Technical report, IDE0891 , November 17, 2008

Introduction to Parabolic Differential Equations

Catherine Bandle

Mathematisches Institut, Universität Basel
Rheinsprung 21, CH-4051 Basel, Switzerland
catherine.bandle@unibas.ch

MASTER'S PROGRAMME IN FINANCIAL MATHEMATICS

School of Information Science, Computer and Electrical Engineering
Halmstad University

Contents

Preface

This lecture course gives a short introduction into the theory of parabolic differential equations. It is part of the lecture course **Introduction to Financial Mathematics**. The main lecture course contains two parts, one of them devoted to the Probability theory (15 lectures and classroom exercises) and the second one to the Theory of parabolic equations, which includes 5 lectures devoted to analytical questions and 10 lectures devoted to numerical methods. This text provides just these 5 lectures devoted to analytical properties of solutions which are the most important for applications.
We started 2006 a one year master program **Master in Financial Mathematics**. The program attracted a lot of students from abroad which have quite different preliminary education in the Theory of parabolic equations or Stochastics. Because the well known Black-Scholes model and its generalizations are parabolic equations we decided to develop such introductory course to give the students a possibility to consolidate their knowledge on this area.
We are very glad that Professor Catherine Bandle (Basel, Switzerland) was able to find time and possibility to came to Halmstad and provide for our students this course.

Ljudmila A. Bordag
Professor in Applied Mathematics
Director of the program
Master in Financial Mathematics
Halmstad University
Halmstad, 2008

1 Some results for ordinary differential equation

An *ordinary differential equation* is a relation of the form

$$F(u^{(n)}(t), u^{(n-1)}(t), \dots, \dot{u}(t), u(t), t) = 0, \tag{1}$$

where $t \in \mathbb{R}$ is a variable and $u(t)$ is a real valued function, $u^{(n)}(t) = \frac{d^n u}{dt^n}$ and $\dot{u} = \frac{du}{dt}$. The highest derivative n is called the *order* of the differential equation. Solving a differential equation means finding a continuous function $u(t)$ satisfying (1) pointwise.

First order equations

$$F(\dot{u}, u, t) = 0. \tag{2}$$

The first order equations are used to describe a relation between the growth rate of a quantity, $\dot{u}(t)$ in terms of the quantity at a fixed time t. One of the simplest evolution laws is the *organic growth*

$$\frac{du}{dt} = a\,u, \qquad u(t) = ce^{at}.$$

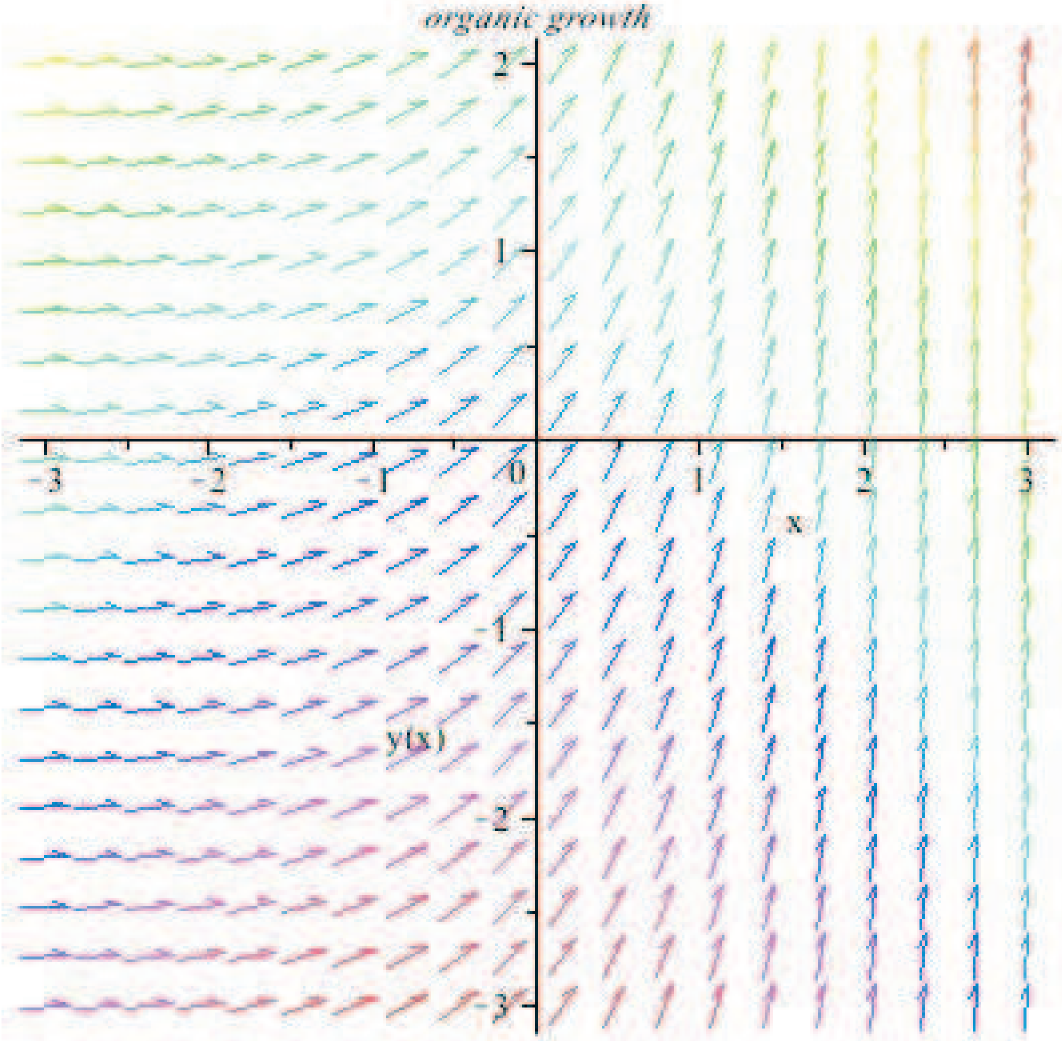
c is an arbitrary parameter which can be determined by prescribing an *initial* condition

$$u(t_0) = u_0.$$

More generally the growth rate of a substance depending only on the time and the total quantity at time t can be expressed as a first order differential equation of the type

$$\frac{du}{dt} = f(t, u(t)). \tag{3}$$

If $f(t, u) > 0$ particle produce new substance (source) whereas if $f(t, u) < 0$ then particles absorb the substance (sink). This equation yields a direction field in which the curve $u(t)$ has to be fitted.



Definition 1 (3) *is called linear if it is of the form*

$$\dot{u}(t) = a(t)u(t) + b(t).$$

It is a homogeneous equation if $b(t) = 0$ *otherwise it is called an inhomogeneous equation.*

If u_h is a solution of the homogeneous equation then the same is true for any multiple cu_h.

Every solution of a linear equation can be represented as

$$u(t) = u_p(t) + u_h(t),$$

where u_p is a particular solution of the inhomogeneous and u_h is a solution of the homogeneous problem. The general solution of the homogeneous problem can be expressed as

$$u_h(t) = e^{\int_{t_0}^{t} a(s)ds}c, \ c \in \mathbb{R} \text{ arbitrary constant number}$$

By means of the *variation of constants* we find

$$u_p(t) = \int_{t_0}^{t} e^{\int_\tau^t a(s)\, ds} f(\tau)\, d\tau.$$

Observation The solution of the linear problem is uniquely determined if we prescribe an *initial condition* $u(t_0) = u_0$.

