Approximation Theory and Proof Assistants: Certified Computations

Nicolas Brisebarre and Damien Pous

Master 2 Informatique Fondamentale
École Normale Supérieure de Lyon, 2023-2024

This course

is about

This course

is about rigourous/validated/reliable/certified numerical computations on a machine

This course

is about rigourous/validated/reliable/certified numerical computations on a machine in mathematical analysis.

(Binary) Floating Point (FP) Arithmetic

Given

$$\begin{cases} \text{a precision} & p \geqslant 1, \\ \text{a set of exponents} & E_{\mathsf{min}}, \cdots, E_{\mathsf{max}}. \end{cases}$$

A finite FP number x is represented by 2 integers:

- integer significand M, $2^{p-1} \leqslant |M| \leqslant 2^p - 1$,
- exponent E, $E_{\mathsf{min}} \leqslant E \leqslant E_{\mathsf{max}}$

such that

$$x = \frac{M}{2^{p-1}} \times 2^E.$$

IEEE Precisions

IEEE 754 standard (1984 then 2008).

See `http://en.wikipedia.org/wiki/IEEE_floating_point`.

	precision p	min. exponent E_{min}	maximal exponent E_{max}
single (binary32)	24	-126	127
double (binary64)	53	-1022	1023
extended double	64	-16382	16383
quadruple (binary128)	113	-16382	16383

We have $x = \frac{M}{2^{p-1}} \times 2^E$ with $2^{p-1} \leqslant |M| \leqslant 2^p - 1$

and $E_{\text{min}} \leqslant E \leqslant E_{\text{max}}$.

Rounding modes

In the IEEE 754 standard, the user defines an *active rounding mode* (or *rounding direction attribute*) among:

- round to nearest (default). If $x \in \mathbb{R}$, $\mathsf{RN}(x)$ is the floating-point number that is the closest to x. In case of a tie, value whose integral significand is even.
- round towards $+\infty$.
- round towards $-\infty$.
- round towards zero.

You Shouldn't Trust Your Computer

S. Rump's example (1988). Consider

$$f(a,b) = 333.75b^6 + a^2\left(11a^2b^2 - b^6 - 121b^4 - 2\right) + 5.5b^8 + \frac{a}{2b},$$

and try to compute $f(a,b)$ for $a = 77617.0$ and $b = 33096.0$. On an IBM 370 computer:

- 1.172603 in single precision;
- 1.1726039400531 in double precision;
- 1.172603940053178 in extended precision.

You Shouldn't Trust Your Computer

S. Rump's example (1988). Consider

$$f(a,b) = 333.75b^6 + a^2\left(11a^2b^2 - b^6 - 121b^4 - 2\right) + 5.5b^8 + \frac{a}{2b},$$

and try to compute $f(a,b)$ for $a = 77617.0$ and $b = 33096.0$. On an IBM 370 computer:

- 1.172603 in single precision;
- 1.1726039400531 in double precision;
- 1.172603940053178 in extended precision.

And yet, the exact result is $-0.8273960599\cdots$.
What about more recent systems? On a Pentium4 (gcc, Linux), Rump's C program returns

- 5.960604×10^{20} in single precision;
- 2.0317×10^{29} in double precision;
- $-9.38724 \times 10^{-323}$ in extended precision.

(Certified ?) Quadrature

W. Tucker. Validated Numerics. Princeton University Press, 2011.

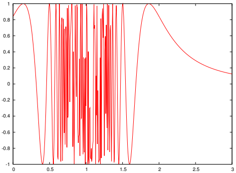
Let $I = \int_0^8 \sin(x + e^x)\mathrm{d}x$. Let's evaluate it using MATLAB.

```
fcn_str = 'sin(x+exp(x))';
f = vectorize(inline(fcn_str));
a = 0; b = 8;
>> q = quad(f,a,b)
q =
    0.251102722027180
```

Actually, $I \in [0.3474, 0.3475]$...

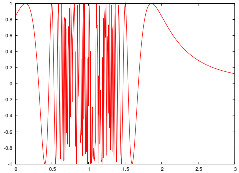
(Certified ?) Quadrature

Let $J = \int_0^3 \sin\left(\frac{1}{(10^{-3}+(1-x)^2)^{3/2}}\right) \mathrm{d}x$.



(Certified ?) Quadrature

Let $J = \int_0^3 \sin\left(\frac{1}{(10^{-3}+(1-x)^2)^{3/2}}\right)\mathrm{d}x$.



- Maple2023 : 10 digits $\to$ $0,7499743685$, 11 digits $\to$ no answer ;
- Pari/GP: 0.7927730971479080755;
- Mathematica and Chebfun fail to answer;
- Sage : $0,7499743685$;
- Chen, '06: 0.7578918118.

  WHAT IS THE CORRECT ANSWER?

How to overcome these problems?

