Course Notes

Nonlinear Diffusion Problems

Odo Diekmann (Utrecht University)

Department of Mathematics
University of Helsinki

Fall 2014

Preface

When looking at the world, we see phenomena and wonder about the underlying mechanisms. Knowing our limitations, we isolate (either experimentally or conceptually) a small part of the world from the rest. For example, we focus our attention on the invasion of the cane toad in Australia. Or on cAMP signalling of Dictyostelium. Or on ... ~~We call~~ The abstraction of such an isolated small part of the world is called a system.

A system evolves in time and when our description takes this into account, we speak about a dynamical system. The precise mathematical definition of a dynamical system depends a bit on the context (e.g., when the aim is to control a system, input and output are incorporated in the definition). A key component of any definition is the concept of "state". The state of the system at a particular moment in time should summarize all information about the past history that is relevant for a prediction of the future (for plant breeding, the genotype of a plant qualifies as state; as demonstrated by Mendel's work, we may introduce this notion without knowing the physical/chemical hard-wiring

of it; also note that it may be hard or even impossible to determine the state in a non-destructive way). A dynamical system specifies how future states depend on an (assumed to be) known current state, so it specifies the dynamics of the state. In abstract mathematical theory, the specification is usually general and implicit. In concrete applications, it may be specific and explicit. As we hope to demonstrate in the current lecture notes, there is a lot of middle ground to be covered before these two extremes are connected. Sometimes we go top down, sometimes bottom up.

In this course we focus on nonlinear reaction-diffusion equations as they arise in the context of models from population biology. (Note that models from many other fields, e.g. combustion, lead to such equations too, but that we shall ignore these applications.) We want to know what kind of spatio-temporal patterns are exhibited by the solutions of such equations. Thus we deal with examples of infinite dimensional dynamical systems and study (in)stability and bifurcation of special solutions like steady states. In addition, we provide biological motivation, explain modelling assumptions and interpret mathematical results in order to deduce biological insights.

The following keywords provide a telegram style outline of the course:

- various derivations of the diffusion equation
- how far do you get? how long does it take?
- linear equations: fundamental solution, separation of variables, superposition, eigenvalues and eigenfunctions, dominant/principal eigenvalue, positivity
- nonlinear equations: steady states for one-dim space via phase plane analysis, minimal patch size, instability of non-constant (in space) steady states, steady states via bifurcation theory
- the maximum principle, sub- and super solutions
- ω-limit set, gradient system, Lyapunov function
- travelling plane wave solutions via phase plane analysis
- hairtrigger effect and asymptotic speed of propagation of an invading species
- pattern formation via Turing instability for activator-inhibitor systems
- chemotaxis, aggregation

Topics that maybe, but most likely are (in fact) not covered:

- forced speed (ecosystem in flowing river, climate change)
- excitability, nerve pulse propagation, FitzHugh-Nagumo, H-H
- spiral patterns in (bio)chemical reactions, Belousov-Zhabotinskii
- Evans function and stability of travelling waves & wave trains
- small parameter asymptotics, (geometric) singular perturbation
- onset of instability in unbounded domains: Ginzburg-Landau
- Hamilton-Jacobi equations via scaling
- cross diffusion
- ~~qou~~ quorum sensing of bacteria
- the human digestive tracle as an ecosystem

1. The dynamical systems perspective

1.1 The state space

The set X of all conceivable states of a system is called the state space (or, alternatively, the phase space). It may happen that X is a finite set. And it may happen that X is a closed subset of a Banach space (so then X is a complete metric space with distance $d(x,y) = \|x - y\|$; the reason to consider only a subset may be, for instance, that the elements are functions defined on a spatial domain and that we want to restrict attention to positive functions).

1.2 Time

We think of time as either discrete and advancing in leaps or as continuous and flowing. In the first case we use integers to label the moments in time, in the second case we use real numbers. In both cases the past is separated from the future by the present, i.e., the order has an interpretation, so carries meaning.

1.3 Describing how the state changes in the course of time

Suppose we know that at time t_0 the state is x_0. What can we deduce from this information about the state at other times?

