

BO1 History of Mathematics
Lecture VIII
Establishing rigorous thinking in analysis

MT 2021 Week 4

Summary

Part 1

- ▶ French institutions
- ▶ Fourier series
- ▶ Early-19th-century rigour

Part 2

- ▶ Limits, continuity, differentiability
- ▶ Mathematics of small quantities
- ▶ The baton passes from France to Germany

France at the turn of the 19th century



The 1798 Egyptian Expedition Under the Command of Bonaparte, Cogniet

France at the turn of the 19th century

French revolution (1789) led to

- ▶ the establishment and prestige of the Grandes écoles

France at the turn of the 19th century

French revolution (1789) led to

- ▶ the establishment and prestige of the Grandes écoles
- ▶ a new concentration of mathematical talent in Paris:
Lagrange, Laplace, Ampère, Fourier, Legendre, Poisson,
Cauchy, ...

France at the turn of the 19th century

French revolution (1789) led to

- ▶ the establishment and prestige of the Grandes écoles
- ▶ a new concentration of mathematical talent in Paris:
Lagrange, Laplace, Ampère, Fourier, Legendre, Poisson,
Cauchy, ...
- ▶ a wealth of personal interactions

France at the turn of the 19th century

French revolution (1789) led to

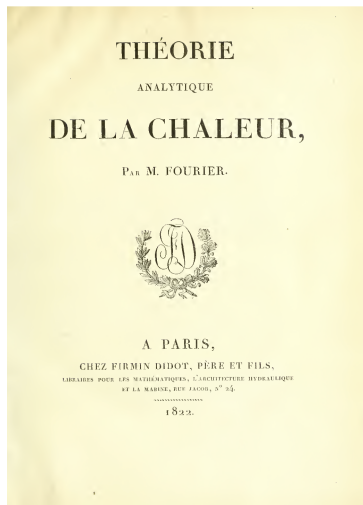
- ▶ the establishment and prestige of the Grandes écoles
- ▶ a new concentration of mathematical talent in Paris:
Lagrange, Laplace, Ampère, Fourier, Legendre, Poisson,
Cauchy, ...
- ▶ a wealth of personal interactions
- ▶ a major new journal: *Journal de l'École polytechnique*

France at the turn of the 19th century

French revolution (1789) led to

- ▶ the establishment and prestige of the Grandes écoles
- ▶ a new concentration of mathematical talent in Paris:
Lagrange, Laplace, Ampère, Fourier, Legendre, Poisson,
Cauchy, ...
- ▶ a wealth of personal interactions
- ▶ a major new journal: *Journal de l'École polytechnique*
- ▶ and a new focus on rigour

Fourier series



Joseph Fourier, *Analytic theory of heat*, 1822



See: [Bernard Maurey, 'Fourier, one man, several lives', *European Mathematical Society Newsletter*, no. 113 \(Sept 2019\), 8–20](#)

Fourier series

Suppose that $\phi(x) = a \sin x + b \sin 2x + c \sin 3x + \dots$

Fourier series

Suppose that $\phi(x) = a \sin x + b \sin 2x + c \sin 3x + \dots$

and also that $\phi(x) = x\phi'(0) + \frac{1}{2}x^2\phi''(0) + \frac{1}{6}x^3\phi'''(0) + \dots$

Fourier series

Suppose that $\phi(x) = a \sin x + b \sin 2x + c \sin 3x + \dots$

and also that $\phi(x) = x\phi'(0) + \frac{1}{2}x^2\phi''(0) + \frac{1}{6}x^3\phi'''(0) + \dots$

After many pages of calculations, multiplying and comparing power series, Fourier found that the coefficient of $\sin nx$ must be

$$\frac{2}{\pi} \int_0^{\pi} \phi(x) \sin nx \, dx$$

Fourier series

Suppose that $\phi(x) = a \sin x + b \sin 2x + c \sin 3x + \dots$

and also that $\phi(x) = x\phi'(0) + \frac{1}{2}x^2\phi''(0) + \frac{1}{6}x^3\phi'''(0) + \dots$

After many pages of calculations, multiplying and comparing power series, Fourier found that the coefficient of $\sin nx$ must be

$$\frac{2}{\pi} \int_0^{\pi} \phi(x) \sin nx \, dx$$

Fourier's derivation was based on 'naive' manipulations of infinite series. It was ingenious but non-rigorous and shaky (see: *Mathematics emerging*, §8.4.1).

