International Summer School: An Introduction to Functional Analysis through Applications

Wiesbaden, September 23-29, 2016

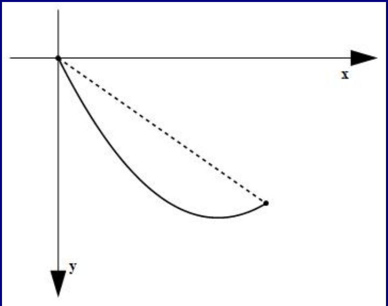
Karlheinz Spindler
Hochschule RheinMain
Kurt-Schumacher-Ring 18
D – 65197 Wiesbaden, Germany
Karlheinz.Spindler@hs-rm.de

The Historical Development of Functional Analysis

- Calculus of variations
- Infinite systems of linear equations
- Integral equations
- Moment problem

The brachistochrone problem (Bernoulli 1696)

Given a point (x_0, y_0) with $x_0 > 0$ and $y_0 > 0$, which curve $x \mapsto y(x)$ satisfying $y(0) = 0$ and $y(x_0) = y_0$ has the property that a point mass, moving from rest and without friction along the curve solely under the influence of gravity, carries out the motion in the shortest possible time?



Mathematical formulation of the brachistochrone problem

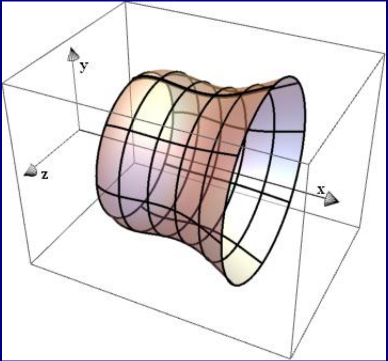
- Motion: $t \mapsto \big(x(t),\, y\big(x(t)\big)\big)$
- Velocity: $v(t) = \sqrt{\dot{x}(t)^2 + y'\big(x(t)\big)\dot{x}(t)^2} = \dot{x}(t)\sqrt{1 + y'(x)^2}$
- Conservation of energy: $(m/2)v(t)^2 = mgy\big(x(t)\big)$
- Consequence: $v(t) = \sqrt{2gy\big(x(t)\big)}$
- Differential equation: $\sqrt{2gy(x)} = \dot{x}\sqrt{1 + y'(x)^2}$
- Separation of variables: $\displaystyle\int_0^T \mathrm{d}t = \int_0^{x_0} \frac{\sqrt{1 + y'(x)^2}}{\sqrt{2gy(x)}}\,\mathrm{d}x$

Hence the brachistochrone problem can be formulated as follows: Find the function $y(x)$ satisfying $y(0) = 0$ and $y(x_0) = y_0$ which minimizes

$$\int_0^{x_0} \sqrt{\frac{1 + y'(x)^2}{y(x)}}\mathrm{d}x.$$

Surfaces of revolution with minimal area (Euler 1744)

Consider two given points (a, y_a) and (b, y_b) in the first quadrant. For each function $y : [a, b] \to (0, \infty)$ satisfying $y(a) = y_a$ and $y(b) = y_b$ we can consider the surface obtained by rotating the curve $y = f(x)$ about the x-axis. For which function y does the surface area become minimal?



Surfaces of revolution with minimal area

Mathematical formulation. Minimize the integral

$$\int_a^b 2\pi y(x)\sqrt{1 + y'(x)^2}\,\mathrm{d}x$$

over all functions y satisfying the boundary conditions $y(a) = y_a$ and $y(b) = y_b$.

Physical interpretation. Hold two wire rings of radii y_a and y_b a distance $b - a$ apart and dip them into a soap solution. Which shape will the soap film have that will form?

Calculus of variations (1)

Generalizing the above problems, we can consider an arbitrary function L in three variables and ask which function y minimizes the expression

$$I[y] := \int_a^b L(x, y(x), y'(x))\,\mathrm{d}x$$

(subject to prescribed boundary conditions). We are used to solving optimization problems in one variable or in several variables, but here we want to minimize an expression I which is not a function of one or several variables, but a function whose argument is itself a function! Such a function is usually called a **functional**. (We may consider I as a function of infinitely many variables, namely all the values $y(x)$ where $a \le x \le b$.) Thus we are led to considering mappings $I : V \to \mathbb{R}$ where V is a space of functions. Formulated in this way, we are already close to basic ideas of functional analysis, but historically, the development went differently.

Calculus of variations (2)

Naively, one simply sought smooth solutions. However, already Euler, in his investigations of vibrations of a string, had to deal with nonsmooth solutions and distinguished between "continuous" (i.e., analytic) and "mechanical" (i.e., piecewise C^2) functions. Similarly, by solving problems in the calculus of variations, one encountered nonsmooth solutions (Goldschmidt solutions for the problem of minimal surfaces of revolution, discontinuities in the derivatives as dealt with by the Weierstraß-Erdmann corner conditions). Very slowly, the need was realized to clearly specify the domain of the functional I (as a well-defined space of functions) and to even reappraise the concept of "function". This process eventually led to concepts like weak solutions of partial differential equations and distributions.

In the classical period, however, no such approach was taken. One simply considered a single function y (of not clearly defined degree of smoothness) and asked which conditions such a function y must satisfy to minimize $I[y]$. Observe that if y_0 minimizes the expression $I[y]$, then, given arbitrary functions $h : [a, b] \to \mathbb{R}$ with $h(a) = h(b) = 0$, the function $\varphi(\varepsilon) := I[y_0 + \varepsilon h]$ takes a minimum at $\varepsilon = 0$. Thus we must have $\varphi'(0) = 0$ and $\varphi''(0) \geq 0$.

The Euler-Lagrange equation

Assume y minimizes $\int_a^b L(x,y,y')\mathrm{d}x$. Letting $\bullet = \big(x, y(x), y'(x)\big)$, we have

$$
\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} \int_a^b L(x, y+\varepsilon h, y'+\varepsilon h')\,\mathrm{d}x \\
&= \int_a^b \left(\frac{\partial L}{\partial y}(\bullet)\cdot h(x) + \frac{\partial L}{\partial y'}(\bullet)\cdot h'(x)\right)\mathrm{d}x \\
&= \int_a^b \left(\frac{\partial L}{\partial y}(\bullet) - \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial L}{\partial y'}(\bullet)\right]\right)\cdot h(x)\,\mathrm{d}x
\end{aligned}
$$

where we used integration by parts in the last step, exploiting the conditions $h(a) = h(b) = 0$. Since this must hold for all possible choices for h, we conclude that

$$
\frac{\partial L}{\partial y} = \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial L}{\partial y'}\right] \quad \text{along the trajectory } x \mapsto \big(x, y(x), y'(x)\big).
$$

This is called the **Euler-Lagrange equation**.

