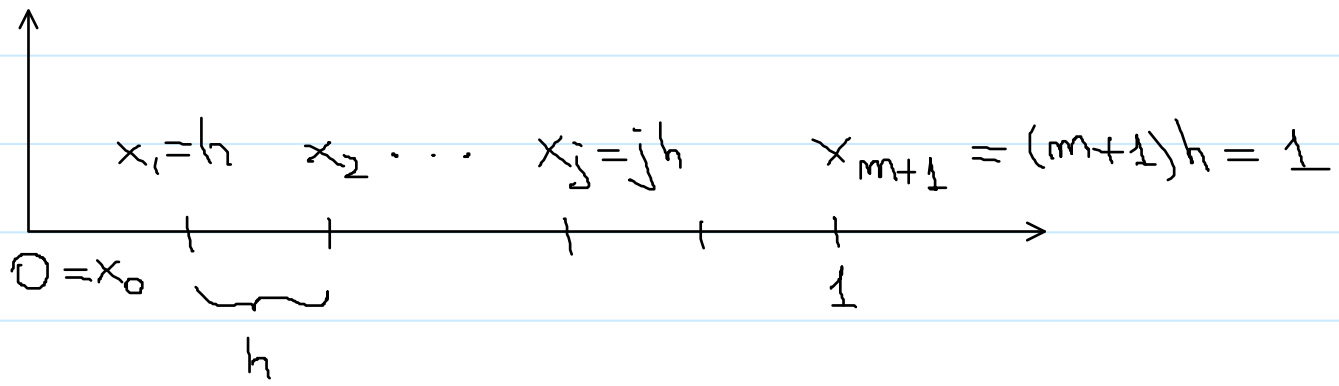
Let's see how accuracy, stability and convergence play out in a simple PDE problem. (General treatment: Lax-Richtmyer theory.)

We do this for the simple PDE (ODE!) boundary value problem

$$\begin{cases} u''(x) = f(x), & 0 < x < 1 \\ u(0) = \alpha, \quad u(1) = \beta \end{cases}$$

Closed-form solution! (Integration constants obtained from α and β.

In a finite difference method we obtain a "grid function" $U_0, U_1, \ldots, U_m, U_{m+1}$, where U_j is an approximation of $u(x_j)$:



$$h = \frac{1}{m+1}$$

$$U_0 = \alpha, \quad U_{m+1} = \beta \qquad (\text{bdry. conds.})$$

We produce a finite-difference approximation by substituting

$$\frac{d^2}{dx^2} \longrightarrow D^2\{U_j\} = \frac{1}{h^2}\left(U_{j-1} - 2U_j + U_{j+1}\right)$$

(Error $\simeq O(h^2)$). Our discrete problem thus reads

$$\frac{1}{h^2}\left(U_{j-1} - 2U_j + U_{j+1}\right) = f(x_j), \quad j = 1, 2, \ldots, m.$$

(1st eqn. involves $U_0 = \alpha$, last eqn. involves $U_{m+1} = \beta$.)

This is a linear system

$$AU = F$$

of m equations and m unknowns, where

$$U = [U_1, U_2, \ldots, U_m]^T$$

and

$$A = \frac{1}{h^2}\begin{bmatrix} -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{bmatrix}, \qquad F = \begin{bmatrix} f(x_1) - \alpha/h^2 \\ f(x_2) \\ f(x_3) \\ \vdots \\ f(x_{m-1}) \\ f(x_m) - \beta/h^2 \end{bmatrix}.$$

It can be solved rapidly: it requires $O(m)$ operations.

Error? Form the vector

$$\hat{U} = [u(x_1), u(x_2), \ldots, u(x_m)]^T$$

and define the error vector

$$E = U - \hat{U}$$

which contains the numerical errors at the grid points.

To express the "error" as a single quantity we

use an error norm, such as

$$\|E\|_\infty = \max_{1 \le j \le m} |U_j - u(x_j)|$$

If we show that, e.g., $\|E\|_\infty = O(h^1)$ then

the pointwise error will be $O(h^2)$ as well.

(We could have used other norms, such as the

1-norm

$$\|E\|_{1,h} = h \sum_{j=1}^{m} |E_j|$$

or the 2-norm

$$\|E\|_{2,h} = \left(h \sum_{j=1}^{m} |E_j|^2 \right)^{1/2}$$

———— ○ ————

. . . Brief detour on matrix norms.

Brief detour on Vector and Matrix Norms

Vector norms we consider in the vector spaces $\mathbb{C}^n$ and $\mathbb{R}^n$:

For a vector $x \in \mathbb{C}^n$ or $x \in \mathbb{R}^n$ we define the

p norm $$\|x\|_p = \left(\sum_{j=1}^{n} |x_j|^p \right)^{1/p} \qquad 1 \leq p \leq \infty$$

and the particular case $p=2$, for which it coincides

with the Euclidean norm

$$\|x\|_2 = \left(\sum_{j=1}^{n} |x_j|^2 \right)^{1/2}.$$

The "infinity norm"

$$\|x\|_\infty = \max_{1 \leq j \leq n} |x_j|$$

is thus called as it equals the limit of the p norm

as $p \to \infty$. (Easy to check.)

Like all norms, these norms satisfy

(i) $\|x\| \geq 0$, and $\|x\| = 0$ iff $x = 0$.

(ii) $\|\lambda x\| = |\lambda| \cdot \|x\|$ for all $\lambda \in \mathbb{C}$ (or $\lambda \in \mathbb{R}$ for real vector spaces)

(iii) $\|x+y\| \leq \|x\| + \|y\|$

We also consider norms on the real and complex vector spaces of matrices $\mathbb{R}^{n\times n}$ and $\mathbb{C}^{n\times n}$.

Some of the norms we consider, namely the p matrix norms for matrices $A \in \mathbb{C}^{n\times n}$, are induced by the p vector norm via the relation

$$\|A\|_p = \max_{\substack{x\in\mathbb{C}^n \\ x\neq 0}} \frac{\|Ax\|_p}{\|x\|_p}, \quad (1 \leq p \leq \infty)$$

with obvious modifications for the real case.

Note that for induced norms we have $\|Ax\| \leq \|A\|\|x\|$.

