A brief introduction to pseudo-spectral methods: application to diffusion problems

Denys Dutykh

Denys Dutykh
CNRS–LAMA, Université Savoie Mont Blanc, France

A brief introduction to pseudo-spectral methods: application to diffusion problems

Last modified: June 17, 2016

A brief introduction to pseudo-spectral methods: application to diffusion problems

Denys Dutykh*

Abstract. The topic of these notes could be easily expanded into a full one-semester course. Nevertheless we shall try to give some flavour along with theoretical bases of spectral and pseudo-spectral methods. The main focus is made on Fourier-type discretizations, even if some indications on how to handle non-periodic problems via Tchebyshev and Legendre approaches are made as well. The applications presented here are diffusion-type problems in accordance with the topics of the PhD school.

Key words and phrases: pseudo-spectral methods; approximation theory; diffusion; parabolic equations

MSC: [2010] 65M70, 65N35 (primary), 80M22, 76M22 (secondary)
PACS: [2010] 47.11.Kb (primary), 44.35.+c (secondary)

Key words and phrases. pseudo-spectral methods; approximation theory; diffusion; parabolic equations.

* Corresponding author.

Contents

1 General introduction . . . 5

2 Introduction to spectral methods . . . 6

2.1 Choice of the basis . . . 7

Periodic problems . . . 7

Non-periodic problems . . . 8

2.2 Determining expansion coefficients . . . 14

3 Aliasing, interpolation and truncation . . . 16

3.1 Example of a second order boundary value problem . . . 18

Tau–Lanczos . . . 19

Galerkin . . . 19

Collocation . . . 20

4 Application to heat conduction . . . 21

4.1 An elementary example . . . 21

4.2 A less elementary example . . . 22

4.3 A real-life example . . . 23

5 Indications for further reading . . . 26

A Some identities involving Tchebyshev polynomials . . . 28

A.1 Compositions of Tchebyshev polynomials . . . 31

B Trefftz method . . . 31

C A brief history of diffusion in Physics . . . 34

D Monte–Carlo approach to the diffusion simulation . . . 39

D.1 Brownian motion generation . . . 43

E An exact non-periodic solution to the 1D heat equation . . . 43

F Some popular numerical schemes for ODEs . . . 45

F.1 Existence and unicity of solutions . . . 49

Counterexamples . . . 51

Acknowledgments . . . 52

References 52

1. General introduction

A numerical simulation is like sex.
If it is good, then it is great.
If it is bad, then it is still better than nothing.

— Dr. D (paraphrasing Dr. Z)

People think they do not understand Mathematics,
but it is all about how you explain it to them.

— I. M. Gelfand

This document represents a collection of brief notes of lectures delivered by the Author at the International PhD school "*Numerical Methods for Diffusion Phenomena in Building Physics: Theory and Practice*" which took place in Pontifical Catholic University of Parana (PUCPR, Curitiba, Brazil) in April, 2016. The Author of the present document is grateful to the Organizing Committee of this school (in particular to Professors Nathan Mendes and Marx Chhay) for giving him an opportunity to lecture there. The present document should be considered as a supplementary material to the Lectures delivered by the Author at this PhD school. The exposition below is partially based on [11] and some other references mentioned in the manuscript.

This lecture is organized as follows. First, we present some theoretical bases behind spectral discretizations in Section 2. An application to a problem stemming from the building physics is given in Section 4. Finally, we give some indications for the further reading in Section 5. This document contains also a certain number of Appendices directly or indirectly related to spectral methods. For instance, in Appendix A we give some useful identities about Tchebyshev polynomials and in Appendix B we give some flavour of Trefftz methods, which are essentially forgotten nowadays. We decided to include also some history of the diffusion in Sciences in general and in Physics in particular. You will find this information in Appendix C. Finally, we prepared also an Appendix D devoted to the Monte–Carlo methods to simulate numerically diffusion processes. The Author admits that it is not related to spectral methods, but nevertheless we decided to include it since these methods remain essentially unknown in the community of numerical methods for Partial Differential Equations (PDEs).

2. Introduction to spectral methods

Consider for simplicity a 1D compact* domain, *e.g.* $\mathcal{U} = [-1, 1]$ an Ordinary or Partial Differential Equation (ODE or PDE) on it

$$\mathscr{L}u = g, \qquad x \in \mathcal{U} \tag{2.1}$$

where $u(x, t)$ (or just $u(x)$ in the ODE case) is the solution† which satisfies some additional (boundary) conditions at $x = \pm 1$ depending on the (linear or nonlinear) operator $\mathscr{L} = \mathscr{L}(\partial_t, \partial_x, \partial_{xx}, \ldots)$. Function $g(x, t)$ is a source term which does not depend on the solution u. For example, if (2.1) is the classical heat equation in the free space (*i.e.* without heat sources), then

$$\mathscr{L} \equiv \partial_t - \nu \partial_{xx}, \qquad g \equiv 0,$$

where $\nu \in \mathbb{R}$ is the diffusion coefficient. If the solution $u(x)$ is steady (*i.e.* time-independent), we deal with an ODE. In any case, in the present document we focus only on the discretization in space. The time discretization is discussed deeper in the lecture of Dr. Marx Chhay. A brief reminder of some most useful numerical techniques for ODEs is given in Appendix F.

The idea behind a spectral method is to approximate a solution $u(x, t)$ by a finite sum

$$u(x, t) \approx u_n(x, t) = \sum_{k=0}^{n} v_k(t)\, \phi_k(x), \tag{2.2}$$

where $\{\phi_k(x)\}_{k=0}^{\infty}$ is the set of basis functions. In ODE case all $v_k(t) \equiv$ const. The main question which arises is how to choose the basis functions? Once the choice of $\{\phi_k(x)\}_{k=0}^{\infty}$ is made, the second question appears: how to determine the expansion coefficients $v_k(t)$?

Here we shall implicitly assume that the function $u(x, t)$ is *smooth*. Only in this case the full potential of spectral methods can be exploited‡. The concept of *smoothness* in Mathematics is ambiguous since there are various classes of smooth functions:

$$C^p(\mathcal{U}) \supseteq C^\infty(\mathcal{U}) \supseteq \mathscr{A}^\infty(\mathcal{D}), \qquad (p \in \mathbb{Z}^+, \quad \mathcal{U} \subseteq \mathcal{D} \subseteq \mathbb{C}).$$

The last sequence of inclusions needs perhaps some explanations:

$C^p(\mathcal{U})$: the class of functions $f : \mathcal{U} \mapsto \mathbb{R}$ having at least $p \geqslant 1$ continuous derivatives.

*A domain $\mathcal{U} \subseteq \mathbb{R}^d$, $d \geqslant 1$ is compact if it is bounded and closed. For a more general definition of compactness we refer to any course in General Topology.

†For simplicity we consider in this Section scalar equations only. The generalizations for systems is straightforward, since one can apply the same discretization for every individual component of the solution.

‡We have to say that pseudo-spectral methods can be applied to problems featuring eventually discontinuous solutions as well (*e.g.* hyperbolic conservation laws). However, it is out of scope of the present lecture devoted rather to parabolic problems. As a side remark, the Author would add that the advantage of pseudo-spectral methods for nonlinear hyperbolic equations is not so clear comparing to modern high-resolution shock-capturing schemes [38].

$C^\infty(\mathcal{U})$: the class of functions $f : \mathcal{U} \mapsto \mathbb{R}$ having infinitely many continuous derivatives.

$\mathscr{A}^\infty(\mathcal{D})$: the class of functions $f : \mathcal{D} \mapsto \mathbb{C}$ analytical (holomorphic) in a domain $\mathcal{D} \subseteq \mathbb{C}$ containing the segment $\mathcal{U}$ in its interior.

For example, the RUNGE* function $f(x) = \dfrac{1}{1+25\,x^2}$ is in $C^\infty\big([-1,\,1]\big)$, but not analytical in the complex plain $\mathscr{A}^\infty(\mathbb{C})$.

2.1. Choice of the basis

A successful expansion basis meets the following requirements:

(1) [**Convergence**] The approximations $u_n(x,\,t)$ should converge rapidly to $u(x,\,t)$ as $n \to \infty$

(2) [**Differentiation**] Given coefficients $\{v_k(t)\}_{k=0}^n$, it should be easy to determine another set of coefficients† $\{v_k'(t)\}_{k=0}^n$ such that

$$\frac{\partial u_n}{\partial x} = \sum_{k=0}^{n} v_k(t)\,\frac{\mathrm{d}\phi_k(x)}{\mathrm{d}x} \rightsquigarrow \sum_{k=0}^{n} v_k'(t)\,\phi_k(x)\,.$$

(3) [**Transformation**] The computation of expansion coefficients $\{v_k\}_{k=0}^n$ from function values $\big\{u(x_i,t)\big\}_{i=0}^n$ and the reconstruction of solution values in nodes from the set of coefficients $\{v_k\}_{k=0}^n$ should be easy, *i.e.* the conversion between two data sets is algorithmically efficient

$$\big\{u(x_i,t)\big\}_{i=0}^n \quad \leftrightarrows \quad \{v_k\}_{k=0}^n\,.$$

2.1.1 Periodic problems

For periodic problems it is straightforward to propose a basis which satisfies the requirements (1)–(3) above. It consists of *trigonometric polynomials*:

$$u_n(x,t) = a_0(t) + \sum_{k=1}^{n} \big\{a_k(t)\,\cos(k\pi x) + b_k(t)\,\sin(k\pi x)\big\}. \tag{2.3}$$

The first two points are explained in elementary courses of analysis, while the requirement (3) is possible thanks to the invention of the Fast FOURIER Transform (FFT) algorithm first by GAUß and later by COOLEY & TUKEY in 1965 [6].

Remark 1. *The use of trigonometric bases such as (2.3) or $\{\mathrm{e}^{\mathrm{i}\,k\pi x}\}_{k\in\mathbb{Z}}$ for heat conduction problems has been initiated by J.* FOURIER‡ *(1822) [12] even if* FOURIER *series have been*

*Carl David Tolmé RUNGE (1856 – 1927) is a German Mathematician and Physicist. A PhD student of Karl WEIERSTRASS.

†Here the prime does not mean a time derivative!

‡Jean-Baptiste Joseph FOURIER (1768 – 1830) is a French Physicist and Mathematician. In particular he accompanied Napoleon BONAPARTE on his campaign to Egypt as a scientific adviser.

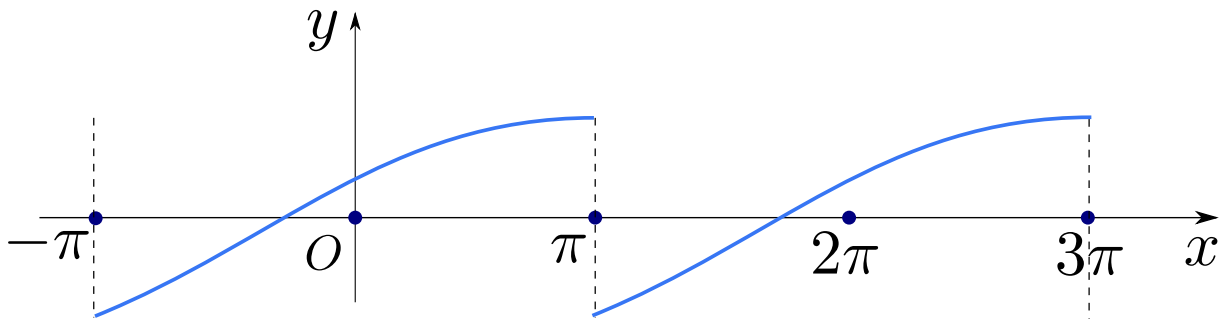


Figure 1. *Periodisation of a smooth continuous function defined on* $[-\pi, \pi]$.

known well before FOURIER. *There is the so-called* ARNOLD*'s*[*] *principle which states that in Mathematics nothing is named after its true inventor. The question the Author would like to rise is whether* ARNOLD*'s principle is applicable to it-self?*

Remark 2. *The term 'Spectral methods' can be now explained. It comes from the fact that the solution* $u(x, t)$ *is expanded into a series of orthogonal eigenfunctions of some linear operator* $\mathcal{L}$ *(with partial or ordinary derivatives). In this way, the numerical solution is related to its spectrum, thus justifying the name 'spectral methods'. For example, if we take the* LAPLACE *operator* $\mathcal{L} = -\boldsymbol{\nabla}^2 \equiv -\sum_{j=1}^{d} \dfrac{\partial^2}{\partial x_j^2}$ *on the periodic domain* $[0, 2\pi)^d$*, its spectrum consists of* FOURIER *modes:*

$$-\boldsymbol{\nabla}^2 \mathrm{e}^{-\mathrm{i}\,\boldsymbol{k}\cdot\boldsymbol{x}} = |\,\boldsymbol{k}\,|^2\,\mathrm{e}^{-\mathrm{i}\,\boldsymbol{k}\cdot\boldsymbol{x}}\,.$$

In this way we obtain naturally the FOURIER *analysis and* FOURIER*-type pseudo-spectral methods.*

2.1.2 Non-periodic problems

The trigonometric basis (2.3) fails to work for general non-periodic problems essentially because of the failure of requirement (1). Indeed, the artificial discontinuities arising after the periodisation (see Figure 1) make the FOURIER coefficients v_n decay as $\mathcal{O}(n^{-1})$ when $n \to \infty$.

The analytic series (or TAYLOR-like expansions) represent another interesting alternative:

$$u_n(x, t) = \sum_{k=1}^{n} v_k(t)\, x^k\,. \tag{2.4}$$

The problem is that TAYLOR-type expansions (2.4) work well only for a very limited class of functions. Namely, they satisfy the requirement (1) only for functions analytical[†] in the

[*] Vladimir ARNOLD (1937 – 2010), a prominent Soviet/Russian mathematician. Please, read his books!

[†] The analyticity property is understood here in the sense of the Complex Analysis.

unit disc $\mathcal{D}_1(\mathbf{0}) \subseteq \mathbb{C}$. For instance, the celebrated RUNGE's function $f(x) = \dfrac{1}{1+25x^2}$ fails to satisfy this condition because of two imaginary poles located at $z = \pm\dfrac{\mathrm{i}}{5}$.

The successful examples of polynomial bases are given by *orthogonal polynomials*, which arise naturally in various contexts:

Numerical integration: Optimal numerical integration formulas achieve a high accuracy by using zeros of orthogonal polynomials as nodes.

Sturm–Liouville problem: JACOBI* polynomials arise as eigenfunctions to singular STURM†–LIOUVILLE‡ problems.

Approximation in L_2: Truncated expansions in LEGENDRE§ polynomials are optimal approximants in the L_2-norm.

Approximation in L_∞: Truncated expansions in TCHEBYSHEV¶ polynomials are optimal approximants in the L_∞-norm.

Each of the topics above deserves a separate course to be covered. Here we content just to provide this information as facts, which can be deepened later, if necessary. We would just like to quote BERNARDI & MADAY (1991):

> *We do think that the corner stone of collocation techniques is the choice of the collocation nodes [...] in spectral methods these are always built from the nodes of a Gauß quadrature formula.*

Consequently, extrema (and zeros) of TCHEBYSHEV (and some other orthogonal) polynomials play a very important rôle in the Numerical Analysis (NA). TCHEBYSHEV nodes are given explicitly by

$$x_k = -\cos\left(\frac{\pi k}{N}\right) \in [-1, 1], \qquad k = 0, 1, 2, \ldots, N. \tag{2.5}$$

There is a result saying that using nodes (2.5) as interpolation points gives an interpolant which is not too far from the optimal polynomial $\mathcal{P}_N^{\mathrm{Opt}}$, *i.e.*

$$\| f - \mathcal{P}_N^{\mathrm{Tch}} \|_\infty \leqslant \left(1 + \Lambda_N^{\mathrm{Tch}}\right) \| f - \mathcal{P}_N^{\mathrm{Opt}} \|_\infty,$$

*Carl Gustav Jacob JACOBI (1804 – 1851) is a German Mathematician.

†Jacques Charles François STURM (1803 – 1855) is a French Mathematician who was born in Geneva which was a part of France at that time.

‡Joseph LIOUVILLE (1809 – 1882) is a French Mathematician who founded the *Journal de Mathématiques Pures et Appliquées*.

§Adrien-Marie LEGENDRE (1752 – 1833) is a French Mathematician. Not to be confused with an obscure politician Louis LEGENDRE. Their portraits are often confused.

¶Pafnuty TCHEBYSHEV (1821 – 1894), a Russian mathematician who contributed to many fields of Mathematics from number theory to probabilities and numerical analysis.

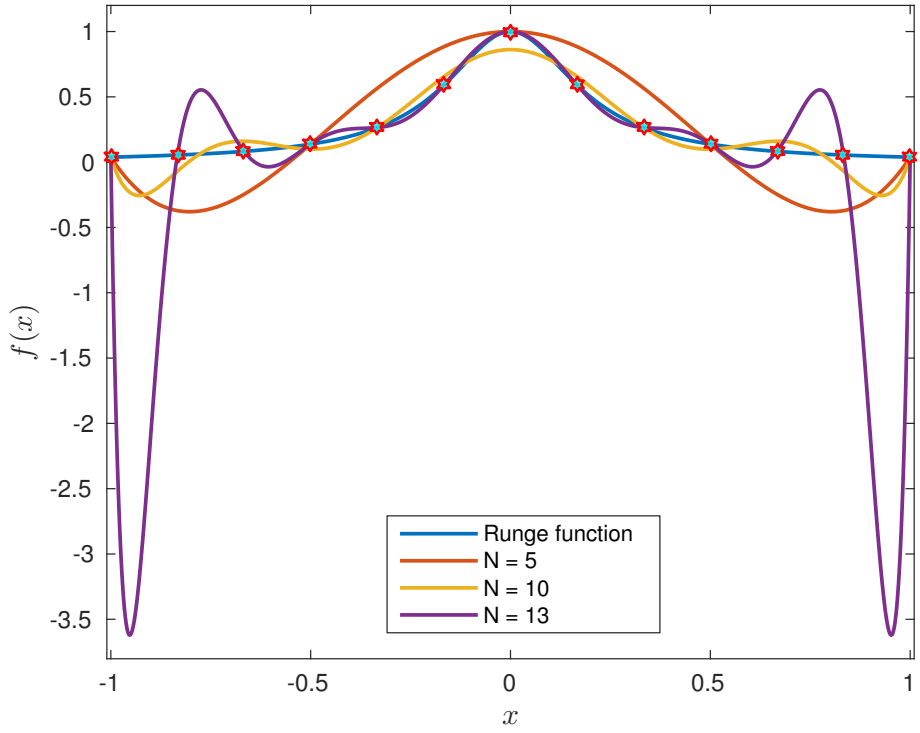


Figure 2. *Interpolation of the* RUNGE *function* $f(x) = \dfrac{1}{1+25x^2}$ *on a sequence of successively refined of uniform grids. One can observe the divergence phenomenon close to the interval boundaries* $x = \pm 1$.

where $f : [-1, 1] \mapsto \mathbb{R}$ is the function that we interpolate and $\mathcal{P}_N^{\mathrm{Tch}}$ is the interpolation polynomial constructed on nodes (2.5). Here Λ_N^{Tch} is the so-called LEBESGUE* constant for TCHEBYSHEV interpolation polynomials of degree N. Notice that Λ_N^{Tch} does not depend on the function f being interpolated. The LEBESGUE constants for various interpolation techniques have the following asymptotic behaviour

$$\Lambda_N^{\mathrm{Tch}} \sim \mathcal{O}(\log N), \qquad \Lambda_N^{\mathrm{Leg}} \sim \mathcal{O}(\sqrt{N}), \qquad \Lambda_N^{\mathrm{Uni}} \sim \mathcal{O}\Big(\frac{2^N}{N \log N}\Big),$$

where Λ_N^{Leg} and Λ_N^{Uni} are LEBESGUE constants for LEGENDRE and uniform nodes distributions correspondingly. In particular one can see that the uniform node distribution is simply disastrous. This phenomenon is another 'avatar' of the so-called RUNGE phenomenon illustrated in Figure 2. The performance of TCHEBYSHEV's nodes is shown in Figure 3.