The Legrendre condition

Similarly, for all h with $h(a) = h(b) = 0$, we must have

$$\begin{aligned} 0 &\leq \left.\frac{\mathrm{d}^2}{\mathrm{d}\varepsilon^2}\right|_{\varepsilon=0} \int_a^b L(x, y + \varepsilon h, y' + \varepsilon h')\,\mathrm{d}x \\ &= \int_a^b \left(L_{yy}(\bullet)h(x)^2 + 2L_{yy'}(\bullet)h(x)h'(x) + L_{y'y'}(\bullet)h'(x)^2\right)\,\mathrm{d}x. \end{aligned}$$

Again, this must hold for all possible choices for h, which can be shown to imply the **Legendre condition**

$$L_{y'y'}\big(x, y(x), y'(x)\big) \geq 0.$$

Calculus of variations and functional analysis

- Euler, Lagrange and Legendre derived only necessary optimality conditions.
- Legendre ran into problems when he tried to derive sufficient conditions for optimality.
- In various problems the existence of optimal solutions could not be guaranteed a priori. (This applied in particular to the famous "Dirichlet Principle" in potential theory.)
- When discussing local extrema, one has to specify what is meant by a "nearby" function. This led to a distinction between "weak" and "strong" extrema.
- It became clear that one had to specify the exact range of functions over which an optimization is performed.
- This was not done in the classical period. The idea to study a totality of functions as a single entity could only occur when the necessary set-theoretical concepts were available (Cantor).

Leonhard Euler
(1707-1783)

Joseph-Louis Lagrange
(1736-1813)

Louis Legendre
(1752-1797)

Adrien-Marie Legendre
(1752-1833)

What did Legendre look like? An inverse problem



Fourier's Problem (1822)

Jean Baptiste Joseph Fourier (1768-1830)

Find a solution of the partial differential equation

$$u_{xx} + u_{yy} = 0$$

on $(0, \infty) \times (-\pi/2, \pi/2)$ satisfying the following boundary conditions:

- $u(0, y) = 1$ for all y;
- $u(x, y) \to 0$ for $x \to \infty$ for all y;
- $u(x, -\pi/2) = u(x, \pi/2) = 0$ for all x.

Fourier's solution (1)

To solve the equation $u_{xx} + u_{yy} = 0$, try $u(x,y) = f(x)g(y)$. This yields $f''(x)g(y) + f(x)g''(y) = 0$ for all x, y, hence

$$\frac{f''(x)}{f(x)} \;=\; -\frac{g''(y)}{g(y)} \;=\; \text{const.}$$

Those solutions which also satisfy the homogenous boundary conditions are the functions

$$u_m(x,y) \;=\; e^{-(2m-1)x}\cos\bigl((2m-1)y\bigr) \quad \text{where } m \in \mathbb{N}.$$

Now linear combinations of those solutions are again solutions; hence try $u(x,y) = \sum_{m=1}^{\infty} a_m e^{-(2m-1)x}\cos\bigl((2m-1)y\bigr)$ and adjust the coefficients a_m in such a way that the inhomogeneous boundary condition $u(0,y) = 1$ is satisfied. This leads to the requirement that

$$1 \;=\; \sum_{m=1}^{\infty} a_m \cos\bigl((2m-1)y\bigr) \quad \text{for all } y \in \bigl(-\pi/2,\, \pi/2\bigr).$$

Fourier's solution (2)

Differentiate the equation $1 = \sum_{m=1}^{\infty} a_m \cos\big((2m-1)y\big)$ infinitely often and plug in $y = 0$ each time. This results in the equations

$$1 = \sum_{m=1}^{\infty} a_m, \quad 0 = \sum_{m=1}^{\infty} (2m-1)^2 a_m, \quad 0 = \sum_{m=1}^{\infty} (2m-1)^4 a_m, \ldots$$

which can be formulated as a system of an infinite number of linear equations in infinitely many unknowns $a_1, a_2, a_3, \ldots$, namely

$$\begin{bmatrix} 1 & 1 & 1 & 1 & \cdots \\ 1^2 & 3^2 & 5^2 & 7^2 & \cdots \\ 1^4 & 3^4 & 5^4 & 7^4 & \cdots \\ 1^6 & 3^6 & 5^6 & 7^6 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}.$$

Fourier's solution (3)

For each k, consider the finite system obtained by cutting off all entries after the k-th row or column. For example, for $k = 4$ this gives the system

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1^2 & 3^2 & 5^2 & 7^2 \\ 1^4 & 3^4 & 5^4 & 7^4 \\ 1^6 & 3^6 & 5^6 & 7^6 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Verify that each of these reduced systems has a unique solution $(a_1^{(k)}, a_2^{(k)}, \ldots, a_k^{(k)})$. Verify that

$$a_m = \lim_{k\to\infty} a_m^{(k)} = \frac{(-1)^{m-1} \cdot (\pi/4)}{2m-1}$$

exists and has the specified value. This yields the solution

$$u(x,y) = \frac{\pi}{4} \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{2m-1} e^{-(2m-1)x} \cos\big((2m-1)y\big).$$

Hill's Problem (1886)

George William Hill (1838-1914)

As part of his theory of the moon's motion, Hill tried to describe the motion of the lunar perigee as a function of the mean motions of the sun and the moon. In doing so, he came up with a differential equation

$$\ddot{u}(t) + \theta(t)u(t) = 0$$

where $\theta(t) = \sum_{n=-\infty}^{\infty} \theta_n e^{int}$ with given coefficients θ_n satisfying $\theta_{-n} = \theta_n$ for all n.

Hill's solution (1)

To solve the equation $\ddot{u} + \theta u = 0$, Hill tried the ansatz $u(t) = \sum_{n=-\infty}^{\infty} \lambda_n e^{i(n+c)t}$ with a constant c and coefficients λ_n to be determined. Plugging into the original differential equation and comparing coefficients results in the equations

$$-\lambda_n(n+c)^2 + \sum_{k=-\infty}^{\infty} \theta_{n-k}\lambda_k \;=\; 0$$

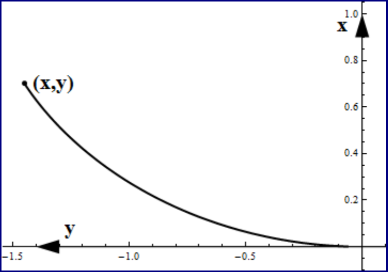
which constitute a doubly infinite system of linear equations. The coefficient matrix is

$$\begin{bmatrix} & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & \theta_1 & \theta_0 - c^2 & \theta_{-1} & \theta_{-2} & \theta_{-3} & \theta_{-4} & \cdots \\ \cdots & \theta_2 & \theta_1 & \theta_0 - (c+1)^2 & \theta_{-1} & \theta_{-2} & \theta_{-3} & \cdots \\ \cdots & \theta_3 & \theta_2 & \theta_1 & \theta_0 - (c+2)^2 & \theta_{-1} & \theta_{-2} & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \end{bmatrix}$$

Abel's Problem (1823)

Niels Henrik Abel (1802-1829)

A particle slides along a curve $y = f(x)$ under the influence of gravity. Determine the shape of the curve for which the time of descent from $(x, f(x))$ takes a prescribed value $T(x)$.