A system is called autonomous if the value of t_0 is irrelevant (in the sense that our predictions for the state at time $t = t_0 + s$ depend on x_0 and s but not on t_0; so the system is not influenced by an external environment (fluctuating,). We now focus on autonomous systems and take, without loss of generality, $t_0 = 0$.

The notion of "state" entails that, in principle, the information that the state is x_0 at time 0 provides us with a probability distribution for the state at future times, so for $t > 0$ (for technical convenience consider X to be finite, so that it is clear what exactly "probability distribution" means).

For deterministic systems the information in fact fixes the state exactly for any $t>0$. So in that case there is a family of maps

$$\varphi^t : X \to X \tag{1.1}$$

parameterized by $t \geq 0$, such that $\varphi^t(x_0)$ is the state at time $t \geq 0$, given that the state at time 0 equals x_0. Consistency requires that

$$\varphi^0(x) = x \quad \text{for all } x \in X \tag{1.2}$$

$$\varphi^{t+s}(x) = \varphi^t(\varphi^s(x)) \quad \text{for } t,s \geq 0 \text{ and } x \in X \tag{1.3}$$

A collection of maps $\{\varphi^t\}_{t\geq 0}$ such that (1.1) - (1.3) hold is called a dynamical system on X (it may also happen that for a given x_0 there is a finite horizon in the sense that $\varphi^t(x_0)$ is only defined for $0 \leq t < T(x_0)$; this is usually called "blow up"; to facilitate the formulation we have excluded this situation).

If φ^t is invertible, we can reconstruct the past from the present. In general this is impossible. In particular this is impossible for time evolution governed by diffusion equations.

1.4 The generator

If time is discrete, (1.3) first implies

$$\varphi^2 = \varphi^1 \circ \varphi^1$$

and by induction that

$$\varphi^k = \underbrace{\varphi^1 \circ \varphi^1 \circ \cdots \circ \varphi^1}_{k\text{-times}}$$

In other words, the dynamical system is generated by φ^1 in the sense that, in order to find the state after k time steps, we have to iterate the map φ^1 k times. The change in the smallest possible time step is the one and only building block for the full story.

In continuous time, there is no smallest time step. The identity (1.2) states that we obtain only trivial (i.e., system unspecific) information when letting the time interval shrink to zero. However, we might compute the rate of change of the state, i.e.,

$$\lim_{t \downarrow 0} \frac{1}{t} \left(\varphi^t(x) - x \right)$$

As a rule, this limit does not exist for all $x \in X$, but still for a sufficiently large (i.e., dense) set of $x \in X$. If we call the limit $A(x)$, we have derived the abstract ordinary differential equation

$$\frac{dx}{dt} = A(x) \tag{1.4}$$

with initial condition

$$x(0) = x_0 \tag{1.5}$$

In fact we often proceed in the other direction. That is, we start from (1.4) – (1.5) and prove that a unique solution $x(t)$ exists, that we denote by $x(t;x_0)$ if we want to express the dependence on the initial condition. If we subsequently define

$$\varphi^t(x_0) = x(t;x_0) \tag{1.6}$$

it is immediate that (1.2) holds, while (1.3) follows from the uniqueness and the fact that there is no explicit t-dependence at the right hand side of (1.4). So the φ^t are the solution operators for the autonomous differential equation (1.4). We call A the <u>generator</u> of the dynamical system $\{\varphi^t\}$.

The description of a dynamical system by way of a differential equation, so by way of its generator, is strongly motivated by modelling considerations.

Many physical, chemical, biological, ..., laws give expressions for the rate of change in time of quantities that describe the state. Moreover, contributions of different mechanisms to the rate of change can simply be added. In contrast, contributions to changes over a finite time interval often interact in subtle ways.

1.5 Orbits

We call

$$\Gamma(x_0) = \{ x : x = \varphi^t(x_0) \text{ for some } t \geq 0 \} \tag{1.7}$$

the (forward) orbit through x_0. We call x_0 a fixed point (or, equilibrium, or steady state) if $\Gamma(x_0) = \{x_0\}$, i.e., $\varphi^t(x_0) = x_0$ for all $t \geq 0$. We call x_0 a periodic point of (minimal) period p if $\varphi^p(x_0) = x_0$ while $\varphi^t(x_0) \neq x_0$ for any $t \in (0, p)$. So the orbit through a periodic point is closed.