General existence and uniqueness result

Definition 2 *A function $f : U \to W, \quad U \subset \mathbb{R}^2, \quad W \subset \mathbb{R}$ satisfies a Lipschitz condition with respect to u if there exist a constant $L > 0$ such that*

$$|f(t, u_1) - f(t, u_2)| \le L|u_1 - u_2|$$

for all $(t, u_1)\ ,\ (t, u_2) \in U$

Theorem 3 *Let the function $f(t, u)$ be a continuous in some set $U = \{(t, u) : |t - t_0| < a, |u - u_0| < b\}$. Assume that it satisfies a Lipschitz condition in U with respect to u. Then the initial value problem*

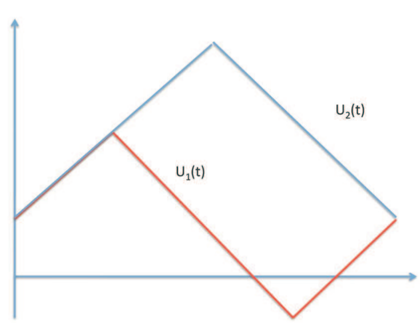
$$\dot{u}(t) = f(t, u(t)), \qquad u(t_0) = u_0$$

has a unique solution in the interval $t_0 - \alpha \le t \le t_0 + \alpha$, $\alpha = \min(a, \frac{b}{\|f\|_\infty})$, where $\|f\|_\infty = \max_U |f|$

Remarks

1. The fact that $\dot{u}$ can be expressed uniquely as a function of u and t is crucial for the validity of Theorem 3.

   Example where this is not the case: $|\dot{u}| = 1$. The solution of the initial value problem is not unique (cf. figure).



2. If in Theorem 3 the function f is only continuous and not Lipschitz, there still exists a solution, but it is not necessarily unique.

   Examples. $\dot{u} = u^p$, $0 < p < 1$, $u(0) = 0$.

Problems

1. Construct the direction fields of

$$\dot{u}(t) = e^{u(t)}, \quad \dot{u}(t) = u(t)(1 - u(t)).$$

2. Solve the two equations. Plot the solutions through $(0, 0.5)$. Discuss the long time behavior of the solutions.

3. Plot the nonlinearity $f(u) = u(1 - u)$. Discuss the role of this function for the solution computed above. When does it have the role of a source or a sink?

4. The existence theorem guarantees the existence of local solutions, i.e. solution in a neighborhood of t_0. Show by means of the equation $\dot{u}(t) = 1 + u^2(t)$ that all solutions cease to exist after a finite time.

References

E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, (1955)
P. Hartman, *Ordinary Differential Equations*, (1982)

2 Partial differential equations

2.1 Introduction

Let $u : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ be a function which depends on the time t and the space variable $x = (x_1, x_2, \ldots, x_N)$. Its first order partial derivatives are

$$u_{x_i} = \lim_{h \to 0} \frac{u(t, x_1, \ldots, x_i + h, \ldots, x_n) - u(x, t)}{h}, \quad u_t = \lim_{h \to 0} \frac{u(t + h, x) - u(x, t)}{h}.$$

The *gradient* ∇u is a vector pointing in the direction of biggest increase of the function $u(t, \cdot)$. It is given by

$$\nabla u(t, \cdot) = (u_{x_1}, u_{x_2}, \ldots, u_{x_N}).$$

Partial derivatives of the second order are defined recursively by

$$u_{x_i t} = (u_{x_i})_t, \quad u_{x_j x_i} = (u_{x_i})_{x_j}, u_{tt} = (u_t)_t,$$

and similarly higher order derivatives. An equation of the form

$$F(t, x, u, u_{x_i}, u_t, u_{x_i x_j}, u_{t x_i}) = 0, \quad i = 1, \ldots, N, j = 1, \ldots, N,$$

is called a *partial differential equation* of the second order.

We shall deal with partial differential equations of the form

$$u_t = \underbrace{\sum_{i=1}^{n} u_{x_i x_i}}_{\triangle u} + f(t, x, u, u_{x_i}), \quad i = 1, \ldots, N.$$

$\triangle u$ is called the *Laplacian* of u. This equation is an example of a *parabolic* differential equation.

Problems

1. Let $f(x) = 3x_1^2 + 5x_1x_2^3$
   Compute
$$\nabla f(x), \quad \triangle f(x).$$

2. If $u(x,t)$ is a solution of the heat equation $u_t = \triangle u$, prove that $v(t,x) = e^{\alpha t}u(x,t)$ is a solution of $v_t = \triangle v + \alpha v$.

3. If $u(x,t)$ solves $u_t = \triangle u$, what kind of equation does $v(t,x) = u(-t,x)$ solve?

2.2 Reaction-diffusion equations

2.2.1

Diffusion mechanism models the movement of a substance (population) in a media (environment).

Let $u : \mathbb{R} \times D \to \mathbb{R}^+$ be a function representing for instance the density of the population or of a substance.
Our goal is to know how $u(x,t)$ changes as time evolves and as the location x varies. We have three possibilities

1. particles can produce new individuals or kill existing ones (reaction)

2. particles can move around (diffusion)

3. both cases occur (reaction-diffusion) .

The case of pure reaction can be modeled by a ordinary differential equation as we have seen in Section 1.

2.2.2 The balance equation

We want to express the change of the total mass in $D \subset \mathbb{R}^N$ per unit time, i.e.

$$\frac{d}{dt}\int_D u(x,t)dx, \qquad dx = dx_1 \cdot \ldots \cdot dx_n$$

in terms of diffusion and reaction. We shall make the assumption that the particles obey the

Fick's Law *Particles move from high density place to a low density place.*

Special case: n=1

The balance equation in the interval $(x, x+dx)$ is per unit time

$$\underbrace{u_t(t,x)dx}_{\text{rate of total mass}} = \underbrace{\sigma(x+dx)u_x(t,x+dx) - \sigma(x)u_x(t,x)}_{\text{diffusion through endpoints}} + \underbrace{f(t,u,u,u_x)dx}_{\text{production rate}},$$

whence

$$u_t(x,t) = (\sigma(x)u_x)_x(t,x) + f(t,x,u,u_x). \tag{4}$$

<u>General case</u>

We consider a flux density $\vec{f_l} = -\sigma(x)\nabla u(x,t)$ where $\sigma(x) > 0$ is the diffusion coefficient. In accordance with Fick's law it points in the direction of the lowest density. The total flux through the boundary ∂D of D is

$$\int_{\partial D} (\vec{f_l}, \vec{n})dS, \quad \vec{n} \text{ outer normal, } (\cdot,\cdot) \text{ scalar product.}$$

The *balance law* reads as

$$\frac{d}{dt}\int_D u(x,t)dx = \oint_{\partial D} \sigma(x)\frac{\partial u}{\partial n}dS + \int_D f(x,t,u(x,t),\nabla u(x,t))dx,$$

where $f(x,t,u,\nabla u)$ is a measure for the production ($f > 0$) or absorption ($f < 0$) per unit volume element. The divergence theorem implies that

$$\oint_{\partial D} \frac{\partial u}{\partial n}dS = \int_D div(\sigma\nabla u)dx, \quad \frac{\partial}{\partial n} \text{ outer normal derivative.}$$

Inserting the last expression into the balance law we obtain

$$\int_D \frac{\partial u(x,t)}{\partial t} = \int_D \left[div(\sigma(x)\nabla u(x,t)) + f(x,t,u(x,t),\nabla u(x,t))\right] dx.$$

Since the choice of the region D is arbitrary, we deduce that the differential equation

$$\frac{\partial u(x,t)}{\partial t} = div(\sigma(x)\nabla u(x,t)) + f(x,t,u(x,t),\nabla u(x,t)) \tag{5}$$

holds for any (t,x).

Equation (5) is called *reaction diffusion equation*. Here $div(\sigma(x)\nabla u(x,t))$ is the diffusion term which describes the movement of the individuals, and $f(x,t,u(x,t),\nabla u(x,t))$ is the reaction term which describes the birth-death or reaction occurring inside the habitat or reactor.

Definition 4 *A reaction-diffusion equation of the form*

$$u_t = div(\sigma \nabla u) + f(x,t)$$

is called linear. If $f = 0$ is is homogeneous and otherwise inhomogeneous.

The diffusion coefficient $\sigma(x)$ is not a constant in general since the environment is usually heterogeneous. But when the region of the diffusion is approximately homogeneous, we can assume that $\sigma(x) \equiv \sigma_0$ and can be simplified to

$$\frac{\partial u}{\partial t} = \sigma_0 \Delta u + f(x,t,u,\nabla u)$$

where $\Delta u = div(\nabla u) = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2}$ is the *Laplacian operator*. When no reaction takes place, the equation becomes

$$\frac{\partial u}{\partial t} = \sigma_0 \Delta u$$

In classical mathematical physics, the equation $u_t = \Delta u$ is called heat equation, where u is temperature function. Therefore (5) is often called *nonlinear heat equation*.

