MATH0024 – Modeling with PDEs

Heat equation

Maarten Arnst and Romain Boman

November 15, 2017

Outline

- Heat equation:
  - Heat operator.
  - Physical examples.
- Heat equation on all of space:
  - Fundamental solution.
  - Superposition formulae.
  - Properties of solutions.
- Heat equation on a portion of space:
  - Separation of variables.
- Summary and conclusion.
- References.

Heat equation

Heat operator

$$\frac{\partial}{\partial t} - \triangle_{\boldsymbol{x}} = \frac{\partial}{\partial t} - \text{div}_{\boldsymbol{x}} \boldsymbol{\nabla}_{\boldsymbol{x}}.$$

Heat equation

$$\frac{\partial u}{\partial t} - \triangle_{\boldsymbol{x}} u = f.$$

Areas of application in mechanics and physics

- The **heat equation** is the **mathematical prototype** of models of **diffusive phenomena**, such as heat conduction, diffusive transport of substances (pollutants, solutes,. . .),. . .
- Of course, we know that there exist also convective and radiative heat transfer and convective transport of substances. However, mathematicians use different mathematical prototypes to study equations that mechanical and physical engineers use to model convection and wave propagation.

Unsteady heat conduction with heat source

- Let us consider a closed system that consists of a fixed amount of mass that occupies an open bounded set $\Omega(t)$ of $\mathbb{R}^3$ with a sufficiently smooth boundary $\partial\Omega(t)$, which depends on time t.

- For such a closed system, **conservation of energy** reads as follows:

$$\frac{d}{dt}\int_{\Omega(t)} \rho\left(\frac{1}{2}\|\boldsymbol{v}\|^2 + e\right) dV = \int_{\Omega(t)} (\boldsymbol{f}_{\mathsf{v}} \cdot \boldsymbol{v} + r)\, dV + \int_{\partial\Omega(t)} \big(\boldsymbol{\sigma}(\boldsymbol{n}) \cdot \boldsymbol{v} - \boldsymbol{q} \cdot \boldsymbol{n}\big) dS.$$

  Here, ρ is the mass density ($[\rho] = \mathrm{kg\,m^{-3}}$), $\boldsymbol{v}$ the velocity($[\boldsymbol{v}] = \mathrm{m\,s^{-1}}$), e the energy density ($[e] = \mathrm{J\,kg^{-1}}$), $\boldsymbol{f}_{\mathsf{v}}$ the volume force ($[\boldsymbol{f}_{\mathsf{v}}] = \mathrm{N\,m^{-3}}$), r the heat source ($[r] = \mathrm{J\,m^{-3}\,s^{-1}}$), $\boldsymbol{\sigma}$ the stress tensor $[\boldsymbol{\sigma}] = \mathrm{N\,m^{-2}}$), and $\boldsymbol{q}$ the heat flux ($[\boldsymbol{q}] = \mathrm{J\,m^{-2}\,s^{-1}}$).

- Differentiating under the integral sign and using conservation of mass $\frac{\partial\rho}{\partial t} + \mathrm{div}_{\boldsymbol{x}}(\rho\boldsymbol{v}) = 0$, conservation of momentum $\mathbf{div}_{\boldsymbol{x}}\boldsymbol{\sigma} + \boldsymbol{f}_{\mathsf{v}} = \rho\frac{d\boldsymbol{v}}{dt}$, and Stokes's theorem, we obtain

$$\int_{\Omega(t)} \rho\frac{de}{dt} dV = \int_{\Omega(t)} r dV + \int_{\Omega(t)} \big(\mathrm{tr}(\boldsymbol{\sigma}\boldsymbol{D}) - \mathrm{div}_{\boldsymbol{x}}\boldsymbol{q}\big) dS.$$

  Here, $\frac{de}{dt} = \frac{\partial e}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla}_{\boldsymbol{x}} e$ is the material derivative of the energy density, $\frac{d\boldsymbol{v}}{dt} = \frac{\partial\boldsymbol{v}}{\partial t} + \mathbf{D}_{\boldsymbol{x}}\boldsymbol{v}(\boldsymbol{v})$ the acceleration, and $\boldsymbol{D} = \frac{1}{2}(\mathbf{D}_{\boldsymbol{x}}\boldsymbol{v} + \mathbf{D}_{\boldsymbol{x}}\boldsymbol{v}^{\mathrm{T}})$ the rate of deformation tensor.

Unsteady heat conduction with heat source (continued)

- Thus, locally, conservation of energy is expressed by the following PDE:

$$\rho \frac{de}{dt} = r + \mathrm{tr}(\boldsymbol{\sigma} \boldsymbol{D}) - \mathrm{div}_{\boldsymbol{x}} \boldsymbol{q}.$$

- **Fourier's law** indicates that $\boldsymbol{q}$ is proportional to the gradient $\boldsymbol{\nabla}_{\boldsymbol{x}} T$ of temperature T but points oppositely because heat conduction is from regions of higher to regions of lower temperature:

$$\boldsymbol{q} = -k \boldsymbol{\nabla}_{\boldsymbol{x}} T, \qquad k\text{: thermal conductivity } \left([k] = \mathrm{J\,m^{-1}\,K^{-1}\,s^{-1}}\right).$$

- For small temperature changes, energy density depends on temperature as follows:

$$e = cT, \qquad c\text{: heat capacity } \left([c] = \mathrm{J\,kg^{-1}\,K^{-1}}\right).$$

- Combining the aforementioned results, we obtain

$$\boxed{\rho\, c \frac{dT}{dt} - \mathrm{tr}(\boldsymbol{\sigma} \boldsymbol{D}) - \mathrm{div}_{\boldsymbol{x}} (k \boldsymbol{\nabla}_{\boldsymbol{x}} T) = r.}$$

- If $\boldsymbol{v} = \boldsymbol{0}$, as is the case, for example, for heat conduction through a rigid material, we obtain

$$\boxed{\rho\, c \frac{\partial T}{\partial t} - \mathrm{div}_{\boldsymbol{x}} (k \boldsymbol{\nabla}_{\boldsymbol{x}} T) = r.}$$

Heat equation on all of space

Notion of fundamental solution for heat equation

- A **fundamental solution** K for the heat equation is a solution that solves the heat equation for a Dirac impulse $\delta_{(\mathbf{0},0)}$ centered at $(\boldsymbol{x}, t) = (\mathbf{0}, 0)$ on the right-hand side,

$$\frac{\partial K}{\partial t}(\boldsymbol{x}, t) - \triangle_{\boldsymbol{x}} K(\boldsymbol{x}, t) = \delta_{(\mathbf{0},0)}(\boldsymbol{x}, t), \qquad (\boldsymbol{x}, t) \in \mathbb{R}^m \times \mathbb{R}.$$

- As in the case of the definition of a fundamental solution for the Laplace/Poisson equation in Lecture 2, the theory of distributions must be used here to define mathematically fully rigorously the derivatives and the Dirac impulse involved in this equation.

