann integrable functions on $[a, b]$. In this case we put

$$\int_a^b A(t)\,\mathrm{d}t = \left\{\int_a^b a_{ij}(t)\,\mathrm{d}t\right\}_{i,j=1}^n.$$

We will often encounter integrals of the type

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

These integrals should be interpreted as iterated integrals, where $x_k \in [a, b]$ and $x_i \in [a, x_{i+1}]$ for $i \in \{1, \ldots, k-1\}$.

Lemma 2.4.2. Let $A : [a, b] \to \mathbf{R}^{n\times n}$ be a Riemann integrable function such that $A(x)A(y) = A(y)A(x)$ for every $x, y \in [a, b]$. Then

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

$$= \frac{1}{k!}\int_a^b\int_a^b\cdots\int_a^b A(x_k)\cdots A(x_1)\,\mathrm{d}x_1\cdots\,\mathrm{d}x_k$$

for every $k \in \mathbf{N}$.

Proof. If $P(k)$ denotes all permutations of the set $\{1, \ldots, k\}$ and

$$M_\pi = \{(x_1, \ldots, x_k) \in [a, b]^k;\ x_{\pi(1)} \le x_{\pi(2)} \le \cdots \le x_{\pi(k)}\}$$

for every $\pi \in P(k)$, then

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

$$= \sum_{\pi\in P(k)}\int\!\!\int\cdots\int_{M_\pi} A(x_k)\cdots A(x_1)\,\mathrm{d}x_1\cdots\,\mathrm{d}x_k.$$

The assumption of commutativity implies

$$\int\!\!\int\cdots\int_{M_\pi} A(x_k)\cdots A(x_1)\,\mathrm{d}x_1\cdots\,\mathrm{d}x_k = \int_a^b\int_a^{x_k}\cdots\int_a^{x_2} A(x_k)\cdots A(x_1)\,\mathrm{d}x_1\cdots\,\mathrm{d}x_k$$

for every permutation $\pi \in P(k)$, which completes the proof. □

Theorem 2.4.3.[1] If $A : [a,b] \to \mathbf{R}^{n\times n}$ is a Riemann integrable function, then both product integrals exist and

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

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

Proof. Volterra's proof goes as follows: Expanding the product $P(A,D)$ gives

$$\prod_{i=m}^{1} (I + A(\xi_i)\Delta t_i) = I + \sum_{k=1}^{m} \left(\sum_{1\le i_1<\cdots<i_k\le m} A(\xi_{i_k})\cdots A(\xi_{i_1})\Delta t_{i_1}\cdots\Delta t_{i_k} \right).$$

Volterra now argues that for $\nu(D) \to 0$ we obtain

$$\sum_{1\le i_1\le m} A(\xi_{i_1})\Delta t_{i_1} \to \int_a^b A(x_1)\,\mathrm{d}x_1,$$

$$\sum_{1\le i_1<i_2\le m} A(\xi_{i_2})A(\xi_{i_1})\Delta t_{i_1}\Delta t_{i_2} \to \int_a^b \int_a^{x_2} A(x_2)A(x_1)\,\mathrm{d}x_1\,\mathrm{d}x_2,$$

and generally

$$\sum_{1\le i_1<\cdots<i_k\le m} A(\xi_{i_k})\cdots A(\xi_{i_1})\Delta t_{i_1}\cdots\Delta t_{i_k} \to \int_a^b \int_a^{x_k} \cdots \int_a^{x_2} A(x_k)\cdots A(x_1)\,\mathrm{d}x_1\cdots\mathrm{d}x_k$$

for every $k \in \{1,\ldots,m\}$. Using the fact that $m \to \infty$ for $\nu(D) \to 0$, Volterra arrived at the result

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

The proof for the right product integral is carried out in a similar way. □

The infinite series expressing the value of the product integrals are often referred to as the Peano series. They were discussed by Giuseppe Peano in his paper [GP] from 1888 dealing with systems of linear differential equations.

Remark 2.4.4. The proof of Theorem 2.4.3 given by Volterra is somewhat unsatisfactory. First, he didn't justify that

$$\sum_{1\le i_1<\cdots<i_k\le m} A(\xi_{i_k})\cdots A(\xi_{i_1})\Delta t_{i_1}\cdots\Delta t_{i_k} \to \int_a^b \int_a^{x_k} \cdots \int_a^{x_2} A(x_k)\cdots A(x_1)\,\mathrm{d}x_1\cdots\mathrm{d}x_k$$

1 [VH], p. 49–52

for $\nu(D) \to 0$. This can be done as follows: Define

$$X^k = \{(x_1, \ldots, x_k) \in \mathbf{R}^k; a \le x_1 < x_2 < \cdots < x_k \le b\}$$

and let χ^k be the characteristic function of the set X^k. Then

$$\lim_{\nu(D)\to 0} \sum_{1\le i_1<\cdots<i_k\le m} A(\xi_{i_k})\cdots A(\xi_{i_1})\Delta t_{i_1}\cdots\Delta t_{i_k} =$$

$$= \lim_{\nu(D)\to 0} \sum_{i_1,\ldots,i_k=1}^{m} A(\xi_{i_k})\cdots A(\xi_{i_1})\chi(\xi_{i_1},\ldots,\xi_{i_k})\Delta t_{i_1}\cdots\Delta t_{i_k} =$$

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

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

The second problem is that Volterra didn't explain the equality

$$\lim_{\nu(D)\to 0}\left(I+\sum_{k=1}^{m}\sum_{1\le i_1<\cdots<i_k\le m} A(\xi_{i_k})\cdots A(\xi_{i_1})\Delta t_{i_1}\cdots\Delta t_{i_k}\right) =$$

$$I+\sum_{k=1}^{\infty}\lim_{\nu(D)\to 0}\sum_{1\le i_1<\cdots<i_k\le m} A(\xi_{i_k})\cdots A(\xi_{i_1})\Delta t_{i_1}\cdots\Delta t_{i_k}.$$

We postpone its justification to Chapter 5, Lemma 5.5.9.

Theorem 2.4.5.[1] If $A : [a,b] \to \mathbf{R}^{n\times n}$ is a Riemann integrable function, then the infinite series

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

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

converge absolutely and uniformly for $x \in [a,b]$.

Proof. We give only the proof for the first series: Its sum is a matrix whose (i,j)-th entry is the number

$$\sum_{k=1}^{\infty}\left(\sum_{l_1,\ldots,l_{k-1}=1}^{n}\int_a^x\int_a^{x_k}\cdots\int_a^{x_2} a_{i,l_1}(x_k)\cdots a_{l_{k-1},j}(x_1)\,\mathrm{d}x_1\,\cdots\,\mathrm{d}x_k\right). \qquad (2.4.1)$$

[1] [VH], p. 51–52

The functions a_{ij} are Riemann integrable, therefore bounded: There exists a positive number $M \in \mathbf{R}$ such that $|a_{ij}(t)| \le M$ for $i, j \in \{1, \ldots, n\}$ and $t \in [a, b]$. Using Lemma 2.4.2 we obtain the estimate

$$\left| \sum_{l_1,\ldots,l_{k-1}=1}^{n} \int_a^x \int_a^{x_k} \cdots \int_a^{x_2} a_{i,l_1}(x_k) \cdots a_{l_{k-1},j}(x_1) \, \mathrm{d}x_1 \cdots \mathrm{d}x_k \right| \le$$

$$\le n^{k-1} M^k \int_a^b \int_a^{x_k} \cdots \int_a^{x_2} \mathrm{d}x_1 \cdots \mathrm{d}x_k = \frac{1}{n} \frac{(nM(b-a))^k}{k!}$$

for every $x \in [a, b]$. Since

$$\sum_{k=1}^{\infty} \frac{1}{n} \frac{(nM(b-a))^k}{k!} = \frac{1}{n} e^{nM(b-a)},$$

we see that (according to the Weierstrass M-test) the infinite series (2.4.1) converges uniformly and absolutely on $[a, b]$. □

Theorem 2.4.6.[1] If $A : [a, c] \to \mathbf{R}^{n\times n}$ is a Riemann integrable function, then

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

and

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

for every $c \in [a, b]$.

Proof. Take two sequences of tagged partitions $\{D_k^1\}_{k=1}^{\infty}$ of interval $[a, b]$ and $\{D_k^2\}_{k=1}^{\infty}$ of interval $[b, c]$ such that

$$\lim_{k\to\infty} \nu(D_k^1) = \lim_{k\to\infty} \nu(D_k^2) = 0.$$

If we put $D_k = D_k^1 \cup D_k^2$, we obtain a sequence of tagged partitions $\{D_k\}_{k=1}^{\infty}$ of interval $[a, c]$ such that $\lim_{k\to\infty} \nu(D_k) = 0$. Consequently

$$\prod_a^c (I + A(x)\,\mathrm{d}x) = \lim_{k\to\infty} P(A, D_k) = \lim_{k\to\infty} P(A, D_k^2) \cdot \lim_{k\to\infty} P(A, D_k^1) =$$

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

The second part is proved in the same way. □

[1] [VH], p. 54–56

Remark 2.4.7. Volterra also offers a different proof of Theorem 2.4.6, which goes as follows[1]: Denote

$$D(i) = \{x \in \mathbf{R}^i;\ a \le x_1 \le \cdots \le x_i \le c\},$$

$$D(j,i) = \{x \in \mathbf{R}^i;\ a \le x_1 \le \cdots \le x_j \le b \le x_{j+1} \le \cdots \le x_i \le c\}$$

for each pair of numbers $i \in \mathbf{N}$ and $j \in \{0, \ldots, i\}$. Clearly

$$D(i) = D(0,i) \cup \left(\bigcup_{j=1}^{i-1} D(i-j, i-j) \times D(0,j) \right) \cup D(i,i) \tag{2.4.2}$$

for every $i \in \mathbf{N}$. We have

$$\prod_a^b (I + A(x)\,\mathrm{d}x) = I + \sum_{i=1}^{\infty} \int_{D(i,i)} A(x_i) \cdots A(x_1)\,\mathrm{d}x_1 \cdots \mathrm{d}x_i,$$

$$\prod_b^c (I + A(x)\,\mathrm{d}x) = I + \sum_{i=1}^{\infty} \int_{D(0,i)} A(x_i) \cdots A(x_1)\,\mathrm{d}x_1 \cdots \mathrm{d}x_i.$$

Since both infinite series converge absolutely, their product is equal to the Cauchy product:

$$\prod_b^c (I + A(x)\,\mathrm{d}x) \prod_a^b (I + A(x)\,\mathrm{d}x) = I + \sum_{i=1}^{\infty} \int_{D(i,i)} A(x_i) \cdots A(x_1)\,\mathrm{d}x_1 \cdots \mathrm{d}x_i +$$

$$+ \sum_{i=1}^{\infty} \int_{D(0,i)} A(x_i) \cdots A(x_1)\,\mathrm{d}x_1 \cdots \mathrm{d}x_i +$$

$$+ \sum_{i=1}^{\infty} \left(\sum_{j=1}^{i-1} \left(\int_{D(0,j)} A(x_j) \cdots A(x_1)\,\mathrm{d}x_1 \cdots \mathrm{d}x_j \right) \cdot \right.$$

$$\left. \cdot \left(\int_{D(i-j,i-j)} A(x_{i-j}) \cdots A(x_1)\,\mathrm{d}x_1 \cdots \mathrm{d}x_{i-j} \right) \right) =$$

$$= I + \sum_{i=1}^{\infty} \int_{D(i)} A(x_i) \cdots A(x_1)\,\mathrm{d}x_1 \cdots \mathrm{d}x_i = \prod_a^c (I + A(x)\,\mathrm{d}x)$$

(we have used Equation (2.4.2)).

If a and b are two real numbers such that $a < b$, we usually define

$$\int_b^a f = -\int_a^b f.$$

[1] [VH], p. 54–56

The following definition assigns a meaning to product integral whose lower limit is greater than its upper limit.

Definition 2.4.8. For any function $A : [a, b] \to \mathbf{R}^{n\times n}$ we define

$$\prod_b^a (I + A(t)\,\mathrm{d}t) = \lim_{\nu(D)\to 0} \prod_{i=1}^m (I - A(\xi_i)\Delta t_i) = (I - A(t)\,\mathrm{d}t) \prod_a^b$$

and

$$(I + A(t)\,\mathrm{d}t) \prod_b^a = \lim_{\nu(D)\to 0} \prod_{i=m}^1 (I - A(\xi_i)\Delta t_i) = \prod_a^b (I - A(t)\,\mathrm{d}t),$$

provided that the integrals on the right hand sides exist.