*Henri LEBESGUE (1875 – 1941), a French mathematician most known for the theory of integration having its name.

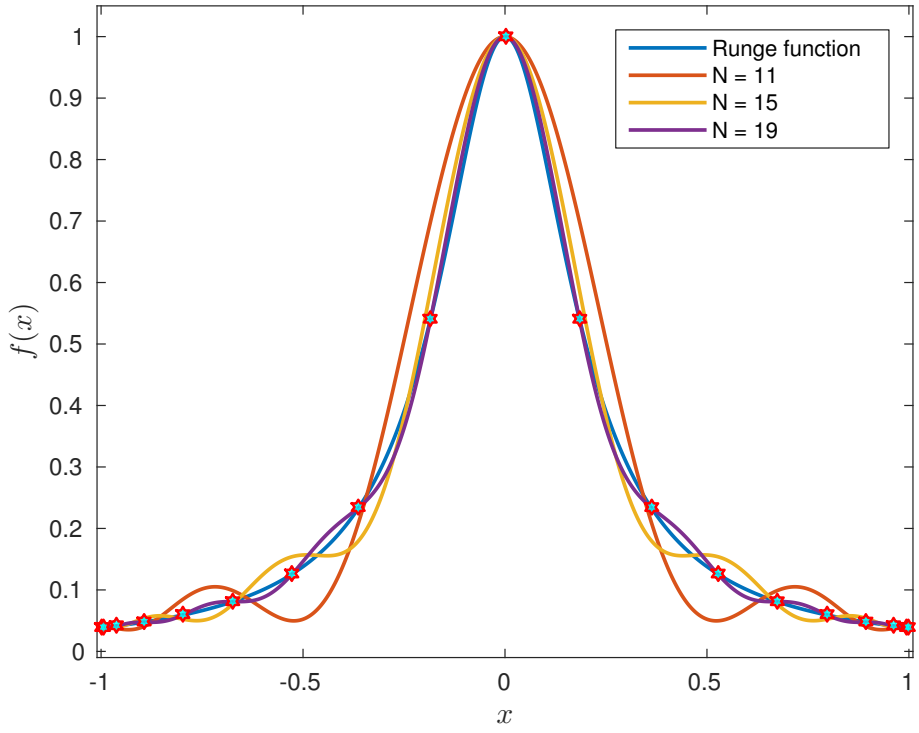


Figure 3. *Interpolation of the* Runge *function* $f(x) = \dfrac{1}{1+25x^2}$ *on a sequence of successively refined of* Tchebyshev*'s grids. One can see that the oscillations present in Figure 2 disappear and the interpolant seems to converge to the interpolated function* $f(x)$ *.*

Remark 3. *As it was shown by* Vértesi *(1990) [39], the* Lebesgue *constant for* Tchebyshev *distribution of nodes is very close for the smallest possible* Lebesgue *constant:*

$$\Lambda_N^{\mathrm{Tch}} = \frac{2}{\pi}\left(\ln N + \gamma + \ln\frac{8}{\pi}\right) + o(1),$$
$$\Lambda_N^{\min} = \frac{2}{\pi}\left(\ln N + \gamma + \ln\frac{4}{\pi}\right) + o(1).$$

where $\gamma \approx 0.5772156649\ldots$ *is the* Euler*–*Mascheroni *constant.*

The Lebesgue constant. It is worth to explain better the important notion of the Lebesgue constant and how it appears in the theory of interpolation (below we follow [34]). Let us take an arbitrary (valid) nodes distribution $\{x_i\}_{i=0}^n$ in domain $\mathcal{U}$, *i.e.*

$$\forall i = 1, 2, \ldots, n : x_i \in \mathcal{U}, \qquad \forall j \neq i : x_i \neq x_j .$$

*Leonhard Euler (1707 – 1783) is a great mathematician who was born in Switzerland and worked all his life in Saint-Petersburg.

†Lorenzo Mascheroni (1750 – 1800) was an Italian mathematician who worked in Pavia.

For any continuous function $u \in C(\mathcal{U})$ there exists a unique interpolating polynomial $\mathcal{P}(x) \in \mathbb{P}_n[\mathbb{R}]$ of degree $n = \deg \mathcal{P}$:

$$\mathcal{P}_n(x_i) = u_i \equiv u(x_i), \qquad i = 1, 2, \ldots, n.$$

This interpolating polynomial $\mathcal{P}_n(x)$ will be denoted as $\mathbb{I}_n[u]$. We introduce also the best possible approximating polynomial $\mathcal{P}^* \in \mathbb{P}_n[\mathbb{R}]$:

$$\| u - \mathcal{P}^* \|_\infty = \inf_{\mathcal{P} \in \mathbb{P}_n[\mathbb{R}]} \| u - \mathcal{P} \|_\infty .$$

It is not obliged that $\mathcal{P}^* \equiv \mathbb{I}_n[u]$. However, since $\mathcal{P}^* \in \mathbb{P}_n[\mathbb{R}]$ we necessarily have $\mathcal{P}^* \equiv \mathbb{I}_n[u]$. Therefore, the following inequalities hold:

$$\begin{aligned}
\| u - \mathbb{I}_n[u] \|_\infty &= \| u - \mathcal{P}^* + \mathcal{P}^* - \mathbb{I}_n[u] \|_\infty \equiv \\
&\| u - \mathcal{P}^* + \mathbb{I}_n[\mathcal{P}^*] - \mathbb{I}_n[u] \|_\infty \leqslant \| u - \mathcal{P}^* \|_\infty + \| \mathbb{I}_n[\mathcal{P}^*] - \mathbb{I}_n[u] \|_\infty \\
&\leqslant \| u - \mathcal{P}^* \|_\infty + \| \mathbb{I}_n \| \cdot \| \mathcal{P}^* - u \|_\infty \leqslant \big(1 + \| \mathbb{I}_n \|\big) \cdot \| u - \mathcal{P}^* \|_\infty .
\end{aligned}$$

The norm of a linear operator $\mathbb{I}_n[\cdot]$ is defined as

$$\| \mathbb{I}_n \| \stackrel{\text{def}}{:=} \sup_{\| u \|_\infty = 1} \| \mathbb{I}_n[u] \|_\infty .$$

The aforementioned norm of the interpolation operator $\mathbb{I}_n[\cdot]$ is called the Lebesgue constant for the set of nodes $\{x_i\}_{i=0}^n$. In multiple dimensions the Lebesgue constant depends also on the shape of domain $\mathcal{U}$ additionally to the nodes distribution $\{x_i\}_{i=0}^n$. As we mentioned above, obvious choices such as the uniform distribution of nodes is disastrous, since it yields the exponential growth of the Lebesgue constant $\Lambda_n^{\text{Uni}} \sim \mathcal{O}\Big(\frac{2^n}{n \log n}\Big)$, where n is the degree of the interpolating polynomial.

The estimation of the Lebesgue constant in multi-dimensional non-Cartesian domains is a problem essentially open nowadays. The most precious information is to find the nodes distribution, for example over a triangle, which minimizes the Lebesgue constant. This knowledge would be crucial for the design of new spectral elements [34]. The locations of nodes which minimize the magnitude of the Lebesgue constant $\Lambda_n \to \min$ for a given polynomial space $\mathbb{P}_n[\mathbb{R}]$ are called Lebesgue points. Nowadays, the best known points on quadrangles are tensor products of Gauss–Lobatto* or Tchebyshev nodes. On triangles the Fekete† points seem currently to be the best choice.

*Rehuel Lobatto is a Dutch Mathematician born in a Portuguese family.

†Michael Fekete (1886 – 1957) is a Hungarian Mathematician who did his PhD under the supervision of Lipót Fejér. M. Fekete gave also some private tutoring to JánosNeumann known today as John von Neumann.

Intermediate conclusions. The remarks above show that TCHEBYSHEV polynomials satisfy the requirement (1) for a useful spectral basis. The requirement (2) is met due to derivative recursion formulas, which can be easily demonstrated (see also Appendix A):

$$T_k'(x) \;=\; 2\,k \sum_{n=0}^{\left[\frac{k-1}{2}\right]} \frac{1}{\delta_{k-1-2\,n}}\, T_{k-1-2\,n}(x)\,, \qquad \delta_k \;=\; \begin{cases} 2\,, & k \;=\; 0\,, \\ 1\,, & k \;\neq\; 0\,. \end{cases}$$

Finally, the requirement (3) is satisfied as well thanks to the Fast Cosine FOURIER Transform (FCFT) (a variant of FFT). It allows to compute spectral coefficients $\{v_k\}_{k=0}^{N}$ from the node values $\left\{u_n(x_i)\right\}_{i=0}^{N}$ and vice versa. Consequently, TCHEBYSHEV polynomials have become an almost universal choice for non-periodic problems. These methods in 1D have been implemented in the MATLAB toolbox `Chebfun`* (and `Chebfun2` in 2D).

Remark 4. *To Author's knowledge,* LEGENDRE *polynomials found some applications in the construction of Spectral Element Method (SEM) bases [34].*

Remark 5. *We have to mention that high-order polynomial interpolants have a bad reputation in the Numerical Analysis (NA) community for the two following reasons:*

(1) Because of the RUNGE *phenomenon*
(2) and because of the following

> **Theorem 1.** *For any node density distribution function there exists a continuous function such that the L_∞-norm of the interpolation error tends to infinity when the number of nodes $N \;\to\; +\infty$.*

However, this bad reputation is not justified. For instance, the RUNGE *phenomenon is completely suppressed by* TCHEBYSHEV *nodes distribution. The low or moderate* LEBESGUE *constant Λ_N ensures that we are not too far from the optimal polynomial.*

Infinite domains. The infinite domains can be handled, first of all, by periodisation in the context of FOURIER-type methods as explained above. However, truly infinite domains can be handled by various approaches:

- HERMITE polynomials
- Sinc functions
- (almost) Rational functions
- Change of variables, *e.g.*

$$x:\; [-\pi,\,\pi] \;\mapsto\; \mathbb{R}\,, \qquad x(q) \;=\; \ell\,\tan\left(\frac{q}{2}\right), \qquad \ell \;\in\; \mathbb{R}^+$$

- Truncation of a domain $(-\infty\,,\infty) \;\rightsquigarrow\; [-\ell,\,\ell]$ followed by a change of variables such as

$$x:\; [-\pi,\,\pi] \;\mapsto\; [-\ell,\,\ell]\,, \qquad x(q) \;=\; \frac{\ell\,q}{\pi}\,.$$

*`Chebfun`'s team is lead by Nick TREFETHEN.

Two last items are followed by a conventional pseudo-spectral discretization on a finite domain $[-\pi, \pi]$.

A HERMITE*-type pseudo-spectral method can be briefly sketched as follows. First of all, we construct recursively the family of HERMITE's polynomials:

$$\mathcal{H}_0(x) = 1, \qquad \mathcal{H}_1(x) = 2x, \qquad \mathcal{H}_{n+1}(x) = 2x\,\mathcal{H}_n(x) - 2n\,\mathcal{H}_{n-1}(x), \qquad n \geqslant 1.$$

Then, we construct HERMITE's functions, which form an *orthonormal* basis on $\mathbb{R}$:

$$\mathscr{H}_n(x) \stackrel{\text{def}}{:=} \frac{1}{\sqrt{2^n \cdot n!}}\,\mathcal{H}_n(x)\,\mathrm{e}^{-\frac{x^2}{2}}.$$

Derivatives of expansions in HERMITE functions can be easily re-expanded using the relation:

$$\mathscr{H}_n'(x) = -\sqrt{\frac{n+1}{2}}\,\mathscr{H}_{n+1}(x) + \sqrt{\frac{n}{2}}\,\mathscr{H}_{n-1}(x).$$

However, the lack of a Fast HERMITE Transform (FHT) implies the use of *differentiation matrices* in numerical implementations. There are also some concerns about a poor convergence rate when the number of modes N is increased. The infinite domains remain very challenging *in silico*.

2.2. Determining expansion coefficients

There are three main techniques to find the spectral expansion coefficients $\{v_k\}_{k=0}^N$. In order to explain them, let us introduce first the residual function on a trial solution $u_n(x, t)$:

$$\mathscr{R}[u_n](x, t) \stackrel{\text{def}}{:=} \Big[\mathscr{L}u_n - g\Big](x, t).$$

Typically, we shall evaluate the residual $\mathscr{R}$ on the expansion (2.2) and the residual norm $\|\mathscr{R}\|$ is generally considered as a measure of the approximate solution quality. The goal is to keep the residual as small as possible across the domain $\mathcal{U}$ (in our particular case $\mathcal{U} = [-1, 1]$).

So, we can mention here at least five approaches to determine the expansion coefficients:

Tau–Lanczos: Spectral coefficients $\{v_k\}_{k=0}^N$ are selected such that the boundary conditions are satisfied identically and the residual $\mathscr{R}[u_n]$ is orthogonal to as many basis functions $\phi_k(x)$ as possible.

Galerkin: First the basis functions are recombined $\{\phi_k(x)\}_{k=0}^N \rightsquigarrow \{\tilde{\varphi}_k(x)\}_{k=0}^N$ so that the boundary conditions are satisfied identically. Then, the coefficients $\{v_k\}_{k=0}^N$ are

*Charles HERMITE (1822 – 1901), a French mathematician whose "*Cours d'analyse*" represent a lot of interest even today. He was also the PhD advisor of another great French mathematician — Henri POINCARÉ.

found so that the residual $\mathscr{R}[u_n]$ be orthogonal to as many of *new* basis functions $\{\tilde{\phi}_k(x)\}_{k=0}^N$ as possible.

Collocation: This approach is similar to the Tau–Lanczos method concerning the boundary conditions: spectral coefficients $\{v_k\}_{k=0}^N$ are selected such that the boundary conditions are satisfied. The rest of coefficients is determined so that the residual $\mathscr{R}[u_n](x, t)$ vanishes at as many (thoroughly chosen) spatial locations as possible.

Petrov–Galerkin: It is a variant of Galerkin method in which the residual $\mathscr{R}[u_n]$ is made orthogonal to a set of functions, which is different from the approximation space basis $\{\phi_k(x)\}_{k=0}^N$.

Least squares: Various least square-type approaches are used when for some reason the number of coefficients to be determined is different from the number of conditions which can be imposed. Below we shall avoid such pathological situations.

The Tau–Lanczos technique was proposed by C. Lanczos* in 1938. What we call the Galerkin† technique was proposed independently first by I. Bubnov‡ and by W. Ritz§ (one more example of the Arnold principle in action!). So, to respect the historical time line, the method should be called Bubnov–Ritz–Galerkin. Today it is the basis of the Finite Element Method (FEM). The collocation technique was called the *pseudo-spectral method* presumably for the first time by S. Orszag¶ in 1972. The Petrov–Galerkin method was proposed by G.I. Petrov‖. It is used up to now in some convection-dominated problems.

Remark 6. *The collocation method can be recast into the* Petrov–Galerkin *framework when we make the residual* $\mathscr{R}[u_n]$ *is made orthogonal to* Dirac** *singular measures* $\{\delta(x - x_k)\}_{k=1}^n$, *where* $\{x_k\}_{k=1}^n$ *are the collocation points.*

*Cornelius Lanczos (1893 – 1974), a Hungarian/American numerical analyst. His books are very pedagogical as well.

†Boris Galerkin (more precisely romanized as Galyorkin) (1871 – 1945), a Russian then Soviet Civil Engineer. The Author suggests to read his biography which is comparable to James Bond (007) movies.

‡Ivan Bubnov (1872 – 1919), a Russian naval engineer.

§Walther Ritz (1878 – 1909), a talented Swiss Physicist, who died young in Göttingen, Germany.

¶Steven Orszag (1943 – 1962), an American numerical analyst, one of the first users of pseudo-spectral methods.

‖Georgii Ivanovich Petrov (1912 – 1987), a Russian fluid mechanician.

**Paul Adrien Maurice Dirac (1902 – 1984) is a British Theoretical Physicist who predicted theoretically using Dirac's equation the existence of the positron. One of the founders of the Quantum Mechanics.

For linear problems all these methods work equally well. However, for nonlinear ones (and in the presence of variable coefficients) the pseudo-spectral (collocation) approach is particularly easy to apply since it involves the products of numbers (solution/variable coefficient's values in collocation points) instead of products of expansions, which are much more difficult to handle.

The convergence of pseudo-spectral approximations for very smooth functions is always geometrical, *i.e.* $\sim \mathcal{O}(q^N)$, where N is the number of modes. This statement is true for any derivative with the same convergence factor $0 < q < 1$. However, the periodic pseudo-spectral method converges always faster than its non-periodic counterpart. This conclusion follows from convergence properties of FOURIER and TCHEBYSHEV series in the complex domain.

The relative resolution ability of various pseudo-spectral methods can be also quantified in terms of the number of points per wavelength needed to resolve a signal. Indeed, this description is more suitable for wave propagation problems. For periodic FOURIER-type methods one needs 2 points per wavelength. For TCHEBYSHEV-type methods one needs about π points. Finally, this number goes up to 6 nodes per wavelength for uniform grids. However, due to the huge LEBESGUE constant Λ_N^{Uni} the uniform grids in pseudo-spectral setting are not usable in practice.

3. Aliasing, interpolation and truncation

Let us take a continuous (and possibly a smooth) function $u(x)$ defined on the interval $\mathcal{I} = (-\pi, \pi)$ and develop it in a FOURIER series. In general it will contain the whole spectrum (*i.e.* the infinite number) of frequencies:

$$u(x) = \sum_{k=-\infty}^{+\infty} v_k \, \mathrm{e}^{\mathrm{i}kx}.$$

Now let us discretize the interval $\mathcal{I}$ with N equispaced collocation points (as we do it in FOURIER-type pseudo-spectral methods. In the following we assume N to be odd, *i.e.* $N = 2m+1$. On this discrete grid all modes $\left\{\mathrm{e}^{\mathrm{i}(k+jN)x}\right\}_{j\in\mathbb{Z}}$ are indistinguishable. See Figure 4 for an illustration of this phenomenon.