In principle we would like to understand how a dynamical system partitions the state space into orbits. In practice we rarely ever succeed. So often we concentrate on a neighbourhood of the various fixed points and periodic orbits that we can

find.

1.6 Stability

A subset S of X is called positively invariant if for any $x_0 \in S$ we have that $\Gamma(x_0) \subset S$. Fixed points as well as closed orbits are examples of positively invariant sets.

Definition A positively invariant set S is called stable if for any neighbourhood U of S there exists a neighbourhood V of S such that $\varphi^t(x_0) \in U$ for all $t > 0$ if $x_0 \in V$. And S is called asymptotically stable if, in addition, there exists a neighbourhood U of S such that $d(\varphi^t(x_0), S) \to 0$ as $t \to \infty$ (for every $x \in U$ we have that)
If we can take $U = X$ we call S globally asymptotically stable.

1.7 Linearization

Often it is possible to determine the stability character of a fixed point or a periodic point from an analysis of the linear problem obtained by linearization. The behaviour of the linear problem can often be ascertained by spectral analysis. In particular

the spectral bound, i.e., $\sup\{\mathrm{Re}\,\lambda : \lambda \in \sigma(A)\}$ ~~belongs~~ where $\sigma(A)$ is the spectrum of A and A is the linearized generator, is decisive: if the spectral bound is positive we have instability and if the spectral bound is negative we have, provided some additional conditions are fulfilled, stability.

1.8 Positivity

In many applications (and in particular in population biology), the variables need to assume positive values in order to be meaningful. So as a check on the correctness of the description of the generator by the right hand side of (1.4), one has to verify that positivity is preserved in time.

The linearization in the steady state of no population at all (called "trivial" by mathematicians) has itself an interpretation as a population model and hence also has this positivity preserving property. In the spectrum this manifests itself by the spectral bound being an eigenvalue, called the dominant eigenvalue, ~~with~~ characterized by the positivity of the corresponding eigenfunction (in the finite dimensional situation this is the so-called Perron-Frobenius theorem; the infinite dimensional analogue is called the Krein-

Rutman theorem).

If we linearize a scalar reaction-diffusion equation at whatever steady state, the maximum principle guarantees that positivity is preserved and hence we can focus our attention on the sign of the dominant eigenvalue.

1.9 Lack of compactness

The preceding discussion suggests that the growth away from an unstable steady state is exponential. For bounded spatial domains this is basically correct. But if the domain is very large (in a spatial scale yet to be specified), we might as well consider all of $\mathbb{R}^n$ (with $n=1, 2,$ or perhaps 3, as the only relevant values). A localized (in space) disturbance of a uniform (= constant in space) unstable steady state now expands as travelling wave fronts and the growth is characterized by the speed of these fronts (and not by an exponent).

2. On individuals and populations

(a primer on bookkeeping and the associated formula manipulation)

[See W. Feller, An Introduction to Probability Theory and Its Applications, Vol. II, Wiley, 1966, Chapter X for inspiration and notation.]

Let $\mathcal{X} \subset \mathbb{R}^n$ denote the i-state space (i for individual) You may (but need not) think of $x \in \mathcal{X}$ as spatial position. In the following Γ denotes a (measurable) subset of $\mathcal{X}$. We consider one individual (because we are interested in population biology, we talk about individuals and not, as physicists would do, about particles). The i-state at time t is denoted by $X(t)$ and is a random variable. The fundamental model ingredient is

$$Q_t(x, \Gamma) = \text{Prob}\{X(t+\tau) \in \Gamma \mid X(\tau) = x\} \tag{2.1}$$

which, by assumption, is independent of τ (in other words, the system is autonomous) and satisfies

$$Q_t(x, \mathcal{X}) \leq 1 \tag{2.2}$$

Note that $Q_t(x, \mathcal{X}) < 1$ means that the individual may die; so $X(t)$ denotes the i-state of an individual that happens to be alive at time t). We may, but need not, add the assumption

$$\lim_{t \downarrow 0} Q_t(x, \Gamma) = \delta_x(\Gamma) \tag{2.3}$$

Exercise Describe what this means

The assumption that x is the i-state corresponds to the assumption that the process is Markovian and necessitates that Q satisfies a consistency condition, the so-called Chapman-Kolmogorov relation

$$(2.4)\quad Q_{t+s}(x,\Gamma) = \int_{\Omega} Q_s(x,dz)\,Q_t(z,\Gamma) \qquad s,t>0$$