Corollary 2.4.9. If $A : [a, b] \to \mathbf{R}^{n\times n}$ is a Riemann integrable function, then

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

$$(I + A(t)\,\mathrm{d}t) \prod_b^a = I + \sum_{k=1}^\infty (-1)^k \int_a^b \int_a^{x_k} \cdots \int_a^{x_2} A(x_k) \cdots A(x_1)\,\mathrm{d}x_1 \cdots \mathrm{d}x_k.$$

The following statement represents a generalized version of Theorem 2.4.6.

Theorem 2.4.10.[1] If $A : [p, q] \to \mathbf{R}^{n\times n}$ is a Riemann integrable function, then

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

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

for every $a, b, c \in [p, q]$.

Proof. If $a \le b \le c$, then the statement reduces to Theorem 2.4.6. Let's have a look at the case $b < a = c$: Denote

$$E(j, i) = \{x \in \mathbf{R}^i;\ b \le x_1 \le \cdots \le x_j \le a \text{ and } a \ge x_{j+1} \ge \cdots \ge x_i \ge b\}$$

for each pair of numbers $i \in \mathbf{N}$ and $j \in \{0, \ldots, i\}$. A simple observation reveals that

$$E(j, i) = E(j, j) \times E(0, i - j) \tag{2.4.3}$$

for every $i \in \mathbf{N}$ and $j \in \{1, \ldots, i - 1\}$. We also assert that

$$E(0, i) \cup E(2, i) \cup \cdots = E(1, i) \cup E(3, i) \cup \cdots \tag{2.4.4}$$

[1] [VH], p. 56–58

for every $i \in \mathbf{N}$. Indeed, if $x \in \mathbf{R}^i$ is a member of the union on the left side, then $x \in E(2k,i)$ for some k. If $2k < i$ and $x_{2k} \le x_{2k+1}$, then $x \in E(2k+1,i)$. If $2k = i$, or $2k < i$ and $x_{2k+1} < x_{2k}$, then $x \in E(2k-1,i)$. In any case, x is a member of the union on the right side; the reverse inclusion is proved similarly.

Now, the Peano series expansions might be written as

$$\prod_b^a (I + A(x)\,\mathrm{d}x) = I + \sum_{i=1}^{\infty} \int_{E(i,i)} A(x_1)\cdots A(x_i)\,\mathrm{d}x_1\cdots\mathrm{d}x_i,$$

$$\prod_a^b (I + A(x)\,\mathrm{d}x) = I + \sum_{i=1}^{\infty} (-1)^i \int_{E(0,i)} A(x_1)\cdots A(x_i)\,\mathrm{d}x_1\cdots\mathrm{d}x_i$$

(we have used Corollary 2.4.9). Since both infinite series converge absolutely, their product is equal to the Cauchy product:

$$\prod_b^a (I + A(x)\,\mathrm{d}x)\cdot \prod_a^b (I + A(x)\,\mathrm{d}x) = I + \sum_{i=1}^{\infty} \int_{E(i,i)} A(x_1)\cdots A(x_i)\,\mathrm{d}x_1\cdots\mathrm{d}x_i +$$

$$+ \sum_{i=1}^{\infty} (-1)^i \int_{E(0,i)} A(x_1)\cdots A(x_i)\,\mathrm{d}x_1\cdots\mathrm{d}x_i +$$

$$+ \sum_{i=1}^{\infty} \left(\sum_{j=1}^{i-1} \left(\int_{E(j,j)} A(x_1)\cdots A(x_k)\,\mathrm{d}x_1\cdots\mathrm{d}x_k \right) \cdot \right.$$

$$\left. \cdot \left((-1)^{i-j} \int_{E(0,i-j)} A(x_1)\cdots A(x_{i-j})\,\mathrm{d}x_1\cdots\mathrm{d}x_{i-j} \right) \right) =$$

$$= I + \sum_{i=1}^{\infty} \sum_{j=0}^{i} (-1)^{i-j} \int_{E(j,i)} A(x_1)\cdots A(x_i)\,\mathrm{d}x_1\cdots\mathrm{d}x_i,$$

where the last equality is a consequence of Equation (2.4.3). Equation (2.4.4) implies

$$\sum_{j=0}^{i} (-1)^{i-j} \int_{E(j,i)} A(x_1)\cdots A(x_i)\,\mathrm{d}x_1\cdots\mathrm{d}x_i = 0$$

for every positive number i, which proves that

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

We see that our statement is true is even in the case $b > a = c$.

The remaining cases are now simple to check: For example, if $a < c < b$, then

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

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

The prove the second part, we calculate

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

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

□

Corollary 2.4.11. If $A : [a, b] \to \mathbf{R}^{n\times n}$ is a Riemann integrable function, then

$$\prod_b^a (I + A(x)\,\mathrm{d}x) = \left(\prod_a^b (I + A(x)\,\mathrm{d}x)\right)^{-1},$$

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

Theorem 2.4.12.[1] If $A : [a, b] \to \mathbf{R}^{n\times n}$ is a Riemann integrable function, then the functions

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

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

satisfy the integral equations

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

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

for every $x \in [a, b]$.

Proof. Theorem 2.4.3 implies

$$A(t)Y(t) = A(t) + \sum_{k=1}^{\infty} \int_a^t \int_a^{x_k} \cdots \int_a^{x_2} A(t)A(x_k)\cdots A(x_1)\,\mathrm{d}x_1 \cdots \mathrm{d}x_k. \qquad (2.4.5)$$

---

[1] [VH], p. 52–53

The Peano series converges uniformly and the entries of A are bounded, therefore the series (2.4.5) also converges uniformly for $t \in [a,b]$ and might be integrated term by term to obtain

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

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

The other integral equation is deduced similarly. □

2.5 Continuous matrix functions

Volterra is now ready to state and prove the fundamental theorem of calculus for product integral. Recall that the ordinary fundamental theorem has two parts:
1) If f is a continuous function on $[a,b]$, then the function $F(x) = \int_a^x f(t)\,\mathrm{d}t$ satisfies $F'(x) = f(x)$ for every $x \in [a,b]$.
2) If f is a continuous function on $[a,b]$ and F its antiderivative, then

$$\int_a^b f(t)\,\mathrm{d}t = F(b) - F(a).$$

The function $\int_a^x f(t)\,\mathrm{d}t$ is usually referred to as the indefinite integral of f; similarly, the functions $\prod_a^x (I + A(t)\,\mathrm{d}t)$ and $(I + A(t)\,\mathrm{d}t)\prod_a^x$ are called the indefinite product integrals of A.

Before proceeding to the fundamental theorem we make the following agreement: A matrix function $A : [a,b] \to \mathbf{R}^{n\times n}$ is called continuous, if the entries a_{ij} of A are continuous functions on $[a,b]$.

Theorem 2.5.1.[1] If $A : [a,b] \to \mathbf{R}^{n\times n}$ is a continuous matrix function, then the indefinite product integrals

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

satisfy the equations

$$Y'(x) = A(x)Y(x),$$
$$Z'(x) = Z(x)A(x)$$

for every $x \in [a,b]$.

Proof. The required statement is easily deduced by differentiating the integral equations obtained in Theorem 2.4.12. □

[1] [VH], p. 60–61

The differential equations from the previous theorem can be rewritten in the form

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

which closely resembles the first part of the ordinary fundamental theorem. We see that the left (or right) derivative is in a certain sense inverse operation to the left (or right) product integral.

Remark 2.5.2. A function $Y : [a, b] \to \mathbf{R}^{n\times n}$ is a solution of the equation

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

and satisfies $Y(a) = I$ if and only if Y solves the integral equation

$$Y(x) = I + \int_a^x A(t)Y(t)\,\mathrm{d}t, \quad x \in [a, b]. \tag{2.5.1}$$

This is a special type of equation of the form

$$y(x) = f(x) + \int_a^x K(x, t)y(t)\,\mathrm{d}t,$$

which is today called the Volterra's integral equation of the second kind. Volterra proved (see e.g. [Kl, VV4]) that such equations may be solved by the method of successive approximations; in case of Equation (2.5.1) we obtain the solution

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

which is exactly the Peano series.

Theorem 2.5.3.[1] Consider a continuous matrix function $A : [a, b] \to \mathbf{R}^{n\times n}$. If there exists a function $Y : [a, b] \to \mathbf{R}^{n\times n}$ such that

$$\frac{\mathrm{d}}{\mathrm{d}x}\,Y(x) = A(x)$$

for every $x \in [a, b]$, then

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

Similarly, if there exists a function $Z : [a, b] \to \mathbf{R}^{n\times n}$ such that

$$Z(x)\frac{\mathrm{d}}{\mathrm{d}x} = A(x)$$

[1] [VH], p. 62–63

for every $x \in [a,b]$, then

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

Proof. We prove the first part: The functions $\prod_a^x(I + A(t)\,\mathrm{d}t)$ and $Y(x)$ have the same left derivative for every $x \in [a,b]$. Theorem 2.3.6 implies the existence of a matrix C such that

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

for every $x \in [a,b]$. Substituting $x = a$ yields $C = Y(a)^{-1}$. □

Theorem 2.5.4.[1] If $A : [a,b] \to \mathbf{R}^{n\times n}$ is a continuous function, then the function

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

is the fundamental matrix of the system of differential equations

$$y_i'(x) = \sum_{j=1}^n a_{ij}(x)y_j(x), \quad i = 1,\dots,n. \tag{2.5.2}$$

Proof. Let y^k denote the k-th column of Y, i.e.

$$y^k(x) = \prod_a^x(I + A(t)\,\mathrm{d}t)\cdot e_k,$$

where e_k is the k-th vector from the canonical basis of $\mathbf{R}^n$. Theorem 2.5.1 implies that each of the vector functions y^k, $k = 1,\dots,n$ yields a solution of the system (2.5.2). Since $y^k(a) = e_k$, the system of functions $\{y^k\}_{k=1}^n$ is linearly independent and represents a fundamental set of solutions of the system (2.5.2). □

Example 2.5.5.[2] Volterra now shows the familiar method of converting a linear differential equation of the n-th order

$$y^{(n)}(x) = p_1(x)y^{n-1}(x) + p_2(x)y^{n-2}(x) + \cdots + p_n(x)y(x)$$

to a system of equations of the first order. If we introduce the functions $z_0 = y$, $z_1 = z_0'$, $z_2 = z_1', \dots, z_{n-1} = z_{n-2}'$, then the above given n-th order equation is equivalent to the system of equations written in matrix form as

$$\begin{pmatrix} z_0' \\ z_1' \\ \vdots \\ z_{n-2}' \\ z_{n-1}' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ p_n & p_{n-1} & p_{n-2} & \cdots & p_1 \end{pmatrix} \begin{pmatrix} z_0 \\ z_1 \\ \vdots \\ z_{n-2} \\ z_{n-1} \end{pmatrix}.$$

1 [VH], p. 69
2 [VH], p. 70

The fundamental matrix of this system can be calculated using the product integral and the solution of the original equation (which corresponds to the function z_0) is represented by the first column of the matrix (we obtain a set of n linearly independent solutions).

Example 2.5.6. If $n = 1$, then the function $A : [a, b] \to \mathbf{R}^{n\times n}$ is in fact a scalar function, and we usually write $\prod_a^b (1 + A(t)\,\mathrm{d}t)$ instead of $\prod_a^b (I + A(t)\,\mathrm{d}t)$. Using Theorem 2.4.3 and Lemma 2.4.2 we obtain

$$y(x) = \prod_a^x (1 + A(t)\,\mathrm{d}t) = 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \left(\int_a^x A(t)\,\mathrm{d}t \right)^k = \exp\left(\int_a^x A(t)\,\mathrm{d}t \right),$$

which is indeed a solution of the differential equation $y'(x) = A(x)y(x)$ and satisfies $y(a) = 1$.