The interpolating trigonometric polynomial on a given grid can be written as

$$\mathbb{I}_N[u] = \sum_{k=-m}^{m} \hat{v}_k \, \mathrm{e}^{\mathrm{i}kx}.$$

Each discrete FOURIER coefficient incorporates the contributions of all modes which looks the same on the considered grid:

$$\hat{v}_k = \sum_{j=-\infty}^{+\infty} v_{k+jN}.$$

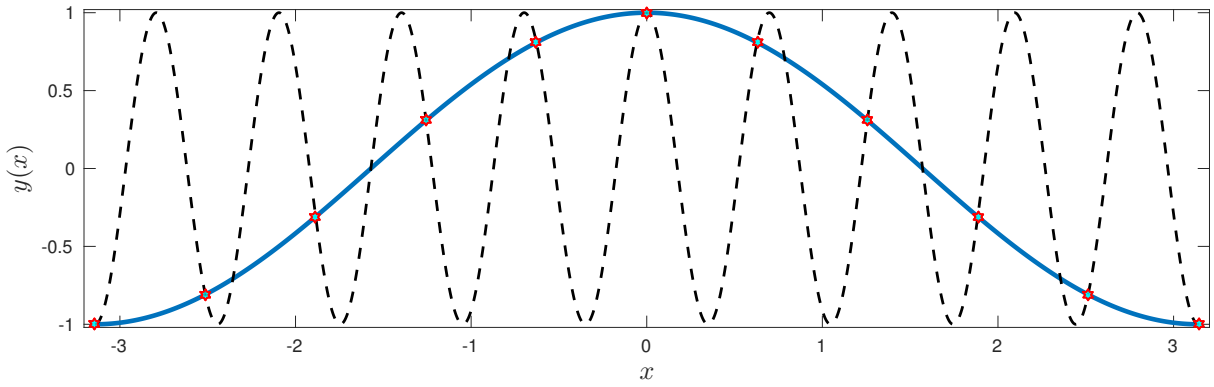


Figure 4. *Illustration of the aliasing phenomenon: two* Fourier *modes are indistinguishable on the discrete grid. The modes represented here are* $\cos(x)$ *and* $\cos(9x)$ *and the discrete grid is composed of* $N = 11$ *equispaced points on the segment* $[-\pi, \pi]$.

Let us recall that the polynomial $\mathbb{I}_N[u]$ takes the prescribed values $\{u(x_k)\}_{k=-m}^{m}$ in the points of the grid $\{x_k\}_{k=-m}^{m}$. This object is fundamentally different from the truncated Fourier series:

$$\mathbb{T}_N[u] \;=\; \sum_{k=-m}^{m} v_k \, \mathrm{e}^{\mathrm{i}kx} \,.$$

The difference between these two quantities is known as the *aliasing error*:

$$\mathscr{R}_N[u] \;\stackrel{\text{def}}{:=}\; \mathbb{I}_N[u] \;-\; \mathbb{T}_N[u] \;=\; \sum_{k=-m}^{m} \sum_{\forall j \neq 0} v_{k+jN} \, \mathrm{e}^{\mathrm{i}kx} \,.$$

After applying the Pythagoras theorem*, we obtain

$$\| \, u \;-\; \mathbb{I}_N[u] \, \|_{L_2}^2 \;=\; \| \, u \;-\; \mathbb{T}_N[u] \, \|_{L_2}^2 \;+\; \| \, \mathscr{R}_N[u] \, \|_{L_2}^2 .$$

Thus, the interpolation error is always larger than the truncation error in the standard L_2 norm. The amount of this difference is precisely equal to the committed *aliasing error*. However, we prefer to use in pseudo-spectral methods the interpolation technique because of the Discrete Fourier Transform (DFT), which allows to transform quickly (thanks to the FFT algorithm) from the set of function values in grid points to the set of its interpolation coefficients. So, it is easier to apply an FFT instead of computing N integrals to determine the Fourier series coefficients.

*We can apply the Pythagoras theorem since the aliasing error $\mathscr{R}_N[u]$ contains the Fourier modes with numbers $|\,k\,| \leqslant m$, while the reminder $u \;-\; \mathbb{T}_N[u]$ contains only the modes with $|\,k\,| > m$. Thus, they are orthogonal.

Nonlinearities. Let us take the simplest possible nonlinearity — the product of two functions $u(x)$ and $v(x)$ defined by their truncated Fourier series containing the modes up to m:

$$u(x) \;=\; \sum_{k=-m}^{m} u_k \, \mathrm{e}^{\mathrm{i}kx}\,, \qquad v(x) \;=\; \sum_{k=-m}^{m} v_k \, \mathrm{e}^{\mathrm{i}kx}\,.$$

The product of these two functions $w(x)$ can be obtained my multiplying the Fourier series:

$$w(x) \;=\; u(x)\cdot v(x) \;=\; \left(\sum_{k=-m}^{m} u_k \, \mathrm{e}^{\mathrm{i}kx}\right)\cdot\left(\sum_{k=-m}^{m} v_k \, \mathrm{e}^{\mathrm{i}kx}\right) \;\equiv\; \sum_{k=-2\,m}^{2\,m} w_k \, \mathrm{e}^{\mathrm{i}kx}\,.$$

It can be clearly seen that the product contains high order harmonics up to $\mathrm{e}^{\pm \mathrm{i}mx}$ which cannot be represented on the initial grid. Thus, they will contribute to the aliasing error explained above.

The aliasing of a nonlinear product can be ingeniously avoided by adopting the so-called $3/2^{\text{th}}$ rule whose Matlab implementation is given below. This function assumes that input vectors are Fourier coefficients of functions $u(x)$ and $v(x)$. The resulting vector contains (anti-aliased) Fourier coefficients of their product $w(x) \;=\; u(x)\cdot v(x)$.

```
function w_hat = AntiAlias(u_hat, v_hat)
  N         = length(u_hat);
  M         = 3*N/2; % 3/2th rule
  u_hat_pad = [u_hat(1:N/2) zeros(1, M-N) u_hat(N/2+1:end)];
  v_hat_pad = [v_hat(1:N/2) zeros(1, M-N) v_hat(N/2+1:end)];
  u_pad     = ifft(u_hat_pad);
  v_pad     = ifft(v_hat_pad);
  w_pad     = u_pad.*v_pad;
  w_pad_hat = fft(w_pad);
  w_hat     = 3/2*[w_pad_hat(1:N/2) w_pad_hat(M-N/2+1:M)];
end % AntiAlias()
```

The main idea behind is to complete vectors of Fourier coefficients by a sufficient number of zeros (*i.e.* the so-called zero padding technique) so that in the physical space the product $u(x)\cdot v(x)$ can be fully resolved. The final step consists in extracting m relevant Fourier coefficients [35, 37].

Remark 7. *To Author's knowledge, the development of efficient and rigorously justified anti-aliasing rules for other types of nonlinearities such as the division, square root,* etc. *is an open problem.*

3.1. Example of a second order boundary value problem

Consider the following second order Boundary Value Problem (BVP) on the interval $\mathcal{I} \;=\; [-1,\,1]$:

$$u_{xx} \;+\; u_x \;-\; 2u \;+\; 2 \;=\; 0\,, \qquad u(-1) \;=\; u(1) \;=\; 0\,. \tag{3.1}$$

This BVP has the exact solution

$$u(x) \;=\; 1 \;-\; \frac{\sinh(2)}{\sinh(3)}\,\mathrm{e}^{x} \;-\; \frac{\sinh(1)}{\sinh(3)}\,\mathrm{e}^{-2x}\,, \qquad x \;\in\; [-1,\,1]\,. \tag{3.2}$$

We shall seek for the numerical solution to (3.1) in the form of a truncated TCHEBYSHEV expansion:

$$u(x) \;\approx\; \sum_{k=0}^{4} a_k\,T_k(x)\,, \qquad \forall k:\; a_k \;\in\; \mathbb{R}$$

Spectral coefficients $\{a_k\}_{k=0}^{4}$ have to be determined. The last expansion is substituted into the governing equation (3.1). There is no reason that (3.1) will be satisfied identically in every point of $\mathcal{I}$. Hence, we can measure the residual

$$\mathscr{R}(x) \;=\; \big(u_{xx} \;+\; u_x \;-\; 2u \;+\; 2\big)(x) \;\rightsquigarrow\; 0\,.$$

The enforcing of boundary conditions $u(-1) \;=\; u(1) \;=\; 0$ leads to two additional relations on spectral coefficients:

$$\begin{aligned} a_0 \;+\; a_1 \;+\; a_2 \;+\; a_3 \;+\; a_4 \;&=\; 0\,,\\ a_0 \;-\; a_1 \;+\; a_2 \;-\; a_3 \;+\; a_4 \;&=\; 0\,. \end{aligned}$$

So, we have two relations coming from boundary conditions and we have five degrees of freedom $\{a_k\}_{k=0}^{4}$. It means that we have to impose three additional conditions to determine uniquely all spectral coefficients. Different approaches prescribe different numerical recipes.

3.1.1 Tau–Lanczos

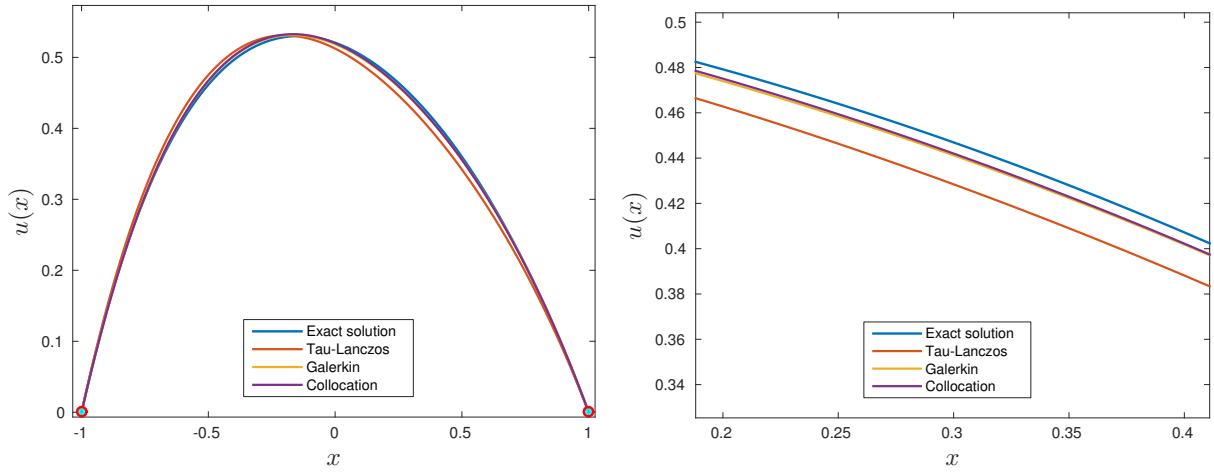
We require that the residual $\mathscr{R}(x)$ be orthogonal to the first three basis functions $T_{0,1,2}(x)$. It gives us three additional relations:

$$\langle \mathscr{R},\,T_k\rangle \;\equiv\; \int_{-1}^{1} \frac{\mathscr{R}(x)\,T_k(x)}{\sqrt{1 \;-\; x^2}}\,\mathrm{d}x \;=\; 0\,, \qquad k \;=\; 0,\,1,\,2\,.$$

3.1.2 Galerkin

We recombine the TCHEBYSHEV polynomials to form a different basis which satisfies identically the boundary conditions:

$$\begin{aligned} \phi_0(x) \;&\stackrel{\text{def}}{:=}\; \big(T_2 \;-\; T_0\big)(x)\,,\\ \phi_1(x) \;&\stackrel{\text{def}}{:=}\; \big(T_3 \;-\; T_1\big)(x)\,,\\ \phi_2(x) \;&\stackrel{\text{def}}{:=}\; \big(T_4 \;-\; T_0\big)(x)\,. \end{aligned}$$



(a) Exact and numerical solutions

(b) Zoom on a portion of domain

Figure 5. *Comparison of various numerical approaches to determine the spectral expansion coefficients for a linear BVP* (3.1).

Then we require that the residual $\mathscr{R}(x)$ is orthogonal to the new basis functions:

$$\langle \mathscr{R},\, \phi_k \rangle \;\equiv\; \int_{-1}^{1} \frac{\mathscr{R}(x)\, \phi_k(x)}{\sqrt{1 \,-\, x^2}}\, \mathrm{d}x \;=\; 0\,, \qquad k \;=\; 0,\, 1,\, 2\,.$$

3.1.3 Collocation

This approach is particularly simple. We require that $\mathscr{R}(x_k) \;\equiv\; 0$ in three interior Tchebyshev points $x_k \;=\; \cos\bigl(\frac{\pi k}{4}\bigr)$, $k = 1,\, 2,\, 3$.

Conclusions. The comparison of three numerical solutions with the exact one (3.2) is shown in Figure 5. In this Figure 5 one can see that the collocation* and Galerkin methods provide nearly identical† numerical solutions which are closer to the exact solution than the prediction of the Tau method. However, in general the Author is impressed by the performance of spectral methods — with only five degrees of freedom $\{a_k\}_{k=0}^{4}$ we capture very well the global behaviour of the solution (3.2) in every point of the interval $\mathcal{I}$. Perhaps, the last point has to be emphasized again: the (pseudo-)spectral methods provide the numerical value of the solution in the *whole domain*, not only in collocation or grid points.

*The advantage of the collocation approach for nonlinear (and variable coefficients) problems becomes even more flagrant.

†With a little advantage towards the collocation method. However, this conclusion is not general at all. It is based only on this particular problem.

4. Application to heat conduction

In this Section we show how to apply the FOURIER-type spectral methods to simulate some simple and not so simple diffusion processes.

4.1. An elementary example

As the simplest example, consider the linear heat equation

$$u_t \;=\; \nu\, u_{xx}\,, \qquad x \;\in\; \mathbb{R}\,, \tag{4.1}$$

completed with the initial condition

$$u(x,\,0) \;=\; u_0(x)\,.$$

We consider this equation on the whole line $\mathbb{R}$ for the sake of simplicity. However, *in silico* it will be truncated to a periodic interval (in the example below it will be $\mathcal{I} \;=\; [-1,\,1]$). Let us apply the FOURIER integral transform to the both sides of equation 4.1:

$$\frac{\mathrm{d}\hat{u}(k,\,t)}{\mathrm{d}t} \;=\; -\nu\, k^2\, \hat{u}(k,\,t)\,, \tag{4.2}$$

where the forward $\hat{u}(k,\,t) \;=\; \mathcal{F}\{u(x,\,t)\}$ and inverse $u(x,\,t) \;=\; \mathcal{F}^{-1}\{\hat{u}(k,\,t)\}$ integral FOURIER transforms are defined below in (4.7) and (4.8) correspondingly. The transformed heat equation (4.2) can be regarded as a linear ODE. Its solution can be readily obtained:

$$\hat{u}(k,\,t) \;=\; \hat{u}_0(k)\,\mathrm{e}^{-\nu\,k^2\,t}\,, \qquad \hat{u}_0(k) \;=\; \mathcal{F}\{u_0(x)\}\,.$$

Consequently, we possess an analytical solution to the heat equation (4.1) in FOURIER space. In order to obtain the solution in physical space, an inverse FOURIER transform has to be computed. The exponential decay of FOURIER coefficients $\hat{u}(k,\,t)$ (except the zero mode $k \;=\; 0$) ensures that the solution $u(x,\,t)$ becomes infinitely smooth C^∞ for $t \;>\; 0$. For example, the initial condition and corresponding solution at $t \;=\; T \;=\; 5\,\mathsf{s}$ are represented in Figure 6. The MATLAB code used to generate Figure 6 is provided below as well.

```
l  = 1.0;     % half-length of the domain
N  = 256;     % number of Fourier modes
dx = 2*l/N;   % distance between two collocation points
x  = (1-N/2:N/2)*dx; % physical space discretization
nu = 0.01;    % diffusion parameter
T  = 5.0;     % time where we compute the solution

dk = pi/l;                  % discretization step in Fourier space
k  = [0:N/2 1-N/2:-1]*dk;   % vector of wavenumbers
k2 = k.^2;                  % almost 2nd derivative in Fourier space

u0     = sech(10.0*x).^2;   % initial condition
```

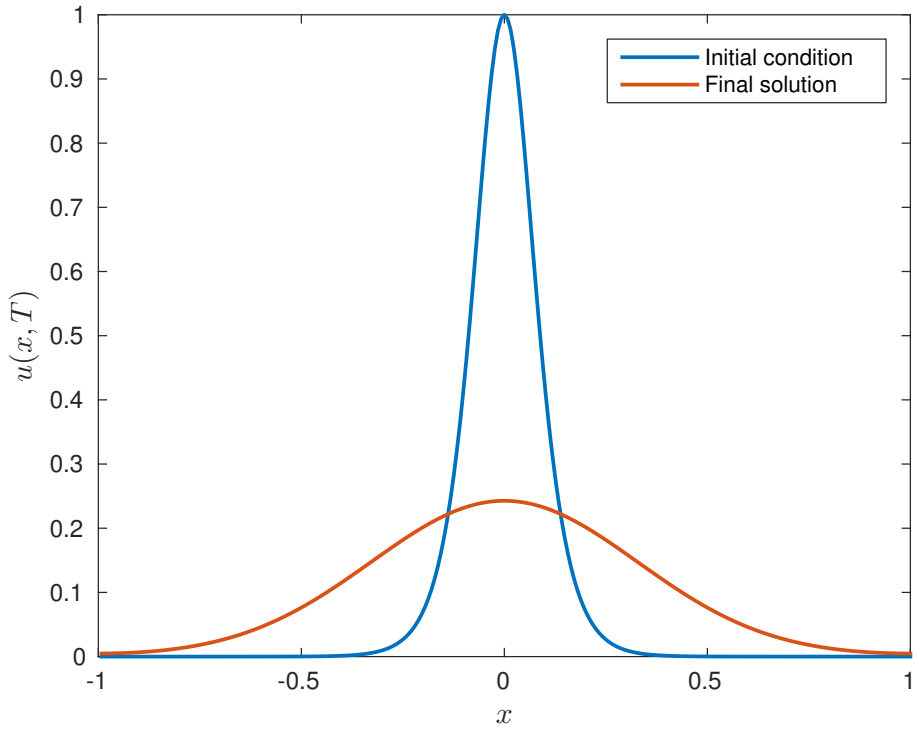


Figure 6. *Spectral solution to the linear heat equation* (4.1) *at* $t = 5.0$, *with* $\nu = 10^{-2}$ *and the initial condition is* $u_0(x) = \operatorname{sech}^2(10\,x)$.

```
13 u0_hat = fft(u0);            % Its Fourier transform
14
15 % and the solution at final time:
16 uT     = real(ifft(exp(-nu*k2*T).*u0_hat));
```

4.2. A less elementary example

The numerical code above is based on the knowledge of the analytical solution to ODE (4.2) in the Fourier space. It is unnecessary to say that an analytical solution is available in very simple situations only. Consequently, we assume that the resulting ODE system in the physical (Fourier) space is formally solved (*i.e.* advanced in time) by the semigroup operator $\mathcal{S}(t)$ ($\hat{\mathcal{S}}(t) \equiv \mathcal{F} \cdot \mathcal{S}(t) \cdot \mathcal{F}^{-1}$ correspondingly):

$$u(x,\,t) \equiv \mathcal{S}(t) \cdot \big[\, u_0(x) \,\big] .$$

In practice this semigroup operator is realized using numerical time-marching techniques described briefly in Appendix F. So, a more general Fourier spectral algorithm is

