INTRODUCTORY REAL ANALYSIS

A. N. KOLMOGOROV
S. V. FOMIN

Revised English Edition
Translated and Edited by

Richard A. Silverman

PRENTICE-HALL, INC.
Englewood Cliffs, N.J.

Problem 6. Give an example of a complete metric space R and a nested sequence $\{A_n\}$ of closed subsets of R such that

$$\bigcap_{n=1}^{\infty} A_n = \varnothing.$$

Reconcile this example with Problem 4.

Problem 7. Prove that a subspace of a complete metric space R is complete if and only if it is closed.

Problem 8. Prove that the real line equipped with the distance

$$\rho(x, y) = |\arctan x - \arctan y|$$

is an incomplete metric space.

Problem 9. Give an example of a complete metric space homeomorphic to an incomplete metric space.

Hint. Consider the example on p. 44.

Comment. Thus homeomorphic metric spaces can have different "metric properties."

Problem 10. Carry out the program discussed in the last sentence of the example on p. 65.

Hint. If $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences of rational numbers serving as "representatives" of real numbers x^* and y^*, respectively, define $x^* + y^*$ as the real number with representative $\{x_n + y_n\}$.

8. Contraction Mappings

8.1. Definition of a contraction mapping. The fixed point theorem. Let A be a mapping of a metric space R into itself. Then x is called a *fixed point* of A if $Ax = x$, i.e., if A carries x into itself. Suppose there exists a number $\alpha < 1$ such that

$$\rho(Ax, Ay) \leqslant \alpha\rho(x, y) \qquad (1)$$

for every pair of points $x, y \in R$. Then A is said to be a *contraction mapping.* Every contraction mapping is automatically continuous, since it follows from the "contraction condition" (1) that $Ax_n \to Ax$ whenever $x_n \to x$.

THEOREM 1 (***Fixed point theorem***[10]). *Every contraction mapping A defined on a complete metric space R has a unique fixed point.*

[10] Often called the *method of successive approximations* (see the remark following Theorem 1) or the *principle of contraction mappings.*

Proof. Given an arbitrary point x_0 R, let[11]

$$x_1 = Ax_0, \quad x_2 = Ax_1 = A^2x_0, \ldots, \quad x_n = Ax_{n-1} = A^nx_0, \ldots \qquad (2)$$

Then the sequence $\{x_n\}$ is fundamental. In fact, assuming to be explicit that $n \leqslant n'$, we have

$$\begin{aligned}\rho(x_n, x_{n'}) &= \rho(A^nx_0, A^{n'}x_0) \leqslant \alpha^n\rho(x_0, x_{n'-n}) \\ &\leqslant \alpha^n[\rho(x_0, x_1) + \rho(x_1, x_2) + \cdots + \rho(x_{n'-n-1}, x_{n'-n})] \\ &\leqslant \alpha^n\rho(x_0, x_1)[1 + \alpha + \alpha^2 + \cdots + \alpha^{n'-n-1}] < \alpha^n\rho(x_0, x_1)\frac{1}{1-\alpha}.\end{aligned}$$

But the expression on the right can be made arbitrarily small for sufficiently large n, since $\alpha < 1$. Since R is complete, the sequence $\{x_n\}$, being fundamental, has a limit

$$x = \lim_{n\to\infty} x_n.$$

Then, by the continuity of A,

$$Ax = A \lim_{n\to\infty} x_n = \lim_{n\to\infty} Ax_n = \lim_{n\to\infty} x_{n+1} = x.$$

This proves the existence of a fixed point x. To prove the uniqueness of x, we note that if

$$Ax = x, \qquad Ay = y,$$

(1) becomes

$$\rho(x, y) \leqslant \alpha\rho(x, y).$$

But then $\rho(x, y) = 0$ since $\alpha < 1$, and hence $x = y$. ∎

Remark. The fixed point theorem can be used to prove existence and uniqueness theorems for solutions of equations of various types. Besides showing that an equation of the form $Ax = x$ has a unique solution, the fixed point theorem also gives a practical method for finding the solution, i.e., calculation of the "successive approximations" (2). In fact, as shown in the proof, the approximations (2) actually converge to the solution of the equation $Ax = x$. For this reason, the fixed point theorem is often called the *method of successive approximations.*

Example 1. Let f be a function defined on the closed interval $[a, b]$ which which maps $[a, b]$ into itself and satisfies a Lipschitz condition

$$|f(x_1) - f(x_2)| \leqslant K|x_1 - x_2|, \qquad (3)$$

with constant $K < 1$. Then f is a contraction mapping, and hence, by

[11] A^2x means $A(Ax)$, A^3x means $A(A^2x) = A^2(Ax)$, and so on.