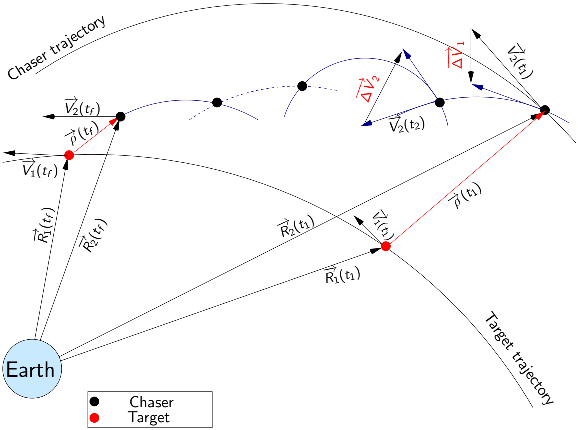
- Use Computer Algebra (Maple, Mathematica) to perform exact computations!

How to overcome these problems?

- Use Computer Algebra (Maple, Mathematica) to perform exact computations! Problem with the speed of computations.
- Interval Arithmetic: replace any number with an interval containing it. Has to be used with extreme caution.

An Example[1]: Tschauner-Hempel Equation

Relative Motion in Keplerian Dynamics



[1] Courtesy of Mioara Joldeș

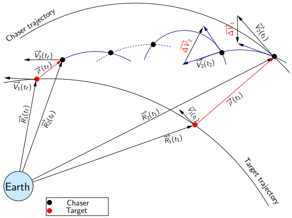
An Example[2]: Tschauner-Hempel Equation

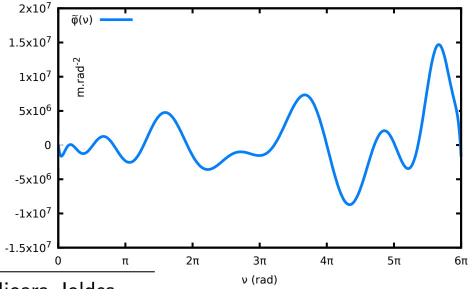
Relative Motion in Keplerian Dynamics

Reduced Equation

$$z''(\nu) + \left(4 - \frac{3}{1 + e\cos\nu}\right) z(\nu) = c$$

c initial conditions, e orbital eccentricity





[2] Courtesy of Mioara Joldeș

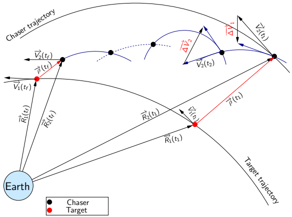
An Example[2]: Tschauner-Hempel Equation

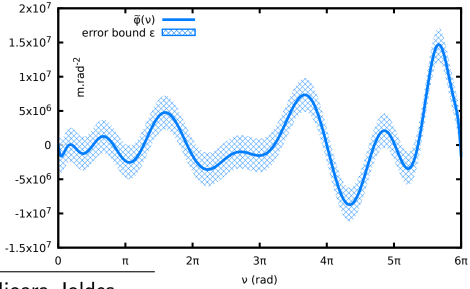
Relative Motion in Keplerian Dynamics

Reduced Equation

$$z''(\nu) + \left(4 - \frac{3}{1 + e\cos\nu}\right) z(\nu) = c$$

c initial conditions, e orbital eccentricity





[2] Courtesy of Mioara Joldeș

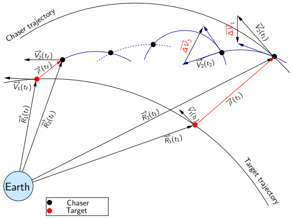
An Example[2]: Tschauner-Hempel Equation

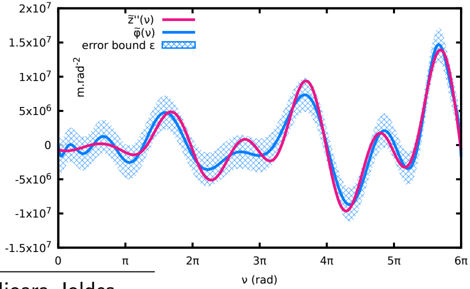
Relative Motion in Keplerian Dynamics

Reduced Equation

$$z''(\nu) + \left(4 - \frac{3}{1 + e\cos\nu}\right) z(\nu) = c$$

c initial conditions, e orbital eccentricity





[2] Courtesy of Mioara Joldeș

Near-Earth Objects

Asteroids, comets, etc. whose orbit can get close to the Earth: risk of collision.

Near-Earth Objects

Asteroids, comets, etc. whose orbit can get close to the Earth: risk of collision.

Monitoring programs: `https://cneos.jpl.nasa.gov/` (NASA) et `https://neo.ssa.esa.int/` (ESA).

Near-Earth Objects

Asteroids, comets, etc. whose orbit can get close to the Earth: risk of collision.

Monitoring programs: `https://cneos.jpl.nasa.gov/` (NASA) et `https://neo.ssa.esa.int/` (ESA).

Torino scale: 0 (no risk of a collision) - 10 (collision is certain, possibly causing a global climatic catastrophe).

Near-Earth Objects

Asteroids, comets, etc. whose orbit can get close to the Earth: risk of collision.

Monitoring programs: `https://cneos.jpl.nasa.gov/` (NASA) et `https://neo.ssa.esa.int/` (ESA).

Torino scale: 0 (no risk of a collision) - 10 (collision is certain, possibly causing a global climatic catastrophe).

Question: certified trajectory?

