ON THE THEORY OF INFINITE SYSTEMS OF DIFFERENTIAL EQUATIONS AND THEIR APPLICATION TO THE THEORY OF STOCHASTIC PROCESSES AND THE PERTURBATION THEORY OF QUANTUM MECHANICS.

By

NIELS ARLEY and VIBEKE BORCHSENIUS.

Institute of Theoretical Physics of the University of COPENHAGEN.

Contents.

Introduction.

In a previous paper[1] one of us has presented the theory of finite systems of simultaneous linear differential equations in such a form that it is formally independent of the dimension, i. e. the number of equations in the system. Formally the theory may, therefore, be immediately generalized to the case in which the dimension is enumerable infinite[2]. Such infinite systems of differential

[1] Arley (1943) §§ 2.2—2.6 and chap. 7.

[2] The theory may even be generalized to the case in which the dimension is non-enumerable. We intend to give such a generalization in a later paper.

equations play an important rôle in many applications, e. g. in the theory of discontinuous stochastic processes (discussed in part II) and in quantum mechanical perturbation theory (discussed in part IV).

In order to make the theory work it is, however, necessary to impose certain convergence conditions[1], which are at any rate automatically fulfilled for every finite dimension. We shall in the present paper investigate these conditions more closely. It is, namely, obvious at beforehand that as soon as we have limit processes at our disposal, any 'pathological' case desired may be constructed by properly utilizing limit processes. Such 'pathologies' are not only of great interest in themselves, but are even met with in the practical applications of the theory mentioned above. We shall, therefore, not only discuss the theory itself (parts I and III), but also the two applications mentioned (parts II and IV).

PART I.

General Theory.

§ 1.

We shall first give a survey of the usual theory of *finite* systems in the form given in the paper quoted.

The most general form of a finite system of simultaneous linear differential equations can be written

$$\begin{gathered} F_{00}(D,x)\,Y_0(x) + \cdots + F_{0,m-1}(D,x)\,Y_{m-1}(x) = B_0(x) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ F_{m-1,0}(D,x)\,Y_0(x) + \cdots + F_{m-1,m-1}(D,x)\,Y_{m-1}(x) = B_{m-1}(x). \end{gathered} \tag{1.1}$$

The letters have here the following meaning:

x: The independent variable, which may be real or complex[2].

$D = \dfrac{d}{dx}$.

$Y_i(x)$, $i = 0, 1, 2, \ldots, m-1$: the unknown functions, which may also be real or complex.

---

[1] Arley (1943) §§ 2.2—2.6 and chap. 7.

[2] We shall in the following tacitly assume x to be real, but the proofs are also valid in case x varies on a regular curve L without double points in the complex plane, if, only, we interpret $|x-x_0|$ as the length of the curve L from x_0 to x.

$B_i(x)$, $i = 0, 1, 2, \ldots, m-1$: given functions, assumed to be *continuous* in a certain region (apart from possible isolated singularities).

$F_{ij}(D,x)$, $i,j = 0, 1, 2, \ldots, m-1$: given polynomials in D with coefficients which are functions of x, assumed to be *continuous* in a certain region (apart from possible isolated singularities):

$$F_{ij}(D,x) = \underset{ij}{F_N}(x)\, D^N + \underset{ij}{F_{N-1}}(x)\, D^{N-1} + \cdots + \underset{ij}{F_1}(x)\, D + \underset{ij}{F_0}(x), \qquad (1.2)$$

$$i,j = 0, 1, 2, \ldots, m-1.$$

m: the dimension of the system.
N: » order » » » .

Using the matrix symbolism our equations can obviously be written in the compact form

$$\boldsymbol{F}(D,x) \cdot \boldsymbol{Y}(x) = \boldsymbol{B}(x) \qquad (1.3)$$

with

$$\boldsymbol{F}(D,x) = \boldsymbol{F}_N(x)\, D^N + \cdots + \boldsymbol{F}_1(x) D + \boldsymbol{F}_0(x). \qquad (1.4)$$

If the matrix $\boldsymbol{F}_N(x)$ has a reciprocal, $\boldsymbol{F}_N^{-1}(x)$, for all x in the region of definition of the system of equations (1.3), this system may, as is well-known, in several ways be transformed into an equivalent system of the first order and dimension $n = m \cdot N$.[1] Let one such system be denoted by

$$D\boldsymbol{Y}(x) = \boldsymbol{Y}'(x) = \boldsymbol{A}(x) \cdot \boldsymbol{Y}(x) + \boldsymbol{B}(x) \qquad \text{(dimension: } n). \qquad (1.5)$$

If $\boldsymbol{F}_N^{-1}(x)$ does not exist for some values of x, (1.5) is only defined for all other values of x. If $\boldsymbol{F}_N^{-1}(x)$ does not exist for any value of x, this fact means that our equations are restricted by a certain number of linear relations, and we may, therefore, in such case write down a system of differential equations containing a smaller number of functions and then transform this system to the form (1.5)[2].

As the result we thus see that we need only consider systems of the form (1.5), socalled *simple* systems. On the other hand we note that a system of the form (1.5) may also in several ways be transformed into a system of the form (1.3), which fact may sometimes be successfully utilized for the actual solution of the equations.

[1] See e. g. Frazer, Duncan and Collar (1938).
[2] Frazer, loc. cit. p. 163.

§ 2.

Let our system be of the form (1. 5), i. e.

$$\boldsymbol{Y}'(x) = \boldsymbol{A}(x)\cdot\boldsymbol{Y}(x) + \boldsymbol{B}(x) \qquad \text{(dimension: } n). \tag{2.1}$$

Here

$$\boldsymbol{Y}(x) = \begin{Bmatrix} Y_0(x) \\ \vdots \\ Y_{n-1}(x) \end{Bmatrix}, \quad \boldsymbol{A}(x) = \begin{Bmatrix} A_{00}(x) & \cdots & A_{0,n-1}(x) \\ \vdots & & \vdots \\ A_{n-1,0}(x) & \cdots & A_{n-1,n-1}(x) \end{Bmatrix} \text{ and } \boldsymbol{B}(x) = \begin{Bmatrix} B_0(x) \\ \vdots \\ B_{n-1}(x) \end{Bmatrix}. \tag{2.2}$$

The two given matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are, as already mentioned, assumed to be continuous in a certain region. We first observe that $\boldsymbol{Y}$, being differentiable, is also continuous, and $\boldsymbol{A}$ being continuous, $\boldsymbol{Y}'$ is, consequently, also continuous, due to the dimension of the system being *finite*. $\boldsymbol{Y}'$ may, therefore, be integrated.

First we prove that if (2. 1) *has* any solution, it can only have *one*, corresponding to a given initial condition

$$\boldsymbol{Y}(x_0) = \boldsymbol{C} \tag{2.3}$$

in which $\boldsymbol{C}$ is an arbitrary constant.

Let $\boldsymbol{Y}_1(x)$ and $\boldsymbol{Y}_2(x)$ be two solutions satisfying the same initial condition (2. 3). Then

$$\boldsymbol{Y}(x) = \boldsymbol{Y}_1(x) - \boldsymbol{Y}_2(x) \tag{2.4}$$

will be a solution of the corresponding *homogeneous* equation

$$\boldsymbol{Y}' = \boldsymbol{A}\cdot\boldsymbol{Y} \tag{2.5}$$

satisfying the initial condition

$$\boldsymbol{Y}(x_0) = \boldsymbol{0}. \tag{2.6}$$

Thus we shall prove that $\boldsymbol{Y}(x) \equiv \boldsymbol{0}$. As both $\boldsymbol{A}(x)$ and $\boldsymbol{Y}(x)$ are continuous, the following two matrices exist[1]

$$\boldsymbol{K} = \max_{x_0 \leqq t \leqq x} |\boldsymbol{A}(t)| \tag{2.7}$$

$$\boldsymbol{G} = \max_{x_0 \leqq t \leqq x} |\boldsymbol{Y}(t)|. \tag{2.8}$$

From (2. 5) we now have that

$$|\boldsymbol{Y}'| \leqq \boldsymbol{K}\cdot|\boldsymbol{Y}| \leqq \boldsymbol{K}\cdot\boldsymbol{G}. \tag{2.9}$$

[1] By $|\boldsymbol{A}| = \{|A_{ik}|\}$ we understand the matrix whose elements are the numerical values of the corresponding elements of $\boldsymbol{A}$. By $\max \boldsymbol{A}$ we understand the matrix whose elements are the maximum values of the corresponding elements of $\boldsymbol{A}$ etc. We note, furthermore, that by $x_0 \leqq t \leqq x$ we shall always denote the interval, also in the case $x < x_0$.

$|Y'|$ being also continuous, (2.9) may be integrated; introducing the result into (2.9) and integrating again it follows by repeating this process that

$$|Y| \leqq K^\nu \frac{|x-x_0|^\nu}{\nu!} \cdot G \quad \text{for all} \quad \nu = 1, 2, 3, \ldots. \tag{2.10}$$

We shall now show that the matrix function $\exp[K(x-x_0)] = \sum_{\nu=0}^{\infty} K^\nu \frac{|x-x_0|^\nu}{\nu!}$ exists, which fact we simply express by saying that A is *absolutely exponentiable* in the interval (x_0, x) and writing

$$\exp[K|x-x_0|] = \sum_{\nu=0}^{\infty} K^\nu \frac{|x-x_0|^\nu}{\nu!} < \infty. \tag{2.11}$$

K given in (2.7) being of finite dimension, there is, namely, among its n^2 elements a greatest one, k. We thus have

$$K \leqq kE \quad \text{with} \quad E_{ij} = 1 \quad \text{for all} \quad i, j = 0, 1, 2, \ldots, n-1, \tag{2.12}$$

i. e.

$$K^\nu \leqq (nk)^\nu E \quad \begin{matrix} \text{for all} & n = 1, 2, 3, \ldots, \\ \text{»} \quad \text{»} & \nu = 0, 1, 2, \ldots. \end{matrix} \tag{2.13}$$

Consequently

$$\exp[K|x-x_0|] = \sum_{\nu=0}^{\infty} K^\nu \frac{|x-x_0|^\nu}{\nu!} \leqq E \sum_{\nu=0}^{\infty} (nk)^\nu \frac{|x-x_0|^\nu}{\nu!} =$$

$$= E \exp[nk|x-x_0|] < \infty. \tag{2.14}$$

Due to the *finite* dimension we next have for an arbitrary non-negative column matrix G that, due to (2.11),

$$(\exp[K|x-x_0|] \cdot G)_i = \sum_{\nu=0}^{\infty} \left(K^\nu \frac{|x-x_0|^\nu}{\nu!} \cdot G \right)_i < \infty \tag{2.15}$$

for all $i = 0, 1, 2, \ldots, n-1$.

(2.15) shows that the right hand side of (2.10) tends to zero

$$K^\nu \frac{|x-x_0|^\nu}{\nu!} \cdot G \underset{\nu \to \infty}{\to} 0. \tag{2.16}$$

(2.10) can, consequently, only be satisfied for all values of ν if

$$Y(x) \equiv 0, \tag{2.17}$$

q. e. d.

§ 3.

Next we prove that the system (2.1) actually has a solution satisfying the arbitrary initial condition (2.3). We first consider the homogeneous equation (2.5). The dimension being *finite*, $\boldsymbol{Y}'$ is continuous and (2.5) is thus, due to (2.3), equivalent with the equation

$$\boldsymbol{Y}(x) = \boldsymbol{C} + \int_{x_0}^{x} \boldsymbol{A}(t) \cdot \boldsymbol{Y}(t)\, dt. \tag{3.1}$$

We now use the method of iteration and put

$$\boldsymbol{Y}(x) = \sum_{\nu=0}^{\infty} \boldsymbol{Y}_\nu(x) \tag{3.2}$$

in which

$$\boldsymbol{Y}_0(x) = \boldsymbol{C}, \qquad \boldsymbol{Y}_\nu(x) = \int_{x_0}^{x} \boldsymbol{A}(t) \cdot \boldsymbol{Y}_{\nu-1}(t)\, dt, \qquad \nu = 1, 2, 3, \ldots. \tag{3.3}$$

The series (3.2) is called the *Peano series.*

It is then easily seen

(a) that (3.2) is absolutely and uniformly convergent, and

(b) that (3.2) is a solution of (3.1).

In fact we have, due to (2.7), (2.11) and the fact that $|\boldsymbol{C}|$ satisfies (2.15),

$$|\boldsymbol{Y}(t)| \leqq \sum_{\nu=0}^{\infty} |\boldsymbol{Y}_\nu(t)| \leqq \sum_{\nu=0}^{\infty} \boldsymbol{K}^\nu \frac{|x - x_0|^\nu}{\nu!} \cdot |\boldsymbol{C}| = \exp\,[\boldsymbol{K}|x - x_0|] \cdot |\boldsymbol{C}| < \infty \tag{3.4}$$

which proves (a). Inserting (3.3) into (3.2) we next find

$$\begin{aligned} \boldsymbol{Y}(x) = \boldsymbol{C} + \sum_{\nu=1}^{\infty} \int_{x_0}^{x} \boldsymbol{A}(t) \cdot \boldsymbol{Y}_{\nu-1}(t)\, dt \underset{(\alpha)}{=} \boldsymbol{C} + \int_{x_0}^{x} \left(\sum_{\nu=1}^{\infty} \boldsymbol{A} \cdot \boldsymbol{Y}_{\nu-1} \right) dt \underset{(\beta)}{=} \\ \boldsymbol{C} + \int_{x_0}^{x} \boldsymbol{A} \cdot \left(\sum_{\nu=0}^{\infty} \boldsymbol{Y}_\nu \right) dt = \boldsymbol{C} + \int_{x_0}^{x} \boldsymbol{A} \cdot \boldsymbol{Y}\, dt. \end{aligned} \tag{3.5}$$

The operation (α) is legitimate due to the series (3.2) being *uniformly* convergent; (β) due to the double-sum $\sum_\nu \boldsymbol{A} \cdot \boldsymbol{Y}_\nu$ being *absolutely* convergent, because using (2.7) and (3.4)

$$|\boldsymbol{A}(t) \cdot \boldsymbol{Y}(t)| \leqq \boldsymbol{K} \cdot \exp\,[\boldsymbol{K}|x - x_0|] < \infty. \tag{3.6}$$

The exponential is, namely, a *power* series in $|x - x_0|$ and its ν'th differential coefficient is, therefore, obtained by term-by-term differentiation for all values of

x within the convergence region of the exponential. Due to (3.5) $\boldsymbol{Y}(x)$ is the integral of a continuous function and, therefore, differentiable with the continuous derivative $\boldsymbol{Y}' = \boldsymbol{A} \cdot \boldsymbol{Y}$, which proves (b).

Putting $\boldsymbol{C} = \mathbf{1}$ in (3.1)—(3.5) we obtain a *quadratic* matrix each of whose columns is a solution of (3.1), i. e. (2.5). This matrix we denote by

$$\boldsymbol{F}(x, x_0) = \mathop{\mathcal{P}}_{x_0}^{x}(\mathbf{1} + \boldsymbol{A}(t)\,dt) = \sum_{\nu=0}^{\infty} \boldsymbol{F}_\nu(x, x_0), \tag{3.7}$$

$$\boldsymbol{F}_0 \equiv \mathbf{1}, \qquad \boldsymbol{F}_\nu(x, x_0) = \int_{x_0}^{x} \boldsymbol{A}(t) \cdot \boldsymbol{F}_{\nu-1}(t, x_0)\,dt =$$

$$= \mathop{\int \cdots \int}_{x \geqq t_\nu \geqq \cdots \geqq t_1 \geqq x_0} \boldsymbol{A}(t_\nu) \cdots \boldsymbol{A}(t_1)\,dt_\nu \cdots dt_1 = \int_{x_0}^{x} \boldsymbol{F}_{\nu-1}(x, t) \cdot \boldsymbol{A}(t)\,dt,$$

$$\nu = 1, 2, 3, \ldots .$$

$\boldsymbol{F}$ is called the *product-integral* (or the *matrizant*). The first name and the symbol $\mathcal{P}(\mathbf{1} + \boldsymbol{A}\,dt)$ refers to the fact that $\boldsymbol{F}$ may also be defined as

$$\boldsymbol{F}(x, x_0) = \lim_{m \to \infty} \prod_{i=0}^{m-1} (\mathbf{1} + \boldsymbol{A}(x_i)\,\Delta_i) \tag{3.8}$$

in which $x_0 < x_1 < \cdots < x_m = x$, $\Delta_i = x_{i+1} - x_i$, is an arbitrary division of the interval (x_0, x).[1]

$\boldsymbol{F}$ satisfies[2]

$$\frac{\partial}{\partial x} \boldsymbol{F}(x, x_0) = \boldsymbol{A}(x) \cdot \boldsymbol{F}(x, x_0) \tag{3.9}$$

$$\frac{\partial}{\partial x_0} \boldsymbol{F}(x, x_0) = - \boldsymbol{F}(x, x_0) \cdot \boldsymbol{A}(x_0) \tag{3.10}$$

and

$$\lim_{x_0 \to x} \boldsymbol{F}(x, x_0) = \lim_{x \to x_0} \boldsymbol{F}(x, x_0) = \boldsymbol{F}(x_0, x_0) = \mathbf{1}. \tag{3.11}$$

Furthermore, $\boldsymbol{F}$ is a *fundamental solution*, i. e. $\boldsymbol{F}$ has the property that *any* solution $\boldsymbol{Y}(x)$ is the product of $\boldsymbol{F}$ and a constant $\boldsymbol{C} = \boldsymbol{Y}(x_0)$. This fact follows from the theorem of uniqueness (§ 2), because (a): $\boldsymbol{F} \cdot \boldsymbol{C}$ will for arbitrary $\boldsymbol{C}$ be a solution of (3.1), i. e. (2.5), and (b): $\boldsymbol{Y}(x)$ and $\boldsymbol{F}(x, x_0) \cdot \boldsymbol{Y}(x_0)$ are both solutions which are equal to $\boldsymbol{C}$ for $x = x_0$. They must, therefore, coincide for all values of x, i. e.

$$\boldsymbol{Y}(x) \equiv \boldsymbol{F}(x, x_0) \cdot \boldsymbol{Y}(x_0). \tag{3.12}$$

[1] Cf. Arley (1943) § 2.5. We note that if $\boldsymbol{ln\,A}$ exists and satisfies $\boldsymbol{exp}\,[\boldsymbol{ln\,A}] = \boldsymbol{A}$, then the product-integral is obviously related to the notion *'productal'* introduced by Reichenbach (1935) as follows: $\mathcal{P}(\mathbf{1} + \boldsymbol{ln\,A}(t)\,dt) = \mathop{\mathcal{P}}_{x_0}^{x}(\boldsymbol{A}(t))^{dt} = \lim\limits_{m \to \infty} \prod\limits_{i=0}^{m-1} (\boldsymbol{A}(x_i))^{\Delta_i}$.

[2] We note that (3.10), which is said to be *adjointed* to (3.9), follows from (3.7) in exactly the same way as (3.9) was proved in (3.5).

For arbitrary x_0, x and x_1 in the region of definition we now have

$$\boldsymbol{Y}(x) = \boldsymbol{F}(x, x_1) \cdot \boldsymbol{Y}(x_1) = \boldsymbol{F}(x, x_1) \cdot \boldsymbol{F}(x_1, x_0) \cdot \boldsymbol{Y}(x_0) = \boldsymbol{F}(x, x_0) \cdot \boldsymbol{Y}(x_0). \quad (3.13)$$

This equation being valid for all values of $\boldsymbol{Y}(x_0)$, it follows from the theorem of uniqueness that

$$\boldsymbol{F}(x, x_0) = \boldsymbol{F}(x, x_1) \cdot \boldsymbol{F}(x_1, x_0). \quad (3.14)$$

We note that this relation also follows from (3.8). Especially we obtain for $x = x_0$ and $x_1 = x$, due to (3.11),

$$\boldsymbol{F}(x_0, x) \cdot \boldsymbol{F}(x, x_0) = \boldsymbol{1} \quad (3.15)$$

which means that $\mathcal{P} = \boldsymbol{F}(x, x_0)$ has an inverse

$$\left(\underset{x_0}{\overset{x}{\mathcal{P}}}\right)^{-1} = \boldsymbol{F}^{-1}(x, x_0) = \boldsymbol{F}(x_0, x) = \underset{x}{\overset{x_0}{\mathcal{P}}}. \quad (3.16)$$

We can now solve the inhomogeneous equation (2.1), i.e.

$$\boldsymbol{Y}' = \boldsymbol{A} \cdot \boldsymbol{Y} + \boldsymbol{B}. \quad (3.17)$$

As $\boldsymbol{B}(x)$ is assumed to be continuous, the matrix

$$\boldsymbol{M} = \max_{x_0 \leqq t \leqq \omega} |\boldsymbol{B}(t)| \quad (3.18)$$

exists (cf. (2.8)) and satisfies, due to the finite dimension, (2.15).