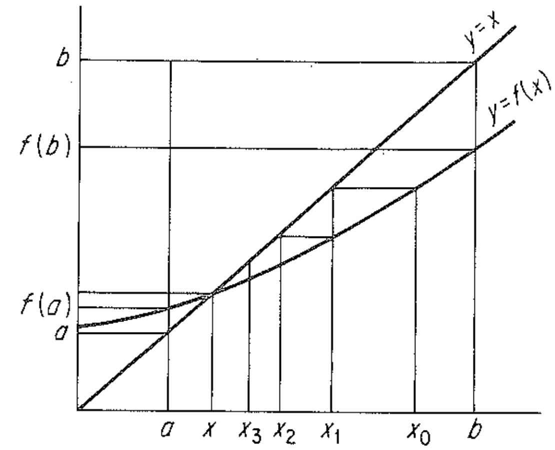

FIGURE 10

Theorem 1, the sequence

$$x_0, \qquad x_1 = f(x_0), \qquad x_2 = f(x_1), \ldots \tag{4}$$

converges to the unique root of the equation $f(x) = x$. In particular, the "contraction condition" (3) holds if f has a continuous derivative f' on $[a, b]$ such that

$$|f'(x)| \leqslant K < 1.$$

The behavior of the successive approximations (4) in the cases $0 < f'(x) < 1$ and $-1 < f'(x) < 0$ is shown in Figures 10 and 11.

***Example* 2.** Consider the mapping A of n-dimensional space into itself given by the system of linear equations

$$y_i = \sum_{j=1}^{n} a_{ij}x_j + b_i \qquad (i = 1, \ldots, n). \tag{5}$$

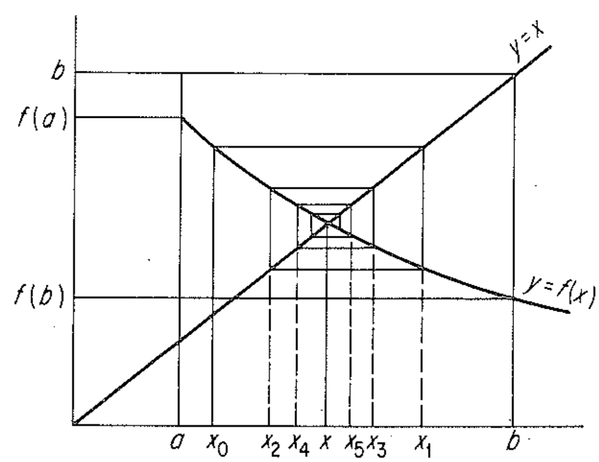

FIGURE 11

If A is a contraction mapping, we can use the method of successive approximations to solve the equation $Ax = x$. The conditions under which A is a contraction mapping depend on the choice of metric. We now examine three cases:

1) The space R_0^n with metric

$$\rho(x, y) = \max_{1 \leqslant i \leqslant n} |x_i - y_i|.$$

In this case,

$$\begin{aligned}\rho(y, \tilde{y}) &= \max_i |y_i - \tilde{y}_i| = \max_i \left| \sum_j a_{ij}(x_j - \tilde{x}_j) \right| \\ &\leqslant \max_i \sum_j |a_{ij}|\, |x_j - \tilde{x}_j| \\ &\leqslant \max_i \sum_j |a_{ij}| \max_j |x_j - \tilde{x}_j| = \left(\max_i \sum_j |a_{ij}| \right) \rho(x, \tilde{x}),\end{aligned}$$

and the contraction condition

$$\sum_j |a_{ij}| \leqslant \alpha < 1 \qquad (i = 1, \ldots, n). \tag{6}$$

2) The space R_1^n with metric

$$\rho(x, y) = \sum_{i=1}^{n} |x_i - y_i|.$$

Here

$$\begin{aligned}\rho(y, \tilde{y}) &= \sum_i |y_i - \tilde{y}_i| = \sum_i \left| \sum_j a_{ij}(x_j - \tilde{x}_j) \right| \\ &\leqslant \sum_i \sum_j |a_{ij}|\, |x_j - \tilde{x}_j| \leqslant \left(\max_j \sum_i |a_{ij}| \right) \rho(x, \tilde{x}),\end{aligned}$$