Near-Earth Objects

`http://en.wikipedia.org/wiki/99942_Apophis`

> *99942 Apophis is a near-Earth asteroid that caused a brief period of concern in December 2004 because initial observations indicated a probability of up to 2.7 % that it would strike the Earth in 2029.*

Estimates: diameter of 330 metres (1,080 ft) and mass of 40 megatonnes (within a factor of three).

P. Di Lizia's PhD[3] gives algorithms to compute a "certified trajectory".

[3] P. Di Lizia, Robust Space Trajectory and Space System Design using Differential Algebra, Politecnico di Milano, 2008.

Near-Earth Objects

http://en.wikipedia.org/wiki/99942_Apophis

> *99942 Apophis is a near-Earth asteroid that caused a brief period of concern in December 2004 because initial observations indicated a probability of up to 2.7 % that it would strike the Earth in 2029.*

Estimates: diameter of 330 metres (1,080 ft) and mass of 40 megatonnes (within a factor of three).

P. Di Lizia's PhD[3] gives algorithms to compute a "certified trajectory".

Question: how to automate this?

---

[3] P. Di Lizia, Robust Space Trajectory and Space System Design using Differential Algebra, Politecnico di Milano, 2008.

Hilbert's 16th problem

Let $n \in \mathbb{N}$, $P, Q \in \mathbb{R}[X,Y]$, $\deg P, \deg Q \leqslant n$.

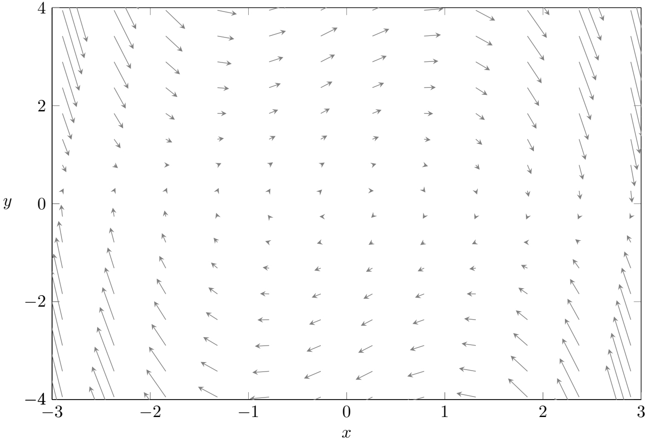
We consider

$$\begin{cases} \dot{x} = P(x,y), \\ \dot{y} = Q(x,y). \end{cases}$$

Limit cycle: periodic orbit whose neighbouring trajectories spiral either towards or away from.

Hilbert's 16th problem (second part): For a given n, maximum number of limit cycles?

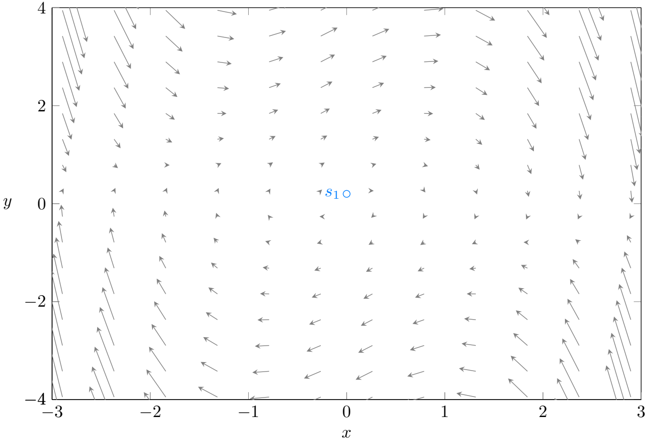
An example of limit cycle[4]



Van der Pol oscillator: $\begin{cases} \dot{x} = y \\ \dot{y} = -x + (1 - x^2)y \end{cases}$

[4] Courtesy of Florent Bréhard

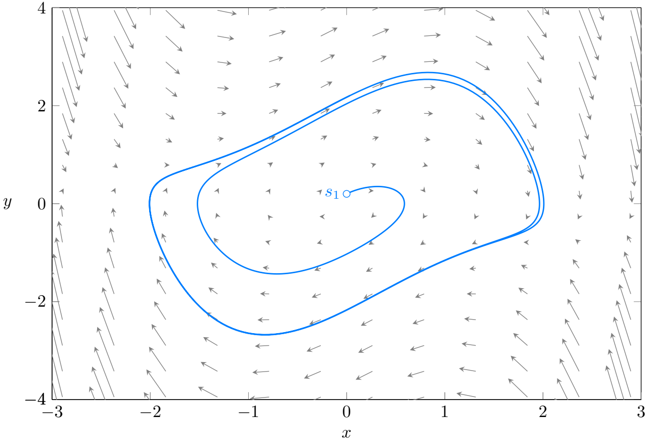
An example of limit cycle[4]



Van der Pol oscillator: $\begin{cases} \dot{x} = y \\ \dot{y} = -x + (1 - x^2)y \end{cases}$