Gaussian kernel

- It can be shown that the **Gaussian kernel**, also called **Gauss-Weierstrass kernel** or **heat kernel**,

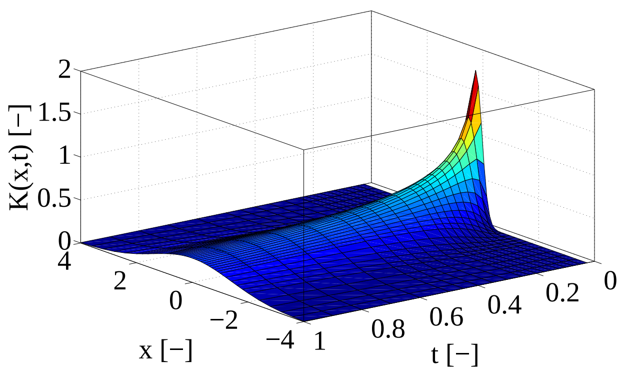
$$K(\boldsymbol{x}, t) = K_t(\boldsymbol{x}) = \begin{cases} 0 & \text{for } t < 0, \\ (4\pi t)^{-m/2} \exp\left(-\dfrac{\|\boldsymbol{x}\|^2}{4t}\right) & \text{for } t > 0, \end{cases}$$

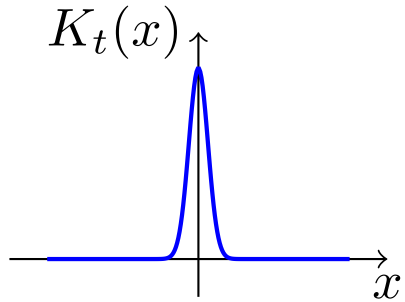
is a fundamental solution for the heat equation.

- The Gaussian kernel K is singular at the origin $(\boldsymbol{x}, t) = (\mathbf{0}, 0)$. Thus, as in the case of the fundamental solutions that we considered in Lecture 2 for the Laplace/Poisson equation, this fundamental solution for the heat equation is a **generalized solution**.

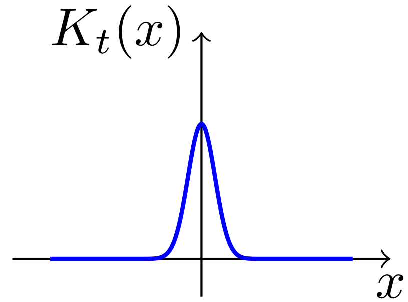
Gaussian kernel in 1D ($m = 1$)

- The following figures show the Gaussian kernel in 1D ($m = 1$):

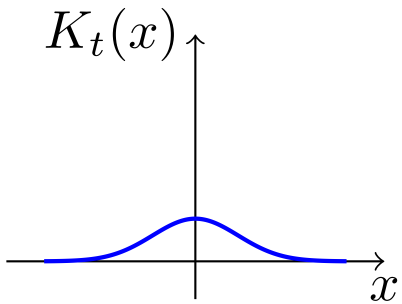




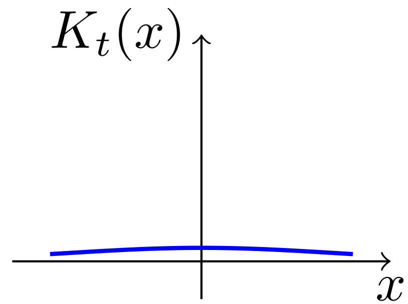
At $t = 0.05$.



At $t = 0.1$.



At $t = 1$.



At $t = 10$.

- We can observe that at any time instant $t > 0$, this fundamental solution is nonzero everywhere.

- Further, we can observe that as time increases, this fundamental solution shrinks and widens, eventually decaying everywhere to zero.

Properties of Gaussian kernel

- As mentioned previously, the Gaussian kernel K is **singular at the origin** $(\boldsymbol{x}, t) = (\mathbf{0}, 0)$.

- The Gaussian kernel is **strictly positive** everywhere in $\mathbb{R}^m$ for $t > 0$, that is,

$$K_t > 0 \qquad \text{for } t > 0.$$

- The Gaussian kernel is **normalized** in that it integrates to 1 w.r.t. the $\boldsymbol{x}$ variable for $t > 0$, that is,

$$\int_{\mathbb{R}^m} K_t(\boldsymbol{x}) d\boldsymbol{x} = 1 \qquad \text{for } t > 0.$$

- The **Fourier transform** of the Gaussian kernel (with respect to the $\boldsymbol{x}$ variable) reads as

$$\widehat{K}_t(\boldsymbol{\xi}) = \int_{\mathbb{R}^m} \exp(-i\boldsymbol{\xi} \cdot \boldsymbol{x}) K_t(\boldsymbol{x}) d\boldsymbol{x} = \exp\left(-t\|\boldsymbol{\xi}\|^2\right) \qquad \text{for } t > 0.$$

  Thus, the Fourier transform of the Gaussian kernel is itself Gaussian (albeit no longer normalized).

- These properties of the Gaussian kernel are standard properties, whose proof can be found in many references. The normalization property can be proven by carrying out a change of variables to polar coordinates. The Fourier transform can be proven by contour integration in the complex plane.

Proof of Gaussian kernel being fundamental solution for heat equation

- A fundamental solution K for the heat equation must satisfy

$$\frac{\partial K}{\partial t}(\boldsymbol{x},t) - \triangle_{\boldsymbol{x}} K(\boldsymbol{x},t) = \delta_{(\mathbf{0},0)}(\boldsymbol{x},t), \qquad (\boldsymbol{x},t) \in \mathbb{R}^m \times \mathbb{R}.$$

- Taking the Fourier transform of this equation with respect to the $\boldsymbol{x}$ variable, we obtain

$$\frac{\partial \widehat{K}}{\partial t}(\boldsymbol{\xi},t) + \|\boldsymbol{\xi}\|^2 \widehat{K}(\boldsymbol{\xi},t) = \delta_0(t), \qquad (\boldsymbol{\xi},t) \in \mathbb{R}^m \times \mathbb{R}.$$

- This equation is solved by the function

$$\widehat{K}(\boldsymbol{\xi},t) = H(t) \exp\left(-t\|\boldsymbol{\xi}\|^2\right),$$

  where H is the Heavyside function such that $H(t) = 1$ if $t \geq 0$ and $H(t) = 0$ otherwise. Indeed, the derivative of $\widehat{K}(\boldsymbol{\xi},t)$ w.r.t. t in the sense of the distributions, refer to Lecture 2, reads as