Example 2.5.7.[1] Recall that if $A \in \mathbf{R}^{n\times n}$, then the exponential of A is defined as

$$\exp A = \sum_{k=0}^{\infty} \frac{A^k}{k!} \qquad (2.5.3).$$

The fundamental matrix of the system of equations

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

is given by

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

$$I + \sum_{k=1}^{\infty} \frac{(x-a)^k A^k}{k!} = e^{(x-a)A}$$

(we have used Theorem 2.4.3 and Lemma 2.4.2), which is a well-known result from the theory of differential equations. We also remark that a similar calculation leads to the relation

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

Example 2.5.8.[2] Volterra is also interested in actually calculating the matrix $e^{A(x-a)}$. Convert A to the Jordan normal form

$$A = C^{-1} \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{pmatrix} C, \quad \text{where } J_i = \begin{pmatrix} \lambda_i & 0 & \cdots & 0 & 0 \\ 1 & \lambda_i & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \lambda_i \end{pmatrix}$$

---

[1] [VH], p. 70–71

[2] [VH], p. 66–68

for $i \in \{1, \dots, k\}$. If

$$S_i(x) = \begin{pmatrix} e^{\lambda_i x} & 0 & 0 & \cdots & 0 & 0 \\ \frac{x^1}{1!} e^{\lambda_i x} & e^{\lambda_i x} & 0 & \cdots & 0 & 0 \\ \frac{x^2}{2!} e^{\lambda_i x} & \frac{x^1}{1!} e^{\lambda_i x} & e^{\lambda_i x} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \end{pmatrix}$$

is a square matrix which has the same dimensions as J_i, it is easily verified that

$$S_i(x)^{-1} = S_i(-x),$$

$$\frac{\mathrm{d}}{\mathrm{d}x} S_i(x) = S_i(x)' S_i(x)^{-1} = J_i.$$

Applying Theorem 2.3.8 to matrix

$$S(x) = \begin{pmatrix} S_1(x) & 0 & \cdots & 0 \\ 0 & S_2(x) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_k(x) \end{pmatrix}$$

we obtain

$$\frac{\mathrm{d}}{\mathrm{d}x} S(x) = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{pmatrix}$$

and Theorem 2.3.7 gives

$$\frac{\mathrm{d}}{\mathrm{d}x} \left(C^{-1} S(x)\right) = C^{-1} \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{pmatrix} C = A.$$

Theorem 2.3.6 implies the existence of a matrix $D \in \mathbf{R}^{n \times n}$ such that $e^{(x-a)A} = C^{-1} S(x) D$ for every $x \in [a, b]$; substituting $x = a$ yields $D = S(a)^{-1} C$.

Remark 2.5.9. Volterra gives no indication how to "guess" the calculation in the previous example. We may proceed as follows: Let again $A = C^{-1} J C$, where

$$J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{pmatrix}, \quad J_i = \begin{pmatrix} \lambda_i & 0 & \cdots & 0 & 0 \\ 1 & \lambda_i & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \lambda_i \end{pmatrix}.$$

The definition of matrix exponential implies

$$\exp(Ax) = \exp(C^{-1} J x C) = C^{-1} \exp(Jx) C$$

for every $x \in \mathbf{R}$ and it suffices to calculate $\exp(Jx)$. We see that

$$J_i x = x \begin{pmatrix} \lambda_i & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix} + x \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

It is easy to calculate an arbitrary power of the matrices on the right hand side (the second one is a nilpotent matrix); the definition of matrix exponential then gives

$$\exp(J_i x) = \begin{pmatrix} e^{\lambda_i x} & 0 & 0 & \cdots & 0 & 0 \\ \frac{x^1}{1!} e^{\lambda_i x} & e^{\lambda_i x} & 0 & \cdots & 0 & 0 \\ \frac{x^2}{2!} e^{\lambda_i x} & \frac{x^1}{1!} e^{\lambda_i x} & e^{\lambda_i x} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \end{pmatrix},$$

$$\exp(Jx) = \begin{pmatrix} \exp(J_1 x) & 0 & \cdots & 0 \\ 0 & \exp(J_2 x) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \exp(J_k x) \end{pmatrix}.$$

Theorem 2.5.10.[1] If $A : [a,b] \to \mathbf{R}^{n\times n}$ is a continuous function and $\varphi : [c,d] \to [a,b]$ a continuously differentiable function such that $\varphi(c) = a$ and $\varphi(d) = b$, then

$$\prod_a^b (I + A(x)\,\mathrm{d}x) = \prod_c^d (I + A(\varphi(t))\varphi'(t)\,\mathrm{d}t).$$

Proof. Define

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

for every $x \in [a,b]$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}(Y \circ \varphi) = Y'(\varphi(t))\varphi'(t)Y(\varphi(t))^{-1} = A(\varphi(t))\varphi'(t)$$

for every $t \in [c,d]$. The fundamental theorem for product integral gives

$$\prod_a^b (I + A(x)\,\mathrm{d}x) = Y(b)Y(a)^{-1} = Y(\varphi(d))Y(\varphi(c))^{-1} = \prod_c^d (I + A(\varphi(t))\varphi'(t)\,\mathrm{d}t).$$

□

[1] [VH], p. 65

Theorem 2.5.11.[1] If $A : [a, b] \to \mathbf{R}^{n\times n}$ is a continuous function, then

$$\det\left(\prod_a^b (I + A(x)\,\mathrm{d}x)\right) = \exp\left(\int_a^b \sum_{i=1}^n a_{ii}(x)\,\mathrm{d}x\right).$$

Proof. Denote

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

The determinant of $Y(x)$ might be interpreted as a function of the n^2 entries $y_{ij}(x)$, $i, j \in \{1, \ldots, n\}$. The chain rule therefore gives

$$(\det Y)'(x) = \sum_{i,j=1}^n \frac{\partial(\det Y)}{\partial y_{ij}} y'_{ij}(x).$$

Formula (2.2.2) for the expansion of determinant by minors implies

$$\frac{\partial(\det Y)}{\partial y_{ij}} = Y_{ij},$$

and consequently

$$(\det Y)'(x) = \sum_{i,j=1}^n y'_{ij}(x) Y_{ij}(x) = \sum_{i,j,k=1}^n a_{ik}(x) y_{kj}(x) Y_{ij}(x) =$$

$$= \sum_{i,j,k=1}^n a_{ik}(x)\delta_{ik}(x) \det Y(x) = \left(\sum_{i=1}^n a_{ii}(x)\right) \det Y(x)$$

(we have used Theorem 2.5.1 and Equation (2.2.3)). However, the differential equation

$$(\det Y)'(x) = \left(\sum_{i=1}^n a_{ii}(x)\right) \det Y(x)$$

has a unique solution that satisfies

$$\det Y(a) = \det\ I = 1.$$

It is given by

$$\det Y(x) = \exp\left(\int_a^x \sum_{i=1}^n a_{ii}(t)\,\mathrm{d}t\right), \quad x \in [a, b].$$

□

[1] [VH], p. 61–62

Theorem 2.5.12.[1] If $A : [a,b] \to \mathbf{R}^{n\times n}$ is a continuous function and $C \in \mathbf{R}^{n\times n}$ a regular matrix, then

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

Proof. Define

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

for every $x \in [a,b]$. Theorem 2.3.7 gives

$$\frac{\mathrm{d}}{\mathrm{d}x}(C^{-1}Y) = C^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}x}Y\right)C = C^{-1}AC,$$

and therefore

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

□

2.6 Multivariable calculus

In this section we turn our attention to matrix functions of several variables, i.e. to functions $A : \mathbf{R}^m \to \mathbf{R}^{n\times n}$, where $m, n \in \mathbf{N}$. We introduce the notation

$$\frac{\partial A}{\partial x_k}(x) = \left\{\frac{\partial a_{ij}}{\partial x_k}(x)\right\}_{i,j=1}^n,$$

provided the necessary partial derivatives exist.

Definition 2.6.1. Let G be a domain in $\mathbf{R}^m$ and $x \in G$. Consider a function $A : G \to \mathbf{R}^{n\times n}$ that is regular at x and such that $\frac{\partial A}{\partial x_k}(x)$ exists. We define the left partial derivative of A at x with respect to the k-th variable as

$$\frac{\mathrm{d}}{\mathrm{d}x_k}A(x) = \frac{\partial A}{\partial x_k}(x)A^{-1}(x).$$

Remark 2.6.2. Volterra also introduces the left differential of A as the matrix

$$\mathrm{d}A = A(x_1 + \mathrm{d}x_1, \ldots, x_m + \mathrm{d}x_m)A^{-1}(x_1, \ldots, x_m) = I + \sum_{k=1}^m \left(\frac{\mathrm{d}}{\mathrm{d}x_k}A(x)\right)\mathrm{d}x_k, \tag{2.6.1}$$

[1] [VH], p. 63

which differs infinitesimally from the identity matrix. He also claims that

$$\mathrm{d}A = \prod_{k=1}^{m} \left(I + \left(\frac{\mathrm{d}}{\mathrm{d}x_k} A(x) \right) \mathrm{d}x_k \right),$$

since the product of infinitesimal quantities can be neglected.

Recall the following well-known theorem of multivariable calculus: If $f_1, \dots, f_m : \mathbf{R}^m \to \mathbf{R}$ are functions that have continuous partial derivatives with respect to all variables, then the following statements are equivalent:

(1) There exists a function $F : \mathbf{R}^m \to \mathbf{R}$ such that $\frac{\partial F}{\partial x_i} = f_i$ for $i = 1, \dots, m$.

(2) $\dfrac{\partial f_i}{\partial x_j} - \dfrac{\partial f_j}{\partial x_i} = 0$ for $i, j = 1, \dots, m$, $i \neq j$.

Volterra proceeds to the formulation of a similar theorem concerning left derivatives.

Definition 2.6.3. Let $A, B : \mathbf{R}^m \to \mathbf{R}^{n \times n}$ be matrix functions that possess partial derivatives with respect to the i-th and j-th variable. We define

$$\Delta(A, B)_{x_i, x_j} = \frac{\partial B}{\partial x_i} - \frac{\partial A}{\partial x_j} + BA - AB.$$

Volterra's proof of the following lemma has been slightly modified to make it more readable. We also require the equality of mixed partial derivatives, whereas Volterra supposes that the mixed derivatives can be interchanged without any comment.

Lemma 2.6.4.[1] Let $m \in \mathbf{N}$, $i, j \in \{1, \dots, m\}$, $i \neq j$. Let G be an open set in $\mathbf{R}^m$ and $x \in G$. Consider a pair of matrix functions $X, Y : G \to \mathbf{R}^{n \times n}$ that possess partial derivatives with respect to x_i and x_j at x, and a function $S : G \to \mathbf{R}^{n \times n}$ that satisfies

$$\frac{\mathrm{d}}{\mathrm{d}x_i} S(x) = X(x), \tag{2.6.2}$$

$$\frac{\partial^2 S}{\partial x_i \partial x_j}(x) = \frac{\partial^2 S}{\partial x_j \partial x_i}(x).$$

Then the equality

$$\frac{\partial}{\partial x_i} \left(S^{-1} \left(Y - \frac{\mathrm{d}}{\mathrm{d}x_j} S \right) S \right) = S^{-1} \Delta(X, Y)_{x_i, x_j} S$$

holds at the point x.