[4]Courtesy of Florent Bréhard

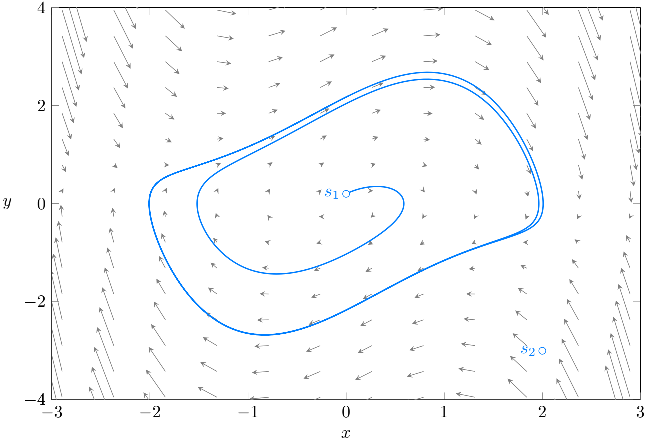
An example of limit cycle[4]



Van der Pol oscillator:
$$\begin{cases} \dot{x} = y \\ \dot{y} = -x + (1 - x^2)y \end{cases}$$

[4] Courtesy of Florent Bréhard

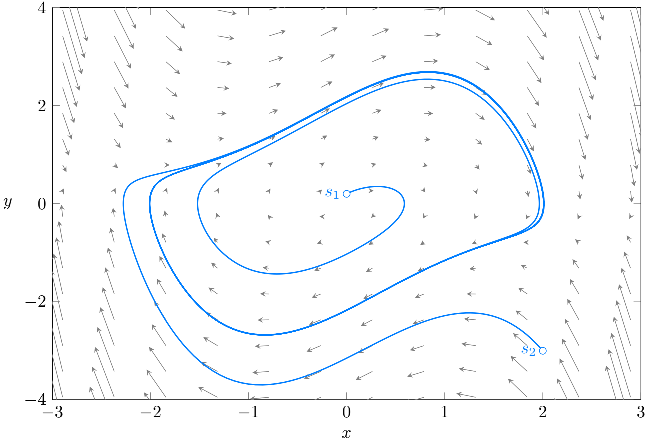
An example of limit cycle[4]



Van der Pol oscillator: $\begin{cases} \dot{x} = y \\ \dot{y} = -x + (1 - x^2)y \end{cases}$

[4] Courtesy of Florent Bréhard

An example of limit cycle[4]

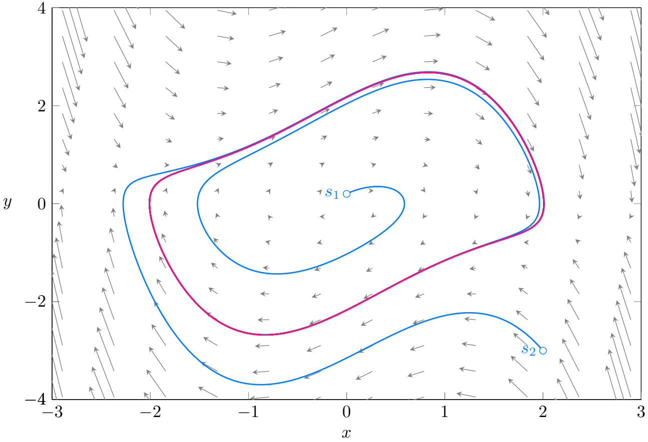


Van der Pol oscillator:

$$\begin{cases} \dot{x} = y \\ \dot{y} = -x + (1 - x^2)y \end{cases}$$

[4] Courtesy of Florent Bréhard

An example of limit cycle[4]



Van der Pol oscillator: $\begin{cases} \dot{x} = y \\ \dot{y} = -x + (1 - x^2)y \end{cases}$

[4] Courtesy of Florent Bréhard

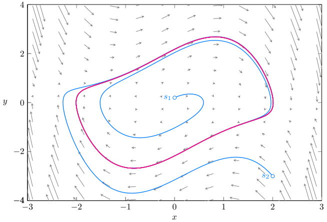
Hilbert's 16th Problem

Hilbert's 16th problem (second part)

For a given integer n, what is the maximum number $\mathcal{H}(n)$ of limit cycles a polynomial vector field of degree at most n in the plane can have?

D. Hilbert, International Congress of Mathematicians, Paris, 1900

- 1923: P. Dulac (incorrectly) proved that a *single* polynomial vector field has a finite number of limit cycles



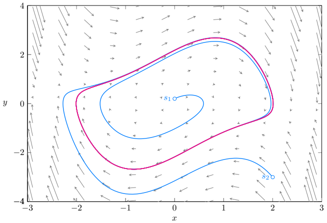
Hilbert's 16th Problem

Hilbert's 16th problem (second part)

For a given integer n, what is the maximum number $\mathcal{H}(n)$ of limit cycles a polynomial vector field of degree at most n in the plane can have?

D. Hilbert, International Congress of Mathematicians, Paris, 1900

- 1923: P. Dulac (incorrectly) proved that a *single* polynomial vector field has a finite number of limit cycles
- 1981: Y. S. Il'Yashenko found a major gap in Dulac's proof
- 1991: New proofs of Dulac's result by Y. S. Il'Yashenko and J. Écalle



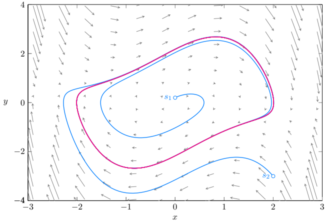
Hilbert's 16th Problem

Hilbert's 16th problem (second part)

For a given integer n, what is the maximum number $\mathcal{H}(n)$ of limit cycles a polynomial vector field of degree at most n in the plane can have?

