LECTURES ON DIFFERENTIAL EQUATIONS

SOLOMON LEFSCHETZ

ANNALS OF MATHEMATICS STUDIES

PRINCETON UNIVERSITY PRESS

LECTURES ON DIFFERENTIAL EQUATIONS

By

SOLOMON LEFSCHETZ

PRINCETON UNIVERSITY PRESS
Princeton, New Jersey

Published by Princeton University Press
41 William Street
Princeton, New Jersey 08540

Library of Congress Card No. 1578
ISBN 0-691-08395-9

Princeton University Press books are printed on acid-free paper and
meet the guidelines for permanence and durability of the
Committee on Production Guidelines for Book
Longevity of the Council on Library Resources

Printed in the United States of America

PREFACE

The subject of differential equations in the large would seem to offer a most attractive field for further study and research. Many hold the opinion that the classical contributions of Poincaré, Liapounoff and Birkhoff have exhausted the possibilities. This is certainly not the opinion of a large school of Soviet physico mathematicians as the reader will find by consulting N. Minorsky's recent Report on Non-Linear Mechanics issued by the David Taylor Model Basin. In recent lectures at Princeton and Mexico, the author endeavored to provide the necessary background and preparation. The material of these lectures is now offered in the present monograph.

The first three chapters are self-explanatory and deal with more familiar questions. In the presentation vectors and matrices are used to the fullest extent. The fourth chapter contains a rather full treatment of the asymptotic behavior and stability of the solutions near critical points. The method here is entirely inspired by Liapounoff, whose work is less well known that it should be. In Chapter V there will be found the Poincaré-Bendixson theory of planar characteristics in the large. The very short last chapter contains an analytical treatment of certain non-linear differential equations of the second order, dealt with notably by Liénard and van der Pol, and of great importance in certain applications.

The author wishes to express his indebtedness to Messrs. Richard Bellman and Jaime Lifshitz for many valuable suggestions and corrections to this monograph. The responsibility, however, for whatever is still required along that line is wholly the author's.

TABLE OF CONTENTS

CHAPTER I

SOME PRELIMINARY QUESTIONS

§1. MATRICES

1. The reader is assumed familiar with the elements of matrix theory. The matrices $\| a_{ij} \|$, $\| x_{ij} \|$, ..., are written A, X, ... The transpose of A is written A'. The matrix diag $(A_1, \ldots, A_r)$ is

$$\begin{Vmatrix} A_1, & 0, & \ldots & \\ 0, & A_2, & \ddots & \\ & & & A_r \end{Vmatrix}$$

where the A_i are square matrices and the zeros stand for zero matrices. Noteworthy special case: $\mathrm{diag}(a_1, \ldots, a_n)$ denotes a square matrix of order n with the scalars a_i down the main diagonal and the other terms zero. In particular if $a_1 = \ldots = a_n = 1$, the matrix is written E_n or E and called a unit-matrix. The terms of E_n are written δ_{ij} and called Kronecker deltas.

2. Suppose now A square and of order n. The determinant of A is denoted by $|A|$. When $|A| = 0$, A is said to be <u>singular</u>. A non-singular matrix A possesses an inverse A^{-1} which satisfies $AA^{-1} = A^{-1}A = E$. The trace of a square matrix A written tr A, is the expression $\sum a_{ii}$. If $A^n = 0$, A is called <u>nilpotent</u>. We recall the relations

$$(AB)^{-1} = B^{-1}A^{-1} \qquad (A^{-1})' = (A')^{-1},$$

where A, B are non-singular.

If $f(\lambda) = a_0 + a_1\lambda + \ldots + a_r\lambda^r$ then $a_0E + a_1A + \ldots + a_rA^r$ has a unique meaning and is written $f(A)$. The polynomial $\phi(\lambda) = |A - \lambda E|$ is known as the <u>characteristic</u> polynomial of A, and its roots as the characteristic roots of A. (See Theorem (3.5) below.)