Exercise Spell out the meaning of (2.4) in words

For ease of formulation, take $\tau = 0$ in (2.1). Now suppose that we do not know $X(0)$ with certainty, but that we do know

$$(2.5)\quad \mu_0(\Gamma) = \text{Prob}\{X(0)\in\Gamma\}$$

If we define

$$(2.6)\quad \mu_t(\Gamma) = \text{Prob}\{X(t)\in\Gamma\}$$

then (2.1) implies that

$$(2.7)\quad \mu_t(\Gamma) = \int_{\Omega} \mu_0(dx)\,Q_t(x,\Gamma)$$

which we rewrite as

$$(2.8)\quad \mu_t = T(t)\mu_0$$

(in other words, (2.7) explains what we mean by $T(t)$)

Exercise Use (2.4) to prove that $T(t+s) = T(t)T(s)$

Sometimes the measure $Q_t(x,\cdot)$ has, for all $t>0$, $x\in\mathcal{S}$, a density. This means that u defined on $\mathcal{S}\times(0,\infty)\times\mathcal{S}$ exists such that

$$(2.9)\qquad Q_t(x,\Gamma)=\int_\Gamma u(x,t,y)\,dy$$

(here and in the following we do not formulate the measurability and integrability assumptions that are needed to make precise sense of the various formulas; these assumptions are implicit)

So $u(x,t,\cdot)$ is the density of the (possibly defective) probability distribution of i-state at time t, given that the i-state at time zero equals x.

Exercise Show that if (2.9) holds, the Chapman-Kolmogorov relation (2.4) can be written in the form

$$(2.10)\qquad u(x,t+s,y)=\int_{\mathcal{S}} u(x,s,z)\,u(z,t,y)\,dz$$

and state the meaning of (2.10) in words.

Exercise Assume (2.9). Show that μ_t, defined for $t>0$ by (2.7), has a density. Also show that if μ_0 has density ϕ then μ_t has density $T(t)\phi$, where

$$(2.11)\qquad (T(t)\phi)(x)=\int_{\mathcal{S}}\phi(y)\,u(y,t,x)\,dy$$

Finally, check that once again $T(t+s)=T(t)T(s)$

Our description above is in terms of one individual, but, acknowledging that life is not entirely predictable, we adopted a statistical description. The information provided by such a description becomes more and more accurate if we consider more and more individuals. Provided we assume that these individuals follow the process described by (2.1) independently of each other, we may conceive of those individuals as forming a population that develops in time. In other words, we can interpret $\mu_t(\Gamma)$ as the fraction of a very large population having, ~~an i-state~~ at time t, an i-state that belongs to Γ. Nota bene that if individuals may die, the population size gets smaller and $\mu_t(\Gamma)$ is more correctly described as the fraction, of the population that exists at time zero, that is still ~~alife~~ alive, and has an i-state, that belongs to Γ. (at time t) (at time t)

From the point of view of modelling, the formalism has two major weaknesses:

1. Only changes of i-state and survival are described, reproduction is not incorporated
2. Interaction is not incorporated (since independence is assumed)

Point 1 is taken care of in J. Math. Biol. (1998) 36: 349–388 and point 2 in J. Math. Biol. (2001) 43: 157–189

Both these papers have as main title "On the formulation and analysis of general deterministic structured population models", to which is added "I. Linear Theory" and "II. Nonlinear Theory", respectively. O. Diekmann, M. Gyllenberg, J.A.J. Metz and H.R. Thieme are authors of both, H. Huang and M. Kirkilionis are additional authors of the second paper.