D. Hilbert, International Congress of Mathematicians, Paris, 1900

- 1923: P. Dulac (incorrectly) proved that a *single* polynomial vector field has a finite number of limit cycles
- 1981: Y. S. Il'Yashenko found a major gap in Dulac's proof
- 1991: New proofs of Dulac's result by Y. S. Il'Yashenko and J. Écalle
- But even $\mathcal{H}(2) < \infty$ is open!



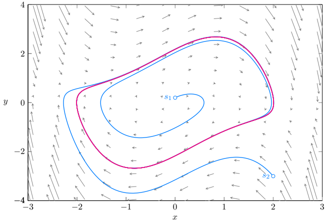
Hilbert's 16th Problem

Hilbert's 16th problem (second part)

For a given integer n, what is the maximum number $\mathcal{H}(n)$ of limit cycles a polynomial vector field of degree at most n in the plane can have?

D. Hilbert, International Congress of Mathematicians, Paris, 1900

- 1923: P. Dulac (incorrectly) proved that a *single* polynomial vector field has a finite number of limit cycles
- 1981: Y. S. Il'Yashenko found a major gap in Dulac's proof
- 1991: New proofs of Dulac's result by Y. S. Il'Yashenko and J. Écalle
- But even $\mathcal{H}(2) < \infty$ is open!
- Some lower bounds: $\mathcal{H}(2) \geqslant 4$, $\mathcal{H}(3) \geqslant 13$, $\mathcal{H}(4) \geqslant 22$.



Hilbert's 16th problem

Hilbert's 16th problem (second part)

For a given integer n, what is the maximum number $\mathcal{H}(n)$ of limit cycles a polynomial vector field of degree at most n in the plane can have?

D. Hilbert, International Congress of Mathematicians, Paris, 1900

V. I. Arnold (1977) suggested to study a restriction of this question, the so-called weak (or infinitesimal) Hilbert's 16th problem.

Moreover, he established a link between the number of limit cycles and the number of zeros of a certain integral.

(Formal) Proof of $\mathcal{H}(4) \geqslant 24$.

Approximation Theory and Proof Assistants: Certified Computations

Nicolas Brisebarre and Damien Pous

Master 2 Informatique Fondamentale
École Normale Supérieure de Lyon, 2023-2024

Chapter 1. Polynomial approximations

In this chapter, we present various theoretical and algorithmic results regarding polynomial approximations of functions.

We will mainly deal with real-valued continuous functions over a compact interval $[a, b]$, $a, b \in \mathbb{R}$, $a \leqslant b$.

We will denote $\mathcal{C}([a, b])$ the real vector space of continuous functions over $[a, b]$.

Polynomial approximations

In the framework of function evaluation one usually works with the two following norms over this vector space, namely

- the least-square norm L^2: given a weight[1] function $w \in \mathcal{C}([a,b])$, if $\mathrm{d}x$ denotes the Lebesgue measure, we write

  $$g \in L^2([a,b], w, \mathrm{d}x) \text{ if } \int_a^b w(x) |g(x)|^2 \mathrm{d}x < \infty,$$

  and then we define

  $$\|g\|_{2,w} = \sqrt{\int_a^b w(x) |g(x)|^2 \mathrm{d}x};$$

[1] Here, we will assume that it means that $w \in \mathcal{C}((a,b))$ and $w > 0$ almost everywhere.

Polynomial approximations

In the framework of function evaluation one usually works with the two following norms over this vector space, namely

- the supremum norm (aka Chebyshev norm, infinity norm, L^∞ norm) : if g is bounded on $[a,b]$, we set

$$\|g\|_\infty = \sup_{x\in[a,b]} |g(x)|$$

  (if f continuous, we have $\|g\|_\infty = \max_{x\in[a,b]} |g(x)|$).

Best polynomial approximations

One of the main questions we are interested in here is the following. We shall consider both the case $\|\cdot\| = \|\cdot\|_2$, and the case $\|\cdot\| = \|\cdot\|_\infty$.

Question. *Given* $f \in \mathcal{C}([a,b])$ *and* $n \in \mathbb{N}$*, minimize* $\|f - p\|$ *where* p *describes the space* $\mathbb{R}_n[x]$ *of polynomials with real number coefficients and degree at most* n*.*

Best polynomial approximations

In the L^2 case, the answer is easy to give. The space $\mathcal{C}\left([a,b]\right) \subset L^2([a,b], w, \mathrm{d}x)$ which is a Hilbert space, i.e. a vector space equipped with an inner product

$$\langle f, g \rangle = \int_a^b f\left(x\right) g\left(x\right) w\left(x\right) \mathrm{d}x,$$

and $\left\|\cdot\right\|_2$ is the associated norm, for which L^2 is complete.

Best polynomial approximations

In the L^2 case, the answer is easy to give. The space $\mathcal{C}([a,b]) \subset L^2([a,b], w, \mathrm{d}x)$ which is a Hilbert space, i.e. a vector space equipped with an inner product

$$\langle f, g \rangle = \int_a^b f(x)\, g(x)\, w(x)\, \mathrm{d}x,$$

and $\|\cdot\|_2$ is the associated norm, for which L^2 is complete.