It can be shown (HW problem) that

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij}|$$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|$$

$$\|A\|_2 = \sqrt{\rho(A \cdot A^H)}$$

(ρ = spectral radius = maximum eigenvalue (abs. value))

We also consider the Frobenius norm

$$\|A\|_F = \left(\sum_{i=1,j} |a_{ij}|^2 \right)^{1/2}$$

which is <u>not</u> an induced norm.

All matrix norms satisfy properties (i), (ii) and (iii) above, with x substituted by A.

Note that

$$\|A\|_F = \sqrt{tr(AA^H)}$$

A matrix norm is said to be "submultiplicative" if it satisfies

$$\|A \cdot B\| \leq \|A\| \cdot \|B\|$$

It is easy to check that the induced norms are submultiplicative.

The Frobenius norm is also submultiplicative:

Proof:

$$\|AB\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\left|\sum_{k=1}^{n} a_{ik}b_{kj}\right|^2$$

$$\leq \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{k=1}^{n}|a_{ik}|^2\sum_{k=1}^{n}|b_{kj}|^2\right)$$ (Cauchy-Schwarz)

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{k,l=1}^{n}|a_{ik}|^2\,|b_{lj}|^2\right)$$

$$= \sum_{i=1}^{n}\sum_{k=1}^{n}|a_{ik}|^2\sum_{l=1}^{n}\sum_{j=1}^{n}|b_{lj}|^2$$

$$= \|A\|_F^2\|B\|_F^2$$

Note that a matrix norm may not be submultiplicative.

Indeed, consider the matrix norm

$$\|M\| = \max_{1\leq i,j\leq m}(|m_{ij}|) \qquad (1)$$

This is a matrix norm (it satisfies (i), (ii), (iii))

but it is not submultiplicative. Indeed, consider the matrix

$$A = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, \text{ for which } A^2 = \begin{pmatrix} 8 & 8 \\ 8 & 8 \end{pmatrix}$$

Using the norm (1) we obtain

$$\|A\,A\| = 8 \not\leq \|A\| \cdot \|A\| = 2 \cdot 2 = 4.$$

This concludes our discussion of matrix norms.

Continuing with our convergence and stability analysis initiated on pp. 126-130 we seek to show that, as indicated on p. 129,

$$\|E\|_{2,h} = O(h^1)$$

To do this we consider first the error inherent in the derivative discretization, and associated "local truncation error".

Local Truncation Error

Continuing from page 129, to study the error in the approximation U, we substitute the exact solution

$$\hat{U} = [u(x_1), u(x_2), \ldots, u(x_m)]^t$$

into the approximate equation, i.e., we seek to estimate

$$z = A\hat{U} - F. \qquad (2)$$

The entries in this difference vector are given by

$$z_j = \frac{1}{h^2}\left(u(x_{j-1}) - 2u(x_j) + u(x_{j+1})\right) - f_j,$$

or, using a Taylor expansion around x_j,

$$z_j = u''(x_j) + \frac{1}{12}h^2 u^{(4)}(x_j) + O(h^4) - f_j.$$

But the differential equation tells us that

$$u''(x_j) = f(x_j) = f_j,$$

and, thus

$$\tau_j = \frac{1}{12} h^2 u^{(4)}(x_j) + O(h^4).$$

This is called the local truncation error.

Since the solution is smooth (provided f is smooth, which we assume) we see that

$$\tau_j = O(h^2) \quad \text{as} \quad h \to 0.$$

We re-express eqn. (2) in the form

$$A\hat{U} = F + \tau \qquad (3)$$

we also have

$$AU = F \qquad (4)$$

and we wish to estimate the global solution error

$$E = U - \hat{U}$$

(p. 129).

In what follows we estimate (bound) the

2-norm $\|E\|_2$ where

$$\|x\|_2 = \left(\sum_{j=1}^{n} |x_j|^2 \right)^{1/2} = \sqrt{\rho(A \cdot A^H)}$$

Subtracting (3) from (4) we obtain

$$AE = -\delta, \qquad (5)$$

or $$\frac{1}{h^2}(E_{j-1} - 2E_j + E_{j+1}) = -\delta_j \quad (j = 1, 2, \ldots, m),$$

with boundary conditions

$$E_0 = E_{m+1} = 0.$$

What to expect from these equations?

Since the RHS tends to zero like $O(h^2)$,

may we expect the same for the error E?

Stability

Let us append an h to the various quantities in (5)

(to emphasize their h-dependence):

$$A^h E^h = -\tau^h$$

(A is an $m \times m$ matrix with $h = \frac{1}{m+1}$. The

dimension grows as $m \to \infty$!)

It follows that

$$E^h = -(A^h)^{-1} \tau^h,$$

so that, taking norms

$$h^{\frac{1}{2}} \| E^h \| = h^{\frac{1}{2}} \| (A^h)^{-1} \tau^h \| \leq \| (A^h)^{-1} \| \underbrace{h^{\frac{1}{2}} \| \tau^h \|}_{\leq C h^2}$$

Need to show: $\| (A^h)^{-1} \|$ is bounded by a constant independent of h:

$$\| (A^h)^{-1} \| \leq C \quad \text{for all } h < h_0.$$

The finite-difference method is said to be stable iff it satisfies this property.

(Not just for this particular equation.)

Consistency

The method is consistent with the differential

equation and the boundary conditions provided

$$\|\epsilon^h\| \xrightarrow[h\to 0]{} 0 .$$

Convergence

$$\|E^h\| \to 0 \quad \text{as} \quad h \to 0 .$$

From the discussion above

$$\text{Consistency} + \text{Stability} \Rightarrow \text{Convergence}$$

$$\underbrace{\|E^h\|}_{\to 0 \text{ (Conv.)}} \leq \overbrace{\|(A^h)^{-1}\|}^{\text{bounded (Stab.)}} \underbrace{\|\epsilon^h\|}_{\to 0 \text{ (Cons.)}}$$

Further, this shows that

$$\|E^h\| = \mathcal{O}(h^p) \quad \text{if} \quad \|\epsilon^h\| = \mathcal{O}(h^p)$$

Let's see how to verify stability for the problem we are considering (In the 2 norm; for the ∞ norm, see Leveque.)

