Capitolo 1

Spazi Metrici

Uno spazio metrico è un insieme X dotato di una *metrica.* La metrica associa ad ogni coppia di elementi (*punti*) di X una *distanza.* La metrica è definita assiomaticamente, gli assiomi essendo suggeriti da alcune proprietà semplici della distanza, così com'è familiarmente definita fra punti della retta reale $\mathbb{R}$ o del piano complesso $\mathbb{C}$. Si tratta come mostrano alcuni esempi basilari di un concetto molto generale. Un'importante proprietà aggiuntiva che uno spazio metrico può possedere è la *completezza.* Un altro concetto di interesse teorico e pratico è la *separabilità* di uno spazio metrico. Gli spazi metrici separabili sono più semplici di quelli non separabili.

1.1 Definizione (Spazio Metrico, Metrica)
Uno *spazio metrico* è una coppia (X, d), dove X è un insieme e d una *metrica su* X (o *distanza* su X), cioè una funzione definita su $X \times X$ tale che per ogni $x, y, z \in X$ si abbia

M1 d è a valori reali, finito e non negativo.

M2 $d(x, y) = 0$ se e solo se $x = y$.

M3 $d(x, y) = d(y, x)$ **(Simmetria)**

M4 $d(x, y) \leq d(x, z) + d(z, y)$ **(Disuguaglianza Triangolare)**

■

Alcuni termini di uso corrente sono i seguenti. X è normalmente chiamato l'*insieme sottostante* a (X, d). I suoi elementi sono chiamati *punti.* Per x, y fissati il numero non negativo $d(x, y)$ si chiama *distanza* fra x e y. Le proprietà da (M1) a (M4) sono gli *assiomi della metrica.* Il nome "disuguaglianza triangolare" è preso a prestito dalla geometria elementare.

Un **sottoinsieme** $(Y, \widetilde{d})$ di (X, d) si ottiene prendendo un sottoinsieme $Y \subset X$ e restringendo d a $Y \times Y$; allora la metrica su Y è data dalla restrizione

$$\widetilde{d} = d|_{Y \times Y}.$$

$\widetilde{d}$ si chiama la metrica **indotta** su Y da d.

Example 1. Setting

$$\rho(x, y) = \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \neq y, \end{cases}$$

where x and y are elements of an arbitrary set X, we obviously get a metric space, which might be called a "discrete space" or a "space of isolated points."

Example 2. The set of all real numbers with distance

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

is a metric space, which we denote by R^1.

Example 3. The set of all ordered n-tuples

$$x = (x_1, x_2, \ldots, x_n)$$

of real numbers $x_1, x_2, \ldots, x_n$, with distance

$$\rho(x, y) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2}, \tag{1}$$

is a metric space denoted by R^n and called *n-dimensional Euclidean space* (or simply *Euclidean n-space*). The distance (1) obviously has properties 1) and 2) in Definition 1. Moreover, it is easy to see that (1) satisfies the triangle inequality. In fact, let

$$x = (x_1, x_2, \ldots, x_n), \quad y = (y_1, y_2, \ldots, y_n), \quad z = (z_1, z_2, \ldots, z_n)$$

be three points in R^n, and let

$$a_k = x_k - y_k, \quad b_k = y_k - z_k \qquad (k = 1, \ldots, n).$$

Then the triangle inequality takes the form

$$\sqrt{\sum_{k=1}^{n} (x_k - z_k)^2} \leqslant \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2} + \sqrt{\sum_{k=1}^{n} (y_k - z_k)^2}, \tag{2}$$

or equivalently

$$\sqrt{\sum_{k=1}^{n} (a_k + b_k)^2} \leqslant \sqrt{\sum_{k=1}^{n} a_k^2} + \sqrt{\sum_{k=1}^{n} b_k^2}. \tag{2'}$$

It follows from the *Cauchy-Schwarz inequality*

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leqslant \sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2 \tag{3}$$

(see Problem 2) that

$$\begin{aligned}\sum_{k=1}^{n} (a_k + b_k)^2 &= \sum_{k=1}^{n} a_k^2 + 2\sum_{k=1}^{n} a_k b_k + \sum_{k=1}^{n} b_k^2 \\ &\leqslant \sum_{k=1}^{n} a_k^2 + 2\sqrt{\sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2} + \sum_{k=1}^{n} b_k^2 = \left(\sqrt{\sum_{k=1}^{n} a_k^2} + \sqrt{\sum_{k=1}^{n} b_k^2}\right)^2\end{aligned}$$

Taking square roots, we get (2′) and hence (2).

Example 4. Take the same set of ordered n-tuples $x = (x_1, \ldots, x_n)$ as in the preceding example, but this time define the distance by the function

$$\rho_1(x, y) = \sum_{k=1}^{n} |x_k - y_k|. \tag{4}$$

It is clear that (4) has all three properties of a distance figuring in Definition 1. The corresponding metric space will be denoted by R_1^n.

Example 5. Take the same set as in Examples 3 and 4, but this time define distance between two points $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ by the formula

$$\rho_0(x, y) = \max_{1 \leqslant k \leqslant n} |x_k - y_k|. \tag{5}$$

Then we again get a metric space (verify all three properties of the distance). This space, denoted by R_0^n, is often as useful as the Euclidean space R^n.

Remark. The last three examples show that it is sometimes important to use a different notation for a metric space than for the underlying set of points in the space, since the latter can be "metrized" in a variety of different ways.

Example 6. The set $C_{[a,b]}$ of all continuous functions defined on the closed interval $[a, b]$, with distance

$$\rho(f, g) = \max_{a \leqslant t \leqslant b} |f(t) - g(t)| \tag{6}$$

is a metric space of great importance in analysis (again verify the three properties of distance). This metric space and the underlying set of "points" will both be denoted by the symbol $C_{[a,b]}$. Instead of $C_{[0,1]}$, we will often write just C. A space like $C_{[a,b]}$ is often called a "function space," to emphasize that its elements are functions.

Example 7. Let l_2 be the set of all infinite sequences[2]

$$x = (x_1, x_2, \ldots, x_k, \ldots)$$

of real numbers $x_1, x_2, \ldots, x_k, \ldots$ satisfying the convergence condition

$$\sum_{k=1}^{\infty} x_k^2 < \infty,$$

[2] The infinite sequence with general term x_k can be written as $\{x_k\}$ or simply as $x_1, x_2, \ldots, x_k, \ldots$ (this notation is familiar from calculus). It can also be written in "point notation" as $x = (x_1, x_2, \ldots, x_k, \ldots)$, i.e., as an "ordered ∞-tuple" generalizing the notion of an ordered n-tuple. (In writing $\{x_k\}$ we have another use of curly brackets, but the context will always prevent any confusion between the *sequence* $\{x_k\}$ and the set whose only element is x_k.)

where distance between points is defined by

$$\rho(x, y) = \sqrt{\sum_{k=1}^{\infty}(x_k - y_k)^2}. \tag{7}$$

Clearly (7) makes sense for all $x, y \in l_2$, since it follows from the elementary inequality

$$(x_k \pm y_k)^2 \leqslant 2(x_k^2 + y_k^2)$$

that convergence of the two series

$$\sum_{k=1}^{\infty} x_k^2, \qquad \sum_{k=1}^{\infty} y_k^2$$

implies that of the series

$$\sum_{k=1}^{\infty}(x_k - y_k)^2.$$

At the same time, we find that if the points $(x_1, x_2, \ldots, x_k, \ldots)$ and $(y_1, y_2, \ldots, y_k, \ldots)$ both belong to l_2, then so does the point

$$(x_1 + y_1, x_2 + y_2, \ldots, x_k + y_k, \ldots).$$

The function (7) obviously has the first two defining properties of a distance. To verify the triangle inequality, which takes the form

$$\sqrt{\sum_{k=1}^{\infty}(x_k - z_k)^2} \leqslant \sqrt{\sum_{k=1}^{\infty}(x_k - y_k)^2} + \sqrt{\sum_{k=1}^{\infty}(y_k - z_k)^2} \tag{8}$$

for the metric (7), we first note that all three series converge, for the reason just given. Moreover, the inequality

$$\sqrt{\sum_{k=1}^{n}(x_k - z_k)^2} \leqslant \sqrt{\sum_{k=1}^{n}(x_k - y_k)^2} + \sqrt{\sum_{k=1}^{n}(y_k - z_k)^2} \tag{9}$$

holds for all n, as shown in Example 3. Taking the limit as $n \to \infty$ in (9), we get (8), thereby verifying the triangle inequality in l_2. Therefore l_2 is a metric space.

