The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

1958-1

Summer School and Conference Mathematics, Algorithms and Proofs

11 - 29 August 2008

Lectures on Constructive Mathematics I

Douglas S. Bridges
Dept. of Mathematics, University of Canterbury
Christchurch, New Zealand

Lectures on Constructive Mathematics

Douglas S. Bridges

Department of Mathematics & Statistics,

University of Canterbury,

Christchurch, New Zealand

Lecture 1

Constructive vs nonconstructive

Nonconstructive proof: an existence proof-by-contraction of this schematic form:

> Suppose that the desired object x does not exist.
>
> Derive a contradiction.
>
> Claim that x must exist after all.

This proves that it is impossible for x not to exist; but it does not tell us how to find/compute/construct x.

> "[Nonconstructive existence proofs] inform the world that a treasure exists without disclosing its location." Hermann Weyl

Constructive proof of the existence of an object x: a proof that embodies an algorithm for the construction/computation of the desired object x.

Note: Not all proofs-by-contradiction are nonconstructive. it is perfectly constructive to prove P false by assuming that P is true and deriving a contradiction. This process just captures the constructive meaning of negation.

A nonconstructive proof:

There exists a digit that appears infinitely often in the decimal expansion of the number π.

Note first that the decimal expansion of π is nonterminating and nonrecurring, since π is irrational.

Suppose that each of the digits $0, 1, 2, \ldots, 9$ occurs only finitely many times in the decimal expansion of π.

Then there exists a positive integer N such that each of $0, 1, 2, \ldots, 9$ appears at most N times in the decimal expansion of π.

So that decimal expansion cannot have more than $10N$ places, which contradicts the "Note first ..." above.

Although the decimal expansion of π has been computed to billions of places, the foregoing proof does not tell us (and nobody knows) which of the digits $0, 1, 2, \ldots, 9$ appears infinitely often in the nonterminating, nonrecurring expansion.

All we know is that it is impossible that each of the ten digits appears only a finite number of times.

For another nonconstructive proof, consider the statement:

> There exist irrational numbers a, b such that a^b is rational.

Either $\sqrt{2}^{\sqrt{2}}$ is rational or it is irrational.

In the first case, take $a = b = \sqrt{2}$.

In the second case, take $a = \left(\sqrt{2}^{\sqrt{2}}\right)$ and $b = \sqrt{2}$.

Why is this proof nonconstructive?

1) It does not tell us which of the two alternatives for $\sqrt{2}^{\sqrt{2}}$ (rational or irrational) actually holds.

2) It therefore does not tell us which of the two choices for a and b actually produces irrational numbers with the desired property.

A constructive proof would produce, unambiguously, two irrational numbers a and b and show us that a^b is rational.

Explicit example of irrational numbers a, b such that a^b is rational:

$$a = \sqrt{2}, \quad b = \log_2 9, \quad a^b = 3.$$

In fact, $\sqrt{2}^{\sqrt{2}}$ is transcendental, by the (classical) Gelfand-Schneider theorem:

> *a^b is transcendental if* (i) *a is algebraic,* (ii) *$a \neq 0, 1$ and* (iii) *b is both algebraic and irrational.*

If we want to prove something constructively, then we must not use the law of excluded middle,

LEM: For any proposition P, either P is true or else P is false.

Allowing the use of **LEM** is tantamount to allowing nonconstructive existence proofs.

Historical note: Existence proofs-by-contradiction go back at least as far as Gauss (Fundamental Theorem of Algebra, 1799).

They became dominant after Hilbert's proof of his "basis theorem" (1888).

"Das ist nicht Mathematik. Das ist Theologie." Paul Gordan

The constructive vs nonconstructive controversy goes back at least to Kronecker's attacks on Cantor's set theory (1877–).

It was strengthened by Brouwer's campaign, from 1907 onwards, to convert all mathematicians to the exclusive use of constructive methods, and culminated in the Grundlagenstreit between Brouwer and Hilbert in the 1920s.

> "Taking the principle of excluded middle from the mathematician would be the same, say, as proscribing the telescope to the astronomer or to the boxer the use of his fists." David Hilbert (1928)

What is constructive mathematics?

Two ways to approach computability in mathematics:

1. Use classical logic:
   - allows "decisions" that cannot be made by any real computer
   - requires clearly specified types of algorithm
2. Use intuitionistic logic:
   - automatically takes care of the problem of noncomputational "decisions"
   - enables us to work, with any mathematical objects, in the familiar style of the analyst, algebraist, geometer, ...

Bishop-style constructive mathematics (**BISH**) is just

mathematics with intuitionistic logic

and some appropriate set-theoretic foundation such as the **CST** of Myhill, Aczel, and Rathjen.

Using intuitionistic logic, we can

— clarify distinctions of meaning obscured by classical logic, and

— allow results to have a wider range of interpretations (including recursive ones) than their counterparts proved with classical logic.