Random walks

The heat equation can also be interpreted as a model for random walks. Suppose that a walker takes steps of lenght $\triangle x$ to the left or to the right along a line, and after each time units, the walker will be with probability 1/2 either at $x_0 - \triangle x$ or $x_0 + \triangle x$. If $P(t,x)$ denotes the number of walkers at time t and location x then

$$P(t_0 + \triangle t, x_0) = \frac{1}{2} P(t_0, x_0 + \triangle x) + \frac{1}{2} P(t_0, x_0 - \triangle x).$$

By the Taylor expansion:

$$P(t_0 + \triangle t, x_0) = P(t_0, x_0) + \frac{\partial P}{\partial t}(t_0, x_0)\triangle t + \frac{1}{2}\frac{\partial^2 P}{\partial t^2}(t_0, x_0)(\triangle t)^2 + \cdots,$$

$$\frac{1}{2} P(t_0, x_0 \pm \triangle x) = \frac{1}{2} P(t_0, x_0) \pm \frac{1}{2}\frac{\partial P}{\partial x}(t_0, x_0)(\triangle x) + \frac{1}{4}\frac{\partial^2 P}{\partial x^2}(t_0, x_0)(\triangle x)^2 + \cdots.$$

Inserting these expression into the above equation we find

$$\frac{\partial P}{\partial t}(t_0, x_0)\triangle t + \cdots = \frac{\partial^2 P}{\partial x^2}(t_0, x_0)(\triangle x)^2 + \cdots.$$

Here we have assumed that both $\triangle x$ and $\triangle t$ are small quantities. If in addition we assume that

$$\frac{(\triangle x)^2}{2\triangle t} \to \sigma_0 \quad \text{as} \quad \triangle t,\ \triangle x \to 0,$$

then we arrive at the one-dimensional diffusion equation

$$\frac{\partial P}{\partial t}(t,x) = \sigma_0 \frac{\partial^2 P}{\partial x^2}(t,x).$$

References

L.C.Evans, *Partial Differential Equations*, AMS (1998)
E. DiBenedetto, *Partial Differential Equations*, Birkhäuser (1995).

2.3 Simple observations

1. Consider the homogeneous diffusion equation $u_t = \Delta u$.

   - **Superposition principle**
     If u_1 and u_2 are solutions then $\alpha u_1 + \beta u_2$ is also solution
   - **translation invariance**
     $u(x + x_0, t + t_0) = v(x, t)$ is also a solution
   - **Scaling invariance**
     $u(\lambda x, \lambda^2 t) = v(x, t)$ is also a solution
   - If $u = u(x)$is independent of t, it is called *steady state* and satisfies $\Delta u = 0$. Such functions are called *harmonic functions*.
   - The function $v(t, x) = u(-t, x)$ solves the backward diffusion equation

     $$v_t + \Delta u = 0.$$

2. Consider the inhomogeneous reaction diffusion equation $u_t = \Delta u + f(x, t)$.

   - The general solution is of the form

     $$u(x, t) = u_h(x, t) + u_p(x, t),$$

     where u_h is the general solution of the homogeneous equation and u_p is a particular solution of the inhomogeneous equation.

Problems

1. Let $f(x) = 3x_1^2 + 5x_1x_2^3$
   Compute
   $$\nabla f(x), \quad \triangle f(x).$$

2. If $u(x, t)$ is a solution of the heat equation $u_t = \triangle u$, prove that $e^{\alpha t}u(x, t)$ is a solution of $v_t = \triangle v + \alpha v$.

3. If $u(x, t)$ solves $u_t = \triangle u$, what kind of equation does $v(x, t) = u(x, -t)$ solve?

4. Let $u(x, t) = t(x_1 + 2x_2)^3$ be a density function. Determine the direction of the flow according to Fick's law in $x = (1, 1)$.

3 Construction of a fundamental solution

Let us consider $u(x,t)$ solution to the equation $u_t = \Delta u$. We will seek for a solution of the form

$$v(\xi) \cdot t^{-\alpha}, \quad \text{where} \quad \xi = \frac{x}{\sqrt{t}} \in \mathbb{R}^N.$$

Substituting this into equation, we obtain

$$\begin{aligned}
&-\alpha t^{-\alpha-1} v + \nabla v t^{-\alpha} \left(-\frac{x}{2(\sqrt{t})^3} \right) = t^{-\alpha-1} \Delta v, \\
&\alpha v + \left(\nabla v, \frac{x}{2\sqrt{t}} \right) + \Delta v = 0, \\
&\alpha v + \frac{1}{2} (\nabla v, \xi) + \Delta v = 0.
\end{aligned}$$

Suppose now that $v(\xi) = v(\rho)$, where $\rho = |\xi|$, then

$$\Delta v = v'' + \frac{(N-1)v'}{\rho} = \frac{1}{\rho^{N-1}} (\rho^{N-1} v')',$$

hence, we obtain the following expressions

$$\begin{aligned}
&\alpha v + \frac{1}{2} v' \rho + \frac{1}{\rho^{N-1}} (\rho^{N-1} v')' = 0, \\
&(\rho^{N-1} v')' = \alpha v \rho^{N-1} + \frac{v'}{2} \rho^N.
\end{aligned}$$

Choose $\alpha = \frac{N}{2}$

$$\begin{aligned}
&\frac{1}{2} (v \rho^N)' + (\rho^{N-1} v')' = 0, \\
&\frac{1}{2} v \rho^N + \rho^{N-1} v' = a.
\end{aligned}$$

We assume further that $a = 0$, then

$$\frac{v'}{v} = -\frac{1}{2} \rho.$$

The solution of this equation is

$$v = C_0 e^{-\frac{1}{4}\rho^2},$$

and the solution of the diffusion equation is

$$u(x,t) = \frac{C_0}{t^{\frac{N}{2}}} e^{-\frac{|x|^2}{4t}}.$$

We want the following normalization

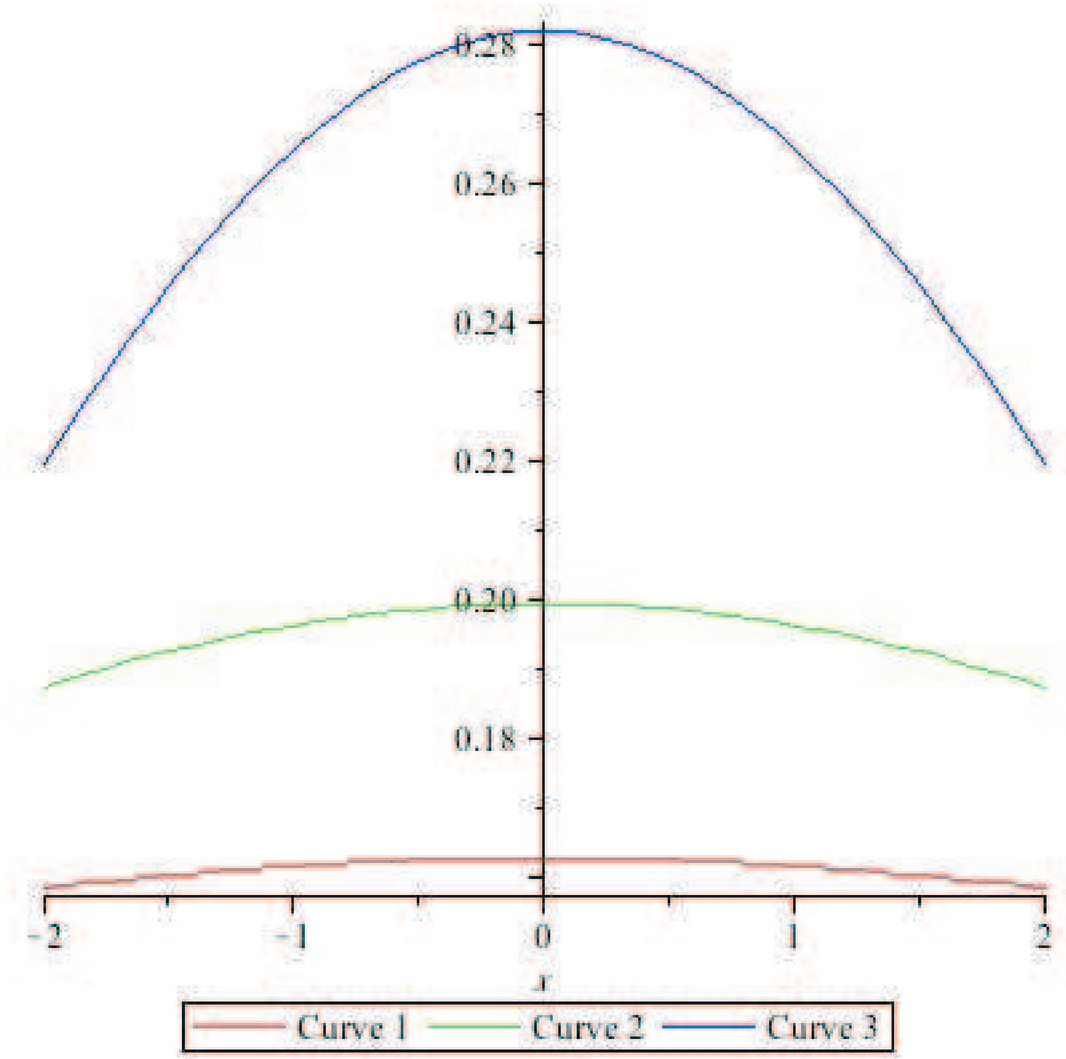
$$\int_{\mathbb{R}^N} u(x,t)dx = 1 \quad \Rightarrow \quad C_0 = \frac{1}{(4\pi)^{\frac{N}{2}}}.$$