Example 8. As in Example 6, consider the set of all functions continuous on the interval $[a, b]$, but this time define distance by the formula

$$\rho(x, y) = \left(\int_a^b [x(t) - y(t)]^2\, dt\right)^{1/2}, \tag{10}$$

instead of (6). The resulting metric space will be denoted by $C^2_{[a,b]}$. The first two properties of the metric are obvious, and the fact that (10) satisfies the triangle inequality is an immediate consequence of *Schwarz's inequality*

$$\left(\int_a^b x(t)y(t)\, dt\right)^2 \leqslant \int_a^b x^2(t)\, dt \int_a^b y^2(t)\, dt \tag{11}$$

(see Problem 3), by the continuous analogue of the argument given in Example 3.

Example 9. Next consider the set of all *bounded* infinite sequences of real numbers $x = (x_1, x_2, \ldots, x_k, \ldots)$, and let[3]

$$\rho(x, y) = \sup_k |x_k - y_k|. \tag{12}$$

This gives a metric space which we denote by m. The fact that (12) has the three properties of a metric is almost obvious.

2.6 Esempio. Consideriamo ancora $X = \mathbb{R}$ con una distanza diversa da quella usuale. Poniamo

$$\rho(x, y) = \frac{|x - y|}{1 + |x - y|}\,.$$

La funzione ρ soddisfa evidentemente le proprietà (SM1)–(SM3). Per verificare la (SM4) considero la funzione $g : [0, \infty) \to [0, \infty)$ definita come

$$g(t) := \frac{t}{1 + t}\,.$$

La funzione g è crescente, infatti

$$g'(t) = \frac{(1 + t) - t}{(1 + t)^2} = \frac{1}{(1 + t)^2} > 0\,.$$

Quindi se $0 \leq t \leq u$ si ha $g(t) \leq g(u)$. D'altra parte sappiamo già che

$$|x - z| \leq |x - y| + |y - z|\,,$$

per cui

$$g(|x - z|) \leq g(|x - y| + |y - z|)\,,$$

che, per definizione di g, significa

$$\frac{|x - z|}{1 + |x - z|} \leq \frac{|x - y| + |y - z|}{1 + |x - y| + |y - z|}\,.$$

Possiamo quindi scrivere

$$\begin{aligned}
\rho(x, z) = \frac{|x - z|}{1 + |x - z|} &\leq \frac{|x - y| + |y - z|}{1 + |x - y| + |y - z|} \\
&= \frac{|x - y|}{1 + |x - y| + |y - z|} + \frac{|y - z|}{1 + |x - y| + |y - z|} \\
&\leq \frac{|x - y|}{1 + |x - y|} + \frac{|y - z|}{1 + |y - z|} = \rho(x, y) + \rho(y, z)\,.
\end{aligned}$$

La disuguaglianza triangolare è verificata, quindi ρ è una metrica su $\mathbb{R}$. Osservo che mentre lo spazio metrico $(\mathbb{R}, d)$ (quello con la metrica usuale) è illimitato, lo stesso insieme con la distanza ρ diventa uno spazio metrico limitato. In particolare il diametro di $(\mathbb{R}, \rho)$ è uguale a 1 (dimostrarlo).

$\boxed{1.1}$ Insiemi Aperti. Insiemi Chiusi. Intorni.

V'è un considerevole numero di concetti ausiliari che giocano un ruolo in connessione con gli spazi metrici. Quelli di cui avremo bisogno sono inclusi in questa sezione. Perciò questa sezione contiene molti concetti, ma il lettore noterà che molti di loro divengono familiari quando vengono applicati agli spazi euclidei.

Consideriamo dapprima alcuni importanti sottoinsiemi di un dato spazio metrico $X = (X, d)$.

1.2 Definizione (Palla e Sfera)
Dato un punto $x_0 \in X$ ed un numero reale $r > 0$ definiamo tre tipi di insiemi

(a) $B(x_0; r) = \{x \in X | d(x, x_0) < r\}$ **(Palla Aperta)**

(b) $\widetilde{B}(x_0; r) = \{x \in X | d(x, x_0) \leq r\}$ **(Palla Chiusa)**

(c) $S(x_0; r) = \{x \in X | d(x, x_0) = r\}$ **(Sfera)**

■

In tutti e tre i casi x_0 è chiamato il *centro* ed r il *raggio.*

Attenzione. Lavorando con gli spazi metrici è assai utile utilizzare una terminologia analoga a quella della geometria Euclidea. Tuttavia bisogna essere coscienti di quanto sia pericoloso assumere che palle e sfere in uno spazio metrico arbitrario soddisfino alle medesime proprietà soddisfatte da palle e sfere in $\mathbb{R}^3$. Ad esempio una possibile proprietà inusuale è che una sfera può essere vuota. Un'altra possibile proprietà inusuale sarà citata più in là.

1.3 Definizione (Insiemi Aperti, Insiemi Chiusi)
Un sottoinsieme M di uno spazio metrico X è detto aperto se contiene una palla centrata in ciascuno dei suoi punti. Un sottoinsieme K è detto chiuso se il suo complemento (in X) è aperto, cioè se $K^C = X - K$ è aperto. ■

La dizione aperta e chiusa utilizzata nella definizione precedente è coerente con questa definizione di insieme aperto e chiuso. Infatti, utilizzando la diseguaglianza triangolare, è facile vedere che per ogni punto $x_1 \in B(x_0; r)$ la palla $B(x_1; r_1)$ centrata in x_1 e di raggio $r_1 = r - d(x_1, x_{0)}$ è contenuta in $B(x_0; r)$ e quindi $B(x_0; r)$ è un insieme aperto. Che $\widetilde{B}(x_0; r)$ sia chiuso si dimostra per assurdo. Sia infatti per assurdo il complemento di $\widetilde{B}(x_0; r)$ non aperto. Allora esiste ameno un punto x_1 esterno a $\widetilde{B}(x_0; r)$, ossia per cui $d(x_1, x_0) > r$, tale che ogni palla centrata in x_1 contiene almeno un punto $x \in \widetilde{B}(x_0; r)$. Consideriamo ora la palla $B(x_1; r_1)$ centrata in x_1 di raggio $r_1 = d(x_1, x_0) - r$. Per l' x comune a $\widetilde{B}(x_0; r)$ e $B(x_1; r_1)$ abbiamo $d(x_1, x_0) \leq d(x_1, x) + d(x, x_0) < r_1 + r = d(x_1, x_0)$, che è impossibile.