and the contraction condition is now

$$\sum_i |a_{ij}| \leqslant \alpha < 1 \qquad (j = 1, \ldots, n). \tag{7}$$

3) Ordinary Euclidean space R^n with metric

$$\rho(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

Using the Cauchy-Schwarz inequality, we have

$$\rho^2(y, \tilde{y}) = \sum_i \left(\sum_j a_{ij}(x_j - \tilde{x}_j) \right)^2 \leqslant \left(\sum_i \sum_j a_{ij}^2 \right) \rho^2(x, \tilde{x}),$$

and the contraction condition becomes

$$\sum_i \sum_j a_{ij}^2 \leqslant \alpha < 1. \tag{8}$$

Thus, if at least one of the conditions (6)–(8) holds, there exists a unique point $x = (x_1, x_2, \ldots, x_n)$ such that

$$x_i = \sum_{j=1}^{n} a_{ij}x_j + b_i \qquad (i = 1, \ldots, n). \tag{9}$$

The sequence of successive approximations to this solution of the equation $x = Ax$ are of the form

$$x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}),$$
$$x^{(1)} = (x_1^{(1)}, x_2^{(1)}, \ldots, x_n^{(1)}),$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}),$$
$$\cdots\cdots\cdots\cdots\cdots$$

where

$$x_i^{(k)} = \sum_{j=1}^{n} a_{ij}x_j^{(k-1)} + b_i,$$

and we can choose any point $x^{(0)}$ as the "zeroth approximation."

Each of the conditions (6)–(8) is *sufficient* for applicability of the method of successive approximations, but none of them is *necessary*. In fact, examples can be constructed in which each of the conditions (6)–(8) is satisfied, but not the other two.

Theorem 1 has the following useful generalization, which will be needed later (see Example 2, p. 75):

THEOREM 1′. *Given a continuous mapping of a complete metric space R into itself, suppose A^n is a contraction mapping (n an integer > 1). Then A has a unique fixed point.*

Proof. Choosing any point $x_0 \in R$, let

$$x = \lim_{k \to \infty} A^{kn}x_0.$$

Then, by the continuity of A,

$$Ax = \lim_{k \to \infty} AA^{kn}x_0.$$

But A^n is a contraction mapping, and hence

$$\rho(A^{kn}Ax_0, A^{kn}x_0) \leqslant \alpha\rho(A^{(k-1)n}Ax_0, A^{(k-1)n}x_0) \leqslant \cdots \leqslant \alpha^k\rho(Ax_0, x_0)$$

where $\alpha < 1$. It follows that

$$\rho(Ax, x) = \lim_{k \to \infty} \rho(AA^{kn}x_0, A^{kn}x_0) = 0,$$

i.e., $Ax = x$ so that x is a fixed point of A. To prove the uniqueness of x, we merely note that if A has more than one fixed point, then so does A^n, which is impossible, by Theorem 1, since A^n is a contraction mapping. ∎

8.2. Contraction mappings and differential equations. The most interesting applications of Theorems 1 and 1′ arise when the space R is a function space. We can then use these theorems to prove a number of existence and uniqueness theorems for differential and integral equations, as shown in this section and the next.

THEOREM 2 (***Picard***). *Given a function* $f(x, y)$ *defined and continuous on a plane domain* G *containing the point* (x_0, y_0),[12] *suppose* f *satisfies a Lipschitz condition of the form*

$$|f(x, y) - f(x, \tilde{y})| \leqslant M\,|y - \tilde{y}|$$

in the variable y. *Then there is an interval* $|x - x_0| \leqslant \delta$ *in which the differential equation*

$$\frac{dy}{dx} = f(x,y) \tag{10}$$

has a unique solution

$$y = \varphi(x)$$

satisfying the initial condition

$$\varphi(x_0) = y_0. \tag{11}$$

Proof. Together the differential equation (10) and the initial condition (11) are equivalent to the integral equation

$$\varphi(x) = y_0 + \int_{x_0}^{x} f(t, \varphi(t))\, dt. \tag{12}$$

By the continuity of f, we have

$$|f(x, y)| \leqslant K \tag{13}$$

in some domain $G' \subset G$ containing the point (x_0, y_0).[13] Choose $\delta > 0$ such that

1) $(x, y) \in G'$ if $|x - x_0| \leqslant \delta$, $|y - y_0| \leqslant K\delta$;
2) $M\delta < 1$,

and let C^* be the space of continuous functions φ defined on the interval

[12] By an n-dimensional *domain* we mean an open connected set in Euclidean n-space R^n (connectedness is defined in Problem 12, p. 55).