3. (3.1) Two real [complex] square matrices A, B of same order n are called similar in the real [complex] domain if there can be found a non-singular square real [complex] matrix P of order n such that $B = PAP^{-1}$. This relation is clearly an equivalence. For if we denote it by $\sim$ then the relation is

symmetric: $A \sim B \longrightarrow B \sim A$, since

$$B = PAP^{-1} \Longrightarrow A = P^{-1}BP$$

reflexive: $A \sim A$, since $A = EAE^{-1}$,

transitive: $A \sim B$, $B \sim C \longrightarrow A \sim C$. For if $A = PBP^{-1}$, $B = QCQ^{-1}$ then $A = PQCQ^{-1}P^{-1} = (PQ)C(PQ)^{-1}$.

(3.2) If $A \sim B$ and $f(\lambda)$ is any polynomial then $f(A) \sim f(B)$. Hence $f(A) = 0 \longrightarrow f(B) = 0$.

For if $B = PAP^{-1}$ then $B^r = PA^rP^{-1}$, $kB = P(kA)P^{-1}$, and $P(A_1 + A_2)P^{-1} = PA_1P^{-1} + PA_2P^{-1}$.

(3.3) Similar matrices have the same characteristic polynomial.

For $B = PAP^{-1} \longrightarrow B - \lambda E = P(A-\lambda E)P^{-1}$, and therefore also $|B-\lambda E| = |A-\lambda E|$.

Since the characteristic polynomials are the same, their coefficients are also the same. Only two are of interest: the determinants, manifestly equal, and the traces. If $\lambda_1, \ldots, \lambda_n$ are the characteristic roots then a ready calculation yields

$$\sum\lambda_i = \sum a_{ii} = \operatorname{tr} A.$$

Therefore

(3.4) Similar matrices have equal traces.

For the proof of the following two classical theorems the reader is referred to the standard treatises on the subject:

(3.5) Theorem. If $\phi(\lambda)$ is the characteristic polynomial of A, then $\phi(A) = 0$.

(3.6) Fundamental Theorem. Every complex square matrix A is similar in the complex domain to a matrix of the form diag $(A_1, \ldots, A_r)$ where A_i is of the form

$$(3.6.1) \qquad A_i = \left\| \begin{matrix} \lambda_j, & 0, & 0, & \ldots\ldots, & 0 \\ 1, & \lambda_j, & 0, & \ldots\ldots & 0 \\ 0, & 1, & \lambda_j, & \ldots\ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0, & \ldots & \ldots & 1, & \lambda_j \end{matrix} \right\|$$

with λ_j one of the characteristic roots. There is at least one A_i for each λ_j and if λ_j is a simple root then there is only one $A_i = \| \lambda_j \|$. Hence if the characteristic roots λ_j are all distinct, A is similar to diag $(\lambda_1, \ldots, \lambda_n)$.

By way of illustration when $n = 2$ and $\lambda_1 = \lambda_2$, we have the two distinct types

$$\left\| \begin{matrix} \lambda, & 0 \\ 0, & \lambda \end{matrix} \right\|, \qquad \left\| \begin{matrix} \lambda, & 0 \\ 1, & \lambda \end{matrix} \right\|.$$

(3.7) Real Matrices. When A is real the λ_i occur in conjugate pairs $\lambda_j, \bar{\lambda}_j$ and hence the matrices A_i occur likewise in conjugate pairs $A_j, \bar{A}_j$ where $\bar{A}_j$ is like A_j with $\bar{\lambda}_j$ instead of λ_j. Thus they may be disposed into a sequence $A_1, \ldots, A_k, \bar{A}_1, \ldots, \bar{A}_k, A_{2k+1}, \ldots, A_s$ where the A_{2k+i} correspond to the real λ_j. We will then say that the canonical form is real.

4. Limits, Series. (4.1) Let $\{A_p\}$, $A_p = \| a^p_{ij} \|$ be a sequence of matrices of order n such that $a_{ij} = \lim \{a^p_{ij}\}$ exists for every pair i, j. We then apply the customary "limit" terminology to the sequence $\{A_p\}$ and call $A = \| a_{ij} \|$ its limit. As a consequence we will naturally say that the infinite series $\sum A_p$ is convergent if the n^2 series

$$(4.2) \qquad a_{ij} = \sum a^p_{ij}$$

are convergent and the sum of the series is by definition the matrix $A = \| a_{ij} \|$.