The main idea of part II is to model interaction in a two step procedure by first explicitly introducing the environmental condition that characterizes the world in which an individual lives (so the system is no longer autonomous, since the rate of change of the i-state depends on the combination of the i-state and the prevailing environmental condition, but "environmental condition" should be chosen in such a way that, if we prescribe how the environmental condition changes over time, the individuals are independent of each other) and next to model how the environmental condition is influenced by feedback from the individuals. As an example, think of "food density" as the environmental condition and of consumption of food as the feedback. ~~(sorry for the somewhat awkward combination, if we you take it very literal~~ The second main example is "predation pressure", with falling victim to a predator as feedback. We might get back to this aspect of modelling when dealing with chemotaxis.

We now sketch the main results of part I, restricting ourselves to the autonomous (i.e., time translation invariant) situation. In the present course, submodels for reproduction will be rather rudimentary and therefore the rest of the text of this chapter is not part of the course material. Note that the notation differs from that in part I.

We supplement Q introduced in (2.1) by a second model ingredient

(2.12) $\Lambda_t(x,\Gamma)$ = the expected # of offspring (daughters) with i-state at birth in Γ produced in the time interval $[0,t)$ by an individual (mother) which at time zero has i-state x

The analogue of the Chapman-Kolmogorov relation (2.4) takes the form of a condition that combines the two ingredients, viz.

$$\Lambda_{t+s}(x,\Gamma) = \Lambda_s(x,\Gamma) + \int_{\Omega} Q_s(x,dz)\,\Lambda_t(z,\Gamma) \tag{2.13}$$

Exercise Spell out the meaning of (2.13)

Exercise Assume (2.9). Put

$$\Lambda_t(x,\Gamma) = \beta \int_{\Gamma[0,t)} u(x,\tau,y)\,d\tau\,dy$$

Check that (2.13) holds.
Also, describe in words the meaning of this expression for $\Lambda_t(x,\Gamma)$

Exercise Consider an individual that at time zero has i-state x. How many granddaughters will she get in $[0,t)$?
Do you agree that the answer is provided by

$$\int_{[0,t)\times\Omega} \Lambda_{ds}(x,dz)\,\Lambda_{t-s}(z,\Gamma) \qquad ? \tag{2.14}$$

Exercise Denote the expression in (2.14) by $\Lambda_t^2(x,\Gamma)$
Define by induction

$$\Lambda_t^{k+1}(x,\Gamma) = \int_{[0,t)\times\Omega} \Lambda_{ds}(x,dz)\,\Lambda_{t-s}^{k}(z,\Gamma) \tag{2.15}$$

Provide the interpretation of $\Lambda_t^k(x,\Gamma)$ and also of

(2.16) $$\Lambda_t^c(x,\Gamma) = \sum_{k=1}^{\infty} \Lambda_t^k(x,\Gamma)$$

where $\Lambda_t^1(x,\Gamma)$ is simply $\Lambda_t(x,\Gamma)$. Hint: c stands for clan

We call positive (= nonnegative) functions of the variables t, x, Γ, with $t>0$, $x \in \mathcal{R}$, $\Gamma \subset \mathcal{R}$, which are increasing in t kernels and define the product of two kernels by

(2.17) $$(\Phi \otimes \Psi)_t(x,\Gamma) = \int_{[0,t)\times\mathcal{R}} \Phi_{ds}(x,dz)\,\Psi_{t-s}(z,\Gamma)$$

Then $\Lambda^2 = \Lambda \otimes \Lambda$, $\Lambda^{k+1} = \Lambda \otimes \Lambda^k$. This product is associative: $\Theta \otimes (\Phi \otimes \Psi) = (\Theta \otimes \Phi) \otimes \Psi$ and

distributive: $$\Theta \otimes (\Phi + \Psi) = \Theta \otimes \Phi + \Theta \otimes \Psi$$

$$(\Phi + \Psi) \otimes \Theta = \Phi \otimes \Theta + \Psi \otimes \Theta$$

Instead of introducing the clan kernel (2.16) directly via (2.1b), we can also first formulate the Renewal Equation

(2.18) $$\Lambda^c = \Lambda + \Lambda^c \otimes \Lambda$$

(saying that a clan member is either a daughter (a daughter of) or a clan member ~~of a daughter~~) or, alternatively,

(2.19) $$\Lambda^c = \Lambda + \Lambda \otimes \Lambda^c$$

(saying that a clan member is either a daughter or a clan member of a daughter)

and then solve this equation by successive approximation, which in the present context is called generation expansion. This yields (2.16).