> "Intuitionistic logic is richer than classical logic, since the former makes distinctions that the latter fails to make."
>
> J.L. Bell & M. Machover

We

- do not restrict to a class of "constructive/computable objects";
- use intuitionistic logic to deal with the normal objects of mathematics.

Ishihara's classification:

- ▷ a constructive theory of real numbers: the usual $\mathbb{R}$ studied with intuitionistic logic.

- ▷ a theory of constructive real numbers: the recursive reals studied with classical logic.

- ▷ a constructive theory of constructive real numbers: the recursive reals studied with intuitionistic logic.

The BHK interpretation

Modern intuitionistic logic is based on the *BHK-interpretation** of the connectives

$$\vee \text{ (or)}, \wedge \text{ (and)}, \rightarrow \text{ (implies)}, \neg \text{ (not)}$$

and quantifiers

$$\exists \text{ (there exists)}, \quad \forall \text{ (for all/each)}.$$

Note that it is provability, rather than an a priori notion of truth, that is fundamental to the constructive approach.

*Brouwer-Heyting-Kolmogorov

- $P \vee Q$: either we have a proof of P or else we have a proof of Q.

- $P \wedge Q$: we have both a proof of P and a proof of Q.

- $P \to Q$: by means of an algorithm we can convert any proof of P into a proof of Q.

- $\neg P$: assuming P, we can derive a contradiction (such as $0 = 1$); equivalently, we can prove $(P \to (0 = 1))$.

- $\exists_x P(x)$: we have (i) an algorithm which computes a certain object x, and (ii) an algorithm which, using the information supplied by the application of algorithm (i), demonstrates that $P(x)$ holds.

- $\forall_{x \in A} P(x)$: we have an algorithm which, applied to an object x and a proof that $x \in A$, demonstrates that $P(x)$ holds.

Note that in the interpretation of the statement $\forall_{x \in A} P(x)$, the proof of $P(x)$ will normally use both the data describing the object x and the information supplied by a proof that x belongs to the set A.

Consider the statement:

LPO For each binary sequence $a \equiv (a_n)_{n \geqslant 1}$ either $a_n = 0$ for all n, or else there exists N such that $a_N = 1$.

This is trivially true under classical logic.

What is its BHK interpretation?

> We have an algorithm which, applied to any binary sequence a, either produces a proof that $a_n = 0$ for each n, or else computes N such that $a_N = 1$.

Claim: Such an algorithm is unlikely to be found.

The Goldbach conjecture (GC, 1742):

Every even integer > 2 is a sum of two primes.

Status still unknown.

Define a binary sequence a as follows.

If $2n + 2$ is a sum of two primes, set $a_n = 0$.

If there exists $k \leqslant n$ such that $2k+2$ is not a sum of two primes, set $a_n = 1$.

Suppose we have an algorithm as in the BHK interpretation of **LPO**. Applied to this binary sequence, this algorithm

either proves that $a_n = 0$ for all n (i.e., proves GC)

or else computes N such that $a_N = 1$ (i.e., gives a counterexample to GC).

The use of GC here is purely illustrative: we could have used any of a multitude of unsolved problems of a certain type (Riemann hypothesis, ...).

Conclusion: the existence of an algorithm as in the BHK interpretation of **LPO** is highly doubtful.

Moreover, **LPO** is provably false in certain models of constructive mathematics (but it is not provably false in Bishop-style constructive mathematics).

We therefore stay clear of **LPO** as a working constructive principle.

Consequence: we also must avoid using any proposition that constructively implies **LPO**.

In particular, we must avoid using the full law of excluded middle.

This has a serious impact on even elementary analysis.

Consider the classically trivial proposition:

$$\forall_{x \in \mathbb{R}} \left(x = 0 \vee x \neq 0 \right),$$

where

$$x \neq 0 \Leftrightarrow \exists_{n \in \mathbb{N}} \left(|x| > 2^{-n} \right).$$

Suppose we have a constructive proof—that is, an algorithm which, applied to any real number x either proves that $x = 0$ or else computes a positive integer N such that $|x| > 2^{-N}$.

Given a binary sequence a, apply this algorithm to the real number

$$x = \sum_{n=1}^{\infty} 2^{-n} a_n = 0 \cdot a_1 a_2 a_3 \ldots \text{ (infinite binary expansion).}$$

The algorithm either proves that $x = 0$, and therefore $a_n = 0$ for all n, or else computes N such that $|x| > 2^{-N}$.

In the second case, $a_n = 0$ for all $n > N$; so, by testing $a_1, a_2, \ldots, a_N$ (a finite test), we can check whether $a_n = 0$ for all n or there exists $n \leqslant N$ with $a_n = 1$.