(1) Decompose the initial condition:

$$\hat{u}_0(k) = \mathcal{F}\big\{u_0(x)\big\} ,$$

(2) Advance in time:

$$\hat{u}_k(t) \;=\; \hat{\mathcal{S}}(t)\big[\,\hat{u}_0(k)\,\big]\,,$$

(3) Synthesize:

$$u(x,\,t) \;=\; \mathcal{F}^{-1}\big\{\hat{u}_k(t)\big\}\,.$$

This algorithm is based on the fact that the following diagram commutes:

$$\begin{array}{ccc} u(x,\,0) & \xrightarrow{\;\mathcal{S}(t)\;} & u(x,\,t) \\ \Big\downarrow{\scriptstyle \mathcal{F}} & & \Big\uparrow{\scriptstyle \mathcal{F}^{-1}} \\ \hat{u}_0(k) & \xrightarrow{\;\hat{\mathcal{S}}(t)\;} & \hat{u}_k(t) \end{array}$$

4.3. A real-life example

In this Section we shall consider a realistic model (to be honest we take a slightly simplified version), which was proposed in [26] to predict heat and moisture transfer through the walls. This model is used in real-world Civil Engineering applications such as the code Domus. So, the model we consider in this Section reads (for simplicity we consider the 1D case):

$$\frac{\partial\theta}{\partial t} \;=\; \frac{\partial}{\partial x}\Big(\mathcal{D}_\theta\,\frac{\partial\theta}{\partial x} \;+\; \mathcal{D}_T\,\frac{\partial T}{\partial x}\Big)\,, \tag{4.3}$$

$$\rho c_m\,\frac{\partial T}{\partial t} \;=\; -\frac{\partial q}{\partial x} \;-\; L(T)\cdot\frac{\partial j_v}{\partial x}\,, \tag{4.4}$$

where θ is moisture volumetric content (*i.e.* moisture density) and T is the temperature. The mass transport coefficients $\mathcal{D}_\theta(\theta,\,T)$ and $\mathcal{D}_T(\theta,\,T)$ may depend nonlinearly on the solution $\big(\theta(x,\,t),\,T(x,\,t)\big)$. Here we assume for simplicity that the mass density ρ and the specific moisture heat c_m are some positive constants*. Finally, the heat flux q and vapor flow j_v can be expressed as

$$\begin{aligned} q \;&=\; -\lambda(\theta,\,T)\,\frac{\partial T}{\partial x}\,, \\ j_v \;&=\; -\mathcal{V}_\theta(\theta,\,T)\,\frac{\partial\theta}{\partial x} \;-\; \mathcal{V}_T(\theta,\,T)\,\frac{\partial T}{\partial x}\,. \end{aligned}$$

In Fourier-type pseudo-spectral methods we usually work with Fourier coefficients. Consequently, we apply the Fourier transform to both sides of equations (4.3), (4.4):

$$\frac{\mathrm{d}\hat{\theta}}{\mathrm{d}t} \;=\; \mathrm{i}k\,\mathcal{F}\Big\{\mathcal{D}_\theta\,\frac{\partial\theta}{\partial x} \;+\; \mathcal{D}_T\,\frac{\partial T}{\partial x}\Big\}\,, \tag{4.5}$$

$$\rho c_m\,\frac{\mathrm{d}\hat{T}}{\mathrm{d}t} \;=\; \mathrm{i}k\,\mathcal{F}\Big\{\lambda(\theta,\,T)\,\frac{\partial T}{\partial x}\Big\} \;-\; \mathcal{F}\Big\{L(T)\cdot\frac{\partial j_v}{\partial x}\Big\}\,, \tag{4.6}$$

*In more realistic modelling $c_m(\theta)$ depends also on the moisture density θ. It does not pose any problems to take it into account in the numerical scheme described in Section 4.

where we introduced some notations for the Fourier transform $\mathcal{F}(\cdot)$:

$$\hat{\theta}(k,\,t) \stackrel{\text{def}}{:=} \mathcal{F}\Big\{\theta(x,\,t)\Big\} \;=\; \int_{-\infty}^{+\infty} \theta(x,\,t)\,\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}x\,, \tag{4.7}$$

where $k \in \mathbb{R}$ is the wave number. Inversely we have

$$\theta(x,\,t) \;=\; \mathcal{F}^{-1}\Big\{\hat{\theta}(k,\,t)\Big\} \;=\; \frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{\theta}(k,\,t)\,\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}k\,. \tag{4.8}$$

Now, system (4.5), (4.6) can be considered as a coupled nonlinear system of Ordinary Differential Equations (ODEs) (not PDEs!) for the Fourier coefficients of solutions $\big(\theta(x,\,t),\,T(x,\,t)\big)$. The numerical methods for systems of ODEs will be explained in a separate course. As general fundamental references on this topic we can recommend [16, 17]. The Author even suggests to employ a well-documented ready-to-use ODE library such as [33], for example. Now we have to explain how to evaluate the right hand side of equations (4.5), (4.6) when only the Fourier coefficients are available. The recipe is very simple:

- All linear operations (*e.g.* additions, subtractions, multiplication by a scalar) can be equally made in Fourier or in a real spaces. The choice has to be done in order to minimize the number of FFT operations

$$\mathcal{F}\big\{\alpha\theta \;+\; \beta T\big\} \;\equiv\; \alpha\hat{\theta} \;+\; \beta\hat{T}$$

- All nonlinear products are made in the real space and then we transfer the result back to the Fourier space using one FFT, *e.g.*

$$\mathcal{F}\big\{\theta(x,\,t)\cdot T(x,\,t)\big\} \;=\; \mathcal{F}\Big\{\mathcal{F}^{-1}\big\{\hat{\theta}(k,\,t)\big\}\cdot\mathcal{F}^{-1}\big\{\hat{T}(k,\,t)\big\}\Big\}$$

- All spatial derivatives are computed only in Fourier space, as

$$\frac{\partial^n}{\partial x^n}\big[\,\cdot\,\big] \quad \leftrightarrows \quad \mathcal{F}^{-1}\Big\{(\mathrm{i}k)^n\,\big[\,\cdot\,\big]\Big\}$$

The Reader can notice that above we used Fourier integrals. It comes from the mathematical tradition. In computer implementations one has to take a finite interval, periodize it and use direct and inverse Discrete Fourier Transforms (DFTs) instead of $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ respectively.

The computation of derivatives using the Fourier collocation spectral method is illustrated below on the following periodic function (with period 2):

$$u(x) \;=\; \sin\big(\pi(x\;+\;1)\big)\,\mathrm{e}^{\sin(\pi(x\;+\;1))}\,, \qquad x\;\in\;[-1,\,1]\,. \tag{4.9}$$

The first three derivatives of this function can be readily computed using any symbolic computation software (the Author used Maple):

$$u'(x) \;=\; \pi\,\cos\big(\pi(x\;+\;1)\big)\Big(1\;+\;\sin\big(\pi(x\;+\;1)\big)\Big)\,\mathrm{e}^{\sin(\pi(x\;+\;1))}\,,$$

$$u''(x) \;=\; \pi^2\Big\{\cos^2\big(\pi(x\;+\;1)\big)\Big(\sin\big(\pi(x\;+\;1)\big)\;+\;3\Big)\;-\;\sin\big(\pi(x\;+\;1)\big)\;-\;1\Big\}\,\mathrm{e}^{\sin(\pi(x\;+\;1))}\,,$$

$$u'''(x) \;=\; \pi^3\,\cos\big(\pi(x+1)\big)\Big\{\cos^2\big(\pi(x+1)\big)\Big(\sin\big(\pi(x+1)\big)+6\Big)-7\,\sin\big(\pi(x+1)\big)-4\Big\}\,.$$

The expressions above are used to assess the accuracy of the computed approximations in collocation points. The analytical derivatives with computed ones (red dots) are shown in Figure 7 for $N = 128$ collocation points. To the graphical accuracy the results are indistinguishable. That is why we computed also the discrete ℓ_∞ norms to compute the relative errors:

$$\varepsilon_N^{(p)} = \frac{\| u_{\text{num}} - u_{\text{exact}} \|_\infty}{\| u_{\text{exact}}(x_i) \|_\infty} .$$

For $N = 32$ FOURIER modes we obtain the following numerical results:

$$\begin{aligned} \varepsilon_{32}^{(1)} &\approx 1.5 \times 10^{-15} , \\ \varepsilon_{32}^{(2)} &\approx 8.4 \times 10^{-15} , \\ \varepsilon_{32}^{(3)} &\approx 4.7 \times 10^{-14} . \end{aligned}$$

The MATLAB code used to generate these results and Figure 7 is provided below.

```
l  = 1.0;     % half-length of the domain
N  = 128;     % number of Fourier modes
dx = 2*l/N;   % distance between two collocation points
x  = (1-N/2:N/2)*dx; % physical space discretization

dk = pi/l;
k  = [0:N/2 1-N/2:-1]*dk; % vector of wavenumbers

arg  = pi*(x + 1);
sar  = sin(arg);
car  = cos(arg); car2 = car.*car;
esa  = exp(sar);
u    = sar.*esa;
uhat = fft(u);

% numerical derivatives:
up   = real(ifft(1i*k.*uhat));
upp  = real(ifft(-k.^2.*uhat));
uppp = real(ifft(-1i*k.^3.*uhat));

% exact derivatives:
pi2  = pi*pi; pi3 = pi2*pi;
u1   = pi*car.*(1 + sar).*esa;
u2   = pi2*(car2.*(sar + 3) - 1 - sar).*esa;
u3   = pi3*car.*(car2.*(sar + 6) - 4 - 7*sar).*esa;
```

Above we exploit the property that the FOURIER transform is a 'change a variables' where the differentiation operator ∂_x becomes diagonal.

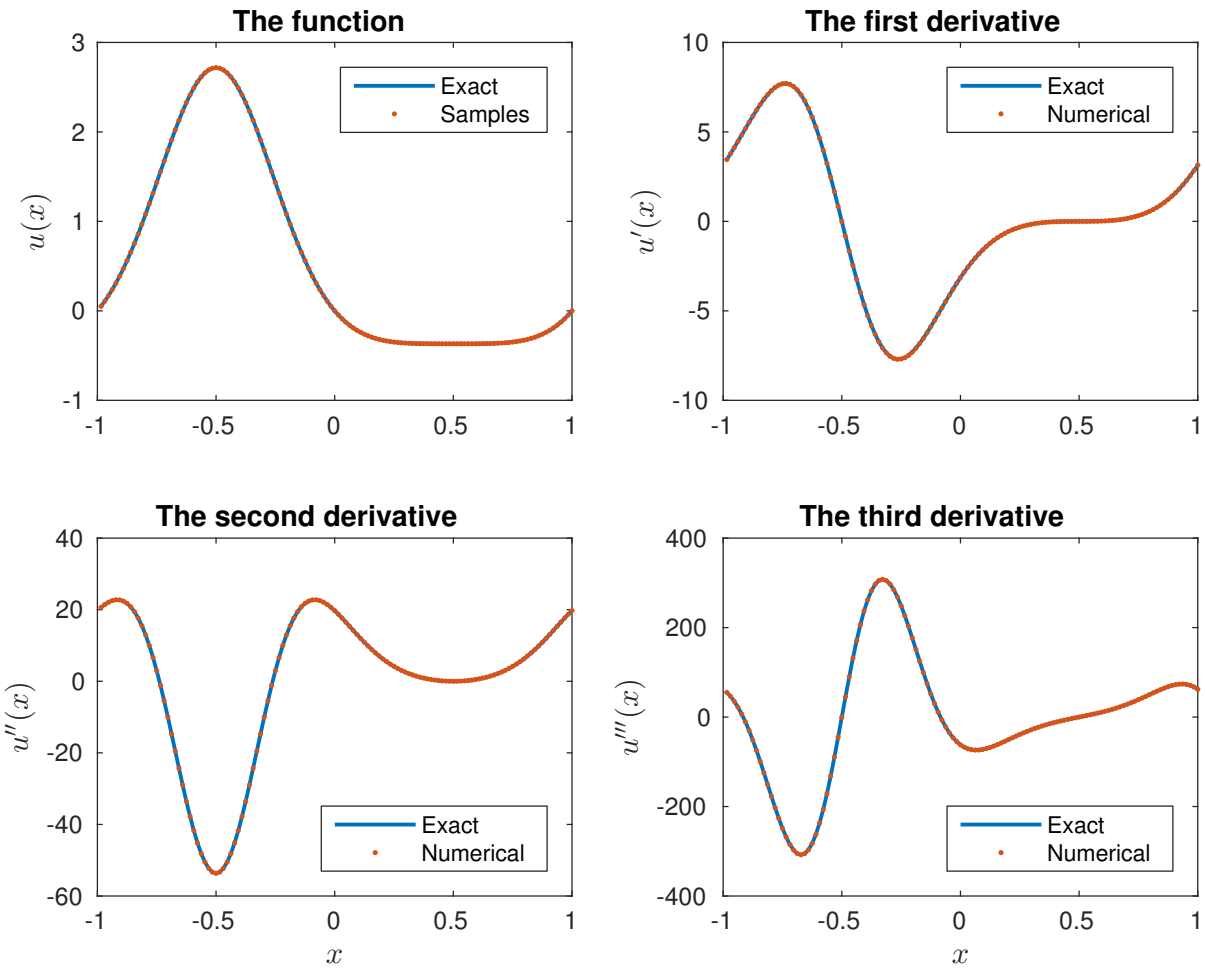


Figure 7. *Numerical differentiation of a periodic function* (4.9) *using the* Fourier *collocation spectral method. We use* 128 *collocation points. The agreement is perfect up to graphic accuracy.*

5. Indications for further reading

If you got interested in the beautiful topic of pseudo-spectral methods, you can find more information in the following books:

- The following book by N. Trefethen* has two major advantages: (*i*) conciseness and (*ii*) collection of Matlab programs which comes along. I would recommend it as the first reading on pseudo-spectral methods. At least you will learn how to program efficiently in Matlab (or in Octave, Scilab, *etc.*) [35]:
  - Trefethen, L. N. (2000). *Spectral methods in MatLab*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

*Lloyd Nick Trefethen (1955 – 20..), an American/British numerical analyst.

- The book by J. Boyd* is probably the most exhaustive one. It covers many topics and applications of spectral methods. The material is presented as a collection of tricks. For example, it is one of seldom books where (semi-)infinite domains and spherical geometries are covered. I am not sure that after reading this book you will know how to program the pseudo-spectral methods, but you will have a broad of view of possible issues and how to address them [2]:
  - Boyd, J. P. (2000). *Chebyshev and Fourier Spectral Methods*. (Dover Publications, New York) (2nd Ed.).

- This book is probably my favourite one. It represents a good balance between the theory and practice and this book arose as an extended version of a previously published review paper. The present lecture notes were inspired in part on this book as well [11]:
  - Fornberg†, B. (1996). *A practical guide to pseudospectral methods*. Cambridge: Cambridge University Press.
- The Author discovered this book only recently. So, he has still to read it, but from the first sight I can already recommend it:
  - Peyret, R. (2002). *Spectral Methods for Incompressible Viscous Flow*. Springer-Verlag New York Inc.
- Finally, I discovered also these very clear and instructive Lecture notes [37]. Lectures were delivered in 2009 at the International Summer School '*Modern Computational Science*' in Oldenburg, Germany. They can be freely downloaded at the following URL address:
  - http://www.staff.uni-oldenburg.de/hannes.uecker/pre/030-mcs-hu.pdf

In general, the Author's suggestions (according to his personal taste and vision) for scientific literature and software are collected in a single document, which is continuously expanded and completed:

https://github.com/dutykh/libs/

Good luck and have a nice reading!

⋆ ⋆ ⋆

*John Boyd (19.. – 20..), an American numerical analyst, meteorologist and occasional science fiction writer.

†Bengt Fornberg (19.. – 20..), a Swedish/American numerical analyst.

A. Some identities involving Tchebyshev polynomials

Tchebyshev polynomials $\{T_n(x)\}_{n=0}^{\infty}$ form a weighted orthogonal system on the segment $[-1, 1]$:

$$\langle T_n, T_m \rangle \equiv \int_{-1}^{1} \frac{T_n(x) \cdot T_m(x)}{\sqrt{1 - x^2}} \, \mathrm{d}x = \begin{cases} 0, & m \neq n, \\ \pi, & m = n = 0, \\ \frac{\pi}{2}, & m = n > 0. \end{cases}$$

The first few Tchebyshev polynomials are

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_2(x) &= 2x^2 - 1, \\ T_3(x) &= 4x^3 - 3x, \\ T_4(x) &= 8x^4 - 8x^2 + 1, \\ T_5(x) &= 16x^5 - 20x^3 + 5x. \end{aligned}$$

They are represented graphically in Figure 8. Higher order Tchebyshev polynomials can be constructed using the three-term recursion relation:

$$T_{n+1}(x) = 2x\,T_n(x) - T_{n-1}(x).$$

The following Matlab code realizes this idea in practice:

```
function P = Chebyshev(n)
    P = cell(n,1);
    if (n == 1)
        P{1} = 1;
        return;
    end % if ()

    P{1} = 1;
    P{2} = [1; 0];
    for j=3:n
        P{j}      = [2*P{j-1}; 0];
        P{j}(3:j) = P{j}(3:j) - P{j-2};
    end % for j

end % Chebyshev()
```

Then, n^{th} Tchebyshev polynomial can be evaluated using the standard Matlab's function `polyval()`, *e.g.*

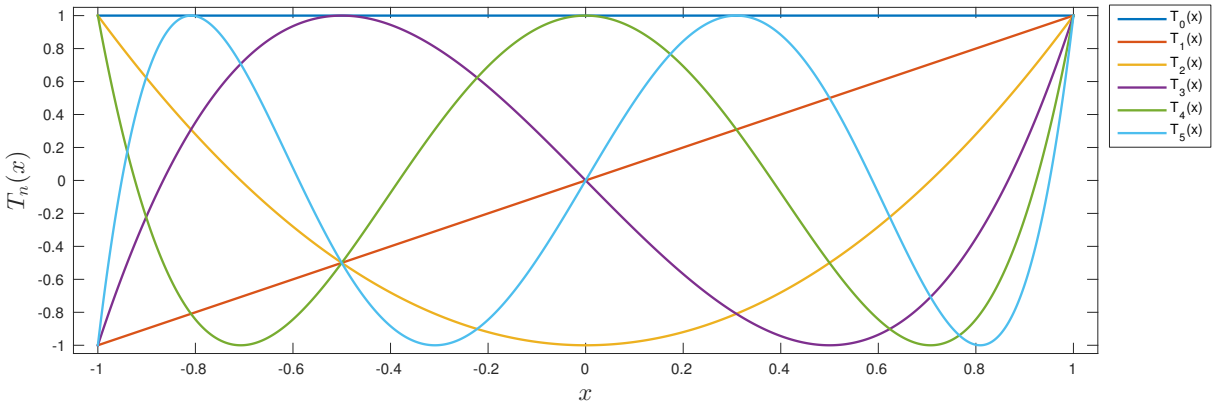


Figure 8. *The first five* TCHEBYSHEV *polynomials. Notice that* $T_n(-1) = (-1)^n$ *and* $T_n(1) \equiv 1$.

```
1 n  = 5;
2 P  = Chebyshev(n+1);
3 x  = linspace(-1, 1, 1000);
4 Tn = polyval(P{n+1}, x);
```