[13] In fact, f is bounded on $[G']$ if $[G'] \subset G$ (cf. Theorem 2, p. 110).

$|x - x_0| \leqslant \delta$ and such that $|\varphi(x) - y_0| \leqslant K\delta$, equipped with the metric

$$\rho(\varphi, \tilde{\varphi}) = \max_x |\varphi(x) - \tilde{\varphi}(x)|.$$

The space C^* is complete, since it is a closed subspace of the space of all continuous functions on $[x_0 - \delta, x_0 + \delta]$. Consider the mapping $\psi = A\varphi$ defined by the integral equation

$$\psi(x) = y_0 + \int_{x_0}^{x} f(t, \varphi(t))\, dt \qquad (|x - x_0| \leqslant \delta).$$

Clearly A is a contraction mapping carrying C^* into itself. In fact, if $\varphi \in C^*$, $|x - x_0| \leqslant \delta$ then

$$|\psi(x) - y_0| = \left| \int_{x_0}^{x} f(t, \varphi(t))\, dt \right| \leqslant \int_{x_0}^{x} |f(t, \varphi(t))|\, dt \leqslant K\,|x - x_0| \leqslant K\delta$$

by (13), and hence $\psi = A\varphi$ also belongs to C^*. Moreover,

$$|\psi(x) - \tilde{\psi}(x)| \leqslant \int_{x_0}^{x} |f(t, \varphi(t)) - f(t, \tilde{\varphi}(t))|\, dt \leqslant M\delta \max_x |\varphi(x) - \tilde{\varphi}(x)|,$$

and hence

$$\rho(\psi, \tilde{\psi}) \leqslant M\delta\rho(\varphi, \tilde{\varphi})$$

after maximizing with respect to x. But $M\delta < 1$, so that A is a contraction mapping. It follows from Theorem 1 that the equation $\varphi = A\varphi$, i.e., the integral equation (12), has a unique solution in the space C^*. ∎

Theorem 2 can easily be generalized to the case of *systems* of differential equations:

THEOREM 2′. *Given n functions $f_i(x, y_1, \ldots, y_n)$ defined and continuous on an $(n + 1)$-dimensional domain G containing the point*

$$(x_0, y_{01}, \ldots, y_{0n}),$$

suppose each f_i satisfies a Lipschitz condition of the form

$$|f_i(x, y_1, \ldots, y_n) - f_i(x, \tilde{y}_1, \ldots, {}_n\tilde{y})| \leqslant M \max_{1 \leqslant i \leqslant n} |y_i - \tilde{y}_i|$$

in the variables $y_1, \ldots, y_n$. Then there is an interval $|x - x_0| \leqslant \delta$ in which the system of differential equations

$$\frac{dy_i}{dx} = f_i(x, y_1, \ldots, y_n) \qquad (i = 1, \ldots, n) \tag{14}$$

has a unique solution

$$y_1 = \varphi_1(x), \ldots, y_n = \varphi_n(x)$$

satisfying the initial conditions

$$\varphi_1(x_0) = y_{01}, \ldots, \varphi_n(x_0) = y_{0n}. \tag{15}$$

Proof. Together the differential equations (14) and the initial conditions (15) are equivalent to the system of integral equations

$$\varphi_i(x) = y_{0i} + \int_{x_0}^{x} f_i(t, \varphi_1(t), \ldots, \varphi_n(t))\, dt \qquad (i = 1, \ldots, n). \tag{16}$$

By the continuity of the functions f_i, we have

$$|f_i(x, y_1, \ldots, y_n)| \leqslant K \qquad (i = 1, \ldots, n) \tag{17}$$

in some domain $G' \subset G$ containing the point $(x_0, y_{01}, \ldots, y_{0n})$. Choose $\delta > 0$ such that

1) $(x, y_1, \ldots, y_n) \in G'$ if $|x - x_0| \leqslant \delta$, $|y_i - y_{0i}| \leqslant K\delta$ for all $i = 1, \ldots, n$;
2) $M\delta < 1$.