Finally

$$\boxed{\Phi(t,x) = \frac{1}{(4\pi t)^{\frac{N}{2}}} e^{-\frac{|x|^2}{4t}}}$$

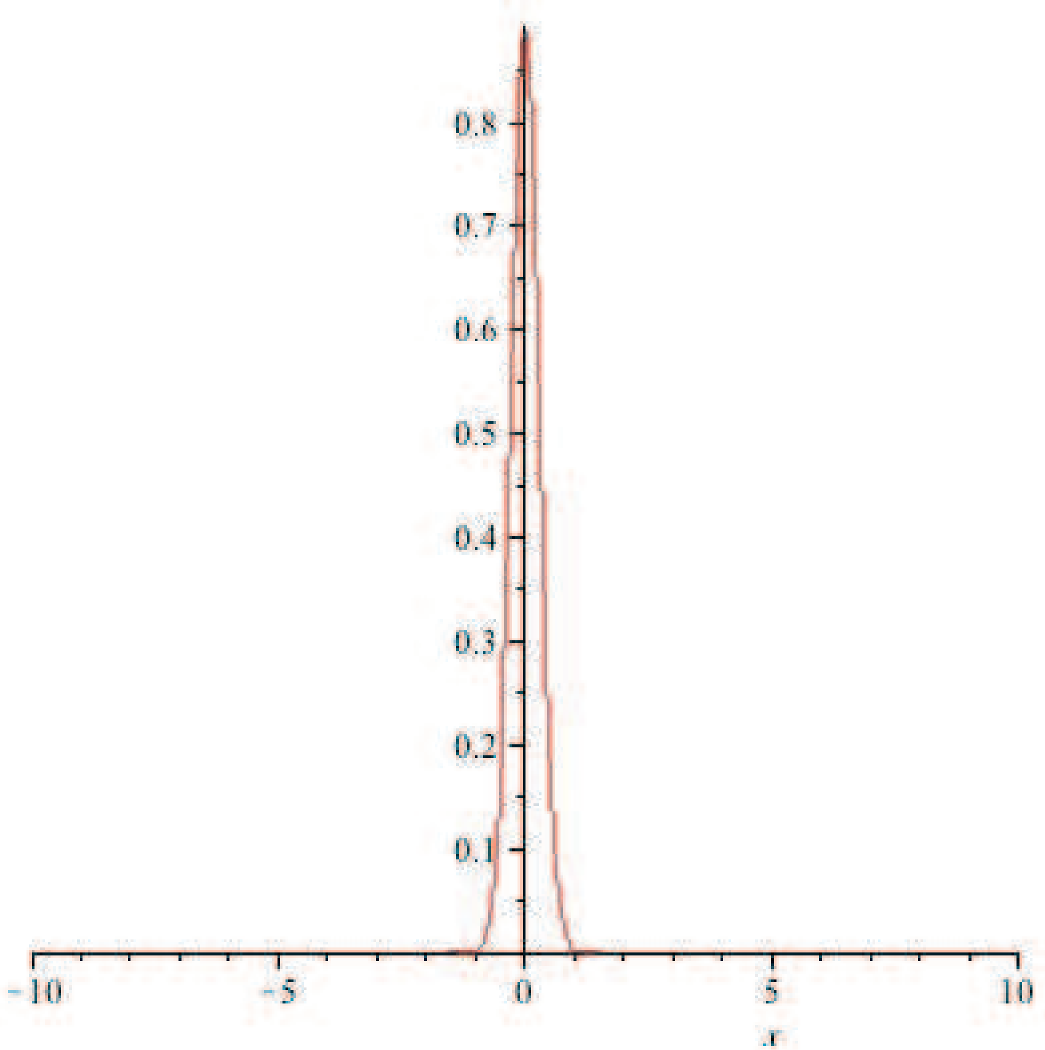
$\Phi(t,x)$ is called *Gaussian kernel*.

Important observation $\Phi(x,t)$ is the normal distribution with mean 0 and variance $\sigma^2 = 2t$



The figure describes the Gaussian kernel at times $t = 3, 2, 1$ (curve 1,2,3). If t is large the Gaussian kernel is flat.

If t decreases the function is more and more peaked (cf. figure) and tends to a Dirac function in the limit as $t \to 0$.



4 Initial value problem (Cauchy problem)

We consider the following initial value problem (IP)

$$\begin{cases} u_t = \Delta u \quad in \quad \mathbb{R}^N \times (0, \infty) \\ u = g(x) \quad in \quad \mathbb{R}^N \times \{t = 0\} \end{cases}$$

Because of the translation invariance $\Phi(x - y, t)$ is also a solution of the heat equation and by the superposition principle so is the Riemann sum

$$\sum_{i=1}^{n} \Phi(x - y_i, t) g(y_i) dy_i$$

for fixed dy_i, $i = 1 \cdots n$.. Now we can formulate the theorem

Theorem 5 *If $g \in C(\mathbb{R}^N) \cap L^\infty(\mathbb{R}^N)$ then*

$$u(x,t) = \int\limits_{\mathbb{R}^N} \Phi(x-y,t) g(y) dy$$

is a solution of (IP).

Proof Since $\Phi(x-y,t)$ is infinitely differentiable for $t > \delta > 0$ and since all its derivatives decrease very rapidly as $x, y \to \infty$ we can differentiate

$$\int\limits_{\mathbb{R}^N} \Phi(x-y,t) g(y) dy$$

under the integral. A straightforward computation, yields

$$u_t - \triangle u = \int_{\mathbb{R}^N} (\Phi_t - \triangle_x \Phi)(x-y,t)] g(y)\, dy.$$

Keeping in mind that Φ is a solution of the heat equation we can conclude that the same is true for u .
The fact that $\lim_{t\to 0} u(x,t) = g(x)$ is technically more involved. A detailed proof can be found e.g. in the book of Evans on Partial Differential Equations p. 47.

Problems

1. Carry out the details in the proof of Theorem 4.1:
   (a) $u \in C^\infty(\mathbb{R}^N \times \mathbb{R}^+)$.
   (b) u solves $u_t = \triangle u$
2. Construct a solution of (IP) in the case where $\triangle$ is replaced by $\sigma_0 \triangle$, $\sigma_0 =$const.

4.1 Nonhomogeneous problem

We consider problems of the following type

$$\begin{cases} u_t - \Delta u = f(x,t) \;\; in \;\; \mathbb{R}^N \times \mathbb{R}^+ \\ \quad u(x,0) = 0 \end{cases}$$

We have seen that

$$u = u(x,t,s) = \int\limits_{\mathbb{R}^N} \Phi(x-y,t-s) f(s,y) dy$$

is a solution of

$$\begin{aligned} u_t &= \Delta u \\ u(x,s) &= f(x,s). \end{aligned}$$

Duhamel's principle

Let

$$u(x,t) = \int_0^t u(x,t,s)ds$$

Then the following theorem holds.