Fourier series

Suppose that $\phi(x) = a \sin x + b \sin 2x + c \sin 3x + \dots$

and also that $\phi(x) = x\phi'(0) + \frac{1}{2}x^2\phi''(0) + \frac{1}{6}x^3\phi'''(0) + \dots$

After many pages of calculations, multiplying and comparing power series, Fourier found that the coefficient of $\sin nx$ must be

$$\frac{2}{\pi} \int_0^{\pi} \phi(x) \sin nx \, dx$$

Fourier's derivation was based on 'naive' manipulations of infinite series. It was ingenious but non-rigorous and shaky (see: *Mathematics emerging*, §8.4.1).

BUT it led to profound results

Establishing rigour

The development of 'rigour':

Establishing rigour

The development of 'rigour':

- ▶ Cauchy sequences

Establishing rigour

The development of 'rigour':

- ▶ Cauchy sequences
- ▶ continuity

Establishing rigour

The development of 'rigour':

- ▶ Cauchy sequences
- ▶ continuity
- ▶ limits

Establishing rigour

The development of 'rigour':

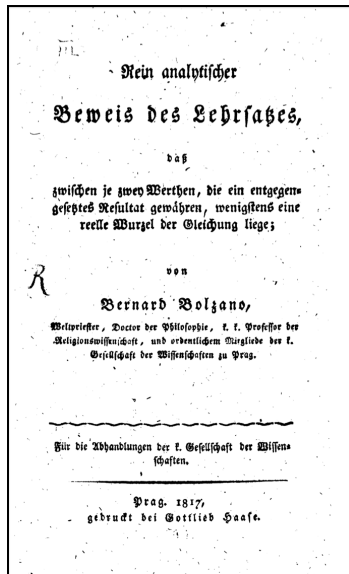
- ▶ Cauchy sequences
- ▶ continuity
- ▶ limits
- ▶ differentiability

Establishing rigour

The development of 'rigour':

- ▶ Cauchy sequences
- ▶ continuity
- ▶ limits
- ▶ differentiability
- ▶ ϵ, δ notation

Cauchy sequences: Bolzano (1817)



Bernard Bolzano, *Purely analytic proof of the theorem that between any two values which give opposite values lies at least one real root of the equation*, 1817



Cauchy sequences: Bolzano (1817)

Began with a discussion of previous proofs of the Intermediate Value Theorem (by “mathematicians of great repute”)

Cauchy sequences: Bolzano (1817)

Began with a discussion of previous proofs of the Intermediate Value Theorem (by “mathematicians of great repute”)

The most common kind of proof depends on a truth borrowed from geometry, namely, that every continuous line of simple curvature of which the ordinates are first positive and then negative (or conversely) must necessarily intersect the x -axis somewhere at a point that lies in between those ordinates.

Cauchy sequences: Bolzano (1817)

Began with a discussion of previous proofs of the Intermediate Value Theorem (by “mathematicians of great repute”)

The most common kind of proof depends on a truth borrowed from geometry, namely, that every continuous line of simple curvature of which the ordinates are first positive and then negative (or conversely) must necessarily intersect the x -axis somewhere at a point that lies in between those ordinates. There is certainly no question concerning the correctness, nor indeed the obviousness, of this geometrical proposition.

Cauchy sequences: Bolzano (1817)

Began with a discussion of previous proofs of the Intermediate Value Theorem (by “mathematicians of great repute”)

The most common kind of proof depends on a truth borrowed from geometry, namely, that every continuous line of simple curvature of which the ordinates are first positive and then negative (or conversely) must necessarily intersect the x -axis somewhere at a point that lies in between those ordinates. There is certainly no question concerning the correctness, nor indeed the obviousness, of this geometrical proposition. But it is clear that it is an intolerable offense against correct method to derive truths of pure (or general) mathematics (i.e., arithmetic, algebra, analysis) from considerations which belong to a merely applied (or special) part, namely, geometry.