Thus the proposition

$$\forall_{x \in \mathbb{R}} \, (x = 0 \vee x \neq 0)$$

implies **LPO** and is therefore essentially nonconstructive!

Here is another essentially nonconstructive principle that is trivially true under classical logic.

LLPO For each binary sequence a with at most one term equal to 1, either $a_n = 0$ for all even n, or else $a_n = 0$ for all odd n.

BHK-interpretation:

> We have an algorithm which, applied to any binary sequence a and the data that $a_n = 1$ for at most one n, either proves that all even-indexed terms of the sequence are 0, or else proves that all odd-indexed terms are 0.

Again, it is extremely unlikely that such an algorithm could be produced.

Moreover, **LLPO**, like **LPO**, is provably false in certain models of constructive mathematics (but it is not provably false in **BISH**).

We therefore avoid using **LLPO** as a working constructive principle.

Note that **LLPO** is a consequence of **LPO**; but **LPO** cannot be derived from **LLPO**.

Consider the classically trivial proposition: For each real number x, either $x \geqslant 0$ or $x \leqslant 0$.

Suppose we have a constructive proof: that is, an algorithm which, applied to any given real number x, either decides that $x \geqslant 0$ or else decides that $x \leqslant 0$.

Given a binary sequence a with at most one term equal to 1, apply this algorithm

$$\begin{aligned} x &= \sum_{n=1}^{\infty} (-1)^{n+1} 2^{-n} a_n \\ &= \frac{a_1}{2} - \frac{a_2}{4} + \frac{a_3}{8} - \frac{a_4}{16} + \cdots . \end{aligned}$$

If $x \geqslant 0$, then $a_n = 0$ for all even n; if $x \leqslant 0$, then $a_n = 0$ for all odd n.

Conclusion: The statement

$$\forall_{x \in \mathbb{R}} \, (x \geqslant 0 \vee x \leqslant 0)$$

implies **LLPO** and is therefore essentially nonconstructive.

The following elementary classical statements also turn out to be nonconstructive.

- ▷ Each real number x is either rational or irrational (that is, $x \neq r$ for each rational number r). To see this, consider

$$x = \sum_{n=1}^{\infty} \frac{1 - a_n}{n!},$$

where a is any increasing binary sequence. This is equivalent to **LPO**.

- ▷ Each real number x has a binary expansion. Note that the standard interval-halving argument for "constructing" binary expansions does not work, since we cannot necessarily decide, for a given number x between 0 and 1, whether $x \geqslant 1/2$ or $x \leqslant 1/2$. In fact, the existence of binary expansions is equivalent to **LLPO**.

- ▷ The intermediate value theorem, which is equivalent to **LLPO**.

- ▷ For all $x, y \in \mathbb{R}$, if $xy = 0$, then either $x = 0$ or $y = 0$. This is equivalent to **LLPO**. The constructive failure of this proposition clearly has implications for the theory of integral domains.

Note: classically valid statements like "each real number is either rational or irrational" that imply omniscience principles are *not false* in constructive mathematics. They cannot be, since **BISH** is consistent with classical mathematics (**CLASS**):

Every theorem in **BISH** *is also a theorem of* **CLASS.**

In fact, we can regard **CLASS** as **BISH** + **LEM**.

Another way of looking at **CLASS**: it is a model of **BISH**.

Brouwer's intuitionistic mathematics (**INT**) and the recursive constructive mathematics (**RUSS**) of the Markov School both use intuitionistic logic, and both are models of **BISH**:

Every theorem in **BISH** *is also a theorem of* **INT** *and of* **RUSS**.

Brauer has shown that **BISH** can be interpreted within Weihrauch's Type 2 Effectivity framework for computable analysis.

We use these models of **BISH** to obtain independence results.

Since

> **INT**/**CLASS** $\vdash$ *Every continuous function* $f : [0,1] \to \mathbb{R}$ *is uniformly continuous*

and

> **RUSS** $\vdash$ *There exists a continuous, real-valued function on* $[0,1]$ *that is not uniformly continuos,*

we see that each of the propositions following "$\vdash$" is neither provable nor disprovable in **BISH**. In other words, each of them is independent of **BISH**.

In place of the essentially nonconstructive propositions

$$\forall_{x\in\mathbb{R}}\,(x = 0 \vee x \neq 0)\,,$$
$$\forall_{x\in\mathbb{R}}\,(x \geqslant 0 \vee x \leqslant 0)\,,$$

we have these constructively valid propositions:

1) If $a < b$, then for each real number x, either $a < x$ or $x < b$.

2) If $(x > 0)$ is impossible, then $x \leqslant 0$.

Note, though, that the statement

> If $(x \geqslant 0)$ is impossible, then $x < 0$

implies (actually, is equivalent to) another constructively dubious principle, *Markov's principle*:

> **MP**: If a is a binary sequence and it is impossible that $a_n = 0$ for all n, then there exists N such that $a_N = 1$.