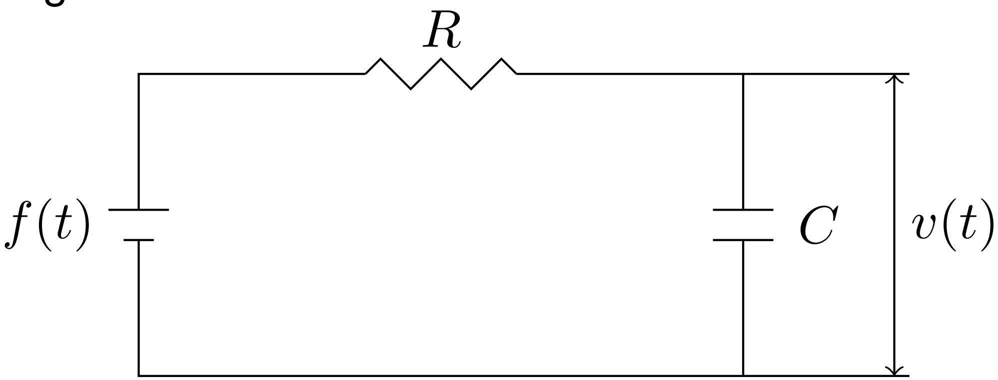
$$\frac{\partial \widehat{K}}{\partial t}(\boldsymbol{\xi},t) = -H(t)\|\boldsymbol{\xi}\|^2 \exp\left(-t\|\boldsymbol{\xi}\|^2\right) + \delta_0(t).$$

- By the aforementioned expression for the Fourier transform of the Gaussian kernel, we obtain

$$K(\boldsymbol{x},t) = H(t)(4\pi t)^{-m/2} \exp\left(-\frac{\|\boldsymbol{x}\|^2}{4t}\right) = \begin{cases} 0 & \text{for } t < 0, \\ (4\pi t)^{-m/2} \exp\left(-\frac{\|\boldsymbol{x}\|^2}{4t}\right) & \text{for } t > 0. \end{cases}$$

Motivating example: RC circuit

- Let us consider the following RC circuit:



The voltage $v(t)$ across the capacitor is governed by the ODE

$$\frac{dv}{dt}(t) + \frac{1}{RC} v(t) = \frac{f(t)}{RC}, \qquad t \in \mathbb{R},$$

where R is the resistance of the resistor and C the capacitance of the capacitor.

- The impulse response function h associated to this ODE is given by

$$h(t) = h_t = H(t) \exp\left(-\frac{t}{RC}\right),$$

where H is the Heavyside function such that $H(t) = 1$ if $t \geq 0$ and $H(t) = 0$ otherwise. Indeed, the derivative of h w.r.t. t in the sense of the distributions, refer to Lecture 2, reads as

$$\frac{dh}{dt} = H(t)\left(-\frac{1}{RC}\right) \exp\left(-\frac{t}{RC}\right) + \delta(t),$$

so that we can readily verify that h indeed solves $\frac{dh}{dt}(t) + \frac{1}{RC} h(t) = \delta(t)$, $t \in \mathbb{R}$.

Motivating example: RC circuit (continued)

- Superposition formulae for determining solutions for general right-hand side:
  - The initial-value problem

$$\begin{cases} \dfrac{dv}{dt}(t) + \dfrac{1}{RC} v(t) = 0, & \text{in }]0,+\infty[, \\ v(0) = g & \text{at } t = 0, \end{cases}$$

is solved by

$$v(t) = h_t g = \exp\left(-\frac{t}{RC}\right) g \quad \text{for } t \geq 0.$$

  - The inhomogeneous problem, for sufficiently regular f,

$$\begin{cases} \dfrac{dv}{dt}(t) + \dfrac{1}{RC} v(t) = \dfrac{f(t)}{RC}, & \text{in }]0,+\infty[, \\ v(0) = 0 & \text{at } t = 0, \end{cases}$$

is solved by

$$v(t) = h \star \frac{f(t)}{RC} = \int_{-\infty}^{t} \exp\left(-\frac{t-s}{RC}\right) \frac{f(s)}{RC} ds \quad \text{for } t \geq 0.$$

  - The inhomogeneous problem with general initial datum, for sufficiently regular f,

$$\begin{cases} \dfrac{dv}{dt}(t) + \dfrac{1}{RC} v(t) = \dfrac{f(t)}{RC}, & \text{in }]0,+\infty[, \\ v(0) = g & \text{at } t = 0, \end{cases}$$

is solved by

$$v(t) = h_t g + h \star \frac{f(t)}{RC} = \exp\left(-\frac{t}{RC}\right) g + \int_{-\infty}^{t} \exp\left(-\frac{t-s}{RC}\right) \frac{f(s)}{RC} ds \quad \text{for } t \geq 0.$$

Superposition formulae

- A fundamental solution allows **superposition formulae** to be established for determining a solution to the heat equation for a general right-hand side:

  - Initial-value problem:
    Given a sufficiently regular function g from $\mathbb{R}^m$ into $\mathbb{R}$, the problem

$$\begin{cases} \dfrac{\partial u}{\partial t} - \triangle_{\boldsymbol{x}} u = 0 & \text{in } \mathbb{R}^m \times]0, +\infty[, \\ u = g & \text{on } \mathbb{R}^m \times \{t = 0\}, \end{cases}$$

    is solved by the convolution of the Gaussian kernel with g:

$$u(\boldsymbol{x}, t) = K_t \star g(\boldsymbol{x}) \qquad \text{for } t > 0.$$

    The solution to the initial-value problem is obtained by convolution of the initial datum with the Gaussian kernel, which becomes progressively wider and wider as t increases.

Superposition formulae (continued)

- Inhomogeneous problem:
  Given a sufficiently regular function f from $\mathbb{R}^m\times]0,+\infty[$ into $\mathbb{R}$, the problem

$$\begin{cases} \dfrac{\partial u}{\partial t} - \triangle_{\boldsymbol{x}} u = f & \text{in } \mathbb{R}^m\times]0,+\infty[, \\ u = 0 & \text{on } \mathbb{R}^m \times \{t=0\}, \end{cases}$$

  is solved by the convolution of the Gaussian kernel with f:

$$u(\boldsymbol{x},t) = K \star f(\boldsymbol{x},t) \qquad \text{for } t>0.$$

- Inhomogeneous problem with general initial datum:
  The previous results can be combined to assert that given sufficiently regular functions f from $\mathbb{R}^m\times]0,+\infty[$ into $\mathbb{R}$ and g from $\mathbb{R}^m$ into $\mathbb{R}$, the problem

$$\begin{cases} \dfrac{\partial u}{\partial t} - \triangle_{\boldsymbol{x}} u = f & \text{in } \mathbb{R}^m\times]0,+\infty[, \\ u = g & \text{on } \mathbb{R}^m \times \{t=0\}, \end{cases}$$

  is solved by the sum of the convolutions of the Gaussian kernel with f and g:

$$u(\boldsymbol{x},t) = K_t \star g(\boldsymbol{x}) + K \star f(\boldsymbol{x},t) \qquad \text{for } t>0.$$