The best polynomial approximation of degree at most n is the projection $p = \mathrm{pr}^{\perp}(f)$ of f over $\mathbb{R}_n[x]$. More details on the L^2 case later on.

The situation in the L^∞ case is more intricate and we will focus on it in the sequel of this chapter.

Section 2.1. Density of the polynomials in $(\mathcal{C}([a,b]), \|.\|_\infty)$

For all $f \in \mathcal{C}([a,b])$ and $n \in \mathbb{N}$, let

$$E_n(f) = \inf_{p \in \mathbb{R}_n[x]} \|f - p\|_\infty .$$

We first show that $E_n(f) \to 0$ as $n \to \infty$ (Weierstraß theorem, 1885).

For all $f \in \mathcal{C}([a,b])$ and $n \in \mathbb{N}$, let

$$E_n(f) = \inf_{p \in \mathbb{R}_n[x]} \|f - p\|_\infty .$$

We first show that $E_n(f) \to 0$ as $n \to \infty$ (Weierstraß theorem, 1885).

Theorem 1

For all $f \in \mathcal{C}([a,b])$ and for all $\varepsilon > 0$, there exists $n \in \mathbb{N}$, $p \in \mathbb{R}_n[x]$ such that $\|p - f\|_\infty < \varepsilon$.

Proofs by Runge (1885), Picard (1891), Lerch (1892 and 1903), Volterra (1897) Lebesgue (1898), Mittag-Leffler (1900), Fejér (1900 and 1916), Landau (1908), la Vallée Poussin (1908), Jackson (1911), Sierpinski (1911), Bernstein (1912), Montel (1918).

Section 2.1. Density of the polynomials in $(\mathcal{C}([a,b]), \|.\|_\infty)$

Note that we only used the values of the $B_n(f,x)$ for $0 \leqslant n \leqslant 2$. In fact, we have the following result.

Theorem 2

(Bohman and Korovkin) Let L_n a sequence of monotone linear operators on $\mathcal{C}(|a,b])$,that is to say: for $\text{all}\, f,g \in \mathcal{C}([a,b])$

- $L_n(\mu f + \lambda g) = \lambda L_n(f) + \mu L_n(g)$ *for all* $\lambda, \mu \in \mathbb{R}$,
- *if* $f(x) \geqslant g(x) \,\text{for all}\, x \in [a,b]$ *then* $L_n f(x) \geqslant L_n g(x) \,\text{for all}\, x \in [a,b]$,

the following conditions are equivalent

1. $L_n f \to f$ *uniformly for all* $f \in \mathcal{C}([a,b])$;
2. $L_n f \to f$ *uniformly for the three functions* $x \mapsto 1, x, x^2$;
3. $L_n 1 \to 1$ *and* $(L_n\varphi_t)(t) \to 0$ *uniformly in* $t \in [a,b]$ *where* $\varphi_t : x \in [a,b] \mapsto (t-x)^2$.

See Cheney's book for a proof.

Section 2.1. Density of the polynomials in $(\mathcal{C}([a,b]),\ \|.\|_\infty)$

A refinement of Weierstraß's theorem that gives the speed of convergence is obtained in terms of the modulus of continuity.

Definition 3

The modulus of continuity of f is the function ω defined as

$$\text{for all}\, \delta > 0, \qquad \omega(\delta) = \sup_{\substack{|x-y| < \delta, \\ x, y \in [a,b]}} |f(x) - f(y)|.$$

Proposition

If f is a continuous function over $[0,1]$, ω its modulus of continuity, then

$$\|f - B_n(f,x)\|_\infty = \tfrac{9}{4}\omega\left(n^{-\frac{1}{2}}\right).$$

Corollary 4

When f is Lipschitz continuous, $E_n(f) = O(n^{-1/2})$.

Remark

For improvements and refinements, see Section 4.6 of Cheney's book or Chapter 16 of Powell's book for a presentation of Jackson theorems.

Section 2.2. Best L^∞ (or minimax) approximation - Existence

The infimum $E_n(f)$ is reached:

Proposition

Let $(E, \|\cdot\|)$ *be a normed* $\mathbb{R}$*-vector space, let* F *be a finite dimensional subspace of* $(E, \|\cdot\|)$*. For all* $f \in E$*, there exists* $p \in F$ *such that* $\|p - f\| = \min_{q \in F} \|q - f\|$*. Moreover, the set of best approximations to a given* $f \in E$ *is convex.*

Section 2.2. Best L^∞ (or minimax) approximation. Uniqueness

The best L^2 approximation is unique, which is not always the case in the L^∞ setting.

Section 2.2. Best L^∞ (or minimax) approximation. Uniqueness

The best L^2 approximation is unique, which is not always the case in the L^∞ setting.

A counter-example?

Section 2.2. Best L^∞ (or minimax) approximation. Uniqueness

The best L^2 approximation is unique, which is not always the case in the L^∞ setting.

A counter-example?

In the case of L^∞, we need to introduce an additional condition known as the Haar condition.