There is an explicit expression for the n^{th} polynomial:

$$T_n(x) = \cos\big(n\theta(x)\big), \qquad \theta(x) \stackrel{\text{def}}{:=} \arccos x. \tag{A.1}$$

The first derivatives of TCHEBYSHEV polynomials can be constructed recursively as well:

$$\frac{T'_{n+1}(x)}{n+1} = \frac{T'_{n-1}(x)}{n-1} + 2\,T_n(x). \tag{A.2}$$

Otherwise, the first derivative $T'_n(x)$ can be found from the following relation:

$$(1-x^2)\,T'_n(x) = -n\,x\,T_n(x) + n\,T_{n-1}(x).$$

TCHEBYSHEV polynomial $T_n(x)$ satisfies the following linear second order differential equation with non-constant coefficients:

$$(1-x^2)\,T''_n - x\,T'_n + n^2\,T_n = 0.$$

Finally, zeros of the n^{th} TCHEBYSHEV polynomial $T_n(x)$ are located at

$$x_k^0 = \cos\left(\frac{2k-1}{2n}\,\pi\right), \qquad k = 1, 2, \ldots, n,$$

and extrema at

$$x_k^{\text{ext}} = \cos\left(\frac{\pi k}{n}\right), \qquad k = 0, 1, \ldots, n.$$

When implementing TCHEBYSHEV-type pseudo-spectral methods for nonlinear problems, one can use the expression of the product of two TCHEBYSHEV polynomials in terms of higher and lower degree polynomials:

$$T_m(x)\,T_n(x) = \frac{1}{2}\left(T_{n+m}(x) + T_{n-m}(x)\right), \qquad \forall n \geqslant m \geqslant 0.$$

To finish this Appendix we give the *generating function* for Tchebyshev polynomials:

$$\frac{1 \;-\; x\,z}{1 \;-\; 2\,x\,z \;+\; z^2} \;=\; \sum_{n=0}^{\infty} T_n(x)\,z^n\,.$$

In order to satisfy the requirement (2) from Section 2.1, we provide here the relations which allow to re-express the derivatives of the Tchebyshev expansion in terms of Tchebyshev polynomials again. Namely, consider a truncated series expansion of a function $u(x)$ in Tchebyshev polynomials:

$$u(x) \;=\; \sum_{k=0}^{N} v_k\,T_k(x)\,.$$

Imagine that we want to compute the first derivative of the expansion above (it is always possible, since the sum is finite and we can differentiate term by term):

$$u'(x) \;=\; \sum_{k=0}^{N} v_k\,\frac{\mathrm{d}T_k(x)}{\mathrm{d}x}\,.$$

Now we re-expand the derivative $u'(x)$ in the same basis functions:

$$u'(x) \;=\; \sum_{k=0}^{N} v'_k\,T_k(x)\,.$$

The main question is how to re-express the coefficients $\{v'_k\}_{k=0}^N$ in terms of coefficients $\{v_k\}_{k=0}^N$? This goal is achieved using basically the recurrence relation (A.2) (even if it does not jump into the eyes). Thanks to it we have an explicit relation between the coefficients:

$$v'_k \;=\; \frac{2}{\delta_k}\;\sum_{\substack{j\,=\,k+1\\ j+k\ \text{odd}}}^{N} j\,v_j\,, \qquad j \;=\; 0,\,1,\,\ldots,\,N-1\,,$$

where

$$\delta_k \;=\; \begin{cases} 2, & k \;=\; 0\,,\\ 0, & k \;\neq\; 0\,. \end{cases}$$

A similar 'trick' can be made for the 2nd derivative as well:

$$u''(x) \;=\; \sum_{k=0}^{N} v_k\,\frac{\mathrm{d}^2T_k(x)}{\mathrm{d}x^2} \;=\; \sum_{k=0}^{N} v''_k\,T_k(x)\,.$$

The connection between coefficients $\{v''_k\}_{k=0}^N$ and $\{v_k\}_{k=0}^N$ is given by the following explicit formula:

$$v''_k \;=\; \frac{1}{\delta_k}\;\sum_{\substack{j=k+2\\ j+k\ \text{even}}}^{N} j\,(j^2 \;-\; k^2)\,v_j\,, \qquad j \;=\; 0,\,1,\,\ldots,\,N-2\,.$$

Finally, in order to construct the spectral coefficients for the n^{th} derivative, one can use the following recurrence relation (also stemming from (A.2)):

$$\delta_{k-1}\, v_{k-1}^{(n)} \;=\; v_{k+1}^{(n)} \;+\; 2\,k\, v_k^{(n-1)}\,, \qquad k \;\geqslant\; 1\,,$$

which has to be completed by starting value $v_N' \;\equiv\; 0\,.$

Remark 8. *Similar identities exist for the family of* Jacobi *polynomials* $\mathcal{P}_n^{\alpha,\beta}(x)$ *as well. However, they are more complicated due to the presence of two arbitrary parameters* α, $\beta > -1$. *The family of* Tchebyshev *polynomials is a particular case of* Jacobi *polynomials when* $\alpha \;=\; \beta \;=\; -\frac{1}{2}$.

A.1. Compositions of Tchebyshev polynomials

We would like to show here another interesting formula involving Tchebyshev polynomials:

Theorem 2 (Composition formula)**.** *If* $T_n(x)$ *and* $T_m(x)$ *are Tchebyshev polynomials* $(m, n \;\geqslant\; 0)$*, then*

$$\big(T_m \circ T_n\big)\,(x) \;\equiv\; T_m\big(T_n\,(x)\big) \;=\; T_{mn}(x)\,. \tag{A.3}$$

Proof. An *eleven* (!) pages combinatorial proof of this result can be found in [1]. Here we shall prove the composition formula (A.3) in one line by using some rudiments of complex variables. Let us introduce the variable $z \overset{\text{def}}{:=} \mathrm{e}^{\mathrm{i}\theta} \;\in \circlearrowleft^*$ such that

$$x \;=\; \frac{1}{2}\left(z \;+\; \frac{1}{z}\right) \quad \leftrightarrows \quad x \;=\; \cos\theta\,.$$

Then, by (A.1) we have a complex representation of the n^{th} Tchebyshev polynomial:

$$T_n\,(x) \;\equiv\; T_n\Big[\tfrac{1}{2}\left(z \;+\; z^{-1}\right)\Big] \;=\; \frac{1}{2}\left(z^n \;+\; \frac{1}{z^n}\right).$$

Now, we have all the elements to show the main result:

$$T_m\left[T_n\Big(\tfrac{1}{2}\left(z \;+\; z^{-1}\right)\Big)\right] \;=\; T_m\Big[\tfrac{1}{2}\left(z^n \;+\; z^{-n}\right)\Big] \;=\; \frac{1}{2}\left(\left(z^n\right)^m \;+\; \frac{1}{\left(z^n\right)^m}\right) \;=$$
$$\frac{1}{2}\left(z^{mn} \;+\; \frac{1}{z^{mn}}\right) \;=\; T_{mn}\Big[\tfrac{1}{2}\left(z \;+\; z^{-1}\right)\Big] \;\equiv\; T_{mn}\,(x)\,.$$

□

B. Trefftz method

In this Appendix I would like to give the flavour of the so-called Trefftz[†] methods, which remain essentially unknown/forgotten nowadays. These methods belong to

*Symbol $\circlearrowleft$ denotes the unit circle $\mathcal{S}_1(\mathbf{0})$ on the complex plane $\mathbb{C}$.

[†]Erich Emmanuel Trefftz (1888 – 1937), a German Mathematician and Mechanical Engineer.

boundary-type solution procedures. Below I shall quote a Professor* from Laboratoire Jacques-Louis Lions (LJLL) at Paris 6 Pierre and Marie Curie University, who makes interesting comments about the knowledge of Trefftz methods in French Applied Mathematics community (the original language and orthography are conserved):

> [...] Sinon j'ai moi aussi fait une petite enquête, à la suite de laquelle il apparaît que personne ne connait Trefftz chez les matheux de Paris 6, hormis Nataf qui en a entendu parler pendant son DEA en méca !!!
>
> Trefftz est un grand oublié car son papier est vraiment extrêmement intéressant, encore plus si tu te rends compte qu'il produit en fait une estimation *a posteriori* (ça doit même être la première). En revanche je me suis aussi persuadé que les méthodes de Trefftz végètent chez les mécano. [...]

So, the situation is rather sad. Fortunately this method continued to live in the Mechanical Engineering community who preserved it for future generations.

In the exposition below we shall follow an excellent review paper [23]. More precisely we describe the so-called *indirect Trefftz method*, which was proposed originally by E. Trefftz (1926) in [36]. There are also *direct Trefftz methods* proposed some sixty years later in [4]. They are much closer to the Boundary Integral Equation Methods (BIEM) and will not be covered here. Interested readers can refer to [23].

Consider a (compact) domain $\Omega \subseteq \mathbb{R}^d$, $d \geqslant 2$ and a (scalar) equation on it (see Figure 9 for an illustration):

$$\mathscr{L} u = 0 . \tag{B.1}$$

You can think, for example, that $\mathscr{L} = -\boldsymbol{\nabla}^2 = -\sum_{i=1}^{d} \dfrac{\partial^2}{\partial x_i^2}$ is the Laplace† operator. Equation (B.1) has to be completed by appropriate boundary conditions:

$$\mathscr{B} u = u^\circ , \qquad \boldsymbol{x} \in \partial\Omega , \tag{B.2}$$

where u° is the boundary data (solution value or its flux) and $\mathscr{B}$ is an operator depending on the type of boundary conditions in use. For example, in the case of Laplace equation the following choices are popular:

Dirichlet: $\mathscr{B} \stackrel{\text{def}}{:=} \mathbb{I}$ (Identity operator, *i.e.* $\mathbb{I}u \equiv u$); in this case the solution value is prescribed on the boundary.

Neumann: $\mathscr{B} \stackrel{\text{def}}{:=} \dfrac{\partial}{\partial n} \equiv \boldsymbol{n} \cdot \boldsymbol{\nabla} = \sum_{i=1}^{d} n_i \cdot \dfrac{\partial}{\partial x_i}$ (Normal derivative, $\boldsymbol{n}$ being the exterior normal to $\partial\Omega$); physically it corresponds to the prescribed flux through the boundary.

*His/Her identity will be hidden in order to avoid any kind of diplomatic incidents.

†Pierre–Simon de Laplace (1749 – 1827) is a French Mathematician whose works were greatly disregarded during his times. The understanding of their importance came much later (il vaut mieux tard que jamais ☺).

Robin: $\mathscr{B} \overset{\text{def}}{:=} \alpha\,\mathbb{I} + \beta\,\dfrac{\partial}{\partial n}$, where $\alpha,\ \beta \in \mathbb{R}$ are some parameters; this case is a mixture of two previous situations. It arises in some problems as well (see *e.g.* [5] for the heat conduction problem in thin liquid films).

Notice, that the boundary $\partial\Omega$ can be divided in some sub-domains where a different boundary condition is imposed:

$$\partial\Omega = \Gamma_1 \cup \Gamma_2 \cup \ldots \cup \Gamma_n \qquad \mu_d\Big(\Gamma_i \cap \Gamma_j\Big) = 0\,, \quad 1 \leqslant i < j \leqslant n\,,$$

where μ_d is the LEBESGUE measure in $\mathbb{R}^d$.

In the indirect TREFFTZ method the numerical solution is sought as a linear combination of $\mathbb{T}$-complete functions $\big\{\phi_k(\boldsymbol{x})\big\}_{k=1}^{n}$ [20], which satisfy exactly the governing equation (B.1):

$$u_n(\boldsymbol{x}) = \sum_{k=1}^{n} v_k\,\phi_k(\boldsymbol{x})\,. \tag{B.3}$$

For example, for the LAPLACE equation in $\mathbb{R}^2$ the $\mathbb{T}$-complete set of functions is [19] $\big\{r^k\,\mathrm{e}^{\mathrm{i}k\theta}\big\}_{k=0}^{\infty}$, where $(r,\,\theta)$ are polar coordinates* on plane. The coefficients $\{v_k\}_{k=1}^{n}$ are to be determined. To achieve this goal the approximate solution (B.3) is substituted into the boundary conditions (B.2) to form the residual:

$$\mathscr{R}[u_n] = \mathscr{B}\,u_n - u^\circ\,.$$

If the residual $\mathscr{R}[u_n]$ is equal identically to zero on the boundary $\partial\Omega$, then we found the exact solution (we are lucky!). Otherwise (we are unlucky and in Numerical Analysis it happens more often), the residual has to be minimized. Very often the boundary operator $\mathscr{B}$ is linear and we can write:

$$\mathscr{R}[u_n] = \sum_{k=1}^{n} v_k\,\mathscr{B}\,\phi_k(\boldsymbol{x}) - u^\circ\,.$$

So, we can apply now the collocation, GALERKIN or least square methods to determine the coefficients $\{v_k\}_{k=1}^{n}$. These procedures were explained above.

Remark 9. *Suppose that a $\mathbb{T}$-complete set of functions for equation (B.1) is unknown. However, there is a way to overcome this difficulty if we know the* GREEN† *function*‡

*$r = \sqrt{x^2 + y^2}$, $\theta = \arctan\dfrac{y}{x}$, with $\arctan(\pm\infty) = \pm\dfrac{\pi}{2}$.

†George GREEN (1793 – 1841), a British Mathematician who made great contributions to the Mathematical Physics and PDEs.

‡The GREEN function is named after George GREEN (see the footnote above), but the notion of this function can be already found first in the works of LAPLACE, then in the works of POISSON. As I said, the works of LAPLACE were essentially disregarded by his "colleagues". POISSON was luckier in this respect. He is also known for (mis)using his administrative resource in order to delay (just for a couple of decades, nothing serious ☺) the publication of his competitors, *e.g.* the young (at that time) CAUCHY.

$\mathcal{G}(\boldsymbol{x};\, Q)$, $\boldsymbol{x} \in \Omega$ *for our equation* (B.1), i.e.

$$\mathscr{L}\,\mathcal{G} \;=\; \delta\big(\boldsymbol{x} \;-\; Q\big)\,, \qquad \boldsymbol{x} \;\in\; \Omega\,, \qquad Q \;\in\; \mathbb{R}^d\,.$$

where $\delta(\boldsymbol{x})$ *is the singular* Dirac *measure. Notice that point* Q *can be inside or outside of domain* Ω*, where the problem is defined. Then, we can seek for an approximate solution* $u_n(\boldsymbol{x})$ *as a linear combination of* Green *functions:*

$$u_n(\boldsymbol{x}) \;=\; \sum_{k=1}^{n} v_k\, \mathcal{G}_k(\boldsymbol{x};\, Q_k)\,,$$

where the points $\big\{Q_k\big\}_{k=1}^n$ *are distributed outside of the computational domain* Ω *in order to avoid the singularities. This method is schematically illustrated in Figure 9. The accuracy of this approach depends naturally on the distribution of points* $\big\{Q_k\big\}_{k=1}^n$*. To Author's knowledge there exist no theoretical indications for the optimal distribution (as it is the case of* Tchebyshev *nodes on a segment). So, it remains mainly an experimental area of the research.*

Conclusions. In the indirect Trefftz methods (historically, the original one) the problem solution is sought as a linear combination of the functions, which satisfy the governing equation identically. Then, the unknown coefficients are chosen so that the approximate solution satisfies the boundary condition(s) as well. It can be done by means of collocation, Galerkin or least square procedures. The main advantage of Trefftz methods is that it allows to reduce problem's dimension by one (*i.e.* 3D $\rightsquigarrow$ 2D and 2D $\rightsquigarrow$ 1D), since the residual $\mathscr{R}[u_n]$ is minimized on the domain boundary $\partial\,\Omega \;\subseteq\; \mathbb{R}^{d-1}$. Readers who are interested in the application of Trefftz methods to their problems should refer to [23] and references therein.

C. A brief history of diffusion in Physics

Since the main focus of the PhD school is set on the diffusion processes (molecular diffusion, heat and moisture conduction through the walls, *etc.*), it is desirable to explain how this research started and why the diffusion is generally modeled by *parabolic PDEs* [10]. The historic part of this Appendix is partially based on [29].

Firs of all, let us make a general, perhaps surprising, remark, which is based on the analysis of historical investigations: understanding the microscopic world is not compulsory to propose a reliable macroscopic law. Indeed, Fourier did not know anything about the nature of heat, Ohm* about the nature of electricity, Fick† about salt solutions and Darcy‡ about the structure of porosity and water therein. In cases of the diffusion, the

*Georg Simon Ohm (1789 – 1854) was a German Physicist.

†Adolf Eugen Fick (1829 – 1901) was a German physician and physiologist.

‡Henry Darcy (1803 - 1858) was a French engineer in Hydraulics.

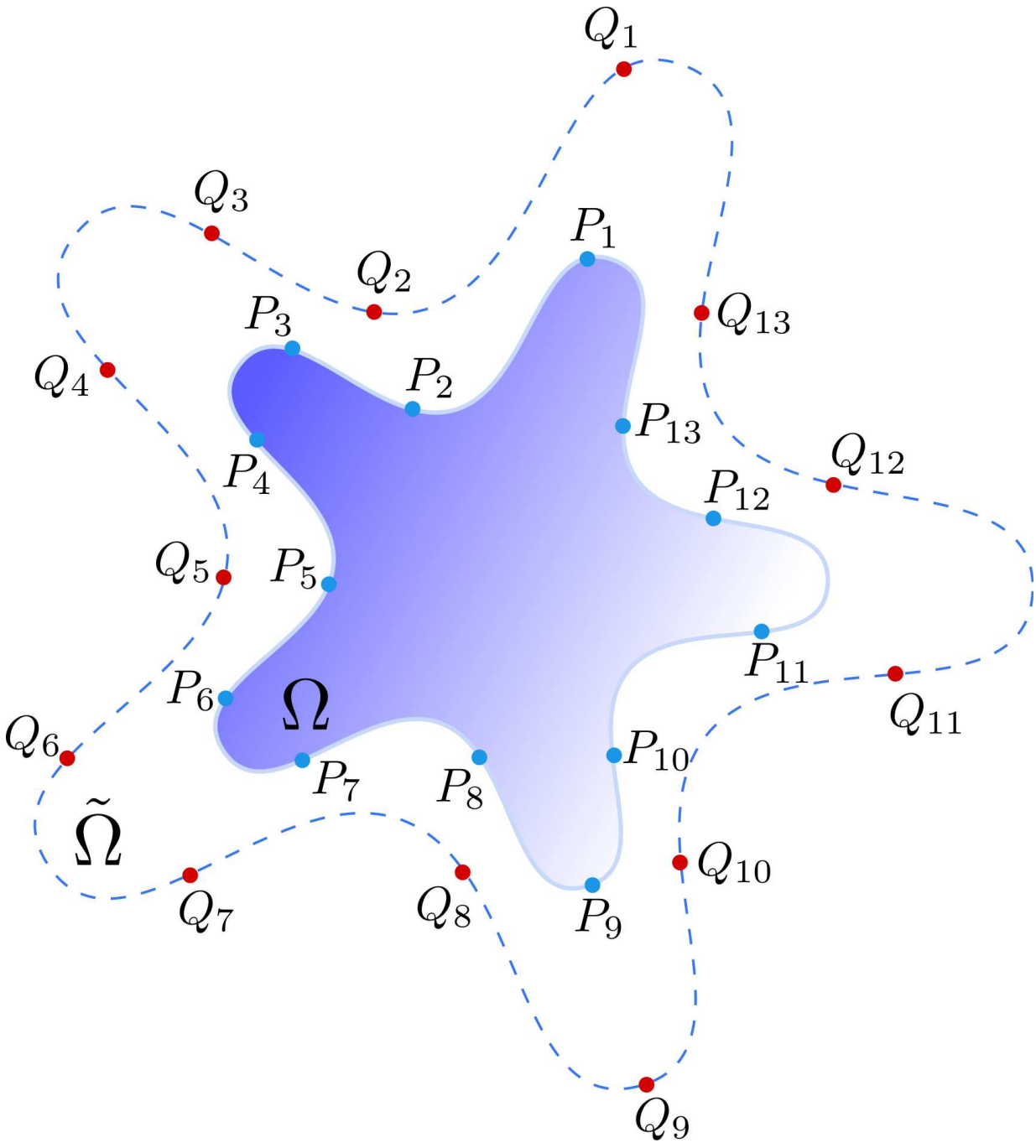


Figure 9. *Collocation Trefftz method using the Green function $\mathcal{G}(P,Q)$: $\{P_i\}_{i=1}^{13} \subseteq \mathbb{R}^d$ are collocation points and $\{Q_i\}_{i=1}^{13} \subseteq \mathbb{R}^d$ are the sources put outside of the domain Ω in order to avoid singularities. Then, the approximate solution is sought as a linear combination of functions $\left\{\mathcal{G}(P,Q_i)\right\}_{i=1}^{13}$. The unknown expansion coefficients are found so that to satisfy the boundary conditions exactly in collocation points $\{P_i\}_{i=1}^{13}$. The outer points $\{Q_i\}_{i=1}^{13}$ are not needed if we know a $\mathbb{T}$-complete set of functions for the operator $\mathscr{L}$.*

bridge between microscopic and macroscopic worlds was built by A. Einstein* in 1905. Namely, he expressed a macroscopic quantity — the diffusion coefficient — in terms of microscopic data (this result will be given below). In general, 1905 was *Annus Mirabilis*† for A. Einstein. During this year he published four papers in *Annalen der Physik* whose aftermath was prodigious.