To obtain stability in the 2 norm we compute explicitly the eigenvalues of the matrix

$$A = A^h$$

Since

$$A = \frac{1}{h^2}\begin{bmatrix} -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{bmatrix},$$

and, thus, its inverse, are symmetric the norm $\|\ \|_2$ equals the spectral radius

eigenvalues

$$\|A\|_2 = \rho(A) = \max_{1\le p\le m} |\lambda_p|$$

and (the one that we care about right now)

$$\|A^{-1}\|_2 = \rho(A^{-1}) = \max_{1\le p\le m} |\lambda_p^{-1}|,$$

(cf. p. 133), or

$$\|A^{-1}\|_2 = \left\{ \min_{1\le p\le m} |\lambda_p| \right\}^{-1}$$

We thus need to compute the eigenvalues of $A = A^h$, and show they remain bounded away from 0.

The eigenvalues are solutions of the difference equation

$$\frac{1}{h^2}\left(u_{j-1}^p - 2u_j^p + u_{j+1}^p\right) = \lambda_p u_j^p,$$

$$u_0 = u_m = 0.$$

The solution is

$$\mu_j^p = \sin(p\pi jh):$$

$$(A\mu^p)_j = \frac{1}{h^2}\left[\sin(p\pi(j-1)h) - 2\sin(p\pi jh) + \sin(p\pi(j+1)h)\right]$$

$$= \underbrace{\frac{2}{h^2}\left(\cos(p\pi h) - 1\right)}_{\lambda_p} \cdot \underbrace{\sin(p\pi jh)}_{\mu_p^j}$$

The eigenvalue of smallest magnitude is

$$\lambda_1 = \frac{2}{h^2}\left(\cos(\pi h) - 1\right) \sim -\pi^2 + O(h^2)$$

← eig. of $u'' = \lambda u$

Bounded away from zero: → stability in $\|\ \|_2$

NOT ∞.

Further,

$$\|E^h\|_{2,h} \leq \frac{1}{\pi^2} \frac{h^2}{12} \|f''\|_{2,h}.$$

($u^{(4)}$)

A similar result can be obtained in the $\|\ \|_\infty$ norm

(see Leveque's text.)

Other discretization approaches

It is easy to see that finite difference formulae of the type we have considered coincide with derivatives of polynomial interpolants at the points used for differentiation. (Slight caveats will be mentioned later.)

To understand this we consider first-order derivatives; it is easy to generalize. We note that, by construction, the error E in a finite difference approximation of order p is proportional to h^p: $E = Ah^p$, where A is a combination of derivatives of order $p+1$ of the differentiated function u (see e.g. pp. 123-124).

It follows that the finite difference formula for differentiation at a given point x is exact for polynomials of degree $\leq p$.

In detail, consider the polynomial interpolant $P(x)$ of degree p,

$$P(x) = \sum_{j=0}^{p} a_j x^j$$

that interpolates the given $p+1$ values

$$u(x_0), u(x_1), \ldots, u(x_p)$$

of a function u.

The error in the finite difference approximation of the first derivative of P on the basis of the values

$$P(x_0), P(x_1), \ldots, P(x_p)$$

at a point $\bar{x}$ equals a combination of derivatives of P of order $p+1$, and, therefore, equals zero: the finite difference formula produces the exact derivative of the polynomial at the point $\bar{x}$.

But, since we have $u(x_0) = P(x_0), \ldots, u(x_p) = P(x_p)$, the finite-difference approximation for $u'(\bar{x})$ equals $P'(\bar{x})$.

Caveat concerning the relation between point numbers $p+1$ and resulting approximation errors $\mathcal{O}(h^p)$.

As an example, consider the expression for D_0:

$$\frac{u(\bar{x}+h)-u(\bar{x}-h)}{2h} = u'(\bar{x}) + \mathcal{O}(h^2)$$

($\mathcal{O}(h^2)$: depends on u''')

Clearly, this formula is exact for polynomials of degree ≤ 2. But the expression uses just two points: $\bar{x}-h$ and $\bar{x}+h$! If we interpolate just using two points we obtain a linear interpolant, which would result in error of order $\mathcal{O}(h)$.

What happens is that in the quadratic interpolant through $\bar{x}-h$, $\bar{x}$ and $\bar{x}+h$, the

coefficient of $u(\bar{x})$ in the finite difference expression is equal to zero. In other words we are really using $p+1=3$ interpolation points, although only two points appear explicitly in the finite difference formula. (Adding a quadratic polynomial which vanishes at $\bar{x}-h$ and $\bar{x}+h$ we can match any desired value at $\bar{x}$ without changing the derivative of the polynomial at $\bar{x}$.)

In view of the equivalence between finite differences and differentiation of polynomials, we consider polynomial interpolation in more detail. Polynomial interpolants also play an important role in numerical integration.

Given the improvement in accuracy we have observed as the degree p of the approximation is increased, how much should (could) we increase the order?

> Note: Issue with finite-difference PDE methods near boundaries: need to use one-sided differences. This can have an impact on stability.

Should we consider polynomial interpolants of arbitrarily high order? ANS: Yes and No.

We consider this question in detail in what follows.

To do this we consider methods to produce polynomial interpolants of a given degree, and the associated

errors they entail.

Methods to produce polynomial interpolants:

1) Lagrange formula.

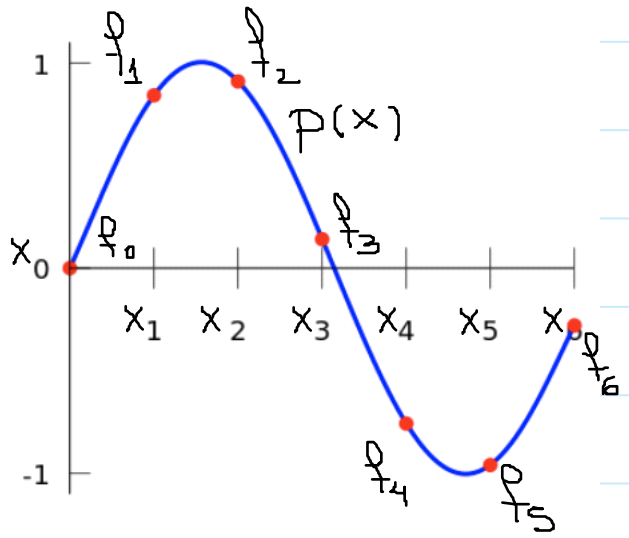
We first note that the polynomial

$$l_j(x) = \frac{\prod_{\substack{k=0 \\ k\neq j}}^{n} (x - x_k)}{\prod_{\substack{k=0 \\ k\neq j}}^{n} (x_j - x_k)}$$

satisfies

$$l_j(x_m) = \begin{cases} 1 & m = j \\ 0 & m \neq j \end{cases} \qquad (m = 0, \ldots, n).$$