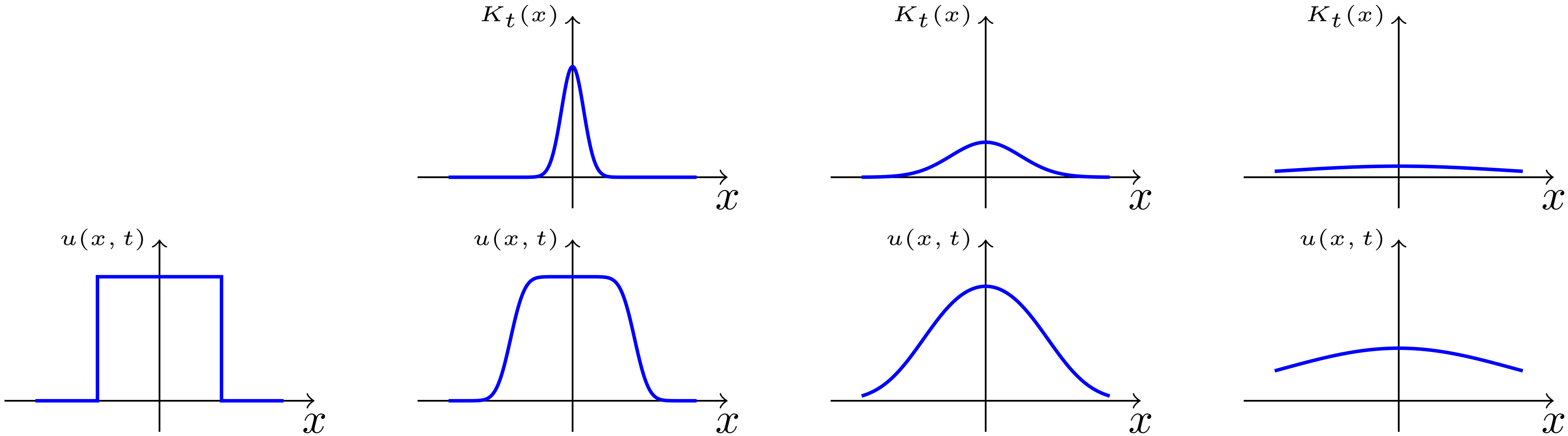
Example of use of superposition formulae

- Let us consider the initial-value problem in 1D ($m = 1$) for the initial value

$$u(x,0) = g(x) = \begin{cases} 1 & \text{for } a \leq x \leq b, \\ 0 & \text{otherwise.} \end{cases}$$

- For this initial-value problem, the aforementioned superposition formula provides the solution

$$u(x,t) = K_t \star g(x) = \int_{\mathbb{R}} K_t(x-y) g(y) dy = \int_{\mathbb{R}} \frac{1}{\sqrt{4\pi t}} \exp\left(-\frac{(x-y)^2}{4t}\right) g(y) dy \quad \text{for } t > 0.$$



At $t = 0$. At $t = 0.1$. At $t = 1$. At $t = 10$.

- We can observe that the convolution integral replaces each initial value by a weighted average of nearby values, the weight being determined by the progressively wider Gaussian kernel. This has the effect of progressively **smoothening** the initial datum.

Example of use of superposition formulae (continued)

- We have seen that the heat equation entails progressive smoothening of the initial datum.
- In fact, to smoothen images, graphics software, such as Photoshop or Instagram, applies so called Gaussian blurring filters, which smoothen images by replacing each pixel value by a weighted average of nearby values, the weight being determined by a Gaussian kernel.

Original image.

Gaussian blur with small pixel radius.

Gaussian blur with medium pixel radius.

Gaussian blur with large pixel radius.

- Please refer, for example, to the Wikipedia article on Gaussian blur.

Proof of superposition formulae using Fourier analysis

- One way of proving the aforementioned superposition formulae involves the use of Fourier analysis:
  - For example, let us consider the initial-value problem
$$\begin{cases} \dfrac{\partial u}{\partial t} - \triangle_{\boldsymbol{x}} u = 0 & \text{in } \mathbb{R}^m \times]0, +\infty[, \\ u = g & \text{on } \mathbb{R}^m \times \{t = 0\}. \end{cases}$$
  - Taking the Fourier transform of these equations with respect to the $\boldsymbol{x}$ variable, we obtain
$$\begin{cases} \dfrac{\partial \hat{u}}{\partial t} + \|\boldsymbol{\xi}\|^2 \hat{u} = 0 & \text{in } \mathbb{R}^m \times]0, +\infty[, \\ \hat{u} = \hat{g} & \text{on } \mathbb{R}^m \times \{t = 0\}. \end{cases}$$
  - The solution to this initial-value problem reads as
$$\hat{u}(\boldsymbol{\xi}, t) = \hat{g}(\boldsymbol{\xi}) \exp(-t\|\boldsymbol{\xi}\|^2) \qquad \text{for } t > 0.$$
  - By the expression for the Fourier transform of the Gaussian kernel, and because the Fourier transform interchanges convolution and multiplication of functions, we obtain:
$$u(\boldsymbol{x}, t) = K_t \star g(\boldsymbol{x}) \qquad \text{for } t > 0.$$
  - Of course, to make this proof mathematically fully rigorous, the function g must be required to be sufficiently regular for the Fourier transform and the convolution to make sense.

Proof of superposition formulae in space-time domain

- Alternatively, the superposition formulae can also be proven directly in the space-time domain.

- Review of theorem relevant to differentiation under integral sign:

  Let $\mathcal{X}$ be an open (bounded or unbounded) subset of $\mathbb{R}^m$.

  Let $\mathcal{Y}$ be an open (bounded or unbounded) subset of $\mathbb{R}^n$.

  Let f be a function from $\mathcal{X} \times \mathcal{Y}$ into $\mathbb{R}$.

  (i) Let f be integrable in that $\int_{\mathcal{X}} |f(\boldsymbol{x}, \boldsymbol{y})| d\boldsymbol{x} < +\infty$ for all $\boldsymbol{y}$ in $\mathcal{Y}$.

  (ii) Let the partial derivative $\partial_{y_j} f$ be integrable in that $\int_{\mathcal{X}} |\partial_{y_j} f(\boldsymbol{x}, \boldsymbol{y})| d\boldsymbol{x} < +\infty$ for all $\boldsymbol{y}$ in $\mathcal{Y}$.

  (iii) Let there exist a function g from $\mathcal{X}$ into $\mathbb{R}$ that is integrable, that is, $\int_{\mathcal{X}} |g(\boldsymbol{x})| d\boldsymbol{x} < +\infty$, and such that $|\partial_{y_j} f(\cdot, \boldsymbol{y})| \leq |g|$ for all $\boldsymbol{y}$ in $\mathcal{Y}$.

