BO1 History of Mathematics
Lecture VIII
Establishing rigorous thinking in analysis
Part 2: Further rigour

MT 2020 Week 4

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

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 <u>limit</u> 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* <u>*limit*</u> *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

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

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

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

led naturally to consideration of

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

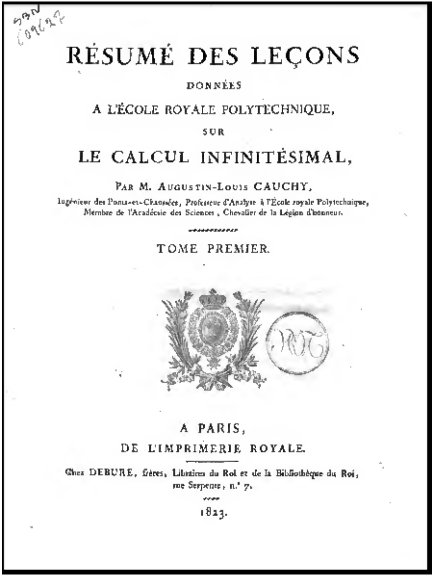
as an approximation to $f'(x)$, for small h

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*



RÉSUMÉ DES LEÇONS
DONNÉES
A L'ÉCOLE ROYALE POLYTECHNIQUE,
SUR
LE CALCUL INFINITÉSIMAL,
PAR M. AUGUSTIN-LOUIS CAUCHY,
Ingénieur des Ponts-et-Chaussées, Professeur d'Analyse à l'École royale Polytechnique, Membre de l'Académie des Sciences, Chevalier de la Légion d'honneur.

TOME PREMIER.

A PARIS,
DE L'IMPRIMERIE ROYALE.
Chez DEBURE, frères, Libraires du Roi et de la Bibliothèque du Roi, rue Serpente, n.° 7.
1823.

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 $x_0, \ x_1, \ldots$ and a decreasing sequence $X, \ X', \ldots$.

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 $x_0,\ x_1, \ldots$ and a decreasing sequence $X,\ X', \ldots$. 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 $x_0,\ x_1, \ldots$ and a decreasing sequence $X,\ X', \ldots$. The difference $X^{(n)} - x_n$ is $(X - x_0)/m^n$, which may be made as small as one wishes. The sequences $x_0,\ x_1, \ldots$ and $X,\ X', \ldots$ 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 $x_0,\ x_1, \ldots$ and a decreasing sequence $X,\ X', \ldots$. The difference $X^{(n)} - x_n$ is $(X - x_0)/m^n$, which may be made as small as one wishes. The sequences $x_0,\ x_1, \ldots$ and $X,\ X', \ldots$ 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 $\epsilon,\ \delta, \ldots$ 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 rigorisation 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 rigorisation 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.