Thus given function values $f_0, \ldots, f_n$ to be interpolated at $x_0, \ldots, x_n$, the interpolating polynomial is given by



$$p(x) = \sum_{j=0}^{n} f_j l_j(x) .$$

2) Matrix inversion

We seek the polynomial

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$$

satisfying

$$p(x_j) = f_j \qquad (j = 0, 1, \cdots, n).$$

Clearly, the coefficients satisfy the system

of equations

$$\underbrace{\begin{bmatrix} x_0^n & x_0^{n-1} & x_0^{n-2} & \dots & x_0 & 1 \\ x_1^n & x_1^{n-1} & x_1^{n-2} & \dots & x_1 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ x_n^n & x_n^{n-1} & x_n^{n-2} & \dots & x_n & 1 \end{bmatrix}}_{V = \text{Vandermonde matrix}} \begin{bmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_0 \end{bmatrix} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{bmatrix}$$

The Vandermonde matrix is invertible (e.g., by uniqueness of interpolating polynomial, or by using the relation

$$\det(V) = \prod_{\substack{i,j=0 \\ i<j}}^{n} (x_i - x_j) \neq 0$$

Unfortunately, this matrix is very ill conditioned, as discussed below.

[Note: In seeking the finite difference coefficients by solving equations the Vandermonde matrix is obtained once again. (See Leveque p. 11.)]

Detours on matrix condition number and polynomial interpolation. We will see that high order can still be achieved, and can be highly beneficial.]

Condition Number

The condition number quantifies the accuracy degradation that occurs as a result of the solution

of a matrix equation problem.

$$Ax = b \qquad (1)$$

The condition number, denoted by $\kappa(A)$ or $\text{cond}(A)$, is defined by

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\| ,$$

where $\|\cdot\|$ denotes a matrix norm, eg.

$$\|\cdot\|_1 , \quad \|\cdot\|_2 , \quad \|\cdot\|_\infty .$$

Letting

$$\varepsilon_x = \frac{\|\delta x\|}{\|x\|}$$

denote the relative error in the solution of (1) caused by relative errors

$$\varepsilon_A = \frac{\|\delta A\|}{\|A\|} \quad \text{and} \quad \varepsilon_b = \frac{\|\delta b\|}{\|b\|}$$

in the given matrix and/or right hand side, then, roughly speaking, the condition number gives the error amplification factor:

$$\varepsilon_x \sim \text{cond}(A) \cdot (\varepsilon_A + \varepsilon_b)$$

(A more precise and rigorous bound is given below.)

Simplest derivation, under the restriction $\delta A = 0$ (only error $\delta b \neq 0$ is assumed).

We have

$$Ax = b, \quad A(x+\delta x) = b + \delta b$$

It follows that

$$A\delta x = \delta b.$$

The growth factor in the relative error is given by

$$\frac{\|\delta x\|}{\|x\|} \Big/ \frac{\|\delta b\|}{\|b\|} = \frac{\|A^{-1}\delta b\|}{\|A^{-1}b\|} \cdot \frac{\|b\|}{\|\delta b\|} =$$

$$= \frac{\|A^{-1}\delta b\|}{\|\delta b\|} \cdot \frac{\|b\|}{\|A^{-1}b\|} = \frac{\|A^{-1}\delta b\|}{\|\delta b\|} \cdot \frac{\|Ax\|}{\|x\|} \leq$$

$$\leq \max_{\delta b \neq 0} \frac{\|A^{-1}\delta b\|}{\|\delta b\|} \cdot \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \|A\| \cdot \|A^{-1}\|$$

$$\Rightarrow \quad \frac{\|\delta x\|}{\|x\|} \leq \underbrace{\|A\| \cdot \|A^{-1}\|}_{\begin{array}{c}\| \\ \text{cond}(A) \\ \| \\ K(A)\end{array}} \frac{\|\delta b\|}{\|b\|}$$

that is to say

$$\frac{\|\delta x\|}{\|x\|} \leq \text{cond}(A) \frac{\|\delta b\|}{\|b\|}$$

Example:

$$x = 1$$

$$x - y = 0$$

$$z = (x - y)/\eta \qquad (\eta \text{ small}).$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & \eta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

a small ε error here results in an error $\frac{\varepsilon}{\eta}$ in z. Large!

In general, under minimally restrictive conditions, we have the following result (for general $\delta b \neq 0$, $\delta A \neq 0$).

Calling

$$\varepsilon = \max\left\{ \frac{\|\delta A\|}{\|A\|}, \frac{\|\delta b\|}{\|b\|} \right\}$$

$$\varepsilon = \max\left\{\frac{\|\delta A\|}{\|A\|}, \frac{\|\delta b\|}{\|b\|}\right\}$$

then

$$\frac{\|\delta x\|}{\|x\|} \leq \text{cond}(A) \cdot \frac{2\varepsilon}{1 - \underbrace{\varepsilon \cdot \text{cond}(A)}_{\text{assumed} < 1}}$$

(This is the minimally restrictive condition mentioned earlier.)

(E.g. Theorem 2.7.2 in Golub-Van Loan, 3rd Ed.)

Note that having a small determinant has little to do with error amplification!

Ex: $\det\begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix} = \varepsilon^2$

can be very small. Yet, the condition number of this matrix equals 1: the best condition number possible.

$$\left(\text{cond}(A) = \|A\| \cdot \|A^{-1}\| \geq \|A \cdot A^{-1}\| = \|I\| = 1. \right)$$

Returning to polynomial interpolation: the Vandermonde matrix is extremely ill conditioned.

For example, for equidistant nodes in $[0,1]$

$\left(x_j = \frac{j-1}{n-1}, \; j = 1, 2, \ldots, n, \right.$ we have

$$\text{Cond}_\infty(V) \sim 8^n \qquad \text{HUGE}$$

(Gautschi (1990): "How (un)stable are Vandermonde matrices.")