Theorem 6 *If $f \in C$, $f_t \in C$, $f_{x_i,x_j} \in C$ then $u(x,t)$ is the solution of the inhomogeneous problem*

Proof Formal differentiation with respect to t implies

$$u_t(x,t) = u(x,t,t) + \int_0^t u_t(x,t,s)\, ds = f(x,t) + \int_{\mathbb{R}^N} \Phi_t(x-y,t-s)\, dy.$$

Here we have used the fact that for small time the Gaussian kernel behaves like a Dirac function,

$$\lim_{t\to 0} \int_{\mathbb{R}^N} \Phi(x,t) f(h(x)\, dx = f(0).$$

Moreover, since Φ satisfies the heat equation,

$$\int_{\mathbb{R}^N} \Phi_t(x-y,t-s)\, dy = -\int_{\mathbb{R}^N} \triangle_x \Phi_t(x-y,t-s)\, dy = -\triangle u.$$

The assumptions are used to guarantee that all integrals exist.

5 Separation of variables

We consider the following problem

$$u_t = u_{xx}$$

Further we look for solutions of this equation that can be represented as $u(x,t) = h(x)T(t)$. Then we obtain following relation

$$\frac{\dot{T}(t)}{T(t)} = \frac{h''(x)}{h(x)} = \alpha,$$

which yields

$$T(t) = e^{\alpha t}$$

- If $\alpha > 0$ then
  $h(x) = e^{\sqrt{\alpha}x}$ and the solution assumes the form

$$u(x,t) = e^{\sqrt{\alpha}x + \alpha t}$$

- if $\alpha < 0$ then the solution can be represented as

$$u(x,t) = e^{\alpha t}[a\cos(\sqrt{\alpha}x) + b\sin(\sqrt{\alpha}x)]$$

6 Boundary conditions

In this section n stands for the outer normal of a domain $D \subset \mathbb{R}^N$. The classical boundary conditions for reaction-diffusion equations are:

1. *Dirichlet's problem*
   - homogeneous problem $u(x,t) = 0$ in $\partial D \times \mathbb{R}^+$,
   - inhomogeneous problem $u(x,t) = \varphi(x,t)$ in $\partial D \times \mathbb{R}^+$.
2. *Neumann's problem*
   - homogeneous problem $\frac{\partial u}{\partial n} = 0$ in $\partial D \times \mathbb{R}^+$(isolation),
   - inhomogeneous problem $\frac{\partial u}{\partial n} = \varphi$ in $\partial D \times \mathbb{R}^+$.

Example
We consider the problem with homogeneous Dirichlet boundary conditions $u(0,t) = u(l,t) = 0$. A particular (separable) solution is

$$u(x,t) = e^{\alpha t}[a\cos(\sqrt{\alpha}x) + b\sin(\sqrt{\alpha}x)]$$

From the boundary conditions we get

$$\begin{aligned} u(0,t) &= 0 \quad\rightarrow\quad a = 0, \\ u(l,t) &= 0 \quad\rightarrow\quad \sqrt{\alpha} = \frac{\pi n}{e}, \end{aligned}$$

Then by the superposition principle

$$u(x,t) = \sum_{n=1}^{\infty} e^{-(\frac{\pi n}{l})^2 t} b_n \sin(\frac{\pi n}{l}x) \tag{6}$$

is also a (formal) solution. (Attention: The infinite series could diverge.) If we choose for b_n the Fourier coefficients of the initial condition $u_0(x)$

$$b_n = \frac{2}{l}\int_0^l u_0(x)\sin(\frac{\pi n}{l}x)dx$$

then the solution constructed in (6) solves (formally) the homogeneous Dirichlet problem with given initial conditions.

Remarks

In the 18th century it was believed that every continuous function has a Fourier series which converges pointwise to the original function. In 1829 Dirichlet established this assertion for Lipschitz functions. 1873 appeared the first example of a continuous function with no convergent Fourier series. In 1966 the Swedish mathematician L. Carlesson proved that a continuous function has almost everywhere a Fourier series.

Consequently (6) does not solve the Dirichlet problem in a classical sense without additional regularity on the initial conditions. A suitable frame for treating such problems is L^2. because every L^2-function has a convergent Fourier series (in the L^2-topology).

L^2- norm of $u(x,t)$:

$$\|u(\cdot,t)\|_{L^2}^2 := \int_0^l u^2(x,t)\,dx = \sum_{n=1}^{\infty} b_n^2 e^{-\left(\frac{\pi n}{l}\right)^2 2t}\frac{l}{2}.$$

Observe that $\|u(\cdot,t)\|_{L^2}$ always decreases as t increases.

6.1 Inhomogeneous problem

Let us consider problems of the following type

$$\begin{aligned}
&u_t = u_{xx} + h(x,t),\\
&u(0,y) = u(l,t) = 0,\\
&u(x,0) = 0,
\end{aligned}$$

where h is given by its Fourier series

$$h(x,t) = \sum h_n(t)\sin\left(\frac{\pi n}{l}x\right).$$

As in the method of the variation of constant we look for a solution of the form $u(x,t) = \sum_{n=1}^{\infty} e^{-\left(\frac{\pi n}{l}\right)^2 t} b_n(t)\sin(\frac{\pi n}{l}x)$. Inserting this expression into the equation we obtain, setting $\lambda_n = -\left(\frac{\pi n}{l}\right)^2$,

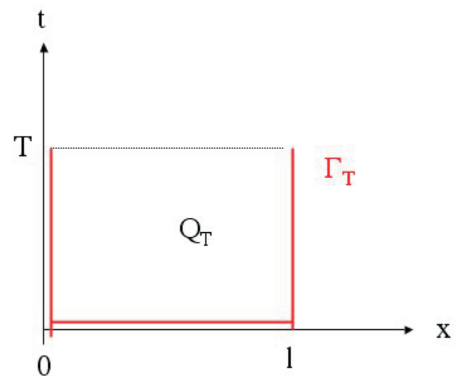
$$\begin{aligned}
&\sum \dot{b}_n(t)e^{-\lambda_n t}\sin\left(\frac{\pi n}{l}x\right) = \sum h_n(t)\sin\left(\frac{\pi n}{l}x\right)\\
&\rightarrow \dot{b}_n(t) = e^{\lambda_n t}h_n(t)\\
&b_n(t) = \int_0^l e^{\lambda_n s}h_n(s)ds
\end{aligned}$$

7 Maximum principle

We consider region $Q_T = (0,l) \times (0,T)$ and its parabolic boundary $\Gamma_T = (0,l) \times \{t = 0\} \cup \{x = 0\} \times (0,T) \cup \{x = l\} \times (0,T)$ and let u satisfy

$$u_t \le a u_{xx} + b u_x,$$

where $a > 0,\ b$ are continuous functions and u is differentiable in Q_T.



Theorem 7 *(Weak maximum principle)*

$$\max_{Q_T} u = \max_{\Gamma_T} u$$

Proof
(i) Suppose $u_t < a u_{xx} + b u_x$. If $\max_{Q_T} u(x,t) = u(x_0,t_0)$, where $(x_0,t_0) \in Q_T$ is some interior point, then $u_x = u_t = 0$ and from the differential inequality $u_{xx} > 0$ in (x_0,t_0). That is impossible because at the maximum $u_{xx} \le 0$. .
If $\max_{Q_T} u(x_0,t) = u(x_0,T)$, where $x_0 \in (0,l)$ then $u_t \ge 0$, $u_x = 0$ again we obtain a contradiction.
(ii) Let us consider now the case

$$u_t \le a u_{xx} + b u_x,$$

and denote $u_\epsilon := u - \epsilon t$, then

$$(u_\epsilon)_t < a(u_\epsilon)_{xx} + b(u_\epsilon)_x$$

by (i) $\max_{Q_T}(u - \epsilon t) = \max_{\Gamma_T}(u - \epsilon t)$ the assertion follows if we let $\epsilon \to 0$.

Remark 1 *If*

$$u_t \leq au_{xx} + bu_x,$$

then we cannot exclude from from the previous argument that the maximum lies both on Γ_T *and in* Q_T*.*

The next result clarifies the situation.