The real line

Starting with the set $\mathbb{N}$ of natural numbers, we can build the sets $\mathbb{Z}$ (of integers) and $\mathbb{Q}$ (of rationals) by elementary algebraic means.

By a *real number* we mean a subset $\mathbf{x}$ of $\mathbb{Q} \times \mathbb{Q}$ such that

- ▷ for all (q, q') in $\mathbf{x}$, $q \leqslant q'$;
- ▷ for all (q, q') and (r, r') in $\mathbf{x}$, the closed intervals $[q, q']$ and $[r, r']$ in $\mathbb{Q}$ intersect in points of $\mathbb{Q}$;
- ▷ for each positive rational ε there exists (q, q') in $\mathbf{x}$ such that $q' - q < \varepsilon$.

The last of these properties ensures that the set x is inhabited—that is, we can construct elements of $\mathbf{x}$).

Any rational number q gives rise to a canonical real number

$$\mathbf{q} = \{(q, q)\}$$

with which the original rational q is identified.

Two real numbers $\mathbf{x}$ and $\mathbf{y}$ are

- *equal*, written $\mathbf{x} = \mathbf{y}$, if for all $(q, q') \in \mathbf{x}$ and all $(r, r') \in \mathbf{y}$, the intervals $[q, q']$ and $[r, r']$ in $\mathbb{Q}$ have a rational point in common;

- *unequal* (or *distinct*), written $\mathbf{x} \neq \mathbf{y}$, if there exist $(q, q') \in \mathbf{x}$ and $(r, r') \in \mathbf{y}$ such that the intervals $[q, q']$ and $[r, r']$ in $\mathbb{Q}$ are disjoint.

Taken with the equality and inequality we have defined above, the collection of real numbers forms a set: the real line $\mathbb{R}$.

Let $\mathbf{x}, \mathbf{y}$ be real numbers. We say that

- $\mathbf{x} > \mathbf{y}$, and that $\mathbf{y} < \mathbf{x}$, if there exist $(q, q') \in \mathbf{x}$ and $(r, r') \in \mathbf{y}$ such that $r' < q$;

- $\mathbf{x} \geqslant \mathbf{y}$, and that $\mathbf{y} \leqslant \mathbf{x}$, if for all $(q, q') \in \mathbf{x}$ and all $(r, r') \in \mathbf{y}$ we have $q' \geqslant r$.

We pass over the (complicated) definitions of the algebraic operations on real numbers.

The set $\mathbb{R}$ is *uncountable*: if $(\mathbf{a}_n)_{n\geqslant 1}$ is a sequence of real numbers, then there exists $\mathbf{x} \in [0,1]$ such that $\mathbf{x} \neq \mathbf{a}_n$ for each n.

The set $\mathbb{R}$ is *complete*: every Cauchy sequence of real numbers converges to a limit in $\mathbb{R}$. (The proof requires the principle of dependent choice.)

What about the order-completeness of $\mathbb{R}$?

Let S be a subset of $\mathbb{R}$.

An *upper bound* of/for S is a real number $\mathbf{b}$ such that $\mathbf{x} \leqslant \mathbf{b}$ for each $\mathbf{x} \in S$. We say that $\mathbf{b}$ is the *supremum*, $\mathbf{sup}\, S$, of S if (i) it is an upper bound for S and (ii) for each $\mathbf{x} < \mathbf{b}$ there exists $\mathbf{s} \in S$ such that $\mathbf{x} < \mathbf{s}$.

We say that S is *upper order located* if for all rational numbers a, b with $a < b$, either $\mathbf{x} \leqslant b$ for all $\mathbf{x} \in S$ or else there exists $\mathbf{x} \in S$ such that $\mathbf{x} > a$.

The *constructive least-upper-bound principle*:

> *Let S be an inhabited set of real numbers that is bounded above. Then* $\mathbf{sup}\, S$ *exists if and only if S is upper order located.*

Analogous definitions and results hold for the infimum, $\mathbf{inf}\, S$, of S.

The upper order locatedness cannot be dropped from the hypotheses of the constructive least-upper-bound principle.

Consider any statement P. The set

$$S \equiv \{0\} \cup \{\mathbf{x} \in \mathbb{R} : x = 1 \wedge (P \vee \neg P)\}$$

is inhabited by 0 and bounded above by 1. Suppose that $\sigma \equiv \mathbf{sup}\, S$ exists. Then $\sigma \leqslant 1$. If $\sigma < 1$, then $\neg (P \vee \neg P)$, which is absurd. Hence $\sigma = 1$ and there exists $\mathbf{s} \in S$ with $\mathbf{s} > 1/2$. It follows that

$$\mathbf{s} \in \{\mathbf{x} \in \mathbb{R} : x = 1 \wedge (P \vee \neg P)\},$$

so $P \vee \neg P$.