Multiplying (3.17) to the left with $\mathcal{P}^{-1}$ we obtain

$$\mathcal{P}^{-1} \cdot \boldsymbol{B} = \mathcal{P}^{-1} \cdot \boldsymbol{Y}' - \mathcal{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{Y} = \frac{d}{dx}(\mathcal{P}^{-1} \cdot \boldsymbol{Y}). \quad (3.19)$$

We have here used that

$$\boldsymbol{0} = \boldsymbol{1}' = (\mathcal{P} \cdot \mathcal{P}^{-1})' = \mathcal{P} \cdot (\mathcal{P}^{-1})' + \mathcal{P}' \cdot \mathcal{P}^{-1}, \quad \text{i.e.} \quad (\mathcal{P}^{-1})' = -\mathcal{P}^{-1} \cdot \mathcal{P}' \cdot \mathcal{P}^{-1}, \quad (3.20)$$

and next that, due to (3.9),

$$\mathcal{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{Y} = \mathcal{P}^{-1} \cdot \boldsymbol{A} \cdot \mathcal{P} \cdot \mathcal{P}^{-1} \cdot \boldsymbol{Y} = \mathcal{P}^{-1} \cdot \mathcal{P}' \cdot \mathcal{P}^{-1} \cdot \boldsymbol{Y} = -(\mathcal{P}^{-1})' \cdot \boldsymbol{Y}. \quad (3.21)$$

From (3.19) we obtain at once by integration a particular integral of the inhomogeneous equation and adding the total integral $\mathcal{P} \cdot \boldsymbol{C}$ of the homogeneous equation we finally have that the total integral of the inhomogeneous equation (3.17) is given by, using (3.14) and (3.16),

$$\boldsymbol{Y}(x) = \underset{x_0}{\overset{x}{\mathcal{P}}} \cdot \boldsymbol{C} + \underset{x_0}{\overset{x}{\mathcal{P}}} \cdot \int_{x_0}^{x} \left(\underset{x_0}{\overset{t}{\mathcal{P}}}\right)^{-1} \cdot \boldsymbol{B}(t)\, dt = \underset{x_0}{\overset{x}{\mathcal{P}}} \cdot \boldsymbol{C} + \int_{x_0}^{x} \underset{t}{\overset{x}{\mathcal{P}}} \cdot \boldsymbol{B}(t)\, dt. \quad (3.22)$$

We observe that this formula is a direct generalization of the well-known formula in case the system reduces to one equation with one unknown function (cf. (d) p. 269).

§ 4.

Apart from the properties given in (3.9)—(3.11) it is directly seen from (3.7) that the product-integral has, furthermore, the properties

$$\underset{x}{\overset{x+\Delta x}{\mathcal{P}}}(\mathbf{1}+\boldsymbol{A}\,dt)=\boldsymbol{F}(x+\Delta x, x)=\mathbf{1}+\boldsymbol{A}(x)\Delta x+\boldsymbol{o}(\Delta x) \tag{4.1}$$

and

$$\left|\underset{x_0}{\overset{x}{\mathcal{P}}}(\mathbf{1}+\boldsymbol{A}\,dt)\right| \leqq \exp\left[K|x-x_0|\right]. \tag{4.2}$$

If $\boldsymbol{A}(x)$ and $\boldsymbol{F}_1(x, x_0)=\int\limits_{x_0}^{x}\boldsymbol{A}(t)\,dt$ commute, i. e.

$$\boldsymbol{A}\cdot\boldsymbol{F}_1=\boldsymbol{F}_1\cdot\boldsymbol{A}, \tag{4.3}$$

we have

$$\boldsymbol{F}_2(x, x_0)=\int\limits_{x_0}^{x}\left(\frac{\partial}{\partial t}\boldsymbol{F}_1(t, x_0)\right)\cdot\boldsymbol{F}_1(t, x_0)\,dt=\frac{1}{2!}(\boldsymbol{F}_1(x, x_0))^2 \tag{4.4}$$

and thus generally

$$\boldsymbol{F}_\nu(x, x_0)=\frac{1}{\nu!}(\boldsymbol{F}_1(x, x_0))^\nu. \tag{4.5}$$

Under the condition (4.3) it then follows that

$$\underset{x_0}{\overset{x}{\mathcal{P}}}(\mathbf{1}+\boldsymbol{A}\,dt)=\exp\left[\int\limits_{x_0}^{x}\boldsymbol{A}(t)\,dt\right]. \tag{4.6}$$

We note that (4.3) is satisfied in three important cases:

(a) $\boldsymbol{A}(x)$ is constant, i. e. independent of x,

(b) $\boldsymbol{A}(x)$ is the product of a constant matrix and a scalar function of x,

(c) $\boldsymbol{A}(x)$ is a diagonal matrix,

(d) $\boldsymbol{A}(x)$ is a one-dimensional matrix, i. e. a scalar function of x.

If $\boldsymbol{A}(z)$ is analytic in some domain Ω in the complex z-plane, i. e. that this is the case for each element of $\boldsymbol{A}(z)$, then it follows that also

$$\boldsymbol{F}(z, z_0)=\underset{z_0}{\overset{z}{\mathcal{P}}}(\mathbf{1}+\boldsymbol{A}\,dz) \tag{4.7}$$

is analytic in every *inner* point of Ω, $\boldsymbol{F}$ being, furthermore, independent of the integration curve L between z_0 and z.

This is easily seen to be true. First we have from (3.7) that all the matrices $\boldsymbol{F}_\nu(x, x_0)$ are independent of L and analytic in Ω. Due to the fact that $\boldsymbol{F}=\mathcal{P}$ is given by a *uniformly* convergent series, the result next follows immediately from a well-known theorem from the theory of analytic functions. Using this theorem once more we thus see that due to $|C| = |Y(z_0)|$ satisfying (2.15), each solution of our homogeneous equation $Y' = A \cdot Y$ is analytic. Next, assuming also B in $Y' = A \cdot Y + B$ to be analytic it is finally seen from (3.14), (3.22) and the fact that M given in (3.18) satisfies (2.15), that also each solution of the inhomogeneous equation is analytic.

If A and B in $Y' = A \cdot Y + B$ are analytic in Ω, we thus see that the solutions may for an arbitrary inner point of Ω be expanded in power series which are convergent in every circle not containing any singular point of any of the elements of $A(z)$ or $B(z)$. For finite systems every regular point of both A and B is, consequently, also a regular point of the equation $Y' = A \cdot Y + B$. Furthermore, these power series may be obtained by the usual method of introducing the power series into the equation and equating corresponding coefficients on both sides.

Finally we shall note the following important transformation property of the product-integral. Let us transform the unknown functions $Y(x)$ to new functions $Z(x)$ by means of

$$Y(x) = T(x) \cdot Z(x) \tag{4.8}$$

in which $T(x)$ is an arbitrary, non-singular matrix-function, i. e. which has a reciprocal T^{-1} satisfying[1]

$$T^{-1}(x) \cdot T(x) = 1. \tag{4.9}$$

From $Y' = A \cdot Y + B$ we then obtain

$$Z' = (T^{-1} \cdot A \cdot T - T^{-1} \cdot T') \cdot Z + T^{-1} \cdot B = A^* \cdot Z + B^* \tag{4.10}$$

in which A^* and B^* denote the matrices

$$\begin{aligned} A^* &= T^{-1} \cdot A \cdot T - T^{-1} \cdot T' \\ B^* &= T^{-1} \cdot B. \end{aligned} \tag{4.11}$$

From (4.8) and (4.10) we thus have in the case of homogeneous equations, i. e. $B = B^* \equiv 0$,

$$\begin{aligned} Y(x) &= \mathscr{P}_{x_0}^{x}(1 + A\,dt) \cdot Y(x_0) = T(x) \cdot \mathscr{P}_{x_0}^{x}(1 + A^*\,dt) \cdot Z(x_0) = \\ &\quad T(x) \cdot \mathscr{P}_{x_0}^{x}(1 + A^*\,dt) \cdot T^{-1}(x_0) \cdot Y(x_0). \end{aligned} \tag{4.12}$$

[1] We note that for finite dimensions we then also have $T \cdot T^{-1} = 1$ and that T^{-1} is uniquely determined by T. For infinite dimensions T_{right}^{-1} need not exist even if T_{left}^{-1} exists. It will be seen, that we use only the left-hand reciprocal of T.

Consequently we have from the theorem of uniqueness

$$\underset{x_0}{\overset{x}{\mathcal{P}}}(1 + \boldsymbol{A}\,dt) = \boldsymbol{T}(x)\cdot \underset{x_0}{\overset{x}{\mathcal{P}}}(1 + \boldsymbol{A}^*\,dt)\cdot \boldsymbol{T}^{-1}(x_0). \tag{4.13}$$

We note that this transformation formula may e. g. be applied for the investigation of the behaviour of the solutions in possible poles of $\boldsymbol{A}$ and $\boldsymbol{B}$[1].

If especially

$$\boldsymbol{A}(x) = \boldsymbol{A}_1(x) + \boldsymbol{A}_2(x) \tag{4.14}$$

and we put

$$\boldsymbol{T}(x) = \underset{x_0}{\overset{x}{\mathcal{P}}}(1 + \boldsymbol{A}_1\,dt) = \underset{x_0}{\overset{x}{\mathcal{P}}}(\boldsymbol{A}_1), \tag{4.15}$$

(4.13) reduces, due to (4.11), (3.11) and (3.9), to

$$\underset{x_0}{\overset{x}{\mathcal{P}}}(\boldsymbol{A}_1 + \boldsymbol{A}_2) = \underset{x_0}{\overset{x}{\mathcal{P}}}(\boldsymbol{A}_1)\cdot \underset{x_0}{\overset{x}{\mathcal{P}}}(\boldsymbol{A}_3) \tag{4.16}$$

in which

$$\begin{aligned}\boldsymbol{A}_3 &= \mathcal{P}^{-1}(\boldsymbol{A}_1)\cdot(\boldsymbol{A}_1 + \boldsymbol{A}_2)\cdot\mathcal{P}(\boldsymbol{A}_1) - \mathcal{P}^{-1}(\boldsymbol{A}_1)\cdot\boldsymbol{A}_1\cdot\mathcal{P}(\boldsymbol{A}_1) = \\ &= \left(\underset{x_0}{\overset{x}{\mathcal{P}}}(\boldsymbol{A}_1)\right)^{-1}\cdot\boldsymbol{A}_2\cdot\underset{x_0}{\overset{x}{\mathcal{P}}}(\boldsymbol{A}_1).\end{aligned} \tag{4.17}$$

The formulae (4.16) and (4.17), which may be said to correspond to the formula for integrating by parts of ordinary integrals, become much simplified in case $\boldsymbol{A}_1$ commutes with $\boldsymbol{A}_2$, because in such case we obtain $\boldsymbol{A}_3 = \boldsymbol{A}_2$.

§ 5.

We shall now give the generalization of our theory to the case of the dimension being enumerable infinite. *It will be seen that all the contents of §§ 1—4 remains valid for infinite systems if we only demand the conditions* (2.11), (2.15) *and the corresponding condition for* $\boldsymbol{B}$ *still to be fulfilled,*[2] i. e.:

(a): The operator matrix $\boldsymbol{A}$ of our equation $\boldsymbol{Y}' = \boldsymbol{A}\cdot\boldsymbol{Y} + \boldsymbol{B}$ we assume to be absolutely exponentiable in the interval (x_0, x) (cf. p. 265)

$$(\alpha) \qquad \exp[\boldsymbol{K}\,|x - x_0|] = \sum_{\nu=0}^{\infty} \boldsymbol{K}^{\nu}\frac{|x - x_0|^{\nu}}{\nu!} < \infty, \tag{5.1}$$

$$\boldsymbol{K} = \max_{x_0 \leq t \leq x} |\boldsymbol{A}(t)|.$$

[1] Cf. e. g. Rasch (1930) p. 59 ff. or Rasch (1934) p. 110 ff.

[2] We note, however, that in the case of *infinite* systems the matrix $\boldsymbol{T}(x)$ occurring in (4.8)—(4.16) may not be quite arbitrary. We leave the discussion of the necessary conditions to the reader. We note, furthermore, that due to the remarks in § 1 we obtain by our generalization also a theory for ordinary linear differential equations of an *infinite* order. Such equations seem, however, not yet to have been met with in practice.

(b): We consider only such solutions of $\boldsymbol{Y}' = \boldsymbol{A} \cdot \boldsymbol{Y} + \boldsymbol{B}$ for which

$$(\beta) \qquad \exp[\boldsymbol{K}|x - x_0|] \cdot \boldsymbol{G} < \infty^{1}, \qquad (5.2)$$

$$\boldsymbol{G} = \max_{x_0 \leqq t \leqq \infty} |\boldsymbol{Y}(t)|.$$

Especially it then follows that (5.2) shall be satisfied for the initial point

$$\exp[\boldsymbol{K}|x - x_0|] \cdot |\boldsymbol{C}| < \infty, \qquad (5.3)$$

$$\boldsymbol{C} = \boldsymbol{Y}(x_0).$$

(c): Finally $\boldsymbol{B}$ is also assumed to satisfy (5.2), i. e.

$$(\gamma) \qquad \exp[\boldsymbol{K}|x - x_0|] \cdot \boldsymbol{M} < \infty, \qquad (5.4)$$

$$\boldsymbol{M} = \max_{x_0 \leqq t \leqq x} |\boldsymbol{B}(t)|.$$

The *essential* difference between the *infinite* and the *finite* case is, however, that in the latter case (α)—(γ) are always automatically fulfilled (cf. § 2), whereas this need not be the case in the former case. In part III we shall give examples showing that the conditions (α)—(γ) are only *sufficient*, but not *necessary* conditions for the theorems of uniqueness and existence to be true and, furthermore, that these two theorems themselves are not generally true. On the other hand we observe, however, that for a general theory the main condition (α) cannot be replaced by any weaker condition ensuring the necessary convergences. As we have seen we have, namely, that the majorizing expression (3.4) for the solution becomes identical with the solution proper, in case $\boldsymbol{A}$ and $\boldsymbol{C}$ are both constant and non-negative matrices.

We shall shortly discuss criteria which are *sufficient* to ensure the main condition (α) to be fulfilled. We have previously given the following four criteria[2], which will presumably cover most cases met with in the applications, at any rate in the theory of stochastic processes:

---

[1] We note that due to (4.2) the product-integral itself satisfies the condition (β), but, if $\boldsymbol{A}$ is only absolutely exponentiable in a finite interval, possibly only in half this interval. In general this fact is, however, irrelevant due to an exponentiable continuation being as a rule possible (cf p. 284). Furthermore, it is seen that if $\boldsymbol{C}$ satisfies (5.3), then $\boldsymbol{Y}(x) = \boldsymbol{F}(x, x_0) \cdot \boldsymbol{C}$ will satisfy (5.2)

[2] Cf. Arley (1943), chap. 7.

Type I: If $\boldsymbol{A}(x)$ is a *bounded* matrix, which means that the column (row) sums of $\boldsymbol{K}$ are uniformly bounded, i. e. there exists a number M, so that

$$\sum_{i=0}^{\infty} \boldsymbol{K}_{iq} < \boldsymbol{M} \quad \text{for all} \quad q = 0, 1, 2, \ldots$$
$$\left(\sum_{j=0}^{\infty} \boldsymbol{K}_{pj} < \boldsymbol{M} \quad \text{»} \quad \text{»} \quad p = 0, 1, 2, \ldots\right), \tag{5.5}$$

then $\boldsymbol{A}(x)$ is absolutely exponentiable in each interval (x_0, x) for which $\boldsymbol{K}$ satisfies (5.5). Furthermore, the exponential is again bounded.

Type II: If $\boldsymbol{A}(x)$ is a row (column) *half-finite* matrix of order N, which means that all the rows (columns) contain only zeros after the N'th column (row) index, then $\boldsymbol{A}(x)$ is absolutely exponentiable in each interval (x_0, x). Furthermore, the exponential $(-\mathbf{1})$ is again row (column) half-finite.

Type III: If $\boldsymbol{A}(x)$ is a row (column) *half* matrix, which means that all the elements above (below) the main diagonal vanish, then $\boldsymbol{A}(x)$ is absolutely exponentiable in each interval (x_0, x). Furthermore, the exponential is again a row (column) half matrix.

Type IV: If $\boldsymbol{A}(x)$ is a column (row) *semi-diagonal* matrix, which means that all the elements below (above) the diagonal lying parallel with and in the distance l below (above) the main diagonal vanish, and the numerical column (row) sums all exist and are bounded by the relation

$$\sum_{i=0}^{\infty} |\boldsymbol{A}_{iq}(x)| \leqq f(x) \cdot q \quad \text{for all} \quad q > 0$$
$$\left(\sum_{j=0}^{\infty} |\boldsymbol{A}_{pj}(x)| \leqq f(x) \cdot p \quad \text{»} \quad \text{»} \quad p > 0\right), \tag{5.6}$$

then $\boldsymbol{A}(x)$ is absolutely exponentiable in each interval (x_0, x) satisfying

$$|x - x_0| < \frac{1}{lC}, \quad C = \max_{x_0 \leqq t \leqq x} f(t). \tag{5.7}$$

Furthermore, we have shown[1] that if the numerical column (row) sums increase stronger than the first power of the column (row) number, it be ever so little, then $\boldsymbol{A}(x)$ need not be absolutely exponentiable in any interval. E. g. we showed that the constant semi-diagonal matrix $(l = 1)$

[1] Arley (1943) ex. 1, p. 203.

$$A = \begin{Bmatrix} 0 & f(1) & 0 & 0 & \cdots \\ f(0) & 0 & f(2) & 0 & \cdots \\ 0 & f(1) & 0 & f(3) & \cdots \\ 0 & 0 & f(2) & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{Bmatrix}, \quad f(q) = q^{1+r} \quad (r > 0), \tag{5.8}$$

is not absolutely exponentiable in any interval as already the diagonal elements

$$(A^{2\nu})_{qq} \frac{|x-x_0|^{2\nu}}{(2\nu)!} > \left(\prod_{\beta=1}^{\nu} f(q+\beta)\right)\left(\prod_{\beta=0}^{\nu-1} f(q+\beta)\right) \frac{|x-x_0|^{2\nu}}{(2\nu)!} \underset{\nu\to\infty}{\to} \infty \tag{5.9}$$

$$\begin{pmatrix} \text{for all } q = 0, 1, 2, \ldots \\ \text{»} \quad \text{»} \quad r > 0 \end{pmatrix}.$$

The criteria I—IV can be shown to be special cases of the following more general criterion — or its analogue operating with row sums — which is due to Cramér[1]:

A sufficient condition for $A(x)$ to be absolutely exponentiable in (x_0, x) is the existence of a non-negative matrix

$$M = \{M_{\nu q}\} \geqq 0 \tag{5.10}$$

satisfying

$$M_{0q} \geqq 1 \qquad \text{for all} \quad q = 0, 1, 2, \ldots \tag{5.11}$$

$$\sum_{\alpha=0}^{\infty} M_{\nu\alpha} K_{\alpha q} \leqq M_{\nu+1, q} \qquad \text{»} \quad \text{»} \quad \nu, q = 0, 1, 2, \ldots \tag{5.12}$$

$$\sum_{\nu=0}^{\infty} M_{\nu q} \frac{|x-x_0|^{\nu}}{\nu!} < \infty \qquad \text{»} \quad \text{»} \quad q = 0, 1, 2, \ldots . \tag{5.13}$$

It is easily seen that due to (5.11) and (5.12) we have for $\nu = 0$ and $\nu = 1$

$$\sum_{i=0}^{\infty} (K^{\nu})_{iq} \leqq M_{\nu q}. \tag{5.14}$$

Let (5.14) be true for some value ν, we then have from (5.12) and (5.14)

$$\sum_{i=0}^{\infty} (K^{\nu+1})_{iq} = \sum_{i=0}^{\infty}\sum_{\alpha=0}^{\infty} (K^{\nu})_{i\alpha} K_{\alpha q} = \sum_{\alpha=0}^{\infty}\left(\sum_{i=0}^{\infty} (K^{\nu})_{i\alpha}\right) K_{\alpha q} \leqq \sum_{\alpha=0}^{\infty} M_{\nu\alpha} K_{\alpha q} \leqq M_{\nu+1, q}. \tag{5.15}$$

[1] Private communication. We wish to express our most sincere thanks to Prof. Cramér for kindly communicating this theorem to us.