Una palla aperta $B(x_0; \varepsilon)$ di raggio ε è spesso chiamata un *ε–intorno* di x_0. Per un **intorno** di x_0 si intende un qualunque sottoinsieme di X che contiene un ε–intorno di x_0.

Vediamo direttamente dalla definizione che ciascun intorno di x_0 contiene x_0; in altre parole x_0 è un punto di ciascuno dei suoi intorni. Se N è un intorno di x_0 e $N \subset M$, allora anche M è un intorno di x_0.

Chiamiamo x_0 un **punto interno** di un insieme $M \subset X$ se M è un intorno di x_0. L'**interno** di M è l'insieme di tutti i punti interni a M e può essere indicato con M^0 o con $\text{Int}(M)$. $\text{Int}(M)$ è aperto ed è l'insieme aperto più grande contenuto in M.

Non è difficile mostrare che la collezione di tutti i sottoinsiemi aperti di X, che possiamo chiamare $\mathcal{T}$, ha le seguenti proprietà

T1 $\emptyset \in \mathcal{T}$, $X \in \mathcal{T}$.

T2 L'unione di membri di $\mathcal{T}$ è un membro di $\mathcal{T}$.

T3 L'intersezione di un numero finito di membri di $\mathcal{T}$ è un membro di $\mathcal{T}$.

Dimostrazione. (T1) segue dall'osservazione che $\emptyset$ non ha elementi e, ovviamente, X è aperto. Proviamo (T2). Un punto qualunque x dell'unione U degli insiemi aperti appartiene ad (almeno) uno di questi insiemi, sia M, ed M contiene una palla B di x poiché M è aperto. Allora $B \subset U$ per definizione di unione. Ciò prova (T2). Infine se y è un punto qualunque dell'intersezione degli insiemi aperti $M_1, \cdots, M_n$ allora ciascun M_j contiene una palla di y e la più piccola di queste palle è contenuta nell'intersezione. Ciò prova (T3). ∎

Osserviamo che le proprietà da (T1) a (T3) sono così fondamentali che vengono di norma inserite in un contesto più generale. Precisamente si definisce come uno **spazio topologico** $(X, \mathcal{T})$ un insieme X ed una collezione $\mathcal{T}$ che soddisfa gli *assiomi* da (T1) a (T3). L'insieme $\mathcal{T}$ è chiamato una *topologia per* X. Da questa definizione segue che

Uno spazio metrico è uno spazio topologico.

Gli insiemi aperti giocano anche un ruolo in connessione con le applicazioni continue, dove la continuità è una naturale generalizzazione della continuità conosciuta dall'analisi ed è definita come segue.

1.4 Definizione (Applicazione Continua)
Siano $X = (X, d)$ e $Y = (Y, \widetilde{d})$ due spazi metrici. Un'applicazione $T : X \to Y$ è detta *continua nel punto* $x_0 \in X$ se per ogni $\varepsilon > 0$ v'è un $\delta > 0$ tale che

$$\widetilde{d}(Tx, Tx_0) < \varepsilon \qquad \text{per ogni } x \text{ che soddisfa a} \qquad d(x, x_0) < \delta.$$

T è detta *continua* se è continua in ogni punto di X. ∎

È importante ed interessante che le applicazioni continue possano essere caratterizzate in termini di insiemi aperti come segue.

1.5 Teorema (Applicazioni Continue)

Un'applicazione T di uno spazio metrico X in uno spazio metrico Y è continua se e solo se l'immagine inversa di un qualunque sottoinsieme aperto di Y è un sottoinsieme aperto di X. ∎

Dimostrazione. **(a)** Supponiamo che T sia continua. Sia $S \subset Y$ aperto ed S_0 l'immagine inversa di S. Se $S_0 = \emptyset$ è aperto. Sia $S_0 \neq \emptyset$. Per un qualunque $x_0 \in S_0$ sia $y_0 = Tx_0$. Poiché S è aperto contiene un ε–intorno N di y_0. Poiché T è continua x_0 ha un δ–intorno N_0 che è applicato in N. Poiché $N \subset S$ abbiamo che $N_0 \subset S_0$ così che S_0 è aperto perché $x_0 \in S_0$ era arbitrario.

(b) Viceversa assumiamo che l'immagine inversa di ogni insieme aperto in Y sia un insieme aperto in X. Allora per ogni $x_0 \in X$ e qualunque ε–intorno N di Tx_0 l'immagine inversa N_0 di N è aperta, poiché N è aperto, e N_0 contiene x_0. Quindi anche N_0 contiene un δ–intorno di x_0 che è applicato in N poiché N_0 è applicato in N. Di conseguenza, per definizione, T è continua in x_0. Poiché $x_0 \in X$ era arbitrario ne segue che T è continua. ∎

Introduciamo ora due altri concetti che sono collegati. Sia M un sottoinsieme di uno spazio metrico X. Allora un punto x_0 di X (che può o può non essere un punto di M) è chiamato un **punto di accumulazione** di M (o *punto limite di* M) se ogni intorno di x_0 contiene almeno un punto $y \in M$ distinto da x_0. L'insieme costituito dai punti di M e dai punti di accumulazione di M è chiamato la **chiusura** *di* M ed è indicato con

$$\overline{M}.$$

L'insieme $\overline{M}$ è chiuso, perché se non lo fosse il suo complemento non sarebbe aperto e conterebbe quindi almeno un punto di accumulazione di M. Inoltre, allo stesso modo, si mostra che $\overline{M}$ è il più piccolo insieme chiuso che contiene M.

Prime di procedere menzioniamo un'altra proprietà inusuale delle palle in uno spazio metrico. Mentre in $\mathbb{R}^3$ la chiusura $\overline{B(x_0;r)}$ di una palla aperta $B(x_0;r)$ è la palla chiusa $\widetilde{B}(x_0;r)$, in un generico spazio metrico ciò può non essere valido.

Usando il concetto di chiusura vogliamo dare una definizione che risulterà di particolare importanza nel seguito.

1.6 Definizione (Insieme Denso, Spazio Separabile)
Un sottoinsieme M di uno spazio metrico X è detto *denso in* X se

$$\overline{M} = X.$$

X è detto *separabile* se ha un sottoinsieme numerabile che è denso in X. ∎

Quindi se M è denso in X ogni palla in X, per quanto piccola, conterrà punti di M; o, in altre parole, non c'è punto $x \in X$ che abbia un intorno che non contiene punti di M.

Vedremo nel seguito che gli spazi metrici separabili sono alquanto più semplici di quelli non separabili.

1.2 Convergenza. Successioni di Cauchy. Completezza.

Sappiamo che le successioni di numeri reali giocano un ruolo importante in analisi ed è la metrica di $\mathbb{R}$ che permette di definire il concetto basilare di convergenza di una tale successione. Lo stesso vale per le successioni di numeri complessi; in questo caso dobbiamo usare la metrica del piano complesso. In uno spazio metrico arbitrario $X = (X, d)$ la situazione è assai simile, cioè possiamo considerare una successione (x_n) di elementi $x_1, x_2, \cdots$ di X ed usare la metrica d per definire la convergenza in maniera analoga a quella dell'analisi.