Mathematical formulation of Abel's problem

Let $\big(\xi(t), \eta(t)\big) = \big(\xi(t), f\big(\xi(t)\big)\big)$ the particle's position at time t. By conservation of energy, we have

$$mgx \;=\; mg\xi + \frac{m}{2}(\dot{\xi}^2 + \dot{\eta}^2) \;=\; mg\xi + \frac{m}{2}\dot{\xi}^2\big(1 + f'(\xi)^2\big).$$

Letting $g(x) := \int_0^x \sqrt{1 + f'(\xi)^2}\,\mathrm{d}\xi$, this becomes

$$\sqrt{2g(x-\xi)} \;=\; -\dot{\xi}\sqrt{1 + f'(\xi)^2} \;=\; -\dot{\xi}\, g'(\xi)$$

and hence

$$\sqrt{2g}\, T(x) = \int_0^{T(x)} \sqrt{2g}\,\mathrm{d}t = -\int_x^0 \frac{g'(\xi)}{\sqrt{x-\xi}}\mathrm{d}\xi = \int_0^x \frac{g'(\xi)}{\sqrt{x-\xi}}\mathrm{d}\xi.$$

Given T, find g.

Abel's solution (1)

Abel immediately considered the more general problem of solving the following equation for u in terms of f (where $0 < \alpha < 1$ is fixed and where $f(a) = 0$):

$$f(x) \;=\; \int_a^x \frac{u(\xi)}{(x-\xi)^\alpha}\,\mathrm{d}\xi.$$

Start with the well-known formula

$$\frac{\pi}{\sin(\alpha\pi)} \;=\; \int_\xi^z \frac{\mathrm{d}x}{(z-x)^{1-\alpha}(x-\xi)^\alpha}.$$

Multiply by $u(\xi)$ and integrate from a to z to get

$$\begin{aligned}
\frac{\pi}{\sin(\alpha\pi)} \int_a^z u(\xi)\,\mathrm{d}\xi \;&=\; \int_a^z \int_\xi^z \frac{u(\xi)}{(z-x)^{1-\alpha}(x-\xi)^\alpha}\mathrm{d}x\,\mathrm{d}\xi \\
&=\; \int_a^z \int_a^x \frac{u(\xi)}{(z-x)^{1-\alpha}(x-\xi)^\alpha}\mathrm{d}\xi\,\mathrm{d}x.
\end{aligned}$$

Abel's solution (2)

On the other hand, multiply the equation

$$f(x) \;=\; \int_a^x \frac{u(\xi)}{(x-\xi)^\alpha}\,\mathrm{d}\xi.$$

by $(z-x)^{\alpha-1}$ and integrate from a to z to get

$$\int_a^z \frac{f(x)}{(z-x)^{1-\alpha}}\mathrm{d}x \;=\; \int_a^z \int_a^x \frac{u(\xi)}{(z-x)^{1-\alpha}(x-\xi)^\alpha}\mathrm{d}\xi\,\mathrm{d}x.$$

A comparison with the previous formula yields

$$\int_a^z \frac{f(x)}{(z-x)^{1-\alpha}}\mathrm{d}x \;=\; \frac{\pi}{\sin(\alpha\pi)}\int_a^z u(\xi)\,\mathrm{d}\xi$$

and hence

$$u(z) \;=\; \frac{\sin(\alpha\pi)}{\pi}\cdot\frac{\mathrm{d}}{\mathrm{d}z}\left(\int_a^z \frac{f(x)}{(z-x)^{1-\alpha}}\mathrm{d}x\right).$$

Sturm and Liouville

Jacques Charles François Sturm (1803-1855)

Joseph Liouville (1809-1882)

Liouville 1837

The linear differential equation $y'' + \rho^2 y = \varphi$ has the general solution

$$y(x) = A\sin\big(\rho(x-a)\big) + B\cos\big(\rho(x-a)\big) + \frac{1}{\rho}\int_a^x \sin\big(\rho(x-a)\big)\varphi(\xi)\,\mathrm{d}\xi.$$

Liouville applied this formula to the initial value problem

$$u''(x) + \big(\rho^2 - \sigma(x)\big)u(x) = 0, \quad u(a) = 1, \quad u'(a) = 0$$

to obtain

$$u(x) \;=\; \cos\big(\rho(x-a)\big) + \frac{1}{\rho}\int_a^x \sin\big(\rho(x-\xi)\big)\sigma(\xi)u(\xi)\,\mathrm{d}\xi,$$

thereby replacing an initial value problem by an integral equation. Liouville wanted to find a series expansion for u which converges rapidly for large values of ρ. He had the fortuitious idea to plug in the right-hand side of this equation into the expression $u(\xi)$ under the integral sign.

Integral equations

Liouville's method works quite generally, not only in the example he considered. Therefore, let us write down the general form of integral equations we want to study.

- $f(x) = \int_a^x K(x,\xi)u(\xi)\,\mathrm{d}\xi$ (type studied by Abel)
- $u(x) = f(x) + \int_a^x K(x,\xi)u(\xi)\,\mathrm{d}\xi$ (type studied by Liouville)

Even more generally, we can consider the following types of integral equations:

- $f(x) = \int_a^b K(x,\xi)u(\xi)\,\mathrm{d}\xi$
- $u(x) = f(x) + \int_a^b K(x,\xi)u(\xi)\,\mathrm{d}\xi$

These reduce to the types studied by Abel and Liouville if $K(x,\xi) = 0$ for $x > \xi$. Liouville's method works for the last kind shown.

Liouville's method

$$
\begin{aligned}
u(x) &= f(x) + \int_a^b K(x,\xi)u(\xi)\,\mathrm{d}\xi \\
u(x) &= f(x) + \int_a^b K(x,\xi)\left[f(\xi) + \int_a^b K(\xi,\widehat{\xi})u(\widehat{\xi})\,\mathrm{d}\widehat{\xi}\right]\mathrm{d}\xi \\
u(x) &= f(x) + \int_a^b K(x,\xi)f(\xi)\,\mathrm{d}\xi + \int_a^b\int_a^b K(x,\xi)K(\xi,\widehat{\xi})u(\widehat{\xi})\mathrm{d}\widehat{\xi}\,\mathrm{d}\xi
\end{aligned}
$$

Iterate the procedure of plugging in the full expression for u into the expression fo u under the integral sign. This gives a series representation

$$
u(x) = f(x) + \sum_{k=1}^{\infty}\int_a^b\cdots\int_a^b K(x,\xi_1)\cdots K(\xi_{k-1},\xi_k)f(\xi_k)\,\mathrm{d}\xi_k\cdots\mathrm{d}\xi_1
$$

("Volterra series"). If this series converges uniformly on some interval, it respresents a solution of the integral equation.