  Then, the following differentiation under the integral sign is permitted:

$$\frac{\partial}{\partial y_j} \int_{\mathcal{X}} f(\boldsymbol{x}, \boldsymbol{y}) d\boldsymbol{x} = \int_{\mathcal{X}} \frac{\partial f}{\partial y_j}(\boldsymbol{x}, \boldsymbol{y}) d\boldsymbol{x}, \qquad \boldsymbol{y} \in \mathcal{Y}.$$

Proof of superposition formulae in space-time domain (continued)

- Initial-value problem:

  - Let us show that given a continuous and bounded function g from $\mathbb{R}^m$ into $\mathbb{R}$, the problem
$$\begin{cases} \dfrac{\partial u}{\partial t} - \triangle_{\boldsymbol{x}} u = 0 & \text{in } \mathbb{R}^m \times]0, +\infty[, \\ u = g & \text{on } \mathbb{R}^m \times \{t = 0\}, \end{cases}$$
is solved by the convolution of the Gaussian kernel with g,
$$u(\boldsymbol{x}, t) = K_t \star g(\boldsymbol{x}) = \int_{\mathbb{R}^m} K_t(\boldsymbol{x} - \boldsymbol{y}) g(\boldsymbol{y}) d\boldsymbol{y} \qquad \text{for } t > 0.$$

  - Using the aforementioned theorem relevant to differentiation under the integral sign, we obtain
$$\frac{\partial u}{\partial t}(\boldsymbol{x}, t) - \triangle_{\boldsymbol{x}} u(\boldsymbol{x}, t) = \int_{\mathbb{R}^m} \underbrace{\left(\frac{\partial K_t}{\partial t}(\boldsymbol{x} - \boldsymbol{y}) - \triangle_{\boldsymbol{x}} K_t(\boldsymbol{x} - \boldsymbol{y}) \right)}_{=0} g(\boldsymbol{y}) d\boldsymbol{y} \quad \text{for } t > 0,$$
thus showing that the PDE involved in the initial-value problem holds in $\mathbb{R}^m \times]0, +\infty[$.

  - It remains to be shown that the initial condition holds, that is,
$$\lim_{\substack{(\boldsymbol{x}, t) \to (\boldsymbol{x}_0, 0) \\ \boldsymbol{x} \in \mathbb{R}^m,\ t > 0}} u(\boldsymbol{x}, t) = g(\boldsymbol{x}_0) \qquad \text{for all } \boldsymbol{x}_0 \text{ in } \mathbb{R}^m.$$

Proof of superposition formulae in space-time domain (continued)

- We will prove this limit by relying essentially on the assumed continuity of g and on $u(\boldsymbol{x}, t)$ being obtained by convolution with a progressively narrower Gaussian kernel as $t \to 0$.

- Because of the normalization property of the Gaussian kernel, we have

$$\big|g(\boldsymbol{x}_0) - u(\boldsymbol{x}, t)\big| = \left|\int_{\mathbb{R}^m} K_t(\boldsymbol{x} - \boldsymbol{y})\big(g(\boldsymbol{x}_0) - g(\boldsymbol{y})\big)d\boldsymbol{y}\right| \qquad \text{for } t > 0.$$

- Given $\boldsymbol{x}_0$ in $\mathbb{R}^m$ and $\epsilon > 0$, there exists, because we assumed g to be continuous, a scalar $\delta > 0$ such that if $\|\boldsymbol{x}_0 - \boldsymbol{y}\| < 2\delta$, then $|g(\boldsymbol{x}_0) - g(\boldsymbol{y})| < \frac{\epsilon}{2}$. Given δ, we can excise a ball $B_{2\delta}(\boldsymbol{x}_0) = \{\boldsymbol{y} \in \mathbb{R}^m : \|\boldsymbol{x}_0 - \boldsymbol{y}\| < 2\delta\}$ around $\boldsymbol{x}_0$ to partition the domain of integration as

$$\left|\underbrace{\int_{B_{2\delta}(\boldsymbol{x}_0)} K_t(\boldsymbol{x} - \boldsymbol{y})\big(g(\boldsymbol{x}_0) - g(\boldsymbol{y})\big)d\boldsymbol{y}}_{=\mathcal{I}} + \underbrace{\int_{\mathbb{R}^m \setminus B_{2\delta}(\boldsymbol{x}_0)} K_t(\boldsymbol{x} - \boldsymbol{y})\big(g(\boldsymbol{x}_0) - g(\boldsymbol{y})\big)d\boldsymbol{y}}_{=\mathcal{J}}\right|.$$

- Because of the normalization property of the Gaussian kernel, the first term, $\mathcal{I}$, is such that

$$\mathcal{I} = \left|\int_{B_{2\delta}(\boldsymbol{x}_0)} K_t(\boldsymbol{x} - \boldsymbol{y})\underbrace{\big(g(\boldsymbol{x}_0) - g(\boldsymbol{y})\big)}_{\leq \frac{\epsilon}{2}} d\boldsymbol{y}\right| \leq \frac{\epsilon}{2} \qquad \text{for } t > 0.$$

Proof of superposition formulae in space-time domain (continued)

- Because we assumed g to be bounded, the second term, $\mathcal{J}$, is such that

$$\mathcal{J} \leq \left(2 \max_{\boldsymbol{y}\in\mathbb{R}^m} |g(\boldsymbol{y})|\right) \int_{\mathbb{R}^m\setminus B_{2\delta}(\boldsymbol{x}_0)} K_t(\boldsymbol{x}-\boldsymbol{y})d\boldsymbol{y} \quad \text{for } t>0.$$

Owing to the triangle inequality, we have $\|\boldsymbol{x}_0-\boldsymbol{y}\| \leq \|\boldsymbol{x}_0-\boldsymbol{x}\| + \|\boldsymbol{x}-\boldsymbol{y}\|$ and therefore if $\|\boldsymbol{x}_0-\boldsymbol{y}\| \geq 2\delta$ and $\|\boldsymbol{x}_0-\boldsymbol{x}\| < \delta$, we have $\|\boldsymbol{x}-\boldsymbol{y}\| > \|\boldsymbol{x}_0-\boldsymbol{y}\|/2$; thus,

$$\begin{aligned}
\mathcal{J} &\leq \left(2 \max_{\boldsymbol{y}\in\mathbb{R}^m} |g(\boldsymbol{y})|\right) \int_{\mathbb{R}^m\setminus B_{2\delta}(\boldsymbol{x}_0)} \frac{1}{\sqrt{(4\pi t)^m)}} \exp\left(-\frac{\|\boldsymbol{x}-\boldsymbol{y}\|^2}{4t}\right) d\boldsymbol{y} \\
&< \left(2 \max_{\boldsymbol{y}\in\mathbb{R}^m} |g(\boldsymbol{y})|\right) \int_{\mathbb{R}^m\setminus B_{2\delta}(\boldsymbol{x}_0)} \frac{1}{\sqrt{(4\pi t)^m}} \exp\left(-\frac{\|\boldsymbol{x}_0-\boldsymbol{y}\|^2}{16t}\right) d\boldsymbol{y} \\
&= \left(2 \max_{\boldsymbol{y}\in\mathbb{R}^m} |g(\boldsymbol{y})|\right) \int_{\mathbb{R}^m\setminus B_{2\delta/\sqrt{t}}(\boldsymbol{0})} \frac{1}{\sqrt{(4\pi)^m}} \exp\left(-\frac{\|\boldsymbol{z}\|^2}{16}\right) d\boldsymbol{z} \\
&\to 0 \quad \text{as } t \to 0.
\end{aligned}$$