Proof. Using the formula for the derivative of an inverse matrix and the assumption (2.6.2) we calculate

$$\frac{\partial S^{-1}}{\partial x_i} \left(Y - \frac{\mathrm{d}}{\mathrm{d}x_j} S \right) S = -S^{-1} \frac{\partial S}{\partial x_i} S^{-1} \left(Y - \frac{\mathrm{d}}{\mathrm{d}x_j} S \right) S =$$

[1] [VH], p. 81

$$= -S^{-1}X\left(Y - \frac{\mathrm{d}}{\mathrm{d}x_j}S\right)S = S^{-1}\left(-XY + X\left(\frac{\mathrm{d}}{\mathrm{d}x_j}S\right)\right)S,$$

further

$$S^{-1}\frac{\partial}{\partial x_i}\left(Y - \frac{\mathrm{d}}{\mathrm{d}x_j}S\right)S = S^{-1}\left(\frac{\partial Y}{\partial x_i} - \frac{\partial}{\partial x_i}\left(\frac{\partial S}{\partial x_j}S^{-1}\right)\right)S =$$

$$= S^{-1}\left(\frac{\partial Y}{\partial x_i} - \frac{\partial^2 S}{\partial x_i \partial x_j}S^{-1} - \frac{\partial S}{\partial x_j}\frac{\partial S^{-1}}{\partial x_i}\right)S =$$

$$= S^{-1}\left(\frac{\partial Y}{\partial x_i} - \frac{\partial (XS)}{\partial x_j}S^{-1} - \frac{\partial S}{\partial x_j}S^{-1}\frac{\partial S}{\partial x_i}S^{-1}\right)S =$$

$$= S^{-1}\left(\frac{\partial Y}{\partial x_i} - \frac{\partial X}{\partial x_j} - X\left(\frac{\mathrm{d}}{\mathrm{d}x_j}S\right) - \left(\frac{\mathrm{d}}{\mathrm{d}x_j}S\right)X\right)S =$$

$$= S^{-1}\left(\frac{\partial Y}{\partial x_i} - \frac{\partial X}{\partial x_j} - X\left(\frac{\mathrm{d}}{\mathrm{d}x_j}S\right) - \left(\frac{\mathrm{d}}{\mathrm{d}x_j}S\right)X\right)S,$$

and finally

$$S^{-1}\left(Y - \frac{\mathrm{d}}{\mathrm{d}x_j}S\right)\frac{\partial S}{\partial x_i} = S^{-1}\left(Y - \frac{\mathrm{d}}{\mathrm{d}x_j}S\right)XS = S^{-1}\left(YX - \left(\frac{\mathrm{d}}{\mathrm{d}x_j}S\right)X\right)S.$$

The product rule for differentiation gives (using the previous three equations)

$$\frac{\partial}{\partial x_i}\left(S^{-1}\left(Y - \frac{\mathrm{d}}{\mathrm{d}x_j}S\right)S\right) = S^{-1}\Bigg(-XY + X\left(\frac{\mathrm{d}}{\mathrm{d}x_j}S\right) + \frac{\partial Y}{\partial x_i} - \frac{\partial X}{\partial x_j} -$$

$$-X\left(\frac{\mathrm{d}}{\mathrm{d}x_j}S\right) - \left(\frac{\mathrm{d}}{\mathrm{d}x_j}S\right)X + YX - \left(\frac{\mathrm{d}}{\mathrm{d}x_j}S\right)X\Bigg)S = S^{-1}\Delta(X,Y)_{x_i,x_j}S.$$

□

Theorem 2.6.5.[1] If $B_1, \ldots, B_m : \mathbf{R}^m \to \mathbf{R}^{n\times n}$ are continuously differentiable with respect to all variables, then the following statements are equivalent:
(1) There exists a function $A : \mathbf{R}^m \to \mathbf{R}^{n\times n}$ such that $B_k = \frac{\mathrm{d}}{\mathrm{d}x_k}A$ for $k = 1, \ldots, m$.
(2) $\Delta(B_i, B_j)_{x_i,x_j} = 0$ for $i, j = 1, \ldots, m$, $i \neq j$.
Proof. We start with the implication (1) ⇒ (2):

$$\frac{\partial B_i}{\partial x_j} - \frac{\partial B_j}{\partial x_i} = \frac{\partial}{\partial x_j}\left(\frac{\partial A}{\partial x_i}A^{-1}\right) - \frac{\partial}{\partial x_i}\left(\frac{\partial A}{\partial x_j}A^{-1}\right) =$$

$$= \frac{\partial}{\partial x_j}\left(\frac{\partial A}{\partial x_i}\right)A^{-1} + \frac{\partial A}{\partial x_i}\frac{\partial A^{-1}}{\partial x_j} - \frac{\partial}{\partial x_i}\left(\frac{\partial A}{\partial x_j}\right)A^{-1} - \frac{\partial A}{\partial x_j}\frac{\partial A^{-1}}{\partial x_i} =$$

[1] [VH], p. 78–85

$$= \frac{\partial A}{\partial x_i}\frac{\partial A^{-1}}{\partial x_j} - \frac{\partial A}{\partial x_j}\frac{\partial A^{-1}}{\partial x_i} = -\frac{\partial A}{\partial x_i}A^{-1}\frac{\partial A}{\partial x_j}A^{-1} + \frac{\partial A}{\partial x_j}A^{-1}\frac{\partial A}{\partial x_i}A^{-1} = B_jB_i - B_iB_j$$

(statement (1) implies that the mixed partial derivatives of A are continuous, and therefore interchangeable).

The reverse implication (2) $\Rightarrow$ (1) is first proved for $m = 2$: Suppose that the function $A : \mathbf{R}^2 \to \mathbf{R}^{n\times n}$ from (2) exists. Choose $x_0 \in \mathbf{R}$ and define

$$S(x,y) = \prod_{x_0}^{x}(I + B_1(t,y)\,\mathrm{d}t).$$

Then

$$\frac{\mathrm{d}}{\mathrm{d}x}S = B_1 = \frac{\mathrm{d}}{\mathrm{d}x}A,$$

which implies the existence of a matrix function $T : \mathbf{R} \to \mathbf{R}^{n\times n}$ such that $A(x,y) = S(x,y)T(y)$ (T is independent on x). We calculate

$$\frac{\mathrm{d}}{\mathrm{d}y}T = \frac{\mathrm{d}}{\mathrm{d}y}(S^{-1}A) = S^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}y}A - \frac{\mathrm{d}}{\mathrm{d}y}S\right)S = S^{-1}\left(B_2 - \frac{\mathrm{d}}{\mathrm{d}y}S\right)S.$$

We now relax the assumption that the function A exists; the function on the right hand side of the last equation is nevertheless independent on x, because Lemma 2.6.4 gives

$$\frac{\partial}{\partial x}\left(S^{-1}\left(B_2 - \frac{\mathrm{d}}{\mathrm{d}y}S\right)S\right) = S^{-1}\Delta(B_1,B_2)_{x,y}S = 0.$$

Thus we define

$$T(y) = \prod_{y_0}^{y}\left(I + S^{-1}(x,t)\left(B_2(x,t) - \frac{\mathrm{d}}{\mathrm{d}y}S(x,t)\right)S(x,t)\,\mathrm{d}t\right)$$

(where x is arbitrary) and $A = ST$. Since

$$\frac{\mathrm{d}}{\mathrm{d}x}A = \frac{\mathrm{d}}{\mathrm{d}x}(ST) = \frac{\mathrm{d}}{\mathrm{d}x}S = B_1$$

and

$$\frac{\mathrm{d}}{\mathrm{d}y}A = \frac{\mathrm{d}}{\mathrm{d}y}(ST) = \frac{\mathrm{d}}{\mathrm{d}y}S + S\left(\frac{\mathrm{d}}{\mathrm{d}y}T\right)S^{-1} = B_2,$$

the proof is finished; we now proceed to the case $m > 2$ by induction: Choose $x_0 \in \mathbf{R}$ and define

$$S(x_1,\ldots,x_m) = \prod_{x_0}^{x_1}(I + B_1(t,x_2,\ldots,x_m)\,\mathrm{d}t).$$

If the function $A : \mathbf{R}^m \to \mathbf{R}^{n\times n}$ exists, we must have

$$\frac{\mathrm{d}}{\mathrm{d}x_1}S = B_1 = \frac{\mathrm{d}}{\mathrm{d}x_1}A$$

and consequently

$$A(x_1,\dots,x_m) = S(x_1,\dots,x_m)T(x_2,\dots,x_m)$$

for some matrix function $T : \mathbf{R}^{m-1} \to \mathbf{R}^{n\times n}$. Then

$$\frac{\mathrm{d}}{\mathrm{d}x_k}T = \frac{\mathrm{d}}{\mathrm{d}x_k}(S^{-1}A) = S^{-1}\left(B_k - \frac{\mathrm{d}}{\mathrm{d}x_k}S\right)S, \quad k = 2,\dots,m.$$

We now relax the assumption that the function A exists and define

$$U_k = S^{-1}\left(B_k - \frac{\mathrm{d}}{\mathrm{d}x_k}S\right)S, \quad k = 2,\dots,m.$$

Each of these functions U_k is indeed independent on x_1, because Lemma 2.6.4 gives

$$\frac{\partial U_k}{\partial x_1} = S^{-1}\Delta(B_1, B_k)_{x_1,x_k}S = 0.$$

Since

$$\Delta(U_i, U_j)_{x_i,x_j} = S^{-1}\Delta(B_i, B_j)_{x_i,x_j}S = 0, \quad i, j = 2,\dots,m,$$

the induction hypothesis implies the existence of a function T of $m-1$ variables $x_2,\dots,x_m$ such that

$$\frac{\mathrm{d}}{\mathrm{d}x_k}T = U_k, \quad k = 2,\dots,m.$$

We now let $A = ST$ and obtain

$$\frac{\mathrm{d}}{\mathrm{d}x_1}A = \frac{\mathrm{d}}{\mathrm{d}x_1}(ST) = \frac{\mathrm{d}}{\mathrm{d}x_1}S = B_1$$

and

$$\frac{\mathrm{d}}{\mathrm{d}x_k}A = \frac{\mathrm{d}}{\mathrm{d}x_k}(ST) = \frac{\mathrm{d}}{\mathrm{d}x_k}S + S\left(\frac{\mathrm{d}}{\mathrm{d}x_k}T\right)S^{-1} = B_k$$

for $k = 2,\dots,m$, which completes the proof. □

Remark 2.6.6. Volterra's proof of Theorem 2.6.5 contains a deficiency: We have applied Lemma 2.6.4 to the function

$$S(x_1,\dots,x_m) = \prod_{x_0}^{x_1}(I + B_1(t, x_2,\dots,x_m)\,\mathrm{d}t)$$

without verifying that

$$\frac{\partial^2 S}{\partial x_i \partial x_1}(x) = \frac{\partial^2 S}{\partial x_1 \partial x_i}(x), \quad i \in \{2,\dots,m\}.$$

This equality follows from the well-known theorem of multivariable calculus provided the derivatives $\frac{\partial S}{\partial x_i}$ exist in some neighbourhood of x for every $i \in \{1, \dots, m\}$, and the derivatives

$$\frac{\partial^2 S}{\partial x_1 \partial x_i}(x)$$

are continuous at x for every $i \in \{2, \dots, m\}$. We have

$$\frac{\partial S}{\partial x_1} = B_1$$

and consequently

$$\frac{\partial^2 S}{\partial x_1 \partial x_i} = \frac{\partial B_1}{\partial x_i},$$

which is a continuous function for every $i \in \{2, \dots, m\}$. The existence of the derivatives $\frac{\partial S}{\partial x_i}$ for $i \in \{2, \dots, m\}$ is certainly not obvious but follows from Theorem 3.6.14 on differentiating the product integral with respect to a parameter, which will be proved in Chapter 3.

Remark 2.6.7. An analogy of Theorem 2.6.5 holds also for right derivatives; the condition $\Delta(B_i, B_j)_{x_i,x_j} = 0$ must be replaced by $\Delta^*(B_i, B_j)_{x_i,x_j} = 0$, where

$$\Delta^*(A, B)_{x_i,x_j} = \frac{\partial B}{\partial x_i} - \frac{\partial A}{\partial x_j} + AB - BA.$$

The second fundamental notion of multivariable calculus is the contour integral. While Volterra introduces only product integrals along a contour φ in $\mathbf{R}^2$, which he denotes by

$$\int_\varphi X \, \mathrm{d}x \cdot Y \, \mathrm{d}y,$$

we give a general definition for curves in $\mathbf{R}^m$; we also use a different notation.

We will always consider curves that are given using a parametrization $\varphi : [a, b] \to \mathbf{R}^m$ that is piecewise continuously differentiable, which means that $\varphi'_-(x)$ exists for $x \in (a, b]$, $\varphi'_+(x)$ exists for $x \in [a, b)$, and $\varphi'_-(x) = \varphi'_+(x)$ except a finite number of points in (a, b).