1.7 Definizione (Convergenza di una Successione, Limite)
Una successione (x_n) in uno spazio metrico $X = (X, d)$ è detta *convergere* od *essere convergente* se v'è un $x \in X$ tale che

$$\lim_{n\to\infty} d(x_n, x) = 0.$$

x è chiamato il *limite* di (x_n) e si scrive

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

o semplicemente

$$x_n \to x.$$

Diciamo che (x_n) *converge a* x o *ammette il limite* x. Se (x_n) non è convergente si dice che è *divergente.* ∎

Come è stata usata la metrica d in questa definizione? d ha fornito la successione di numeri reali $a_n = d(x_n, x)$ la cui convergenza definisce quella di (x_n). Quindi se $x_n \to x$, dato un $\varepsilon > 0$, esiste un $N = N(\varepsilon)$ tale che tutti gli x_n con $n > N$ giacciono in un ε–intorno $B(x; \varepsilon)$ di x.

Per evitare incomprensioni osserviamo che il limite di una successione convergente deve essere un punto dello spazio X.

Mostriamo ora che due proprietà delle successioni convergenti (unicità del limite e limitatezza), che risultano familiari dall'analisi, si mantengono in questo contesto molto più generale.

Chiamiamo un sottoinsieme non vuoto $M \subset X$ un *insieme limitato* se il suo *diametro*

$$\delta(M) = \sup_{x,y \in M} d(x,y)$$

è finito. Chiamiamo una successione (x_n) in X una **successione limitata** se l'insieme dei suoi punti è un sottoinsieme limitato di X.

Ovviamente se M è limitato allora $M \subset B(x_0; r)$, dove $x_0 \in X$ è un qualunque punto ed r è un numero reale (sufficientemente grande) e viceversa.

La nostra asserzione è allora formulata come segue.

1.8 Lemma (Limitatezza, Limite)

Sia $X = (X, d)$ uno spazio metrico. Allora

(a) *Una successione convergente in X è limitata ed il suo limite è unico.*

(b) *Se $x_n \to x$ e $y_m \to y$ in X, allora $d(x_n, y_m) \to d(x, y)$.* ∎

Dimostrazione. **(a)** Supponiamo che $x_n \to x$. Allora prendendo $\varepsilon = 1$ possiamo trovare un N tale che $d(x_n, x) < 1$ per tutti gli $n > N$. Quindi per tutti gli n abbiamo che $d(x_n, x) < 1 + a$ dove

$$a = \max\{d(x_1, x), \cdots, d(x_N, x)\}.$$

Ciò mostra che (x_n) è limitata. Assumendo che $x_n \to x$ e che $\mathrm{x}_n \to z$ abbiamo dalla (M4)

$$0 \leq d(x, z) \leq d(x, x_n) + d(x_n, z) \to 0 + 0$$

e l'unicità $x = z$ del limite segue dalla (M2).

(b) Dalla (M4) abbiamo che

$$d(x_n, y_m) \leq d(x_n, x) + d(x, y) + d(y, y_m).$$

Da cui otteniamo

$$d(x_n, y_m) - d(x, y) \leq d(x_n, x) + d(y_m, y)$$

ed una diseguaglianza simile scambiando x_n con x e y_m con y e moltiplicando per -1. Assieme forniscono

$$|d(x_n, y_m) - d(x, y)| \leq d(x_n, x) + d(y_m, y) \to 0$$

per $n, m \to \infty$. ∎

Definiremo ora il concetto di completezza di uno spazio metrico, che risulterà basilare nel seguito. La completezza *non* segue dagli assiomi (M1) sino a (M4), poiché vi sono spazi metrici *incompleti.* In altre parole, la completezza è una proprietà addizionale che gli spazi metrici possono avere o non avere. Essa ha varie conseguenze che rendono gli spazi metrici completi "migliori e più semplici" di quelli incompleti.

Ricordiamo dapprima dall'analisi che una successione (x_n) di numeri reali o complessi converge sulla retta reale $\mathbb{R}$ o nel piano complesso $\mathbb{C}$ se e solamente se soddisfa il *criterio di convergenza di Cauchy,* cioè se e solo se per ogni $\varepsilon > 0$ v'è un $N = N(\varepsilon)$ tale che

$$|x_m - x_n| < \varepsilon \qquad \text{per tutti gli } m, n > N.$$

Qui $|x_m - x_n|$ è la distanza $d(x_m, x_n)$ da x_m a x_n sulla retta reale $\mathbb{R}$ o nel piano complesso $\mathbb{C}$. Quindi possiamo scrivere la diseguaglianza del criterio di Cauchy nella forma

$$d(x_m, x_n) < \varepsilon \qquad (\ m, n > N).$$

Se una successione (x_n) soddisfa alla condizione del criterio di Cauchy potremo chiamarla una *successione di Cauchy.* Allora il criterio di Cauchy dice semplicemente che una successione di numeri reali o complessi converge in $\mathbb{R}$ o $\mathbb{C}$ se e solamente se è una successione di Cauchy. Sfortunatamente in spazi più generali la situazione può essere più complicata e vi possono essere successioni di Cauchy che non convergono.

1.9 Definizione (Successione di Cauchy, Completezza)
Una successione (x_n) in uno spazio metrico $X = (X, d)$ è detta di *Cauchy* (o *fondamentale*) se per ogni $\varepsilon > 0$ v'è un $N = N(\varepsilon)$ tale che

$$d(x_m, x_n) < \varepsilon \qquad \text{per ogni } m, n > N. \tag{1.1}$$

Lo spazio X è detto *completo* se ogni successione di Cauchy in X converge (cioè se ha un limite che è un elemento di X). ■

A prescindere dalla completezza o meno di X la condizione (1.1) è necessaria per la convergenza di una successione. Infatti si ottiene facilmente il risultato seguente.

1.10 Teorema (Successione Convergente)

Ogni successione convergente in uno spazio metrico è una successione di Cauchy. ■

Dimostrazione. Se $x_n \to x$ allora per ogni $\varepsilon > 0$ v'è un $N = N(\varepsilon)$ tale che

$$d(x_n, x) < \frac{\varepsilon}{2} \qquad \text{per tutti gli } n > N.$$

Quindi dalla disuguaglianza triangolare otteniamo per $m, n > N$

$$d(x_m, x_n) \leq d(x_m, x) + d(x, x_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Ciò mostra che (x_n) è di Cauchy. ■

Se lo spazio X è completo la condizione di Cauchy (1.1) diventa necessaria e sufficiente per la convergenza e si parla di criterio di Cauchy per la convergenza. Il teorema che afferma la validità del criterio di Cauchy in $\mathbb{R}$ e in $\mathbb{C}$ può essere riespresso in termini di completezza come segue.

1.11 Teorema (Retta Reale, Piano Complesso)

La retta reale ed il piano complesso sono spazi metrici completi. ■

Example 1. The set of all rational points is dense in the real line R^1.

Example 2. The set of all points $x = (x_1, x_2, \ldots, x_n)$ with rational coordinates is dense in each of the spaces R^n, R_0^n and R_1^n introduced in Examples 3–5, pp. 38–39.

Example 3. The set of all points $x = (x_1, x_2, \ldots, x_k, \ldots)$ with only finitely many nonzero coordinates, each a rational number, is dense in the space l_2 introduced in Example 7, p. 39.

Example 4. The set of all polynomials with rational coefficients is dense in both spaces $C_{[a,b]}$ and $C^2_{[a,b]}$ introduced in Examples 6 and 8, pp. 39 and 40.

DEFINITION. *A metric space is said to be* ***separable*** *if it has a countable everywhere dense subset.*

Example 5. The spaces R^1, R^n, R_0^n, R_1^n, l_2, $C_{[a,b]}$, and $C^2_{[a,b]}$ are all separable, since the sets in Examples 1–4 above are all countable.