Would it be better to use the Lagrange interpolating polynomial? The evaluation of the coefficients

from it is similarly problematic:

$$p(x) = \sum_{j=0}^{n} f_j \ell_j(x)$$

with

$$\ell_j(x) = \frac{\prod_{\substack{k=0 \\ k \neq j}}^{n} (x - x_k)}{\prod_{\substack{k=0 \\ k \neq j}}^{n} (x_j - x_k)}.$$

To compute the coefficients we would have to expand products.

Example:

$$(x - x_0)^n = \sum_{k=0}^{n} \binom{n}{k} x^k x_0^{n-k} (-1)^{n-k}$$

Huge accuracy loss if x and x_0 are large but $x - x_0$ is small and n is large.

The evaluation of the point values of the interpolant via the Lagrange interpolating polynomial is manifestly much better (see review paper Berrut-Trefethen 2004). Also, Newton form of the interpolating polynomial.

In spite of all this, convergence as $n \to \infty$ generally does not work. Much confusion has arisen Re. ill conditioning and convergence.

Next time: convergence does not take place, unless special considerations are incorporated. If they are → success!

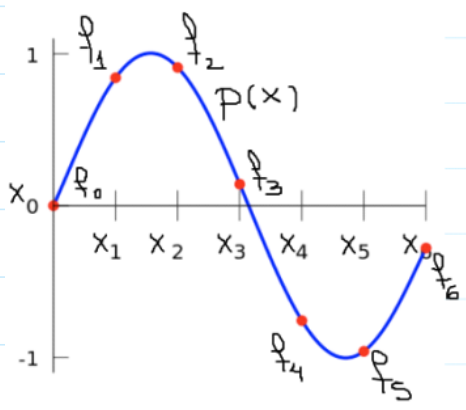
Re. Fornberg, "Practical guide to pseudospectral methods"

Convergence / Divergence of polynomial interpolants

To study the convergence of the polynomial interpolants of a given function we derive an expression analogous to Cauchy's formula, but for the interpolation error

$$R_n(z) = f(z) - \underbrace{\sum_{j=0}^{n} f(x_j)\, l_j(z)}_{p(z)}$$

where (p. 153) $$l_j(z) = \frac{\prod_{\substack{k=0 \\ k\neq j}}^{n} (z - x_k)}{\prod_{\substack{k=0 \\ k\neq j}}^{n} (x_j - x_k)}.$$



May interpolate at complex points z as well!

Defining $$w(z) = \prod_{k=0}^{n} (z - x_k);$$

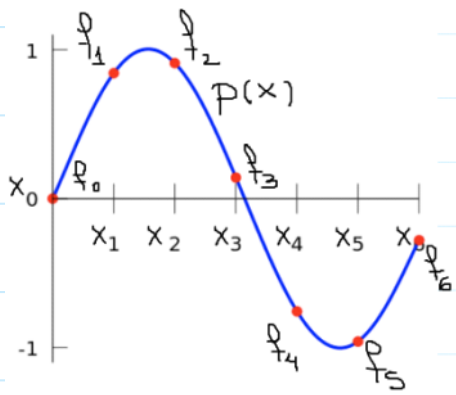
it is easily checked that

$$R_n(z) = f(z) - \sum_{j=0}^{n} f(x_j) \cdot \frac{\omega(z)}{(z-x_j)\,\omega'(x_j)} .$$

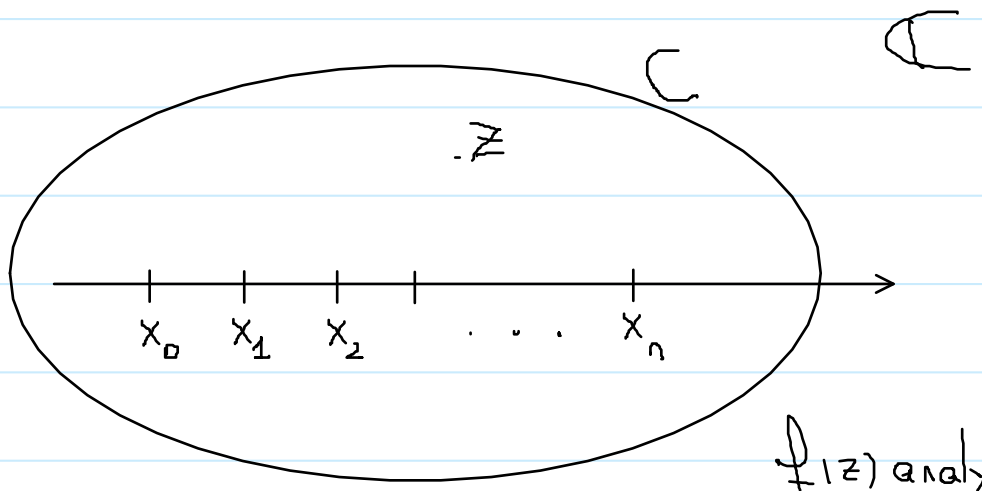
In view of the residue theorem we then obtain

$$R_n(z) = \frac{\omega(z)}{2\pi i} \int_C \frac{f(t)\,dt}{\omega(t)(t-z)} \qquad (1)$$

where C is any closed curve which, together with its interior, is contained within the domain of analyticity of the function f, and which contains z as well as the interpolation points $x_1, \ldots, x_n$.



C simple closed curve.



$f(z)$ analytic within C and continuous up to and including C

To check the validity of (1) we apply the residue theorem. The poles of the integrand are simple, and they are located at

$$t = z \quad \text{and} \quad t = x_j \quad (j = 0, \cdots, n)$$

unless f vanishes at one or more of these points.

Residue at $t = z$: $f(z)/w(z)$

Residue at $t = x_j$: $f(x_j)/(w'(x_j)(x_j - z))$

By the residue theorem the RHS in (1) equals

$$\frac{w(z)}{2\pi i} \, 2\pi i \sum (\text{residues within } C) =$$

$$= f(z) - \sum_{j=0}^{n} f(x_j) \cdot \frac{w(z)}{(z - x_j) w'(x_j)} = R_n(z)$$

as claimed.