From now on, we drop boldface notation for real numbers.

Lecture 2

Metric spaces

The elementary constructive notions associated with a metric space (X, ρ) are more or less as in classical mathematics.

Note that when there are alternative classical definitions, these may not be equivalent constructively. In that case, we choose the most computationally informative notion.

For example, we do not define closed sets as complements of open sets: a closed set S is defined to be one that equals it closure $\overline{S}$; in other words, S is closed in X if and only if all limits of sequences in S belong to S.

We denote the open and closed balls in X with centre a and radius $r > 0$ by $B(a, r)$ and $\overline{B}(a, r)$ respectively.

As in the classical theory, both X and $\varnothing$ are open, unions of open sets are open, and finite intersections of open sets are open; in other words, the open sets form a topology on X.

Likewise, both X and $\varnothing$ are closed, and arbitrary intersections of closed sets are closed.

We cannot prove that the union of two closed sets is closed: in the metric space $\mathbb{R}$, the intervals $[-1, 0]$ and $[0, 1]$ are closed and their union is dense in $[-1, 1]$; but if that union is closed, then

$$\forall_{x \in \mathbb{R}} \left(x \leqslant 0 \vee x \geqslant 0 \right),$$

a proposition equivalent to **LLPO**.

Decent sets come equipped with an inequality as well as an equality. The *inequality* on a metric space is defined by

$$x \neq y \Leftrightarrow \rho(x, y) > 0.$$

Note that

$$\neg\,(x \neq y) \Rightarrow (x = y)$$

but without Markov's principle we cannot prove that

$$\neg(x = y) \Rightarrow x \neq y.$$

The *complement* of a set S in X is the set

$$\sim S \equiv \{x \in X : \forall_{s \in S} (x \neq s)\}.$$

In the absence of Markov's principle, this is not the same as the logical complement

$$\neg S \equiv \{x \in X : x \notin S\}$$

of S.

Proposition: *If S is an open subset of X, then $\sim S$ is closed in X, and $\sim S = \neg S$.*

The complement of a closed set need not be open.

Given $\varepsilon > 0$, by an *ε-approximation* to a subset S of a metric space X we mean an inhabited subset T of S such that for each $s \in S$ there exists $t \in T$ with $\rho(s,t) < \varepsilon$.

If for each $\varepsilon > 0$ there exists a finitely enumerable ε-approximation to S, then we say that S is *totally bounded.*

The closure of a totally bounded subset of X is totally bounded.

If a subset S of X contains a dense totally bounded set, then S itself is totally bounded.

The product of finitely many totally bounded spaces is totally bounded.

Total boundedness is very important in constructive analysis because

- it helps us to compute suprema and infima in many important situations, and

- coupled with completeness, total boundedness gives the only one of three classically equivalent notions of compactness that can be used **BISH**.

Proposition: *If $S \subset \mathbb{R}$ is totally bounded, then* sup S *and* inf S *exist.*

Recall that a mapping $f : X \to Y$ between metric spaces is *uniformly continuous* if for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\forall_{x,x' \in X} \left(\rho(x, x') < \delta \Rightarrow \rho\left(f(x), f(x')\right) < \varepsilon \right).$$

Proposition: *If X is totally bounded, and $f : X \to Y$ is a uniformly continuous mapping of X into a metric space Y, then $f(X)$ is totally bounded.*

Another extremely important, though classically vacuous, property: locatedness.

An inhabited subset S of a metric space X is *located* in X if for each $x \in X$ the *distance*

$$\rho\,(x, S) \equiv \mathsf{inf}\,\{\rho(x, s) : s \in S\}$$

exists.

Proposition: *A totally bounded subset of a metric space is located. A located subset of a totally bounded metric space is totally bounded.*

The proposition

> Every inhabited subset of $\mathbb{R}$ is located

implies **LEM**. How?

Let P be any proposition, and

$$S \equiv \{0\} \cup \{x \in \mathbb{R} : (x = 1) \wedge P\}\,.$$

If S is located, then either $\rho(1, S) > 0$ or $\rho(1, S) < 1$.

In the first case we have $\neg P$. In the second, choosing $s \in S$ such that $\rho(1, s) < 1$, we see that $s \notin \{0\}$, so $s = 1$ and P holds.

Let X be a totally bounded metric space. The next results provide us with a rich supply of totally bounded, and hence located, subsets of X.

Lemma: *For each $x_0 \in X$ and each $r > 0$, there exists a closed, totally bounded subset K of X such that*

$$B(x_0, r) \subset K \subset \overline{B}(x_0, 8r).$$

Proposition: *For each $\varepsilon > 0$ there exist totally bounded subsets $K_1, \ldots, K_n$ each of diameter $\leqslant \varepsilon$, such that $X = \bigcup_{i=1}^{n} K_i$.*

A property P, applicable to certain elements of a set S, is said to hold *for all but countably many* x in S if there exists a sequence $(x_n)_{n\geqslant 1}$ in S such that $P(x)$ holds whenever $x \in S$ and $x \neq x_n$ for each n. The sequence $(x_n)_{n\geqslant 1}$ is then called the *excluded sequence*, and the elements x such that $x \neq x_n$ for each n are said to be *admissible*, for the property P.