Example 6. The "discrete space" M described in Example 1, p. 38 contains a countable everywhere dense subset and hence is separable if and only if it is itself a countable set, since clearly $[M] = M$ in this case.

Example 7. There is no countable everywhere dense set in the space m of all bounded sequences, introduced in Example 9, p. 41. In fact, consider the set E of all sequences consisting exclusively of zeros and ones. Clearly, E has the power of the continuum (recall Theorem 6, Sec. 2.5), since there is a one-to-one correspondence between E and the set of all subsets of the set $Z_+ = \{1, 2, \ldots, n, \ldots\}$ (describe the correspondence). According to formula (12), p. 41, the distance between any two points of E equals 1. Suppose we surround each point of E by an open sphere of radius $\frac{1}{2}$, thereby obtaining an uncountably infinite family of pairwise disjoint spheres. Then if some set M is everywhere dense in m, there must be at least one point of M in each of the spheres. It follows that M cannot be countable and hence that m cannot be separable.

6.4. Closed sets. We say that a subset M of a metric space R is *closed* if it coincides with its own closure, i.e., if $[M] = M$. In other words, a set is called closed if it contains all its limit points (see Problem 2).

Example 1. The empty set $\varnothing$ and the whole space R are closed sets.

Example 2. Every closed interval $[a, b]$ on the real line is a closed set.

Example 3. Every closed sphere in a metric space is a closed set. In particular, the set of all functions f in the space $C_{[a,b]}$ such that $|f(t)| \leqslant K$ (where K is a constant) is closed.

Example 4. The set of all functions f in $C_{[a,b]}$ such that $|f(t)| < K$ (an open sphere) is not closed. The closure of this set is the closed sphere in the preceding example.

Example 5. Any set consisting of a finite number of points is closed.

The first assertion holds for open sets in R^n (and in fact is susceptible to further generalizations), while the second assertion pertains specifically to the real line.

7. Complete Metric Spaces

7.1. Definitions and examples. The reader is presumably already familiar with the notion of the completeness of the real line. The real line is, of course, a particularly simple example of a metric space. We now make the natural generalization of the notion of completeness to the case of an arbitrary metric space.

DEFINITION 1. *A sequence $\{x_n\}$ of points in a metric space R with metric ρ is said to satisfy the **Cauchy criterion** if, given any $\varepsilon > 0$, there is an integer N_ε such that $\rho(x_n, x_{n'}) < \varepsilon$ for all $n, n' > N_\varepsilon$.*

DEFINITION 2. *A subsequence $\{x_n\}$ of points in a metric space R is called a **Cauchy sequence** (or a **fundamental sequence**) if it satisfies the Cauchy criterion.*

THEOREM 1. *Every convergent sequence $\{x_n\}$ is fundamental.*

Proof. If $\{x_n\}$ converges to a limit x, then, given any $\varepsilon > 0$, there is an integer N_ε such that

$$\rho(x_n, x) < \frac{\varepsilon}{2}$$

for all $n > N_\varepsilon$. But then

$$\rho(x_n, x_{n'}) \leqslant \rho(x_n, x) + \rho(x_{n'}, x) < \varepsilon$$

for all $n, n' > N_\varepsilon$. ∎

DEFINITION 3. *A metric space R is said to be **complete** if every Cauchy sequence in R converges to an element of R. Otherwise R is said to be **incomplete**.*

Example 1. Let R be the "space of isolated points" considered in Example 1, p. 38. Then the Cauchy sequences in R are just the "stationary sequences," i.e., the sequences $\{x_n\}$ all of whose terms are the same starting from some index n. Every such sequence is obviously convergent to an element of R. Hence R is complete.

Example 2. The completeness of the real line R^1 is familiar from elementary analysis.

Example 3. The completeness of Euclidean n-space R^n follows from that of R^1. In fact, let

$$x^{(p)} = (x_1^{(p)}, \ldots, x_n^{(p)}) \qquad (p = 1, 2, \ldots)$$

be a fundamental sequence of points of R^n. Then, given any $\varepsilon > 0$, there exists an N_ε such that

$$\sum_{k=1}^{n} (x_k^{(p)} - x_k^{(q)})^2 < \varepsilon^2$$

for all $p, q > N_\varepsilon$. It follows that

$$|x_k^{(p)} - x_k^{(q)}| < \varepsilon \qquad (k = 1, \ldots, n)$$

for all $p, q > N_\varepsilon$, i.e., each $\{x_k^{(p)}\}$ is a fundamental sequence in R^1. Let

$$x = (x_1, \ldots, x_n),$$

where

$$x_k = \lim_{p \to \infty} x_k^{(p)}.$$

Then obviously

$$\lim_{p \to \infty} x^{(p)} = x.$$

This proves the completeness of R^n. The completeness of the spaces R_0^n and R_1^n introduced in Examples 4 and 5, p. 39 is proved in almost the same way (give the details).

Example 4. Let $\{x_n(t)\}$ be a Cauchy sequence in the function space $C_{[a,b]}$ considered in Example 6, p. 39. Then, given any $\varepsilon > 0$, there is an N_ε such that

$$|x_n(t) - x_{n'}(t)| < \varepsilon \tag{1}$$

for all $n, n' > N_\varepsilon$ and all $t \in [a, b]$. It follows that the sequence $\{x_n(t)\}$ is uniformly convergent. But the limit of a uniformly convergent sequence of continuous functions is itself a continuous function (see Problem 1). Taking the limit as $n' \to \infty$ in (1), we find that

$$|x_n(t) - x(t)| \leqslant \varepsilon$$

for all $n > N_\varepsilon$ and all $t \in [a, b]$, i.e., $\{x_n(t)\}$ converges in the metric of $C_{[a,b]}$ to a function $x(t) \in C_{[a,b]}$. Hence $C_{[a,b]}$ is a complete metric space.

Example 5. Next let $x^{(n)}$ be a sequence in the space l_2 considered in Example 7, p. 39, so that

$$x^{(n)} = (x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)}, \ldots),$$

$$\sum_{k=1}^{\infty} (x_k^{(n)})^2 < \infty \qquad (n = 1, 2, \ldots).$$

Suppose further that $\{x^{(n)}\}$ is a Cauchy sequence. Then, given any $\varepsilon > 0$, there is a N_ε such that

$$\rho^2(x^{(n)}, x^{(n')}) = \sum_{k=1}^{\infty}(x_k^{(n)} - x_k^{(n')})^2 < \varepsilon \tag{2}$$

if $n, n' > N_\varepsilon$. It follows that

$$(x_k^{(n)} - x_k^{(n')})^2 < \varepsilon \qquad (k = 1, 2, \ldots),$$

i.e., for every k the sequence $\{x_k^{(n)}\}$ is fundamental and hence convergent. Let

$$x_k = \lim_{n\to\infty} x_k^{(n)},$$

$$x = (x_1, x_2, \ldots, x_k, \ldots).$$

Then, as we now show, x is itself a point of l_2 and moreover $\{x^{(n)}\}$ converges to x in the l_2 metric, so that l_2 is a complete metric space.