If the a^p_{ij} are functions of a parameter t and the n series $\sum a^p_{ij}$ are uniformly convergent as to t over a certain range then $\sum A_p$ is said to be <u>uniformly convergent</u> as to t over the same range.

(4.3) Let us apply to the A_p's the simultaneous operation $B_p = PA_pQ$ where P, Q are fixed. If we set

$$S^{mr}_{ij} = \sum_{p=m+1}^{r} a^p_{ij}$$

then clearly the corresponding T^{mr}_{ij} for the B's is related to the S^{mr}_{ij} by

$$T^{mr}_{ij} = \sum p_{ih} S^{mr}_{hk} q_{kj}.$$

Now a n.a.s.c. of convergency of (4.2) may be phrased thus: for every $\varepsilon > 0$ there is an N such that $m > N \Longrightarrow |S^{mr}_{ij}| < \varepsilon$ whatever r. If $\alpha = \sup \{p_{ih}, q_{kj}\}$ then

$$\sum_{i,j} |T^{mr}_{ij}| < n^2\alpha^2 \sum_{i,j} |S^{mr}_{ij}| .$$

Hence the convergence of (4.2) implies the convergence of

$$\sum B_p = \sum PA_pQ$$

whose limit is clearly $B = PAQ$. In particular

(4.4) <u>If</u> $\{A_p\}$ <u>converges to</u> A <u>and if</u> $B_p = PA_pP^{-1}$ <u>then</u> $\{B_p\}$ <u>converges to</u> $B = PAP^{-1}$.

5. Consider a power series with complex coefficients

$$f(z) = a_0 + a_1 z + a_2 z^2 + \dots \tag{5.1}$$

whose radius of convergence $\rho > 0$. If

$$X = \| x_{ij} \| \tag{5.2}$$

is a square matrix of order n we may form the series

$$(5.3) \qquad a_0E + a_1X + a_2X^2 + \dots$$

and if it converges its limit will be written $f(X)$.

(5.4) Suppose that $X = \operatorname{diag}(X_1, X_2)$. If $g(z)$ is a scalar polynomial then $g(X) = \operatorname{diag}(g(X_1), g(X_2))$. Hence in this case (5.3) converges when and only when the same series in the X_i converge and its limit is then $f(X) = \operatorname{diag}(f(X_1), f(X_2))$.

(5.5) Theorem. Sufficient conditions for the convergence of (5.3) are that X is nilpotent or else that its characteristic roots are all less than ρ in absolute value.

Whatever the radius of convergence, when X is nilpotent the series is finite and hence evidently convergent. In the general case, X is similar to a matrix of the type described in (3.6). Remembering (5.4) we only need to consider the type (3.6.1). In other words, we may assume that the matrix is (3.6.1) itself. Thus λ is its sole characteristic root and so we will merely have to prove

(5.6) If X is (3.6.1) and $|\lambda| < \rho$ then (5.3) is convergent.

Let us set

$$(5.6.1) \qquad Z = \begin{Vmatrix} 0 & & & & \\ 1 & 0 & & & \\ 0 & 1 & \ddots & & \\ & & \ddots & & \\ & & & 1 & 0 \end{Vmatrix}$$

We verify by direct multiplication that Z^r is obtained by moving the diagonal of units so that it starts at the term in the $(r+1)^{st}$ row and first column (the term $z_{r+1,1}$). Hence

$$(5.6.2) \qquad Z^n = 0.$$

Now $X = \lambda E + Z$, and since E commutes with every matrix:

$$(5.6.3) \quad X^p = \lambda^p E + \frac{p}{1!}\lambda^{p-1}Z + \dots + \binom{p}{n-1}\lambda^{p-n+1}Z^{n-1}$$

Hence

(5.7)
$$f(X) = \begin{vmatrix} f(\lambda), & 0, & \dots & & 0 \\ \frac{f'(\lambda)}{1!}, & f(\lambda), & 0 & \dots & 0 \\ \frac{f''(\lambda)}{2!}, & \frac{f'(\lambda)}{1!}, & f(\lambda), 0, & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \frac{f^{(n-1)}(\lambda)}{(n-1)!}, & \dots & \dots & \dots & f(\lambda) \end{vmatrix}$$

Since $|\lambda| < \rho$, $f(z)$ and all its derivatives converge for $z = \lambda$. Hence $f(X)$ converges. This proves (5.6) and therefore also (5.5).