Theorem: *If* $f : X \to \mathbb{R}$ *is a uniformly continuous mapping, then for all but countably many* $r \in \mathbb{R}$ *the set*

$$f^{-1}(-\infty, r] \equiv \{x \in X : f(x) \leqslant r\}$$

is either totally bounded or empty.

A complete, totally bounded metric space X is said to be *compact.*

The bounded closed intervals $[a, b]$ in $\mathbb{R}$, and the closed balls in $\mathbb{C}$, are compact.

The product of finitely many compact spaces is compact.

A compact subset of a metric space is both closed and located.

A closed, located subset of a compact space is compact.

An inhabited metric space X is said to be

- *locally totally bounded* if each bounded subset of X is contained in a totally bounded subset;

- *locally compact* if it is both locally totally bounded and complete.

Every compact space is locally compact.

The spaces $\mathbb{R}$ and $\mathbb{C}$, and the product spaces $\mathbb{R}^n$ and $\mathbb{C}^n$, are locally compact.

A metric space X is locally compact if and only if every bounded subset of X is contained in a compact set.

Proposition: *Let Y be an inhabited subset of a metric space X.*

(i) *If Y is locally totally bounded, then it is located.*

(ii) *If X is locally totally bounded and Y is located, then Y is locally totally bounded.*

Normed linear spaces

Let X be a linear space over the field $\mathbb{K}$ (either $\mathbb{R}$ or $\mathbb{C}$). An inequality relation $\neq$ on X is said to be *compatible with the linear structure* on X if, for all $x, y \in X$ and $t \in \mathbb{K}$,

$$\begin{aligned} x \neq y &\Leftrightarrow x - y \neq 0, \\ x + y \neq 0 &\Rightarrow x \neq 0 \vee y \neq 0, \\ tx \neq 0 &\Rightarrow t \neq 0 \wedge x \neq 0. \end{aligned}$$

Then

$$x \neq y \Rightarrow \forall_{z \in X} \left(x + z \neq y + z \right).$$

From now on, "linear space" means "linear space with a compatible inequality".

A *seminorm* on a linear space X is mapping $x \rightsquigarrow \|x\|$of X into the nonnegative real line $\mathbb{R}^{0+}$ such that for all x, y in X and all t in $\mathbb{K}$,

- $\|x\| > 0 \Rightarrow x \neq 0$,
- $\|tx\| = |t| \, \|x\|$, and
- $\|x + y\| \leqslant \|x\| + \|y\|$.

Then $(X, \| \, \|)$—or just X itself—is a *seminormed (linear) space* over $\mathbb{K}$. If the inequality satisfies

$$x \neq 0 \Leftrightarrow \|x\| > 0,$$

then $\| \, \|$ is called a *norm* on X.

Let X be a normed space. Then the mapping $(x, y) \rightsquigarrow \|x - y\|$ of $X \times X$ into $\mathbb{R}$ is provides the associated metric ρ on X.

The *unit ball* of X is the closed ball with centre 0 and radius 1,

$$B_X = \overline{B}_X(0, 1) = \overline{B}(0, 1) = \{x \in X : \|x\| \leqslant 1\},$$

relative to that metric. This ball, like any open or closed ball in a normed space, is located.

We pass over most of the standard examples, notions, and elementary properties familiar from the classical theory of normed spaces.

A mapping u between vector spaces X, Y is *linear* if

$$u(x+y) = u(x) + u(y) \quad \text{and} \quad u(tx) = tu(x)$$

whenever $x, y \in X$ and $t \in \mathbb{K}$.

If $X = Y$, then u is called an *operator* on X.

If $Y = \mathbb{K}$, then u is called a *linear functional* on X.

A linear mapping $u : X \to Y$ between normed spaces is continuous on X if and only if it is *bounded*, in the sense that there exists $c > 0$ such that $\|u(x)\| \leqslant c\,\|x\|$ for each $x \in X$.

This is not enough to ensure that u is *normed/normable*, in the sense that

$$\|u\| \equiv \sup\{\|u(x)\| : x \in X, \|x\| \leqslant 1\}$$

exists.

There is a criterion for normability of nonzero bounded linear functionals.

Proposition: *A nonzero linear functional u on a normed space X is normed if and only if*

$$\ker u \equiv \{x \in X : u(x) = 0\}$$

is located in X.

Basic idea of the proof: for each $x \in X$,

$$\rho(x, \ker u) = \frac{|u(x)|}{\|u\|},$$

provided either $\rho(x, \ker u)$ or $\|u\|$ exists.