Convergence of the Volterra series (1)

Consider first the case that $K(x,\xi) = 0$ for $x > \xi$ so that $u(x) = f(x) + \int_a^x K(x,\xi)u(\xi)\,\mathrm{d}\xi$. The Volterra series is

$$f(x) + \int_a^x K(x,\xi_1)f(\xi_1)\,\mathrm{d}\xi_1 + \int_a^x \int_a^{\xi_1} K(x,\xi_1)K(\xi_1,\xi_2)f(\xi_2)\,\mathrm{d}\xi_2\,\mathrm{d}\xi_1$$
$$+ \int_a^x \int_a^{\xi_1} \int_a^{\xi_2} K(x,\xi_1)K(\xi_1,\xi_2)K(\xi_2,\xi_3)f(\xi_3)\,\mathrm{d}\xi_3\,\mathrm{d}\xi_2\,\mathrm{d}\xi_1 + \cdots .$$

This series is majorized by

$$\|f\| + \|K\|\,\|f\|(x-a) + \|K\|^2\|f\|\frac{(x-a)^2}{2!} + \|K\|^3\|f\|\frac{(x-a)^3}{3!} + \cdots$$

which is $\|f\|\exp\big(\|K\|(x-a)\big)$. Hence the series is uniformly convergent and thus represents a solution of the integral equation.

Convergence of the Volterra series (2)

The general case $u(x) = f(x) + \int_a^b K(x,\xi)u(\xi)\,\mathrm{d}\xi$ is similar. The Volterra series is

$$f(x) + \int_a^b K(x,\xi_1)f(\xi_1)\,\mathrm{d}\xi_1 + \int_a^b \int_a^b K(x,\xi_1)K(\xi_1,\xi_2)f(\xi_2)\,\mathrm{d}\xi_2\,\mathrm{d}\xi_1 + \int_a^b \int_a^b \int_a^b K(x,\xi_1)K(\xi_1,\xi_2)K(\xi_2,\xi_3)f(\xi_3)\,\mathrm{d}\xi_3\,\mathrm{d}\xi_2\,\mathrm{d}\xi_1 + \cdots .$$

This series is majorized by

$$\|f\| + \|K\|\,\|f\|(b-a) + \|K\|^2\|f\|(b-a)^2 + \|K\|^3\|f\|(b-a)^3 + \cdots$$

which converges if $\|K\|(b-a) < 1$ (geometric series). With hindsight, we recognize glimpses of the the contraction property, the Banach fixed point theorem and the Neumann series.

Vito Volterra (1860-1940)

Volterra's 1896 solution (1)

Again, consider $u(x) = f(x) + \int_a^b K(x,\xi)u(\xi)\,\mathrm{d}\xi$. Let $K_1 := K$ and

$$K_{i+1}(x,y) \;\; := \;\; \int_a^b K(x,\xi)K_{i-1}(\xi,y)\,\mathrm{d}\xi.$$

Inductively, we see that

$$K_i(x,y) \;\; = \;\; \int_a^b \cdots \int_a^b K(x,\xi_1)\cdots K(\xi_{i-1},y)\,\mathrm{d}\xi_{i-1}\cdots\mathrm{d}\xi_1.$$

The kernels (K_i) satisfy the semigroup property

$$K_{i+j}(x,y) \;\; = \;\; \int_a^b K_i(x,\xi)K_j(\xi,y)\,\mathrm{d}\xi.$$

Note that if $K(x,\xi)$ vanishes for $x > \xi$ then so does K_i for any i, which entails that $K_{i+j}(x,y) = \int_y^x K_i(x,\xi)K_j(\xi,y)\,\mathrm{d}\xi$. Only this case was considered by Volterra.

Volterra's 1896 solution (2)

Assume that $K^\star := -\sum_{i=1}^{\infty} K_i$ exists. Then the remainder $R_n := \sum_{i=n+1}^{\infty} K_i$ satisfies

$$(\star) \quad R_n(x,y) \;=\; \sum_{i=n+1}^{\infty} \int_a^b K_{i-j}(x,\xi) K_j(\xi,y)\,\mathrm{d}\xi \quad \big(j = j(i)\big)$$

Since $-K^\star = \sum_{i=n+1}^{\infty} K_{i-n}$, we have

$$-K^\star(x,\xi) \;=\; \sum_{i=n+1}^{\infty} K_{i-n}(x,\xi).$$

Multiply by $K_n(\xi,y)$ and integrate over $a \le \xi \le b$ to obtain

$$-\int_a^b K^\star(x,\xi) K_n(\xi,y)\,\mathrm{d}\xi = \sum_{i=n+1}^{\infty} \int_a^b K_{i-n}(x,\xi) K_n(\xi,y)\,\mathrm{d}\xi = R_n(x,y)$$

using $j(i) = n$ for all i in $(\star)$.

Volterra's 1896 solution (3)

Recall

$$(\star) \quad R_n(x,y) \;=\; \sum_{i=n+1}^{\infty} \int_a^b K_{i-j}(x,\xi)K_j(\xi,y)\,\mathrm{d}\xi \quad \bigl(j = j(i)\bigr)$$

Since $-K^\star = \sum_{i=n+1}^{\infty} K_{i-n}$, we have

$$-K^\star(\xi,y) \;=\; \sum_{i=n+1}^{\infty} K_{i-n}(\xi,y).$$

Multiply by $K_n(x,\xi)$ and integrate over $a \le \xi \le b$ to obtain

$$-\int_a^b K_n(x,\xi)K^\star(\xi,y)\,\mathrm{d}\xi = \sum_{i=n+1}^{\infty} \int_a^b K_n(x,\xi)K_{i-n}(\xi,y)\,\mathrm{d}\xi = R_n(x,y)$$

using $j(i) = i - n$ for all i in $(\star)$.

Volterra's 1896 solution (4)

Equating the two formulas, we have

$$-R_n(x,y) = \int_a^b K_n(x,\xi)K^\star(\xi,y)\,\mathrm{d}\xi = \int_a^b K^\star(x,\xi)K_n(\xi,y)\,\mathrm{d}\xi.$$

For $n=1$ we have $K_1 = K$ and $R_1 = -K^\star - K$. Hence

$$\begin{aligned} K(x,y)+K^\star(x,y) &= \int_a^b K(x,\xi)K^\star(\xi,y)\,\mathrm{d}\xi \\ &= \int_a^b K^\star(x,\xi)K(\xi,y)\,\mathrm{d}\xi. \end{aligned}$$

We call $K^\star$ reciprocal to K if this relation holds. Note that then also K is reciprocal to $K^\star$.

