# Functional analysis

**Lecture 12:** The open mapping and inverse mapping theorems; the closed graph theorem; corollaries, e.g. the Hellinger–Toeplitz theorem; trigonometric polynomials

Our first corollary from the Banach–Steinhaus theorem should be compared with **Problem 3.20**, where we showed that every weakly convergent sequence in $C(K)$ is bounded in the supremum norm (and it was the only difficult thing about that problem). We used there a 'sliding hump' argument and, in fact, in this way H. Hahn proved a special form of the uniform boundedness principle a few years before it appeared in its full form in the paper by Banach and Steinhaus.

**Corollary 7.3.** *A subset $A \subseteq X$ of a normed space is weakly bounded (which means that $\{\langle x, x^* \rangle : x \in A\}$ is bounded for each $x^* \in X^*$) if and only if it is bounded in norm.*

*Proof.* We regard A as a subset of the bidual $X^{**}$ by the canonical embedding $\kappa\colon X \to X^{**}$ given by $\langle x^*, \kappa(x) \rangle = \langle x, x^* \rangle$. If $A$ is weakly bounded, it is pointwise bounded as a set of functionals on the Banach space $X^*$. By Theorem 7.2, $\sup\{\|\kappa(x)\|_{X^{**}} : x \in A\} < \infty$ which is the same as $\sup\{\|x\| : x \in A\} < \infty$, as $\kappa$ is an isometry. □

**Corollary 7.4.** *Let $X$ be a Banach space and $Y$ a normed space. Assume that a sequence $(T_n)_{n=1}^{\infty} \subset \mathscr{L}(X, Y)$ is pointwise convergent on $X$. Then the formula $Tx = \lim_{n\to\infty} T_n x$ defines an operator $T \in \mathscr{L}(X, Y)$ satisfying $\|T\| \leq \liminf_{n\to\infty} \|T_n\|$.*

*Proof.* Linearity of $T$ is obvious. Note that $(T_n)_{n=1}^{\infty}$ is pointwise bounded and hence uniformly bounded by the Banach–Steinhaus theorem. Let $M = \liminf_{n\to\infty} \|T_n\| < \infty$ and observe that

$$\|Tx\| = \lim_{n\to\infty} \|T_n x\| = \liminf_{n\to\infty} \|T_n x\| \leq M\|x\| \quad \text{for every } x \in X. \qquad \square$$

Now, we proceed to the second fundamental result based on a category argument, namely, the open mapping theorem. We will formulate it in a bit more general form than usually stated just assuming that the range of a given operator $T$ is of second category. In this situation we do not need to appeal to the Baire category theorem. However, when applying this result in the case where the codomain is a Banach space and $T$ is surjective, we use Baire's theorem to conlude that the range is of second category and hence the open mapping theorem below implies that $T$ is open.

Recall that a mapping $f\colon X \to Y$ between topological spaces $X$ and $Y$ is *open* at a point $p \in X$ if for every open neighborhood $U$ of $p$, $f(U)$ contains an open neighborhood of $f(p)$. We say $f$ is an *open* map provided it is open at every point of $X$, equivalently: the range of every open set is open.

**Theorem 7.5 (Open mapping theorem).** *Let $X$ be a Banach space, $Y$ a normed space and let $\Lambda \in \mathscr{L}(X, Y)$ be an operator with a range $\Lambda(X)$ being of second category in $Y$. Then:*

(a) $\Lambda(X) = Y$;

(b) $\Lambda$ *is an open mapping;*

(c) $Y$ *is a Banach space.*

*Proof.* Notice that (a) follows easily from (b), since $\Lambda(X)$ is a linear subspace of $Y$ and if it contains any ball of positive radius, it must cover the whole of $Y$.

Fix any zero neighborhood $V \subset X$. We will show that $\Lambda(V)$ contains a zero neighborhood in $Y$. It will then follow that $\Lambda$ is open, because $\Lambda$ is linear and any open set in $X$ is a translation of a zero neighborhood. Take $r > 0$ so small that $rB_X \subset V$. Define

$$V_n = \{x \in X \colon \|x\| < 2^{-n} r\} \quad (n \in \mathbb{N})$$

and note that

$$\overline{\Lambda(V_2)} - \overline{\Lambda(V_2)} \subseteq \overline{\Lambda(V_2) - \Lambda(V_2)} \subseteq \overline{\Lambda(V_1)}. \tag{7.1}$$

*Claim.* There exists a zero neighborhood $W \subset Y$ such that

$$W \subset \overline{\Lambda(V_1)} \subset \Lambda(V).$$

In view of (7.1), for proving the first inclusion it is enough to show that $\operatorname{int} \overline{\Lambda(V_2)} \neq \varnothing$, i.e. that $\Lambda(V_2)$ is not nowhere dense. Since

$$\Lambda(X) = \bigcup_{k=1}^{\infty} k\Lambda(V_2)$$

and $\Lambda(X)$ is of second category, there exists $k \in \mathbb{N}$ for which $k\Lambda(V_2)$ is not nowhere dense. Of course, the map $y \mapsto ky$ is a homeomorphism, thus simply $\Lambda(V_2)$ is not nowhere dense, as desired.