Finite-dimensional spaces

Let X be a linear space.

Vectors $e_1, \ldots, e_n$ in X are *linearly independent* if $\sum_{i=1}^{n} \lambda_i e_i \neq 0$ for all scalars $\lambda_1, \ldots, \lambda_n$ such that $\sum_{i=1}^{n} |\lambda_i| > 0$.

We say that X is *finite-dimensional* if either $X = \{0\}$ or else it contains finitely many linearly independent vectors $e_1, \ldots, e_n$ such that for each $x \in X$ there exist scalars $\lambda_1, \ldots, \lambda_n$ for which $x = \sum_{i=1}^{n} \lambda_i e_i$.

In the first case, X is 0-*dimensional*.

In the second, X is n-*dimensional* and $\{e_1, \ldots, e_n\}$ is a *basis* of X. The *coordinates* $u_i(x) \equiv \lambda_i$ are uniquely determined by x, and the *coordinate functionals* $u_i : X \longrightarrow \mathbb{K}$ are linear mappings.

Inducting on the dimension, we can prove, in turn, that

(i) the coordinate functionals on a finite-dimensional normed space are bounded, and

(ii) every linear mapping of a finite-dimensional normed space into a normed space is bounded and normed.

Proposition: *A finite-dimensional normed space is complete and locally totally bounded (hence locally compact).*

Proposition: *A normed space is finite-dimensional if and only if its closed unit ball is compact.*

A subspace Y of a metric space X is called *proximinal* if each element of X has a *best approximation* in Y: that is, if for each $a \in X$ there exists $b \in Y$ such that $\rho(x, b) \geqslant \rho(x, y)$ for all $y \in Y$. In that case, Y is located in X.

Classical fundamental theorem of approximation theory: a finite-dimensional subspace V of a real normed space X is *proximinal*.

This result implies **LLPO**.

For a constructive version of the theorem, we introduce a new notion: quasi-proximinality.

We say that a has *at most one best approximation* in Y if for all distinct points y, y' in Y, there exists $z \in Y$ such that

$$\max\left\{\rho(a,y), \rho(a,y')\right\} > \rho(a,z).$$

We call Y *quasiproximinal* if each point of X with at most one best approximation in Y actually has a (perforce unique) best approximation in Y.

Proximinal implies quasiproximinal.

The converse cannot be proved in **BISH** but can be proved using **LEM**:

> Suppose that $a \in X$ has no best approximation in a quasiproximinal subspace Y of X. Then (classically!) a has at most one best approximation in Y; so, by quasiproximinality, a has a best approximation in Y, which is a contradiction.

The next lemma is crucial for the proof of our approximation theorem, and uses the extremely important λ-technique (of which more later).

Lemma: *Let x, e be elements of a real normed space X with $e \neq 0$, and let $d \geqslant 0$. Suppose that*

$$\max\left\{\|x - te\|, \left\|x - t'e\right\|\right\} > d$$

whenever t, t' are distinct real numbers. Then there exists $\tau \in \mathbb{R}$ such that if $\|x - \tau e\| > d$, then $\rho(x, \mathbb{R}e) > 0$.

Constructive fundamental theorem of approximation theory: *Every finite-dimensional subspace of a real normed space is quasiproximinal.*

Proved by induction on the dimension n of the subspace. The case $n = 0$ is trivial; the case $n = 1$ follows easily from the lemma. The lemma is also used in the induction step.

Hilbert spaces

We assume familiarity with the elementary properties of an inner product and the corresponding norm, an inner product space, and a Hilbert space (a complete inner product space).

The classical proof of the proximinality of closed, located subspaces of a Hilbert space is constructively sound.

Proposition: *Let S be a closed, located subspace of a Hilbert space H. Then for each $x \in H$, there exists a unique element Px of S such that $\|x - Px\| = \rho(x, S)$. Moreover, Px is the unique element y of S such that $\langle x - y, s\rangle = 0$ for all $s \in S$.*

The mapping P is bounded and linear, and is called the *projection* of H on the subspace S.

Two subsets S, T of an inner product space are said to be *orthogonal* if $\langle x, y \rangle = 0$ for all $x \in S$ and $y \in T$; we then write $S \perp T$.

The *orthogonal complement* of a subset S of X is

$$S^{\perp} = \{x \in X : x \perp S\},$$

a closed linear subspace of X.

If S is a closed, located subspace of a Hilbert space H, with P the corresponding projection, then $I - P$ is the projection of H on $S^{\perp}$, where I is the identity operator $x \rightsquigarrow x$ on H.

A family $(e_i)_{i\in I}$ of vectors in a Hilbert space H is said to be *orthonormal* if

- $e_i \perp e_j$ whenever $i \neq j$, and
- for each i, either $\|e_i\| = 1$ or $e_i = 0$.