Hence, if $\|\boldsymbol{x}_0-\boldsymbol{x}\| < \delta$, there exists a scalar $\kappa > 0$ such that if $0 < t < \kappa$, then $\mathcal{J} < \frac{\epsilon}{2}$.

- We can conclude that for all $\boldsymbol{x}_0$ in $\mathbb{R}^m$, there exists for all $\epsilon > 0$ a $\delta > 0$ and a $\kappa > 0$ such that if $\|\boldsymbol{x}_0-\boldsymbol{x}\| < \delta$ and $0 < t < \kappa$, then $|g(\boldsymbol{x}_0 - u(\boldsymbol{x},t)| < \epsilon$, as asserted.

Proof of superposition formulae in space-time domain (continued)

- Inhomogeneous problem:

  - Let us show that given a continuous function f from $\mathbb{R}^m \times [0, +\infty[$ into $\mathbb{R}$ with continuous partial derivatives $\triangle_{\boldsymbol{x}} f$ and $\partial_t f$ and with closed and bounded support, the problem

$$\begin{cases} \dfrac{\partial u}{\partial t} - \triangle_{\boldsymbol{x}} u = f & \text{in } \mathbb{R}^m \times]0, +\infty[, \\ u = 0 & \text{on } \mathbb{R}^m \times \{t = 0\}, \end{cases}$$

    is solved by the convolution of the Gaussian kernel with f,

$$u(\boldsymbol{x}, t) = K \star f(\boldsymbol{x}, t) = \int_0^t \int_{\mathbb{R}^m} K(\boldsymbol{x} - \boldsymbol{y}, t - s) f(\boldsymbol{y}, s) d\boldsymbol{y} ds \qquad \text{for } t > 0.$$

  - Because of the singularity of K at the origin, we cannot directly differentiate under the integral sign. To circumvent this difficulty, as in the case of the integral representation theorem in Lecture 2, we proceed by first isolating the singularity in a small neighborhood of the origin and then considering a limit as this neighborhood becomes smaller and smaller.

Proof of superposition formulae in space-time domain (continued)

- After the change of variables

$$u(\boldsymbol{x},t) = \int_0^t \int_{\mathbb{R}^m} K(\boldsymbol{y},s) f(\boldsymbol{x}-\boldsymbol{y},t-s) d\boldsymbol{y} ds \qquad \text{for } t>0,$$

we can invoke the aforementioned theorem to justify that the assumption of f being continuous from $\mathbb{R}^m \times [0,+\infty[$ into $\mathbb{R}$ with continuous partial derivatives $\boldsymbol{\nabla}_{\boldsymbol{x}} f$, $\mathrm{D}_{\boldsymbol{x}}\boldsymbol{\nabla}_{\boldsymbol{x}} f$, and $\partial_t f$ and with closed and bounded support allows the following differentiations under the integral sign:

$$\frac{\partial u}{\partial t}(\boldsymbol{x},t) = \int_0^t \int_{\mathbb{R}^m} K(\boldsymbol{y},s) \frac{\partial f}{\partial t}(\boldsymbol{x}-\boldsymbol{y},t-s) d\boldsymbol{y} ds + \int_{\mathbb{R}^m} K(\boldsymbol{y},t) f(\boldsymbol{x}-\boldsymbol{y},0) d\boldsymbol{y} \quad \text{for } t>0,$$

$$\triangle_{\boldsymbol{x}} u(\boldsymbol{x},t) = \int_0^t \int_{\mathbb{R}^m} K(\boldsymbol{y},s) \triangle_{\boldsymbol{x}} f(\boldsymbol{x}-\boldsymbol{y},t-s) d\boldsymbol{y} ds \qquad \text{for } t>0,$$

thus leading to

$$\left(\frac{\partial}{\partial t} - \triangle_{\boldsymbol{x}}\right) u(\boldsymbol{x},t) = \int_0^t \int_{\mathbb{R}^m} K(\boldsymbol{y},s) \left(-\frac{\partial}{\partial s} - \triangle_{\boldsymbol{y}}\right) f(\boldsymbol{x}-\boldsymbol{y},t-s) d\boldsymbol{y} ds$$
$$+ \int_{\mathbb{R}^m} K(\boldsymbol{y},t) f(\boldsymbol{x}-\boldsymbol{y},0) d\boldsymbol{y} \qquad \text{for } t>0.$$

Proof of superposition formulae in space-time domain (continued)

- To deal with the singularity of K at the origin, we excise a small neighborhood of the origin to partition the domain of integration as

$$\underbrace{\int_0^\epsilon \int_{\mathbb{R}^m} K(\boldsymbol{y},s)\left(-\frac{\partial}{\partial s}-\triangle_{\boldsymbol{y}}\right) f(\boldsymbol{x}-\boldsymbol{y},t-s)d\boldsymbol{y}ds}_{=\mathcal{I}}$$
$$+\underbrace{\int_\epsilon^t \int_{\mathbb{R}^m} K(\boldsymbol{y},s)\left(-\frac{\partial}{\partial s}-\triangle_{\boldsymbol{y}}\right) f(\boldsymbol{x}-\boldsymbol{y},t-s)d\boldsymbol{y}ds}_{=\mathcal{J}}.$$

- Because of the normalization property of the Gaussian kernel and because we assumed that the partial derivatives $\triangle_{\boldsymbol{x}} f$ and $\partial_t f$ of f are continuous from $\mathbb{R}^m \times [0,+\infty[$ into $\mathbb{R}$ and have a closed and bounded support, hence, are bounded, the first term, $\mathcal{I}$, is such that

$$|\mathcal{I}| \leq \epsilon \left(\max_{(\boldsymbol{y},s)\in\mathbb{R}^m\times[0,+\infty[} \left|\frac{\partial f}{\partial s}(\boldsymbol{y},s)\right| + \max_{(\boldsymbol{y},s)\in\mathbb{R}^m\times[0,+\infty[} |\triangle_{\boldsymbol{y}} f(\boldsymbol{y},s)| \right),$$

hence,

$$\lim_{\epsilon\to 0} |\mathcal{I}| = 0.$$

Proof of superposition formulae in space-time domain (continued)