In the theory of RE (Renewal Equations) the kernel Λ^c is called the resolvent of the kernel Λ. The reason is, that, if we consider the equation

$$B = F + B \otimes \Lambda \tag{2.20}$$

where Λ and F are given and B is the unknown, then we have explicitly

$$B = F + F \otimes \Lambda^c \tag{2.21}$$

Indeed, if we multiply (2.20) from the right by Λ^c then for the second term at the right hand side we have

$$B \otimes \Lambda \otimes \Lambda^c \overset{(2.19)}{=} B \otimes (\Lambda^c - \Lambda) = B \otimes \Lambda^c - B \otimes \Lambda$$

The first of those two terms cancels the left hand side. Thus we deduce $0 = F \otimes \Lambda^c - B \otimes \Lambda$ which allows us to rewrite (2.20) as (2.21). We will need (2.21) below.

The kernels Λ and Λ^c describe the production of offspring in time and specify the i-state at birth of the offspring. But of course the i-state changes during the life of this offspring, as described by Q. If we want to describe the clan of an individual at a certain time in terms of the i-states of all the individuals at that time, we need to take these changes into account.

Definition (2.22)
$$Q^c_t(x,\Gamma) = Q_t(x,\Gamma) + \int_{[0,t)\times\mathbb{R}} \Lambda^c_{ds}(x,dz)\, Q_{t-s}(z,\Gamma)$$

Exercise Describe the meaning of Q^c, first in terms of one individual and her offspring and next, in the spirit of the first part of page 2.4, in terms of a large population. Do you agree that Q^c should satisfy the Chapman-Kolmogorov relation (2.4)?

In order to prove that Q^c does indeed satisfy the Chapman-Kolmogorov relation, it is helpful to first check that (2.13) also holds if we provide both Q and Λ with the index c.

Lemma
$$\Lambda^c_{t+s}(x,\Gamma) = \Lambda^c_s(x,\Gamma) + \int_{\mathbb{R}} Q^c_s(x,dz)\,\Lambda^c_t(z,\Gamma) \quad (2.23)$$

Proof We first write out the equation (2.18) with time $= t+s$ and then re-organize the equation such that it becomes (2.20) with time $= t$ and $B = \Lambda^c_s$. We then apply (2.21) and re-organize a bit more in order to arrive at (2.23)

⊛ $\Lambda^c_{t+s}(x,\Gamma) =$ (1) + (2) with

$$(1) = \Lambda_{t+s}(x,\Gamma) \overset{(2.13)}{=} \underbrace{\Lambda_s(x,\Gamma)}_{(3)} + \int_{\mathbb{R}} Q_s(x,dz)\,\underbrace{\Lambda_t(z,\Gamma)}_{(4)}$$

$$(2) = \int_{[0,t+s)\times\mathbb{R}} \Lambda^c_{d\sigma}(x,dz)\,\Lambda_{s+t-\sigma}(z,\Gamma) = \int_{[0,s)\times\mathbb{R}} \dots + \int_{[s,t+s)\times\mathbb{R}} \dots$$

$$\int_{[0,s)\times\mathbb{R}} \dots \overset{(2.13)}{=} \int_{[0,s)\times\mathbb{R}} \Lambda^c_{d\sigma}(x,dz)\Big\{\underbrace{\Lambda_{s-\sigma}(z,\Gamma)}_{(5)} + \int_{\mathbb{R}} Q_{s-\sigma}(z,dy)\,\underbrace{\Lambda_t(y,\Gamma)}_{(6)}\Big\}$$

$$(3)+(5) \overset{(2.18)}{=} \Lambda_s^c(x,\Gamma)$$

$$(4)+(6) = \int_{\mathbb{R}} \underbrace{\left\{ Q_s(x,dz) + \int_{[0,s)\times\mathbb{R}} \Lambda_{d\sigma}^c(x,dy)\, Q_{s-\sigma}(y,dz) \right\}}_{\overset{(2.22)}{=} Q_s^c(x,dz)} \Lambda_t(z,\Gamma)$$

replace integration variable y by z and vice versa

Define $F_t(x,\Lambda) = \Lambda_s^c(x,\Gamma) + \int_{\mathbb{R}} Q_s^c(x,dz)\,\Lambda_t(z,\Gamma)$