(5.8) Returning now to the arbitrary matrix X we note that we can define in particular

$$e^X = E + \frac{X}{1!} + \frac{X^2}{2!} \dots$$

for every X, and also

$$\log (E + X) = \frac{X}{1} - \frac{X^2}{2} + \dots$$

when the characteristic roots are less than one in absolute value.

(5.9) The usual rules for adding, multiplying, differentiating and generally combining series in X hold here also. However those for multiplying series in X by series in Y hold only when X and Y are commutative. Thus we may prove $e^{X+Y} = e^X \cdot e^Y$ when X and Y commute, but not so in the contrary case.

(5.10) <u>If</u> $f(X)$ <u>converges and</u> $Y = PXP^{-1}$, $|P| \neq 0$, <u>then</u> $f(Y)$ <u>converges also</u> (4.4) <u>and</u> $f(Y) = P(f(X))P^{-1}$.

6. (6.1) <u>If</u> $|X| \neq 0$ <u>there is a</u> Y <u>such that</u> $e^Y = X$.

Since Y need not be unique we do not insist on designating it by $\log X$.

Referring to (5.4) if $X = \operatorname{diag}(X_1, X_2)$ and if we can find Y_1, Y_2 such that $X_1 = e^{Y_1}$ then $X = \operatorname{diag}(e^{Y_1}, e^{Y_2})$ and so $Y = \operatorname{diag}(Y_1, Y_2)$ answers the question. Hence as in the proof of (5.5) we need only consider X of the type (3.6.1). Here $|X| \neq 0 \longrightarrow \lambda \neq 0$. In the same notations as before $X = \lambda E + Z = \lambda(E + \frac{1}{\lambda} Z)$. Since Z is nilpotent so is $\frac{1}{\lambda} Z$, and therefore we may define by (5.5) the function

$$Y_1 = \log (E + \frac{1}{\lambda} Z).$$

By (5.9):

$$E + \frac{1}{\lambda} Z = e^{Y_1}, \qquad X = \lambda E + Z = \lambda e^{Y_1}.$$

Since $\lambda \neq 0$ we may find a scalar μ such that $\lambda = e^{\mu}$. Then

$$X = (Ee^{\mu})\, e^{Y_1} = (e^{\mu E}) \cdot e^{Y_1} = e^{\mu E + Y_1}.$$

Therefore $Y = \mu E + Y_1$ answers the question.

(6.2) <u>If the series</u> (5.3) <u>for</u> f(X) <u>converges then the determinant</u>

(6.2.1) $$|f(X)| = \prod f(\lambda_j).$$

<u>Hence</u>

(6.2.2) $$|e^X| = e^{\sum \lambda_j} = e^{\operatorname{tr} X}.$$

<u>Thus</u> $Y = e^X$ <u>is never singular, and so it has an inverse</u> Y^{-1}. <u>Now</u> e^{-X} <u>exists likewise and</u> $e^X \cdot e^{-X} = e^{-X}e^X = E$, <u>so</u> $Y^{-1} = e^{-X}$.

Referring to (3.6) X is similar to a matrix

$$Y = \left\| \begin{matrix} \lambda_1 & & & & \\ \alpha_{12} & \lambda_2 & & & \\ & \ddots & \ddots & & \\ & & & \ddots & \\ & & & \alpha_{n-1,n}, & \lambda_n \end{matrix} \right\|$$

with terms above the diagonal all zero. The λ_j are the characteristic roots each repeated as often as its multiplicity. This follows immediately from $|Y - \lambda E| = \prod (\lambda_j - \lambda)$.

By (4.4) $f(X) \sim f(Y)$ and so we may replace everywhere X, by Y, i.e. we merely need to prove (6.2) for Y. Now f(Y) is of the same form as Y with λ_j replaced by $f(\lambda_j)$. This implies that the $f(\lambda_j)$ are the characteristic roots of f(Y). Since the determinant is the product of the characteristic roots (6.2.1) holds for Y, hence also for X, and the rest follows.