This time let C^* be the space of ordered n-tuples

$$\varphi = (\varphi_1, \ldots, \varphi_n)$$

of continuous functions $\varphi_1, \ldots, \varphi_n$ defined on the interval $|x - x_0| \leqslant \delta$ such that $|\varphi_i(x) - y_{0i}| \leqslant K\delta$ for all $i = 1, \ldots, n$, equipped with the metric

$$\rho(\varphi, \tilde{\varphi}) = \max_{x,i} |\varphi_i(x) - \tilde{\varphi}_i(x)|.$$

Clearly C^* is complete. Moreover, the mapping $\psi = A\varphi$ defined by the system of integral equations

$$\psi_i(x) = y_{0i} + \int_{x_0}^{x} f_i(t, \varphi_1(t), \ldots, \varphi_n(t))\, dt$$

$$(|x - x_0| \leqslant \delta, i = 1, \ldots, n)$$

is a contraction mapping carrying C^* into itself. In fact, if

$$\varphi = (\varphi_1, \ldots, \varphi_n) \in C^*, \qquad |x - x_0| \leqslant \delta,$$

then

$$|\psi_i(x) - y_{0i}| = \left| \int_{x_0}^{x} f_i(t, \varphi_1(t), \ldots, \varphi_n(t))\, dt \right| \leqslant K\delta \qquad (i = 1, \ldots, n)$$

by (17), so that $\psi = (\psi_1, \ldots, \psi_n) = A\varphi$ also belongs to C^*. Moreover,

$$|\psi_i(x) - \tilde{\psi}_i(x)| = \int_{x_0}^{x} |f_i(t, \varphi_1(t), \ldots, \varphi_n(t)) - f_i(t, \tilde{\varphi}_1(t), \ldots, \tilde{\varphi}_n(t))|\, dt$$

$$\leqslant M\delta \max_i \varphi_i(t) - |\tilde{\varphi}_i(t)|,$$

and hence

$$\rho(\psi, \tilde{\psi}) \leqslant M\delta\rho(\varphi, \tilde{\varphi})$$

after maximizing with respect to x and i. But $M\delta < 1$, so that A is a contraction mapping. It follows from Theorem 1 that the equation $\varphi = A\varphi$, i.e., the system of integral equations (16), has a unique solution in the space C^*. ∎

8.3. Contraction mappings and integral equations. We now show how the method of successive approximations can be used to prove the existence and uniqueness of solutions of integral equations.

Example 1. By a *Fredholm equation* (of the second kind) is meant an integral equation of the form

$$f(x) = \lambda \int_a^b K(x, y) f(y)\, dy + \varphi(x), \tag{18}$$

involving two given functions K and φ, an unknown function f and an arbitrary parameter λ. The function K is called the *kernel* of the equation, and the equation is said to be *homogeneous* if $\varphi \equiv 0$ (but otherwise *non-homogeneous*).

Suppose $K(x, y)$ and $\varphi(x)$ are continuous on the square $a \leqslant x \leqslant b$, $a \leqslant y \leqslant b$, so that in particular

$$|K(x, y)| \leqslant M \qquad (a \leqslant x \leqslant b, a \leqslant y \leqslant b).$$

Consider the mapping $g = Af$ of the complete metric space $C_{[a,b]}$ into itself given by

$$g(x) = \lambda \int_a^b K(x, y) f(y)\, dy + \varphi(x).$$

Clearly, if $g_1 = Af_1$, $g_2 = Af_2$, then

$$\rho(g_1, g_2) = \max_x |g_1(x) - g_2(x)| \leqslant |\lambda|\, M(b - a) \max_x |f_1(x) - f_2(x)|$$
$$= |\lambda|\, M(b - a)\rho(f_1, f_2),$$

so that A is a contraction mapping if

$$|\lambda| < \frac{1}{M(b - a)}. \tag{19}$$

It follows from Theorem 1 that the integral equation (18) has a unique solution for any value of λ satisfying (19). The successive approximations $f_0, f_1, \ldots, f_n, \ldots$ to this solution are given by

$$f_n(x) = \lambda \int_a^b K(x, y) f_{n-1}(y)\, dy + \varphi(x) \qquad (n = 1, 2, \ldots),$$

where any function continuous on $[a, b]$ can be chosen as f_0. Note that the method of successive approximations can be applied to the equation (18) only for sufficiently small $|\lambda|$.