(5.14) is, thus, generally true. From (5.13) and (5.14) it then finally follows that

$$\sum_{i=0}^{\infty}(\exp[\boldsymbol{K}|x-x_0|])_{iq}=\sum_{\nu=0}^{\infty}\left(\sum_{i=0}^{\infty}(\boldsymbol{K}^{\nu})_{iq}\right)\frac{|x-x_0|^{\nu}}{\nu!}\leqq\sum_{\nu=0}^{\infty}\boldsymbol{M}_{\nu q}\frac{|x-x_0|^{\nu}}{\nu!}<\infty. \quad (5.16)$$

The exponential thus exists, having, furthermore, convergent column sums.

Of course we have also a corresponding criterion operating with the row sums instead of the column sums (which is e. g. necessary if $\boldsymbol{A}$ is a *row* half matrix). By considering the transposed matrix $\tilde{\boldsymbol{A}}$ instead of $\boldsymbol{A}$ we see, however, that we need only a criterion working with the column sums. (In the application to the theory of stochastic processes (cf. p. 286) it is, namely, the column sums which enter.)

Finally, it may be of interest to note that, as also pointed out by Cramér, the conditions of type IV may be weakened so that $\boldsymbol{A}$ need not be semi-diagonal, if only the elements in each column of $\boldsymbol{A}$ decrease sufficiently rapidly:

Type V: Let $m_0, m_1, m_2, \ldots$ be a non-decreasing sequence of positive numbers and $g(x)\geqq 0$ so that

$$\sum_{i=0}^{\infty}\exp[m_s-m_q]\,|\boldsymbol{A}_{iq}(x)|\leqq g(x)m_q \quad \text{for all} \quad q=0,1,2,\ldots \quad (5.17)$$

in which

$$s=s(i,q)=\begin{cases} q & \text{for } i=0,1,2,\ldots,q \\ i & \text{»} \quad i>q.\end{cases} \quad (5.18)$$

We then have for all $i, q, \nu\geqq 0$

$$\left(\frac{m_i+\nu}{m_q+\nu}\right)^{\nu}=\left(1+\frac{m_i-m_q}{m_q+\nu}\right)^{\nu}\leqq\exp\left[\frac{\nu}{m_q+\nu}(m_i-m_q)\right]\leqq\exp[m_s-m_q]. \quad (5.19)$$

Consequently it follows from (5.17) that

$$\sum_{i=0}^{\infty}(g(x))^{\nu}(m_i+\nu)^{\nu}|\boldsymbol{A}_{iq}(x)|\leqq(g(x))^{\nu+1}(m_q+\nu+1)^{\nu+1} \text{ for all } \nu, q=0,1,2,\ldots. \quad (5.20)$$

Putting now

$$\boldsymbol{M}_{\nu q}=c^{\nu}(m_q+\nu)^{\nu}, \qquad c=\max_{x_0\leqq t\leqq x} g(t), \quad (5.21)$$

we thus see that the exponentiability conditions (5.10)—(5.13) are satisfied for all values of $|x-x_0|$ for which

$$\lim_{\nu\to\infty}\frac{c(m_q+\nu+1)^{\nu+1}}{(m_q+\nu)^{\nu}}\cdot\frac{|x-x_0|}{\nu+1}=ce|x-x_0|<1, \quad (5.22)$$

i. e.

$$|x-x_0|<\frac{1}{ce}. \quad (5.23)$$

It will be seen that the condition (5.17) is a generalization of the conditions in type IV. Putting, namely,

$$m_q = k q \tag{5.24}$$

and

$$g(x) = \frac{1}{k} e^{kl} f(x) \tag{5.25}$$

we obtain from (5.6) and the fact that A is semi-diagonal

$$\sum_{i=0}^{q+l} |A_{iq}(x)| \leqq \sum_{i=0}^{q+l} \exp[m_s - m_q] A_{iq}(x) \leqq$$

$$\leqq \exp[m_{q+l} - m_q] \sum_{i=0}^{q+l} |A_{iq}(x)| \leqq e^{kl} f(x)\, q = g(x)\, m_q, \tag{5.26}$$

i. e. (5.17). Next we obtain from (5.21), (5.25) and (5.7),

$$c = \frac{1}{k} e^{kl}\, C. \tag{5.27}$$

For $k = \frac{1}{l}$, c is seen to become as small as possible, i. e. (5.23) cannot generally give any greater interval of exponentiability than

$$|x - x_0| < \frac{1}{C l e^2}. \tag{5.28}$$

This interval is, however, e^2 times smaller than the interval obtained from (5.7). The reason for this difference is, of course, that different majorizations are applied, a factor of the type $\nu!$ entering in the proof of (5.7), but of the type ν^ν in the proof of (5.23).

Finally it may, however, be seen that if the column sums of $|A(x)|$ increase more rapidly with q than in (5.6), i. e. linearly with q, (5.17) cannot be fulfilled except in very special cases. From (5.17) it follows, namely, that

$$\exp[m_{q+1} - m_q] \sum_{i=q+1}^{\infty} |A_{iq}(x)| \leqq \sum_{i=0}^{\infty} \exp[m_s - m_q] |A_{iq}(x)| \leqq g(x)\, m_q. \tag{5.29}$$

If, now, the column sums of $|A(x)|$ increase more rapidly than linearly with q, we see from (5.17) that also the numbers m_q increase more rapidly than linearly with q. In this case we see, however, from (5.29) that unless $\sum_{i=q+1}^{\infty} |A_{iq}(x)|$ decreases sufficiently with increasing values of q, we obtain a contradiction because the left hand side will increase more rapidly with q than the right hand side.

PART II.

Application to the Theory of Stochastic Processes.

§ 6.

A discontinuous, stochastically definite process[1] in which the stochastic variable can assume only an enumerable manifold of values is characterized analytically by a *relative probability function* of the type $P(n, t; n', s)$, $n, n' = 0, 1, 2, \ldots, t \geqq s$, denoting the conditioned probability of a stochastic variable assuming the value n at the time t, relative to the hypothesis, that it assumes the value n' at the time s. We note the essential fact that $t \geqq s$, as in all probability questions the time can move only in the *forward* direction (cf. p. 280). By the expression *stochastically definite*[2] we mean that the function $P(n, t; n', s)$ is independent of any knowledge of the antecedent of the process, i. e. of the development of the process before the time s. The exact statement of this fact is the following: Let $a < s_1 < s_2 < \cdots < s_p = s \leqq t = t_1 < t_2 < \cdots < t_q < b$. Next, let us consider the simultaneous conditioned probability distribution of the values of the stochastic variable at the times $t_1, t_2, t_3, \ldots, t_q$ relative to the hypothesis that it assumes a certain value $n(s)$ at the time s. *If* now, this simultaneous probability distribution is *independent* of the further hypothesis (the antecedent of the process) that the variable has assumed certain values $n(s_1), n(s_2), \ldots, n(s_{p-1})$ at the times $s_1, s_2, \ldots, s_{p-1}$ and this holds true for *arbitrary* values of $p, q, s_1, \ldots, s_p, t_1, \ldots, t_q$ and $n(s_1), \ldots, n(s_p)$, *then* the process is called stochastically definite in the interval (a, b).

We note that even in simple practical applications[3] we may meet with stochastic processes, which are *not* stochastically definite, the antecedent entering in a decisive way. In the examples just mentioned the relevant probability

[1] Arley (1943), part I. In part II of this paper we have discussed various special stochastic processes of both one and two dimensions and their application to the theory of cosmic ray cascade showers. For the mathematical theory see also Kolmogoroff (1931), Feller (1937), Lundberg (1940) and Fréchet (1938). In the paper of Lundberg special attention is paid to the application of the theory to sickness and accident statistics.

[2] Khintchine (1934) has suggested the expression »Markoff process« instead of »stochastically definite process«. We think, however, that the latter expression is already so widely adopted, that an alteration in the terminology would rather be confusing. Furthermore, the former expression is generally used to express the fact that the stochastic variable can assume only an enumerable manifold of values.

[3] Cf. e. g. Arley (1943) §§ 4.5 and 4.9.

distributions were, however, simply the marginal distributions of a multi-dimensional stochastic process, i. e. a process in which several mutually dependent stochastic variables enter. Another way in which a non-definite process may be reduced to a definite one is to take the knowledge of the antecedent into account by introducing some further quantities, parameters, as e. g. the velocities in classical physics.[1]

The specification of the process is now given through the introduction of two functions, the *intensity function* $p(n, t)$ and the *relative transition probability function* $\Pi(n; n', t)$, both assumed to be *continuous*[2]. Here $p(n, t)\,dt$ is an asymptotic expression for the probability of a stochastic change of the variable taking place in the interval between t and $t + dt$ when the variable assumes the value n at the time t. Next $\Pi(n; n', t)$ is the conditioned probability of the variable assuming the value n at the time $t + dt$ relative to the hypothesis that a stochastic change of the variable from the state n' has taken place during the interval between t and $t + dt$. From the definition of the p and Π functions it follows that

$$p(n, t) \geqq 0 \tag{6.1}$$

$$0 \leqq \Pi(n; n', t) \leqq 1; \qquad \Pi(n'; n', t) \equiv 0 \tag{6.2}$$

$$\sum_{n=0}^{\infty} \Pi(n; n', t) \equiv 1. \tag{6.3}$$

Next it follows from the definitions that the P functions must satisfy the following five *fundamental conditions*:

$$\lim_{t \to s} P(n, t; n', s) = \lim_{s \to t} P(n, t; n', s) = \delta_{n n'} \tag{6.4}$$

$$P(n, t + \Delta t; n', t) = (1 - p(n', t)\Delta t)\,\delta_{n n'} + \Pi(n; n', t)\,p(n', t)\,\Delta t + o(\Delta t) \tag{6.5}$$
$$(o(\Delta t) = f(n, t, \Delta t, n')).$$

$$P(n, t;\ n', s) = \sum_{n''=0}^{\infty} P(n, t; n'', \tau)\,P(n'', \tau; n', s) \text{ for all } \tau \text{ in } s \leqq \tau \leqq t \tag{6.6}$$

$$0 \leqq P(n, t; n', s) \leqq 1 \tag{6.7}$$

$$\sum_{n=0}^{\infty} P(n, t;\ n', s) \equiv 1. \tag{6.8}$$

The relation (6.6) is called the *Chapman-Kolmogoroff equation*.

---

[1] Kolmogoroff (1931).

[2] We note that it is possible to give up this assumption of continuity and thus obtain a more general theory of stochastic processes. We intend to deal with this problem in a later paper.

§ 7.

Introducing the *distribution matrix*

$$\boldsymbol{P}(t,s) = \{P(n,t;n',s)\}, \tag{7.1}$$

the diagonal *intensity matrix*

$$\boldsymbol{p}(t) = \{p(n,t)\delta_{nn'}\} \tag{7.2}$$

and the *relative transition matrix*

$$\boldsymbol{\Pi}(t) = \{\Pi(n;n',t)\}, \tag{7.3}$$

we see that the fundamental conditions (6.4)—(6.8) may be written in the compact form

$$\lim_{t\to s} \boldsymbol{P}(t,s) = \lim_{s\to t} \boldsymbol{P}(t,s) = \mathbf{1} \tag{7.4}$$

$$\boldsymbol{P}(t+\Delta t, t) = \mathbf{1} + \boldsymbol{A}(t)\Delta t + \boldsymbol{o}(\Delta t) \tag{7.5}$$

$$\boldsymbol{P}(t,s) = \boldsymbol{P}(t,\tau)\cdot\boldsymbol{P}(\tau,s) \tag{7.6}$$

$$\mathbf{0} \leqq \boldsymbol{P}(t,s) \leqq \{1\} \tag{7.7}$$

$$\boldsymbol{\Sigma}\cdot\boldsymbol{P}(t,s) = \sum_{n=0}^{\infty} P(n,t;n',s) \equiv \{1\} \tag{7.8}$$

in which

$$\boldsymbol{A}(t) = -\boldsymbol{p}(t) + \boldsymbol{\Pi}(t)\cdot\boldsymbol{p}(t) = (\boldsymbol{\Pi}-\mathbf{1})\cdot\boldsymbol{p}. \tag{7.9}$$

Due to (6.3) $\boldsymbol{A}$ satisfies

$$\boldsymbol{\Sigma}\cdot\boldsymbol{A} = \sum_{n=0}^{\infty} \boldsymbol{A}_{nn'} \equiv \mathbf{0}. \tag{7.10}$$

Introducing (7.5) for the first, respectively the second, factor in (7.6), we obtain

$$\Delta_t\boldsymbol{P}(t,s) = \boldsymbol{P}(t+\Delta t,s) - \boldsymbol{P}(t,s) = \boldsymbol{A}(t)\cdot\boldsymbol{P}(t,s)\Delta t + \boldsymbol{o}(\Delta t)\cdot\boldsymbol{P}(t,s) \tag{7.11}$$

and

$$\Delta_s\boldsymbol{P}(t,s) = \boldsymbol{P}(t,s+\Delta s) - \boldsymbol{P}(t,s) = -\boldsymbol{P}(t,s)\cdot\boldsymbol{A}(s)\Delta s - \boldsymbol{P}(t,s)\cdot\boldsymbol{o}(\Delta s). \tag{7.12}$$

Making now the natural assumptions that (a) $\boldsymbol{A}\cdot\boldsymbol{P}$ and $\boldsymbol{P}\cdot\boldsymbol{A}$ are both convergent and (b) $\boldsymbol{o}(\Delta t)\cdot\boldsymbol{P} = \boldsymbol{o}(\Delta t)$ and $\boldsymbol{P}\cdot\boldsymbol{o}(\Delta s) = \boldsymbol{o}(\Delta s)$[1], we see that $\boldsymbol{P}$ has partial differential coefficients both with respect to t and s, which satisfy the *fundamental equations*

[1] We have not been able to decide whether or not these two assumptions follow from the previous ones. As will be seen they are in any case necessary for the theory in the present form (cf., however, the remarks at the end of p. 283).

$$\frac{\partial}{\partial t} \boldsymbol{P}(t, s) = \boldsymbol{A}(t) \cdot \boldsymbol{P}(t, s) \tag{7.13}$$

$$\frac{\partial}{\partial s} \boldsymbol{P}(t, s) = - \boldsymbol{P}(t, s) \cdot \boldsymbol{A}(s). \tag{7.14}$$

It is now the object of our theory to show that if $\boldsymbol{A}$ is given by (7.9) *and* $\boldsymbol{p}$ *and* $\boldsymbol{\Pi}$ *satisfy* (6.1)—(6.3) *and are assumed to be continuous for all* $t \geqq s$, *then* (7.13) *and* (7.14) *has each one and only one solution which is the same for both systems and which satisfies the fundamental conditions* (7.4)—(7.8).

Assuming now the matrix $\boldsymbol{A}(t)$ to be absolutely exponentiable in some interval $s \leqq \tau \leqq t$ it follows from part I that (7.13) has one and only one solution, satisfying (5.2), given by the product-integral

$$\boldsymbol{P}(t, s) = \underset{s}{\overset{t}{\mathcal{P}}} (\mathbf{1} + \boldsymbol{A}(\tau)\, d\tau) \tag{7.15}$$

which will, due to (3.10), also be the — unique — solution of the *adjointed* equation (7.14). Furthermore, due to (3.11), (4.1) and (3.14) this unique solution will automatically satisfy the three first fundamental conditions (7.4)—(7.6). We thus see, that the product-integral is the ideal mathematical tool for the theory of stochastic, discontinuous processes.

The left hand side of the fourth fundamental condition (7.7) follows immediately from the definition of the product-integral, mentioned in (3.8), and the essential fact that $t \geqq s$.[1] Due to (6.1) and (6.2) we have, namely, that all the non-diagonal elements of $\boldsymbol{A}$ are non-negative, because all $\Delta_i > 0$, and the diagonal elements of the form $A_{nn} = 1 - p(n, t) \Delta t$. Thus in the limit also the diagonal elements become non-negative. We note the important fact that this statement need not be true if $t < s$ (cf. p. 277). (7.7) may, however, also be seen by means of the transformation formula (4.16). Putting in this formula $\boldsymbol{A}_1 = -\boldsymbol{p}$ and $\boldsymbol{A}_2 = \boldsymbol{\Pi} \cdot \boldsymbol{p}$ we obtain, due to the fact that $\boldsymbol{p}$ is a diagonal matrix and its product-integral thus given by the diagonal matrix (4.6),

$$\boldsymbol{P}(t, s) = \exp\left[-\int_s^t \boldsymbol{p}(\tau)\, d\tau\right] \cdot \underset{s}{\overset{t}{\mathcal{P}}} (\mathbf{1} + \boldsymbol{A}_3(\tau)\, d\tau) \tag{7.16}$$

in which, from (4.17),

$$\boldsymbol{A}_3(t) = \exp\left[\int_s^t \boldsymbol{p}(\tau)\, d\tau\right] \cdot \boldsymbol{\Pi}(t) \cdot \boldsymbol{p}(t) \cdot \exp\left[-\int_s^t \boldsymbol{p}(\tau)\, d\tau\right] \geqq 0. \tag{7.17}$$

[1] We have not been able to decide whether there may exist a stochastic process admitting of other non-negative solutions than (7.15), i.e. which do not satisfy (5.2) (cf. p. 299).

All the matrix elements occurring in (7. 16) and (7. 17) now being non-negative, the left hand side of (7. 7) follows at once. We note that from the Peano series (3. 7) for the product-integral of a non-negative matrix it follows that all the diagonal elements of the product-integral are *positive*. From (7. 16) it is next seen that also the diagonal elements of $P(t, s)$ have the same property

$$P_{nn}(t, s) > 0 \quad \text{for all} \quad t \geqq s. \tag{7.18}$$

From the Chapman-Kolmogoroff equation (7.6) we next have, due to the left hand side of (7. 7), for the non-diagonal elements

$$P_{nn'}(t, s) \geqq P_{nn}(t, \tau)\, P_{nn'}(\tau, s) \geqq 0. \tag{7.19}$$

(7. 19) combined with (7. 18) shows that if $P_{nn'} = 0$ for some value $t_0 > s$, then this is the case for all times in $s \leqq t \leqq t_0$, i. e. we have generally

$$P_{nn'}(t, s) \begin{cases} \text{either} \equiv 0 \text{ for all } t \text{ in } s \leqq t \leqq t_0 \\ \text{or} \quad > 0 \quad \text{»} \quad \text{»} \quad t > s. \end{cases} \tag{7.20}$$

The right hand side of (7. 7) now follows immediately from the fundamental equation (7. 13), (7. 4) and the relation (7. 10), because

$$\sum_{n=0}^{\infty} P_{nn'}(t,s) = \lim_{N \to \infty} \sum_{n=0}^{N} P_{nn'}(t, s) = 1 + \lim_{N \to \infty} \int_{s}^{t} \left[\sum_{n''=0}^{\infty} \left(\sum_{n=0}^{N} A_{nn''}(\tau) \right) P_{n''n'}(\tau, s) \right] d\tau. \tag{7.21}$$

For each fixed value of N we have, namely, a finite number of convergent series and we may, therefore, first interchange $\sum\limits_{n}$ and $\int d\tau$, and next $\sum\limits_{n}$ and $\sum\limits_{n''}$. Firstly, the left hand side is, due to (7. 20), non-decreasing with increasing N and the right hand side, therefore, either tends to a finite limit or to ∞. Secondly, from (6. 2), (6. 3), (7. 9) and (7. 20) it follows that for each fixed value of n'' and τ

$$\left(\sum_{n=0}^{N} A_{nn''}(\tau) \right) P_{n''n'}(\tau, s) \leqq 0 \quad \begin{matrix} \text{for all} & N \geqq n'' \\ \text{»} \quad \text{»} & \tau \geqq s. \end{matrix} \tag{7.22}$$

Consequently we have from (7. 21) that

$$\sum_{n=0}^{\infty} P_{nn'}(t, s) \leqq 1 \tag{7.23}$$

from which the right hand side of (7. 7) follows immediately. We note that (7. 23) is just what may be expected to hold true generally. It is, namely, a priori possible that the stochastic variable can increase so strongly with the time, that it may reach the value 'infinity' with a positive probability for a *finite* value of t, i. e.

$$P(\infty, t; n', s) = 1 - \sum_{n=0}^{\infty} P(n, t; n', s) > 0. \tag{7.24}$$

§ 8.