Cauchy sequences: Bolzano (1817)

If a series of quantities has the property that the difference between its n -th term and every later one remains smaller than any given quantity ... then there is always a certain constant quantity ... which the terms of this series approach.

Cauchy sequences: Bolzano (1817)

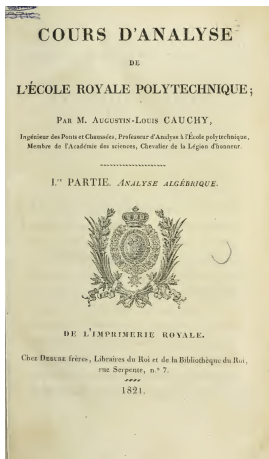
If a series of quantities has the property that the difference between its n -th term and every later one remains smaller than any given quantity ... then there is always a certain constant quantity ... which the terms of this series approach.

Proof: The hypothesis that there exists a quantity X which the terms of this series approach ... contains nothing impossible ...

(See: *Mathematics emerging*, §16.1.1; for a full translation, see: S. B. Russ, A translation of Bolzano's paper on the intermediate value theorem, *Historia Mathematica* 7(2) (1980), 156–185)

Cauchy's *Cours d'analyse*

Augustin-Louis Cauchy, *Cours d'analyse de l'École royale polytechnique* (1821)



(Annotated translation by Robert E. Bradley and C. Edward Sandifer, Springer, 2009)

Cauchy sequences: Cauchy (1821)

Augustin-Louis Cauchy, *Cours d'analyse* (1821), Ch. VI, pp. 124, 125:

In order for the series u_0, u_1, u_2, \dots [that is, $\sum u_i$] to be convergent ... it is necessary and sufficient that the partial sums

$$s_n = u_0 + u_1 + u_2 + \&c. \dots + u_{n-1}$$

converge to a fixed limit s : in other words, it is necessary and sufficient that for infinitely large values of the number n , the sums

$$s_n, s_{n+1}, s_{n+2}, \&c. \dots$$

differ from the limit s , and consequently from each other, by infinitely small quantities.

(See: *Mathematics emerging*, §16.1.2.)

More from Cauchy (1821)

Further results from Cauchy's *Cours d'analyse*:

More from Cauchy (1821)

Further results from Cauchy's *Cours d'analyse*:

- ▶ ratio test;

More from Cauchy (1821)

Further results from Cauchy's *Cours d'analyse*:

- ▶ ratio test;
- ▶ root test;

More from Cauchy (1821)

Further results from Cauchy's *Cours d'analyse*:

- ▶ ratio test;
- ▶ root test;
- ▶ alternating series test (proof uses Cauchy sequences);

More from Cauchy (1821)

Further results from Cauchy's *Cours d'analyse*:

- ▶ ratio test;
- ▶ root test;
- ▶ alternating series test (proof uses Cauchy sequences);
- ▶ and many more.

(See *Mathematics emerging*, §16.1.2)

Cauchy sequences concluded

Early uses of Cauchy sequences:

Cauchy sequences concluded

Early uses of Cauchy sequences:

- ▶ in da Cunha's *Principios mathematicos* (1782)

Cauchy sequences concluded

Early uses of Cauchy sequences:

- ▶ in da Cunha's *Principios mathematicos* (1782)
- ▶ in Bolzano's proof of the existence of a least upper bound (1817)

Cauchy sequences concluded

Early uses of Cauchy sequences:

- ▶ in da Cunha's *Principios mathematicos* (1782)
- ▶ in Bolzano's proof of the existence of a least upper bound (1817)
- ▶ in Cauchy's further results on sequences and series (1821)

Cauchy sequences concluded

Early uses of Cauchy sequences:

- ▶ in da Cunha's *Principios mathematicos* (1782)
- ▶ in Bolzano's proof of the existence of a least upper bound (1817)
- ▶ in Cauchy's further results on sequences and series (1821)
- ▶ in Abel's proof of the general binomial theorem (1826)

Cauchy sequences concluded

Early uses of Cauchy sequences:

- ▶ in da Cunha's *Principios mathematicos* (1782)
- ▶ in Bolzano's proof of the existence of a least upper bound (1817)
- ▶ in Cauchy's further results on sequences and series (1821)
- ▶ in Abel's proof of the general binomial theorem (1826)

BUT the convergence of Cauchy sequences themselves remained unproved

Part 2: Further Rigour

Continuity

Early definitions of continuity:

Continuity

Early definitions of continuity:

Wallis (1656): a curve that doesn't 'jump about'

Continuity

Early definitions of continuity:

Wallis (1656): a curve that doesn't 'jump about'

Euler (1748): a curve described by a single expression

Continuity

Early definitions of continuity:

Wallis (1656): a curve that doesn't 'jump about'

Euler (1748): a curve described by a single expression

Later definitions of continuity:

Bolzano (1817): *$f(x + \omega) - f(x)$ can be made smaller than any given quantity, provided ω can be taken as small as we please*

Continuity

Early definitions of continuity:

Wallis (1656): a curve that doesn't 'jump about'

Euler (1748): a curve described by a single expression

Later definitions of continuity:

Bolzano (1817): *$f(x + \omega) - f(x)$ can be made smaller than any given quantity, provided ω can be taken as small as we please*

Cauchy (1821): $f(x + a) - f(x)$ decreases with a

Continuity

Early definitions of continuity:

Wallis (1656): a curve that doesn't 'jump about'

Euler (1748): a curve described by a single expression

Later definitions of continuity:

Bolzano (1817): *$f(x + \omega) - f(x)$ can be made smaller than any given quantity, provided ω can be taken as small as we please*

Cauchy (1821): $f(x + a) - f(x)$ decreases with a

[Question: dependence? plagiarism? or a common source?]

Limits: early definitions

Wallis (1656): a quantity 'less than any assignable' quantity is zero

Limits: early definitions

Wallis (1656): a quantity 'less than any assignable' quantity is zero

Newton (1687): adopted and 'proved' Wallis's definition; also used 'limit' in the sense of a 'bound' or 'ultimate value'; developed theory of 'first and last ratios'

Limits: early definitions

Wallis (1656): a quantity 'less than any assignable' quantity is zero

Newton (1687): adopted and 'proved' Wallis's definition; also used 'limit' in the sense of a 'bound' or 'ultimate value'; developed theory of 'first and last ratios'

D'Alembert (1751): 'one may approach a limit as closely as one wishes ... but never surpass it'; example: polygons and circle; he assumed that $\lim AB = \lim A \times \lim B$; a dictionary definition only — no theory
[Encyclopedie online \(in French\)](#)

Limits: a later definition

Cauchy, *Cours d'analyse* (1821), p. 4:

When the values successively given to a variable approach indefinitely to a fixed value, so as to finish by differing from it by as little as one would wish, the latter is called the limit of all the others.

Limits: a later definition

Cauchy, *Cours d'analyse* (1821), p. 4:

When the values successively given to a variable approach indefinitely to a fixed value, so as to finish by differing from it by as little as one would wish, the latter is called the limit of all the others.

Examples:

- ▶ an irrational number is a limit of rationals;
- ▶ in geometry a circle is a limit of polygons.

Limits: a later definition

Cauchy, *Cours d'analyse* (1821), p. 4:

When the values successively given to a variable approach indefinitely to a fixed value, so as to finish by differing from it by as little as one would wish, the latter is called the limit of all the others.

Examples:

- ▶ an irrational number is a limit of rationals;
- ▶ in geometry a circle is a limit of polygons.

BUT still no formal definition of

- ▶ 'as small as one wishes',
- ▶ 'as closely as one wishes', ...