The image of the curve is then defined as

$$\langle \varphi \rangle = \varphi([a, b]) = \{\varphi(t); t \in [a, b]\}.$$

Definition 2.6.8. Consider a piecewise continuously differentiable function $\varphi : [a, b] \to \mathbf{R}^m$ and a system of m matrix functions $B_1, \dots, B_m : \langle \varphi \rangle \to \mathbf{R}^{n \times n}$. The contour product integral of these functions along φ is defined as

$$\prod_\varphi (I + B_1 \, \mathrm{d}x_1 + \cdots + B_m \, \mathrm{d}x_m) = \prod_a^b (I + (B_1(\varphi(t))\varphi'_1(t) + \cdots + B_m(\varphi(t))\varphi'_m(t)) \, \mathrm{d}t).$$

Given an arbitrary curve $\varphi : [a, b] \to \mathbf{R}^m$, we define the curve $-\varphi$ as

$$(-\varphi)(t) = \varphi(-t), t \in [-b, -a].$$

This curve has the same image as the original curve, but is traversed in the opposite direction.
For any pair of curves $\varphi_1 : [a_1, b_1] \to \mathbf{R}^m$, $\varphi_2 : [a_2, b_2] \to \mathbf{R}^m$ such that $\varphi_1(b_1) = \varphi_2(a_2)$ we define the composite curve $\varphi_1 + \varphi_2$ by

$$(\varphi_1 + \varphi_2)(t) = \begin{cases} \varphi_1(t), & t \in [a_1, b_1], \\ \varphi_2(t - b_1 + a_2), & t \in [b_1, b_1 + b_2 - a_2]. \end{cases}$$

Theorem 2.6.9.[1] Contour product integral has the following properties:
(1) If $\varphi_1 + \varphi_2$ is a curve obtained by joining two curves φ_1 and φ_2, then

$$\prod_{\varphi_1+\varphi_2} (I + B_1 \, \mathrm{d}x_1 + \cdots + B_m \, \mathrm{d}x_m) =$$

$$= \prod_{\varphi_2} (I + B_1 \, \mathrm{d}x_1 + \cdots + B_m \, \mathrm{d}x_m) \cdot \prod_{\varphi_1} (I + B_1 \, \mathrm{d}x_1 + \cdots + B_m \, \mathrm{d}x_m).$$

(2) If $-\varphi$ is a curve obtained by reversing the orientation of φ, then

$$\prod_{-\varphi} (I + B_1 \, \mathrm{d}x_1 + \cdots + B_m \, \mathrm{d}x_m) = \left(\prod_{\varphi} (I + B_1 \, \mathrm{d}x_1 + \cdots + B_m \, \mathrm{d}x_m) \right)^{-1}.$$

Proof. Let $\varphi_1 : [a_1, b_1] \to \mathbf{R}^m$, $\varphi_2 : [a_2, b_2] \to \mathbf{R}^m$. Then

$$\prod_{-\varphi} (I + B_1 \, \mathrm{d}x_1 + \cdots + B_m \, \mathrm{d}x_m) =$$

$$\prod_{b_1}^{b_1+b_2-a_2} (I + (B_1(\varphi(t - b_1 + a_2))\varphi_1'(t) + \cdots + B_m(\varphi(t - b_1 + a_2))\varphi_m'(t)) \, \mathrm{d}t) \cdot$$

$$\cdot \prod_{a_1}^{b_1} (I + (B_1(\varphi(t))\varphi_1'(t) + \cdots + B_m(\varphi(t))\varphi_m'(t)) \, \mathrm{d}t).$$

The change of variables Theorem 2.5.10 gives

$$\prod_{b_1}^{b_1+b_2-a_2} (I + (B_1(\varphi(t - b_1 + a_2))\varphi_1'(t) + \cdots + B_m(\varphi(t - b_1 + a_2))\varphi_m'(t)) \, \mathrm{d}t) =$$

[1] [VH], p. 91

$$= \prod_{a_2}^{b_2} (I + (B_1(\varphi(t))\varphi_1'(t) + \cdots + B_m(\varphi(t))\varphi_m'(t))\,\mathrm{d}t),$$

which proves the first statement. The second one is also a direct consequence of Theorem 2.5.10. Note that the change of variables theorem was proved only for continuously differentiable functions, while our contours are piecewise continuously differentiable. It is however always possible to partition the integration interval in such a way that the integrated functions are continuously differentiable on every subinterval. □

Definition 2.6.10. Let G be a subset of $\mathbf{R}^m$ and $B_1, \ldots, B_m : G \to \mathbf{R}^{n\times n}$. The contour product integral $\prod(I + B_1\,\mathrm{d}x_1 + \cdots + B_m\,\mathrm{d}x_m)$ is called path-independent in G if

$$\prod_{\varphi}(I + B_1\,\mathrm{d}x_1 + \cdots + B_m\,\mathrm{d}x_m) = \prod_{\psi}(I + B_1\,\mathrm{d}x_1 + \cdots + B_m\,\mathrm{d}x_m)$$

for each pair of curves $\varphi, \psi : [a,b] \to G$ such that $\varphi(a) = \psi(a)$ and $\varphi(b) = \psi(b)$.

Using Theorem 2.6.9 it is easy to see that the contour product integral is path-independent in G if and only if

$$\prod_{\varphi}(I + B_1\,\mathrm{d}x_1 + \cdots + B_m\,\mathrm{d}x_m) = I$$

for every closed curve φ in G.

As already mentioned, Volterra is concerned especially with curves in $\mathbf{R}^2$. His effort is directed towards proving the following theorem:

Theorem 2.6.11.[1] Let G be a simply connected domain in $\mathbf{R}^2$. Consider a pair of functions $A, B : G \to \mathbf{R}^{n\times n}$ such that $\Delta(A,B)_{x,y} = 0$ at every point of G. Then

$$\prod_{\varphi}(I + A\,\mathrm{d}x + B\,\mathrm{d}y) = I$$

for every piecewise continuously differentiable closed curve φ in G.

Although Volterra's proof is somewhat incomplete, we try to indicate its main steps in the rest of the section. Theorem 2.6.11 is of great importance for Volterra as he uses it to prove an analogue of Cauchy theorem for product integral in complex domain; this topic will be discussed in the next section.

Definition 2.6.12. A set S in $\mathbf{R}^2$ is called simple in the x-direction, if the set $S \cap \{(x, y_0);\ x \in \mathbf{R}\}$ is connected for every $y_0 \in \mathbf{R}$. Similarly, S is simple in the y-direction, if the set $S \cap \{(x_0, y);\ y \in \mathbf{R}\}$ is connected for every $x_0 \in \mathbf{R}$.

[1] [VH], p. 95

Equivalently said, the intersection of S and a line parallel to the x-axis (or the y-axis) is either an interval (possibly degenerated to a single point), or an empty set.

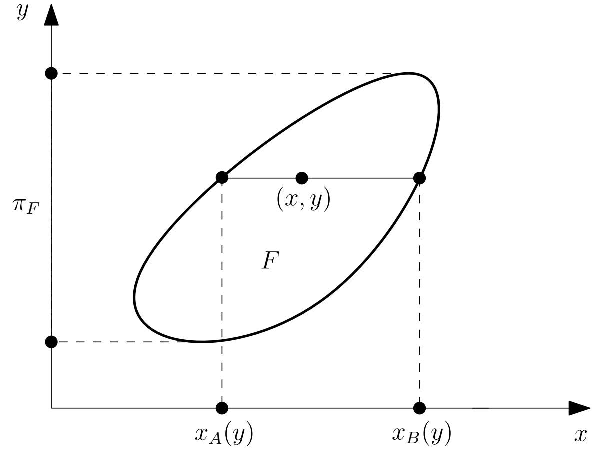
Definition 2.6.13. Let F be a closed bounded subset of $\mathbf{R}^2$ that is simple in the x-direction. For every $y \in \mathbf{R}$ denote

$$\pi_F = \{y \in \mathbf{R};\ \text{there exists } x \in \mathbf{R} \text{ such that } (x,y) \in G\}. \tag{2.6.3}$$

Further, for every $y \in \pi_F$ let

$$x_A(y) = \inf\{x;\ (x,y) \in F\}, \quad x_B(y) = \sup\{x;\ (x,y) \in F\}.$$

The meaning of these symbols is illustrated by the following figure. Note that the segment $[x_A(y), x_B(y)] \times \{y\}$ is contained in F for every $y \in \pi_F$, i.e. the set F is enclosed between the graphs of the functions $y \mapsto x_A(y)$ and $y \mapsto x_B(y)$, $y \in \pi_F$.



Definition 2.6.14. Let F be a closed bounded subset of $\mathbf{R}^2$ that is simple in the x-direction. The double product integral of a continuous function $A : F \to \mathbf{R}^{n\times n}$ is defined as

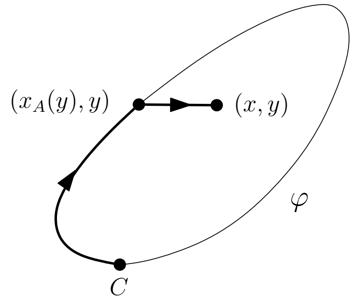
$$\prod_F (I + A(x,y)\,\mathrm{d}x\,\mathrm{d}y) = \prod_{\inf \pi_F}^{\sup \pi_F} \left(I + \left(\int_{x_A(y)}^{x_B(y)} A(x,y)\,\mathrm{d}x \right) \mathrm{d}y \right).$$

Before proceeding to the next theorem we recall that a Jordan curve is a closed curve with no self-intersections. Formally written, it is a curve with parametrization $\varphi : [a,b] \to \mathbf{R}^2$ that is injective on $[a,b)$ and $\varphi(a) = \varphi(b)$. It is known that a Jordan curve divides the plane in two components – the interior and the exterior of φ. In the following text we denote the interior of φ by $\operatorname{Int}\varphi$.

Theorem 2.6.15.[1] Consider a piecewise continuously differentiable Jordan curve $\varphi : [a,b] \to \mathbf{R}^2$ such that the set $F = \langle\varphi\rangle \cup \operatorname{Int}\varphi$ is simple in the x-direction. Assume that φ starts at a point $C = (c_x, c_y)$ such that $c_y = \inf \pi_F$ and $c_x = x_A(c_y)$. Denote

$$S(x,y) = \prod_{x_A(y)}^{x} (I + X(t,y)\,\mathrm{d}t) \prod_{C}^{(x_A(y),y)} (I + X\,\mathrm{d}x + Y\,\mathrm{d}y),$$

where the second integral is taken over that part of φ that joins the points C and $(x_A(y), y)$ (see the figure below).



Let G be an open neighbourhood of the set F. Then the equation

$$\prod_{\varphi} (I + X\,\mathrm{d}x + Y\,\mathrm{d}y) = \prod_{F} (I + S^{-1}\Delta(X,Y)_{x,y} S\,\mathrm{d}x\,\mathrm{d}y)$$

holds for each pair of continuously differentiable functions $X, Y : G \to \mathbf{R}^{n\times n}$.

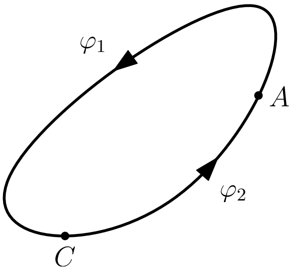
Theorem 2.6.15 might be regarded as an analogy of Green's theorem, since it provides a relationship between the double product integral over F and the contour product integral over the boundary of F. The proof in [VH] is somewhat obscure (mainly because of Volterra's calculations with infinitesimal quantities) and will not be reproduced here. A statement similar to Theorem 2.6.15 will be proved in Chapter 3, Theorem 3.7.4.

Theorem 2.6.16.[2] Consider a piecewise continuously differentiable Jordan curve $\varphi : [a,b] \to \mathbf{R}^2$ such that the set $F = \langle\varphi\rangle \cup \operatorname{Int}\varphi$ is simple in the x-direction. Let G be an open neighbourhood of the set F. If $A, B : G \to \mathbf{R}^{n\times n}$ is a pair of continuously differentiable functions such that $\Delta(A,B)_{x,y} = 0$ at every point of G, then

$$\prod_{\varphi} (I + A\,\mathrm{d}x + B\,\mathrm{d}y) = I.$$

[1] [VH], p. 92–94
[2] [VH], p. 95

Proof. Let $C = (c_x, c_y)$ be the point such that $c_y = \inf \pi_F$ and $c_x = x_A(c_y)$. If A is the starting point of φ, we may write $\varphi = \varphi_1 + \varphi_2$, where φ_1 is the part of the curve between the points A, C and φ_2 is the part between C, A provided we travel along φ in direction of its orientation.