Before discussing the main problem of the theory, namely whether (7.8) holds true or not, we shall consider the problem of the *absolute probability distribution*. If $\boldsymbol{P}(s)$ is an arbitrary matrix function consisting of only one column which satisfies

$$\mathbf{0} \leqq \boldsymbol{P}(s) \leqq \{1\} \tag{8.1}$$

and

$$\boldsymbol{\Sigma} \cdot \boldsymbol{P} = \sum_{n=0}^{\infty} P_n(s) = 1, \tag{8.2}$$

$P_n(s)$ can be interpreted as the absolute probability of the stochastic variable assuming the value n at the time s. From the definition of $\boldsymbol{P}(t, s)$ and $\boldsymbol{P}(s)$ it follows that the absolute probability distribution at the time t is given by

$$\boldsymbol{P}(t) = \boldsymbol{P}(t, s) \cdot \boldsymbol{P}(s) \tag{8.3}$$

in which $\boldsymbol{P}(t)$ is also a solution of the fundamental equation (7.13) and satisfies (8.1) and

$$\lim_{t \to s} \boldsymbol{P}(t) = \boldsymbol{P}(s). \tag{8.4}$$

Firstly, it follows immediately from (7.7) and (8.2) that $\boldsymbol{P}(t)$ given in (8.3) exists and is non-negative, the convergence being, furthermore, uniform in t. Secondly, the convergence being, of course, absolute it next follows from (7.23) that we have

$$\sum_{n=0}^{\infty} P_n(t) = \sum_{n'=0}^{\infty} \left(\sum_{n=0}^{\infty} P_{nn'}(t, s) \right) P_{n'}(s) \leqq \sum_{n'=0}^{\infty} P_{n'}(s) = 1, \tag{8.5}$$

the sign of equality holding true if and only if the same is the case in (7.23). Consequently the right hand side of (8.1) is fulfilled.

Finally, we shall prove that $\boldsymbol{P}(t)$ given in (8.3) satisfies

$$\boldsymbol{P}(t) = \boldsymbol{P}(s) + \int_s^t (\boldsymbol{A}(\tau) \cdot \boldsymbol{P}(\tau, s)) \cdot \boldsymbol{P}(s)\, d\tau = \boldsymbol{P}(s) + \int_s^t \boldsymbol{A}(\tau) \cdot \boldsymbol{P}(\tau)\, d\tau, \tag{8.6}$$

in which the associative rule holds true because $\boldsymbol{P}(\tau, s)$ and $\boldsymbol{P}(s)$ are non-negative and $\boldsymbol{A}(\tau)$ has in each fixed row at most one negative element. (8.6) is proved in the following way. As $\boldsymbol{P}(t,s)$ satisfies $\boldsymbol{P}(t,s)=\mathbf{1}+\int\limits_s^t \boldsymbol{A}(\tau)\cdot\boldsymbol{P}(\tau,s)\,d\tau$, we have for all values of $N \geqq n$

$$\sum_{n'=0}^{N} \boldsymbol{P}_{nn'}(t,\,s)\,\boldsymbol{P}_{n'}(s) = \boldsymbol{P}_n(s) + \sum_{n'=0}^{N}\Big(\int\limits_s^t \boldsymbol{A}\quad \boldsymbol{P}\,d\tau\Big)_{nn'}\boldsymbol{P}_{n'}(s) =$$

$$\boldsymbol{P}_n(s) + \int\limits_s^t \left(\sum_{n'=0}^{N}(\boldsymbol{A}(\tau)\cdot\boldsymbol{P}(\tau,s))_{nn'}\,\boldsymbol{P}_{n'}(s)\right)d\tau =$$

$$= \boldsymbol{P}_n(s) + \int\limits_s^t \left(\sum_{n''=0}^{\infty}\boldsymbol{A}_{nn''}(\tau)\left(\sum_{n'=0}^{N}\boldsymbol{P}_{n''n'}(\tau,s)\,\boldsymbol{P}_{n'}(s)\right)\right)d\tau =$$

$$\boldsymbol{P}_n(s) + {}^{(1)}\int\limits_s^t \boldsymbol{A}_{nn}\left(\sum_{n'=0}^{N}\boldsymbol{P}_{nn'}(\tau,s)\,\boldsymbol{P}_{n'}(s)\right)d\tau +$$

$$+ {}^{(2)}\int\limits_s^t \sum_{n''\neq n}\boldsymbol{A}_{nn''}(\tau)\left(\sum_{n'=0}^{N}\boldsymbol{P}_{n''n'}(\tau,s)\,\boldsymbol{P}_{n'}(s)\right)d\tau. \qquad (8.7)$$

Firstly we have that, due to (7.9) and $\boldsymbol{P}(\tau,s)$ and $\boldsymbol{P}(s)$ being non-negative, the integrands in both terms (1) and (2) are monotonously decreasing, respectively increasing, functions with increasing values of N. Going to the limit, $N\to\infty$, we secondly obtain from (8.3) that both terms (1) and (2) are convergent. Consequently it follows from a well-known theorem[1] that we may in (8.7) go to the limit before we integrate, which fact proves (8.6).

$\boldsymbol{P}(t)$ thus being an integral it follows[2] that $\boldsymbol{P}'(t)$ exists *almost everywhere* and satisfies

$$\boldsymbol{P}'(t) = \boldsymbol{A}(t)\cdot\boldsymbol{P}(t) = (\boldsymbol{A}(t)\cdot\boldsymbol{P}(t,s))\cdot\boldsymbol{P}(s) = \boldsymbol{P}'(t,s)\cdot\boldsymbol{P}(s). \qquad (8.8)$$

If especially $\boldsymbol{P}(s)$ besides (8.1) and (8.2) satisfies the condition (5.3) and, consequently, $\boldsymbol{P}(t)$ the condition (β), (5.2), (cf. [1] p. 272), then it follows from part I that (8.8) holds true everywhere. We note, however, that in general this is not the case (cf. ex. (9.II) and ex. (13.I)). *In this connection it may be worth while to observe that the examples mentioned show that the assumptions underlying the deduction of the fundamental equations* (7.13) *and* (7.14), *namely the convergence of*

---

[1] Cf. e. g. Titchmarsh (1932) § 10.82.

[2] Cf. e. g. Titchmarsh (1932) § 11.5.

$\boldsymbol{A}\cdot\boldsymbol{P}$ *and* $\boldsymbol{P}\cdot\boldsymbol{A}$ (cf. [1] p. 279), *are too narrow for a general theory of stochastic processes as we may very well in practice meet with processes of just the type mentioned, i. e. in which the rate of change of the probability of certain values becomes infinite at certain times* (and in which even the condition (7.5) is possibly no more fulfilled).[1]

It now follows that $\boldsymbol{P}(t,s)$ given in (7.15) exists and is the unique solution in the *whole* region of definition independent of whether $\boldsymbol{A}$ is exponentiable in $s\leqq\tau<\infty$ or only in a finite interval $s\leqq\tau\leqq t$. In the last case there exists a finite 'radius of exponentiability', i. e. a radius of convergence of the series (3.7). Let t_1 be a point within this radius. We then simply start once more from this point and repeat the iteration process, calculating the product-integral $\boldsymbol{P}(t,t_1)$ in a new interval $t_1\leqq t\leqq t_2$. Each column of $\boldsymbol{P}(t_1,s)$ satisfying (8.1) and (8.2) it follows at once from the above discussion of the absolute probability distribution that

$$\boldsymbol{P}(t,s)=\boldsymbol{P}(t,t_1)\cdot\boldsymbol{P}(t_1,s) \tag{8.9}$$

exists and is — at any rate *almost* everywhere — the solution in the *whole* interval $s\leqq t\leqq t_2$. This procedure, which we call *exponentiable continuation*, we may now repeat ad infinitum, obtaining a series of continuation points $t_0=s<t_1<t_2<\cdots<T$. $\boldsymbol{P}(t,s)$ given in (7.15) thus exists and is the unique solution in the whole interval $s\leqq t\leqq T$. It is, of course, possible that the exponentiable continuation stops *within* the region of definition of $\boldsymbol{A}$. In such case our theory would turn out to be too narrow. It is, however, easily seen that the continuation may be carried through to arbitrarily high values for matrices of the types I—V. In the case of the types I—III there is no problem as $\boldsymbol{A}$ is absolutely exponentiable for all intervals. In the case of the types IV and V the distance between two consecutive continuation-points is limited by a relation of the form

$$t_n-t_{n-1}=\frac{\theta}{\text{const. } C_n},\quad 0<\theta<1,\qquad C_n=\max_{t_{n-1}\leqq\tau\leqq t_n} f(\tau). \tag{8.10}$$

If, now, the continuation process should stop, i. e. $t_n\underset{n\to\infty}{\rightarrow}T<\infty$, we obtain a contradiction, as the left hand side of (8.10) then tends to 0, and the right hand side does not, θ being a constant and $f(t)$ being finite for all values of t.

[1] We intend in a later paper to put the theory in a more general form comprising such processes.

Example (8. I).

In the general case we may, on the other hand, very well meet with matrices for which the continuation process does stop. Let us, namely, consider the matrix given by[1]

$$A_{pq}(x) = K_{pq} = \left(\frac{x}{4}\right)^p 2^q, \qquad p, q = 0, 1, 2, \ldots, \qquad x \geqq 0. \tag{8.11}$$

By induction we can prove that

$$(K^\nu)_{pq} = \left(\frac{x}{4}\right)^p 2^q \left(\frac{2}{2-x}\right)^{\nu-1}, \quad \nu \geqq 1, \qquad x < 2. \tag{8.12}$$

(8. 12) is obviously satisfied for $\nu = 1$. Let it be true for ν, we then have

$$(K^{\nu+1})_{pq} = \sum_{\alpha=0}^{\infty} (K^\nu)_{p\alpha} K_{\alpha q} = \left(\frac{x}{4}\right)^p \left(\frac{2}{2-x}\right)^{\nu-1} \sum_{\alpha=0}^{\infty} 2^\alpha \left(\frac{x}{4}\right)^\alpha 2^q =$$
$$\begin{cases} \left(\frac{x}{4}\right)^p 2^q \left(\frac{2}{2-x}\right)^{(\nu+1)-1} & \text{for} \quad x < 2 \\ \infty & \quad » \quad x \geqq 2. \end{cases} \tag{8.13}$$

Consequently $A(x)$ is absolutely exponentiable in the interval (x_0, x)[2],

$$0 \leqq x_0 \leqq x < 2, \tag{8.14}$$

because

$$(\exp[K(x - x_0)])_{pq} = \delta_{pq} + \left(\frac{x}{4}\right)^p 2^{q-1}(2 - x) \exp\left[\frac{2(x - x_0)}{2-x}\right]. \tag{8.15}$$

(8. 15) shows in fact that we cannot continue exponentiably beyond the critical point $x = 2$.

Finally we shall discuss the last fundamental condition (7. 8). As already mentioned (cf. the end of p. 281) this condition is, in contrast to the preceding four conditions, not generally fulfilled. In fact it is easy to indicate processes for which (7. 24) holds true (cf. p. 288).

[1] We note that this A does not belong to a stochastic process. It could, however, easily be modified so that this were the case.

[2] We observe that putting $M_{\nu q} = \left(\frac{2}{2-x}\right)^\nu 2^q$ it may be seen that $A(x)$ is covered by Cramér's criterion (cf. p. 274).

Let us first assume that the matrix of our process is covered by Cramér's criterion (cf. p. 274)[1]. We then have, due to (5. 1), (4. 2), (5. 14) and (5. 13),

$$\left|\sum_{n=0}^{\infty} \frac{\partial}{\partial \tau} P_{n n'}(\tau, s)\right| \leqq \sum_{n=0}^{\infty} \sum_{n''=0}^{\infty} |A_{n n''}(\tau)\, P_{n'' n'}(\tau, s)| \leqq \sum_{n=0}^{\infty} \sum_{\nu=0}^{\infty} (K^{\nu+1})_{n n'} \frac{(t-s)^{\nu}}{\nu!} \leqq$$
$$\sum_{\nu=0}^{\infty} M_{\nu+1, n'} \frac{(t-s)^{\nu}}{\nu!} = \frac{d}{d t} \sum_{\nu=0}^{\infty} M_{\nu, n'} \frac{(t-s)^{\nu}}{\nu!} < \infty \text{ for all } \tau \text{ in } s \leqq \tau \leqq t. \tag{8. 16}$$

(8. 16) shows:

(a) that the product $A \cdot P$ is *uniformly* convergent and thus, A and P being continuous, that also $\frac{\partial}{\partial t} P$ is continuous,

(b) that $\sum_{n=0}^{\infty} \frac{\partial}{\partial t} P_{n n'}$ is a uniformly convergent series of continuous functions, from which fact it follows that

$$\frac{\partial}{\partial t} \sum_{n=0}^{\infty} P_{n n'} = \sum_{n=0}^{\infty} \frac{\partial}{\partial t} P_{n n'}, \tag{8. 17}$$

(c) that $\sum_{n} \sum_{n''}$ is an absolutely convergent double-sum, i. e. the order of summation may be inverted. Consequently we have, due to (7. 10),

$$\frac{\partial}{\partial t} \sum_{n=0}^{\infty} P_{n n'} = \sum_{n=0}^{\infty} \frac{\partial}{\partial t} P_{n n'} = \sum_{n=0}^{\infty} \sum_{n''=0}^{\infty} A_{n n''} P_{n'' n'} = \sum_{n''=0}^{\infty} \left(\sum_{n=0}^{\infty} A_{n n''}\right) P_{n'' n'} \equiv 0, \tag{8. 18}$$

i. e., due to (7. 4),

$$\sum_{n=0}^{\infty} P_{n n'}(t, s) = \text{const.} = \sum_{n=0}^{\infty} P_{n n'}(s, s) = 1, \tag{8. 19}$$

q. e. d.

As a consequence the last fundamental condition (7. 8) is fulfilled for processes of the type I, IV, V (cf. p. 273 ff.), as Cramér's criterion may be directly applied in these cases. As regards type II there is two possibilities: either the matrix is row or column half-finite. In the latter case it is, however, seen from (7. 9) that the matrix is then also row half-finite (i. e. finite). In the former case the matrix is bounded, i. e. of type I. As regards type III there is also two possibilities: either the matrix is row or column half. In the latter case

[1] We observe that at this place it is essential that the criterion operates with columns and not with rows (cf. p. 275). Consequently the following proof does not apply if A is e. g. a *row* half matrix (cf. p. 273).

Cramér's criterion may be directly applied. *In the former case it fails*, however, notwithstanding the fact that it may still be applied for the proof of the exponentiability of the matrix (if we, namely, only consider the transposed matrix). Even though $\boldsymbol{A}$ is known always to have convergent numerical column sums because, due to (6. 3),

$$\sum_{n=0}^{\infty} |\boldsymbol{A}_{nn'}| = \sum_{n=0}^{\infty} ((\boldsymbol{\Pi} + \boldsymbol{1}) \cdot \boldsymbol{p})_{nn'} = 2\, p(n', t) < \infty, \tag{8.20}$$

already $\boldsymbol{A}^2$ need not have convergent column sums in the case of a row half matrix, as shown by the following example.

Example (8. II).

Let $\boldsymbol{A}$ be the following row half matrix

$$\begin{gathered} \boldsymbol{A}_{pq} = p^{-1-\varepsilon} q^{2\varepsilon} \varDelta(p \geqq q), \quad p, q = 1, 2, 3, \ldots, \quad 0 < \varepsilon < 1, \\ \varDelta(p \geqq q) = \begin{cases} 1 & \text{for } p \geqq q \\ 0 & \text{» } p < q. \end{cases} \end{gathered} \tag{8.21}$$

We then have

$$\sum_{i=1}^{\infty} \boldsymbol{A}_{iq} = \sum_{i=q}^{\infty} i^{-1-\varepsilon} q^{2\varepsilon} \leqq \zeta(1+\varepsilon) \cdot q^{2\varepsilon} < \infty \tag{8.22}$$

and

$$\begin{gathered} (\boldsymbol{A}^2)_{pq} = \sum_{\alpha=1}^{\infty} p^{-1-\varepsilon} \alpha^{2\varepsilon} \alpha^{-1-\varepsilon} q^{2\varepsilon} \varDelta(p \geqq \alpha) \varDelta(\alpha \geqq q) = p^{-1-\varepsilon} q^{2\varepsilon} \sum_{\alpha=q}^{p} \alpha^{-1+\varepsilon} \varDelta(p \geqq q) \geqq \\ p^{-1-\varepsilon} q^{2\varepsilon} (p - q + 1) \varDelta(p \geqq q) p^{-1+\varepsilon}. \end{gathered} \tag{8.23}$$

Consequently

$$\sum_{i=1}^{\infty} (\boldsymbol{A}^2)_{iq} \geqq q^{2\varepsilon} \sum_{i=q}^{\infty} \frac{1}{i} - q^{2\varepsilon}(q-1) \sum_{i=q}^{\infty} \frac{1}{i^2} = \infty, \tag{8.24}$$

q. e. d.

In the general case we can obtain a *sufficient* condition for the fundamental condition (7. 8) in the following way. Integrating the fundamental equation (7. 13), summing over n and rearranging somewhat we obtain for an *arbitrary* process which is assumed to have a solution satisfying (3. 1), (7. 4)—(7. 7), due to (7. 23),

$$0 \leqq 1 - \sum_{n=0}^{N} \boldsymbol{P}_{nn'}(t, s) = \left(\sum_{n=N+1}^{\infty} \sum_{n''=0}^{N} - \sum_{n=0}^{N} \sum_{n''=N+1}^{\infty} \right) \int_s^t (\boldsymbol{\Pi}(\tau) \cdot \boldsymbol{p}(\tau))_{nn''} \boldsymbol{P}_{n''n'}(\tau, s)\, d\tau \tag{8.25}$$

for all $N \geqq n'$.

This equation is simply a continuity equation for the 'probability mass' of the values $n = 0, 1, 2, \ldots, N$, the first and second term on the right hand side being the 'probability mass' which in the time from s to t has flowed 'upwards', respectively 'downwards'. (8.25) shows immediately that we have the following *sufficient* condition for (7.8) to hold true

$$\lim_{N\to\infty} \sum_{n=N+1}^{\infty} \sum_{n''=0}^{N} (\overline{\Pi} \cdot \overline{p})_{nn''} = 0,$$
$$\overline{\Pi} = \max_{s\leq\tau\leq t} \Pi(\tau), \qquad \overline{p} = \max_{s\leq\tau\leq t} p(\tau). \tag{8.26}$$

This condition is, however, not very useful as it will in practice seldom be fulfilled. In the case of A being an arbitrary *semi-diagonal* matrix (cf. p. 273) we have previously proved the following *sufficient* condition, called the Feller-Lundberg condition[1]

$$\sum_{n=0}^{\infty} \frac{1}{\bar{p}(n)} = \infty,$$
$$\bar{p}(n) = \max_{s\leq\tau\leq t} \{p(n', \tau)\} \text{ for } n' = n - l + 1,\ n - l + 2, \ldots, n. \tag{8.27}$$

Furthermore, we have in the case of A being an arbitrary *row half* matrix proved the following *necessary* condition [2]

$$\sum_{n=0}^{\infty} \frac{1}{\underline{p}(n)} = \infty,$$
$$\underline{p}(n) = \min_{s\leq\tau\leq t} p(n, \tau). \tag{8.28}$$

If e.g. we put $p(n, t) = n^2$ (and e.g. $\Pi_{nn'} = \delta_{n, n'+1}$) we see that (8.28) is *not* fulfilled, i. e. that

$$P(\infty, t; n', s) = 1 - \sum_{n=0}^{\infty} P(n, t; n', s) > 0. \tag{8.29}$$

---

[1] Arley (1943), p. 63. This condition and the following one are generalizations of results obtained by Feller and Lundberg, see Lundberg (1940).