Volterra's 1896 solution (5)

The integral equation to be solved is $u(x) = f(x) + \int_a^b K(x,\xi)u(\xi)\,\mathrm{d}\xi$, i.e,

$$u(\xi) = f(\xi) + \int_a^b K(\xi,\xi_1)u(\xi_1)\,\mathrm{d}\xi_1.$$

Assume $K^\star$ is reciprocal to K. Multiply by $K^\star(x,\xi)$ and integrate over $a \le \xi \le b$ to see that $\int_a^b K^\star(x,\xi)u(\xi)\,\mathrm{d}\xi$ equals

$$\begin{aligned}
& \int_a^b K^\star(x,\xi)f(\xi)\,\mathrm{d}\xi + \int_a^b\int_a^b K^\star(x,\xi)K(\xi,\xi_1)u(\xi_1)\,\mathrm{d}\xi_1\,\mathrm{d}\xi \\
= & \int_a^b K^\star(x,\xi)f(\xi)\,\mathrm{d}\xi + \int_a^b\int_a^b K^\star(x,\xi)K(\xi,\xi_1)u(\xi_1)\,\mathrm{d}\xi\,\mathrm{d}\xi_1 \\
= & \int_a^b K^\star(x,\xi)f(\xi)\,\mathrm{d}\xi + \int_a^b \big(K^\star(x,\xi_1) + K(x,\xi_1)\big)u(\xi_1)\,\mathrm{d}\xi_1 \\
= & \int_a^b K^\star(x,\xi)f(\xi)\,\mathrm{d}\xi + \int_a^b K^\star(x,\xi)u(\xi)\,\mathrm{d}\xi + \int_a^b K(x,\xi)u(\xi)\,\mathrm{d}\xi.
\end{aligned}$$

Volterra's 1896 solution (6)

Subtracting $\int_a^b K^\star(x,\xi)u(\xi)\,\mathrm{d}\xi$ yields

$$0 = \int_a^b K^\star(x,\xi)f(\xi)\,\mathrm{d}\xi + \int_a^b K(x,\xi)u(\xi)\,\mathrm{d}\xi.$$

Hence

$$u(x) = f(x) + \int_a^b K(x,\xi)u(\xi)\,\mathrm{d}\xi$$

if and only if

$$u(x) = f(x) - \int_a^b K^\star(x,\xi)f(\xi)\,\mathrm{d}\xi.$$

The contributions of Fredholm and Hilbert

Ivar Fredholm (1866-1927)

David Hilbert (1862-1943)

Discretization (1)

We want to solve the integral equation

$$u(x) \quad = \quad f(x) + \int_a^b K(x,\xi)u(\xi)\,\mathrm{d}\xi.$$

Subdivide the interval $[a,b]$ as $a = x_0 < x_1 < x_2 < \cdots < x_n = b$ where $x_k - x_{k-1} = (b-a)/n =: \delta$. Replace the integral by a Riemann sum to get

$$u_n(x) \quad = \quad f(x) + \sum_{i=1}^{n} K(x,x_j)u_n(x_j)\,\delta.$$

Let $k_{ij} := \delta K(x_i,x_j)$ and evaluate at the points $x = x_i$ to get

$$\begin{bmatrix} 1-k_{11} & -k_{12} & \cdots & -k_{1n} \\ -k_{21} & 1-k_{22} & \cdots & -k_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -k_{n1} & -k_{n2} & \cdots & 1-k_{nn} \end{bmatrix} \begin{bmatrix} u_n(x_1) \\ u_n(x_2) \\ \vdots \\ u_n(x_n) \end{bmatrix} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{bmatrix}$$

Discretization (2)

Solve this system by using Cramer's Rule:

$$u_n(x_i) = \frac{f(x_1)D_n(x_i,x_1) + f(x_2)D_n(x_i,x_2) + \cdots + f(x_n)D_n(x_i,x_n)}{\Delta_n}$$

where Δ_n is the coefficient determinant and where $D_n(x_i,x_j)$ is the cofactor of this determinant with respect to the ij-entry. Expansion shows that

$$\begin{aligned}\Delta_n &= 1 - \sum_{i=1}^{n} \delta\, K(x_i,x_i) + \frac{1}{2!}\sum_{i,j=1}^{n} \delta^2 \begin{vmatrix} K(x_i,x_i) & K(x_i,x_j) \\ K(x_j,x_i) & K(x_j,x_j) \end{vmatrix} \\ &\quad - \frac{1}{3!}\sum_{i,j,k=1}^{n} \delta^3 \begin{vmatrix} K(x_i,x_i) & K(x_i,x_j) & K(x_i,x_k) \\ K(x_j,x_i) & K(x_j,x_j) & K(x_j,x_k) \\ K(x_k,x_i) & K(x_k,x_j) & K(x_k,x_k) \end{vmatrix} \pm \cdots .\end{aligned}$$

Discretization (3)

... and also

$$D_n(x_i, x_j)/\delta \;=\; K(x_i, x_j) - \sum_{k=1}^{n} \delta \begin{vmatrix} K(x_i, x_j) & K(x_i, x_k) \\ K(x_k, x_i) & K(x_k, x_k) \end{vmatrix} + \frac{1}{2!} \sum_{k,\ell=1}^{n} \delta^2 \begin{vmatrix} K(x_i, x_j) & K(x_i, x_k) & K(x_i, x_\ell) \\ K(x_k, x_i) & K(x_k, x_k) & K(x_k, x_\ell) \\ K(x_\ell, x_i) & K(x_\ell, x_j) & K(x_\ell, x_\ell) \end{vmatrix} \pm \cdots .$$

for $i \neq j$ whereas $D_n(x_i, x_i) \approx \Delta_n$. We now let $n \to \infty$ while allowing i and j to vary in such a way that $(x_i, x_j) \to (x, y)$.