For the second inclusion of our Claim, fix any $y_1 \in \overline{\Lambda(V_1)}$. We are going to define recursively a sequence $(y_n)_{n=1}^{\infty} \subset Y$ as follows. Assume $n \in \mathbb{N}$ and we have already defined $y_n \in \overline{\Lambda(V_n)}$. By the above argument, $\overline{\Lambda(V_{n+1})}$ contains a zero neighborhood, hence

$$\big(y_n - \overline{\Lambda(V_{n+1})}\,\big) \cap \Lambda(V_n) \neq \varnothing.$$

Therefore, there exists $x_n \in V_n$ such that $\Lambda(x_n) \in y_n - \overline{\Lambda(V_{n+1})}$. We then define

$$y_{n+1} = y_n - \Lambda(x_n) \in \overline{\Lambda(V_{n+1})}.$$

The idea of this induction is that after the $n^{\text{th}}$ step we arrived at a similar problem as the initial one, that is, to represent $y_1 \in \overline{\Lambda(V_1)}$ in the form $y_1 = \Lambda(x)$ with some $x \in V$. But now we have a point $y_{n+1}$ which lies in a much smaller set $\overline{\Lambda(V_{n+1})}$ and each difference $y_j - y_{j+1}$ $(1 \leq j \leq n)$ is of the form $\Lambda(x_j)$ which allows us to represent $y_1 - y_{n+1}$ as a value of $\Lambda$ by using a telescoping sum argument. What remains to do is to extend the induction to infinity and use the completeness of $X$.

Since $\|x_n\| \leq 2^{-n} r$ for $n \in \mathbb{N}$, the partial sums $(x_1 + \ldots + x_n)_{n=1}^{\infty}$ form a Cauchy sequence in $X$. Hence, we can define $x = \sum_{n=1}^{\infty} x_n \in X$. Plainly, $\|x\| < r$ which means that $x \in V$. Also, notice that for each $N \in \mathbb{N}$ we have

$$\sum_{n=1}^{N} \Lambda(x_n) = \sum_{n=1}^{N} (y_n - y_{n+1}) = y_1 - y_{N+1}.$$

Obviously, $\|y_N\| \to 0$ as $N \to \infty$ and since $\Lambda$ is continuous, we get $y_1 = \Lambda(x) \in \Lambda(V)$ which completes the proof of the Claim.

It remains to show assertion (c). To this end, we consider the quotient space $X/\ker \Lambda$ of all cosets $x + \ker \Lambda$ equipped with the distance norm

$$\|x + \ker \Lambda\| = \inf\{\|x + y\| \colon y \in \ker \Lambda\}.$$

This is indeed a norm on $X/\ker\Lambda$ as $\ker\Lambda$ is closed. It is also easy to show that $X/\ker\Lambda$ is a Banach space (recall that $X$ is complete). Thus, it is enough to exhibit a linear isomorphism of $X/\ker\Lambda$ onto $Y$.

Define $\Phi\colon X/\ker\Lambda\to Y$ by $\Phi(x+\ker\Lambda)=\Lambda(x)$. Then, $\Phi$ is well-defined, linear and surjective. Moreover, $\Lambda=\Phi\circ\pi$, where $\pi\colon X\to X/\ker\Lambda$ is the canonical quotient map which is open. Indeed, the topology in $X/\ker\Lambda$ consists of all sets $E\subseteq X/\ker\Lambda$ for which $\pi^{-1}(E)$ is open in $X$. Since for any open set $U\subseteq X$ we have $\pi^{-1}(\pi(U))=U+\ker\Lambda$, we infer that $\pi(U)$ is open. To see that $\Phi$ is an isomorphism, let $V\subseteq Y$ by any open set. Then $\Phi^{-1}(V)=\pi(\Lambda^{-1}(V))$ is open as $\Lambda$ is continuous and $\pi$ is open. Similarly, if $U\subseteq X/\ker\Lambda$ is open, then $\Phi(U)=\Lambda(\pi^{-1}(U))$ is open as $\Lambda$ is open (which has been already proved) and $\pi$ is continuous. Consequently, $\Phi$ is a linear homeomorphism, that is, an isomorphism. □

The following result is by far the most important corollary from the open mapping theorem and it is known as Banach's inverse mapping theorem.