[2] Arley (1943), p. 67.

PART III.

'Pathologies' in the Theory of Infinite Systems of Differential Equations.

§ 9.

We shall now return to the general theory of part I, investigating in more detail the conditions (α) and (β) (cf. p. 271—272) by discussing the 'pathological' cases arising when we go from finite to infinite dimensions. Our principle by the construction of these examples will be simply to split the system (2. 5), i. e.

$$\boldsymbol{Y}' = \boldsymbol{A} \cdot \boldsymbol{Y}, \tag{9. 1}$$

into two parts, one containing $Y_1, Y_2, Y_3, \ldots$ and which may be solved successively, the other containing only Y_0' expressed as a series in $Y_1, Y_2, Y_3, \ldots$. It will be seen that for this purpose we need only take the first column of $\boldsymbol{A}$ equal to 0 throughout. Furthermore, we shall as far as possible choose our examples in such a way that (9. 1) represents a stochastic process, i. e. that $\boldsymbol{A}$ is of the type given in (7. 9).

Firstly, we observe that in the equation (9. 1) the dot now represents an *infinite* sum. We cannot, consequently, from the fact that $\boldsymbol{A}$ and $\boldsymbol{Y}$ are continuous functions now conclude that $\boldsymbol{Y}'$ is continuous, as the sum defining $\boldsymbol{Y}'$ need not be uniformly convergent. In fact $\boldsymbol{Y}'$ can even be so discontinuous that it is not absolutely integrable (even in the sense of Lebesgue). Before we give an example of this fact we shall show that we cannot even from (9. 1) conclude to (2. 9), i. e. in the equation

$$|\boldsymbol{Y}'| \leqq \boldsymbol{K} \cdot \boldsymbol{G} \tag{9. 2}$$

the product $\boldsymbol{K} \cdot \boldsymbol{G}$ need not necessarily exist.

Example (9. I). ((α): +. (β): −. un: +. ex: +)[1].

$$\boldsymbol{Y}' = \begin{Bmatrix} 0 & 1 & -2 & 3 & \cdots \\ 0 & -1 & 0 & 0 & \cdots \\ 0 & 0 & -2 & 0 & \cdots \\ 0 & 0 & 0 & -3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{Bmatrix} \cdot \boldsymbol{Y} = \boldsymbol{A} \cdot \boldsymbol{Y}, \qquad \boldsymbol{Y}(x_0) = \begin{Bmatrix} 0 \\ \frac{1}{1^2} \\ \frac{1}{2^2} \\ \frac{1}{3^2} \\ \vdots \end{Bmatrix}. \tag{9. 3}$$

[1] The symbols (α): +, (β): −, un: +, ex: + and so on denote, respectively: the condition (α) is fulfilled, (β) is not, the theorem of uniqueness is fulfilled, that of existence is also, and so on.

It is easily seen that for this equation[1] the condition (α) is fulfilled, A being a half matrix (cf. p. 273). Next the equation is readily seen to have the unique solution

$$Y_n(x) = \frac{1}{n^2} e^{-n(x-x_0)} \quad \text{for} \quad n \geqq 1 \tag{9.4}$$

and

$$Y_0'(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} e^{-n(x-x_0)} = \ln\left(1 + e^{-(x-x_0)}\right) \quad (\mathrm{Re}(x-x_0) \geqq 0), \tag{9.5}$$

i. e.

$$Y_0(x) = \int_{x_0}^{x} Y_0'(x)\, dx = \sum_{n=1}^{\infty} (-1)^{n+1} \cdot \frac{1 - e^{-n(x-x_0)}}{n^2} \quad (\mathrm{Re}(x-x_0) \geqq 0). \tag{9.6}$$

As

$$K = \begin{Bmatrix} 0 & 1 & 2 & 3 \cdots \\ 0 & 1 & 0 & 0 \cdots \\ 0 & 0 & 2 & 0 \cdots \\ 0 & 0 & 0 & 3 \cdots \\ \vdots & \vdots & \vdots & \vdots \end{Bmatrix} \quad \text{and} \quad \underset{(x_0 \leqq x < \infty)}{G} = \begin{Bmatrix} \frac{\pi^2}{12} \\ \frac{1}{1^2} \\ \frac{1}{2^2} \\ \frac{1}{3^2} \\ \vdots \end{Bmatrix}^{2} \tag{9.7}$$

we see in fact that $K \cdot G$ is divergent, q. e. d. Consequently the condition (β) is not fulfilled. *We observe, however, that a pathological case of this type cannot occur under our assumption* (β), *because it then follows that* $K \cdot G$ *is convergent.*

Example (9. II). ((α): +. (β): —. un: +. ex: (+)).

It may be of interest to observe that by omitting the negative signs in the first row of A in the preceding example and multiplying $Y(x_0)$ by $\frac{6}{\pi^2}$ $\left(\text{in order to make } \sum_{n=0}^{\infty} Y_n(x_0) = 1\right)$ we obtain a stochastic process of just the type

[1] We note that due to the alternating signs in the first row of A (9.3) does not represent a stochastic process (cf., however, the following example).

[2] As $Y_0' \geqq 0$, $Y_0(x)$ is monotonously increasing, i. e.

$$\max Y_0(x) = Y_0(\infty) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} = \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} - \sum_{n=1}^{\infty} \frac{1}{(2n)^2} = \frac{\pi^2}{8} - \frac{1}{4}\frac{\pi^2}{6} = \frac{\pi^2}{12},$$

which result follows from the theory of the Riemann ζ-function.

announced above (cf. the end of p. 283) in which Y' does not exist *everywhere*, but only almost everywhere. In fact we now find

$$Y_n(x) = \frac{6}{\pi^2}\frac{1}{n^2}e^{-n(x-x_0)} \quad \text{for} \quad n \geqq 1 \tag{9.8}$$

and instead of (9.5)

$$Y_0'(x) = \frac{6}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n}e^{-n(x-x_0)} = \begin{cases} -\frac{6}{\pi^2}\ln\left(1-e^{-(x-x_0)}\right) & \operatorname{Re}(x-x_0) > 0 \\ \infty & \operatorname{Re}(x-x_0) = 0. \end{cases} \tag{9.9}$$

Nevertheless the function (9.9) is integrable, and integrating we obtain

$$Y_0(x) = \int_{x_0}^{x} Y_0'(x)\,dx = \frac{6}{\pi^2}\sum_{n=1}^{\infty}\frac{1-e^{-n(x-x_0)}}{n^2} \qquad \operatorname{Re}(x-x_0) \geqq 0. \tag{9.10}$$

It is easily seen (a) that the functions given in (9.8) and (9.10) constitute a probability distribution as they are non-negative and have the sum 1, and (b) that they satisfy the equation $\boldsymbol{Y}' = \boldsymbol{A}\cdot\boldsymbol{Y}$ except in the initial point $x = x_0$, because here the rate of change of Y_0 is $+\infty$[1].

Finally we observe that due to the fact that the solution given in (9.8) and (9.10) is only a solution almost everywhere, it is *not* a solution in every limit-point of points in which it is a solution. We shall later (cf. § 13) return to this interesting point.

We shall next give an example showing that Y' need not be absolutely integrable, the integration being taken even in the sense of Lebesgue. (We note that (2.9) then shows that $\boldsymbol{K}\cdot\boldsymbol{G}$ cannot be convergent). In such case we could not, consequently, generally perform our proof of the theorem of uniqueness by the method of iteration applied in § 2.

Example (9. III). ((α): +. (β): −. un: +. ex: +).

Firstly, we consider the well-known function

$$y = x^2 \sin\frac{1}{x^2} \qquad (x \geqq 0) \tag{9.11}$$

[1] It may be interesting to remark that in a stochastic process the rate of change of the probabilities can never assume the value $-\infty$, so long as the intensity function $p(t)$ is finite.

for which

$$y' = \begin{cases} 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2} & x > 0 \\ 0 & x = 0. \end{cases} \tag{9.12}$$

Due to the term $\frac{2}{x}$ the function y', being discontinuous in $x = 0$, is just seen to be non-integrable in the sense of Lebesgue. In fig. 1 we have indicated the graph of y' in the interval $0 < x < \infty$. The function $y'(x)$ has an infinity of zeros, given by the equation

$$\operatorname{tg} \frac{1}{x^2} = \frac{1}{x^2}, \quad \text{i.e.} \quad x_p \sim \left(\frac{\pi}{2} + p\pi\right)^{-\frac{1}{2}}, \quad p = 1, 2, 3, \ldots \tag{9.13}$$

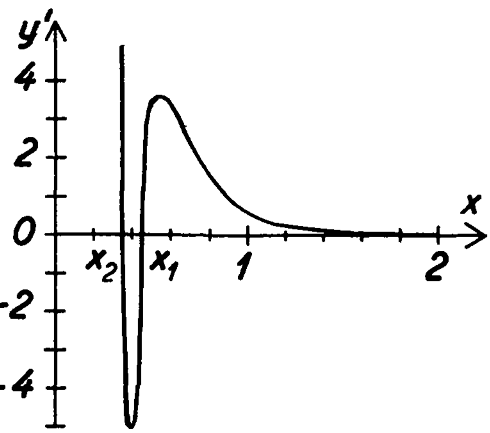

Fig. 1.

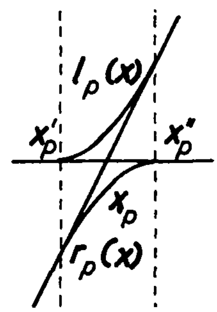

Fig. 2.

In the neighbourhood of each of these zeros we replace the curve of $y'(x)$ by two suitable, monotonous, smooth curves $y = l_p(x)$ (left hand curves) and $y = r_p(x)$ (right hand curves), as indicated in fig. 2:

$$\left.\begin{array}{l} y = l_p(x) \\ y = r_p(x) \end{array}\right\} \quad \text{for} \quad x'_p \leqq x \leqq x''_p. \tag{9.14}$$

Here

$$0 < \cdots < x'_p < x_p < x''_p < \cdots < x'_1 < x_1 < x''_1$$

$$l_p(x'_p) = r_p(x''_p) = 0 \tag{9.15}$$

$$l_p(x) + r_p(x) = y'(x).$$

This replacement has to be done in such a way that the following functions are differentiable with continuous derivatives in the whole interval $0 \leqq x < \infty$:

$$
z_1(x) = \begin{cases} 0 & \text{for } 0 \leqq x \leqq x_1' \\ l_1(x) & \text{»} \quad x_1' \leqq x \leqq x_1'' \\ 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2} & \text{»} \quad x_1'' \leqq x \end{cases}
$$

$$
z_p(x) = \begin{cases} 0 & \text{»} \quad 0 \leqq x \leqq x_p' \\ l_p(x) & \text{»} \quad x_p' \leqq x \leqq x_p'' \\ 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2} & \text{»} \quad x_p'' \leqq x \leqq x_{p-1}' \qquad p = 2, 3, 4, \ldots \\ r_p(x) & \text{»} \quad x_{p-1}' \leqq x \leqq x_{p-1}'' \\ 0 & \text{»} \quad x_{p-1}'' \leqq x. \end{cases} \tag{9.16}
$$

Finally we put

$$
y_n(x) = z_n(x) + (-1)^{n+1} \frac{1}{2^n}, \qquad n = 1, 2, 3, \ldots \tag{9.17}
$$

and define a matrix $A(x)$ with the following elements:

$$
\begin{aligned}
A_{nn}(x) &= \frac{1}{y_n} \frac{d y_n}{d x} \quad \text{for } n = 1, 2, 3, \ldots \\
A_{0n}(x) &= 1 \qquad\quad \text{»} \quad n = 1, 2, 3, \ldots \\
A_{nm}(x) &= 0 \qquad\quad \text{» all other values of } n, m = 0, 1, 2, \ldots .
\end{aligned} \tag{9.18}
$$

We now consider the equation

$$
Y'(x) = A(x) \cdot Y(x), \tag{9.19}
$$

where $A(x)$ is given in (9.18). $A(x)$ is thus continuous in $0 \leqq x < \infty$ and is, being a half matrix, furthermore exponentiable in this interval (cf. p. 273). The condition (α) is thus satisfied. Corresponding to the initial value

$$
Y(0) = \begin{Bmatrix} 0 \\ \frac{1}{2} \\ -\frac{1}{4} \\ \frac{1}{8} \\ \vdots \end{Bmatrix} \tag{9.20}
$$

(9.19) has thus the unique solution

$$Y_n(x) = y_n(x) \text{ given in (9.17)}$$

and, due to (9.15),

$$Y_0'(x) = \sum_{n=1}^{\infty} Y_n(x) = \begin{cases} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{2^n} = \frac{1}{3} & \text{for } x = 0 \\ 2x \sin\frac{1}{x^2} - \frac{2}{x}\cos\frac{1}{x^2} + \frac{1}{3} & \text{» } x > 0, \end{cases} \tag{9.21}$$

i. e., due to (9.20),

$$Y_0(x) = x^2 \sin\frac{1}{x^2} + \frac{1}{3}x. \tag{9.22}$$

(9.21), consequently, shows that we have in fact obtained an equation (9.19) for which Y' is not absolutely integrable, q. e. d.

Finally we observe that our example could, obviously, just as well have been constructed in such a way that the matrix A becomes not only continuous — as in our example — but furthermore differentiable an arbitrary, but finite, number of times.

We stress, however, that a pathological case of this type is excluded in case our second condition (β) *is fulfilled.* From (β) it follows, namely, that $K \cdot G$ is convergent, i. e. that all the series defining $A \cdot Y$ are uniformly — and absolutely — convergent. A and Y being continuous, Y' is, therefore, also continuous i. e. absolutely integrable and satisfies, furthermore, $\int_{x_0}^{x} Y'\,dx = Y(x) - Y(x_0)$, which relation was the starting point for the proof of the theorem of existence (cf. § 3).

§ 10.

In this paragraph we shall show that our conditions (α) and (β) are no *necessary* conditions for the theorem of *uniqueness* to hold true.

Example (10.1). ((α): —. un: +. ex: +).

$$Y' = \begin{Bmatrix} 0 & \frac{1}{1\cdot 2}\lambda & 2\lambda & 3\lambda & \cdots \\ 0 & -1\lambda & 0 & 0 & \cdots \\ 0 & \frac{1}{2\cdot 3}\lambda & -2\lambda & 0 & \cdots \\ 0 & \frac{1}{3\cdot 4}\lambda & 0 & -3\lambda & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{Bmatrix} \cdot Y = A \cdot Y, \quad \lambda > 0. \tag{10.1}$$

It will be seen that A satisfies (7.9) and (7.10) and that (10.1), therefore, represents a stochastic process.

Since

$$(A^2)_{01} = \lambda^2\left(-\frac{1}{2}+\sum_{n=3}^{\infty}\frac{1}{n}\right) = \infty, \tag{10.2}$$

already A^2 does not exist, i. e. our first condition (α) is not fulfilled. Nevertheless, it is easily seen that (10.1) admits of at most one solution (in § 11 we shall show that it *has* in fact a solution) because (10.1) simply means

$$\begin{aligned} Y_1' &= -\lambda Y_1 \\ Y_n' &= -n\lambda Y_n + \frac{\lambda}{n(n+1)}Y_1, \quad n=2,3,4,\ldots \end{aligned} \tag{10.3}$$

and

$$Y_0' = \frac{\lambda}{1\cdot 2}Y_1 + \lambda\sum_{n=2}^{\infty} n Y_n.$$

We next show that nor our second condition (β) is necessary for the theorem of uniqueness to hold true. For this purpose we shall utilize the theory of Fourier series.

Example (10.II). ((α): $+$. (β): $-$. un: $+$. ex: $\underset{(-)}{+}$).

$$Y' = \begin{Bmatrix} Y_0' \\ Y_1' \\ Y_2' \\ Y_3' \\ Y_4' \\ Y_5' \\ \vdots \end{Bmatrix} = \begin{Bmatrix} 0 & a_0 & a_1 & a_{-1} & a_2 & a_{-2} & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & i & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & -i & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 2i & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & -2i & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \end{Bmatrix} \cdot Y = A\cdot Y. \tag{10.4}$$

In this case (α) is fulfilled, because A is simply a half matrix (cf. p. 273). In fact we find

$$\exp\,[K|x-x_0|] = \tag{10.5}$$

$$= \begin{Bmatrix} 1, & |a_0|\,|x-x_0|, & |a_1|(e^{|x-x_0|}-1), & |a_{-1}|(e^{|x-x_0|}-1), & \frac{|a_2|}{2}(e^{2|x-x_0|}-1), & \frac{|a_{-2}|}{2}(e^{2|x-x_0|}-1) & \cdots \\ 0 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & e^{|x-x_0|} & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & e^{|x-x_0|} & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & e^{2|x-x_0|} & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & e^{2|x-x_0|} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \end{Bmatrix}.$$

Our second condition (β) need, however, not be fulfilled. If we e.g. as initial condition choose

$$Y(x_0) = C = \begin{Bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ \vdots \end{Bmatrix} \tag{10.6}$$

and e.g. put

$$a_0 = 0, \quad a_\nu = \frac{1}{|\nu|^2}, \quad \nu = \pm 1, \ \pm 2, \ldots, \tag{10.7}$$

we see that

$$(\exp[K|x - x_0|] \cdot C)_0 = 2 \sum_{\nu=1}^{\infty} \frac{1}{\nu^3}(e^{\nu |x - x_0|} - 1) = \infty. \tag{10.8}$$

Nevertheless, it is again easily seen that (10.4) has at most one solution[1] since (10.4) simply means

$$\begin{aligned} Y'_{2n+1} &= -ni Y_{2n+1}, \qquad n = 0, 1, 2, \ldots, \\ Y'_{2n} &= \quad ni Y_{2n}, \qquad n = 1, 2, 3, \ldots, \\ Y'_0 &= \quad a_0 Y_1 + a_1 Y_2 + a_{-1} Y_3 + \cdots. \end{aligned} \tag{10.9}$$

In § 11 we shall consider this example in more detail.

Finally we show that in the case of our equation $Y' = A \cdot Y$ being of infinite dimension, the theorem of uniqueness itself need not be true in contrast to the case of finite dimensions.

Example (10. III). ((α): + . (β): —. un: —. ex: $\genfrac{}{}{0pt}{}{+}{(-)}$).

$$Y' = \begin{Bmatrix} 0 & 1 & 0 & 0 \cdots \\ 0 & 0 & 1 & 0 \cdots \\ 0 & 0 & 0 & 1 \cdots \\ 0 & 0 & 0 & 0 \cdots \\ \vdots & \vdots & \vdots & \vdots \end{Bmatrix} \cdot Y = A \cdot Y. \tag{10.10}$$

---

[1] We see, however, that (10.4) need not have any solutions at all. By choosing $a_0, a_1, a_{-1}, \ldots$ in a suitable way, Y'_0 given in (10.9) need, namely, not exist. Even if it exists, it may not be integrable. E.g. we may choose $a_0, a_1, a_{-1}, \ldots$ so that $Y'_0 = 1$ for $x > x_0$, $Y'_0 = -1$ for $x < x_0$ and $Y'_0 = 0$ for $x = 0$. In this case Y'_0 cannot be the derivative of any function.

This equation means, simply,

$$\begin{aligned} &Y_1 = Y_0' \\ &Y_2 = Y_1' = Y_0'' \\ &\dots \\ &Y_n = Y_{n-1}' = Y_0^{(n)} \\ &\dots . \end{aligned} \tag{10.11}$$

We thus see that $Y_n(x)$ is uniquely determined by $Y_0(x)$ and that $Y_0(x)$ may be an arbitrary function having, only, derivatives of arbitrary high order.