7. (7.1) <u>Matrix functions of scalars</u>. Let

$$X = \| x_{ij}(t) \|$$

be an $m \times n$ matrix whose terms are [real or complex] scalar functions of a [real or complex] variable t differentiable over a certain range R. Under the rules of operation on matrices we have, if Δ denotes increments:

$$\frac{\Delta X}{\Delta t} = \left| \frac{\Delta x_{ij}}{\Delta t} \right| .$$

Hence if $\lim \frac{\Delta X}{\Delta t}$ exists it is defined as the derivative of X, written $\frac{dX}{dt}$, and exists over R, its expression being

$$\frac{dX}{dt} = \left\| \frac{dx_{ij}}{dt} \right\| .$$

The rules for the derivatives of scalars, for addition and for multiplication by scalars follow as usual. Similar limit arguments yield the definition of the Riemann integral:

$$Y(t) = \int_{t_0}^{t} X \, dt = \left\| \int_{t_0}^{t} x_{ij} \, dt \right\| .$$

If the x_{ij} are continuous on the path of integration then evidently

$$\frac{dY}{dt} = X.$$

(7.2) Suppose now that $X(t)$, $Y(t)$ are any two square matrices of the same order, differentiable over the same range R. Then XY is differentiable over the same range and an elementary argument yields

$$\frac{d(XY)}{dt} = X\frac{dY}{dt} + \frac{dX}{dt}Y. \tag{7.3}$$

Care must be taken here to keep X, Y always in the same order. From (7.3) we deduce readily

$$\frac{d(X_1 \ldots X_r)}{dt} = \sum X_1 \ldots X_{q-1}\frac{dX_q}{dt}X_{q+1} \ldots X_r \tag{7.3.1}$$

and therefore

$$\frac{dX^r}{dt} = \sum X^{q-1}\frac{dX}{dt}X^{r-q}. \tag{7.3.2}$$

If we differentiate both sides of $XX^{-1} = E$, $X \neq 0$, we obtain

$$\frac{d(X^{-1})}{dt} = -X^{-1}\frac{dX}{dt}X^{-1}.$$

(7.4) Observe explicitly that the application to $1 \times m$ matrices yields the derivatives and integrals of vector functions of scalars.

(7.5) If all the $x_{ij}(t)$ are continuous or analytic at a point or a given set, we will say for convenience that $X(t)$ is <u>continuous</u> or <u>analytic</u> at the same point or on the same set.

(7.6) Let A be a constant square matrix and set

(7.6.1) $X(t) = e^{tA} = e^{At} = E + \frac{t}{1!} A + \frac{t^2}{2!} A^2 + \dots .$

By differentiation we obtain

(7.6.2) $$\frac{dX}{dt} = Ae^{At}.$$

Notice that owing to the form of (7.3.2) we can only prove $\frac{d}{dt} e^{X(t)} = e^{X(t)} \frac{dX}{dt}$, when X and $\frac{dX}{dt}$ commute.

Combining (7.6.2) with (7.3), and setting for convenience $\frac{d}{dt} = D$, we obtain for any matrix X the analogue of the well known elementary relation:

(7.7) $$(D - A)X = e^{At} \cdot D \cdot e^{-At} X.$$

As an application consider the matrix differential equation

(7.8) $$\frac{dX}{dt} = AX, \text{ A constant.}$$

Owing to (7.7) it reduces to

$$e^{At} \cdot D \cdot e^{-At} X = 0.$$

Multiplying both sides by e^{-At} (see 6.2) we have

$$De^{-At} \cdot X = 0$$

and hence $e^{-At} \cdot X = C$, an arbitrary constant matrix. Hence the complete solution of (7.8) is $X = e^{At} \cdot C$. We will return to this later.