**Corollary 7.6.** *Let $X$ and $Y$ be Banach spaces. Then, every one-to-one surjective operator $T\in\mathscr{L}(X,Y)$ is invertible, i.e. it is bounded below and hence $T^{-1}\in\mathscr{L}(Y,X)$.*

*Proof.* Since $T(X)=Y$ is of second category due to Baire's category theorem, Theorem 7.5 says that $T$ is open. Hence, there is $\delta>0$ such that $T(B_X)\supseteq\delta B_Y$. This means that $\|Tx\|\geq\delta$ for $\|x\|\geq 1$, that is, $\|T^{-1}\|\leq\delta^{-1}$. □

The next corollary is a bit striking. It turns out that nonequivalent norms on a Banach space cannot be comparable.

**Corollary 7.7.** *Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on a vector space $X$ for which both $(X,\|\cdot\|_1)$ and $(X,\|\cdot\|_2)$ are Banach spaces. If there exists a constant $c>0$ such that $\|\cdot\|_1\leq c\|\cdot\|_2$, then the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.*

*Proof.* Consider the identity mapping $\iota\colon(X,\|\cdot\|_2)\to(X,\|\cdot\|_1)$. By the assumption, it is a bounded linear operator. Since it is plainly bijective, Corollary 7.6 implies that $\iota^{-1}$ is bounded which means that $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent. □

Now, we proceed to the closed graph theorem. It is in fact another quite simple corollary from the open mapping theorem, but it is so important that deserves a special name. In many occasions it provides a very convenient way of verifying whether a given linear operator is bounded. Recall that for any map $f\colon X\to Y$ between topological spaces $X$ and $Y$ we define its *graph* by

$$\mathrm{Gr}(f)=\{(x,f(x))\colon x\in X\}.$$

It is easily seen that if $f$ is continuous, then $\mathrm{Gr}(f)$ is continuous as a subset of $X\times Y$ with the product topology. The converse implication in general fails but, as we will see below, not for linear operators between Banach spaces.

**Theorem 7.8 (Closed graph theorem).** *Let $X$ and $Y$ be Banach spaces and $T\colon X\to Y$ be a linear operator. If the graph $\mathrm{Gr}(T)$ is closed in $X\times Y$, then $T$ is continuous, i.e. $T\in\mathscr{L}(X,Y)$.*

*Proof.* Consider the direct sum $X \oplus Y$ equipped with the norm $\|(x,y)\| = \|x\|_X + \|y\|_Y$. Of course, such a norm generates the product topology on $X \times Y$ as $\|(x_n, y_n) - (x,y)\| \to 0$ if and only if $\|x_n - x\|_X \to 0$ and $\|y_n - y\|_Y \to 0$.

Observe that $\mathrm{Gr}(T)$ forms a linear subspace of $X \oplus Y$ and the latter is a Banach space, because so are $X$ and $Y$. Define projections $p\colon \mathrm{Gr}(T) \to X$ and $q\colon \mathrm{Gr}(T) \to Y$ by $p(x, Tx) = x$ and $q(x, Tx) = Tx$, respectively. Obviously, $p$ and $q$ are bounded operators and $p$ is moreover bijective. By Corollary 7.6, $p^{-1} \in \mathscr{L}(X, \mathrm{Gr}(T))$ which implies that $T = q \circ p^{-1}$ is bounded, as desired. □

To understand what is the real profit from the closed graph theorem, suppose we are to prove that a given linear operator $T\colon X \to Y$ is bounded. Theorem 7.8 reduces this problem to the following situation: Assume we have a convergent sequence $(x_n)_{n=1}^{\infty} \subset X$, say $x_n \to x_0$, such that the sequence of values $(Tx_n)_{n=1}^{\infty}$ is *also convergent* in $Y$, say $Tx_n \to y_0$ (i.e. the elements $(x_n, Tx_n)$ of $\mathrm{Gr}(T)$ converge in $X \oplus Y$). We just need to show that $y_0 = Tx_0$. In many cases it makes the problem of continuity of $T$ almost trivial, as we can assume that $(Tx_n)_{n=1}^{\infty}$ is convergent. Note, however, that it is important that both $X$ and $Y$ are complete (see **Problem 2.6**).