From the initial condition

$$\boldsymbol{Y}(x_0) = \boldsymbol{C} \tag{10.12}$$

and (10.11) it follows that

$$Y_0^{(n)}(x_0) = C_n. \tag{10.13}$$

As is well-known the function $Y_0(x)$ is, however, not uniquely determined by (10.13), i. e. by its Taylor series. If, namely, Y_0 is a function for which (10.13) is fulfilled, e. g. the function

$$Y_0 = \sum_{\nu=0}^{\infty} C_\nu \frac{(x-x_0)^\nu}{\nu!}, \tag{10.14}$$

then for instance the function

$$Y_0^* = Y_0 + k \exp\left[-\frac{1}{(x-x_0)^2}\right] \tag{10.15}$$

will for arbitrary values of the constant k also satisfy (10.13). *The equation* (10.10) *has thus, corresponding to an arbitrary initial condition* (10.12) *for which* (10.14) *has a non-vanishing radius of convergence, an infinity of solutions.*[1]

We note the very important fact, that this property is characteristic for a whole class of *infinite* equations $\boldsymbol{Y}' = \boldsymbol{A} \cdot \boldsymbol{Y}$. By a row semi-finite matrix we understand an infinite matrix which has in each row only a *finite* number of non-vanishing elements, but not necessarily the *same* number for different rows. Let us by ${}_pn$ denote the maximum *column* index in the p'th *row* of $\boldsymbol{A}$, i. e.

$$A_{pj}\begin{cases} \neq 0 & \text{for } j = {}_pn \\ = 0 & \text{» all } j > {}_pn. \end{cases} \tag{10.16}$$

[1] The equation (10.10) need not, however, have any solutions at all, as (10.14) may only be convergent for $x = x_0$, e. g. if $C_\nu = (\nu!)^2$.

If especially ${}_pn$ satisfies

$$ {}_pn > p \quad \text{for all} \quad p = 0, 1, 2, \ldots \text{[1]} \tag{10.17} $$

we obtain the class just mentioned.[2]

Each of the single equations in $\boldsymbol{Y}' = \boldsymbol{A} \cdot \boldsymbol{Y}$ we may, namely, under the conditions (10.16) and (10.17) solve with respect to $Y_{{}_pn}(x)$, thus obtaining

$$ Y_{{}_pn}(x) = \frac{1}{A_{p,\,{}_pn}(x)} \left[Y_p'(x) - \sum_{j=0}^{{}_pn-1} A_{pj}(x)\, Y_j(x) \right] \quad \text{for all } p = 0, 1, 2, \ldots. \tag{10.18} $$

This equation means that each Y_p may either be chosen as an arbitrary function, having, only, derivatives of arbitrary high order, or may be determined uniquely and successively from the lower Y_p-functions, i. e. those with lower index p. If the equation $\boldsymbol{Y}' = \boldsymbol{A} \cdot \boldsymbol{Y}$ has any solution at all, $\boldsymbol{Y}$, satisfying the initial condition (10.12), it will thus be seen that making again the substitution (10.15) we may obtain a different solution $\boldsymbol{Y}^*$ satisfying the same initial condition (10.12). *Even for such simple equations with row semi-finite matrices satisfying* (10.17) *it is, consequently, necessary to impose certain restrictions — e. g. our condition* (β) *— on the solutions considered in order to maintain a theorem of uniqueness.* This fact is especially interesting because in the practical statistical applications of stochastic processes we often meet with processes governed by equations of just this type.

Example (10.IV). ((α): +. (β): ±. un: —. ex: +).

Let us as an example consider a stochastic process[3] with the intensity function $p(n, t) = n$ and the relative transition probability matrix Π having only the non-vanishing elements

$$ \Pi_{n'+1,\,n'} = \frac{\lambda}{\lambda + \gamma} \quad \text{and} \quad \Pi_{n'-1,\,n'} = \frac{\gamma}{\lambda + \gamma}, \ \text{i. e.} $$

[1] We observe that a row semi-finite matrix of the type (10.17) need not be absolutely exponentiable, cf. (5.8) and (5.9).

[2] In the other extreme case, ${}_pn \leqq p$ for all $p = 0, 1, 2, \ldots$, $\boldsymbol{A}$ is simply a row half matrix, and our first condition (α) is thus fulfilled (cf. p. 273). The equation $\boldsymbol{Y}' = \boldsymbol{A} \cdot \boldsymbol{Y}$ is, therefore, in this case covered by our theory. Furthermore, it will be seen from the proof of the theorem of uniqueness (§ 2) that in this case our second condition (β) is automatically fulfilled, as the exponential will again be a row half matrix. *Under this condition the theorems of uniqueness and existence thus both hold generally true without any further conditions,* i. e. the equation behaves exactly as a *finite* equation.

[3] Cf. Arley (1943), §§ 4.6—4.8.

$$P' = \begin{Vmatrix} 0 & \gamma\cdot 1 & 0 & 0 & \cdots \\ 0 & -(\lambda+\gamma)\cdot 1 & \gamma\cdot 2 & 0 & \cdots \\ 0 & \lambda\cdot 1 & -(\lambda+\gamma)\cdot 2 & \gamma\cdot 3 & \cdots \\ 0 & 0 & \lambda\cdot 2 & -(\lambda+\gamma)\cdot 3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{Vmatrix} \cdot P = A\cdot P, \quad P_n(0)=\delta_{n,1}. \tag{10.19}$$

By means of the generating function it is possible to deduce the following exact probability solution[1]

$$\begin{aligned} P_0(t) &= 1 + \frac{\gamma-\lambda}{\lambda-\gamma\exp[(\gamma-\lambda)t]} \\ P_n(t) &= \frac{(\gamma-\lambda)^2}{\lambda^2}\exp[(\gamma-\lambda)t]\frac{(1-\exp[(\gamma-\lambda)t])^{n-1}}{\left(1-\frac{\gamma}{\lambda}\exp[(\gamma-\lambda)t]\right)^{n+1}} \qquad n=1,2,3,\ldots. \end{aligned} \tag{10.20}$$

It is easily seen that (10.19) has apart from the solution (10.20), which is easily seen to satisfy (7.7) and (7.8), an infinity of solutions satisfying the same initial condition $P_n(0)=\delta_{n,1}$. In fact we find, putting

$$P_0^*(t) = P_0(t) + k\exp\left[-\frac{1}{t^2}\right], \tag{10.21}$$

successively that

$$\begin{aligned} P_1^*(t) &= P_1(t) + \frac{k}{\gamma}\exp\left[-\frac{1}{t^2}\right]\frac{2}{t^3} \\ P_2^*(t) &= P_2(t) + \frac{k}{\gamma^2 t^6}\exp\left[-\frac{1}{t^2}\right]((\lambda+\gamma)t^3 - 3t^2 + 2) \end{aligned} \tag{10.22}$$

and so on.

The solution given in (10.21) and (10.22) is, however, no probability solution because even if $k>0$ we may obtain *negative* values for some of the $P_n^*(t)$ functions at various times t. If e.g. $\lambda=\gamma=\frac{1}{2}$ it is seen that the second term of $P_2^*(t)$ is negative in the interval $1<t<1+\sqrt{2}$. The first term being at most 1, we thus obtain $P_2^*(t)<0$ in an interval $1+\varepsilon\leqq t\leqq 1+\sqrt{2}-\varepsilon$ for sufficiently high values of k. (We note that in this example it may be shown that *no* solution other than (10.20) can be non-negative throughout (cf. p. 280)).

[1] This calculation is due to tekn. d:r. Conny Palm. We wish to express our most sincere thanks to Dr. Palm for communicating this solution to us. We have, however, later succeeded in obtaining the solution by a much simpler method, which may even be generalized to processes for which the method of the generating function cannot be carried through. We intend to discuss this method in another paper.

§ 11.

In the preceding paragraph we have discussed pathologies concerning the theorem of *uniqueness*. We shall now discuss pathologies concerning the theorem of *existence*. First we give an example showing that although the condition (α) is not fulfilled, the theorem of existence may nevertheless hold true.

Example (11. I). (= ex. (10. I). (α): —. un: +. ex: +).

Let us again consider the example (10. I). As shown there the equation (10. 1) admits of only one solution obtained by solving successively the equations (10. 3). We thus find, corresponding to the special initial value

$$Y(x_0) = \{Y_{nn'}(x_0)\} = \{\delta_{nn'}\}, \qquad n' = 0, 1, 2, \ldots, \tag{11.1}$$

$$\begin{aligned} Y_{1n'}(x) &= \delta_{1n'} e^{-\lambda(x-x_0)} \\ Y_{nn'}(x) &= \delta_{nn'} e^{-n\lambda(x-x_0)} + \frac{\delta_{1n'}}{n(n+1)(n-1)}\left(e^{-\lambda(x-x_0)} - e^{-n\lambda(x-x_0)}\right) \quad (n > 1) \end{aligned} \tag{11.2}$$

and

$$\begin{aligned} Y'_{0n'}(x) &= \frac{\lambda}{1\cdot 2} Y_{1n'} + \lambda \sum_{n=2}^{\infty} n\, Y_{nn'} = \frac{\lambda}{1\cdot 2}\delta_{1n'} e^{-\lambda(x-x_0)} + \lambda\sum_{n=2}^{\infty} n\,\delta_{nn'} e^{-n\lambda(x-x_0)} + \\ &\delta_{1n'} e^{-\lambda(x-x_0)} \frac{\lambda}{2}\sum_{n=2}^{\infty}\left(\frac{1}{n-1} - \frac{1}{n+1}\right) - \delta_{1n'}\frac{\lambda}{2}\sum_{n=2}^{\infty}\left(\frac{e^{-n\lambda(x-x_0)}}{n-1} - \frac{e^{-n\lambda(x-x_0)}}{n+1}\right) = \\ &\lambda n' e^{-n'\lambda(x-x_0)} \Delta(n' \geqq 2) + \\ &\quad + \lambda\delta_{1n'}\left[e^{-\lambda(x-x_0)} - \frac{1}{2} - \frac{1}{2}\left(e^{\lambda(x-x_0)} - e^{-\lambda(x-x_0)}\right)\ln\left(1 - e^{-\lambda(x-x_0)}\right)\right], \end{aligned} \tag{11.3}$$

$$\operatorname{Re}\lambda(x-x_0) \geqq 0, \quad \Delta(n' \geqq 2) = \begin{cases} 1 \text{ for } n' \geqq 2 \\ 0 \;\;\text{»}\;\; n' < 2. \end{cases}$$

The last series in (11. 3) is seen to be uniformly convergent, thus having a continuous sum for $\operatorname{Re}\lambda(x-x_0) \geqq 0$, but to be divergent for $\operatorname{Re}\lambda(x-x_0) < 0$. The point $x = x_0$ is thus a singular point as also shown by the result of the summation containing a term $\ln(1 - e^{-\lambda(x-x_0)})$. This fact means that although the ma trix A is so regular as we may demand, i. e. constant, one of the functions, $Y_{0n'}(t)$, does not even exist everywhere, but only in the complex half-plane $\operatorname{Re}\lambda(x-x_0) \geqq 0$ and is, consequently, not analytic in $x = x_0$.

We thus have the very interesting fact that the singular points of an infinite equation need not at all coincide with those of the matrix A of the equation and *may, not, in fact, even be read off from A by an immediate inspection. This fact shows in a most striking way the fundamental difference between infinite systems and finite systems, which last systems can only have the singular points of A as singular points (cf. § 4).*

As $Y'_{0n'}(x)$ given in (11.3) is a continuous function for $\operatorname{Re}\lambda(x-x_0)\geqq 0$, it may be integrated and we thus find that $Y_{0n'}(x)$ is uniquely given by (substituting $t=e^{-\lambda(x-x_0)}$)

$$\begin{aligned} Y_{0n'}(x) &= \delta_{0n'} + \int_0^x Y'_{0n'}(x)\,dx = \delta_{0n'} + \Delta(n'\geqq 2)\left(1-e^{-n'\lambda(x-x_0)}\right) + \\ &\quad \delta_{1n'}\left[\left(1-e^{-\lambda(x-x_0)}\right)-\frac{\lambda(x-x_0)}{2}-\frac{1}{2}\int_{e^{-\lambda(x-x_0)}}^{1}\left(\frac{\ln(1-t)}{t^2}-\ln(1-t)\right)dt\right] = \\ &\delta_{0n'} + \Delta(n'\geqq 2)\left(1-e^{-n'\lambda(x-x_0)}\right) + \\ &\quad + \delta_{1n'}\left[\frac{1}{2}\left(1-e^{-\lambda(x-x_0)}\right)-\frac{1}{2}\left(1-e^{-\lambda(x-x_0)}\right)^2 e^{\lambda(x-x_0)}\ln\left(1-e^{-\lambda(x-x_0)}\right)\right], \\ &\qquad \operatorname{Re}\lambda(x-x_0)\geqq 0. \end{aligned} \tag{11.4}$$

We see that $Y_{0n'}(x)$ has for $x=x_0$ a very serious singularity of the same type as the function $y=x^2\ln x$, namely an infinite branch-point.

It will be seen that all the functions in (11.2) and (11.4) are non-negative and satisfy

$$\sum_{n=0}^{\infty} Y_{nn'}(x)\equiv 1 \text{ for all } x\geqq x_0 \text{ and } n'=0,1,2,\ldots, \tag{11.5}$$

as should be the case because our equation represents a stochastic process.

The column matrices $Y_n(x)=Y_{nn'}(x-x_0)$ corresponding to $n'=0,1,2,\ldots$ are now seen to constitute a fundamental solution. Taking together these columns to form a quadratic matrix, which we shall again denote as a product-integral

$$\underset{x_0}{\overset{x}{\mathcal{P}}}(\mathbf{1}+\boldsymbol{A}\,dt)=\{Y_{nn'}(x-x_0)\},\qquad x\geqq x_0, \tag{11.6}$$

we see, namely, that our equation has for an arbitrary initial value

$$\boldsymbol{Y}(x_0)=\boldsymbol{C}=\{C_{n'}\} \text{ with } \sum_{n'=0}^{\infty}|C_{n'}|<\infty \tag{11.7}$$

a unique solution for all $x\geqq x_0$ which is given by

$$\boldsymbol{Y}(x)=\underset{x_0}{\overset{x}{\mathcal{P}}}(\mathbf{1}+\boldsymbol{A}\,dt)\cdot\boldsymbol{C}, \tag{11.8}$$

i. e. formally identical with (3.12). In our case (11.7) will always be satisfied as $Y(x_0)$ denotes the absolute probabilities at the time $x = x_0$. Due to (11.1), (11.5) and (11.7) the matrix (11.6) is seen to have both the properties (3.11) and (3.14) when $x_0 \leqq x_1 \leqq x$. As (11.6) exists only for $x \geqq x_0$ it will, however, not have the property (3.16) i. e. a reciprocal, in accordance with the fact that in all probability problems the time variable will move only in the forward direction.

As $Y(x)$ is not analytic for $x = x_0$ it can, consequently, not be expanded in a power series in $x - x_0$ from the initial value x_0. Nevertheless, as $Y(x)$ is differentiable to the right also for $x = x_0$, Y being a solution of a differential equation, the matrix (11.6) is seen also to satisfy (4.1), i.e. we may for small values of $x - x_0$ obtain a good approximation by putting

$$Y(x) = (1 + A\,(x - x_0)) \cdot Y(x_0) \qquad (|x - x_0| \ll 1) \qquad (11.9)$$

i. e. using the first terms in the power series (4.6), *in spite of the fact that this series itself is divergent, as already* A^2 *does not exist* (cf. (10.2)).

We may, furthermore, make the interesting observation that in spite of the divergence of A^2 and, consequently, of all higher powers of A, we can, nevertheless, *in this example* obtain the power series for any *other* initial point by the usual method. As $Y(x)$ given in (11.2) and (11.4) is analytic for $\mathrm{Re}\,\lambda(x - x_0) > 0$, it may be expanded in a power series from any point x_1 in this region

$$Y(x) = \sum_{\nu=0}^{\infty} Y_\nu \frac{(x - x_1)^\nu}{\nu!}, \qquad Y_0 = Y(x_1) \quad (\mathrm{Re}\, x_1 > \mathrm{Re}\, x_0, \;\; \mathrm{Re}\, x > \mathrm{Re}\, x_0). \quad (11.10)$$

Introducing (11.10) into $Y' = A \cdot Y$ we obtain, due to the convergence of $A \cdot Y$ being uniform,

$$Y' = \sum_{\nu=1}^{\infty} Y_\nu \frac{(x - x_1)^{\nu-1}}{(\nu - 1)!} = A \cdot Y = \sum_{\nu=0}^{\infty} A \cdot Y_\nu \frac{(x - x_1)^\nu}{\nu!}. \qquad (11.11)$$

Equating corresponding coefficients we next obtain the well-known expressions

$$Y_1 = A \cdot Y_0, \quad Y_2 = A \cdot Y_1 = A \cdot (A \cdot Y_0) \neq A^2 \cdot Y_0, \quad \text{and so on}, \qquad (11.12)$$

in which the associative rule may *not* be applied due to the divergence of A^ν, $\nu = 2, 3, 4, \ldots$. The relations (11.12) do not, however, hold generally true — except the first one $Y_1 = A \cdot Y_0$ — because the inversion of the order of summation in (11.11) is not legitimate in general. If it were legitimate it would, namely, implicate that the sum of the infinite series represented by $A \cdot Y$ could be differentiated arbitrarily often term by term, but as is well-known, this is in general not legitimate even if the sum of the series is an analytic function — as in our case (cf. ex. 12. I).

In contrast to the case of our equation $Y' = A \cdot Y$ being of a *finite* dimension we shall now show that the theorem of existence proper may fail in the case of infinite dimension.

Example (11. II). ((α): —. un: +. ex: $(\bar{+})$).

$$Y' = \begin{Vmatrix} 0 & 1 & 1 & 1 \cdots \\ 1 & 0 & 0 & 0 \cdots \\ 1 & 0 & 0 & 0 \cdots \\ 1 & 0 & 0 & 0 \cdots \\ \vdots & \vdots & \vdots & \vdots \end{Vmatrix} \cdot Y = A \cdot Y. \tag{11.13}$$

Obviously this A is not exponentiable as already A^2 is divergent since

$$(A^2)_{00} = 1 + 1 + 1 + \cdots = \infty. \tag{11.14}$$

Next it is seen that (11. 13) admits of at most one solution given by, due to Y_0 being continuous,

$$Y_n' = Y_0 \quad \text{for} \quad n \geqq 1,$$
$$\text{i. e.} \ Y_n = \int_{x_0}^{x} Y_0 \, dx + C_n \tag{11.15}$$

and

$$Y_0 = \int_{x_0}^{x} \sum_{i=1}^{\infty} Y_i \, dx + C_0 = \int_{x_0}^{x} \left(\sum_{i=1}^{\infty} \int_{x_0}^{x} Y_0 dx \right) dx + \left(\sum_{i=1}^{\infty} C_i \right) (x - x_0) + C_0. \tag{11.16}$$

(11. 16) shows that for all initial values not satisfying the very special conditions

$$C_0 = 0, \quad \sum_{i=1}^{\infty} C_i = 0, \tag{11.17}$$

our equation (11. 13) *has no solution at all.* For the only allowed initial condition given in (11. 17) we find, however, the trivial solution

$$Y = C \text{ for all } x. \tag{11.18}$$

In the two examples hitherto discussed in this paragraph our first condition (α) has not been fulfilled. We shall now discuss examples in which our condition (α) is fulfilled, but our second condition (β), i. e.

$$\exp\,[\boldsymbol{K}\,|x-x_0|]\cdot\boldsymbol{G}<\infty, \tag{11.19}$$

is not.

We first give an example of this kind, in which the theorem of existence does hold true.

Example (11. III). ($\sim$ ex. (10. II). (α): $+$. (β): $-$. un: $+$. ex: $+$).

We again consider the example (10. II), i. e. equation (10. 4). Let us now put $x_0=0$,

$$\boldsymbol{Y}(0)=\boldsymbol{C}=\begin{Bmatrix}0\\1\\1\\1\\\vdots\end{Bmatrix} \tag{11.20}$$

and

$$a_0=\frac{\pi}{2},\qquad a_{\pm 2(\nu+1)}=0,\qquad a_{\pm(2\nu+1)}=-\frac{2}{\pi}\frac{1}{(2\nu+1)^2}. \tag{11.21}$$

As is well-known from the theory of Fourier series, our equation has thus in fact a — unique — solution corresponding to the initial value (11.20), viz.

$$\begin{aligned}
&Y_0=\int_0^x |x|\,dx=\begin{cases}\dfrac{x^2}{2} & \text{for}\quad 0\leqq x<\pi\\[2mm] -\dfrac{x^2}{2} & \text{»}\quad -\pi<x\leqq 0\\[2mm] 0 & \text{»}\quad x=\pm\pi\end{cases}\\
&\left.\begin{aligned}Y_1&\equiv 1\\ Y_{2n+1}&=e^{-inx}\\ Y_{2n}&=e^{inx}\end{aligned}\right\}-\pi\leqq x\leqq\pi,\qquad n=1,2,3,\ldots.
\end{aligned} \tag{11.22}$$

In fact it is seen that (β) is in this case *not* fulfilled, since, due to (10. 5) and $x_0=0$,

$$(\exp\,[\boldsymbol{K}\,|x|]\cdot\boldsymbol{G})_0\geqq(\exp\,[\boldsymbol{K}\,|x|]\cdot\boldsymbol{C})_0=\frac{\pi}{2}|x|+\frac{4}{\pi}\sum_{\nu=0}^{\infty}\frac{1}{(2\nu+1)^3}\left(e^{(2\nu+1)|x|}-1\right)=\infty. \tag{11.23}$$

We note the interesting fact that our equation given in (10. 4) and (11. 21) has only solutions for *real* values of x, as the series defining Y_0' is divergent for non-real values of x.