The first scientific study of diffusion was performed by a Scottish chemist Thomas Graham (1805 – 1869). His research on diffusion was conducted between 1828 and 1833. Here we quote his first paper:

> [...] the experimental information we possess on the subject amounts to little more than the well established fact, that gases of different nature, when brought into contact, do not arrange themselves according to their density, the heaviest undermost, and the lighter uppermost, but they spontaneously diffuse, mutually and equally, through each other, and so remain in the intimate state of mixture for any length of time.

In 1867 James Maxwell‡ estimated the diffusion coefficient of CO_2 in the air using the results (measurements) of Graham. The resulting number was obtained within 5% accuracy. It is pretty impressive.

A. Fick§ hold a chair of physiology in Würzburg for 31 years. His main contributions to Physics were made during a few years around 1855. When he was 26 years old he published a paper on diffusion establishing the now classical Fick's diffusion law. Fick did not realize that dissolution and diffusion processes result from the movement of separate entities of salt and water. He deduced the quantitative law proceeding in analogy with the work of J. Fourier [12] who modeled the heat conduction:

> [...] It was quite natural to suppose that this law for diffusion of a salt in its solvent must be identical with that according to which the diffusion of heat in a conducting body takes place; upon this law Fourier founded his celebrated theory of heat, and it is the same that Ohm applied [...] to the conduction of electricity [...] according to this law, the transfer of salt and water occurring in a unit of time between two elements of space filled with two different solutions of the same salt, must be, *ceteris partibus*, directly proportional to the difference of concentrations, and inversely proportional to the distance of the elements from one another.

*Albert Einstein (1879 – 1955) was a German theoretical Physicist. This personality does not need to be introduced.

†*Annus Mirabilis* comes from Latin and stands for "extraordinary year" in English or "Wunderjahr" in German.

‡James Clerk Maxwell (1831 – 1879) was a Scottish Physicist, famous for Maxwell equations.

§Adolf Fick was the author of the first treatise of *Die medizinische Physik* (Medical Physics) (1856) where he discussed biophysical problems, such as the mixing of air in the lungs, the work of the heart, the heat economy of the human body, the mechanics of muscular contraction, the hydrodynamics of blood circulation, *etc.*

Going along this analogy, FICK assumed that the flux of matter is proportional to its concentration gradient with a proportionality factor κ, which he called "*a constant dependent upon the nature of the substances*". Actually, FICK made an error in (minus) sign which introduces the anti-diffusion and leads to an ill-posed problem. This error was not a source of difficulty for FICK since he analyzed only steady states in accordance with available experimental conditions at that time.

The fundamental article [8] devoted to the study of Brownian* motion and entitled "On the Motion of Small Particles Suspended in a Stationary Liquid, as Required by the Molecular Kinetic Theory of Heat" by EINSTEIN was published on 18th July 1905 in *Annalen der Physik*. By the way, it is the most cited EINSTEIN's paper among four works published during his *Annus Mirabilis*. This manuscript is of capital interest to us as well. Let us quote a paragraph from [8]:

> In this paper it will be shown that, according to the molecular kinetic theory of heat, bodies of a microscopically visible size suspended in liquids must, as a result of thermal molecular motions, perform motions of such magnitudes that they can be easily observed with a microscope. It is possible that the motions to be discussed here are identical with so-called Brownian molecular motion; however, the data available to me on the latter are so imprecise that I could not form a judgment on the question [...]

EINSTEIN was the first to understand that the main quantity of interest is the mean square displacement $\langle X^2(t)\rangle$ and not the average velocity of particles. Taking into account the discontinuous nature of particle trajectories, the velocity is meaningless.

Before publication of [8] atoms in Physics were considered as a useful, but purely theoretical concept. Their reality was seriously debated. W. OSTWALD†, one of the leaders of the anti-atom school, later told A. SOMMERFELD‡ that he had been convinced of the existence of atoms by EINSTEIN's complete explanation of Brownian motion.

Without knowing it, EINSTEIN in answered a question posed the same year by K. PEARSON§ in a Letter published in NATURE [28]:

> Can any of your readers refer me to a work wherein I should find a solution of the following problem, or failing the knowledge of any existing solution provide me with an original one? I should be extremely grateful for aid in the matter.
>
> A man starts from a point O and walks ℓ yards in a straight line; he then turns through any angle whatever and walks another ℓ yards in a second straight line. He repeats this process n times. I require the probability that

*Robert BROWN (1773 – 1858) was a Scottish Botanist.

†Friedrich Wilhelm OSTWALD (1853 – 1932) was a Latvian Chemist. He received a Nobel prize in Chemistry in 1909 for his works on catalysis.

‡Arnold Johannes Wilhelm SOMMERFELD (1868 – 1951) was a German theoretical Physicist. He served as the PhD advisor for several Nobel prize winners.

§Karl PEARSON (1857 – 1936) was an English Statistician.

> after these n stretches he is at a distance between r and $r + \mathrm{d}r$ from his starting point, O.
>
> The problem is one of considerable interest, but I have only succeeded in obtaining an integrated solution for two stretches. I think, however, that a solution ought to be found, if only in the form of a series in powers of $\frac{1}{n}$, when n is large.

The expression "*random walk*" was probably coined in Pearson's Letter [28].

Let us follow Einstein's study of the Brownian motion. Consider a long thin tube filled with water (it allows us to consider only one spatial dimension). At the initial time $t = 0$ we inject a unit amount of ink at $x = 0$. Let $u(x, t)$ denote the density of ink particles in location $x \in \mathbb{R}$ and at time $t \geqslant 0$. So, initially we have

$$u(x, 0) = \delta(x), \tag{C.1}$$

where $\delta(x)$ is the Dirac distribution centered at zero.

Then, suppose that the probability density of the event that an ink particle moves from x to $x + \Delta x$ in a small time Δt is $\rho(\Delta x, \Delta t)$. Then, we have

$$u(x, t + \Delta t) = \int_{-\infty}^{+\infty} u(x - \Delta x, t)\, \rho(\Delta x, \Delta t)\, \mathrm{d}(\Delta x).$$

Assuming that solution $u(x, t)$ is smooth, we can apply the Taylor* formula:

$$u(x, t + \Delta t) = \int_{-\infty}^{+\infty} \Big(u - u_x \cdot \Delta x + \frac{1}{2} u_{xx} \cdot (\Delta x)^2 + \cdots \Big) \rho(\Delta x, \Delta t)\, \mathrm{d}(\Delta x). \tag{C.2}$$

Now let us recall that ρ is a Probability Density Function (PDF). Thus,

Normalisation: $\int_{-\infty}^{+\infty} \rho(\Delta x, \Delta t)\, \mathrm{d}(\Delta x) = 1$

Symmetry: we can assume that the diffusion process is symmetric in space, *i.e.*

$$\rho(\Delta x, \Delta t) = \rho(-\Delta x, \Delta t), \qquad \forall \Delta x,\ \Delta t \geqslant 0.$$

Consequently,

$$\int_{-\infty}^{+\infty} \Delta x\, \rho(\Delta x, \Delta t)\, \mathrm{d}(\Delta x) = 0.$$

Variance: We can also assume that the variance if finite and Einstein assumed additionally that the variance is *linear* in Δt:

$$\int_{-\infty}^{+\infty} (\Delta x)^2\, \rho(\Delta x, \Delta t)\, \mathrm{d}(\Delta x) = \mathcal{D} \Delta t,$$

where $\mathcal{D} > 0$ is the so-called *diffusion constant*.

*Brook Taylor (1685 – 1731) was an English Mathematician.

By incorporating these results into equation (C.2) and rearranging the terms, it becomes

$$\frac{u(x,\, t+\Delta t) \;-\; u(x,\, t)}{\Delta t} \;=\; \frac{1}{2}\,\mathcal{D}\,u_{xx}\,(x,\, t) \;+\; \cdots\,.$$

Taking the limit $\Delta t \;\to\; 0$ we obtain straightforwardly

$$u_t \;=\; \frac{1}{2}\,\mathcal{D}\,u_{xx}\,. \tag{C.3}$$

The Initial Value Problem (C.1) for this linear parabolic equation (C.3) can be solved exactly:

$$u(x,\, t) \;=\; \frac{1}{\sqrt{2\pi\mathcal{D}t}}\,\mathrm{e}^{-\frac{x^2}{2\mathcal{D}t}}\,.$$

The last solution is also known as the Green function.

However, the main result of Einstein's paper [8] is the following formula:

$$\mathcal{D} \;=\; \frac{RT}{fN_a}\,, \tag{C.4}$$

where

R: the ideal gas constant, *i.e.* $R \;\approx\; 8.3144598\;\dfrac{\mathsf{J}}{\mathsf{K}\cdot\mathsf{mol}}$

T: the absolute temperature

f: the friction coefficient

N_a: Avogadro's* number, *i.e.* $N_a \;\approx\; 6.022140857\;\mathsf{mol}^{-1}$

Formula (C.4) along with the observation of the Brownian motion enabled J. Perrin† to produce the first historical estimation of Avogadro's constant $N_a \;\approx\; 6\;\mathsf{mol}^{-1}$.

Exercise 1. *Read a fascinating paper by A.* Einstein *on the Hydrodynamics of tea leaves [9]. By the way, the Author attests that Tea is very stimulating for intellectual activities.* ☺

D. Monte–Carlo approach to the diffusion simulation

From considering the history and physical modelling of Brownian motion we naturally come to the numerical simulation of parabolic PDEs using stochastic processes. It falls

*Lorenzo Romano Amedeo Carlo Avogadro di Quaregna e di Cerreto (1776 – 1856) was an Italian scientist.

†Jean Baptiste Perrin (1870 – 1942) was a French Physicist who was honoured with the Nobel prize in Physics in 1926 for the confirmation of the atomic nature of matter (through the observation of Brownian motion).

into the large class of Monte–Carlo and quasi-Monte–Carlo methods [3]. They are based on the so-called *Law of large numbers*, which can be informally stated* as

$$\lim_{N\to+\infty} \frac{1}{N} \sum_{n=1}^{N} g(\xi_n) \;=\; \mathbb{E}\left[g(\xi)\right],$$

where ξ_n are independent, identically distributed random variables and $g(\cdot)$ is a real-valued continuous function. Here $\mathbb{E}[\cdot]$ stands for the mathematical expectation and the convergence is understood in the sense *convergence in probability distribution*†. The proof is based on the Tchebyshev inequality in Probabilities, but we do not enter into these details. The main advantages of Monte–Carlo methods are

- Simplicity of implementation
- Independent of the problem dimension (they do not suffer of the curse of dimensionality)

On the other hand, the convergence rate is given by the Central Limit Theorem [22] and it is rather slow, *i.e.* $\mathcal{O}(N^{-1/2})$, even if some acceleration is possible thanks to some adaptive procedures [25] (such as low-discrepancy sequences, variance reduction and multi-level methods). The large deviation theory guarantees that the probability of falling out of a fixed tolerance interval decays exponentially fast.

To the Author knowledge, the method we are going to describe below was first proposed by R. Feynman‡ in order to solve *numerically* the linear Schrödinger§ equation in 1940s. In fact, he noticed that the Schrödinger equation can be solved by a kind of average over trajectories. This observation led him to a far-reaching reformulation of the quantum theory in terms of *path integrals*. Upon learning Feynman's ideas, M. Kac¶ (Feynman's colleague at Cornell University) understood that a similar method can work for the heat equation as well. Later it became Feynman–Kac method. Unfortunately, this method is not implemented in any commercial software (for PDEs, the financial industry is using these methods since many decades) and it is not described in classical textbooks on PDEs and in the books on numerical methods for PDEs.

Consider first the following simple heat equation:

$$u_t \;=\; \frac{1}{2}\, u_{xx} \;-\; V(x)\cdot u\,, \tag{D.1}$$

*We choose this particular form, since it is suitable for our exposition.

†A sequence of random variables $\{\xi_n\}_{n=1}^{\infty}$ *converges in probability distribution* towards the random variable ξ if for $\forall\,\varepsilon \;>\; 0$

$$\lim_{n\to\infty} \mathbb{P}\big\{|\,\xi_n \;-\; \xi\,| \;\geqslant\; \varepsilon\big\} \;=\; 0\,.$$

This fact can be denoted as $\xi_n \overset{\mathbb{P}}{\rightsquigarrow} \xi$.

‡Richard Phillips Feynman (1918 – 1988) was an American theoretical Physicist. Nobel Prize in Physics (1965) and Author's hero. Please, do not hesitate to read any of his books!

§Erwin Rudolf Josef Alexander Schrödinger (1887 – 1961) was an Austrian theoretical Physicist. Nobel prize in Physics (1933) for the formulation of what is known now as the Schrödinger equation.

¶Mark Kac (1914 – 1984) was a Polish mathematician who studied in Lviv University, Ukraine and immigrated later to USA.

where $V(x)$ is a function representing the amount of external cooling (if $V(x) \geqslant 0$, and external heating if $V(x) < 0$, but we do not consider this case) at point x. Then, we have the following

Theorem 3 (Feynman–Kac formula)**.** *Let $V(x)$ be a non-negative continuous function and let $u_0(x)$ be bounded and continuous. Suppose that $u(x, t)$ is a bounded function that satisfies equation (D.1) along with the initial condition*

$$u(x, 0) \;=\; u_0(x)\,, \tag{D.2}$$

then,

$$u(x, t) \;=\; \mathbb{E}\left[\exp\Big\{-\int_0^t V\big(\mathcal{W}_s\big)\,\mathrm{d}s\Big\}\,u_0\big(\mathcal{W}_t\big)\right], \tag{D.3}$$

where $\{\mathcal{W}_t\}_{t\geqslant 0}$ is a Brownian motion starting at x.

We have to define rigorously also what the Brownian motion is:

Definition 1 (Brownian motion)**.** *A real-valued stochastic process $t \mapsto \mathcal{W}(t)$ (or a random curve) is called a* Brownian motion *(or* Wiener process*) if*

(1) Almost surely $\mathcal{W}(0) = 0$
(2) $\mathcal{W}(t) - \mathcal{W}(s) \sim \mathcal{N}(0, t - s) \equiv \sqrt{t - s}\,\mathcal{N}(0, 1), \quad \forall t \geqslant s \geqslant 0$
(3) For any instances of time $v > u > t > s \geqslant 0$ the increments $\mathcal{W}(v) - \mathcal{W}(u)$ and $\mathcal{W}(t) - \mathcal{W}(s)$ are independent random variables.

In particular, from point (2) by choosing $s = 0$ it follows directly that

$$\mathbb{E}\big[\mathcal{W}(t)\big] \;=\; 0, \qquad \mathbb{E}\big[\mathcal{W}^2(t)\big] \;=\; t\,.$$

More informally we can say that the Brownian motion is a continuous random curve with the *largest* possible amount of randomness.

The Theorem above can be proved even under more general assumptions, but the formulation given above suffices for most practical situations. For instance, the functions $V(x)$ and $u_0(x)$ may have isolated discontinuities and the Feynman–Kac formula will be still valid. Another interesting corollary of Feynman–Kac's formula is the uniqueness of solutions to the Initial Value Problem (IVP) (D.1):

Corollary 1. *Under the assumptions of Theorem 3, there is at most one solution to the heat equation (D.1), which satisfies the initial condition (D.2). Namely, it is given by* Feynman–Kac *formula (D.3).*

The main (computational) advantage of Feynman–Kac formula is that it can be straightforwardly generalized to the arbitrary dimension:

Theorem 4 (Feynman–Kac formula in d dimensions)**.** *Let $V : \mathbb{R}^d \mapsto [0, +\infty)$ and $u_0 : \mathbb{R}^d \mapsto \mathbb{R}$ be continuous functions with bounded initial condition $u_0(\boldsymbol{x})$. Suppose that $u(\boldsymbol{x}, t)$ is also a bounded function, which satisfies the following Partial Differential Equation (PDE):*

$$u_t \;=\; \frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2 u}{\partial x_i^2} \;-\; V(\boldsymbol{x})\cdot u\,,$$

and the initial condition

$$u(\boldsymbol{x},\,0) \;=\; u_0(\boldsymbol{x})\,.$$

Then,

$$u(\boldsymbol{x},\,t) \;=\; \mathbb{E}\left[\exp\Big\{-\int_0^t V(\boldsymbol{\mathcal{W}}_s)\;\mathrm{d}s\Big\}\;u_0(\boldsymbol{\mathcal{W}}_t)\right],$$

where $\{\boldsymbol{\mathcal{W}}_t\}_{t\geqslant 0}$ *is a* d*-dimensional Brownian motion starting at* $\boldsymbol{x}$*.*

We give here another useful generalization of the Feynman–Kac method. For the sake of simplicity we return to the one-dimensional case. Consider a stochastic process X_t, which satisfies the following *Stochastic Differential Equation* (SDE):

$$\mathrm{d}X_t \;=\; \alpha(X_t)\,\mathrm{d}t \;+\; \sigma(X_t)\,\mathrm{d}\mathcal{W}_t\,, \qquad X_0 \;=\; x\,, \tag{D.4}$$

where $\alpha(x)$ is the local drift and $\sigma(x)$ is called the *local volatility* in financial applications. We assume here that $\alpha(x)$ and $\sigma(x)$ are globally Lipschitz continuous and grow linearly in space at most. Then, the function given by formula (D.3) satisfies the generalized diffusion equation

$$u_t \;=\; \alpha(x)\,u_x \;+\; \frac{1}{2}\,\sigma^2(x)\,u_{xx} \;-\; V(x)\,u(x)\,,$$

together with the initial condition (D.2).