To estimate $R_n(z)$ we first estimate $\omega(z)$:

$$|\omega(z)| = \prod_{k=0}^{n} |z - x_k| = e^{\sum_{k=0}^{n} \log|z - x_k|} =$$

$$\sim e^{-n\phi(z)}$$

where

$$-\phi(z) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} \log|z - x_k| =$$

$$= \int_{-1}^{1} \mu(x) \log|z - x| dx$$

where $\mu(x)$ is the node-density function.

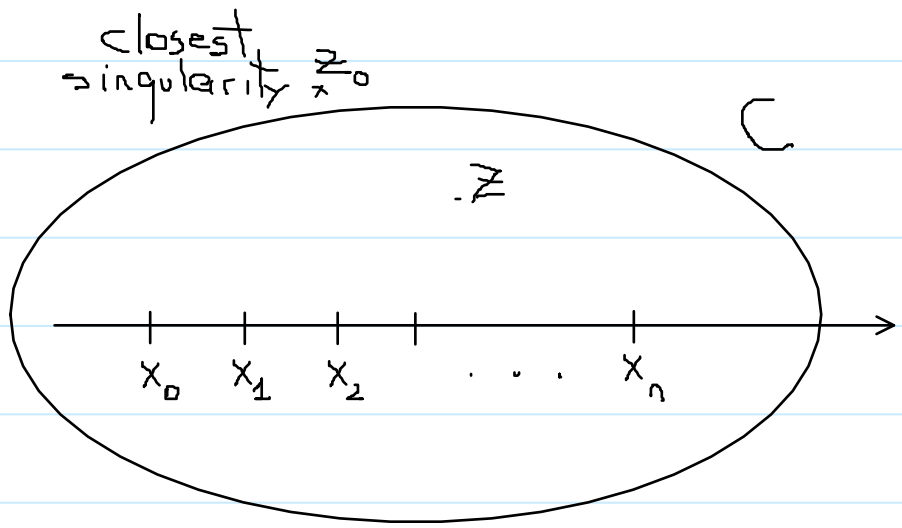
(For example, $\mu(x) = \frac{1}{2}$ for equidistant nodes in $[-1, 1]$.)

We now choose C to be a level curve for

ϕ that barely leaves z_0 out:

$$\phi(z) = M > \phi(z_0) + \varepsilon$$

($\phi(z)$: on C; M: constant; ε: small)



(Curves $\phi = M$ "grow" as M grows.)

Then, considering (1), we obtain

$$\left| \int_C \frac{f(t)\,dt}{w(t)(t-z)} \right| \leq \frac{1}{\min_{z \in C} |w(z)|} \int_C \left| \frac{f(t)}{(t-z)} \right| dt \leq$$

$$\leq C(\varepsilon)\, e^{n(\phi(z_0)+\varepsilon)},$$

and, thus, from (1)

$$|R_n(z)| \leq \tilde{C}(\varepsilon)\, e^{n[(\phi(z_0)+\varepsilon)-\phi(z)]}$$

Therefore

$$|R_n(z)|^{1/n} \xrightarrow[n\to\infty]{} e^{-[\phi(z)-\phi(z_0)]}$$

If z is inside the contour then $\phi(z)$ is "less negative" than $\phi(z_0) \Rightarrow$ exponent is negative $\Rightarrow R_n(z) \xrightarrow[n\to\infty]{} 0$. If z is outside the contour that touches z_0, then $R_n(z)$ diverges.

Interesting!

As in the case of Taylor expansions, we have convergence inside a certain curve, and divergence outside that curve.

For uniform interpolation in the interval $[-1,1]$ $\left(\mu = \frac{1}{2}\right)$ we have

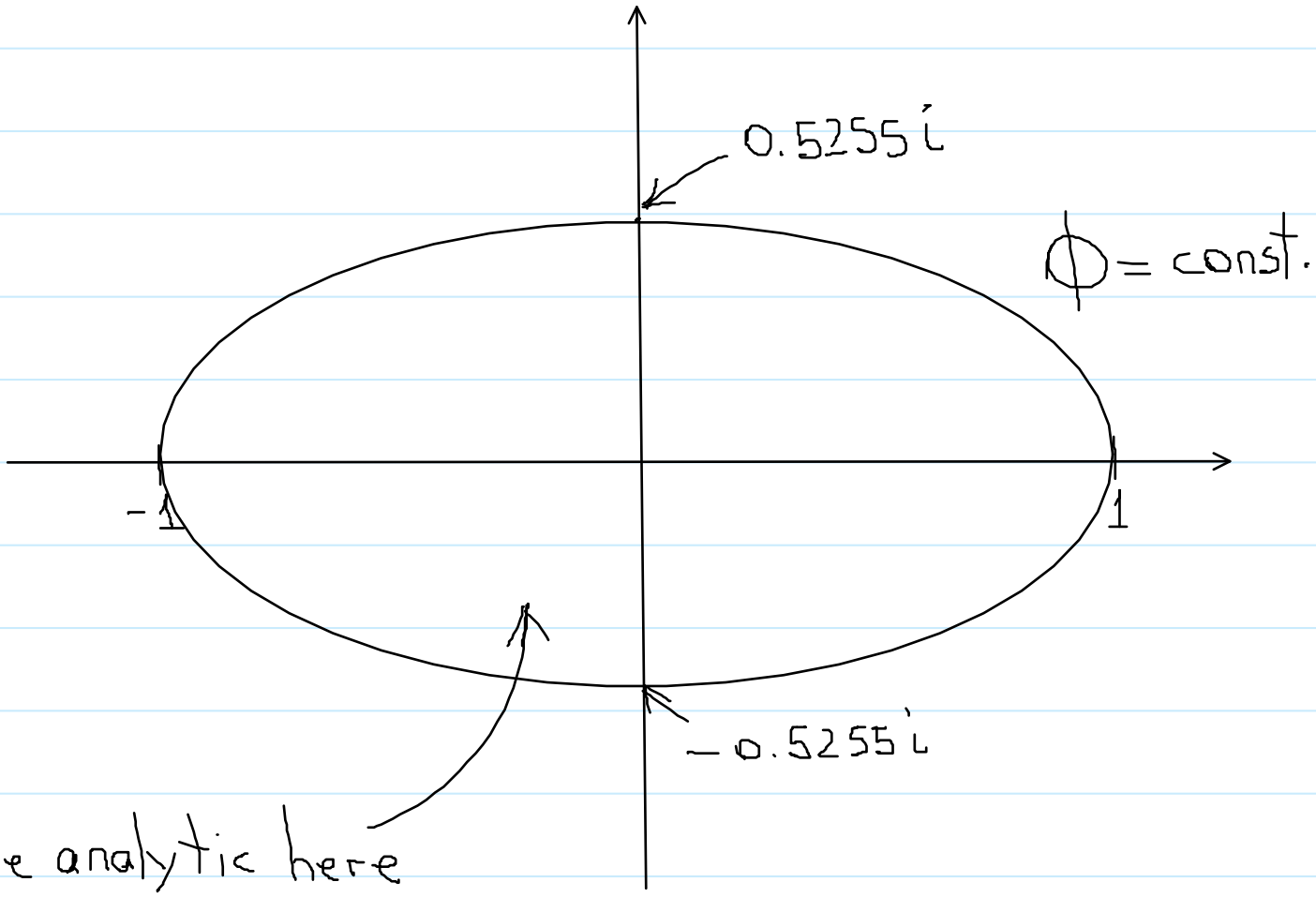
$$\Phi(z) = -\frac{1}{2}\int_{-1}^{1} \log|z-x|\,dx =$$