Thus every point of the complex plane except the real axis is a singular point of the equation in spite of the fact that A is constant, i. e. analytic in the whole complex plane.

Finally we consider an equation for which the condition (α) is satisfied, and the theorem of existence does not hold true.

Example (11. IV). ((α): +. (β): $\overline{(+)}$. un: +. ex: $\overline{(+)}$).

$$Y' = \begin{Bmatrix} 0 & a_1 & a_2 & a_3 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{Bmatrix} \cdot Y = A \cdot Y. \tag{11.24}$$

As A is a half matrix, (α) is in fact fulfilled with

$$\exp\,[K\,|x-x_0|] = \begin{Bmatrix} 1 & |a_1|\,(e^{|x-x_0|}-1) & |a_2|\,(e^{|x-x_0|}-1) & \cdots \\ 0 & e^{|x-x_0|} & 0 & \cdots \\ 0 & 0 & e^{|x-x_0|} & \cdots \\ \vdots & \vdots & \vdots & \end{Bmatrix}. \tag{11.25}$$

Obviously

$$\begin{aligned} Y_n(x) &= C_n e^{x-x_0}, \qquad n \geqq 1, \\ Y_0(x) &= C_0 + (e^{x-x_0}-1) \sum_{n=1}^{\infty} a_n C_n. \end{aligned} \tag{11.26}$$

Thus we see that if the initial point C does not satisfy the condition that the series

$$\sum_{n=1}^{\infty} a_n C_n \text{ is convergent,} \tag{11.27}$$

our equation (11.24) *has no solution at all.*

§ 12.

In fact, we have now given examples of all types of pathologies — in respect to our conditions (α) and (β) — which may arise in the case of *infinite* systems of simultaneous linear differential equations in contrast to the case of finite systems. There remain, however, still a few questions which it may be interesting to discuss. Firstly, we shall remind of another proof of the theorem of uniqueness of the finite equation (2. 1), i. e. $\boldsymbol{Y}'=\boldsymbol{A}\cdot\boldsymbol{Y}+\boldsymbol{B}$, in case $\boldsymbol{A}$ and $\boldsymbol{B}$ are assumed to be analytic in a certain region. If there were two analytic solutions $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ of (2. 1) for the same initial condition (2. 3), then $\boldsymbol{Y}=\boldsymbol{Y}_1-\boldsymbol{Y}_2$ would be a solution of the homogeneous equation (2. 5), i. e. $\boldsymbol{Y}'=\boldsymbol{A}\cdot\boldsymbol{Y}$, satisfying the initial condition (2. 6), i. e. $\boldsymbol{Y}(x_0)=\boldsymbol{0}$. $\boldsymbol{A}$ and $\boldsymbol{Y}$ being analytic and the system being *finite*, we may differentiate the series in $\boldsymbol{Y}'=\boldsymbol{A}\cdot\boldsymbol{Y}$ term by term arbitrarily often, thus obtaining, due to $\boldsymbol{Y}(x_0)=\boldsymbol{0}$,

$$\begin{aligned} \boldsymbol{Y}'(x_0) &= \boldsymbol{A}(x_0)\cdot\boldsymbol{Y}(x_0)=\boldsymbol{0} \\ \boldsymbol{Y}''(x_0) &= \boldsymbol{A}'(x_0)\cdot\boldsymbol{Y}(x_0)+\boldsymbol{A}(x_0)\cdot\boldsymbol{Y}'(x_0)=\boldsymbol{0} \\ &\text{and so on.} \end{aligned} \tag{12. 1}$$

From (12. 1) and Taylor's theorem it then follows that

$$\boldsymbol{Y}(x)\equiv\boldsymbol{0} \tag{12. 2}$$

q. e. d., i. e. $\boldsymbol{Y}'=\boldsymbol{A}\cdot\boldsymbol{Y}+\boldsymbol{B}$ has at most one *analytic* solution.

This proof may, however, not be generalized to the case of *infinite* systems. Firstly, such systems need, namely, not at all admit of solutions for complex values of x or even of solutions being only real-analytic[1], even if $\boldsymbol{A}$ is constant (cf. ex. (11. III)).

Secondly, the equation

$$\boldsymbol{Y}''=\frac{d}{dx}(\boldsymbol{A}\cdot\boldsymbol{Y})=\boldsymbol{A}'\cdot\boldsymbol{Y}+\boldsymbol{A}\cdot\boldsymbol{Y}' \tag{12. 3}$$

and its analogues in (12. 1) need not at all hold true in the case of *infinite* systems. The dot represents, namely, in this case an infinite sum, and the process of differentiating term by term need not, consequently, always be legitimate, as shown by the following example.

Example (12. I). (=ex. (11. III). (α): + . (β): —. un: + . ex: + . $Y_0''(0)$ does not exist).
We again consider example (11. III). As in this example we have

$$Y_0'=|x|=\frac{\pi}{2}-\frac{4}{\pi}\left[\frac{\cos x}{1^2}+\frac{\cos 3x}{3^2}+\frac{\cos 5x}{5^2}+\cdots\right], \quad -\pi\leqq x\leqq\pi, \tag{12. 4}$$

[1] By a real-analytic function we mean a function of a real variable, the Taylor series of which is convergent with a sum equal to the function.

we see that Y_0'' does not exist at all for $x=0$. Differentiating the series term by term we obtain, however,

$$(\boldsymbol{A}\cdot\boldsymbol{Y}'(0))_0=\frac{4}{\pi}\cdot\left[\frac{\sin x}{1}+\frac{\sin 3x}{3}+\frac{\sin 5x}{5}+\cdots\right]_{x=0}=0\neq Y_0''(0). \qquad (12.5)$$

In this connection it may be interesting to observe that we may for Y_0' obtain the well-known function of Weierstrass[1].

Example (12. II). ($\sim$ ex. (10. II). (α): $+$. (β): $-$. un: $+$. ex: $+$. Y_0'' does not exist for any x.)

In ex. (10. II) we may choose the Fourier coefficients $a_0, a_1, a_{-1}, \ldots$ of Y_0' in such a way that Y_0' becomes equal to the Weierstrass function

$$Y_0'=\sum_{n=0}^{\infty}a^n\cos(p^n\pi x) \qquad (12.6)$$

in which a is an arbitrary number in $0<a<1$ and p is an odd, positive integer satisfying $ap>1+\frac{3}{2}\pi$. This function has just the property of being continuous in $-\infty<x<\infty$, but not differentiable for any value of x.

§ 13.

Finally we shall discuss an equation which has the following interesting property. A certain matrix function $\boldsymbol{Y}(z)$ is analytic in the whole complex z-plane. Furthermore, $\boldsymbol{Y}(z)$ is in a certain *open* region Ω a solution of an equation $\boldsymbol{Y}'=\boldsymbol{A}\cdot\boldsymbol{Y}$, in which $\boldsymbol{A}$ is analytic in the whole complex z-plane except in certain points in which $\boldsymbol{A}$ has simple poles, but which points do *not* lie on the boundary of Ω (see fig. 3). $\boldsymbol{Y}(z)$ is thus not a solution in limit-points of points in which it is a solution, a behaviour which is excluded in the case of *finite* dimensions (cf. ex. (9. II)).

Example (13. I). ((α): $+$. (β): $\pm$. un: ?. ex: ?. The regularity region of the equation may be *open*).

Let us consider the functions

$$Y_n(z)=e^{-nz}-e^{-(n-1)z}, \qquad n=1,2,3,\ldots, \qquad (13.1)$$

---

[1] Cf. e. g. Titchmarsh (1932), p. 351.

which are analytic in the whole complex z-plane. Next, we consider the equation

$$Y' = \begin{Bmatrix} Y_1' \\ Y_2' \\ Y_3' \\ Y_4' \\ \vdots \end{Bmatrix} = \begin{Bmatrix} -1 & w & w^2 & w^3 & \cdots \\ 0 & -2 & w & w^2 & \cdots \\ 0 & 0 & -3 & w & \cdots \\ 0 & 0 & 0 & -4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{Bmatrix} \cdot Y = A \cdot Y \tag{13.2}$$

in which

$$w = \frac{e^{2z}}{2e^z - 1}. \tag{13.3}$$

In this equation we see that A is in fact analytic in the whole complex z-plane except in the simple poles

$$2e^z - 1 = 0 \quad \text{i. e.} \quad z = -\ln 2 + p \cdot 2\pi i, \qquad p = 0, \pm 1, \pm 2, \ldots. \tag{13.4}$$

Next it is seen that A, being a half matrix (cf. p. 273), is absolutely exponentiable along any regular curve L between x_0 and x. In fact we have

$$\begin{aligned} K_{pq} &= q\,\delta_{pq} + M^{q-p}\Delta(q>p), \\ M &= \max_{x_0 \leq z \leq x} |w(z)|^1 \end{aligned} \tag{13.5}$$

in which the Δ-symbol is defined in (8.21). By induction we find

$$(K^\nu)_{pq} = q^\nu \delta_{pq} + M^{q-p}\Delta(q>p)(q^\nu - (q-1)^\nu), \quad \nu = 0, 1, 2, \ldots, \tag{13.6}$$

and thus

$$\begin{aligned} (\exp[K\,|x-x_0|])_{pq} &= \delta_{pq} \exp[q\,|x-x_0|] + \\ &+ M^{q-p}\Delta(q>p)(\exp[q\,|x-x_0|] - \exp[(q-1)\,|x-x_0|]). \end{aligned} \tag{13.7}$$

Consequently the condition (α) is fulfilled and the equation (13.2) thus covered by our theory.

It is easily verified that (13.1) satisfies our equation (13.2):

$$\begin{aligned} Y_n'(z) &= -\,n e^{-nz} + n e^{-(n-1)z} - e^{-(n-1)z} \\ (A \cdot Y)_n &= -\,n Y_n + \sum_{i=1}^{\infty} \left(\frac{e^{2z}}{2e^z - 1}\right)^i Y_{n+i} = -\,n\,(e^{-nz} - e^{-(n-1)z}) + \\ &\qquad + \sum_{i=1}^{\infty} \left(\frac{e^{2z}}{2e^z - 1}\right)^i (e^{-(n+i)z} - e^{-(n+i-1)z}) = \\ -ne^{-nz} &+ ne^{-(n-1)z} - e^{-nz}(e^z - 1) \sum_{i=1}^{\infty} \left(\frac{e^z}{2e^z - 1}\right)^i = -ne^{-nz} + ne^{-(n-1)z} - e^{-(n-1)z}, \end{aligned} \tag{13.8}$$

[1] If the integration curve from x_0 to x should be complex, we understand by this symbol $\max |w(z)|$ along this curve (cf. [2] p. 262).

q. e. d. The convergence of the series in (13.8) is, however, limited by the condition that

$$\left|\frac{e^z}{2e^z-1}\right| < 1, \quad \text{i. e. } e^{2x} - \frac{4}{3}e^x \cos y + \frac{1}{3} > 0 \qquad (z = x + iy), \tag{13.9}$$

the solution of which is

$$\begin{aligned} x &> \ln\left(\frac{2}{3}\cos y + \frac{1}{3}\sqrt{4\cos^2 y - 3}\right) \\ x &< \ln\left(\frac{2}{3}\cos y - \frac{1}{3}\sqrt{4\cos^2 y - 3}\right). \end{aligned} \tag{13.10}$$

In fig. 3 we have shown the closed regions of divergence (hatched regions) and the *open* region of convergence Ω (the rest of the complex plane) as given by (13.10). Furthermore, we have in the same figure shown the singular points of A as given by (13.4) (the points denoted by $\oplus$). We thus see that Y given in (13.1) satisfies the equation (13.2) in all points of the complex z-plane except in the points of the *closed*, hatched regions of fig. 3; e. g. in the point $z = 0$ Y is *not* a solution although (1): $z = 0$ is a limit-point of points in which Y *is* a solution, (2): $z = 0$ is an *inner* point of the regularity region of A and (3): Y is analytic in the whole complex z-plane. Consequently, we see again that in the case of *infinite* systems the singular points of the equation + initial condition need not be singular points of the *matrix* and cannot even be read off at all from the matrix by an immediate inspection (cf. the observation p. 300—301).

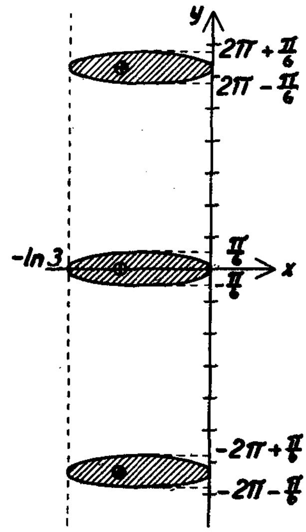

Fig. 3.

These facts are, of course, also met with by the consideration of the condition (β). If, namely, we take as initial point a point z_0 in the regularity region given by (13.10) and as initial condition the corresponding values of Y given in (13.1), the condition (β) turns out to be fulfilled only in a region which lies entirely within the regularity region. Let us e. g. consider the simplest case, viz. z_0 real and positive, $z_0 = x_0 > 0$. We then have from (13.1), considering only values $0 \leqq x \leqq x_0$,

$$G_n = \max_{x \leqq t \leqq x_0} |Y_n(t)| = e^{-(n-1)x} - e^{-nx_0} \quad (x \geqq 0), \tag{13.11}$$

from (13.5)

$$M = \max_{x \leqq t \leqq x_0} |w(t)| = \frac{e^{2x_0}}{2e^{x_0} - 1} \tag{13.12}$$

and, consequently, from (13.7)

$$(\exp[K|x - x_0|] \cdot G)_p =$$

$$\sum_{\alpha=0}^{\infty} \left(\delta_{p\alpha} \exp[\alpha(x_0 - x)] + \left(\frac{e^{2x_0}}{2e^{x_0} - 1} \right)^{\alpha - p} \Delta(\alpha > p) (\exp[\alpha(x_0 - x)] - \exp[(\alpha - 1)(x_0 - x)]) \right) \cdot$$
$$\cdot (\exp[-(\alpha - 1)x] - \exp[-\alpha x_0]) =$$
$$\exp[p x_0 - (2p - 1)x] - \exp[-px] + \tag{13.13}$$
$$\left(\frac{e^{2x_0}}{2e^{x_0} - 1} \right)^{-p} \sum_{\alpha=p+1}^{\infty} \left(\frac{1}{2e^{x_0} - 1} \right)^{\alpha} \Big(-\exp[\alpha(3x_0 - 2x) + 2x - x_0] +$$
$$+ \exp[\alpha(2x_0 - x) + x - x_0] + \exp[\alpha(3x_0 - 2x) + x] - \exp[\alpha(2x_0 - x)] \Big).$$

These series are obviously convergent for such values of x for which

$$\frac{\exp[3x_0 - 2x]}{2e^{x_0} - 1} < 1, \qquad 0 < x \leqq x_0, \tag{13.14}$$

and we just see that this relation is fulfilled for $x = x_0$, but *not* for $x = 0$.

PART IV.

Application to the Perturbation Theory of Quantum Mechanics.

§ 14.

In quantum mechanical perturbation theory stochastically definite processes are also met with, but here the probabilities in question are described by the numerical squares of certain complex functions, the *probability amplitudes*. Let us shortly review the usual perturbation method, the *variation of parameters*, introduced by Dirac[1]. This method underlies every application of the quantum field theories to practical problems in which the perturbation H_1 is considered

[1] See any textbook on quantum theory, e. g. Heitler (1936) chap. III § 9.3. See also Heisenberg (1938) in which paper the theory is presented in such a form that the relativistic invariance is conspicuous.

as causing transitions of the unperturbed system. Let this system have the Hamiltonian H_0, eigenvalues E_n and eigenfunctions ψ_n[1]:

$$H_0 \psi_n = E_n \psi_n. \tag{14.1}$$

Here the ψ_n's are assumed to form a complete ortho-normalized set. Furthermore, it is assumed that all the following formal operations are legitimate. The total Hamiltonian of the perturbed system is now

$$H = H_0 + H_1 \tag{14.2}$$

in which H, H_0 and H_1 are Hermitian operators and H_1, which may or may not contain the time explicitly, is assumed to be small compared with H_0. We develop the solution ψ of the actual Schrödinger-equation

$$i\hbar \frac{\partial}{\partial t} \psi = (H_0 + H_1)\psi \tag{14.3}$$

in a series of the eigenfunctions ψ_n

$$\psi = \sum_n a_n(t)\, \psi_n \exp\left[-\frac{i}{\hbar} E_n t\right]. \tag{14.4}$$

The amplitudes $a_n(t)$ are functions of the time only and the ψ_n's only of the various space and spin coordinates of the unperturbed system. By scalar multiplication of (14.4) by ψ_n we have

$$a_n(t) = \exp\left[\frac{i}{\hbar} E_n t\right] \int \psi_n^* \psi \, d\tau. \tag{14.5}$$

(The integration includes here and in the following also a summation over all spin variables.) Assuming ψ to be normalized to one it follows, furthermore, from (14.4), due to the ψ_n's forming a complete ortho-normalized set, the Parseval relation.

$$\int \psi^* \psi \, d\tau = \sum_n |a_n(t)|^2 = 1. \tag{14.6}$$

The amplitudes $a_n(t)$ have the physical significance that $|a_n(t)|^2$ denotes the probability at the time t of finding the system in the state ψ_n, which interpretation is in agreement with (14.6). From this interpretation it follows that the time variable t moves only in the positive direction, as a probability statement can refer only to the future, not to the past (cf. p. 277).

[1] This notation is used whether the energy spectrum is discrete, continuous or mixed.

Inserting (14.4) into (14.3) and forming the scalar product of both sides with ψ_n we just obtain an infinite system of linear differential equations for the amplitudes $a_n(t)$

$$i\hbar \frac{d}{dt} a_n(t) = \sum_{n'} H_{1_{nn'}} a_{n'}(t) \exp\left[-\frac{i}{\hbar}(E_{n'} - E_n)\, t\right], \tag{14.7}$$

where $H_{1_{nn'}}$ denotes the matrix element

$$H_{1_{nn'}} = \int \psi_n^* H_1 \psi_{n'}\, d\tau \tag{14.8}$$

and represents a transition *from* the state n' *to* the state n. We stress the important fact that the form of the fundamental perturbation equations (14.7) is quite independent of whether they describe a physical system with only a *finite* number of degrees of freedom (point-mechanics) or with an *infinite* number (field-mechanics). In the first case it follows, however, from the theory proper of the wave equation that all our formal operations are legitimate, and that (14.7) has in fact a solution of the form required, but in the second case these statements do no longer hold true. In fact a mathematical theory has, so far as we know, not yet been given for a partial differential equation with an infinity of independent variables.