Differentiability: early ideas

For Leibniz and his immediate followers, any 'function' you could write down was automatically differentiable (by the usual rules).

Differentiability: early ideas

For Leibniz and his immediate followers, any 'function' you could write down was automatically differentiable (by the usual rules).

For Lagrange, the 'Taylor' series

$$f(x + h) = f(x) + f'(x)h + \dots$$

led naturally to consideration of

$$\frac{f(x + h) - f(x)}{h}$$

as an approximation to $f'(x)$, for small h

Differentiability: early ideas

For Leibniz and his immediate followers, any 'function' you could write down was automatically differentiable (by the usual rules).

For Lagrange, the 'Taylor' series

$$f(x + h) = f(x) + f'(x)h + \dots$$

led naturally to consideration of

$$\frac{f(x + h) - f(x)}{h}$$

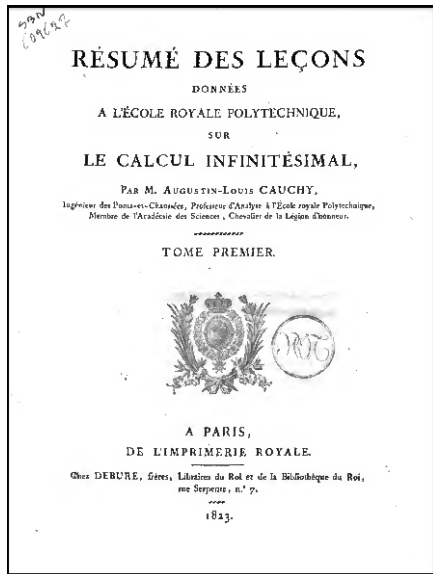
as an approximation to $f'(x)$, for small h

Ampère (1806) struggled with the meaning of

$$\frac{f(x + h) - f(x)}{h}$$

— why isn't it just zero or infinite?

Differentiability: Cauchy's *Résumé*



Cauchy, *Résumé des leçons
données à l'École royale
polytechnique sur le calcul
infinitésimal*, 1823

(Translation by Dennis
M. Cates, Fairview Academic
Press, 2012)

Differentiability: Cauchy's *Résumé*

... those who read my book will I hope be convinced that the principles of the differential calculus and its most important applications can easily be set out without the use of series.

Differentiability: Cauchy's *Résumé*

... those who read my book will I hope be convinced that the principles of the differential calculus and its most important applications can easily be set out without the use of series.

Defined the derivative as the limit of

$$\frac{f(x + h) - f(x)}{h}$$

with many particular examples: ax , a/x , $\sin x$, $\log x$, ...

Differentiability: Cauchy's *Résumé*

... those who read my book will I hope be convinced that the principles of the differential calculus and its most important applications can easily be set out without the use of series.

Defined the derivative as the limit of

$$\frac{f(x+h) - f(x)}{h}$$

with many particular examples: ax , a/x , $\sin x$, $\log x$, ...

but no concerns about existence in general

(See: *Mathematics emerging*, §14.1.4.)

Arbitrarily small intervals

A theorem of Lagrange (1797):

If the first derived function of a function f is strictly positive on an interval $[a, b]$, then $f(b) > f(a)$.

Arbitrarily small intervals

A theorem of Lagrange (1797):

If the first derived function of a function f is strictly positive on an interval $[a, b]$, then $f(b) > f(a)$.

Proof: Divide the interval $[a, b]$ into n subintervals, taking n as large as necessary ...

Arbitrarily small intervals

A theorem of Lagrange (1797):

If the first derived function of a function f is strictly positive on an interval $[a, b]$, then $f(b) > f(a)$.

Proof: Divide the interval $[a, b]$ into n subintervals, taking n as large as necessary ...

Unconvincing to modern eyes, but a useful technique.

(See: *Mathematics emerging*, §11.2.3.)