$$= -\frac{1}{2}\,\mathrm{Re}\int_{-1}^{1} \log(x-z)\,dx =$$

$$= -\frac{1}{2}\left[(x-z)\log(x-z)\right]_{-1}^{1} + C =$$

$$= -\frac{1}{2}\,\mathrm{Re}\left[(1-z)\log(1-z) - (-1-z)\log(-1-z)\right]$$

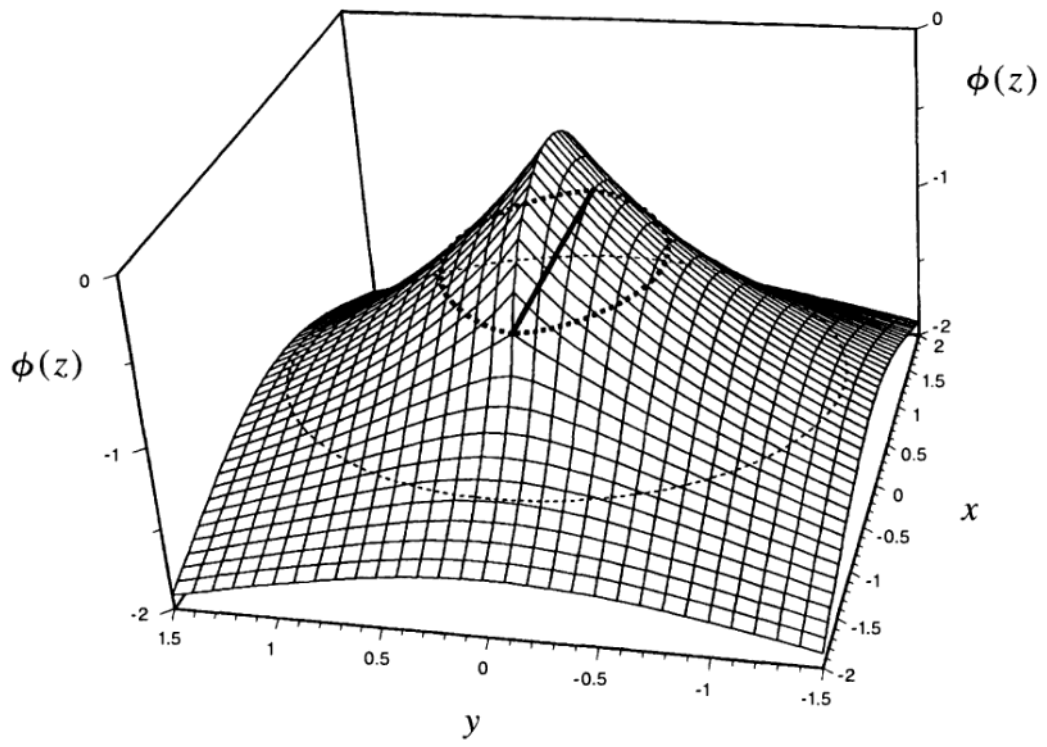
$$+ C$$



for the interpolants

to converge

in $[-1,1]$



Similarly, consider the "Runge" function

$$f(x) = \frac{1}{1+16x^2}$$

(singularities at $\pm 0.25i$).

The curve

$$\phi(z) = \phi(0.25i)$$

cuts the real axis at ± 0.7942

Convergence of the interpolant only in the interval $(-0.7942, 0.7942)$.

Convergence for

$$f(x) = \frac{1}{1+16x^2}$$

shown below.