§2. VECTOR SPACES

8. We will assume familiarity with the first concepts: dimension, base, coordinates relative to a base. When the scalars are all real [complex] the space is said to be a <u>real</u> [<u>complex</u>] vector space. The only vector spaces which we shall encounter are finite dimensional. Let $\mathcal{V}$ be such a space. Its vectors will be denoted by an arrow over small latin characters with possible super-

scripts as: $\vec{a}, \vec{x}, \vec{a}^1, \ldots$. Let in particular $\{\vec{e}^1, \ldots, \vec{e}^n\}$ be a base for $\mathcal{V}$. If we denote by $\vec{x}$ any element of $\mathcal{V}$ then we will have

$$(8.1) \qquad \vec{x} = \overrightarrow{x_1 e}^1 + \ldots + \overrightarrow{x_n e}^n.$$

The x_i are the coordinates or components of $\vec{x}$. If we adopt $\vec{x}$, x_i for the arbitrary vectors and their coordinates we will often denote $\mathcal{V}$ by $\mathcal{V}_x$. Similarly if say vectors and coordinates were $\vec{u}$, u_i we would write $\mathcal{V}_u$ for the space. In $\mathcal{V}_x$ the coordinates of $\vec{x}^h$ will be usually written x_{ih} (exceptionally and then explicitly stated x_{hi}).

The metrization of $\mathcal{V}_x$ will be done in the customary manner by means of a norm $\|\vec{x}\|$. We choose here for convenience

$$(8.2) \qquad \|\vec{x}\| = \sum |x_i|$$

and accordingly define the distance in $\mathcal{V}_x$ as

$$(8.3) \qquad d(x,x') = \|(\vec{x}-\vec{x}')\| = \sum |x_i - x_i'| .$$

As is well known this distance has the usual properties:

$$d(\vec{x},\vec{x}') = 0 \longrightarrow \vec{x} = \vec{x}';$$
$$d(\vec{x},\vec{x}') = d(\vec{x}',\vec{x}) \geq 0;$$
$$d(\vec{x},\vec{x}'') \leq d(\vec{x},\vec{x}') + d(\vec{x}',\vec{x}'').$$

With this specification of distance $\mathcal{V}_x$ is turned into a complete metric space which is topologically Euclidean space. We may show in fact that the above distance-function induces the same topology as the Euclidean distance $[\sum (x_i-x_i')^2]^{1/2}$. The completeness property of $\mathcal{V}_x$ implies that every Cauchy sequence has a limit.

(8.4) Let $\{\vec{e}^1\}$ be a base for $\mathcal{V}_x$. A square matrix A of order n defines a linear transformation of $\mathcal{V}_x$ into itself whereby $\vec{x}$ goes into $\vec{x}'$ designated by $A\vec{x}$, and whose coordinates are given by

$$x_i' = \sum a_{ij}x_j. \tag{8.4.1}$$

If we identify $\vec{x}$ with the one-column matrix of its coordinates then $A\vec{x}$ is merely matrix multiplication.

Consider now a new n-space $\mathcal{V}_y$ referred to a base $\{\vec{f}^i\}$ and let P be a linear transformation $\mathcal{V}_y \to \mathcal{V}_x$ whereby $\vec{y}$ goes into $\vec{x}$ whose coordinates are given by

$$x_i = \sum p_{ij}y_j$$

or in matrix notation $\vec{x} = P\vec{y}$. If $\vec{x}'$ goes into $\vec{y}$ then $\vec{x}' = P\vec{y}'$ and so, $\vec{x}$, $\vec{x}'$ being related as before we have $P\vec{y}' = AP\vec{y}$. Assuming now P non-singular we will have $\vec{y}' = P^{-1}AP\vec{y}$. In other words the transformation of $\mathcal{V}_x$ into itself by A corresponds to a transformation of $\mathcal{V}_y$ into itself by a matrix $\sim A$. Clearly also every matrix $\sim A$ is related to A in this manner. We may therefore interpret the properties of A invariant under similitude, as those of the transformations which are invariant under a non-singular linear transformation from space to space.

(8.5) Let $A \sim \operatorname{diag}(A_1,\dots,A_r)$ where the A_i are like (3.6.1), and let $\{\vec{f}^i\}$ be the base such that on passing to it A goes into the canonical form. Denote by ρ_i the order of A_i and set $\sigma_i = \rho_1 + \dots + \rho_i$. Let also $A_i^* = \operatorname{diag}(0, \dots, 0, A_i, 0, \dots, 0)$. The units

$\vec{f}^{\sigma_{i-1}+1}, \dots, \vec{f}^{\sigma_i}$ may be characterized as follows. First they are all annulled by the A_j^*, $j \neq i$. Then

$$\begin{aligned} A_i^* \vec{f}^{\sigma_{i-1}+1} &= \lambda_i \vec{f}^{\sigma_{i-1}+1} \\ A_i^* \vec{f}^{\sigma_{i-1}+h} &= \vec{f}^{\sigma_{i-1}+h-1} + \lambda_i \vec{f}^{\sigma_{i-1}+h}, \quad h > 1, \end{aligned} \tag{8.5.1}$$

which provides a complete description.