Such a family is an *orthonormal basis* if each vector $x \in H$ can be written uniquely in the form

$$x = \sum_{i\in I} \lambda_i e_i$$

where for each $i \in I$, $\lambda_i \in \mathbb{K}$ and if $e_i = 0$, then $\lambda_i = 0$.

Classically, using (an equivalent of) the axiom of choice, we can prove that every Hilbert space has an orthonormal basis of unit vectors.

Constructively we avoid the axiom of choice by adding separability to the hypotheses on H, by relaxing the requirements to allow basis vectors to be 0, and by using the Gram–Schmidt orthogonalisation process to prove that every separable Hilbert space has a countable orthonormal basis.

We say that a normed space X is infinite-dimensional if the complement of each finite-dimensional subspace of X is inhabited.

Proposition: *Let $(e_n)_{n\geqslant 1}$ be an orthonormal basis of a separable Hilbert space H. Then*

(i) *H is finite-dimensional if and only if $e_n = 0$ for all sufficiently large n.*

(ii) *H is infinite-dimensional if and only if $e_n \neq 0$ for infinitely many n.*

The Riesz representation theorem for linear functionals on a Hilbert space uses the classically redundant condition of normability:

Theorem: *A bounded linear functional u on a Hilbert space is normed if and only if there exists a unique vector $a \in H$ such that $u(x) = \langle x, a \rangle$ for each $x \in H$.*

Proving "only if" is the harder part, in which the classical argument goes through if $\|u\| > 0$. For the general case, we use a little trick.

We consider the direct sum $H \oplus \mathbb{K}$, a Hilbert space with the inner product

$$\left\langle (x, \zeta), \left(x', \zeta'\right) \right\rangle \equiv \left\langle x, x' \right\rangle + \zeta\zeta',$$

on which we define a nonzero bounded linear functional v by

$$v(x, \zeta) = u(x) + \zeta.$$

A little work shows that v is normed. By the first part of the proof, there exists $a \in X$ such that

$$v(x, \zeta) = \langle (x, \zeta), (a, 1) \rangle$$

for each $(x, \zeta) \in H \oplus \mathbb{K}$. Then $u(x) = \langle x, a \rangle$ for each $x \in H$.

An *operator* on a Hilbert space H is a linear mapping of H into itself.

The set of bounded operators on H is denoted by $\mathcal{B}(H)$.

For any not-necessarily-bounded operator T on H, we define the *adjoint* T^*, if it exists, by the equation

$$\langle Tx, y\rangle = \langle x, T^*y\rangle \quad (x, y \in H)\,, \tag{1}$$

in which case we refer to T as *jointed*.

Classically, the Riesz representation theorem enables us to prove the existence of T^* for any element T of $\mathcal{B}(H)$.s

Constructively, the universal existence of adjoints implies **LPO**.

Can we characterise those operators for which the adjoint exists? Yes, by the following result of Ishihara and Richman.

Proposition. *A bounded operator on a Hilbert space H is jointed if and only if it maps the unit ball of H to a located set.*

If T is jointed, then T^* is an operator and T is its adjoint.

Any bound (in particular, the norm if it exists) for T is one for T^*, and vice versa.

Moreover, if S, T are jointed operators, then for each $\lambda \in \mathbb{K}$, so are $\lambda S + T$ and ST, and

$$(\lambda S + T)^* = \lambda^* S^* + T^*,$$
$$(ST)^* = T^* S^*.$$

An operator T is *selfadjoint*, or *Hermitian*, if T^* exists and equals T.

Projections are selfadjoint.

Conversely, if P is any bounded, idempotent, selfadjoint operator on H, then P is the projection of H on the (located) subspace

$$\{y \in H : Py = y\}.$$

A linear mapping T between normed spaces X, Y is said to be *compact* if $T(\overline{B}_X(0,1))$ is a totally bounded subset of Y.

In that case, the norm of T exists, since

$$\Big\{\|Tx\| : x \in \overline{B}_X(0,1)\Big\}$$

is a totally bounded subset of $\mathbb{R}$.

Every bounded linear mapping on a finite-dimensional normed space is compact.

We end this part of the lectures with two early results of Ishihara. The proof of the first uses the Riesz representation theorem to cut things down to a finite-dimensional subspace of the Hilbert space.

Proposition: *Let T be a bounded linear mapping of a Hilbert space H into $\mathbb{C}^n$, and for $1 \leqslant k \leqslant n$ let $P_k : \mathbb{C}^n \to \mathbb{C}$ be defined by*

$$P_i\left(z_1, \ldots, z_n\right) \equiv z_k.$$

Then T is compact if and only if $P_k \circ T$ is normed for each k.

The second result is easily proved classically by a sequential compactness argument.

Proposition: *The sum of two compact operators on a Hilbert space is compact.*