In fact, (2) implies

$$\sum_{k=1}^{M}(x_k^{(n)} - x_k^{(n')})^2 < \varepsilon \tag{3}$$

for any fixed M. Holding n fixed in (3) and taking the limit as $n' \to \infty$, we get

$$\sum_{k=1}^{M}(x_k^{(n)} - x_k)^2 \leqslant \varepsilon. \tag{4}$$

Since (4) holds for arbitrary M, we can in turn take the limit of (4) as $M \to \infty$, obtaining

$$\sum_{k=1}^{\infty}(x_k^{(n)} - x_k)^2 \leqslant \varepsilon. \tag{5}$$

Just as on p. 40, the convergence of the two series

$$\sum_{k=1}^{\infty}(x_k^{(n)})^2, \qquad \sum_{k=1}^{\infty}(x_k^{(n)} - x_k)^2$$

implies that of the series

$$\sum_{k=1}^{\infty} x_k^2.$$

This proves that $x \in l_2$. Moreover, since ε is arbitrarily small, (5) implies

$$\lim_{n\to\infty} \rho(x^{(n)}, x) = \lim_{n\to\infty} \sqrt{\sum_{k=1}^{\infty}(x_k^{(n)} - x_k)^2} = 0,$$

i.e., $\{x^{(n)}\}$ converges to x in the l_2 metric, as asserted.

Example 6. It is easy to show that the space $C^2_{[a,b]}$ of Example 8, p. 40 is *incomplete*. If

$$\varphi_n(t) = \begin{cases} -1 & \text{if } -1 \leqslant t \leqslant -\dfrac{1}{n}, \\[2ex] nt & \text{if } -\dfrac{1}{n} \leqslant t \leqslant \dfrac{1}{n}, \\[2ex] 1 & \text{if } \dfrac{1}{n} \leqslant t \leqslant 1, \end{cases}$$

then $\{\varphi_n(t)\}$ is a fundamental sequence in $C^2_{[-1,1]}$, since

$$\int_{-1}^{1}[\varphi_n(t) - \varphi_{n'}(t)]^2\,dt \leqslant \frac{2}{\min\{n, n'\}}$$

However, $\{\varphi_n(t)\}$ cannot converge to a function in $C^2_{[-1,1]}$. In fact, consider the discontinuous function

$$\psi(t) = \begin{cases} -1 & \text{if } t < 0, \\ 1 & \text{if } t \geqslant 0. \end{cases}$$

Then, given any function $f \in C^2_{[-1,1]}$, it follows from Schwarz's inequality (obviously still valid for piecewise continuous functions) that

$$\left(\int_{-1}^{1}[f(t) - \psi(t)]^2\right)^{1/2} \leqslant \left(\int_{-1}^{1}[f(t) - \varphi_n(t)]^2\,dt\right)^{1/2} + \left(\int_{-1}^{1}[\varphi_n(t) - \psi(t)]^2\,dt\right)^{1/2}.$$

But the integral on the left is nonzero, by the continuity of f, and moreover it is clear that

$$\lim_{n\to\infty} \int_{-1}^{1}[\varphi_n(t) - \psi(t)]^2\,dt = 0.$$

Therefore

$$\int_{-1}^{1}[f(t) - \varphi_n(t)]^2\,dt$$

cannot converge to zero as $n \to \infty$.

7.2. The nested sphere theorem. A sequence of closed spheres

$$S[x_1, r_1],\ S[x_2, r_2], \ldots, S[x_n, r_n], \ldots$$

in a metric space R is said to be *nested* (or *decreasing*) if

$$S[x_1, r_1] \supset S[x_2, r_2] \supset \cdots \supset S[x_n, r_n] \supset \cdots.$$

Using this concept, we can prove a simple criterion for the completeness of R:

Completiamo questa sezione con tre teoremi che sono legati alla convergenza e alla completezza e che saranno necessari nel seguito.

1.12 Teorema (Chiusura, Insieme Chiuso)

Sia M un sottoinsieme non vuoto di uno spazio metrico (X, d) e $\overline{M}$ la sua chiusura così come definita nella sezione precedente. Allora

(a) $x \in \overline{M}$ *se e solo se esiste una successione* (x_n) *in* M *tale che* $x_n \to x$.

(b) M *è chiuso se e solo se ogni successione convergente* (x_n) *di punti di* M *converge ad un punto di* M*, ossia se e solo se per ogni successione di punti* $x_n \in M$ *tale che* $x_n \to x$ *è* $x \in M$. ∎

Dimostrazione. **(a)** Sia $x \in \overline{M}$. Se $x \in M$ una successione di questo tipo è $(x, x, \cdots)$. Se $x \notin M$ è un punto di accumulazione di M. Quindi per ogni $n = 1, 2, \cdots$ la palla $B(x; 1/n)$ contiene un $x_n \in M$ e $x_n \to x$ perché $1/n \to 0$ per $n \to \infty$.

Viceversa se (x_n) è in M e $x_n \to x$ allora se $x \in M$ non v'è nulla da dimostrare. Se $x \notin M$ ogni intorno di x contiene punti $x_n \neq x$, così che $x \in \overline{M}$ per definizione di chiusura.

(b) M è chiuso se e solo se $M = \overline{M}$ così che *(b)* segue facilmente da *(a).* ∎

1.13 Teorema (Sottospazio Completo)

Un sottospazio M di uno spazio metrico completo X è esso stesso completo se e solo se l'insieme M è chiuso in X. ∎

Dimostrazione. Sia M completo. Grazie a 1.12(a) per ogni $x \in \overline{M}$ v'è una successione (x_n) in M che converge in M, il limite essendo unico per l' 1.8. Quindi $x \in M$. Questo prova che M è chiuso perché $x \in \overline{M}$ era arbitrario.

Viceversa sia M chiuso e (x_n) di Cauchy in M. Allora $x_n \to x \in X$ ciò che implica $x \in \overline{M}$ per l'1.12(a) e $x \in M$ poiché $M = \overline{M}$ per assunzione. Quindi la successione arbitraria di Cauchy (x_n) converge in M ciò che prova la completezza di M. ∎

Questo teorema è molto utile e ne avremo molto spesso bisogno nel seguito.

L'ultimo dei tre teoremi annunciati mostra l'importanza della convergenza delle successioni in connessione con la continuità di un'applicazione.

1.14 Teorema (Applicazione Continua)

Un'applicazione $T : X \to Y$ di uno spazio metrico (X, d) in uno spazio metrico $(Y, \widetilde{d})$ è continua in un punto $x_0 \in X$ se e solo se

$$x_n \to x_0 \qquad \textit{implica} \qquad Tx_n \to Tx_0. \qquad \blacksquare$$

Dimostrazione. Si assuma che T sia continua in x_0. Allora per un dato $\varepsilon > 0$ v'è un $\delta > 0$ tale che

$$d(x, x_0) < \delta \qquad \text{implica} \qquad \widetilde{d}(Tx, Tx_0) < \varepsilon.$$

Sia $x_n \to x_0$. Allora v'è un N tale che per ogni $n > N$ sia

$$d(x_n, x_0) < \delta.$$

Quindi per ogni $n > N$

$$\widetilde{d}(Tx_n, Tx_0) < \varepsilon.$$

Per definizione ciò significa che $Tx_n \to Tx_0$.

Viceversa assumiamo che

$$x_n \to x_0 \qquad \text{implichi} \qquad Tx_n \to Tx_0$$

e proviamo che allora T è continua in x_0. Supponiamo che ciò sia falso. Allora v'è un $\varepsilon > 0$ tale che per ogni $\delta > 0$ v'è un $x \neq x_0$ che soddisfa a

$$d(x, x_0) < \delta \qquad \text{ma tale che} \qquad \widetilde{d}(Tx, Tx_0) \geq \varepsilon.$$

In particolare per $\delta = 1/n$ v'è un x_n che soddisfa a

$$d(x_n, x_0) < \tfrac{1}{n} \qquad \text{ma tale che} \qquad \widetilde{d}(Tx_n, Tx_0) \geq \varepsilon.$$

Chiaramente $x_n \to x$ ma (Tx_n) non converge a Tx_0. Ciò contraddice $Tx_n \to Tx_0$ e prova il teorema. ∎

In particolare da questo teorema e dal Lemma 1.8 al punto b) segue la seguente proposizione.