- Integrating by parts, we find for the second term, $\mathcal{J}$, that

$$|\mathcal{J}| = \int_{\epsilon}^{t} \int_{\mathbb{R}^m} \left(\frac{\partial}{\partial s} - \triangle_{\boldsymbol{y}} \right) K(\boldsymbol{y}, s) f(\boldsymbol{x} - \boldsymbol{y}, t - s) d\boldsymbol{y} ds$$
$$+ \int_{\mathbb{R}^m} K(\boldsymbol{y}, \epsilon) f(\boldsymbol{x} - \boldsymbol{y}, t - \epsilon) d\boldsymbol{y}$$
$$- \int_{\mathbb{R}^m} K(\boldsymbol{y}, t) f(\boldsymbol{x} - \boldsymbol{y}, 0) d\boldsymbol{y},$$

where, as in the proof the superposition formula for the initial-value problem, we have

$$\lim_{\epsilon \to 0} \int_{\mathbb{R}^m} K(\boldsymbol{y}, \epsilon) f(\boldsymbol{x} - \boldsymbol{y}, t - \epsilon) d\boldsymbol{y} = f(\boldsymbol{x}, t).$$

- We can conclude that $\partial_t u - \triangle_{\boldsymbol{x}} u = 0$ in $\mathbb{R}^m \times]0, +\infty[$, as asserted.

- Because of the normalization property of K and because we assumed f is continuous from $\mathbb{R}^m \times [0, +\infty[$ into $\mathbb{R}$ and has a closed and bounded support, hence, is bounded, we have

$$\max_{\boldsymbol{y} \in \mathbb{R}^m} |u(\boldsymbol{y}, t)| \leq t \max_{(\boldsymbol{y}, s) \in \mathbb{R}^m \times [0, +\infty[} |f(\boldsymbol{y}, s)|.$$

Consequently, the initial condition holds, that is,

$$\lim_{t \to 0} \max_{\boldsymbol{y} \in \mathbb{R}^m} |u(\boldsymbol{y}, t)| = 0.$$

Infinite propagation speed

- Let us consider again the initial-value problem

$$\begin{cases} \dfrac{\partial u}{\partial t} - \triangle_{\boldsymbol{x}} u = 0 & \text{in } \mathbb{R}^m \times]0, +\infty[, \\ u = g & \text{on } \mathbb{R}^m \times \{t = 0\}. \end{cases}$$

- If the initial value is positive, that is, $g \geq 0$ everywhere in $\mathbb{R}^m$, and strictly positive at least somewhere, that is, $g \neq 0$ everywhere in $\mathbb{R}^m$, then, because the fundamental solution K is strictly positive everywhere in $\mathbb{R}^m$ at any later time $t > 0$ no matter how small, the solution

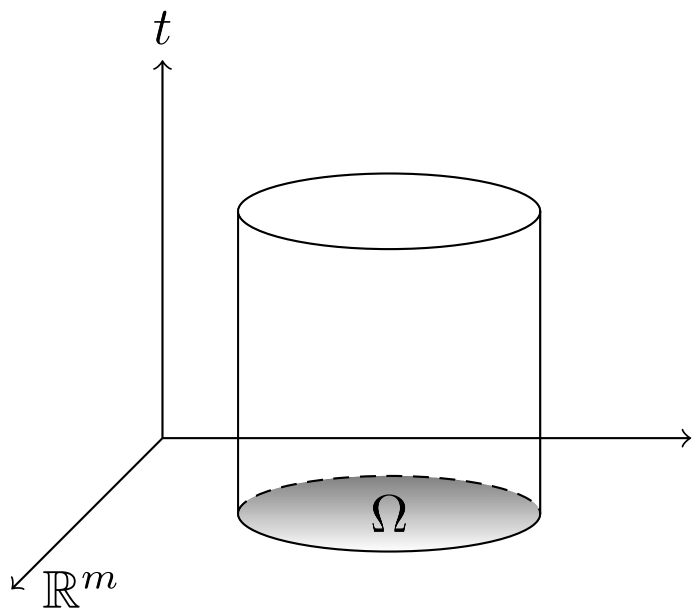
$$u(\boldsymbol{x}, t) = K_t \star g(\boldsymbol{x}) = \int_{\mathbb{R}^m} K_t(\boldsymbol{x} - \boldsymbol{y}) g(\boldsymbol{y}) d\boldsymbol{y} \qquad \text{for } t > 0$$

is in fact strictly positive everywhere in $\mathbb{R}^m$ at any later time $t > 0$ no matter how small. This observation can be interpreted by saying that the heat equation entails **infinite propagation speed**.

- In spite of this striking, and decidedly **nonphysical**, artifact of the heat equation, equations similar to the heat equation are very useful as mathematical models of diffusive phenomena in many applications in science and engineering, as well as of great mathematical importance.

Maximum property

- Let Ω be an open bounded subset of $\mathbb{R}^m$ and $0 < \tau < +\infty$. Let u be a continuous function from $\overline{\Omega} \times [0, \tau]$ into $\mathbb{R}$ that satisfies the homogeneous heat equation $\partial_t u - \triangle_{\boldsymbol{x}} u = 0$ on $\Omega \times]0, \tau[$. Then, u assumes its maximum either on $\Omega \times \{t = 0\}$ or on $\partial\Omega \times [0, \tau]$.



- Consequences of the maximum property:
  This maximum property has consequences for uniqueness and stability properties of solutions. For example, it indicates the “overdeterminedness” of the Dirichlet problem: a “terminal condition” on $\Omega \times \{t = \tau\}$ cannot be specified.

Heat equation on a portion of space

Notion of initial-boundary-value problem

- If one wishes to solve the heat equation in an open bounded subset Ω of $\mathbb{R}^m$, it is appropriate to specify not only an initial value $u(\boldsymbol{x}, 0)$ on Ω but also a boundary condition, for example,
  - a Dirichlet boundary condition: $u = g$,
  - a Neumann boundary condition: $\boldsymbol{\nabla}_{\boldsymbol{x}} u \cdot \boldsymbol{n} = g$,

  the interpretation of which depends on the application. For example, in a heat conduction problem, a Dirichlet-type boundary condition would specify the temperature on the boundary, and a Neumann-type boundary condition would specify the heat flux through the boundary.