Theorem 2.6.15 gives

$$\prod_{\varphi_2+\varphi_1} (I + A\,\mathrm{d}x + B\,\mathrm{d}y) = I,$$

and consequently

$$\prod_{\varphi}(I + A\,\mathrm{d}x + B\,\mathrm{d}y) = \prod_{\varphi_1+\varphi_2} (I + A\,\mathrm{d}x + B\,\mathrm{d}y) =$$

$$= \prod_{\varphi_2}(I + A\,\mathrm{d}x + B\,\mathrm{d}y) \prod_{\varphi_2+\varphi_1} (I + A\,\mathrm{d}x + B\,\mathrm{d}y) \left(\prod_{\varphi_2}(I + A\,\mathrm{d}x + B\,\mathrm{d}y)\right)^{-1} = I.$$

□

Remark 2.6.17. In case the set G in statement of the last theorem is simple both in the x direction and in the y direction, there is a simpler alternative proof of Theorem 2.6.16. It is based on Theorem 2.6.5, which we proved for $G = \mathbf{R}^2$, but the proof is exactly the same also for sets G which are simple in x as well as in y direction. Consequently, the assumption $\Delta(A, B)_{x,y} = 0$ and Theorem 2.6.5 imply the existence of a function $T : G \to \mathbf{R}^2$ such that

$$A(x, y) = \frac{\mathrm{d}}{\mathrm{d}x}\, T(x, y), \quad B(x, y) = \frac{\mathrm{d}}{\mathrm{d}y}\, T(x, y)$$

for every $(x, y) \in G$. Thus for arbitrary closed curve $\varphi : [a, b] \to G$ we have

$$\prod_{\varphi}(I + A\,\mathrm{d}x + B\,\mathrm{d}y) = \prod_{\varphi} \left(I + \frac{\mathrm{d}}{\mathrm{d}x}\, T\,\mathrm{d}x + \frac{\mathrm{d}}{\mathrm{d}y}\, T\,\mathrm{d}y\right) =$$

$$= \prod_{a}^{b} \left(I + \left(\frac{\partial T}{\partial x}(\varphi(t))\, T(\varphi(t))^{-1}\varphi_1'(t) + \frac{\partial T}{\partial y}(\varphi(t))\, T(\varphi(t))^{-1}\varphi_2'(t)\right)\mathrm{d}t\right) =$$

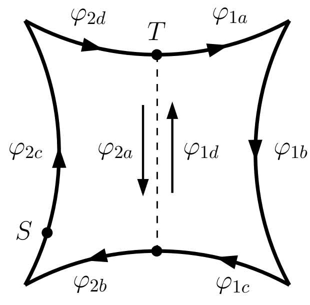
$$= \prod_a^b \left(I + \frac{\mathrm{d}}{\mathrm{d}t}(T \circ \varphi)(t)\,\mathrm{d}t \right) = T(\varphi(b))T(\varphi(a))^{-1} = I.$$

Both the statement and its proof is easily generalized to the case of curves in $\mathbf{R}^m$, $m > 2$.

Volterra now comes to the justification of Theorem 2.6.12: Let G be a simply connected domain in $\mathbf{R}^2$ and $A, B : G \to \mathbf{R}^{n\times n}$ such that $\Delta(A,B)_{x,y} = 0$ at every point of G. We have to verify that

$$\prod_{\varphi}(I + A\,\mathrm{d}x + B\,\mathrm{d}y) = I \tag{2.6.4}$$

for every piecewise continuously differentiable closed curve φ in G. Theorem 2.6.16 ensures that the statement is true, if the set $F = \langle\varphi\rangle \cup \mathrm{Int}\,\varphi$ is simple in the x-direction. Volterra first notes that it remains true, if F can be split by a curve in two parts each of which is simple in the x-direction.



Indeed, using the notation from the above figure, if $\varphi_1 = \varphi_{1a} + \varphi_{1b} + \varphi_{1c} + \varphi_{1d}$ and $\varphi_2 = \varphi_{2a} + \varphi_{2b} + \varphi_{2c} + \varphi_{2d}$, then

$$\prod_{\varphi_1}(I + A\,\mathrm{d}x + B\,\mathrm{d}y) = I, \quad \prod_{\varphi_2}(I + A\,\mathrm{d}x + B\,\mathrm{d}y) = I,$$

and thus

$$\prod_{\varphi_1+\varphi_2}(I + A\,\mathrm{d}x + B\,\mathrm{d}y) = \prod_{\varphi_2}(I + A\,\mathrm{d}x + B\,\mathrm{d}y) \cdot \prod_{\varphi_1}(I + A\,\mathrm{d}x + B\,\mathrm{d}y) = I.$$

Now if S denotes the initial point of φ, then

$$\prod_{\varphi}(I + A\,\mathrm{d}x + B\,\mathrm{d}y) =$$

$$= \prod_S^T(I + A\,\mathrm{d}x + B\,\mathrm{d}y) \cdot \prod_{\varphi_1+\varphi_2}(I + A\,\mathrm{d}x + B\,\mathrm{d}y) \cdot \prod_T^S(I + A\,\mathrm{d}x + B\,\mathrm{d}y) = I,$$

where $\prod_S^T$ denotes the contour product integral taken along the part of φ that connects the points S, T, and $\prod_T^S$ is taken along the same curve with reversed orientation (it is thus the inverse matrix of $\prod_S^T$).
By induction it follows that (2.6.4) holds if the set $F = \langle\varphi\rangle \cup \operatorname{Int}\varphi$ can be decomposed into a finite number of subsets which are simple in the x-direction. Volterra now states that this is possible for every curve φ in consideration, and so Theorem 2.6.12 is proved. He however gave no justification of the last statement, so his proof remains incomplete.

2.7 Product integration in complex domain

So far we have been concerned with real matrix functions defined on a real interval, i.e. with functions $A : [a, b] \to \mathbf{R}^{n\times n}$. Most of our results can be, without greater effort, generalized to complex-valued matrix functions, i.e. to functions $A : [a, b] \to \mathbf{C}^{n\times n}$. However, in the following two sections, we will focus our interest to matrix functions of a complex variable, i.e. $A : G \to \mathbf{C}^{n\times n}$, where G is a subset of the complex plane.
A matrix function $A = \{a_{jk}\}_{j,k=1}^n$ will be called differentiable at a point $z \in \mathbf{C}$, if its entries a_{jk} are differentiable at that point. We will use the notation

$$A'(z) = \{a'_{jk}(z)\}_{j,k=1}^n.$$

The function A is called holomorphic in an open domain $G \subseteq \mathbf{C}$, if it is differentiable everywhere in G.

Definition 2.7.1. The left derivative of a complex matrix function A at a point $z \in \mathbf{C}$ is defined as

$$\frac{\mathrm{d}}{\mathrm{d}z}A = A'(z)A^{-1}(z),$$

provided that A is differentiable and regular at the point z.

Each matrix function A of a complex variable z might be interpreted as a function of two real variables x, y, where $z = x + iy$. The Cauchy-Riemann equation states that

$$A'(z) = \frac{\partial A}{\partial x}(x+iy) = \frac{1}{i}\frac{\partial A}{\partial y}(x+iy),$$

thus the left derivative satisfies

$$\frac{\mathrm{d}}{\mathrm{d}z}A = \frac{\mathrm{d}}{\mathrm{d}x}A = \frac{1}{i}\frac{\mathrm{d}}{\mathrm{d}y}A,$$

provided all the derivatives exist.
We now proceed to the definition of product integral along a contour in the complex domain. We again restrict ourselves to contours with a piecewise continuously differentiable parametrization $\varphi : [a, b] \to \mathbf{C}$, i.e. $\varphi'_-(x)$ exists for all $x \in (a, b]$,

$\varphi'_+(x)$ exists for all $x \in [a,b)$, and $\varphi'_-(x) = \varphi'_+(x)$ except a finite number of points in (a,b).

Definition 2.7.2. Let $\varphi : [a,b] \to \mathbf{C}$ be a piecewise continuously differentiable contour in the complex plane and A a matrix function which is defined and continuous on $\langle\varphi\rangle$. The left product integral along φ is defined as

$$\prod_{\varphi}(I + A(z)\,\mathrm{d}z) = \prod_{a}^{b}(I + A(\varphi(t))\varphi'(t)\,\mathrm{d}t). \tag{2.7.1}$$

As Volterra remarks, the left contour product integral is equal to the limit

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

where D is a tagged partition of $[a,b]$ with division points t_i, tags $\xi_i \in [t_{i-1}, t_i]$ and

$$P(A,D) = \prod_{i=m}^{1}\left(I + A(\varphi(\xi_i))(\varphi(t_i) - \varphi(t_{i-1}))\right). \tag{2.7.2}$$

Instead of our $\prod_{\varphi}(I + A(z)\,\mathrm{d}z)$ he uses the notation $\int_{\varphi} A(z)\,\mathrm{d}z$.
The product integral $\prod_{\varphi}(I+A(z)\,\mathrm{d}z)$ can be converted to a contour product integral taken along a contour $\tilde{\varphi}$ in $\mathbf{R}^2$ with the parametrization

$$\tilde{\varphi}(t) = (\mathrm{Re}\ \varphi(t), \mathrm{Im}\ \varphi(t)), \quad t \in [a,b].$$

Indeed, define $A_1(x,y) = A(x+iy)$ and $A_2(x,y) = iA(x+iy)$. Then

$$\prod_{\varphi}(I + A(z)\,\mathrm{d}z) = \prod_{a}^{b}(I + A(\varphi(t))\varphi'(t)\,\mathrm{d}t) =$$

$$\prod_{a}^{b}(I + (A_1(\varphi(t))\mathrm{Re}\ \varphi'(t) + A_2(\varphi(t))\mathrm{Im}\ \varphi'(t))\,\mathrm{d}t),$$

thus

$$\prod_{\varphi}(I + A(z)\,\mathrm{d}z) = \prod_{\tilde{\varphi}}(I + A(x+iy)\,\mathrm{d}x + iA(x+iy)\,\mathrm{d}y). \tag{2.7.3}$$

The following theorem is an analogy of Theorem 2.6.9. It can be proved directly in the same way as Theorem 2.6.9, or alternatively by using the relation (2.7.3) and applying Theorem 2.6.9.

Theorem 2.7.3.[1] The left contour product integral has the following properties:

[1] [VH], p. 107

(1) If $\varphi_1 + \varphi_2$ is a curve obtained by joining two curves φ_1 and φ_2, then

$$\prod_{\varphi_1+\varphi_2} (I + A(z)\,\mathrm{d}z) = \prod_{\varphi_2} (I + A(z)\,\mathrm{d}z) \cdot \prod_{\varphi_1} (I + A(z)\,\mathrm{d}z).$$

(2) If $-\varphi$ is a curve obtained by reversing the orientation of φ, then

$$\prod_{-\varphi} (I + A(z)\,\mathrm{d}z) = \left(\prod_{\varphi} (I + A(z)\,\mathrm{d}z) \right)^{-1}.$$

Our interest in product integral of a matrix function $A : [a,b] \to \mathbf{R}^{n\times n}$ stems from the fact that it provides a solution of the differential equation (or a system of equations)

$$y'(x) = A(x)y(x),$$

where $y : [a,b] \to \mathbf{R}^n$. The situation is similar in the complex domain: Since the contour product integral is a limit of the products (2.7.2), we expect that the solution of the differential equation

$$y'(z) = A(z)y(z)$$

that satisfies $y(z_0) = y_0$ will be given by

$$y(z) = \left(\prod_{\varphi} (I + A(w)\,\mathrm{d}w) \right) y_0,$$

where $\varphi : [a,b] \to \mathbf{C}$ is a contour connecting the points z_0 and z. However, this definition of y is correct only if the product integral is independent on the choice of a particular contour, i.e. if

$$\prod_{\varphi} (I + A(z)\,\mathrm{d}z) = \prod_{\psi} (I + A(z)\,\mathrm{d}z),$$

whenever φ and ψ are two curves with the same initial points and the same endpoints. From Theorem 2.7.3 we see that $\prod_{\varphi+(-\psi)} (I + A(z)\,\mathrm{d}z)$ should be the identity matrix. Equivalently said,

$$\prod_{\varphi} (I + A(z)\,\mathrm{d}z) = I$$

should hold for every closed contour φ.

Volterra proves that the last condition is satisfied in every simply connected domain G provided that the function A is holomorphic in G. He first uses the formula (2.7.3) to convert the integral in complex domain to an integral in $\mathbf{R}^2$. Then, since

$$\Delta(A, iA)_{x,y} = \frac{\partial iA}{\partial x} - \frac{\partial A}{\partial y} + iAA - AiA = 0,$$

Theorem 2.6.11 implies that the contour product integral along any closed curve in G is equal to the identity matrix. Because we didn't prove Theorem 2.6.11, we offer a different justification taken over from [DF].

Theorem 2.7.4.[1] If $G \subseteq \mathbf{C}$ is a simply connected domain and $A : G \to \mathbf{C}^{n\times n}$ a holomorphic function in G, then the contour product integral of A is path-independent in G.