1.15 Proposizione (Continuità della distanza)
La distanza $d(x, y)$ in X è continua in x ed in y. ■

1.3 Completamento di uno Spazio metrico

Sappiamo che la retta razionale $\mathbb{Q}$ non è completa ma può essere "allargata" alla retta reale $\mathbb{R}$ che è completa. Questo "completamento" $\mathbb{R}$ di $\mathbb{Q}$ è tale che $\mathbb{Q}$ è denso in $\mathbb{R}$. È molto importante che un arbitrario spazio metrico incompleto possa essere "completato" in una maniera simile. Per una formulazione precisa e conveniente useremo i due concetti seguenti collegati fra loro e che hanno anche diverse altre applicazioni.

1.16 Definizione (Applicazione isometrica, Spazi Isometrici)
Siano $X = (X, d)$ e $\widetilde{X} = (\widetilde{X}, \widetilde{d})$ spazi metrici. Allora

(a) Un'applicazione T di X in $\widetilde{X}$ è detta *isometrica* o una *isometria* se T conserva le distanze, cioè se per ogni $x, y \in X$

$$\widetilde{d}(Tx, Ty) = d(x, y),$$

dove Tx e Ty sono le immagini di x e y, rispettivamente.

(b) Lo spazio X è detto *isometrico* allo spazio $\widetilde{X}$ se esiste un'isometria biiettiva di X su $\widetilde{X}$. Gli spazi X e $\widetilde{X}$ sono allora chiamati *spazi isometrici.* ■

Si noti che una isometria è sempre iniettiva.

Due spazi isometrici possono differire al più per la natura dei loro punti ma sono indistinguibili dal punto di vista della metrica. In uno studio, in cui la natura dei punti non abbia importanza, i due spazi si possono considerare identici — ovvero come due copie del medesimo spazio "astratto".

Possiamo ora formulare e provare il teorema che asserisce che ogni spazio metrico può essere completato. Lo spazio $\widehat{X}$ che appare in questo teorema è chiamato **completamento** dello spazio dato X.

1.17 Teorema (Completamento)
Per ogni spazio metrico $X = (X, d)$ esiste uno spazio metrico completo $\widehat{X} = (\widehat{X}, \widehat{d})$ che ha un sottospazio W che è isometrico a X e che è denso in $\widehat{X}$. Questo spazio $\widehat{X}$ è unico a meno di isometrie, cioè, se $\widetilde{X}$ è un qualunque spazio metrico completo che ha un sottospazio denso $\widetilde{W}$ isometrico a X, allora $\widetilde{X}$ e $\widehat{X}$ sono isometrici. ■

Dimostrazione. La dimostrazione è piuttosto lunga ma diretta. La suddividiamo in quattro passi da *(a)* a *(d).* Costruiamo

(a) $\widehat{X} = (\widehat{X}, \widehat{d})$
(b) un'isometria T di X su W, dove $\overline{W} = \widehat{X}$.

Poi proviamo

(c) la completezza di $\widehat{X}$,
(d) l'unicità di $\widehat{X}$ a meno di isometrie.

Parlando rozzamente possiamo dire che il nostro compito è quello di assegnare dei limiti convenienti a quelle successioni di Cauchy in X che non convergono. Tuttavia non dobbiamo introdurre “troppi” limiti, ma dobbiamo tener conto che certe successioni “alla fine divengono arbitrariamente vicine le une alle altre”. Questa idea intuitiva può essere espressa matematicamente in termini di una conveniente relazione di equivalenza [vedi (1.2) qui di seguito].

(a) *Costruzione di* $\widehat{X} = (\widehat{X}, \widehat{d})$. Siano (x_n) e (x'_n) successioni di Cauchy in X. Definiamo (x_n) *equivalente* a (x'_n) e scriviamo $(x_n) \sim (x'_n)$ se

$$\lim_{n\to\infty} d(x_n, x'_n) = 0. \tag{1.2}$$

Sia $\widehat{X}$ l'insieme delle classi di equivalenza $\widehat{x}, \widehat{y}, \cdots$ di successioni di Cauchy così ottenute. Scriviamo $(x_n) \in \widehat{x}$ per indicare che (x_n) è un membro di $\widehat{x}$ (un *rappresentate* della classe $\widehat{x}$). Poniamo ora

$$\widehat{d}(\widehat{x}, \widehat{y}) = \lim_{n\to\infty} d(x_n, y_n) \tag{1.3}$$

dove $(x_n) \in \widehat{x}$ e $(y_n) \in \widehat{y}$. Mostriamo che questo limite esiste. Abbiamo

$$d(x_n, y_n) \leq d(x_n, x_m) + d(x_m, y_m) + d(y_m, y_n);$$

quindi otteniamo

$$d(x_n, y_n) - d(x_m, y_m) \leq d(x_n, x_m) + d(y_m, y_n)$$

e una disuguaglianza simile con m ed n scambiati. Da entrambe

$$|d(x_n, y_n) - d(x_m, y_m)| \leq d(x_n, x_m) + d(y_m, y_n). \tag{1.4}$$

Poiché (x_n) e (y_n) sono di Cauchy possiamo rendere il membro a destra piccolo a piacere. Ciò implica che il limite in (1.3) esiste perché $\mathbb{R}$ è completo.

Dobbiamo anche mostrare che il limite in (1.3) è indipendente dalla particolare scelta del rappresentante. Infatti se $(x_n) \sim (x'_n)$ e $(y_n) \sim (y'_n)$ allora per la (1.2)

$$|d(x_n, y_n) - d(x'_n, y'_n)| \leq d(x_n, x'_n) + d(y_n, y'_n) \to 0$$

per $n \to \infty$, che implica l'asserzione che

$$\lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(x'_n, y'_n).$$

Proviamo che $\widehat{d}$ in (1.3) è una metrica in $\widehat{X}$. Ovviamente $\widehat{d}$ soddisfa (M1) così come $\widehat{d}(\widehat{x}, \widehat{x}) = 0$ e (M3). Inoltre

$$\widehat{d}(\widehat{x}, \widehat{y}) = 0 \Longrightarrow (x_n) \sim (y_n) \Longrightarrow \widehat{x} = \widehat{y}$$

fornisce (M2) e (M4) per $\widehat{d}$ segue da

$$d(x_n, y_n) \leq d(x_n, z_n) + d(z_n, y_n)$$

per $n \to \infty$.

(b) *Costruzione di un isometria* $T : X \to W \subset \widehat{X}$. A ciascun $b \in X$ associamo la classe $\widehat{b} \in \widehat{X}$ che contiene la successione costante di Cauchy $(b, b, \cdots)$. Ciò definisce un'applicazione

$T : X \to W$ sul sottospazio $W = T(X) \subset \widehat{X}$. L'applicazione T è data da $b \longmapsto \widehat{b} = Tb$, dove $(b, b, \cdots) \in \widehat{b}$. Vediamo che T è un'isometria perché (1.3) diviene semplicemente

$$\widehat{d}(\widehat{b}, \widehat{c}) = d(b, c);$$

qui $\widehat{c}$ è la classe di (y_n) dove $y_n = c$ per tutti gli n. Una qualunque isometria è iniettiva e $T : X \to W$ è surgettiva perché $T(X) = W$. Quindi W e X sono isometrici.