Theorem 8 *(Strong maximum principle)*
If u *attains its maximum at an interior point in* Q_T*, then* $u \equiv const.$

7.1 Applications

7.1.1 Uniqueness

(A). We consider the following problem

$$\text{(BVP1)} \begin{cases} u_t = au_{xx} + bu_x + f(x,t) \;\; in \;\; (0,l) \times \mathbb{R}^+ \\ u(0,t) = \varphi_1(t) \;,\; u(l,t) = \varphi_2(t) \\ u(x,0) = u_0(x) \end{cases}$$

Theorem 9 *(BVP1) has at most one solution.*

Proof
Let u_1 and u_2 be two solutions and $d(x,t) := u_1 - u_2$, then

$$d_t = ad_{xx} + bd_x \;\; in \;\; (0,l) \times \mathbb{R}^+$$

and $d(0,t) = d(l,t) = 0, \;\; d(x,0) = 0$. By the maximum principle

$$\max_{Q_T} d = \min_{Q_T} d = 0$$

(B). The next problem contains an additional linear term

$$\text{(BVP2)} \begin{cases} u_t = au_{xx} + bu_x + c(x,t)u + f(x,t) \;\; in \;\; (0,l) \times \mathbb{R}^+ \\ u(0,t) = \varphi_1(t) \;,\; u(l,t) = \varphi_2(t) \\ u(x,0) = u_0(x) \end{cases}$$

Theorem 10 *If* $|c(x,t)| \leq c_0$ *then there exist at most one solution to (BVP2).*

Proof
Let u_1 and u_2 be two solutions and $d(x,t) := u_1 - u_2$, then

$$d_t = ad_{xx} + bd_x + cd \;\; in \;\; (0,l) \times \mathbb{R}^+$$

Consider $e^{-c_0 t}d = \sigma$ then

$$\begin{aligned}\sigma_t &= -c_0 e^{-c_0 t}d + e^{-c_0 t}d_t \\ &= -c_0\sigma + a\sigma_{xx} + b\sigma_x + c\sigma \\ &= a\sigma_{xx} + b\sigma_x + (c - c_0)\sigma\end{aligned}$$

it is significant that $c - c_0 < 0$. The claim σ cannot have negative minimum in Q_T, however σ cannot reach its maximum in Q_T, consequently

$$\max_{Q_T}\sigma = \min_{Q_T}\sigma = 0$$

(C). Let us consider

$$(\text{BVP3})\begin{cases} u_t = au_{xx} + bu_x + c(x,t)u + f(x,t,u) \;\; in \;\; (0,l)\times\mathbb{R}^+ \\ u(0,t) = \varphi_1(t)\ ,\ u(l,t) = \varphi_2(t) \\ u(x,0) = u_0(x) \end{cases}$$

Theorem 11 *If f is differentiable in u and if $|f_u|$ is bounded then there is at most one solution to (BVP3).*

Proof Let u_i, $i = 1,2$ be two solutions. The difference d satisfies

$$\begin{gathered} d_t = ad_{xx} + bd_x + c(x,t)d + f_u(x,t,u_1 + \theta u_2)d \;\; in \;\; (0,l)\times\mathbb{R}^+ \\ d(0,t) = 0\ ,\ d(l,t) = 0 \\ d(x,0) = 0. \end{gathered}$$

If $|f'|$ is bounded then we can apply the previous result and conclude that $d = 0$.

Exercise Show that the theorem (10.5) remains valid for Lipschitz functions.

7.1.2 Estimates

Let

$$\begin{cases} \underline{u}_t \le a\underline{u}_{xx} + b\underline{u}_x + f \;\; in \;\; (0,l)\times\mathbb{R}^+ \\ \underline{u}(0,t) \le 0\ ,\ \underline{u}(l,t) \le 0 \\ \underline{u}(x,0) \le u_0(x) \end{cases}$$

Theorem 12 *$\underline{u}$ is called a subsolution of (BVP3). Analogously we define a supersolution $\overline{u}$ with reversed inequality signs .*

Theorem 13 *Let f be Lipschitz in u and let $\underline{u}$, $\overline{u}$ be sub and super solutions. Then*

$$\underline{u}(x,t) \le u(x,t) \le \bar{u}(x,t)$$

Example Consider the problem

$$u_t = u_{xx} + e^u \quad \text{in} \quad (0, \pi) \times \mathbb{R}^+,\ u(0,t) = u(\pi,t) = 0,\ u(x,0) = u_0(x).$$

If $u_0(x) \geq 0$, $\underline{u} = 0$ is a sub solution. As a super solution we take the solution to the ordinary differential equation $\dot{u} = e^u$, that is

$$\overline{u}(t) = \log\left[\frac{1}{e^{-\max u_0} - t}\right].$$

This super solution blows up at $t^* = e^{-\max u_0}$. It is interesting to observe that if $u_0 \geq 0$ the solution of the reaction-diffusion problem blows up in finite time. A way to see it, is to multiply (test) the equation with $\sin x$ and to integrate. Then

$$\int_0^\pi u_t \sin x \, dx = -\int_0^\pi u \sin x \, dx + \int_0^\pi e^u \sin x \, dx \tag{7}$$

Jensen's inequality implies

$$\int_0^\pi e^u \sin x \, dx \geq \frac{1}{2} e^{\int_0^\pi u \sin x \, dx}.$$

Put

$$\Phi(t) := \int_0^\pi u \sin x \, dx.$$

From (7) we deduce the differential inequality

$$\dot{\Phi} + \Phi \geq \frac{1}{2} e^{\Phi}.$$

An elementary argument shows, that Φ cannot exist for all t. It tends to infinity at some finite time t^*.

Exercise Prove the last statement.