Proof. Let $\varphi : [a, b] \to G$ be a curve in G. We expand the product integral of A along φ to the Peano series

$$\prod_{\varphi}(I + A(z)\,\mathrm{d}z) = \prod_{a}^{b}(I + A(\varphi(t))\varphi'(t)\,\mathrm{d}t) =$$

$$= I + \int_a^b A(\varphi(t))\varphi'(t)\,\mathrm{d}t + \int_a^b \int_a^{t_2} A(t_2)A(t_1)\varphi'(t_2)\varphi'(t_1)\,\mathrm{d}t_1\,\mathrm{d}t_2 + \cdots$$

This infinite series might be written as

$$\prod_{\varphi}(I + A(z)\,\mathrm{d}z) = I + \int_{\varphi(a)}^{\varphi(b)} A(z)\,\mathrm{d}z + \int_{\varphi(a)}^{\varphi(b)}\int_{\varphi(a)}^{z_2} A(z_2)A(z_1)\,\mathrm{d}z_1\,\mathrm{d}z_2 + \cdots$$

where the contour integrals are all taken along φ (or its initial segment). However, since ordinary contour integrals of holomorphic functions are path-independent in G, the sum of the infinite series depends only on the endpoints of φ. □

In case the product integral is path-independent in a given domain G, we will occasionally use the symbol

$$\prod_{z_1}^{z_2}(I + A(z)\,\mathrm{d}z)$$

to denote product integral taken along an arbitrary curve in G with initial point z_1 and endpoint z_2.

Volterra now claims that if G is a simply connected domain and A is a holomorphic matrix function in G, then the function

$$Y(z) = \prod_{z_0}^{z}(I + A(w)\,\mathrm{d}w)$$

provides a solution of the differential equation $Y'(z) = A(z)Y(z)$ in G.

Theorem 2.7.5. If $G \subseteq \mathbf{C}$ is a simply connected domain and $A : G \to \mathbf{C}^{n\times n}$ a holomorphic function in G, then the function

$$Y(z) = \left(\prod_{z_0}^{z}(I + A(w)\,\mathrm{d}w)\right)$$

[1] [DF], p. 62–63

satisfies $Y'(z) = A(z)Y(z)$ in G.

Proof. The statement is obtained by differentiating the series

$$\prod_{z_0}^{z}(I + A(w)\,\mathrm{d}w) = I + \int_{z_0}^{z} A(w)\,\mathrm{d}w + \int_{z_0}^{z}\int_{z_0}^{w_2} A(w_2)A(w_1)\,\mathrm{d}w_1\,\mathrm{d}w_2 + \cdots$$

with respect to z. □

Corollary 2.7.6. Let $G \subseteq \mathbf{C}$ be a simply connected domain and $A : G \to \mathbf{C}^{n\times n}$ a holomorphic function. If $z_0 \in G$ and $y_0 \in \mathbf{C}^n$, then the function $y : G \to \mathbf{C}^n$ defined by

$$y(z) = \left(\prod_{z_0}^{z}(I + A(w)\,\mathrm{d}w)\right) y_0$$

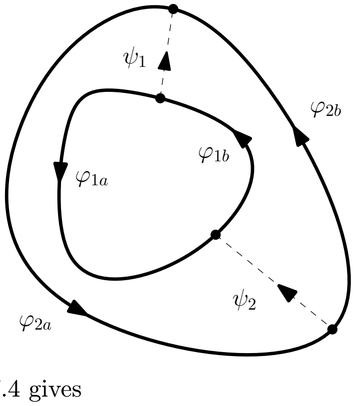
satisfies $y'(z) = A(z)y(z)$ in G and $y(z_0) = y_0$.

Theorem 2.7.7.[1] Let $G \subseteq \mathbf{C}$ be a domain and $A : G \to \mathbf{C}^{n\times n}$ a holomorphic matrix function in G. If $\varphi_1, \varphi_2 : [a, b] \to G$ are two positively oriented Jordan curves such that $\varphi_1 \subset \operatorname{Int}\varphi_2$ and $\operatorname{Int}\varphi_2 \setminus \operatorname{Int}\varphi_1 \subset G$, then

$$\prod_{\varphi_1}(I + A(z)\,\mathrm{d}z) \quad \text{and} \quad \prod_{\varphi_2}(I + A(z)\,\mathrm{d}z)$$

are similar matrices.

Proof. We introduce two disjoint auxiliary segments ψ_1, ψ_2 that connect the curves φ_1, φ_2 (see the figure).



$$\varphi_1 = \varphi_{1a} + \varphi_{1b},$$

$$\varphi_2 = \varphi_{2a} + \varphi_{2b}$$

Theorem 2.7.4 gives

$$\prod_{\varphi_{2a}}(I + A(z)\,\mathrm{d}z) \cdot \prod_{\psi_1}(I + A(z)\,\mathrm{d}z) \cdot \left(\prod_{\varphi_{1a}}(I + A(z)\,\mathrm{d}z)\right)^{-1} = \left(\prod_{\psi_2}(I + A(z)\,\mathrm{d}z)\right)^{-1}$$

[1] [VH], p. 114–116

and

$$\left(\prod_{\varphi_{1b}}(I+A(z)\,\mathrm{d}z)\right)^{-1}\cdot\left(\prod_{\psi_1}(I+A(z)\,\mathrm{d}z)\right)^{-1}\cdot\prod_{\varphi_{2b}}(I+A(z)\,\mathrm{d}z)=\prod_{\psi_2}(I+A(z)\,\mathrm{d}z).$$

Multiplying the first equality by the second from left yields

$$\left(\prod_{\varphi_{1b}}(I+A(z)\,\mathrm{d}z)\right)^{-1}\cdot\left(\prod_{\psi_1}(I+A(z)\,\mathrm{d}z)\right)^{-1}\cdot\prod_{\varphi_{2b}}(I+A(z)\,\mathrm{d}z)\cdot\prod_{\varphi_{2a}}(I+A(z)\,\mathrm{d}z)\cdot$$

$$\cdot\prod_{\psi_1}(I+A(z)\,\mathrm{d}z)\cdot\left(\prod_{\varphi_{1a}}(I+A(z)\,\mathrm{d}z)\right)^{-1}=I,$$

which can be simplified to

$$\left(\prod_{\psi_1}(I+A(z)\,\mathrm{d}z)\right)^{-1}\cdot\prod_{\varphi_2}(I+A(z)\,\mathrm{d}z)\cdot\prod_{\psi_1}(I+A(z)\,\mathrm{d}z)=\prod_{\varphi_1}(I+A(z)\,\mathrm{d}z). \quad (2.7.4)$$

□

Remark 2.7.8. Volterra offers a slightly different proof of the previous theorem: From Theorem 2.7.4 he deduces that

$$\left(\prod_{\varphi_2}(I+A(z)\,\mathrm{d}z)\right)^{-1}\cdot\left(\prod_{\psi_1}(I+A(z)\,\mathrm{d}z)\right)^{-1}\cdot\prod_{\varphi_2}(I+A(z)\,\mathrm{d}z)\prod_{\psi_1}(I+A(z)\,\mathrm{d}z)=I,$$

which implies (2.7.4). This argument is however incorrect, because the domain bounded by $\psi_1+\varphi_2-\psi_1-\varphi_2$ need not be simply connected.

Definition 2.7.9. Let $R>0$, $G=\{z\in\mathbf{C};0<|z-z_0|<R\}$. Suppose $A:G\to\mathbf{C}^{n\times n}$ is holomorphic in G. Let $\varphi:[a,b]\to G$ be a positively oriented Jordan curve, $z_0\in\operatorname{Int}\varphi$. Then

$$\prod_{\varphi}(I+A(z)\,\mathrm{d}z)=CJC^{-1},$$

where J is certain Jordan matrix, which is, according to Theorem 2.7.7, independent on the choice of φ. This Jordan matrix is called the residue of A at the point z_0.

Example 2.7.10.[1] We calculate the residue of a matrix function

$$T(z)=\frac{A}{z-z_0}+B(z)$$

[1] [VH], p. 117–120

at the point z_0, where $A \in \mathbf{C}^{n\times n}$ and B is a matrix function holomorphic in the neighbourhood $U(z_0,R)$ of point z_0. We take the product integral along the circle $\varphi_r(t) = z_0 + re^{it}$, $t \in [0,2\pi]$, $r < R$ and obtain

$$\prod_{\varphi_r}(I + T(z)\,\mathrm{d}z) = \prod_0^{2\pi}(I + iA + ire^{it}B(z_0 + re^{it})\,\mathrm{d}t).$$

Volterra suggests the following procedure (which is however not fully correct, see Remark 2.7.11): Since $iA + ire^{it}B(z_0 + re^{it}) \to iA$ for $r \to 0$, we have

$$\prod_{\varphi_r}(I + T(z)\,\mathrm{d}z) \to \prod_0^{2\pi}(I + iA\,\mathrm{d}t).$$

The integrals $\prod_{\varphi_r}(I+T(z)\,\mathrm{d}z)$, $r \in (0,R)$, are all similar to a single Jordan matrix. Their limit $\prod_0^{2\pi}(I + iA\,\mathrm{d}t)$ is thus similar to the same matrix and it is sufficient to find its Jordan normal form. By the way, this integral is equal to $e^{2\pi iA}$, giving an analogy of the residue theorem:

$$\text{The matrix } \prod_{\varphi_r}(I + T(z)\,\mathrm{d}z) \text{ is similar to } e^{2\pi iA}. \tag{2.7.5}$$

Consider the Jordan normal form of A:

$$A = C^{-1}\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_k \end{pmatrix} C, \quad \text{where } A_j = \begin{pmatrix} \lambda_j & 0 & \cdots & 0 & 0 \\ 1 & \lambda_j & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \lambda_j \end{pmatrix}.$$

Using the result of Example 2.5.8 we obtain

$$\prod_0^{2\pi}(I + iA\,\mathrm{d}t) = C^{-1}\begin{pmatrix} S_1(2\pi) & 0 & \cdots & 0 \\ 0 & S_2(2\pi) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_k(2\pi) \end{pmatrix} \cdot$$

$$\cdot \left(C^{-1}\begin{pmatrix} S_1(0) & 0 & \cdots & 0 \\ 0 & S_2(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_k(0) \end{pmatrix}\right)^{-1} = C^{-1}\begin{pmatrix} S_1(2\pi) & 0 & \cdots & 0 \\ 0 & S_2(2\pi) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_k(2\pi) \end{pmatrix} C,$$

where

$$S_j(x) = \begin{pmatrix} e^{i\lambda_j x} & 0 & 0 & \cdots & 0 & 0 \\ \frac{x^1}{1!}e^{i\lambda_j x} & e^{i\lambda_j x} & 0 & \cdots & 0 & 0 \\ \frac{x^2}{2!}e^{i\lambda_j x} & \frac{x^1}{1!}e^{i\lambda_j x} & e^{i\lambda_j x} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & & \vdots & \vdots \end{pmatrix}$$

is a square matrix of the same dimensions as A_j. The Jordan normal form of the matrix

$$\begin{pmatrix} S_1(2\pi) & 0 & \cdots & 0 \\ 0 & S_2(2\pi) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_k(2\pi) \end{pmatrix},$$

and therefore also the demanded residue, is

$$\begin{pmatrix} V_1 & 0 & \cdots & 0 \\ 0 & V_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & V_k \end{pmatrix}, \text{ where } V_j = \begin{pmatrix} e^{2\pi i\lambda_j} & 0 & \cdots & 0 & 0 \\ 1 & e^{2\pi i\lambda_j} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & e^{2\pi i\lambda_j} \end{pmatrix}$$

is a square matrix of the same dimensions as S_j.