IVT revisited

Cauchy, *Cours d'analyse* (1821), Note III, p. 460 (On the numerical solution of equations):

Theorem: Let f be a real function of the variable x , which remains continuous with respect to this variable between the limits $x = x_0$, $x = X$. If the two quantities $f(x_0)$, $f(X)$ are of opposite signs, the equation $f(x) = 0$ will be satisfied by one or more real values of x contained between x_0 and X .

(See: *Mathematics emerging*, §11.2.6.)

IVT revisited

Cauchy's proof:

Choose $m > 1$. Divide the interval $[x_0, X]$ into m equal parts;

IVT revisited

Cauchy's proof:

Choose $m > 1$. Divide the interval $[x_0, X]$ into m equal parts; find neighbouring division points x_1, X' such that $f(x_1), f(X')$ are of opposite signs.

IVT revisited

Cauchy's proof:

Choose $m > 1$. Divide the interval $[x_0, X]$ into m equal parts; find neighbouring division points x_1, X' such that $f(x_1), f(X')$ are of opposite signs. Subdivide the interval $[x_1, X']$ into m equal parts;

IVT revisited

Cauchy's proof:

Choose $m > 1$. Divide the interval $[x_0, X]$ into m equal parts; find neighbouring division points x_1, X' such that $f(x_1), f(X')$ are of opposite signs. Subdivide the interval $[x_1, X']$ into m equal parts; find neighbouring division points x_2, X'' such that $f(x_2), f(X'')$ are of opposite signs.

IVT revisited

Cauchy's proof:

Choose $m > 1$. Divide the interval $[x_0, X]$ into m equal parts; find neighbouring division points x_1, X' such that $f(x_1), f(X')$ are of opposite signs. Subdivide the interval $[x_1, X']$ into m equal parts; find neighbouring division points x_2, X'' such that $f(x_2), f(X'')$ are of opposite signs. Continue in this way to obtain an increasing sequence x_0, x_1, \dots and a decreasing sequence X, X', \dots

IVT revisited

Cauchy's proof:

Choose $m > 1$. Divide the interval $[x_0, X]$ into m equal parts; find neighbouring division points x_1, X' such that $f(x_1), f(X')$ are of opposite signs. Subdivide the interval $[x_1, X']$ into m equal parts; find neighbouring division points x_2, X'' such that $f(x_2), f(X'')$ are of opposite signs. Continue in this way to obtain an increasing sequence x_0, x_1, \dots and a decreasing sequence X, X', \dots . The difference $X^{(n)} - x_n$ is $(X - x_0)/m^n$, which may be made as small as one wishes.

IVT revisited

Cauchy's proof:

Choose $m > 1$. Divide the interval $[x_0, X]$ into m equal parts; find neighbouring division points x_1, X' such that $f(x_1), f(X')$ are of opposite signs. Subdivide the interval $[x_1, X']$ into m equal parts; find neighbouring division points x_2, X'' such that $f(x_2), f(X'')$ are of opposite signs. Continue in this way to obtain an increasing sequence x_0, x_1, \dots and a decreasing sequence X, X', \dots . The difference $X^{(n)} - x_n$ is $(X - x_0)/m^n$, which may be made as small as one wishes. The sequences x_0, x_1, \dots and X, X', \dots therefore converge to a common limit a , at which $f(a) = 0$.

IVT revisited

Cauchy's proof:

Choose $m > 1$. Divide the interval $[x_0, X]$ into m equal parts; find neighbouring division points x_1, X' such that $f(x_1), f(X')$ are of opposite signs. Subdivide the interval $[x_1, X']$ into m equal parts; find neighbouring division points x_2, X'' such that $f(x_2), f(X'')$ are of opposite signs. Continue in this way to obtain an increasing sequence x_0, x_1, \dots and a decreasing sequence X, X', \dots . The difference $X^{(n)} - x_n$ is $(X - x_0)/m^n$, which may be made as small as one wishes. The sequences x_0, x_1, \dots and X, X', \dots therefore converge to a common limit a , at which $f(a) = 0$.

Note: Cauchy offered this as a fast method of approximation to roots of equations.