$= (3)+(4)+(5)+(6)$

and

$$B_t(x,\Gamma) = \Lambda_{t+s}^c(x,\Gamma)$$

Note that $\int_{[s,t+s)\times\mathbb{R}} \dots = \int_{[0,t)\times\mathbb{R}} B_{d\sigma}(x,dz)\,\Lambda_{t-\sigma}(z,\Gamma)$

So we can rewrite the identity $\circledast$ as

$$B = F + B \otimes \Lambda$$

It follows that $B = F + F \otimes \Lambda^c$ (↑(2.21)) or, written out in detail

$$B_t(x,\Gamma) = F_t(x,\Gamma) + \int_{[0,t)\times\mathbb{R}} F_{d\sigma}(x,dz)\,\Lambda_{t-\sigma}^c(z,\Gamma)$$

Since the first term in the defining expression of F is independent of t, we have [once more switch $z \to y$ as integration variable]

$$F_{d\sigma}(x,dz) = \int_{\mathbb{R}} Q_s^c(x,dy)\,\Lambda_{d\sigma}(y,dz)$$

Hence

$$B_t(x,\Gamma) = F_t(x,\Gamma) + \int_{\mathbb{R}} Q_s^c(x,dy) \int_{[0,t)\times\mathbb{R}} \Lambda_{d\sigma}(y,dz)\,\Lambda_{t-\sigma}^c(z,\Gamma)$$

$$\overset{(2.19)}{\downarrow}$$

$$= F_t(x,\Gamma) + \int_{\Omega} Q_s^c(x,dy)\left\{\Lambda_t^c(y,\Gamma) - \Lambda_t(y,\Gamma)\right\}$$

$$= \Lambda_s^c(x,\Gamma) + \int_{\Omega} Q_s^c(x,dy)\,\Lambda_t^c(y,\Gamma) \qquad \square$$

Theorem $Q_{t+s}^c(x,\Gamma) = \int_{\Omega} Q_s^c(x,dz)\,Q_t^c(z,\Gamma)$

Proof From (2.22) we have

$$Q_{t+s}^c(x,\Gamma) = \underset{(1)}{Q_{t+s}(x,\Gamma)} + \underset{(2)}{\int_{[0,t+s)\times\Omega} \Lambda_{d\sigma}^c(x,dz)\,Q_{t+s-\sigma}(z,\Gamma)}$$

$$(1) \overset{(2.4)}{=} \int_{\Omega} Q_s(x,dy)\,Q_t(y,\Gamma) \qquad (2) = \underset{(3)}{\int_{[0,s)\times\Omega} \cdots} + \underset{(4)}{\int_{[s,t+s)} \cdots}$$

$$(3) \overset{(2.4)}{=} \int_{[0,s)\times\Omega} \Lambda_{d\sigma}^c(x,dz) \int_{\Omega} Q_{s-\sigma}(z,dy)\,Q_t(y,\Gamma)$$

$$\text{So } (1)+(3) = \int_{\Omega}\left\{Q_s(x,dy) + \int_{[0,s)\times\Omega} \Lambda_{d\sigma}^c(x,dz)\,Q_{s-\sigma}(z,dy)\right\} Q_t(y,\Gamma)$$

$$\overset{(2.22)}{=} \int_{\Omega} Q_s^c(x,dy)\,Q_t(y,\Gamma)$$

$$(4) \overset{\tau = \sigma - s}{=} \int_{[0,t)\times\Omega} \Lambda_{s+d\tau}^c(x,dz)\,Q_{t-\tau}(z,\Gamma)$$

$$\overset{(2.23)}{=} \int_{[0,t)\times\Omega} \int_{\Omega} Q_s^c(x,dy)\,\Lambda_{d\tau}^c(y,dz)\,Q_{t-\tau}(z,\Gamma)$$

Hence $Q^c_{t+s}(x,\Gamma) = (1)+(3)+(4) =$ (2.12

$$\int_{\Omega} Q^c_s(x,dy)\left\{ Q_t(y,\Gamma) + \int_{[0,t)\times\Omega} \Lambda^c_{d\tau}(y,dz)\, Q_{t-\tau}(z,\Gamma)\right\}$$