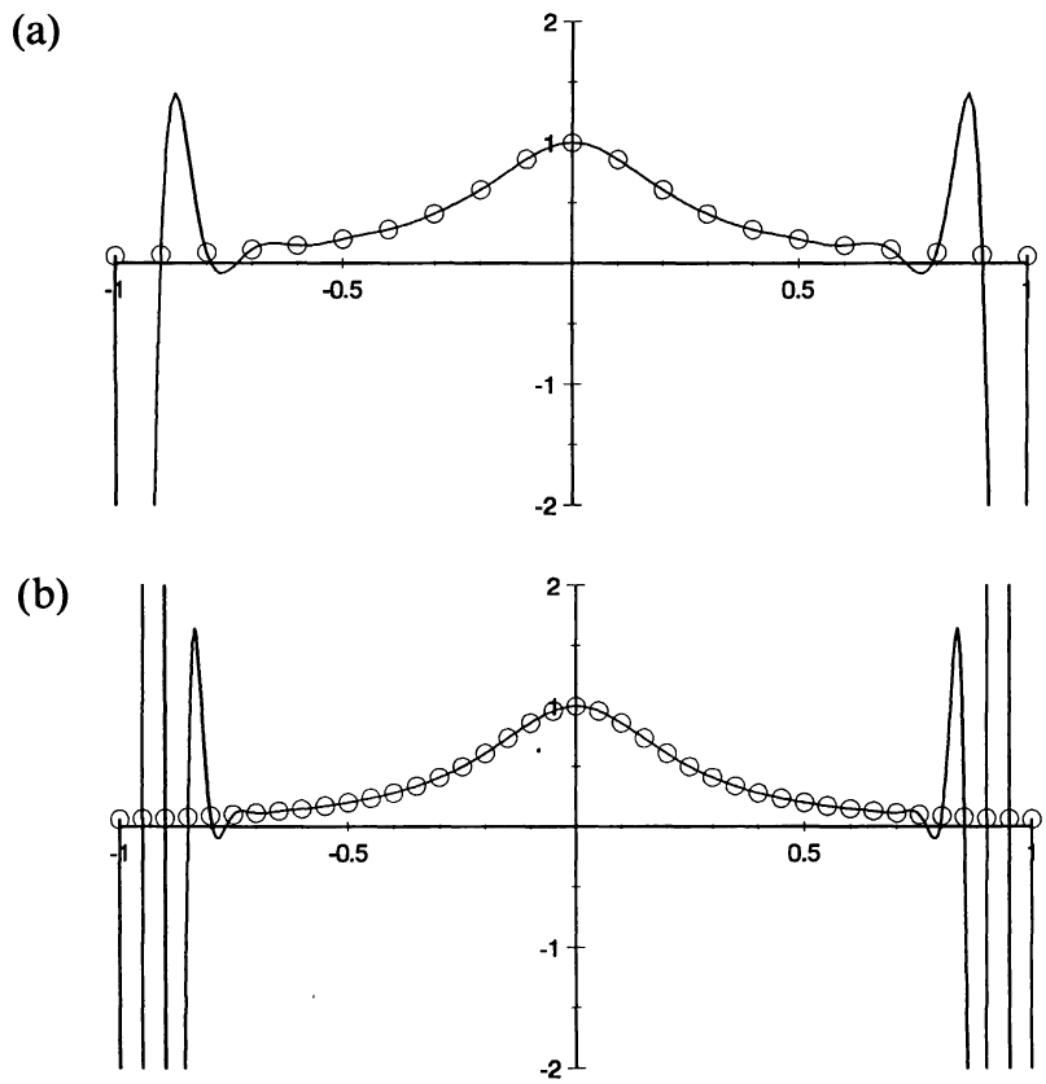


Figure 3.4-3. Results of equi-spaced interpolation on $[-1, 1]$ in the case of **(a)** $N = 20$ and **(b)** $N = 40$.

Complete failure? No!

For interpolation by Chebyshev polynomials (or, more generally, Jacobi polynomials we will consider), there is a different set of

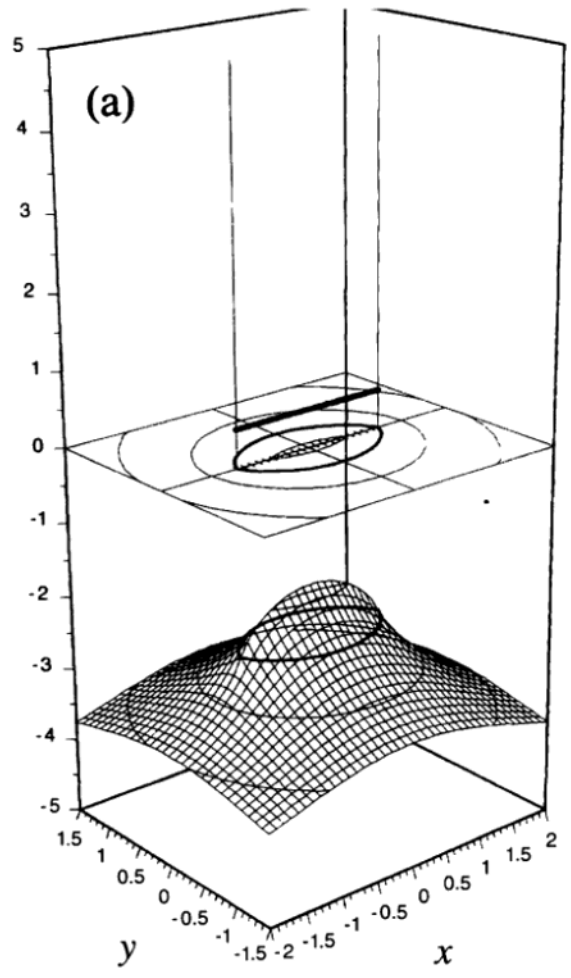
interpolation points. For such cases we have

$\mu(x) = \frac{1}{\pi}(1-x^2)^{-1/2}$, and it follows that

$$\Phi(z) = -\log|z + \sqrt{z^2-1}|$$

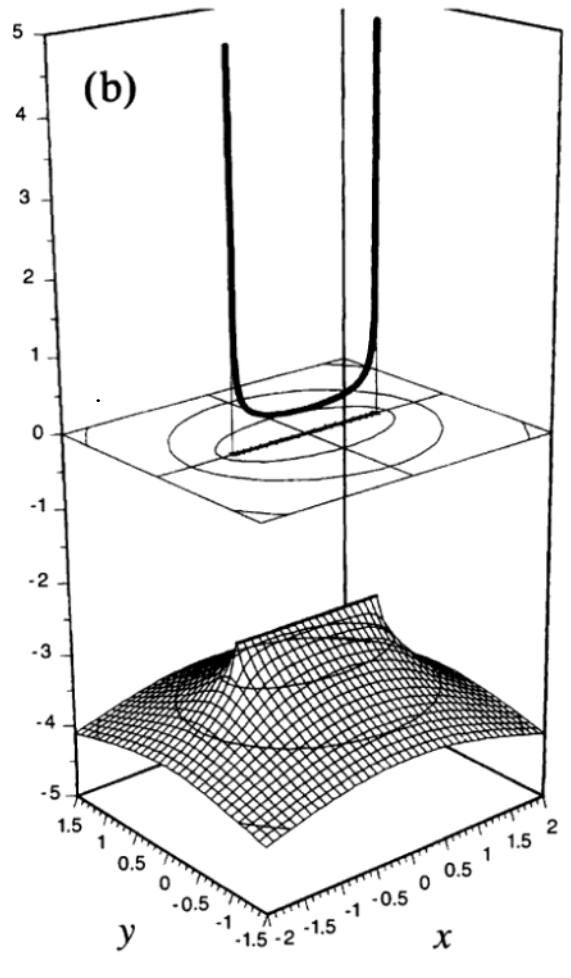
Function Φ

Equispaced



Convergence region shrinks

Chebyshev/Jacobi



Convergence region always contains $[-1,1]$!