Example of initial-boundary-value problem

- As an example, let us look more closely at the initial-boundary-value problem

$$\begin{cases} \dfrac{\partial u}{\partial t} - \triangle_{\boldsymbol{x}} u = 0 & \text{in } \Omega \times]0, +\infty[, \\ u(\boldsymbol{x}, 0) = g(\boldsymbol{x}) & \text{on } \Omega \times \{t = 0\} \quad \text{(initial condition)}, \\ u(\boldsymbol{x}, t) = 0 & \text{on } \partial\Omega \times]0, +\infty[\quad \text{(boundary condition)}. \end{cases}$$

- This problem can be solved by the method of separation of variables.

Example of initial-boundary-value problem (continued)

- **Eigenproblem**: With reference to the boundary condition $u(\boldsymbol{x},t)=0$ on $\partial\Omega\times]0,+\infty[$, we begin by solving the following eigenproblem:

$$\begin{cases} -\triangle_{\boldsymbol{x}}\varphi_k(\boldsymbol{x}) = \lambda_k\varphi_k(\boldsymbol{x}) & \text{in } \Omega, \\ \varphi_k(\boldsymbol{x}) = 0 & \text{on } \partial\Omega. \end{cases}$$

  Let us assume that we can find a sequence of eigenvalues λ_k with corresponding eigenfunctions φ_k, which form an orthonormal basis for $L^2(\Omega)$.

- **Function series**: Given the eigenfunctions $\{\varphi_k\}_{k=1}^{+\infty}$, we seek a solution of the following form:

$$u(\boldsymbol{x},t) = \sum_{k=1}^{+\infty} b_k(t)\varphi_k(\boldsymbol{x});$$

  thus, we seek a solution of the form of a series of products of functions of fewer independent variables, namely, a solution of the form of a series of products of functions b_k of only t (yet to be determined) and the eigenfunctions φ_k of only $\boldsymbol{x}$ (already known).

Example of initial-boundary-value problem (continued)

- **System of uncoupled equations ("diagonalization")**: Inserting this function series into the IBVP and assuming the validity of term-by-term differentiation, we obtain

$$\begin{cases} \displaystyle\sum_{k=1}^{+\infty} \left(\frac{db_k}{dt}(t)\varphi_k(\boldsymbol{x}) - b_k(t)\triangle_{\boldsymbol{x}}\varphi_k(\boldsymbol{x}) \right) = 0 & \text{in } \Omega\times]0,+\infty[, \\ \displaystyle\sum_{k=1}^{+\infty} b_k(0)\varphi_k(\boldsymbol{x}) = g(\boldsymbol{x}) & \text{on } \Omega \times \{t=0\}, \\ \displaystyle\sum_{k=1}^{+\infty} b_k(t)\varphi_k(\boldsymbol{x}) = 0 & \text{on } \partial\Omega\times]0,+\infty[. \end{cases}$$

  - If g is square-integrable, we have $g(\boldsymbol{x}) = \sum_{k=1}^{+\infty} g_k\varphi_k(\boldsymbol{x})$ where $g_k = \int_\Omega g(\boldsymbol{x})\varphi_k(\boldsymbol{x})d\boldsymbol{x}$ for $k = 1, 2, 3, \ldots$; thus,

$$\begin{cases} \displaystyle\sum_{k=1}^{+\infty} \left(\frac{db_k}{dt}(t)\varphi_k(\boldsymbol{x}) + b_k(t)\lambda_k\varphi_k(\boldsymbol{x}) \right) = 0 & \text{in } \Omega\times]0,+\infty[, \\ \displaystyle\sum_{k=1}^{+\infty} b_k(0)\varphi_k(\boldsymbol{x}) = \sum_{k=1}^{+\infty} g_k\varphi_k(\boldsymbol{x}) & \text{on } \Omega \times \{t=0\}, \\ \displaystyle\sum_{k=1}^{+\infty} b_k(t)\varphi_k(\boldsymbol{x}) = 0 & \text{on } \partial\Omega\times]0,+\infty[. \end{cases}$$

Example of initial-boundary-value problem (continued)

- Clearly, the previous equations hold if the following system of uncoupled IVPs holds:

$$\begin{cases} \dfrac{db_k}{dt}(t) + \lambda_k b_k(t) = 0 & \text{in }]0, +\infty[\\ b_k(0) = g_k & \text{at } t = 0 \end{cases}, \qquad \text{where } k = 1, 2, 3, \ldots.$$

  It can be easily verified that the solution to this system of uncoupled IVPs is given by

$$b_k(t) = g_k \exp(-\lambda_k t), \quad \text{where } k = 1, 2, 3, \ldots.$$

- As a conclusion, we obtain the following representation of the solution to the aforementioned IBVP:

$$u(\boldsymbol{x}, t) = \sum_{k=1}^{+\infty} g_k \exp(-\lambda_k t)\varphi_k(\boldsymbol{x}).$$

- The convergence and term-by-term differentiation still require justification, but we omit these details.
- We find the solution as a superposition of eigenfunctions φ_k whose coefficients decay to zero exponentially with corresponding rates λ_k.
- Because eigenfunctions φ_k associated with larger eigenvalues λ_k typically oscillate more rapidly in space, we can conclude from this representation that **components of the solution that oscillate more rapidly in space decay more rapidly in time**.

Summary and conclusion

- The heat equation is the mathematical prototype of models of diffusive phenomena, such as heat conduction, diffusive transport of substances (pollutants, solutes,...),...
- A fundamental solution for the heat equation is a solution that solves the heat equation for a Dirac impulse centered at the origin on the right-hand side. It can be shown that the Gaussian kernel, also called Gauss-Weierstrass kernel or heat kernel, is a fundamental solution for the heat equation.
- A fundamental solution allows superposition formulae to be established for determining solutions to initial-value and inhomogeneous problems with general right-hand sides.
- The notion of fundamental solution allows interesting properties to be deduced:
  - The heat equation entails infinite propagation speed. In spite of this striking, and decidedly nonphysical, artifact of the heat equation, equations similar to the heat equation are very useful as mathematical models of diffusive phenomena in many applications in science and engineering, as well as of great mathematical importance.
  - The heat equation entails progressive smoothening of the initial datum.
  - Maximum property.
- Certain initial-boundary-value problems involving the heat equation on a bounded portion of space can be solved by using the method of separation of variables.

References

Suggested reading material

- P. Olver. Introduction to Partial Differential Equations. Springer, 2014. Sections 4.1, 8.1, 11.2, 11.4, 11.5, and 12.4.

Additional references also consulted to prepare this lecture

- H. Brezis. Functional Analysis, Sobolev Spaces and Partial Differential Equations, Springer, 2010.
- L. Evans. Partial differential equations. American Mathematical Society, 2010.
- G. Folland. Introduction to partial differential equations. Princeton University Press, 1995.
- S. Lang. Real and functional analysis. Springer, 1993.
- P.-A. Raviart and J.-M. Thomas. Introduction à l'analyse numérique des équations aux dérivées partielles. Dunod, 1993.
- C. Zuily. Eléments de distributions et d'équations aux dérivées partielles. Dunod, 2002.