Mostriamo che W è denso in $\widehat{X}$. Consideriamo un qualunque $\widehat{x} \in \widehat{X}$. Sia $(x_n) \in \widehat{x}$. Per ogni $\varepsilon > 0$ v'è un N tale che

$$d(x_n, x_N) < \frac{\varepsilon}{2} \qquad (n > N).$$

Sia $(x_N, x_N, \cdots) \in \widehat{x}_N$. Allora $\widehat{x}_N \in W$. Per la (1.3)

$$\widehat{d}(\widehat{x}, \widehat{x}_N) = \lim_{n\to\infty} d(x_n, x_N) \le \frac{\varepsilon}{2} < \varepsilon.$$

Ciò mostra che ogni ε–intorno dell'arbitrario $\widehat{x} \in \widehat{X}$ contiene un elemento di W. Quindi W è denso in $\widehat{X}$.

(c) *Completezza di* $\widehat{X}$. Sia $(\widehat{x}_n)$ una qualunque successione di Cauchy in $\widehat{X}$. Poiché W è denso in $\widehat{X}$ per ogni $\widehat{x}_n$ v'è un $\widehat{z}_n \in W$ tale che

$$\widehat{d}(\widehat{x}_n, \widehat{z}_n) < \frac{1}{n}. \tag{4}$$

Quindi per la disuguaglianza triangolare

$$\begin{aligned}\widehat{d}(\widehat{z}_m, \widehat{z}_n) &\le \widehat{d}(\widehat{z}_m, \widehat{x}_m) + \widehat{d}(\widehat{x}_m, \widehat{x}_n) + \widehat{d}(\widehat{x}_n, \widehat{z}_n)\\ &< \frac{1}{m} + \widehat{d}(\widehat{x}_m, \widehat{x}_n) + \frac{1}{n}\end{aligned}$$

e ciò è minore di ogni dato $\varepsilon > 0$ per m ed n sufficientemente grandi perché $(\widehat{x}_m)$ è di Cauchy. Quindi $(\widehat{z}_m)$ è di Cauchy. Poiché $T : X \to W$ è isometrica e $\widehat{z}_m \in W$ la successione (z_m) dove $z_m = T^{-1}\widehat{z}_m$ è di Cauchy in X. Sia $\widehat{x} \in \widehat{X}$ la classe a cui (z_m) appartiene. Mostriamo che $\widehat{x}$ è il limite di $(\widehat{x}_n)$. Per la (4)

$$\begin{aligned}\widehat{d}(\widehat{x}_n, \widehat{x}) &\le \widehat{d}(\widehat{x}_n, \widehat{z}_n) + \widehat{d}(\widehat{z}_n, \widehat{x})\\ &< \frac{1}{n} + \widehat{d}(\widehat{z}_n, \widehat{x}).\end{aligned} \tag{1.5}$$

Poiché $(z_n) \in \widehat{x}$ e $\widehat{z}_n \in W$, così che $(z_n, z_n, \cdots) \in \widehat{z}_n$, grazie alla definizione di distanza in $\widehat{X}$ data nella (1.3), la disuguaglianza (1.5) diviene

$$\widehat{d}(\widehat{x}_n, \widehat{x}) < \frac{1}{n} + \lim_{m\to\infty} d(z_n, z_m)$$

ed il membro a destra diviene più piccolo di un qualunque dato $\varepsilon > 0$ per n sufficientemente grandi. Quindi la successione arbitraria di Cauchy $(\widehat{x}_n)$ in $\widehat{X}$ ha il limite $\widehat{x} \in \widehat{X}$ e $\widehat{X}$ è completo.

(d) *Unicità di* $\widehat{X}$ *a meno di isometrie.* Sia $(\widetilde{X}, \widetilde{d})$ un altro spazio metrico completo con un sottospazio $\widetilde{W}$ denso in $\widetilde{X}$ e isometrico a X. Grazie alla proprietà transitiva della isometria W e $\widetilde{W}$ sono isometrici e si può utilizzare questa isometria per definire una applicazione biiettiva T di $\widehat{X}$ in $\widetilde{X}$. Precisamente per ogni $\widehat{x} \in \widehat{X}$ consideriamo una successione $(\widehat{x}_n)$ in

W tale che $\widehat{x}_n \to \widehat{x}$. Se $(\widetilde{x}_n)$ è la successione corrispondente nella isometria in $\widetilde{W}$ vi sarà un $\widetilde{x} \in \widetilde{X}$ tale che $\widetilde{x}_n \to \widetilde{x}$ e poniamo allora $T\widehat{x} = \widetilde{x}$.

Consideriamo ora una coppia $\widehat{x}, \widehat{y} \in \widehat{X}$ e la coppia corrispondente $\widetilde{x}, \widetilde{y} \in \widetilde{X}$ nell'applicazione T. Abbiamo

$$\widehat{d}(\widehat{x}, \widehat{y}) = \lim_{n\to\infty} \widehat{d}(\widehat{x}_n, \widehat{y}_n)$$

e

$$\widetilde{d}(\widetilde{x}, \widetilde{y}) = \lim_{n\to\infty} \widetilde{d}(\widetilde{x}_n, \widetilde{y}_n)$$

dove le successioni appartengono rispettivamente a $\widetilde{W}$ e a W e sono definite come indicato sopra. Poiché W e $\widetilde{W}$ sono isometrici è $\widehat{d}(\widehat{x}_n, \widehat{y}_n) = \widetilde{d}(\widetilde{x}_n, \widetilde{y}_n)$ e le distanze in $\widetilde{X}$ e $\widehat{X}$ sono le medesime. Quindi $\widetilde{X}$ e $\widehat{X}$ sono isometrici. ∎

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

$$\rho(Ax, Ay) \leqslant \alpha\rho(x, y) \tag{1}$$

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]

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

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

$$|f(x_1) - f(x_2)| \leqslant K\,|x_1 - x_2|, \tag{3}$$

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.

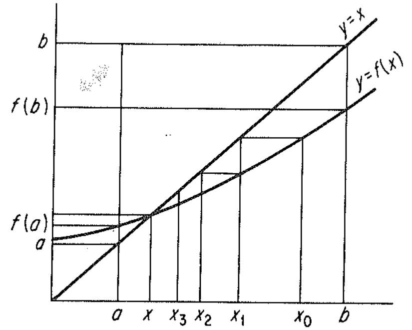

FIGURE 10

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

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

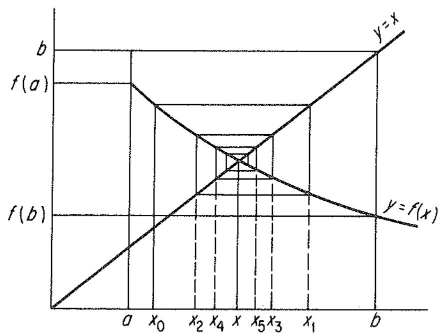

FIGURE 11

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,

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

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

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

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

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

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

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

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^*. ∎

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

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

and in general,

$$|A^nf_1(x) - A^nf_2(x)| \leqslant \lambda^nM^n\frac{(x - a)^n}{n!} \max_x |f_1(x) - f_2(x)|$$
$$\leqslant \lambda^nM^n\frac{(b - a)^n}{n!} \max_x |f_1(x) - f_2(x)|,$$

which implies

$$\rho(A^nf_1, A^nf_2) \leqslant \lambda^nM^n\frac{(b - a)^n}{n!}\rho(f_1, f_2).$$

[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.