So, the resulting numerical algorithm is very simple:

(1) Generate N trajectories of the Brownian motion $\{\mathcal{W}_s^k\}_{s\,\in\,[0,\,t]}^{1\,\leqslant\,k\,\leqslant\,N}$
(2) Compute N solutions to the SDE (D.4) using the Euler–Maruyama* method [21], for example
(3) Compute the solution value using representation (D.3). The mathematical expectation $\mathbb{E}\,[\cdot]$ is replaced by the simple arithmetic average according to the Monte–Carlo approach.

Remark 10. *Notice that the step (2) can be omitted if we solve the linear parabolic equation without the drift term and with constant diffusion coefficient. The simplest version of the* Feynman–Kac *method for the simplest initial value problem*

$$u_t \;=\; \frac{1}{2}\,u_{xx}\,, \qquad u(x,\,0) \;=\; u_0(x)\,,$$

looks like

$$u(x,\,t) \;\approx\; \frac{1}{N}\sum_{n\,=\,1}^{N} u_0(\xi_n)\,, \qquad \forall n:\;\; \xi_n \;\sim\; \mathscr{N}(x,\,\sqrt{t})\,.$$

It is probably the simplest way to estimate solution value in a given point $(x,\,t)$*.*

The Feynman–Kac method inherits all advantages (and disadvantages) of Monte–Carlo methods. Comparing to grid-based methods it enjoys also the locality property. Namely, if you want to compute your solution in a given point $(x,\,t)$, you do not need to know the solution in neighbouring sites (locations). So, if you want to compute numerically the

*Gisiro Maruyama (1916 – 1986) was a Japanese Mathematician with notable contributions to the theory of stochastic processes.

solution only in a few specific points, please, consider this method. It can be competitive with more conventional approaches. Moreover, the financial industry already appreciated the power of these methods.

D.1. Brownian motion generation

In this Section we provide a simple MATLAB code to generate M sample Brownian motion realizations. It can be used as a building block for SDE solvers and practical implementations of the FEYNMAN–KAC method described above:

```
M  = 100;  % number of paths
N  = 1000; % number of steps
T  = 1;    % final simulation time
dt = T/N;  % time step
dW = sqrt(dt)*randn(M, N);
W  = cumsum(dW, 2);
```

A sample output result (with precisely the same parameters) of this script is depicted in Figure 10.

It is quite straightforward to employ the same MATLAB script to generate Brownian paths in d dimensions (just take $M = m \times d$, where $m \in \mathbb{N}$ and then regroup matrix W in m d-dimensional trajectories). A few realizations of the Brownian motion in two spatial dimensions are depicted in Figure 11.

E. An exact non-periodic solution to the 1D heat equation

In this Appendix we provide an exact solution (see [11, Section §7.2]) to the (dimensionless) heat equation

$$u_t - \frac{1}{9\pi^2} u_{xx} = 0, \qquad x \in [0, 1], \qquad t \in \mathbb{R}^+, \tag{E.1}$$

subject to the initial condition

$$u(x, 0) = 0, \qquad x \in [0, 1],$$

and the following non-periodic and non-homogeneous boundary conditions:

$$u(0, t) = \sin t, \qquad u_x(1, t) = 0, \qquad t \in \mathbb{R}^+.$$

The first boundary condition says that the temperature is prescribed at the left boundary and zero heat flux is imposed on the right boundary. So, the unique solution to the problem

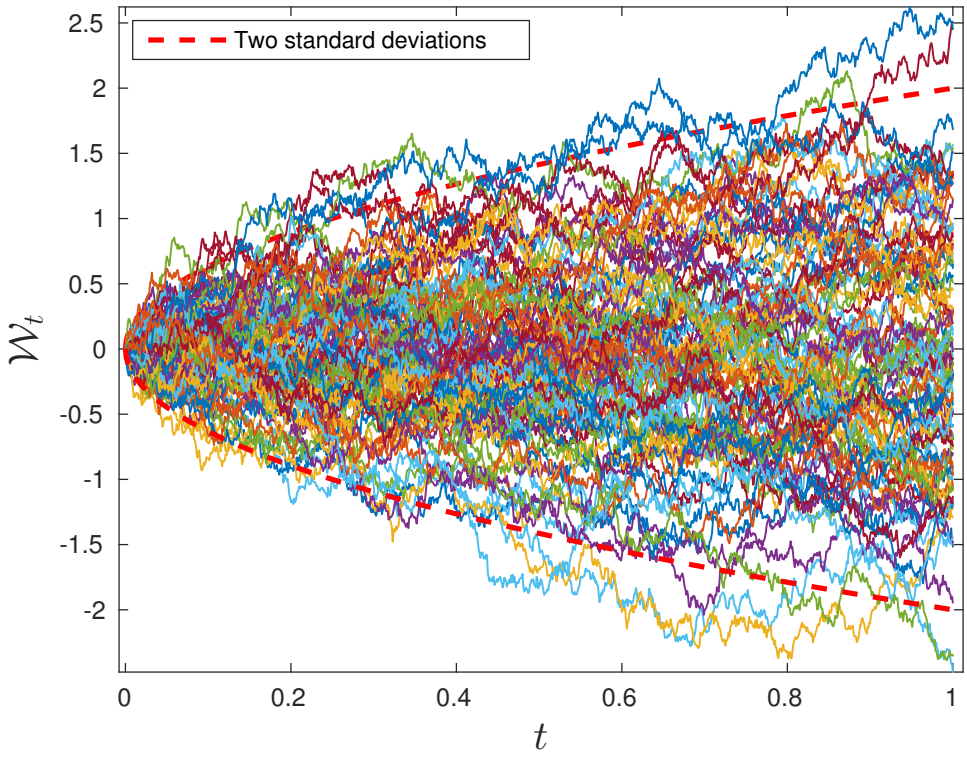


Figure 10. *A collection of sample Brownian paths generated by the code given in Appendix D.1. Two standard deviations should contain about 99% of trajectories. On this picture it looks like it is the case.*

(E.1) described above for $\forall t > 0$ is given by*

$$u(x,\,t) \;=\; \underbrace{\sinh^{-1}\frac{3\pi}{2}\,\Big[\cos\frac{3\pi x}{2}\,\sinh\frac{3\pi(1-x)}{2}\,\sin t \;-\; \sin\frac{3\pi x}{2}\,\cosh\frac{3\pi(1-x)}{2}\,\cos t\Big]}_{u_{\circlearrowleft}(x,\,t)}$$
$$+\;\underbrace{\frac{72}{\pi}\,\sum_{n=1}^{\infty}\frac{(2n-1)\,\mathrm{e}^{-\frac{(2n-1)^2}{18}\,t}}{\big[9\;+\;4(n-2)^2\big]\big[9\;+\;4(n+1)^2\big]}\,\sin(n-\tfrac{1}{2})\pi x}_{u_{\Sigma}(x,\,t)}\,. \qquad \text{(E.2)}$$

This analytical solution can be used as the *reference solution* in order to validate your non-periodic numerical codes. Notice that the second term $u_\Sigma(x,\,t)$ (*i.e.* the infinite series) vanishes uniformly in space, *i.e.*

$$u_\Sigma(x,\,t) \;\rightrightarrows\; 0 \quad \text{as} \quad t \;\to\; +\infty\,.$$

This term u_Σ is needed to enforce the initial condition. Consequently, the long time behaviour of the solution (E.2) is given by $u_{\circlearrowleft}(x,\,t)$.

*This solution can be derived using the classical method of separation of variables [14] [or see an excellent publication of Professor (and my favourite colleague and friend) Marguerite Gisclon [13], if you are not scared by French].

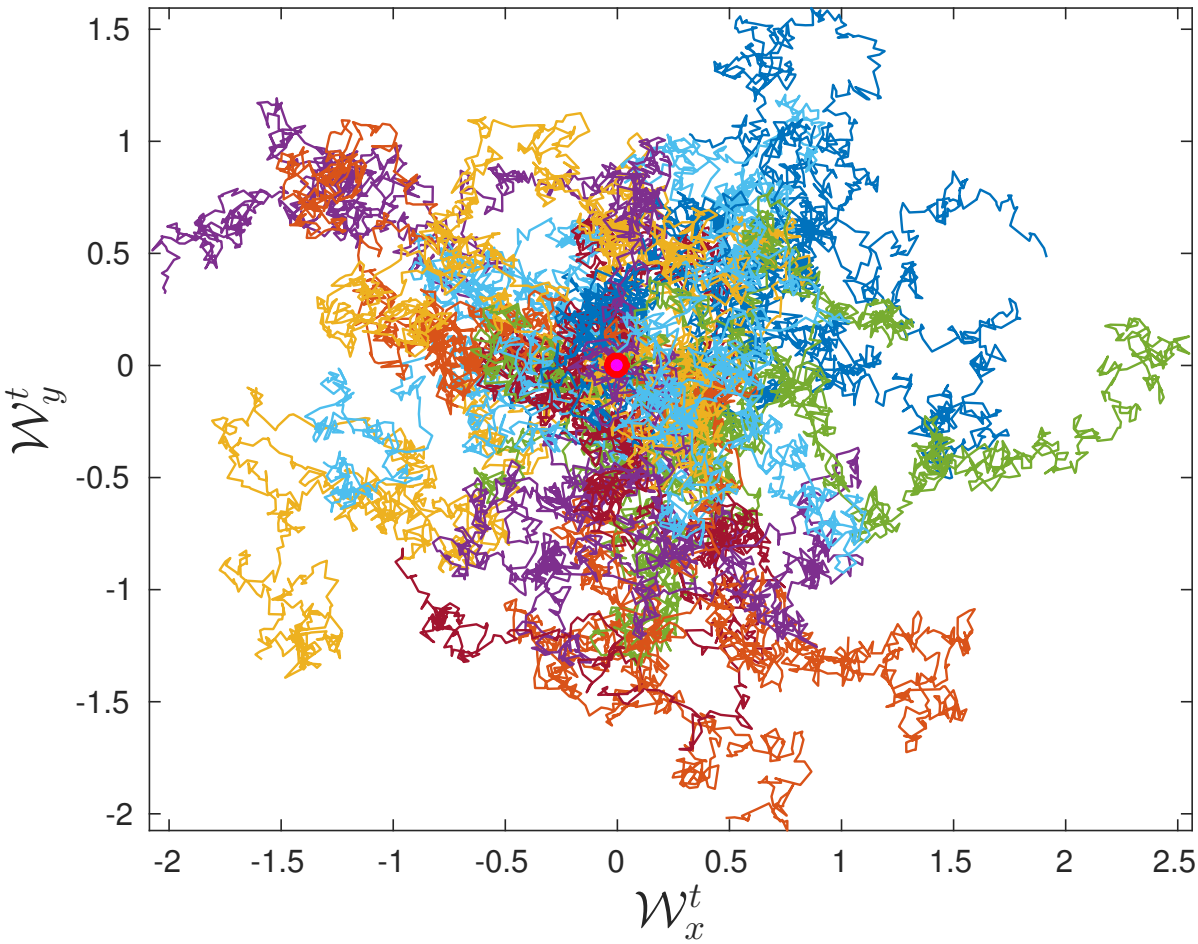


Figure 11. *A few Brownian paths in two spatial dimensions. The red circle depicts the starting point* (0, 0).

Remark 11. *Equation* (E.1) *(along with boundary conditions) models the temperature variation in soil under periodic boundary forcing (modeling seasonal temperature variations). Solution* (E.2) *explains also that for many types of soils, there is a phase shift of seasonal temperature at certain depths (warmest in winter and coolest in summer). See Figure 13 for an illustration. It provides a theoretical explanation why a cave works in practice.*

F. Some popular numerical schemes for ODEs

The discussion of numerical methods for PDEs cannot be complete if we do not provide some basic techniques for the time marching. Basically, if we follow the Method Of Lines (MOL) [24, 30–32], the PDE is discretized first in space, then the resulting system of coupled ODEs has to be solved numerically. This document would not be complete if we did not provide any indications on how to do it. Consider for simplicity an Initial Value

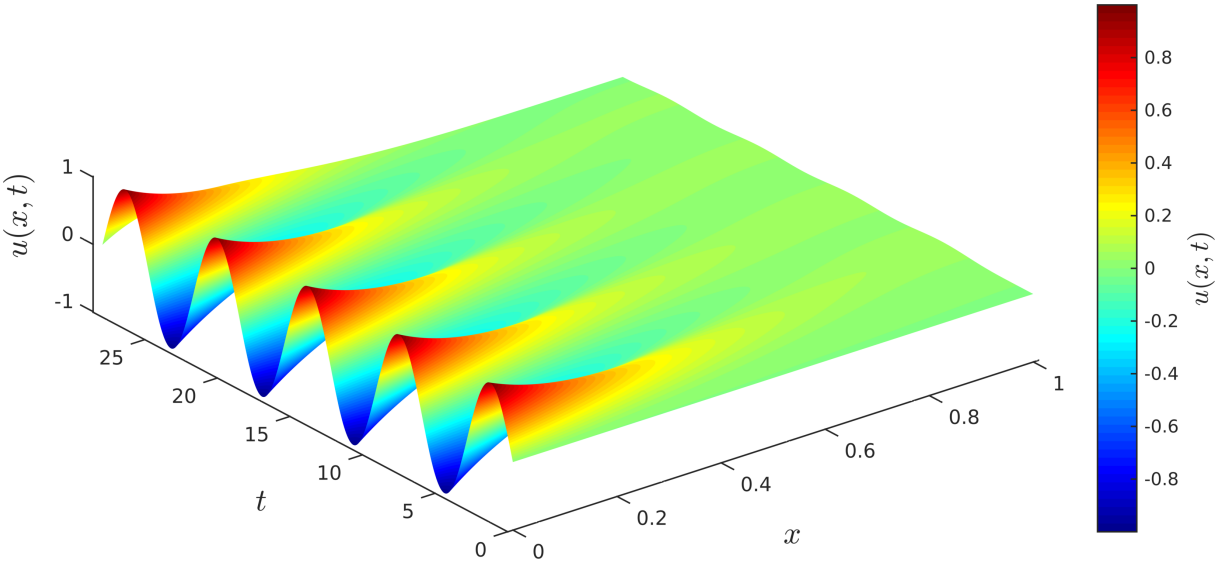


Figure 12. *Space-time plot of the solution* (E.2) *to the linear heat equation* (E.1). *The time window is* $t \in [0, 9\pi]$.

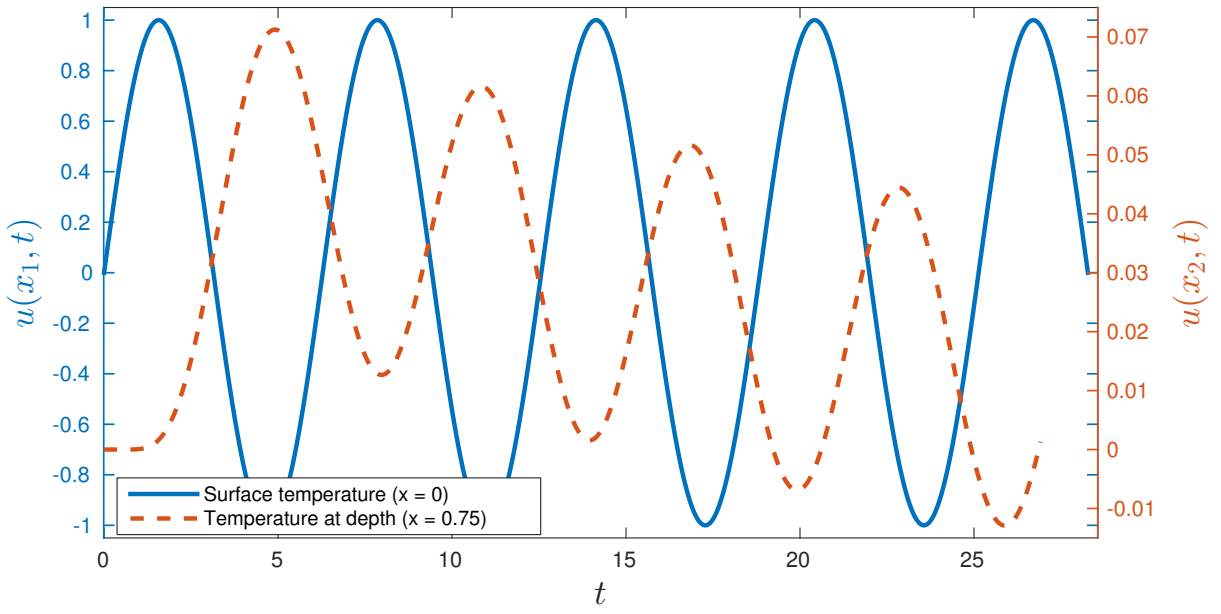


Figure 13. *Solution* (E.2) *shown at the surface* $x = 0$ *and at depth* $x = \frac{3}{4}$. *The goal is to illustrate the phase shift between two curves.*

Problem (IVP), which is sometimes also referred to as the CAUCHY problem:

$$\dot{u} = f(u), \qquad u(0) = u_0.$$

where the dot over a function denotes the derivative with respect to time, *i.e.* $\dot{u} \equiv \dfrac{\mathrm{d}u}{\mathrm{d}t}$. There is a number of time marching schemes proposed in the literature. We refer to [16–18]

as exhaustive references on this topic. Below we provide some most popular (subjectively) schemes.

Forward Euler: (explicit, first order accurate)

$$u_{n+1} \;=\; u_n \;+\; \Delta t\, f(u_n)\,.$$

Backward Euler: (implicit, first order accurate)

$$u_{n+1} \;=\; u_n \;+\; \Delta t\, f(u_{n+1})\,.$$

Adams*–Bashforth-2: (explicit, second order accurate)

$$u_{n+1} \;=\; u_n \;+\; \Delta t \left[\, \frac{3}{2}\, f(u_n) \;-\; \frac{1}{2}\, f(u_{n-1}) \right].$$

Adams–Bashforth†-3: (explicit, third order accurate)

$$u_{n+1} \;=\; u_n \;+\; \Delta t \left[\, \frac{23}{12}\, f(u_n) \;-\; \frac{4}{3}\, f(u_{n-1}) \;+\; \frac{5}{12}\, f(u_{n-2}) \right].$$

Adams–Moulton‡-1 or the trapezoidal rule: (implicit, second order accurate)

$$u_{n+1} \;=\; u_n \;+\; \frac{1}{2}\, \Delta t \left[\, f(u_{n+1}) \;+\; f(u_n) \right].$$

Adams–Moulton-2: (implicit, third order accurate)

$$u_{n+1} \;=\; u_n \;+\; \Delta t \left[\, \frac{5}{12}\, f(u_{n+1}) \;+\; \frac{2}{3}\, f(u_n) \;-\; \frac{1}{12}\, f(u_{n-1}) \right].$$

*John Couch Adams (1819 – 1892), a British Mathematician and Astronomer. He predicted the existence and position of the planet Neptune.