References

M. H. Protter, H. F. Weinberger, *Maximum Principles in Differential Equations*, Prentice Hall (1967).

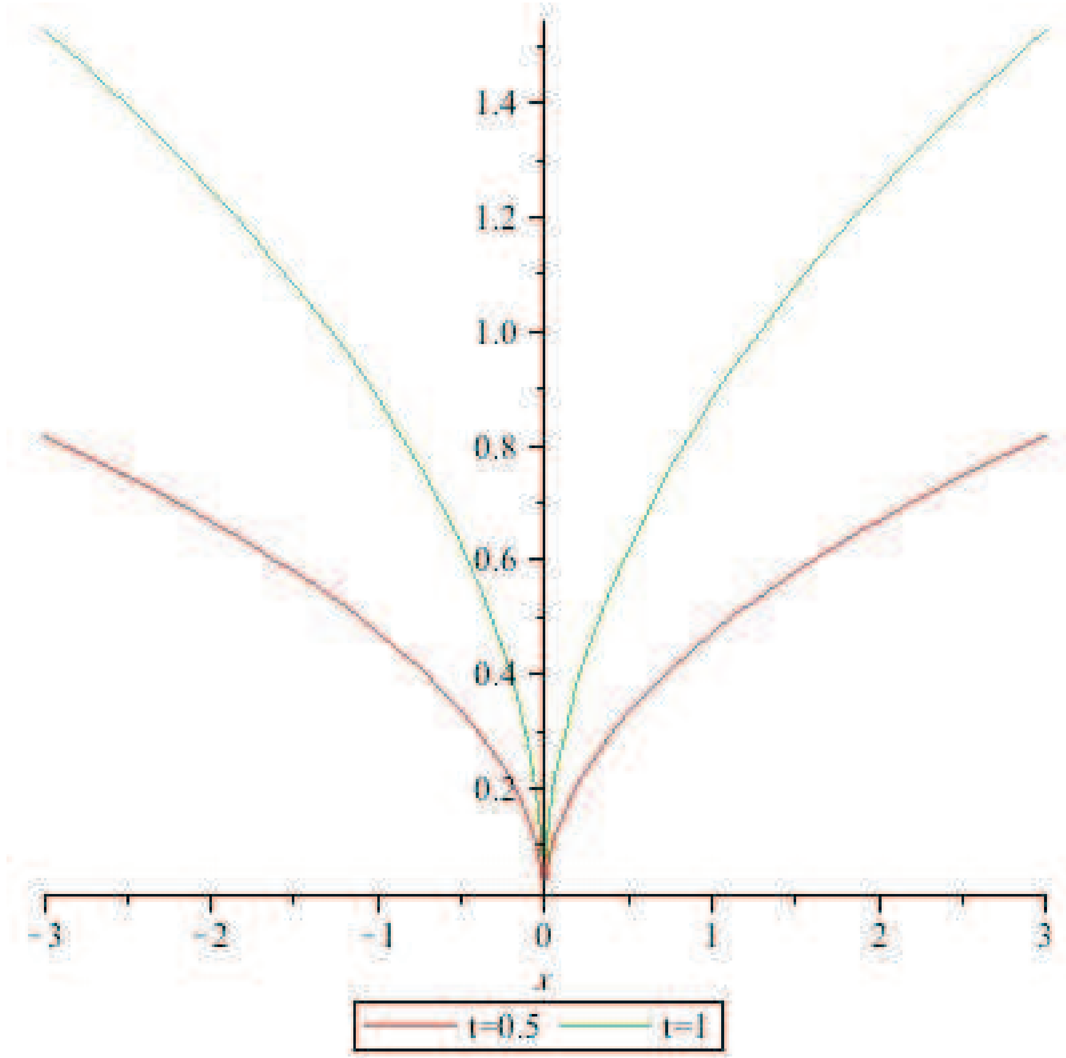
8 Viscosity solutions

The existence theory for parabolic equations fails if instead of the diffusion operator $\triangle$ we have $a(t,x)\triangle$ where $a(t,x)$ vanishes somewhere.

Example

$$u_t = x^2 u_{xx} \quad \text{in} \quad \mathbb{R} \times \mathbb{R}^+, \quad u(x,0) = |x|^\alpha, \quad 0 < \alpha < 1.$$

By means of separation of variables one obtains a formal solution of the form $u(x,t) = |x|^\alpha e^{\alpha(\alpha-1)t}$. Notice that for the particular choice of α this is not a solution in the classical sense because it is not differentiable in $x = 0$. The figure below shows the graph of $u(x, t)$ for $\alpha = 1/2$.



In order to include such less regular solutions, a different concept has been introduced which is described next.

Let us consider an equation of the general form

$$H(t, x, u, u_t, u_x, u_{xx}) = 0, \qquad (x, t) \in Q_T. \tag{8}$$

For simplicity we take $x \in \mathbb{R}$. It satisfies the *parabolicity condition*, in the sense that if $M \geq N$ then

$$H(t, x, u, p_1, p_2, M) \leq H(t, x, u, p_1, p_2, N) \quad \forall t, x, u, p_1, p_2.$$

Examples The reaction-diffusion equations satisfy the parabolicity condition.

Observation Let $u \in C^2$ be a classical solution of (8) and let $\varphi \in C^2$ be an arbitrary function of x and t. If (x_0, t_0) is a local maximum (minimum) of $u - \varphi$ then $u_t(x_0.t_0) = \varphi_t(x_0.t_0)$, $u_x(x_0.t_0) = \varphi_x(x_0.t_0)$ and $u_{xx}(x_0, t_0) - \varphi_{xx}(x_0.t_0) \leq 0$, (≥ 0). Consequently by the parabolicity condition

$$H(t_0, x_0, u(x_0, t_0), \varphi_t(x_0, t_0), \varphi_x(x_o, t_0), \varphi_{xx}(x_0, t_0)) \leq 0 \text{ (resp. } \geq 0).$$

If those two inequalities are satisfied for all φ then u is a solution (choose $\varphi = u$)

Definition 14 *(Continuous viscosity solution)*
If u is continuous and if for all $\varphi \in C^2$ such that $u - \varphi$ takes its maximum at (x_0, t_0) we have

$$H(t_0, x_0, u(x_0, t_0), \varphi_t(x_0, t_0), \varphi_x(x_o, t_0), \varphi_{xx}(x_0, t_0)) \leq 0$$

and if in addition, at a minimum point (x_1, t_1) of $u - \varphi$ we have

$$H(t_0, x_0, u(x_0, t_0), \varphi_t(x_0, t_0), \varphi_x(x_o, t_0), \varphi_{xx}(x_0, t_0)) \geq 0,$$

then $u(x, t)$ is called a continuous viscosity solution.

If u is locally bounded instead of continuous it is called a *discontinuous viscosity* solution.

Exercise Prove that the solution in Example 2 is a viscosity solution.

Boundary value problems

Consider the parabolic boundary value problem

$$H(t, x, u, u_t, u_x, u_{xx}) = 0 \text{ in } Q_T, \quad u = g \text{ on } \Gamma_T.$$

Here Q_T and Γ_T have the same meaning as in Sec. 7. For a viscosity solution satisfying the boundary conditions, the following inequalities must be satisfied:

If the maximum (minimum) point (x_0, t_0), $((x_1, t_1)$ lies on the parabolic boundary γ_T then

$$\min\{H(t, x, u, \varphi_t, \varphi_x, \varphi_{xx}), u - g|_{(x_0,t_0)}\} \leq 0 \quad \text{or}$$
$$\max\{H(t, x, u, \varphi_t, \varphi_x, \varphi_{xx}), u - g|_{(x_1,t_1)}\} \geq 0.$$

References

G. Barles, B. Perthame, *Exit time problems in optimal control and the vanishing viscosity method*, SIAM J. Control Optim., 26 (1988), 1133-1148.

M. G. Crandall, L. C. Evans and P.-L. Lions, *Some properties of viscosity solutions of Hamilton-Jacobi equations*, TRAMS 282 (1984), 487-502.

9 Parabolic equations of more general type

Let us consider a domain (=open connected set) D in $\mathbb{R}^N$, x is a generic point and $t > 0$ stands for the time. Moreover we introduce a symmetric positive $(N \times N)$-matrix A $(a_{ij} = a_{ji})$ such that

$$\sum a_{ij}\xi_i\xi_j \geq \lambda|\xi|^2, \quad \lambda > 0,$$

and functions b_i, $i = 1 \ldots N$ and c. Further we introduce the differential operator L given by

$$Lu := \sum_{i,j=1}^{N} (a_{ij}u_{x_j})_{x_i} + \sum_{i=1}^{N} b_i u_{x_i} + cu$$

Definition 15 *The equation*

$$u_t(x,t) = Lu + f \quad in \quad D \times \mathbb{R}^+ \tag{9}$$

is called a parabolic equation.

We distinguish between

- linear equations where

$$\begin{aligned} a_{ij} &= a_{ij}(x,t) \\ b_i &= b_i(x,t) \\ c &= c(x,t) \\ f &= f(x,t) \end{aligned}$$

- nonlinear equations where

$$\begin{aligned} a_{ij} &= a_{ij}(x,t,u,\nabla u) \\ b_i &= b_i(x,t,u,\nabla u) \\ c &= c(x,t,u,\nabla u) \\ f &= f(x,t,u,\nabla u) \end{aligned}$$

Examples of nonlinear equations: $N = 1$

1. $u_t = u_{xx} + au(1-u)$ Fisher's equation
2. $u_t = \frac{\partial}{\partial x}(u^m u_x)$ fast $(m < 1)$/slow $(m > 1)$ diffusion
3. $u_t = u_{xx} + uu_x$ Burger's equation

And we consider two types of problems

- *Cauchy problem*. In this case $D = \mathbb{R}^N$
  $u_t = Lu + f$ in $\mathbb{R}^N \times (0,T)$
  $u(x,0) = g(x)$ initial conditions (no boundary conditions)

- *Boundary value problem*. In this case $D \subset \mathbb{R}^N$
  $u_t = Lu + f$ in $D \times (0,T)$

$$\text{(BC)} \begin{cases} u = \varphi \quad on \quad \partial D \times (0,T) & \text{(Dirichlet problem)} \\ \dfrac{\partial u}{\partial n} = \varphi \quad on \quad \partial D \times (0,T) & \text{(Neumann problem)} \end{cases}$$

(IC) $u(x,0) = g(x)$

10 Weak solution

Let us consider the case $N = 1$ then we can rewrite equation (9) as

$$u_t = (au_x)_x + bu_x + cu + f.$$

Let $\varphi(x)$ be a smooth function then

$$\underbrace{\int\limits_0^l u_t \varphi dx}_{(u_t,\varphi)} = -\int\limits_0^l a\varphi_x u_x dx + \varphi a u_x|_0^l + \int\limits_0^l (bu_x + c)\varphi dx + \underbrace{\int\limits_0^l f\varphi dx}_{(f,\varphi)}$$

$$(u_t, \varphi) + \int\limits_0^l a\varphi_x u_x dx - \int\limits_0^l (bu_x + c)\varphi dx = \varphi a u_x|_0^l + (f, \varphi)$$

Suppose that $\varphi(0) = \varphi(l) = 0$. Then

$$(u_t, \varphi) + a(u, \varphi) = (f, \varphi),$$

where $a(u, \varphi)$ is a bilinear form. We want to interpret this equation in a Hilbert space.