Suppose in particular A real and among the λ_j let there be found r pairs which are complex conjugates which for convenience we may assume to be $(\lambda_1, \bar{\lambda}_1)$, ..., $(\lambda_r, \bar{\lambda}_r)$ and the rest $\lambda_{2r+1}, \ldots, \lambda_n$ real. If we obtain the vectors $\vec{f}$ for $\lambda_1, \ldots, \lambda_r$ then their conjugates will do for $\bar{\lambda}_1, \ldots, \bar{\lambda}_r$. The coordinates of a real vector referred to this base will then be $x_1, \ldots, x_r, \bar{x}_1, \ldots, \bar{x}_y, x_{2r+1}, \ldots, x_n$ where the x_{2r+i} are real. We will then say that the coordinate system is <u>real</u>.

(8.6) If $\vec{x}(t)$ depends upon t then the derivatives and integrals of $\vec{x}(t)$ may be defined in the usual manner and are written

$$\frac{d\bar{x}}{dt}, \quad \int_{t^0}^{t} \vec{x}(t)dt.$$

Both are vectors, their components being respectively

$$\frac{dx_i}{dt}, \quad \int_{t_0}^{t} x_i(t)dt.$$

Clearly

$$\frac{d}{dt} \int_{t_0}^{t} \vec{x}(t)dt = \vec{x}(t).$$

We also note the following useful inequality. Suppose t real and $\| x(t) \| < M$ for $t_0 \leq t \leq t_1$. Then

$$\| \int_{t_0}^{t_1} x(t)dt \| < n\, M\, |t_1 - t_0| . \tag{8.6.1}$$

(8.7) Let $\vec{y} = (y_1, \ldots, y_r)$ be a real or complex vector in some $\mathcal{V}_y$. Suppose the x_i analytic in the y_j at $\vec{y}^0 = (y_1^0, \ldots, y_r^0)$, i. e. representable by power series in the y_j valid in a neighborhood of $\vec{y}^0$ in $\mathcal{V}_y$. We will then say: $\vec{x}$ is analytic in $\vec{y}$ at $\vec{y}^0$. This is the prototype of a readily understood terminology used extensively later.

(8.8) In dealing with n dimensional spaces $\mathcal{V}_x$ it will be convenient to define as a sphere of center $\vec{x}^0$ and radius ρ, written $\mathcal{S}(x^0, \rho)$ the set $\| \vec{x} - x^0 \| < \rho$.

(8.9) Frequently besides the vector variables $\vec{x}$, $\vec{y}$, ..., there will occur a real parameter t referred to as the time, and whose range is the real line L. Instead of $\mathcal{V}_x$ for instance we shall have a product space $\mathcal{W} = L \times \mathcal{V}_x$ and new spheres $\sum(\vec{x}^0, t^0, \rho)$ defined by $\|\vec{x} - \vec{x}^0\| + |t - t^0| < \rho$.

The relation of the spheres to open sets, limits, etc., are as in real variables and need not be discussed here.

(8.10) Whenever $\mathcal{V}_x$ is two-dimensional it will be convenient to revert to the more usual "spheres", namely the circular regions of Euclidean geometry. As is well known this does not affect the standard concepts of open sets, ...

§3. ANALYTIC FUNCTIONS OF SEVERAL VARIABLES

9. We shall have repeated occasion to consider analytic functions of several real or complex variables as well as mixed functions analytical in some, but not in all, the variables.

(9.1) Consider first $\mathcal{V}_x$ complex. Write $\vec{x} = \vec{y} + i\vec{z}$, viz. $x_j = y_j + iz_j$ for $j = 1, \ldots, n$. A function $f(\vec{x})$ is said to be analytic in a region Ω of $\mathcal{V}_x$ if it has first partial derivatives relative to all y_j and z_j which are continuous in $\vec{y}$ and $\vec{z}$ at all points of Ω and if it satisfies the Cauchy-Riemann differential equations relative to each pair of y_j and z_j at all points of Ω. The function f is said to be <u>holomorphic</u> in Ω if it is analytic and single-valued in Ω; f is said to be analytic or holomorphic in a closed set F in $\mathcal{V}_x$ if it is analytic or holomorphic is some neighborhood of F, (some region $\supset$ F).