IVT revisited

Cauchy's proof:

Choose $m > 1$. Divide the interval $[x_0, X]$ into m equal parts; find neighbouring division points x_1, X' such that $f(x_1), f(X')$ are of opposite signs. Subdivide the interval $[x_1, X']$ into m equal parts; find neighbouring division points x_2, X'' such that $f(x_2), f(X'')$ are of opposite signs. Continue in this way to obtain an increasing sequence x_0, x_1, \dots and a decreasing sequence X, X', \dots . The difference $X^{(n)} - x_n$ is $(X - x_0)/m^n$, which may be made as small as one wishes. The sequences x_0, x_1, \dots and X, X', \dots therefore converge to a common limit a , at which $f(a) = 0$.

Note: Cauchy offered this as a fast method of approximation to roots of equations.

But it also provides a much more convincing proof of the Intermediate Value Theorem than that appearing earlier in Cauchy's text (*Cours d'analyse*, Ch. II, Theorem 4: p. 44).

ε and δ appear

A theorem of Cauchy, *Résumé* (1823):

Suppose that in the interval $[x_0, X]$ we have $A < f'(x) < B$. Then we also have

$$A < \frac{f(X) - f(x_0)}{X - x_0} < B$$

ϵ and δ appear

A theorem of Cauchy, *Résumé* (1823):

Suppose that in the interval $[x_0, X]$ we have $A < f'(x) < B$. Then we also have

$$A < \frac{f(X) - f(x_0)}{X - x_0} < B$$

Proof: Choose two quantities ϵ, δ, \dots such that for $i < \delta$

$$f'(x) - \epsilon < \frac{f(x+i) - f(x)}{i} < f'(x) + \epsilon$$

etc.

(See: *Mathematics emerging*, §14.1.5.)

Hints of a broader class of functions

If a Taylor series exists for a given function, and all the coefficients vanish, then surely the function itself must vanish . . .

Hints of a broader class of functions

If a Taylor series exists for a given function, and all the coefficients vanish, then surely the function itself must vanish . . .

However, Cauchy gave the example $f(x) = e^{-x^2} + e^{-x^{-2}}$, which is clearly never zero, but all of its derivatives vanish

Hints of a broader class of functions

If a Taylor series exists for a given function, and all the coefficients vanish, then surely the function itself must vanish . . .

However, Cauchy gave the example $f(x) = e^{-x^2} + e^{-x^{-2}}$, which is clearly never zero, but all of its derivatives vanish

So not every function can be expanded into a Taylor series,

Hints of a broader class of functions

If a Taylor series exists for a given function, and all the coefficients vanish, then surely the function itself must vanish ...

However, Cauchy gave the example $f(x) = e^{-x^2} + e^{-x^{-2}}$, which is clearly never zero, but all of its derivatives vanish

So not every function can be expanded into a Taylor series, and it appears to be possible to conceive of functions to which the calculus is not immediately or naturally applicable ...

Modern rigour in analysis



Karl Weierstrass (1815–1897):

Modern rigour in analysis



Karl Weierstrass (1815–1897):

- ▶ taught at University of Berlin from 1856 onwards

Modern rigour in analysis



Karl Weierstrass (1815–1897):

- ▶ taught at University of Berlin from 1856 onwards
- ▶ completed the rigourisation of calculus via systematic use of ε/δ methods

Modern rigour in analysis



Karl Weierstrass (1815–1897):

- ▶ taught at University of Berlin from 1856 onwards
- ▶ completed the rigourisation of calculus via systematic use of ε/δ methods

BUT we have no direct sources, only lecture notes or books by his pupils and followers

From France to Germany

By the later 19th century the mathematical centre of gravity in Europe had moved from the Parisian Écoles to the German universities:

Göttingen (est. 1734): Gauss, Dirichlet, [Dedekind], Riemann, Klein, Hilbert, ...

Berlin (est. 1810): Crelle (editor), Dirichlet, Eisenstein, Kummer, [Jacobi], Kronecker, Weierstrass, ...

with a focus on both research and teaching.