Section 2.2. Best L^∞ (or minimax) approximation

Definition 5

Consider $n+1$ functions $\varphi_0, \ldots, \varphi_n$ defined over $[a,b]$. We say that $\varphi_0, \ldots, \varphi_n$ satisfy the Haar condition iff

1. the φ_i are continuous;
2. and the following equivalent statements hold:
   - for all $x_0, x_1, \ldots, x_n \in [a,b]$,

$$\left|\varphi_i\left(x_j\right)\right|_{0 \leqslant i,j \leqslant n} = 0 \quad \Leftrightarrow \quad \exists i \neq j, x_i = x_j,$$

   - given pairwise distinct $x_0, \ldots, x_n \in [a,b]$ and values $y_0, \ldots, y_n$, there exists a unique interpolant

$$p = \sum_{k=0}^{n} \alpha_k \varphi_k, \text{ with } \alpha_k \in \mathbb{R}, \forall k = 0, \ldots, n, \text{ s.t. } p\left(x_i\right) = y_i,$$

   - any $p = \sum_{k=0}^{n} \alpha_k \varphi_k \neq 0$ has at most n distinct zeros.

Section 2.2. Best L^∞ (or minimax) approximation

A set of functions that satisfy the Haar condition is called a Chebyshev system. The prototype example is $\varphi_i\left(x\right) = x^i$, for which we have

$$\begin{vmatrix} \varphi_0\left(x_0\right) & \cdots & \varphi_n\left(x_0\right) \\ \vdots & & \vdots \\ \varphi_0\left(x_n\right) & \cdots & \varphi_n\left(x_n\right) \end{vmatrix} = \begin{vmatrix} 1 & \cdots & x_0^n \\ \vdots & & \vdots \\ 1 & \cdots & x_n^n \end{vmatrix} = V_n = \prod_{i<j}\left(x_i - x_j\right).$$

Other examples include (exercise: prove that it is indeed the case!):

- $\left\{e^{\lambda_0 x}, \ldots, e^{\lambda_n x}\right\}$ for $\lambda_0 < \lambda_1 < \cdots < \lambda_n$;
- $\{1, \cos x, \sin x, \ldots, \cos(nx), \sin(nx)\}$ over $[a,b]$ where $0 \leqslant a < b < 2\pi$;
- $\{x^{\alpha_0}, \ldots, x^{\alpha_n}\}$, $\alpha_0 < \cdots < \alpha_n$, over $[a,b]$ with $a > 0$.

Section 2.2. Best L^∞ (or minimax) approximation

Let E be a real vector space, $e_1, e_2, \ldots, e_m \in E$, we will denote $\mathrm{Span}_{\mathbb{R}} \{e_1, \ldots, e_m\}$ the set $\{\sum_{k=1}^{m} \alpha_k e_k; e_{1,} \ldots, e_m \in \mathbb{R}\}$.

If $\{\varphi_0, \ldots, \varphi_n\}$ is a Chebyshev system over $[a, b]$, any element of $\mathrm{Span}_{\mathbb{R}} \{\varphi_0, \ldots, \varphi_n\}$ will be called generalized polynomial.

Section 2.2. Best L^∞ (or minimax) approximation

Theorem 6

[Alternation Theorem. Chebyshev? Borel (1905)? Kirchberger (1902)]

Let $\{\varphi_0, \ldots, \varphi_n\}$ be a Chebyshev system over $[a, b]$. Let $f \in \mathcal{C}([a, b])$. A generalized polynomial $p = \sum_{k=0}^{n} \alpha_k \varphi_k$ is the best approximation (or minimax approximation) of f iff there exist $n+2$ points $x_0, \ldots, x_{n+1}$, $a \leqslant x_0 < x_1 < \cdots < x_{n+1} \leqslant b$ such that, for all k,

$$f(x_k) - p(x_k) = (-1)^k (f(x_0) - p(x_0)) = \pm \|f - p\|_\infty .$$

Section 2.2. Best L^∞ (or minimax) approximation

Theorem 6

[Alternation Theorem. Chebyshev? Borel (1905)? Kirchberger (1902)]

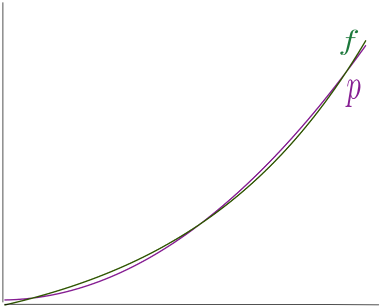
Let $\{\varphi_0, \ldots, \varphi_n\}$ be a Chebyshev system over $[a, b]$. Let $f \in \mathcal{C}([a, b])$. A generalized polynomial $p = \sum_{k=0}^{n} \alpha_k \varphi_k$ is the best approximation (or minimax approximation) of f iff there exist $n+2$ points $x_0, \ldots, x_{n+1}$, $a \leqslant x_0 < x_1 < \cdots < x_{n+1} \leqslant b$ such that, for all k,

$$f(x_k) - p(x_k) = (-1)^k (f(x_0) - p(x_0)) = \pm \|f - p\|_\infty .$$

In other words, $p = \sum_{k=0}^{n} \alpha_k \varphi_k$ is the best approximation if and only if the error function $f - p$ has (at least) $n+2$ extrema, all global (of the same absolute value) and with alternating signs.