Example 2. Next consider the *Volterra equation*

$$f(x) = \lambda \int_a^x K(x, y) f(y)\, dy + \varphi(x), \tag{20}$$

which differs from the Fredholm equation (18) by having the variable x rather than the fixed number b as the upper limit of integration.[14] It is easy to see that the method of successive approximations can be applied to the Volterra equation (20) for *arbitrary* λ, not just for sufficiently small $|\lambda|$ as in the case of the Fredholm equation (18). In fact, let A be the mapping of $C_{[a,b]}$ into itself defined by

$$Af(x) = \lambda \int_a^x K(x, y) f(y)\, dy + \varphi(x),$$

and let $f_1, f_2 \in C_{[a,b]}$. Then

$$|Af_1(x) - Af_2(x)| = \lambda \int_a^x K(x, y)[f_1(y) - f_2(y)]\, dy$$
$$\leqslant \lambda M(x - a) \max_x |f_1(x) - f_2(x)|,$$

where

$$M = \max_{x,y} |K(x, y)|.$$

It follows that

$$|A^2 f_1(x) - A^2 f_2(x)| \leqslant \lambda^2 M^2 \max_x |f_1(x) - f_2(x)| \int_a^x (x - a)\, dx$$
$$= \lambda^2 M^2 \frac{(x - a)^2}{2} \max_x |f_1(x) - f_2(x)|,$$

and in general,

$$|A^n f_1(x) - A^n f_2(x)| \leqslant \lambda^n M^n \frac{(x - a)^n}{n!} \max_x |f_1(x) - f_2(x)|$$
$$\leqslant \lambda^n M^n \frac{(b - a)^n}{n!} \max_x |f_1(x) - f_2(x)|,$$

which implies

$$\rho(A^n f_1, A^n f_2) \leqslant \lambda^n M^n \frac{(b - a)^n}{n!} \rho(f_1, f_2).$$

---

[14] Equation (20) can be regarded formally as a special case of (18) by extending the definition of the kernel, i.e., by setting

$$K(x, y) = 0 \quad \text{if} \quad y > x.$$

But, given any λ, we can always choose n large enough to make

$$\lambda^n M^n \frac{(b-a)^n}{n!} < 1,$$

i.e., A^n is a contraction mapping for sufficiently large n. It follows from Theorem 1′ that the integral equation (20) has a unique solution for arbitrary λ.

Problem 1. Let A be a mapping of a metric space R into itself. Prove that the condition

$$\rho(Ax, Ay) < \rho(x, y) \qquad (x \neq y)$$

is insufficient for the existence of a fixed point of A.

Problem 2. Let $F(x)$ be a continuously differentiable function defined on the interval $[a, b]$ such that $F(a) < 0$, $F(b) > 0$ and

$$0 < K_1 \leqslant F'(x) \leqslant K_2 \qquad (a \leqslant x \leqslant b).$$

Use Theorem 1 to find the unique root of the equation $F(x) = 0$.

Hint. Introduce the auxiliary function $f(x) = x - \lambda F(x)$, and choose λ such that the theorem works for the equivalent equation $f(x) = x$.

Problem 3. Devise a proof of the implicit function theorem based on the use of the fixed point theorem.[15]

Problem 4. Prove that the method of successive approximations can be used to solve the system (9) if $|a_{ij}| < 1/n$ (for all i and j), but not if $|a_{ij}| = 1/n$.

Problem 5. Prove that the condition (6) is necessary for the mapping (5) to be a contraction mapping in the space R_0^n.

Problem 6. Prove that any of the conditions (6)–(8) implies

$$\begin{vmatrix} a_{11}-1 & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-1 & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn}-1 \end{vmatrix} \neq 0.$$

Comment. Hence the fact that the system (5) has a unique solution (under suitable conditions) follows from Cramer's rule as well as from the fixed point theorem.

[15] See e.g., I. G. Petrovski, *Ordinary Differential Equations* (translated by R. A. Silverman), Prentice-Hall, Inc., Englewood Cliffs, N.J. (1966), p. 47.

Problem 7. Consider the nonlinear integral equation

$$f(x) = \lambda \int_a^b K(x, y; f(y))\, dy + \varphi(x) \tag{21}$$

with continuous K and φ, where K satisfies a Lipschitz condition of the form

$$|K(x, y; z_1) - K(x, y; z_2)| \leqslant M\, |z_1 - z_2|$$

in its "functional" argument. Prove that (21) has a unique solution for all

$$|\lambda| < \frac{1}{M(b - a)}.$$

Write the successive approximations to this solution.