Discretization (4)

Then

$$\Delta_n = 1 - \sum_{i=1}^{n} \delta\, K(x_i, x_i) + \frac{1}{2!} \sum_{i,j=1}^{n} \delta^2 \begin{vmatrix} K(x_i, x_i) & K(x_i, x_j) \\ K(x_j, x_i) & K(x_j, x_j) \end{vmatrix}$$

$$- \frac{1}{3!} \sum_{i,j,k=1}^{n} \delta^3 \begin{vmatrix} K(x_i, x_i) & K(x_i, x_j) & K(x_i, x_k) \\ K(x_j, x_i) & K(x_j, x_j) & K(x_j, x_k) \\ K(x_k, x_i) & K(x_k, x_j) & K(x_k, x_k) \end{vmatrix} \pm \cdots$$

$$\to \; 1 - \int_a^b K(\xi_1, \xi_1)\, \mathrm{d}\xi_1 + \frac{1}{2!} \int_a^b \int_a^b \begin{vmatrix} K(\xi_1, \xi_1) & K(\xi_1, \xi_2) \\ K(\xi_2, \xi_1) & K(\xi_2, \xi_2) \end{vmatrix} \mathrm{d}\xi_2\, \mathrm{d}\xi_1$$

$$- \frac{1}{3!} \int_a^b \int_a^b \int_a^b \begin{vmatrix} K(\xi_1, \xi_1) & K(\xi_1, \xi_2) & K(\xi_1, \xi_3) \\ K(\xi_2, \xi_1) & K(\xi_2, \xi_2) & K(\xi_2, \xi_3) \\ K(\xi_3, \xi_1) & K(\xi_3, \xi_2) & K(\xi_3, \xi_3) \end{vmatrix} \mathrm{d}\xi_3\, \mathrm{d}\xi_2\, \mathrm{d}\xi_1 \pm \cdots$$

Discretization (5)

... and

$$
\begin{aligned}
D_n(x_i, x_j)/\delta &= K(x_i, x_j) - \sum_{k=1}^{n} \delta \begin{vmatrix} K(x_i, x_j) & K(x_i, x_k) \\ K(x_k, x_i) & K(x_k, x_k) \end{vmatrix} \\
&\quad + \frac{1}{2!} \sum_{k,\ell=1}^{n} \delta^2 \begin{vmatrix} K(x_i, x_j) & K(x_i, x_k) & K(x_i, x_\ell) \\ K(x_k, x_i) & K(x_k, x_k) & K(x_k, x_\ell) \\ K(x_\ell, x_i) & K(x_\ell, x_j) & K(x_\ell, x_\ell) \end{vmatrix} \pm \cdots \\
&\to \; K(x, y) - \int_a^b \begin{vmatrix} K(x, y) & K(x, \xi_1) \\ K(\xi_1, y) & K(\xi_1, \xi_1) \end{vmatrix} \mathrm{d}\xi_1 \\
&\quad + \frac{1}{2!} \int_a^b \int_a^b \begin{vmatrix} K(x, y) & K(x, \xi_1) & K(x, \xi_2) \\ K(\xi_1, y) & K(\xi_1, \xi_1) & K(\xi_1, \xi_2) \\ K(\xi_2, y) & K(\xi_2, \xi_1) & K(\xi_2, \xi_2) \end{vmatrix} \mathrm{d}\xi_2 \, \mathrm{d}\xi_1 \mp \cdots \\
&=: D(x, y) \text{ for } i \neq j.
\end{aligned}
$$

Discretization (6)

On the other hand, $D_n(x_i, x_i) \to \Delta$ as $n \to \infty$. Let $n \to \infty$ in the equation

$$u_n(x_i) \quad = \quad \frac{f(x_1)D_n(x_i, x_1) + f(x_2)D_n(x_i, x_2) + \cdots + f(x_n)D_n(x_i, x_n)}{\Delta_n}$$

to obtain

$$u(x) \quad = \quad f(x) + \frac{1}{\Delta} \int_a^b D(x, \xi) f(\xi) \, \mathrm{d}\xi.$$

This was the solution Fredholm found for the integral equation $u(x) = f(x) + \int_a^b K(x, \xi) u(\xi) \, \mathrm{d}\xi$. Fredholm used the above arguments only heuristically to find his solution und justified the solution differently. It was Hilbert who showed that all steps in the derivation can be made rigorous.

The moment problem

Chebysheff 1855: Is the normal distribution characterized by its moments? More concretely, can we conclude from

$$\int_{-\infty}^{\infty} f(x)x^n \mathrm{d}x = \int_{-\infty}^{\infty} e^{-x^2} x^n \mathrm{d}x \text{ for all } n \in \mathbb{N}_0$$

that $f(x) = e^{-x^2}$?

Stieltjes 1894: Given a sequence (c_n) of real numbers, is there an increasing function g such that $\int_a^b x^n \mathrm{d}g(x) = c_n$ for all $n \in \mathbb{N}$?

Riesz 1910: Let $(1/p) + (1/q) = 1$. Given functions $f_k \in L^q$ and real numbers c_k, is there a function $g \in L^p$ such that $\int_a^b f_k g = c_k$ for all $k \in \mathbb{N}$?

The moment problem and linear functionals (1)

With hindsight, the moment problem can be seen as a problem of the following kind: Given a Banach space X, a sequence (x_n) in X and a sequence (c_n) in $\mathbb{K}$, is there a continuous linear functional $f : X \to \mathbb{K}$ such that $f(x_n) = c_n$ for all n? Similarly, given an infinite system of linear equations

$$\sum_{k=1}^{\infty} a_{ik} x_k = b_i,$$

we may consider the "row vectors" $a_i = (a_{i1}, a_{i2}, a_{i3}, \ldots)$ as elements of a Banach space X and identify $f := (x_1, x_2, x_3, \ldots)$ with an element of $X^\star$. Solving the system then amounts to finding $f \in X^\star$ such that $f(a_i) = b_i$ for all i. Then for all $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ we have

$$\left| \sum_{i=1}^{n} \lambda_i b_i \right| = \left| f\left(\sum_{i=1}^{n} \lambda_i a_i \right) \right| \leq \|f\|_{\mathrm{op}} \left\| \sum_{i=1}^{n} \lambda_i a_i \right\|.$$

The moment problem and linear functionals (2)

Hence a necessary condition for the solvability of the system is the existence of a constant $M \geq 0$ such that

$$\left| \sum_{i=1}^{n} \lambda_i b_i \right| \leq M \cdot \left\| \sum_{i=1}^{n} \lambda_i a_i \right\|$$

for all $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$. The fact that this condition is also sufficient was established step by step:

- for ℓ_2 by Schmidt in 1908;
- for $L_p[a, b]$ where $1 < p < \infty$ by Riesz in 1909;
- for $C[a, b]$ by Riesz (1911) and Helly (1912);
- for ℓ_p where $1 < p < \infty$ by Riesz in 1913;
- for arbitrary sequence spaces by Helly in 1921;
- in the general case by Hahn (1927) and Banach (1929).

The moment problem and linear functionals (3)

The idea of the sufficiency proof was to choose $f_n \in X^\star$ with minimal norm $\|f_n\| \leq M$ such that $f_n(a_i) = b_i$ for $1 \leq i \leq n$. In the separable case, the sought functional f is obtained by taking the limit of a weakly* convergent subsequence of (f_n). However, already Riesz (1909) realized how the nonseparable case could be handled: "Unsere Resultate bleiben auch für Systeme bestehen, die mehr als abzählbar viele Gleichungen enthalten." To satisfy $f(a_i) = b_i$ for an arbitrary index set I, there must be a constant $M \geq 0$ such that

$$|\sum_{i \in I_0} \lambda_i b_i| \leq M \cdot \|\sum_{i \in I_0} \lambda_i a_i\|$$

for all $\lambda_i \in \mathbb{K}$ over a finite index set I_0. If X_0 is the span of the elements a_i where $i \in I_0$ we define $f_0(\sum_{i \in I_0} \lambda_i a_i) := \sum_{i \in I_0} \lambda_i b_i$ and have $f_0 \in X_0^\star$ with $\|f_0\| \leq M$. Hence solving the moment problem is equivalent to extending f_0 in a norm-preserving way. This led eventually to the Hahn-Banach extension theorem.