Let us present one important consequence of the closed graph theorem which has some meaning in the mathematical formulation of quantum mechanics. In that theory, observables, such as velocity or energy, are identified with symmetric operators on a Hilbert space. Since some of them are unbounded, the result below implies that they cannot be defined on the whole Hilbert space. Therefore, one need to consider *densely defined* unbounded operators.

**Corollary 7.9** (**Hellinger–Toeplitz theorem**). *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a linear symmetric operator on a Hilbert space $\mathcal{H}$, that is,*

$$(Tx, y) = (x, Ty) \quad \textit{for all } x, y \in \mathcal{H}.$$

*Then $T$ is bounded.*

*Proof.* In view of Theorem 7.8, we need to show that $\mathrm{Gr}(T)$ is closed in $\mathcal{H} \oplus \mathcal{H}$, where on the direct sum we consider e.g. the Hilbertian norm $\|(x,y)\| = (\|x\|^2 + \|y\|^2)^{1/2}$ (or any other equivalent norm). Suppose a sequence of elements $(x_n, Tx_n)$ of $\mathrm{Gr}(T)$ converges in $\mathcal{H} \oplus \mathcal{H}$. By a suitable translation, we can assume that $x_n \to 0$ and let $y = \lim_{n\to\infty} Tx_n$. By the assumption and the Cauchy–Schwarz inequality, we have

$$|(Tx_n, y)| = |(x_n, Ty)| \le \|x_n\| \|Ty\| \xrightarrow[n\to\infty]{} 0 \quad \text{for every } y \in \mathcal{H}. \tag{7.2}$$

By the Riesz representation theorem (Theorem 5.7), every functional from $\mathcal{H}^*$ is of the form $(\cdot, y)$ with some $y \in \mathcal{H}$. Hence, (7.2) yields that $Tx_n \to 0$ weakly. But since $(Tx_n)_{n=1}^{\infty}$ is norm convergent to $y$, we must have $y = 0$. Therefore, $(x_n, Tx_n) \to (0,0) \in \mathrm{Gr}(T)$ which shows that the graph of $T$ is closed and completes the proof. □

## 8 An application: Convergence of Fourier series

Denote by $\mathbb{T} = \{z \in \mathbb{C} \colon |z| = 1\}$ the unit circle. Naturally, every function $F\colon \mathbb{T} \to \mathbb{C}$ can be identified with a $2\pi$-periodic function on the real line $f\colon \mathbb{R} \to \mathbb{C}$ by the formula $f(t) = F(e^{\mathrm{i}t})$. In this section, we will consider the Banach spaces $C(\mathbb{T})$ of complex-valued continuous functions on $\mathbb{T}$ and $L_p(\mathbb{T})$ (basically for $p \in \{1,2\}$), where $\mathbb{T}$ is equipped with the normalized Lebesgue measure. The latter space can be thus identified with $L_p[-\pi,\pi]$ or $L_p[0,2\pi]$, where the measure considered is the normalized Lebesgue measure on an interval of length $2\pi$, i.e. these are $L_p(\mu)$-spaces with the measure $\mathrm{d}\mu = (2\pi)^{-1}\mathrm{d}x$. Hence, $L_2[0,2\pi]$ is equipped with a norm and an inner product given by the formulas:

$$\|f\|_2 = \left\{\frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2\,\mathrm{d}t\right\}^{1/2}, \qquad (f,g) = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\,\mathrm{d}t.$$

The identification between functions defined on $\mathbb{T}$ and $2\pi$-periodic functions on $\mathbb{R}$ will be used throughout without mentioning. Hence, an argument $e^{\mathrm{i}\theta}$ of a given function $f$ on $\mathbb{T}$ will be often identified with the angle $\theta$, so there should be nothing misleading in using symbols like $f(\theta)$.

The basic observation for this section is that the sequence of complex exponents $(e^{\mathrm{i}nt})_{n\in\mathbb{Z}}$ forms an orthonormal system in the Hilbert space $L_2[0,2\pi]$. For proving this, denote $u_n(t) = e^{\mathrm{i}nt}$ and observe that for $n \neq 0$ the primitive function is $(\mathrm{i}n)^{-1}e^{\mathrm{i}nt}$ which integrates to zero over the whole period $[0,2\pi]$. Hence,

$$(u_m,u_n) = \frac{1}{2\pi}\int_0^{2\pi} e^{\mathrm{i}(m-n)t}\,\mathrm{d}t = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n. \end{cases}$$

This is probably the most important orthonormal set in the whole mathematics. In fact, it is complete (i.e. linearly dense) in $L_2[0,2\pi]$, but it is by no means trivial and we shall derive this fundamental fact from Fejér theorem. This is only one of several possible methods of proving the completeness of $(u_n)_{n\in\mathbb{Z}}$, but it gives a nice quantitative information about approximating continuous $2\pi$-periodic functions by linear combinations of the exponents $u_n$.