By means of our matrix symbolism (14.7) may be written in the compact form

$$\frac{d}{dt}\boldsymbol{a}(t) = \boldsymbol{a}'(t) = \boldsymbol{A}(t)\cdot\boldsymbol{a}(t) \tag{14.9}$$

where $\boldsymbol{a}(t)$ is the matrix formed by the probability amplitudes

$$\boldsymbol{a}(t) = \{a_n(t)\} \tag{14.10}$$

and the matrix $\boldsymbol{A}(t)$ is given by

$$\boldsymbol{A}(t) = \{A_{nn'}(t)\} = \left\{-\frac{i}{\hbar} H_{1_{nn'}} \exp\left[-\frac{i}{\hbar}(E_{n'} - E_n)\, t\right]\right\}. \tag{14.11}$$

By means of the diagonal matrix

$$\exp\left[\frac{i}{\hbar}\boldsymbol{E}t\right] = \left\{\exp\left[\frac{i}{\hbar}E_n t\right]\delta_{nn'}\right\} \tag{14.12}$$

$\boldsymbol{A}$ may also be written as

$$\boldsymbol{A} = -\frac{i}{\hbar}\exp\left[\frac{i}{\hbar}\boldsymbol{E}t\right]\cdot\boldsymbol{H}_1\cdot\exp\left[-\frac{i}{\hbar}\boldsymbol{E}t\right]. \tag{14.13}$$

$\boldsymbol{H}_1$ being a Hermitian operator it follows from (14.13) that $\boldsymbol{A}$ is anti-Hermitian

$$\boldsymbol{A}^* = -\boldsymbol{A}. \tag{14.14}$$

Assuming now $\boldsymbol{A}$ to be absolutely exponentiable, i. e. the condition (α), (5.1), to be fulfilled, it follows immediately from part I that the equation (14.9), i. e. (14.7), has for each initial condition

$$\boldsymbol{a}(s) = \{a_n(s)\} \tag{14.15}$$

satisfying (5.3) a unique solution given by

$$\boldsymbol{a}(t) = \mathcal{P}_s^t(\mathbf{1} + \boldsymbol{A}\,dt) \cdot \boldsymbol{a}(s). \tag{14.16}$$

Here the matrix

$$\boldsymbol{a}(t,s) = \mathcal{P}_s^t(\mathbf{1} + \boldsymbol{A}\,dt) \tag{14.17}$$

may be interpreted as a *relative transition probability amplitude*, in analogy with the matrix $\boldsymbol{P}(t,s)$ in part II, because it satisfies the conditions being analogous to (7.4)—(7.8), viz.

$$\lim_{t\to s}\{|\boldsymbol{a}_{nn'}(t,s)|^2\} = \lim_{s\to t}\{|\boldsymbol{a}_{nn'}(t,s)|^2\} = \mathbf{1} \tag{14.18}$$

$$|\boldsymbol{a}_{nn'}(t,s)|^2 = \sum_{n''}|\boldsymbol{a}_{nn''}(t,\tau)|^2\,|\boldsymbol{a}_{n''n'}(\tau,s)|^2 \tag{14.19}$$

$$0 \leqq |\boldsymbol{a}_{nn'}(t,s)|^2 \leqq 1 \tag{14.20}$$

$$\sum_n |\boldsymbol{a}_{nn'}(t,s)|^2 = 1. \tag{14.21}$$

(14.18) follows immediately from (3.11). Next we have, due to the fact that $\mathcal{P}$ satisfies Chapman-Kolmogoroff's equation i. e. (3.14),

$$|\boldsymbol{a}_{nn'}(t,s)|^2 = \Big|\sum_{n''}\boldsymbol{a}_{nn''}(t,\tau)\,\boldsymbol{a}_{n''n'}(\tau,s)\Big|^2 = \sum_{n''}|\boldsymbol{a}_{nn''}(t,\tau)|^2\,|\boldsymbol{a}_{n''n'}(\tau,s)|^2 +$$

$$2\sum_{n''>n'''}|\boldsymbol{a}_{nn''}(t,\tau)\,\boldsymbol{a}_{nn'''}(t,\tau)|\,|\boldsymbol{a}_{n''n'}(\tau,s)\,\boldsymbol{a}_{n'''n'}(\tau,s)|\cos(\varphi^{t\tau}_{nn'''} - \varphi^{t\tau}_{nn''} + \varphi^{\tau s}_{n'''n'} - \varphi^{\tau s}_{n''n'})$$

$$\left(\boldsymbol{a}_{nn''}(t,\tau) = |\boldsymbol{a}_{\bar{n}n''}(t,\tau)|\exp[i\varphi^{t\tau}_{nn''}],\ldots\right). \tag{14.22}$$

Obviously the right hand side of (14.22) is, however, not equal to the right hand side of (14. 19) unless the interference term $2\sum\limits_{n''>n'''}$ vanishes. Whether or not this is the case depends partly on the problem itself and partly on the experimental arrangement[1]. This may, namely, be of such a kind that it averages over all the phases φ in the intermediate states n'', n''' and as $\overline{\cos\varphi}=0$, (14. 19) therefore holds true in such cases. (We observe that this phase averaging is caused by the fact that in quantum theory every observation means an interaction between observer and object which brings about uncontrollable changes in the system observed. The fact that the different probability amplitudes may interfere with each other, i. e. that the intermediate states do not exclude each other two and two, and that (14. 19) is, consequently, not generally true is just one of the most essential features of the quantum theory.

Thirdly it follows from (14. 14), (3. 8) and (3. 16) that $\boldsymbol{a}(t,s)$ is a *unitary* matrix:

$$\boldsymbol{a}^*(t,s)=\left(\overset{t}{\underset{s}{\mathscr{P}}}(1+\boldsymbol{A}\,dt)\right)^*=\lim_{m\to\infty}\prod_{i=m-1}^{0}(1-\boldsymbol{A}(t_i)\Delta_i)=\overset{s}{\underset{t}{\mathscr{P}}}(1+\boldsymbol{A}\,dt)=\boldsymbol{a}^{-1}(t,s). \quad (14.\,23)$$

Consequently we have

$$\sum_n |\boldsymbol{a}_{nn'}(t,s)|^2=(\boldsymbol{a}^*\cdot\boldsymbol{a})_{n'n'}=(\boldsymbol{a}^{-1}\cdot\boldsymbol{a})_{n'n'}=1 \quad (14.\,24)$$

which proves both (14. 20) and (14. 21).

It may be interesting to observe that the fact that $\boldsymbol{a}(t,s)$ is a *unitary* matrix can also be seen directly. If we, namely, transform $\boldsymbol{a}(t,s)$ from the Heisenberg-representation used above to the Schrödinger-representation $\boldsymbol{f}(t,s)$ by means of the transformation

$$\boldsymbol{a}(t,s)=\exp\left[\frac{i}{\hbar}\boldsymbol{E}\,t\right]\cdot\boldsymbol{f}(t,s), \quad (14.\,25)$$

we see from (4. 10) that $\boldsymbol{f}$ satisfies

$$\frac{\partial}{\partial t}\boldsymbol{f}(t,s)=-\frac{i}{\hbar}(\boldsymbol{H}_1+\boldsymbol{E})\cdot\boldsymbol{f}(t,s). \quad (14.\,26)$$

Assuming $\boldsymbol{H}_1$ to be independent of the time, which is usually the case, we have from (4. 6) and (14. 25) that

$$\boldsymbol{a}(t,s)=\exp\left[\frac{i}{\hbar}\boldsymbol{E}\,t\right]\cdot\exp\left[-\frac{i}{\hbar}(\boldsymbol{H}_1+\boldsymbol{E})\,(t-s)\right] \quad (14.\,27)$$

which shows immediately that $\boldsymbol{a}(t,s)$ is a unitary matrix.

[1] Cf. the discussion in Heisenberg (1930) chap. IV § 2. Cf. also Dirac (1930) chap. I.

Finally we see that if $\boldsymbol{a}(s)$ in (14.15) satisfies (5.3) and is besides a unitary column matrix

$$\boldsymbol{a}^*(s)\cdot\boldsymbol{a}(s)=1, \tag{14.28}$$

i. e. $\boldsymbol{a}(s)$ may be interpreted as an absolute probability amplitude, then $\boldsymbol{a}(t)$ given in (14.16) is, besides being a solution of (14.9), a unitary matrix, i. e. may also be interpreted as an absolute probability amplitude. If, namely, $\boldsymbol{a}(s)$ is an arbitrary unitary matrix, we have from (14.23), (14.28) and Schwarz' inequality that

$$|(\boldsymbol{a}(t,s)\cdot\boldsymbol{a}(s))_n| \leqq \sqrt{(\boldsymbol{a}^*(t,s)\cdot\boldsymbol{a}(t,s))_{nn}}\,\sqrt{\boldsymbol{a}^*(s)\cdot\boldsymbol{a}(s)}=1. \tag{14.29}$$

which shows that $\boldsymbol{a}(t)=\boldsymbol{a}(t,s)\cdot\boldsymbol{a}(s)$ exists. Next we have

$$\boldsymbol{a}^*(t)\cdot\boldsymbol{a}(t)=\boldsymbol{a}^*(s)\cdot\boldsymbol{a}^*(t,s)\cdot\boldsymbol{a}(t,s)\cdot\boldsymbol{a}(s)=\boldsymbol{a}^*(s)\cdot\boldsymbol{a}(s)=1 \tag{14.30}$$

q. e. d.

§ 15.

From the Peano series (3.7) for the exact solution (14.16) of the perturbation equations (14.7) we now obtain the well-known expressions (in the case of no resonance) for the probability amplitudes $a_{nn_0}(t,0)$ — giving essentially the transition probabilities from the initial state $n'=n_0$ at the time $s=0$ to the final state n at the time t — in the first, second and higher approximations

$$a_{nn_0}(t,0)=\delta_{nn_0}+a^{(1)}_{nn_0}(t,0)+a^{(2)}_{nn_0}(t,0)+\cdots,\qquad a_{nn_0}(0,0)=\delta_{nn_0} \tag{15.1}$$

in which

$$a^{(1)}_{nn_0}(t,0)=\Big(\int_0^t \boldsymbol{A}(t)\,dt\Big)_{nn_0}=H_{1nn_0}\,\frac{\exp\left[-\frac{i}{\hbar}(E_{n_0}-E_n)\,t\right]-1}{E_{n_0}-E_n}\qquad (n\neq n_0) \tag{15.2}$$

$$a^{(2)}_{nn_0}(t,0)=\Big(\int_0^t dt'\boldsymbol{A}(t')\int_0^{t'}dt''\boldsymbol{A}(t'')\Big)_{nn_0}=$$

$$\sum_{n'}\frac{H_{1nn'}\,H_{1n'n_0}}{E_{n_0}-E_{n'}}\left(\frac{\exp\left[-\frac{i}{\hbar}(E_{n_0}-E_n)\,t\right]-1}{E_{n_0}-E_n}-\frac{\exp\left[-\frac{i}{\hbar}(E_{n'}-E_n)\,t\right]-1}{E_{n'}-E_n}\right) \tag{15.3}$$

and so on. $\qquad (n\neq n_0)$

Now the perturbation H_1 always contains an interaction parameter — e. g. the electric charge e in electro-dynamics or the various f and g factors in the meson theory — and the Peano series (15.1) consequently consists in an expansion in a *power series of this interaction parameter*. This procedure to be legitimate it is, however — quite independent of the numerical magnitude of the parameters in question i. e. whether or not H_1 is small compared with H_0 — a *necessary* condition that the probability amplitudes governing the transition probabilities we are looking for are *analytic* functions in the parameters, and this is by no means always the case. Although A given in (14.11) may be an analytic function of any parameter contained in H_1 we can, as discussed in part I (cf. p. 269), even in the case when A fulfills our essential condition of being absolutely exponentiable conclude only that the solutions of our perturbation equations (14.9), i. e. (14.7), are analytic in every *inner* point of their convergence region. Just the initial point from which we expand our series may, namely, be a singular point in which the first, but not the higher derivatives giving the coefficients of our expansion exist. This fact is in a most striking way illustrated by the equation in examples (10.I) and (11.I) in which one of the functions, given in (11.4), has in the initial point a very serious singularity of the same type as the function $y = x^2 \ln x$, namely an infinite branch-point.

As is well-known we are in the application of quantum mechanics to the field theories — both electro-dynamics and the various meson theories — just faced with this peculiar situation that the theories lead, when applied to practical problems, in the first approximation always to convergent results (*simply because the probability amplitudes satisfy differential equations of the first order, viz.* (14.9)) *which agree with experimental results, but that the higher approximations often give divergent results, which fact means that no physical meaning can be attached to them.* Usually this difficulty is simply overcome by various artificial methods such as »cutting off« the divergent integrals at some suitably chosen point. Such a procedure is, of course, highly unsatisfactory, quite apart from the fact that it spoils the relativistic invariance of the theory. From our general theory it is, furthermore, obvious that we may not expect results obtained in such ways to have much physical meaning in accordance with what is found to be the case, especially in the meson theories, which give quite wrong results for very high energies.

The question thus naturally arises whether these divergence difficulties are due to deficiences of the present quantum theory or to our usual perturbation methods failing. The last possibility has previously been suggested from time to time[1] and arguments may also be given in its favour. Firstly it may be said that the existence

[1] Cf. e.g. Rosenfeld (1935).

of solutions of the perturbation equations (14.7) follows by means of (14.5) from the theory of the wave equation. Secondly such examples as that discussed in examples (10.I) and (11.I) show that even if our main condition of the operator matrix A of the equations being absolutely exponentiable is not fulfilled, i. e. that the usual method of solving by means of the Peano series diverges, the equations may still have a unique — although non-analytic — solution. In fact it may not be wondered at that the usual perturbation methods may fail because these methods of solution have been formally carried over from finite systems of equations which, as shown in part I, can have no other singularities than the singular points of the matrix A itself, to infinite systems in which, as discussed in part III, the singular points may arise through the limiting processes proper defining the system itself and need not at all be singular points of A or even be detectable by an immediate inspection of A. Just as the divergence difficulties which arose in the theory of collision problems by a too rough application of the Born approximation at low velocities were later removed by the more suitable perturbation methods of e. g. Faxén and Holtsmark, we ought perhaps at present rather look for better mathematical methods of solving the perturbation equations (14.7) than for better physical theories. Although for this purpose eventually quite new, and perhaps hitherto unknown, mathematical methods have to be invented for dealing with infinite systems of differential equations not admitting of iteration solutions, i. e. which are not covered by the condition of A being absolutely exponentiable, it may not be premature to suggest such methods to consist simply in new ways of expanding our solutions in series. Bearing in mind how partial differential equations are solved in problems of heat conduction or diffusion it is an obvious idea to suggest the application e. g. of Fourier analysis on our solutions. This method seems specially promising in as much as it is well-known that alone the existence of the first derivative of a function is enough to ensure its Fourier series to be convergent. Also the work of Poincaré on the application of infinite determinants in the perturbation theory of astronomy may perhaps turn out to be useful[1].

In spite of the arguments just discussed, *the first of the above mentioned possibilities of understanding the divergence difficulties, viz. that they are more deep-rooted, being due to deficiences of the present quantum theory itself, must now be favoured* by the following reasons. As regards the first argument in the discussion

[1] We intend to investigate these problems more closely.

above it must not be forgotten, firstly that the Schrödinger equations (14. 3) occurring in the *field* theories describe physical systems with an *infinity* of degrees of freedom, and secondly that the perturbing interaction term H_1 between the atomic systems and the wave fields in question must, due to the relativistic invariance, involve the highly discontinuous Dirac δ-functions. Rigorous proofs of the existence of finite eigenvalues and eigenfunctions and of the completeness of the latter have, however, as already mentioned (p. 312) not yet been given. (In fact such systems may be constructed which have infinite eigenvalues.) These facts underlying the deduction of (14. 5) we cannot, consequently, conclude that the existence of solutions of the perturbation equations follows from the theory of the wave equation.

Secondly our discussion in part III shows equally well that our perturbation equations may have no solutions at all. In fact this may be the case in spite of A being anti-Hermitian as shown by example (11. II) if, only, we multiply the — symmetric — matrix of equation (11. 13) by i.

Thirdly Heisenberg[1] has given strong arguments showing that in order that the present quantum mechanics shall give a consistent description of nature, the perturbation equations *must* not at all give convergent results. If this were the case, this fact would, e. g., imply that the theory would yield convergent expressions for the self-energies of all the elementary particles. Consequently the masses of these particles would be given by the theory itself in spite of the fact that these masses enter also in the theory as arbitrary parameters, the values of which we may ourselves dispose of freely. The whole present quantum theory being just founded on the correspondence principle as shown by the way the Hamiltonian (14. 2) itself is built up from a 0-approximation term, H_0, and a 1-approximation term, H_1, we cannot, consequently, expect the present theory to give convergent results beyond the first approximation — and if it did, this fact would, as mentioned, even lead to contradictions in the interpretation of the theory.

Notwithstanding the fact that a more general theory of infinite systems of differential equations than the present theory will certainly be created in the future, and that such a generalization is much needed in e. g. the theory of stochastic processes, we must conclude from the above discussion that such a generalization may not be expected to overcome any of the divergence difficulties of the present quantum theory, these difficulties being far more deep-rooted in this theory itself.

[1] We wish to thank prof. Heisenberg for valuable discussions on these questions.

Summary.

In part I (§§ 1—5) we first review (§§ 1—4) the usual theory of *finite* systems of simultaneous linear differential equations of arbitrary order. In § 1 we present the theory in matrix form, the theory becoming thus independent of the dimension of the system. In §§ 2—3 we prove the theorems of uniqueness and existence, respectively. In § 4 we give some properties of the product-integral representing the exact solutions. In § 5 we perform the transition to *infinite* systems, giving conditions which are *sufficient* to allow us of maintaining the whole theory of §§ 1—4 for infinite systems.

In part II (§§ 6—8) we give the application of *infinite* equations to the theory of stochastic, discontinuous processes. It is shown that a wide class of such processes, being most important in the practical statistical applications of this theory, is covered by our theory and satisfies all the requirements being necessary for an interpretation of the solutions as probabilities being possible.

In part III (§§ 9—13) we investigate the conditions of § 5, ensuring the necessary convergences. By means of suitably constructed examples we show, partly that our conditions are only *sufficient*, but not *necessary* to maintain the theorems of uniqueness and existence, and partly the important fact that these theorems themselves do not generally hold true for *infinite* systems. Especially we discuss in §§ 10—11 the questions regarding the theorems of uniqueness and existence, respectively. As a result the important fact turns out that in contrast to *finite* systems an *infinite* system may have singular points other than the singular points of the matrix of the equation and that the former singular points may not always be read off from the matrix itself by an immediate inspection.

In part IV (§§ 14—15) we give the application of infinite equations to the perturbation theory of quantum mechanics. It is shown that the fact that the usual perturbation method gives in the first approximation always convergent results, being in agreement with experiments, in spite of the higher approximations diverging, is simply explained by the fact that the solutions of the perturbation equations — if they exist at all — need not be *analytic* functions in the parameters, but that the initial point from which we expand our series may be a singular point in which the first, but not the higher derivatives, giving the coefficients of our expansion, exist. Finally we shortly discuss whether these divergence difficulties are due to deficiences of the present quantum theory or to our usual perturbation methods failing. Notwithstanding the fact that a more

general theory of infinite systems of differential equations than the present theory is required (and possibilities for such generalizations are suggested in the form of other ways of expanding the solutions in series) we conclude with Heisenberg that the divergence difficulties are more deep-rooted. They are, namely, a consequence of the present quantum theory being based on the correspondence principle, a consequence which is not to be regretted, but on the contrary necessary for the consistent interpretation of the theory.

List of pathological examples in part III.

(The symbols (α): +, (β): —, un: +, ex: — and so on denote, respectively, the condition (α) is fulfilled, (β) is not, the theorem of uniqueness is fulfilled, that of existence is not, and so on.)

List of references.

ARLEY 1943, On the theory of stochastic processes and their application to the theory of cosmic radiation, Copenhagen (Dissertation).

DIRAC 1930, The principles of quantum mechanics, Oxford.

FELLER 1937, Math. Ann. **113**, 113.

FRAZER, DUNCAN and COLLAR 1938, Elementary Matrices, Cambridge.

FRÉCHET 1938, Traité du calcul des probabilités et de ses applications, **I 3**, Livre II, Paris.

HEISENBERG 1930, Die physikalischen Prinzipien der Quantenmechanik, Leipzig.

—— 1938, Z. Phys. **110**, 251.

HEITLER 1936, Quantum theory of radiation, Oxford.

KHINTCHINE 1934, Math. Ann. **109**, 604.

KOLMOGOROFF 1931, Math. Ann. **104**, 415.

LUNDBERG 1940, On random processes and their application to sickness and accident statistics, Uppsala (Dissertation).

RASCH 1930, Om Matrixregning og dens Anvendelse paa Differens- og Differentialligninger, København (Dissertation).

—— 1934, Journ. rein. angew. Math. **171**, 65.

REICHENBACH 1935, Wahrscheinlichkeitslehre, Leiden.

ROSENFELD 1935, Fys. Tidsskr. N. Bohr anniversary 120.

TITCHMARSH 1932, Theory of functions, Oxford.