†Francis Bashforth (1819 – 1912) is a British Applied Mathematician working in the field of Ballistics. However, his famous numerical scheme was proposed in collaboration with J. C. Adams to study the drop formation.

‡Forest Ray Moulton (1872 – 1952) is an American Astronomer. There is a crater on the Moon named after him.

*Martin Wilhelm Kutta (1867 – 1944) is a German Mathematician who worked in the field of Fluid Mechanics and Numerical Analysis.

Two-stage Runge–Kutta*: (explicit, second order accurate)

$$\begin{aligned} k_1 &= \Delta t\, f(u_n)\,, \\ k_2 &= \Delta t\, f\bigl(u_n + \alpha\, k_1\bigr)\,, \\ u_{n+1} &= u_n + \Bigl(1 - \frac{1}{2\,\alpha}\Bigr)\, k_1 + \frac{1}{2\,\alpha}\, k_2\,. \end{aligned}$$

- $\alpha = \frac{1}{2}$: mid-point scheme
- $\alpha = 1$: Heun's† method
- $\alpha = \frac{2}{3}$: Ralston's method

Runge–Kutta-4 (RK4): (explicit, fourth order accurate)

$$\begin{aligned} k_1 &= \Delta t\, f(u_n)\,, \\ k_2 &= \Delta t\, f\Bigl(u_n + \frac{1}{2}\, k_1\Bigr)\,, \\ k_3 &= \Delta t\, f\Bigl(u_n + \frac{1}{2}\, k_2\Bigr)\,, \\ k_4 &= \Delta t\, f(u_n + k_3)\,, \\ u_{n+1} &= u_n + \frac{1}{6}\,\bigl[\, k_1 + 2\,k_2 + 2\,k_3 + k_4\,\bigr]\,. \end{aligned}$$

We do not discuss here the questions related to the stability of these schemes. This topic is of uttermost importance but out of scope of these Lecture notes. Moreover, we skip also the adaptive embedded Runge–Kutta schemes which can choose the time step to meet some prescribed accuracy requirements [7].

In order to illustrate the usage of RK4 scheme (sometimes called *the* Runge–Kutta scheme) we take a simple nonlinear logistic equation:

$$\dot{u} = u\cdot(1 - u)\,, \qquad u(0) = 2\,. \tag{F.1}$$

It is not difficult to check that the exact solution to this IVP is

$$u(t) = \frac{2}{2 - \mathrm{e}^{-t}}\,.$$

The Matlab code used to study numerically the convergence of the RK4 scheme on the nonlinear equation (F.1) is given below:

†Karl Heun (1859 – 1929) is a German Mathematician.

```
T   = 2.0; % final simulation time
uex = 2/(2 - exp(-T));
rhs = @(u) u*(1 - u); % RHS of the logistic equation

% list of successfully refined grids:
NN  = [100; 150; 200; 250; 350; 500; 750; 900; 1000; 1250; 1500;
       2000; 2500; 3000; 3500; 4000; 4500; 5000; 5500; 6000];

Err = zeros(size(NN));
for n = 1:length(NN)
    N  = NN(n); % number of time steps
    dt = T/N;   % one time step
    u  = 2.0;   % initial condition on the solution
    for j = 1:N
        k1 = dt*rhs(u);
        k2 = dt*rhs(u + 0.5*k1);
        k3 = dt*rhs(u + 0.5*k2);
        k4 = dt*rhs(u + k3);
        u  = u + (k1 + 2*k2 + 2*k3 + k4)/6;
    end % for j
    Err(n) = abs(u - uex);
end % for n
```

The numerical result is shown in Figure 14. One can clearly observe the 4^{th} order convergence of the RK4 scheme applied to a nonlinear example. The convergence is broken only by the rounding effects inherent to the floating point arithmetics.

F.1. Existence and unicity of solutions

Normally, before attempting to solve a differential equation, one has to be sure that the mathematical problem is well-posed. The mathematical notion of the well-posedness was proposed by J. Hadamard* [15] and it includes three points to be checked:

(1) Existence
(2) Uniqueness
(3) Continuous dependence on the initial condition and other problem parameters

The last point is usually more difficult to be proven theoretically. However, for 3D Navier†–Stokes‡ equations even the global existence and smoothness of solutions (*i.e.*the first point) poses already a Millennium problem:

http://www.claymath.org/millennium-problems/navier-stokes-equation

*Jacques Hadamard (1865 – 1963), a French Mathematician who made seminal contributions to several fields of Mathematics including the Number Theory, Complex Analysis and the theory of PDEs.

†Claude Louis Marie Henri Navier (1785 – 1836) is a French Engineer (Corps des Ponts et Chaussées) and Physicist who made contributions to Mechanics.

‡George Gabriel Stokes (1819 – 1903) was an Irish Physicist and Mathematician.

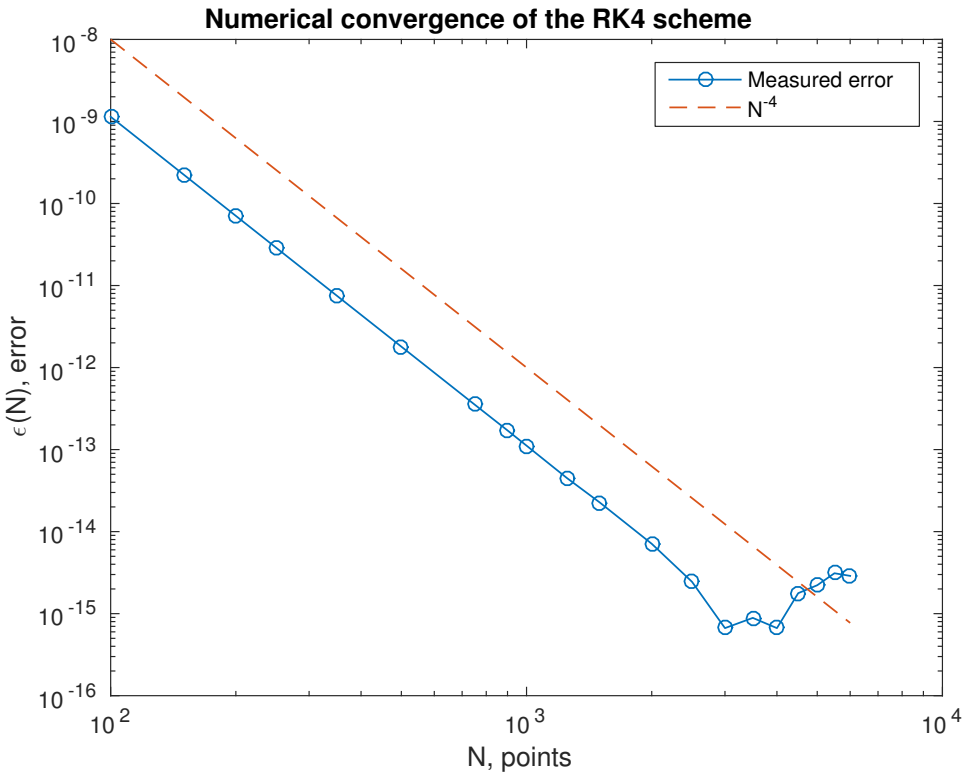


Figure 14. *Numerical convergence of the 4^{th} order* Runge–Kutta *scheme when applied to the logistic equation.*

Below we give some theoretical results which address the well-posedness conditions for an Initial Value Problem (IVP) for a scalar equation:

$$\dot{u} \;=\; f(t,\,u)\,, \qquad u(t_0) \;=\; u_0\,. \tag{F.2}$$

The generalization to systems of Ordinary Differential Equations (ODEs) is straightforward.

Theorem 5 (Existence)**.** *If function $f(t,\,u)$ is continuous and bounded in a domain $(t,\,u) \;\in\; \mathcal{D} \;\subseteq\; \mathbb{R}_t \times \mathbb{R}_u$, then through every point $(t_0,\,u_0) \;\in\; \mathcal{D}$ passes at least one integral curve of equation (F.2).*

Theorem 6 (Prolongation)**.** *Let function $f(t,\,u)$ is defined and continuous in domain $\mathcal{D}$. If a solution $u \;=\; \phi(t)$ to problem (F.2) exists in the interval $t \;\in\; [t_0,\,\alpha)$ cannot be prolongated beyond the point $t \;=\; \alpha$, then it can happen only for one of three following reasons:*

(1) $\alpha \;=\; +\infty\,,$
(2) When $t \;\to\; \alpha \;-\; 0,\; |\,\phi(t)\,| \;\to\; +\infty\,,$
(3) When $t \;\to\; \alpha \;-\; 0$, *distance between the point* $\big(t,\,\phi(t)\big)$ *to the boundary* $\partial\mathcal{D}$ *goes to zero.*

Theorem 7 (Uniqueness, [27]). *If function $f(t,\,u)$ is defined in domain $\mathcal{D}$ and for any pair of points $(t_1,\,u_1)\,,\,(t_2,\,u_2)\;\in\;\mathcal{D}$ satisfies the condition*

$$|\,f(t_1,\,u_1)\;-\;f(t_2,\,u_2)\,|\;\leqslant\;\omega\left(|\,u_1\;-\;u_2\,|\right),\tag{F.3}$$

where $\omega(u)\;>\;0$ is continuous for $0\;<\;u\;\leqslant\;\mathcal{U}$ and

$$\int_{t_0\;+\;\varepsilon}^{\mathcal{U}}\frac{\mathrm{d}u}{\omega(u)}\;\to\;+\infty\,,\qquad as\qquad\varepsilon\;\to\;0\,,$$

then through any point $(t_0,\,u_0)\;\in\;\mathcal{D}$ passes at most one integral curve of equation (F.2).

Suitable functions $\omega(u)$ which satisfy the conditions of the Theorem are

$$\begin{aligned}\omega(u)\;&=\;\mathcal{K}\,u\,,\\ \omega(u)\;&=\;\mathcal{K}\,u\,|\ln u\,|\,,\\ \omega(u)\;&=\;\mathcal{K}\,u\,|\ln u\,|\cdot\ln|\ln u\,|\,,\\ \omega(u)\;&=\;\mathcal{K}\,u\,|\ln u\,|\cdot\ln|\ln u\,|\cdot\ln\ln|\ln u\,|\,,\qquad etc.\end{aligned}$$

Above $\mathcal{K}$ is a positive constant. If we take $\omega(u)\;=\;\mathcal{K}\,u$ then condition (F.3) becomes the well-known Lipschitz* condition for function $f(t,\,u)$ in the second variable u. In order to satisfy the Lipschitz condition, it is sufficient for function $f(t,\,u)$ to be defined in a domain $\mathcal{D}$ convex in u and to have a bounded derivative $\dfrac{\partial f}{\partial u}$ in this domain. Later, Wintner showed that Osgood†'s Theorem conditions are sufficient for the convergence of Picard‡'s iterations to a local solution on a sufficiently small interval [40].

F.1.1 Counterexamples

The Theorem above gives the existence of solutions only *locally* in time $\left[\,t_0,\,t_0\;+\;\delta\,\right)$. The only reason for a solution not to exist beyond the given interval is the *blow-up* (*i.e.* the solution becomes unbounded). Otherwise, the solution exists globally. For instance, the following problem

$$\dot{u}\;=\;1\;+\;u^2\,,\qquad u(0)\;=\;0\,,$$

has the unique exact solution $u(t)\;=\;\tan(t)$ which exists only in the interval $\left[\,0,\,\frac{\pi}{2}\right)$. At time $t\;=\;\frac{\pi}{2}$ the solution blows up.

The uniqueness property is sometimes violated as well. For instance, the scalar equation

$$\dot{u}\;=\;\sqrt{|\,u\,|}\,,\qquad u(0)\;=\;0\,,$$

*Rudolf Otto Sigismund Lipschitz (1832 – 1903) is a German Mathematician who made contributions to Mathematical Analysis. He studied with Gustav Dirichlet at the University of Berlin.

†William Fogg Osgood (1864 – 1943) was an American Mathematician born un Boston, MA. He studied in Universities of Göttingen and Erlangen, Germany. He made contributions to Mathematical and Complex Analysis, Conformal Mappings.

‡Charles Émile Picard (1856 – 1941) was a French Mathematician who made contributions to the Mathematical Analysis. He was married to Marie, a daughter of his Professor Charles Hermite.

has actually infinitely many solutions. Two examples are $u(t) \equiv 0$ and $u(t) = \frac{t^2}{4}$. Obviously, the right-hand side does not satisfy the Lipschitz condition which guarantees the uniqueness.

Acknowledgments

The Author would like to thank Professor Didier Clamond (University of Nice Sophia Antipolis, France) for introducing me to Fourier-type pseudo-spectral methods. I would like also to thank Professor Laurent Gosse (CNR, Italy) who brought Author's attention to Trefftz methods. Special thanks go to my colleague Prof. Paul-Éric Chaudru De Raynal (LAMA, Université Savoie Mont Blanc, France) for his help with Monte–Carlo simulations. Finally, the Author would like to thank Professors Nathan Mendes and Marx Chhay for giving me an opportunity to deliver these lectures in Brazil.

References

[1] A. T. Benjamin and D. Walton. Combinatorially composing Chebyshev polynomials. *Journal of Statistical Planning and Inference*, 140(8):2161–2167, aug 2010. 31

[2] J. P. Boyd. *Chebyshev and Fourier Spectral Methods*. New York, 2nd edition, 2000. 27

[3] R. E. Caflisch. Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*, 7:1–49, 1998. 40

[4] Y. K. Cheung, W. G. Jin, and O. C. Zienkiewicz. Direct solution procedure for solution of harmonic problems using complete, non-singular, Trefftz functions. *Comm. Appl. Num. Meth.*, 5:159–169, 1989. 32

[5] M. Chhay, D. Dutykh, M. Gisclon, and C. Ruyer-Quil. Asymptotic heat transfer model in thin liquid films. *Submitted*, pages 1–24, 2015. 33

[6] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation*, 19(90):297–297, may 1965. 7

[7] J. R. Dormand and P. J. Prince. A family of embedded Runge-Kutta formulae. *J. Comp. Appl. Math.*, 6:19–26, 1980. 48

[8] A. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 322(8):549–560, 1905. 37, 39

[9] A. Einstein. Die Ursache der Mäanderbildung der Flußläufe und des sogenannten Baerschen Gesetzes. *Die Naturwissenschaften*, 14(11):223–224, 1926. 39

[10] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, Providence, Rhode Island, 2 edition, 2010. 34

[11] B. Fornberg. *A practical guide to pseudospectral methods*. Cambridge University Press, Cambridge, 1996. 5, 27, 43

[12] J. Fourier. *Théorie analytique de la chaleur*. Didot, Paris, 1822. 7, 36

[13] M. Gisclon. A propos de l'équation de la chaleur et de l'analyse de Fourier. *Le journal de maths des élèves*, 1(4):190–197, 1998. 44

[14] D. F. Griffiths, J. W. Dold, and D. J. Silvester. Separation of Variables. In *Essential Partial Differential Equations*, pages 129–159. Springer International Publishing, 2015. 44
[15] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52, 1902. 49
[16] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*, volume 31 of *Spring Series in Computational Mathematics*. Springer-Verlag, Berlin, Heidelberg, second edition, 2006. 24, 46
[17] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations: Nonstiff problems*. Springer, 2009. 24
[18] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*. Springer Series in Computational Mathematics, Vol. 14, 1996. 46
[19] I. Herrera. Boundary methods: development of complete systems of solutions. In T. Kawai, editor, *Finite Elements Flow Analysis*, pages 897–906, Tokyo, 1982. University of Tokyo Press. 33
[20] I. Herrera. *Boundary Methods: An Algebraic Theory*. Pitman, 1984. 33
[21] D. J. Higham. An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations. *SIAM Review*, 43(3):525–546, jan 2001. 42
[22] E. T. Jaynes. *Probability Theory*. Cambridge University Press, Cambridge, 2003. 40
[23] E. Kita and N. Kamiya. Trefftz method: an overview. *Advances in Engineering Software*, 24:3–12, 1995. 32, 34
[24] H. O. Kreiss and G. Scherer. Method of lines for hyperbolic equations. *SIAM Journal on Numerical Analysis*, 29:640–646, 1992. 45
[25] B. Lapeyre and J. Lelong. A framework for adaptive Monte-Carlo procedures. *Monte Carlo Methods Appl.*, 17(1):77–98, 2011. 40
[26] N. Mendes and P. C. Philippi. A method for predicting heat and moisture transfer through multilayered walls based on temperature and moisture content gradients. *Int. J. Heat Mass Transfer*, 48(1):37–51, 2005. 23
[27] W. F. Osgood. Beweis der Existenz einer Lösung der Differentialgleichung $$\frac{{dy}}{{dx}} = f\left({x,y} \right)$$ ohne Hinzunahme der Cauchy-Lipschitz'schen Bedingung. *Monatshefte für Mathematik und Physik*, 9(1):331–345, dec 1898. 51
[28] K. Pearson. The Problem of the Random Walk. *Nature*, 72:294, 1905. 37, 38
[29] J. Philibert. One and a half century of diffusion: Fick, Einstein, before and beyond. In J. Kärger, F. Grinberg, and P. Heitjans, editors, *Diffusion Fundamentals*, pages 8–17. Leipzig Universtätsverlag, Leipzig, 2005. 34
[30] S. C. Reddy and L. N. Trefethen. Stability of the method of lines. *Numerische Mathematik*, 62(1):235–267, 1992. 45
[31] W. E. Schiesser. Method of lines solution of the Korteweg-de vries equation. *Computers Mathematics with Applications*, 28(10-12):147–154, 1994.
[32] L. F. Shampine. ODE solvers and the method of lines. *Numerical Methods for Partial Differential Equations*, 10(6):739–755, 1994. 45
[33] L. F. Shampine and M. W. Reichelt. The MATLAB ODE Suite. *SIAM Journal on Scientific Computing*, 18:1–22, 1997. 24
[34] P. Solin. *Partial Differential Equations and the Finite Element Method*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2005. 11, 12, 13
[35] L. N. Trefethen. *Spectral methods in MatLab*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000. 18, 26

[36] E. Trefftz. Gegenstück zum ritzschen Verfahren. In *Proc. 2nd Int. Cong. Appl. Mech.*, pages 131–137, Zürich, 1926. 32

[37] H. Uecker. A short ad hoc introduction to spectral methods for parabolic PDE and the Navier-Stokes equations. Technical report, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany, 2009. 18, 27

[38] B. van Leer. Upwind and High-Resolution Methods for Compressible Flow: From Donor Cell to Residual-Distribution Schemes. *Commun. Comput. Phys.*, 1:192–206, 2006. 6

[39] P. Vértesi. Optimal Lebesgue constant for Lagrange interpolation. *SIAM J. Numer. Anal.*, 27:1322–1331, 1990. 11

[40] A. Wintner. On the Convergence of Successive Approximations. *American Journal of Mathematics*, 68(1):13, jan 1946. 51

LAMA, UMR 5127 CNRS, Université Savoie Mont Blanc, Campus Scientifique, 73376 Le Bourget-du-Lac Cedex, France

E-mail address: Denys.Dutykh@univ-savoie.fr

URL: http://www.denys-dutykh.com/