Duality

Given a normed space V (or, more generally, a topological vector space V), we denote by $V^\star$ the space of all continuous linear forms on V. The idea to deduce information on $x \in V$ from the values $f(x)$ where $f \in V^\star$ corresponds to treating the state x of a system as a "black box" about which nothing is known except the values of measurements performed on the system. This approach has analogies to the treatment of observers in quantum mechanics (which is one of the first fields to which functional analytic methods were applied) and also to the theory of distributions (where a distribution is determined by its effect on test functions).

Notations for the effect of a linear form on an element of a function space:

- Hadamard 1903: $U[f(x)]$
- Riesz 1911: $A[f(x)]$
- Helly 1912: $U[f]$

The genesis of functional analysis around 1900 (1)

- Problems were studied which could be tackled by a transition from a finite-dimensional to an infinite-dimensional setting.
- In particular, spaces of sequences and spaces of functions were studied (without initially calling them so). Sequences and functions were treated as vectors with an infinite number of components.
- It became necessary to consider families of objects (such as functions) as entities in their own right. (A good example is the equicontinuity of a family of functions.) It was of fundamental importance that set-theoretical terminology and concepts had been developed (Cantor).
- Typically, special problems were treated in an ad hoc manner, but structural similarities between different kinds of problems were soon discovered and stimulated an interest in "mathematical structures".

The genesis of functional analysis around 1900 (2)

- It became apparent that problems could be made more accessible by ignoring special features of the concrete problem under consideration and inserting the problem into a more general context.
- Since many problems were solved by approximation procedures, the concept of convergence needed to be corroborated, and new concepts such as weak convergence evolved. Eventually, this needed the definition of topological concepts.
- There was a tendency to question basic concepts such as that of a function (Euler, Goldschmidt, Cantor) or that of a number (Dedekind) and to revise mathematical foundations such as logic (Frege, Russell), set theory (Cantor), or geometry (Hilbert). This certainly helped to lay the foundations of topology (Frechet, Hausdorff).
- The simultaneous development of Lebesgue's theory of integration helped enormously the study of integral equations (for example due to the completeness of L^p spaces).

Linear algebra and functional analysis

In our teaching, we can prepare functional analytic ideas from the very first semester on, obviously in analysis courses, but also in linear algebra courses, by regularly incorporating examples involving infinite-dimensional spaces (typically spaces of sequences and functions). However, this does not reflect the historical development.

Jean Dieudonné: History of Functional Analysis:
Unfortunately, linear algebra, as it was understood in the XIXth century (and even much later) did not readily lend itself to affording a good guidance to such generalizations. Its own evolution had been very slow and painful, stretching over 130 years, and in a succession of stages which, to our eyes, is exactly the **reverse** of the **logical** sequences of notions, namely

linear equations $\rightarrow$ determinants $\rightarrow$ linear and bilinear forms
$\rightarrow$ matrices $\rightarrow$ vector spaces and linear maps

Stefan Banach's Thesis (1)

The thesis was submitted to the University of Lwów in June 1920 and published two years later. (Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales, Fundamenta Mathematicae **3** (1922), pp. 133-181.)

Soit E une classe composée tout au moins de deux éléments, d'ailleurs arbitraires, que nous désignerons p.e. par $X, Y, Z, \ldots$ a, b, c désignant les nombres réels quelquonques, nous définissons pour E deux opérations suivantes:

1) l'addition des éléments de E

$$X + Y, X + Z, \ldots$$

2) la multiplication des éléments de E par un nombre réel

$$a \cdot X, b \cdot Y, \ldots$$

Admettons que les propriétés suivantes sont réalisées:

Stefan Banach's Thesis (2)

- I_1 $X + Y$ est un élément bien déterminé de la classe E,
- I_2 $X + Y = Y + X$,
- I_3 $X + (Y + Z) = (X + Y) + Z$,
- I_4 $X + Y = X + Z$ entraîne $Y = Z$,
- I_5 Il existe un élément de la classe E déterminé θ et tel qu'on ait toujours $X + \theta = X$,
- I_6 $a \cdot X$ est un élément bien déterminé de la classe E,
- I_7 $a \cdot X = \theta$ équivaut à $X = \theta$ ou $a = 0$,
- I_8 $a \neq 0$ et $a \cdot X = a \cdot Y$ entraînent $X = Y$,
- I_9 $X \neq \theta$ et $a \cdot X = b \cdot X$ entraînent $a = b$,
- I_{10} $a \cdot (X + Y) = a \cdot X + a \cdot Y$,
- I_{11} $(a + b) \cdot X = a \cdot X + b \cdot X$,
- I_{12} $1 \cdot X = X$,
- I_{13} $a \cdot (b \cdot X) = (a \cdot b) \cdot X$.

Stefan Banach's Thesis (3)

Admettons ensuite qu'il existe une opération appelée norme (nous la désignerons par le symbole $\|X\|$), définie dans le champ E, ayant pour contre-domaine l'ensemble de nombres réels et satisfaisant aux conditions suivantes:

II_1 $\|X\| \geq 0$,

II_2 $\|X\| = 0$ équivaut à $X = \theta$,

II_3 $\|a \cdot X\| = |a| \cdot \|X\|$,

II_4 $\|X + Y\| \leq \|X\| + \|Y\|$,

III Si

1^0 $\{X_n\}$ est une suite d'éléments de E

2^0 $lim_{\substack{r=\infty \\ p=\infty}} \|X_r - X_p\| = 0$,

il existe un élément X tel que $\lim_{n=\infty} \|X - X_n\| = 0$.

Norbert Wiener (1894-1964)

On the theory of sets of points in terms of continuous transformations, C. R. Congrès Int. Math., Strasbourg 1920, pp. 312-315

A **vector system** is defined as a system K of elements correlated with a system σ of entities and the operations $\oplus$, $\odot$, and $\|\cdot\|$ in a manner indicated by the following propositions:

(1) If ξ and η belong to σ, $\xi \oplus \eta$ belongs to σ,

(2) If ξ belongs to σ, and n is a real number ≥ 0, $n \odot \xi$ belongs to σ,

(3) If ξ belongs to σ, $\|\xi\|$ is a real number ≥ 0,

(4) $n\odot(\xi\oplus\eta) = (n\odot\xi)\oplus(n\odot\eta)$,

(5) $(m \odot \xi) \oplus (n \odot \xi) = (m + n) \odot \xi$,

(6) $\|n \odot \xi\| = n\|\xi\|$,

(7) $\|\xi \oplus \eta\| \leq \|\xi\| + \|\eta\|$,

(8) $m \odot (n \odot \xi) = mn \odot \xi$,

(9) If A and B belong to K, there is associated with them a single member AB of σ,

(10) $\|AB\| = \|BA\|$,

(11) Given an element A of K and an element ξ of σ, there is an element B of K such that $AB = \xi$,

(12) $AC = AB \oplus BC$,

(13) $\|AB\| = 0$ when and only when $A = B$,

(14) If $AB = CD$, $BA = DC$.