Remark 2.7.11. The calculation from the previous example contains two deficiencies: First, Volterra interchanges the order of limit and product integral to obtain

$$\lim_{r\to 0} \prod_0^{2\pi} (I + iA + ire^{it}B(z_0 + re^{it})\,\mathrm{d}t) = \prod_0^{2\pi} (I + iA\,\mathrm{d}t)$$

without any further comment. However, the convergence $iA+ire^{it}B(z_0+re^{it}) \to iA$ for $r \to 0$ is uniform and in this case, as we will prove in Chapter 5, Theorem 5.6.4, the statement is in fact true.
The second deficiency is more serious. Volterra seems to have assumed that if some matrices $S(r)$, $r \in (0, R)$ (in our case $S(r)$ is the product integral taken along φ_r), are all similar to a single Jordan matrix J, then $\lim_{r\to 0} S(r)$ is also similar to J. This statement is incorrect, as demonstrated by the example

$$S(r) = \begin{pmatrix} 1 & 0 \\ r & 1 \end{pmatrix},$$

where $S(r)$ is similar to

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

for $r > 0$, but

$$\lim_{r\to 0} S(r) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

isn't. The mentioned statement can be proved only under additional assumptions on $S(r)$. For example, if the matrices $S(r)$, $r > 0$, have n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, then the limit matrix $\lim_{r\to 0} S(r)$ has the same eigenvalues, because

$$\det(S(0) - \lambda I) = \lim_{r\to 0} \det(S(r) - \lambda I) = (\lambda - \lambda_1)\cdots(\lambda - \lambda_n).$$

This means that all the matrices $S(r)$, $r \geq 0$, are similar to a single Jordan matrix

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

A more detailed discussion of the residue theorem for product integral can be found in the book [DF]; for example, if the set

$$\sigma(A) - \sigma(A) = \{\lambda_1 - \lambda_2;\ \lambda_1 \text{ and } \lambda_2 \text{ are eigenvalues of matrix } A\}$$

doesn't contain any positive integers, then the statement (2.7.5) is shown to be true.

2.8 Linear differential equations at a singular point

In this section we assume that the reader is familiar with the basics of the theory of analytic functions (see e.g. [EH] or [VJ]). We are interested in studying the differential equation

$$Y'(z) = A(z)Y(z), \tag{2.8.1}$$

where the function A is holomorphic in the ring $P(z_0, R) = \{z \in \mathbf{C}; 0 < |z - z_0| < R\}$ and $R > 0$. If we choose $z_1 \in P(z_0, R)$ and denote $r = \min(|z_1 - z_0|, R - |z_1 - z_0|)$, then the function

$$Y_1(z) = \prod_{z_1}^{z} (I + A(w)\,\mathrm{d}w)$$

provides a solution of (2.8.1) in $B(z_1, r) = \{z \in \mathbf{C}; |z - z_1| < r\}$; the product integral is path-independent, because A is holomorphic in $U(z_1, r)$. The holomorphic function Y_1 can be continued along an arbitrary curve φ v $P(z_0, R)$; this procedure leads to a (multiple-valued) analytic function $\mathcal{Y}$, which will be denoted by

$$\mathcal{Y}(z) = \prod_{z_1}^{z} (I + A(w)\,\mathrm{d}w), \quad z \in P(z_0, R).$$

If the element $(z_1, Y_1) \in \mathcal{Y}$ is continued along a curve φ in $P(z_0, R)$ to an element $(z_2, Y_2) \in \mathcal{Y}$ (we write this as $(z_1, Y_1) \xrightarrow{\varphi} (z_2, Y_2)$), then (using $Y_1(z_1) = I$)

$$Y_2(z_2) = \prod_{\varphi} (I + A(w)\,\mathrm{d}w) \cdot Y_1(z_1) = \prod_{\varphi} (I + A(w)\,\mathrm{d}w).$$

Let φ be the circle with center z_0 which passes through the point z_1, i.e.

$$\varphi(t) = z_0 + (z_1 - z_0)\exp(it), \qquad t \in [0, 2\pi].$$

If $(z_1, Y_2) \in \mathcal{Y}$ is the element such that $(z_1, Y_1) \stackrel{\varphi}{\to} (z_1, Y_2)$, then

$$\frac{\mathrm{d}}{\mathrm{d}z} Y_1 = \frac{\mathrm{d}}{\mathrm{d}z} Y_2 = A(z)$$

for $z \in B(z_1, r)$. Consequently, there is a matrix $C \in \mathbf{C}^{n \times n}$ such that $Y_2(z) = Y_1(z) \cdot C$. Substituting $z = z_1$ gives

$$C = \prod_{\varphi} (I + A(w)\,\mathrm{d}w).$$

Volterra refers[1] to the point z_0 as *point de ramification abélien* of the analytic function $\mathcal{Y}$; this means that it is the branch point of $\mathcal{Y}$, but not of its derivative A, which is a single-valued function. Volterra proceeds to prove that $\mathcal{Y}$ can be written in the form

$$\mathcal{Y} = \mathcal{S}_1 \cdot \mathcal{S}_2,$$

where $\mathcal{S}_1$ is single-valued in $P(z_0, R)$ and $\mathcal{S}_2$ is an analytic function that is uniquely determined by the matrix $C = \prod_{\varphi}(I + A(w)\,\mathrm{d}w)$.
Here is the proof[2]: We write $C = M^{-1}TM$, where T is a Jordan matrix. Then

$$T = \begin{pmatrix} T_1 & 0 & \cdots & 0 \\ 0 & T_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & T_k \end{pmatrix}, \quad \text{kde } T_j = \begin{pmatrix} e^{2\pi i \lambda_j} & 0 & \cdots & 0 & 0 \\ 1 & e^{2\pi i \lambda_j} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & e^{2\pi i \lambda_j} \end{pmatrix},$$

where we have expressed the eigenvalues of T in the form $\exp(2\pi i \lambda_j)$; this is certainly possible as the matrices C and consequently also T are regular, and thus have nonzero eigenvalues. We now define the analytic function

$$\mathcal{V}(z) = \prod_{z_1}^{z} \left(I + \frac{U}{w - z_0}\,\mathrm{d}w \right), \quad z \in P(z_0, R),$$

where

$$U = \begin{pmatrix} U_1 & 0 & \cdots & 0 \\ 0 & U_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & U_k \end{pmatrix} \quad \text{and} \quad U_j = \begin{pmatrix} \lambda_j & 0 & \cdots & 0 & 0 \\ 1 & \lambda_j & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \lambda_j \end{pmatrix}$$

and U_j has the same dimensions as T_j for $j \in \{1, \ldots, k\}$. Consider a function element (z_1, V_1) of $\mathcal{V}$; what happens if we continue it along the circle φ? As in the case of function $\mathcal{Y}$ we obtain the result

$$(z_1, V_1) \stackrel{\varphi}{\to} (z_1, V_1 \cdot D),$$

[1] [VH], p. 121
[2] [VH], p. 122–124

where

$$D = \prod_{\varphi} \left(I + \frac{U}{w - z_0}\,\mathrm{d}w \right) = \prod_{0}^{2\pi} (I + iU\,\mathrm{d}w).$$

In Example 2.7.10 we have calculated the result

$$D = S(2\pi) \cdot S(0)^{-1} = S(2\pi) = \begin{pmatrix} S_1(2\pi) & 0 & \cdots & 0 \\ 0 & S_2(2\pi) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_k(2\pi) \end{pmatrix},$$

where

$$S_j(x) = \begin{pmatrix} e^{i\lambda_j x} & 0 & 0 & \cdots & 0 & 0 \\ \frac{x^1}{1!} e^{i\lambda_j x} & e^{i\lambda_j x} & 0 & \cdots & 0 & 0 \\ \frac{x^2}{2!} e^{i\lambda_j x} & \frac{x^1}{1!} e^{i\lambda_j x} & e^{i\lambda_j x} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \end{pmatrix}.$$

The matrix D is similar to the Jordan matrix T; thus

$$D = N^{-1} T N,$$

and consequently

$$(z_1, V_1) \stackrel{\varphi}{\to} (z_1, V_1 N^{-1} T N),$$

$$(z_1, Y_1) \stackrel{\varphi}{\to} (z_1, Y_1 M^{-1} T M).$$

We now consider the analytic function

$$\mathcal{S}_1(z) = \mathcal{Y}(z) M^{-1} N \mathcal{V}(z)^{-1}$$

and continue its element along φ:

$$(z_1, Y_1 M^{-1} N V_1^{-1}) \stackrel{\varphi}{\to} (z_1, Y_1 M^{-1} T M M^{-1} N (V_1 N^{-1} T N)^{-1}) = (z_1, Y_1 M^{-1} N V_1^{-1}).$$

Thus the analytic function $\mathcal{S}_1$ is in fact single-valued in $P(z_0, R)$. The proof is finished by putting

$$\mathcal{S}_2(z) = \mathcal{V}(z) N^{-1} M.$$

Consequently

$$\mathcal{Y} = \mathcal{S}_1 \cdot \mathcal{S}_2$$

and $\mathcal{S}_2$ is uniquely determined by the matrix C.

We now briefly turn our attention to the analytic function $\mathcal{V}$. Assume that

$$(z_1, V_1) \stackrel{\psi}{\to} (z, V_2),$$

where $\psi : [a,b] \to P(z_0, R)$, $\psi(a) = z_1$, $\psi(b) = z$. It is known from complex analysis that given the curve φ, we can find a function a function $g : [a,b] \to \mathbf{C}$ such that

$$\exp(g(t)) = \psi(t) - z_0,$$

for every $t \in [a,b]$; g is a continuous branch of logarithm of the function $\psi - z_0$. For convenience we will use the notation

$$g(t) = \log(\psi(t) - z_0)$$

with the understanding that g is defined as above. We also have

$$g'(t) = \frac{\psi'(t)}{\psi(t) - z_0}$$

for every $t \in [a,b]$. We now calculate

$$V_2(z) = \prod_{\psi} \left(I + \frac{U}{z - z_0}\,\mathrm{d}z \right) = \prod_{a}^{b} \left(I + \frac{U}{\psi(t) - z_0}\,\psi'(t)\,\mathrm{d}t \right).$$

Substituting $v = g(t)$ gives

$$V_2(z) = \prod_{g(a)}^{g(b)} (I + U\,\mathrm{d}v) = S(g(b))S(g(a))^{-1},$$

where

$$S(z) = \begin{pmatrix} S_1(z) & 0 & \cdots & 0 \\ 0 & S_2(z) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_k(z) \end{pmatrix}$$

is a block diagonal matrix composed of the matrices

$$S_j(z) = \begin{pmatrix} e^{\lambda_j z} & 0 & 0 & \cdots & 0 & 0 \\ \frac{z^1}{1!}e^{\lambda_j z} & e^{\lambda_j z} & 0 & \cdots & 0 & 0 \\ \frac{z^2}{2!}e^{\lambda_j z} & \frac{z^1}{1!}e^{\lambda_j z} & e^{\lambda_j z} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \end{pmatrix}.$$

Consequently, the solution of Equation (2.8.1), i.e. the analytic function $\mathcal{Y}$, can be expressed as

$$\mathcal{Y}(z) = \mathcal{S}_1(z)\mathcal{S}_2(z) = \mathcal{S}_1(z)S(g(b))S(g(a))^{-1}N^{-1}M, \tag{2.8.2}$$

where

$$S(g(b)) = S(\log(z - z_0)),$$

$$S_j(\log(z-z_0)) =$$

$$= \begin{pmatrix} (z-z_0)^{\lambda_j} & 0 & 0 & \cdots & 0 & 0 \\ (z-z_0)^{\lambda_j}\log(z-z_0) & (z-z_0)^{\lambda_j} & 0 & \cdots & 0 & 0 \\ (z-z_0)^{\lambda_j}\frac{\log^2(z-z_0)}{2!} & (z-z_0)^{\lambda_j}\log(z-z_0) & (z-z_0)^{\lambda_j} & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots & \vdots \end{pmatrix}.$$

The above result can be applied to obtain the general form of solution of the differential equation

$$y^{(n)}(z) + p_1(z)y^{(n-1)}(z) + \cdots + p_n(z)y(z) = 0, \tag{2.8.3}$$

where the functions p_i are holomorphic in $P(z_0, R)$. We have seen that this equation of the n-th order is easily converted to a system of linear differential equations of first order, which can be written in the vector form as

$$y'(z) = A(z)y(z),$$

where A is a holomorphic matrix function in $P(z_0, R)$. The fundamental matrix of this system is given by (2.8.2); the first row of this matrix then yields the fundamental system of solutions (composed of n analytic functions) of Equation (2.8.3). From the form of Equation (2.8.2) we infer that every solution of Equation (2.8.3) can be expressed as a linear combination of analytic functions of the form

$$(z-z_0)^{\lambda_j}\left(\varphi_0^j(z) + \varphi_1^j(z)\log(z-z_0) + \cdots + \varphi_{n_j}^j(z)\log^{n_j}(z-z_0)\right),$$

where φ_k^j are holomorphic functions in $P(z_0, R)$.

Thus we see that Volterra was able to obtain the result of Lazarus Fuchs (see Chapter 1) using the theory of product integration. The next chapters of Volterra's book [VH] are concerned with the study of analytic functions on Riemann surfaces; the topic is rather special and we don't follow Volterra's treatment here.