**Definition 8.1.** Any $2\pi$-periodic function on $\mathbb{R}$ of the form

$$f(t) = a_0 + \sum_{n=1}^{N}(a_n\cos nt + b_n\sin nt), \tag{8.1}$$

where $N \in \mathbb{N}$, $a_0, a_1, \ldots, a_N, b_1, \ldots, b_N \in \mathbb{C}$, is called a *trigonometric polynomial.*

*Remark.* By Euler's formulas, any trigonometric polynomial (8.1) can be written in a complex form

$$f(t) = \sum_{n=-N}^{N} c_n e^{\mathrm{i}nt} \tag{8.2}$$

with some $c_0, c_{\pm1}, \ldots, c_{\pm N} \in \mathbb{C}$. In fact, we have

$$\cos nt = \frac{1}{2}(e^{\mathrm{i}nt} + e^{-\mathrm{i}nt}), \quad \sin nt = \frac{1}{2\mathrm{i}}(e^{\mathrm{i}nt} - e^{-\mathrm{i}nt}),$$

whence

$$\begin{aligned} a_n\cos nt + b_n\sin nt &= \Big(\frac{1}{2}a_n + \frac{1}{2\mathrm{i}}b_n\Big)e^{\mathrm{i}nt} + \Big(\frac{1}{2}a_n - \frac{1}{2\mathrm{i}}b_n\Big)e^{-\mathrm{i}nt} \\ &= \frac{1}{2}(a_n - \mathrm{i}b_n)e^{\mathrm{i}nt} + \frac{1}{2}(a_n + \mathrm{i}b_n)e^{-\mathrm{i}nt}. \end{aligned}$$

Therefore, the connection between (8.1) and (8.2) is given by the formulas

$$c_n = \begin{cases} a_0 & \text{if } n = 0 \\ a_n - \mathrm{i}b_n & \text{if } n > 0 \\ a_n + \mathrm{i}b_n & \text{if } n < 0. \end{cases}$$

From this, one can readily derive the following conditions for some natural properties of the trigonometric polynomial $f$ given by (8.2):

- $f$ is real-valued if and only if $c_{-n} = \overline{c_n}$ $(0 \le n \le N)$;
- $f$ is even (i.e. all $b_n$'s are zero) if and only if $c_{-n} = c_n$ $\quad(0 \le n \le N)$;
- $f$ is odd (i.e. all $a_n$'s are zero) if and only if $c_{-n} = -c_n$ $(0 \le n \le N)$.

There is another way of calculating the coefficients $c_n$ $(-N \le n \le N)$ directly in terms of $f$. Since $(u_n)_{n\in\mathbb{Z}}$ is orthonormal, these coefficients must be equal to the Fourier coefficients of $f$ (see the Remark after Definition 5.8). Hence,

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-\mathrm{i}nt}\,\mathrm{d}t \quad \text{for every } -N \le n \le N.$$

*Remark.* Instead of the orthonormal system $(u_n)_{n\in\mathbb{Z}}$ we could consider an equivalent system of real-valued functions

$$1,\ \sqrt{2}\cos x,\ \sqrt{2}\sin x, \sqrt{2}\cos 2x,\ \sqrt{2}\sin 2x,\ \ldots \tag{8.3}$$

The factor $\sqrt{2}$ is there to guarantee that the sequence is normalized, i.e. the norm of each of the functions above in $L_2[0, 2\pi]$ equals 1. Recall that we consider the normalized Lebesgue measure $\mathrm{d}\mu = (2\pi)^{-1}\,\mathrm{d}x$ on $[0, 2\pi]$; for the usual Lebesgue measure every sine and cosine should be divided by $\sqrt{\pi}$ and the constant 1 function should be replaced by $(2\pi)^{-1/2}$. Of course, every function above can be uniquely expressed in terms of $u_n$'s and *vice versa*. Therefore, the completeness of (8.3) will follow automatically when we prove the completeness of $(u_n)_{n\in\mathbb{Z}}$ which will be our next goal.