10.1 Small excursion to functional analysis

Let us introduce some notation: $\mathcal{H}$ Hilbert space
$(\, , \,)$ scalar product
$\| \; \|$ norm

Example

We consider the space L^2 *with a scalar product* $(f,g) = \int\limits_0^l f \cdot g \; dx$ *and norm* $\|f\| = \left(\int\limits_0^l f^2 dx\right)^{\frac{1}{2}}$.

A *linear functional* $A : \mathcal{H} \to \mathbb{R}$ is a mapping such that

$$A(\alpha f + \beta g) = \alpha A f + \beta A g \quad \forall \alpha, \beta \in \mathbb{R}.$$

A is *bounded* if $\frac{\|Ax\|}{\|x\|} \leq \|A\|$ for any $x \in \mathcal{H}$

Theorem 16 *A linear operator is continuous if and only if it is bounded.*

10.2 Generalized derivatives

Let us consider a set $\Gamma = \{\varphi(x) \in C_0^\infty(0,l)\}$. An example of an element in Γ is $\varphi_0(x) = e^{-\frac{1}{(x-l/2)^2 - l^2/4}}$.

Exercise: Prove that φ_0 is infinitely often differentiable.

Definition 17 *Let* f , h *be integrable. If*

$$\int\limits_0^l f \varphi' dx = -\int\limits_0^l h \varphi dx, \quad \forall \varphi \in \Gamma$$

then h *is a generalized derivative of* f*.*

Generalized derivatives exist even if the function is not differentiable. Consider for instance the function $f = |x|$ in $(-1,1)$. we have

$$\int_{-1}^1 |x| \phi' \, dx = -\int_{-1}^0 x\phi' \, dx + \int_0^1 x\phi' \, dx = \int_{-1}^0 \phi \, dx + \int_0^1 \phi \, dx.$$

The generalized derivative is therefore

$$f'(x) = \begin{cases} -1 & \text{if } x < 0 \\ 1 & \text{if } x > 0 \end{cases}$$

Consider the Heaviside function

$$f(x) = \begin{cases} 0 & \text{in } (0, l/2) \\ 1 & \text{in } (l/2, l). \end{cases}$$

Then

$$\int_0^l f\phi' \, dx = \phi(l/2) \quad \forall \phi \in \Gamma.$$

We cannot associate to f' a function h but a linear functional $T_{f'}(\phi) = \phi(l/2)$.

Definition 18 *A distribution is a bounded linear functional* $T : \Gamma \to \mathbb{R}$.

The Heaviside function possesses a derivative in the distributional sense.

Definition 19 *The functional* $T_{\delta_y}\phi = \phi(y)$ *is called Dirac distribution.*

References

R. A. Adams, *Sobolev Spaces*, Academic Press (1975).

10.3 Sobolev spaces

Definition 20 *The Sobolev space* $\mathcal{H}^1(0,l)$ *is a Hilbert space consisting of the elements*

$$\{u : (0,l) \to \mathbb{R},\ u_x \in L^2(0,l)\}$$

and whose norm is defined as follows:

$$\|u\|_{\mathcal{H}^1} = \left(\int_0^l u_x^2 dx + \int_0^l u^2 dx \right)^{\frac{1}{2}}$$

We say that $u_n \to u$ in $\mathcal{H}^1$ if $\|u_n - u\|_{\mathcal{H}^1} \to 0$. The definition above generalises immediately to higher dimensions. The next Sobolev space takes into account that the functions vanish in a generalised sense at the boundary.

Definition 21 $\mathcal{H}_0^1(0,l) = \{u \in \mathcal{H}^1(0,l) : \exists \{u_n\}_{n=1}^\infty \subset C_0^\infty(0,l),\quad u = \lim_{n\to\infty} u_n \ \ in\ \ \mathcal{H}^1\}$.

$\mathcal{H}^1$ and $\mathcal{H}_0^1$ are Hilbert spaces.

Exercises

1. Prove that a function in $\mathcal{H}^1(0,l)$ is absolutely continuous.
2. Take $D = B_1(0)$ and $f = |x|^\alpha$. Prove that f belongs to $\mathcal{H}(B_1(0))$ if and only if $\alpha < \frac{N-2}{2}$.

10.4 Definition of weak solutions

Let us go back to the parabolic equation

$$\boxed{(u_t, \varphi) + a(u,\varphi) = (f,\varphi)} \quad \forall \varphi \in \mathcal{H}_0^1(0,l), \tag{8.1}$$

and let us switch viewpoint. We interprete

$$u : [0,T] \to \mathcal{H}_0^1(0,l)$$

and denote by u' the distributional derivative with respect to t.Then u is a weak solution of (8.1) if

$$u \in L^2(0,T;\mathcal{H}_0^1)$$

and

$$u_t = u' \in L^2(0,T;\mathcal{H}^{-1})$$

and (8.1) is satisfied. Here $\mathcal{H}^{-1}$ denotes the dual space of $\mathcal{H}_0^1(0,l)$. More precisely

Definition 22 *A function f belongs to the dual space, $\mathcal{H}^{-1}$ if there exists $f^0 \dots f^N$ in L^2 such that*

$$(f,v) = (f^0,v) + \sum_{i=1}^{N}(f^i, v_{x_i})$$

for any $v \in \mathcal{H}_0^1(D)$. The $\mathcal{H}^{-1}$ is called a dual space to $\mathcal{H}_0^1$.

The space $L^2(0,T;X)$, X Banach space, consists of elements $u:(0,T)\to X$ such that

$$\left(\int_0^T \|u\|_X^2 dt\right)^{\frac{1}{2}} < \infty.$$

11 Galerkin approximation

The Galerkin approximation provides a method to construct numerical solutions to the boundary value problem

$$u_t = (au_x)_x + bu_x + c + f \quad \text{in } (0,l)\times\mathbb{R}^+,\ u(0,t) = u(l,t) = 0, \quad u(x,0) = u_0(x).$$

It is an extension of the method of Fourier series considered in Sect. 6. We shall be looking for approximations to the weak solutions, Let $\{w_k\}_{k=1}^{\infty}$ be a basis of $\mathcal{H}_0^1(0,l)$ such that the scalar product in $L^2(0,l)$ satisfies $(w_k,w_j) = \delta_{kj}$. Set

$$u_m(x,t) = \sum_{k=1}^{m} d_m^k(t) w_k(x)$$
$$u_m(x,0) = \sum_{k=1}^{m} d_m^k(0) w_k(x)$$

where $d_m^k(0) = (u_0,w_k)$. We are looking for a solution of

$$(u_m', w_j) + a(u_m,w_j) = (f,w_j) \quad \forall j = 1\cdots m.$$

This equation can be seen as the projection of the original problem into the space generated by $\{w_k\}$. The above equation leads to a first order system of differential equations for the unknown $d_k(t)$, namely

$$
\begin{aligned}
&d_m^{j'}(t) + \sum_{k=1}^{m} d_m^k(t) a(w_k, w_j) = (f, w_j) \\
&d_m^j(0) = (u_0, w_j) \\
&j = 1 \ldots k
\end{aligned}
$$

We obtained the first order system of linear ODE's with constant coefficients consequently there exists an unique solution.

Remark This method is also used to prove analytically the existence of solutions to parabolic problems. It applies to very general problems and to higher dimensions.

References

O.A. Ladyzenskajia, N.N. Ural'tzeva, *Linear and Quasilinear Equations of Parabolic Type*, Transl. Math, Monographs, 23, AMS, Providence R.I., (1968)
A. Friedman, *Partial Differential Equations of Parabolic Type*, New York (1964).