A n.a.s.c. for analyticity in Ω is that f may be expanded in Taylor series around each point $\vec{\xi}$ of Ω. The series will be convergent in a set

$$\mathcal{J}(\vec{\xi}, \alpha) : |x_j - \xi_j| < \alpha, \qquad j = 1, 2, \ldots, n,$$

which we call conveniently a toroid of center $\vec{\xi}$ and radius α. Moreover α may be chosen such that the toroid is in Ω.

(9.2) Suppose now $\mathcal{V}_x$ real and let $[\mathcal{V}_x]$ be its complex extension, i. e. the complex vector space obtained by allowing the coordinates x_j to take complex values. A real function $f(\vec{x})$ will be said to be analytic or holomorphic in a region Ω of $\mathcal{V}_x$ under the following conditions: there exists a region $[\Omega]$ of $[\mathcal{V}_x]$ and a function $[f]$ on $[\mathcal{V}_x]$ analytic or holomorphic in $[\Omega]$, and such that: (a) $[\Omega] \supset \Omega$; (b) the values of $[f]$ on Ω are those of f.

(9.3) The definitions just given for the real case may seem indirect. They have however the advantage to guarantee that the following important property holds:

(9.4) Theorem. *If a series of real or complex functions analytic or holomorphic in a region Ω, is uniformly convergent in Ω, then the limit is analytic or holomorphic in Ω.*

For the complex case this is a standard theorem due to Weierstrass (see Osgood II p. 15) and for the real case it is a consequence of our definition.

(9.5) An analytical vector is a vector $\vec{f}(\vec{x})$ whose components $f_i(x_1, \ldots, x_n)$ are analytical.

(9.6) Given two series

$$a = a_1 + a_2 + \ldots, \qquad b = b_1 + b_2 + \ldots,$$

the second is said to be a majorante of the first, written $a \ll b$ (Poincaré's notation) whenever $|a_m| \leq |b_m|$ for every m. More generally, if the multiple series

$$a = \sum a_{m_1,\ldots,m_p}, \qquad b = \sum b_{m_1,\ldots,m_p}$$

are such that $|a_{m_1,\ldots,m_p}| \leq |b_{m_1,\ldots,m_p}|$ for every combination $m_1,\ldots,m_p$ then b is called a majorante of a, written as before $a \ll b$.

(9.6.1) If $m = \sum m_i$ then it is often convenient to denote by (m) the set $\{m_1, \ldots, m_p\}$. Thus we would write above:

$$(9.6.2) \qquad a = \sum a_{(m)}, \qquad b = \sum b_{(m)}.$$

(9.7) Suppose that $F(x_1, \ldots, x_n)$ is holomorphic in the closed region Ω : $|x_i| \leq A_i$, $i = 1, 2, \ldots, n$, where the A_i are positive constants. Since Ω is compact $|F|$ has an upper bound M in Ω. It is then shown in the treatises on the subject (see for instance Picard, Traité d'Analyse, vol. III, Ch. 9) that F admits in Ω the McLaurin expansion

$$F = \sum F^{(m)} x_1^{m_1} \ldots x_n^{m_n}$$

with the following estimate for the coefficients:

$$(9.8) \qquad F^{(m)} < \frac{M}{A_1^{m_1} \ldots A_n^{m_n}}.$$

If we identify F with the series we have therefore

$$(9.9) \qquad F \ll \frac{M}{\prod(1 - \frac{x_i}{A_i})}.$$

If $A = \inf A_i$ then another useful relation of the same type is

$$(9.10) \qquad F \ll \frac{M}{1 - \frac{1}{A}\sum x_i}.$$

It is in fact a simple matter to show that

$$\frac{1}{\prod(1 - \frac{x_i}{A_i})} \ll \frac{1}{1 - \frac{1}{A}\sum x_i}$$

from which (9.10) follows