Linear algebra and functional analysis (1)

- It is amazing to see that even around 1920 the concept of "vector space" was not a well-established one, despite work by Cayley, Hamilton, Graßmann, Bellavitis, Cauchy, Saint-Venant, Frobenius, Jordan, Tait, Peirce, Maxwell, Clifford, Gibbs, Heaviside, Peano, Pincherle, and others.
- Nowadays, we teach linear algebra in a first-semester course. Therefore, we have the luxury to prepare functional analytic ideas by incorporating examples in which infinite-dimensional spaces occur.
- Historically, the need to put functional analytic ideas on a firm basis in fact helped to formulate an abstract theory of vector spaces and linear mappings.

Linear algebra and functional analysis (2)

"All this was to weigh heavily on the evolution of linear Functional Analysis; in particular it followed (over a shorter span of years) the same unfortunate succession of states through which Linear Algebra had to go; and it was only after it was realized that the current conception of vectors as "*n*-tuples" could not possible be extended to infinite-dimensional function spaces, that this conception was finally abandoned and that genuinely geometrical notions won the day." (Dieudonné, History of Functional Analysis)

Hamilton and Grassmann

Sir William Rowan Hamilton
(1805-1865)

Hermann Günter Grassmann
(1809-1877)

Title page of Grassmann's "Ausdehnungslehre" (1844):
Hermann Grassmann,
Lehrer an der Friedrich-Wilhelms-Schule zu Stettin

Title page of Hamilton's "Lectures on Quaternions" (1853):
Sir William Rowan Hamilton, LL. D., M.R.I.A., Fellow of the American Society of Arts and Sciences; of the Society of Arts for Scotland; of the Royal Astronomical Society of London; and of the Royal Northern Society of Antiquaries at Copenhagen; Corresponding Member of the Institute of France; Honorary or Corresponding Member of the Imperial or Royal Academies of St. Petersburgh, Berlin, and Turin; of the Royal Societies of Edinburgh and Dublin; of the Cambridge Philosophical Society; the New York Historical Society; the Society of Natural Sciences at Lausanne; and of Other Scientific Societies in British and Foreign Countries; Andrews' Professor of Astronomy in the University of Dublin; and Royal Astronomer of Ireland.

D. Fearnley-Sander: Hermann Grassmann and the creation of linear algebra, Amer. Math. Monthly 86 (10) (1979), 809-817:

Beginning with a collection of "units" e_1, e_2, e_3, ... he effectively defines the free linear space which they generate; that is to say, he considers formal linear combinations $a_1e_1 + a_2e_2 + a_3e_3 + \ldots$ where the a_j are real numbers, defines addition and multiplication by real numbers [in what is now the usual way] and formally proves the linear space properties for these operations. ... He then develops the theory of linear independence in a way which is astonishingly similar to the presentation one finds in modern linear algebra texts. He defines the notions of subspace, independence, span, dimension, join and meet of subspaces, and projections of elements onto subspaces. He is aware of the need to prove invariance of dimension under change of basis, and does so. He proves the Steinitz Exchange Theorem, named for the man who published it in 1913 ... Among other such results, he shows that any finite set has an independent subset with the same span and that any independent set extends to a basis, and he proves the important identity $\dim(U+W) = \dim(U) + \dim(W) - \dim(U \cap W)$. He obtains the formula for change of coordinates under change of basis, defines elementary transformations of bases, and shows that every change of basis (equivalently, in modern terms, every invertible linear transformation) is a product of elementaries.

Uniform Boundedness Principle (1)

Stefan Banach
(1892-1945)

Hugo Dyonizy
Steinhaus (1887-1972)

Uniform Boundedness Principle (2)

Theorem. Let $\mathfrak{M}$ be a family of operators $T : X \to Y$. Assume that for each $x \in X$ there is a constant $c(x)$ such that $\|Tx\| \leq c(x)$ for all $T \in \mathfrak{M}$. Then there is a constant $C > 0$ such that $\|T\|_{op} \leq C$ for all $T \in \mathfrak{M}$.

Proof. Assume not. Then we can choose $T_1, T_2, T_3, \ldots \in \mathfrak{M}$ and $x_1, x_2, x_3, \ldots \in X$ such that

$$\|T_n\|_{op} \geq 4 \cdot 3^n \left[n + \sum_{k<n} c(x_k) \right], \; \|x_n\| \leq \frac{1}{3^n}, \; \|T_n x_n\| \geq \frac{3\|T_n\|_{op}}{4 \cdot 3^n}.$$

Namely, once T_n is chosen, pick ξ_n with $\|\xi_n\| \leq 1$ and $\|T_n \xi_n\| \geq (3/4) \cdot \|T_n\|_{op}$ and let $x_n := \xi_n / 3^n$. As a consequence, $x := \sum_{k=1}^{\infty} x_k$ exists, and $T_n x = \sum_k T_n x_k$ gets its main contribution from $T_n x_n$ ("gliding hump").

Uniform Boundedness Principle (3)

More precisely, $\|\sum_{k<n} T_n x_k\| \leq \sum_{k<n} \|T_n x_k\| \leq \sum_{k<n} c(x_k)$ and

$$\left\|\sum_{k>n} T_n x_k\right\| \leq \sum_{k>n} \|T_n x_k\| \leq \sum_{k>n} \|T_n\|_{op} \|x_k\| \leq \sum_{k>n} \frac{\|T\|_{op}}{3^k} = \frac{3\|T_n\|_{op}}{2 \cdot 3^n}.$$

Since $T_n x = \sum_{k=1}^{\infty} T_n x_k = \sum_{k<n} T_n x_k + T_n x_n + \sum_{k>n} T_n x_k$, we have

$$\begin{aligned} \|T_n x\| &\geq -\sum_{k<n} c(x_k) + \|T_n x_n\| - \frac{\|T_n\|_{op}}{2 \cdot 3^n} \\ &\geq -\sum_{k<n} c(x_k) + \frac{3\|T_n\|_{op}}{4 \cdot 3^n} - \frac{\|T_n\|_{op}}{2 \cdot 3^n} \geq n. \end{aligned}$$

Contradiction, because we should have $\|T_n x\| \leq c(x)$ for all n.