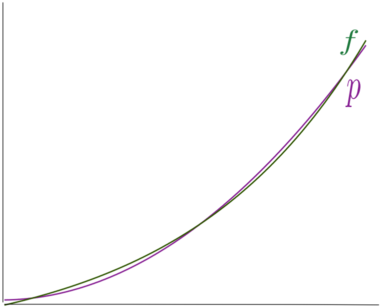
Section 2.2. Best L^∞ (or minimax) approximation

$f(x) = e^{1/\cos(x)}$, $x \in [0,1]$, $p(x) = \sum_{i=0}^{10} c_i x^i$ its minimax approximation,



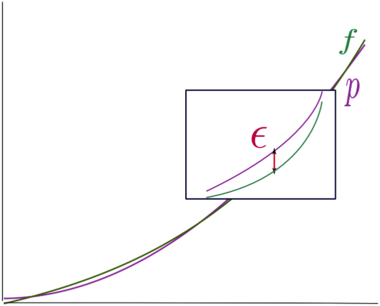
Section 2.2. Best L^∞ (or minimax) approximation

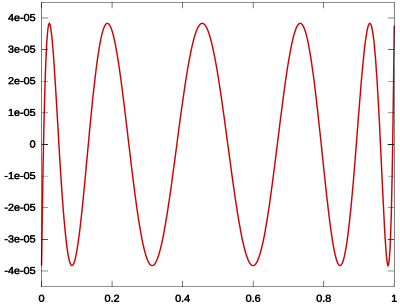
$f(x) = e^{1/\cos(x)},\ x \in [0,1],\ p(x) = \sum_{i=0}^{10} c_i x^i$ its minimax approximation, $\varepsilon(x) = f(x) - p(x)$



Section 2.2. Best L^∞ (or minimax) approximation

$f(x) = e^{1/\cos(x)}$, $x \in [0, 1]$, $p(x) = \sum_{i=0}^{10} c_i x^i$ its minimax approximation, $\varepsilon(x) = f(x) - p(x)$ s.t. $\|\varepsilon\|_\infty = \sup_{x\in[a,\,b]}\{|\varepsilon(x)|\}$ is as small as possible



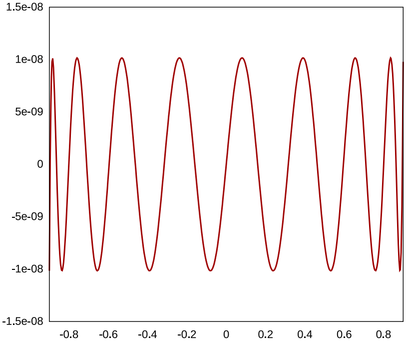


Section 2.2. Best L^∞ (or minimax) approximation

Example:

$f(x) = \arctan(x)$ over $[-0.9, 0.9]$
$p(x) =$ minimax, degree 15
$\varepsilon(x) = p(x) - f(x)$

$\|\varepsilon\|_\infty \simeq 10^{-8}$



Section 2.2. Best L^∞ (or minimax) approximation

Can you tell me what is the best approximation of $\cos x$ over $[0, 10\pi]$ on the Chebyshev system $\{1, x, x^2\}$? on $\{1, x, \dots, x^h\}$ up to and including $h = 9$?

Section 2.2. Best L^∞ (or minimax) approximation

Theorem 7 (Alternation Theorem. Chebyshev? Borel (1905)? Kirchberger (1902))

Let $\{\varphi_0, \ldots, \varphi_n\}$ *be a Chebyshev system over* $[a, b]$*. Let* $f \in \mathcal{C}([a, b])$*. A generalized polynomial* $p = \sum_{k=0}^{n} \alpha_k \varphi_k$ *is the best approximation (or minimax approximation) of* f *iff there exist* $n+2$ *points* $x_0, \ldots, x_{n+1}$*,* $a \leqslant x_0 < x_1 < \cdots < x_{n+1} \leqslant b$ *such that, for all* k*,*

$$f(x_k) - p(x_k) = (-1)^k \left(f(x_0) - p(x_0)\right) = \pm \|f - p\|_\infty .$$

Section 2.2. Best L^∞ (or minimax) approximation

Theorem 7 (Alternation Theorem. Chebyshev? Borel (1905)? Kirchberger (1902))

Let $\{\varphi_0, \ldots, \varphi_n\}$ *be a Chebyshev system over* $[a, b]$. *Let* $f \in \mathcal{C}([a, b])$. *A generalized polynomial* $p = \sum_{k=0}^{n} \alpha_k \varphi_k$ *is the best approximation (or minimax approximation) of* f *iff there exist* $n+2$ *points* $x_0, \ldots, x_{n+1}$, $a \leqslant x_0 < x_1 < \cdots < x_{n+1} \leqslant b$ *such that, for all* k,

$$f(x_k) - p(x_k) = (-1)^k (f(x_0) - p(x_0)) = \pm \|f - p\|_\infty .$$

In other words, $p = \sum_{k=0}^{n} \alpha_k \varphi_k$ *is the best approximation if and only if the error function* $f - p$ *has* $n+2$ *extrema, all global (of the same absolute value) and with alternating signs.*