(2.22)
$$\stackrel{\downarrow}{=} \int_{\Omega} Q^c_s(x,dy)\, Q^c_t(y,\Gamma) \qquad \square$$

Conclusion A general linear autonomous population model requires two ingredients, Q and Λ, that satisfy (2.4) and (2.13). We first define Λ^c by (2.16) or, in other words, we first solve (2.18) and, equivalently, (2.19). Next we define Q^c by (2.22). Subsequently we define a semigroup of operators that describes the population development by the analogue of (2.8) with Q in (2.7) replaced by Q^c

3. A first derivation of the diffusion equation

Brownian motion refers to the movement of pollen grains in water, as observed under a microscope by the English botanist R. Brown in 1928. The underlying physical mechanism is that a single grain has very many collisions with randomly moving water molecules. The mathematical theory was initiated by Einstein in 1905 and subsequently a more rigorous mathematical foundation and description was provided by N. Wiener in 1923. The derivation of diffusion equations as presented below goes back to A.N. Kolmogorov (1931) and W. Feller (1936). Let $S = \mathbb{R}$.

In terms of the fundamental model ingredient $Q_t(x, \Gamma)$ introduced in (2.1), we <u>postulate</u> that for all $\varepsilon > 0$ as $h \downarrow 0$

i) $$\frac{1}{h} \int_{|y-x| \geq \varepsilon} Q_h(x, dy) \longrightarrow 0 \qquad (3.1)$$

ii) $$\frac{1}{h} \int_{|y-x| < \varepsilon} (y-x) Q_h(x, dy) \longrightarrow b(x) \qquad (3.2)$$

iii) $$\frac{1}{h} \int_{|y-x| < \varepsilon} (y-x)^2 Q_h(x, dy) \longrightarrow a(x) \qquad (3.3)$$

(note: since i) holds for all $\varepsilon > 0$, we may in ii) and iii) replace ε by any fixed quantity; under regularity/smoothness conditions on Q, ii) and iii) are a consequence of i))

For a given bounded continuous function on Ω we define

$$v(t,x) = \int_\Omega Q_t(x,dy)\,\psi(y) \tag{3.4}$$

[So $v(t,x)$ is the expectation of $\psi(X(t))$, given $X(0)=x$]

The Chapman-Kolmogorov relation (2.4) implies that

$$v(t+s,x) = \int_\Omega Q_s(x,dz)\,v(t,z) \tag{3.5}$$

So provided $Q_t(x,\Omega)=1$ we can write

$$\frac{v(t+h,x)-v(t,x)}{h} = \frac{1}{h}\int_\Omega Q_h(x,dy)\,[v(t,y)-v(t,x)] \tag{3.6}$$

We want to determine the limit as of the right hand side of (3.6) for $h \downarrow 0$. The postulates tell us that, for computing the limit, we may replace Ω by a ball of radius δ centered at x, where we can choose $\delta>0$ as we like. We write, assuming that $v(t,\cdot)$ has two bounded continuous derivatives

$$v(t,y)-v(t,x) = (y-x)\frac{\partial v}{\partial x}(t,x) + \tfrac{1}{2}(y-x)^2\frac{\partial^2 v}{\partial x^2}(t,x) + R(t,x,y)$$

and for any $\varepsilon>0$ choose $\delta=\delta(\varepsilon,x)>0$ such that

$$|R(t,x,y)| < \varepsilon\,|y-x|^2 \quad \text{for } |y-x|\le\delta$$

Using the postulates, we deduce that for small $h>0$

$$\left|\frac{1}{h}\int_\Omega Q_h(x,dy)\,R(t,x,y)\right| < \varepsilon\,a(x)$$

Since $\varepsilon>0$ is arbitrary, $\lim_{h\downarrow 0}\frac{1}{h}\int_\Omega Q_h(x,dy)\,R(t,x,y)=0$.

So by taking the limit $h\downarrow 0$ in (3.6) we find the partial differential equation (pde)

$$\frac{\partial v}{\partial t} = \frac{1}{2} a \frac{\partial^2 v}{\partial x^2} + b \frac{\partial v}{\partial x} \tag{3.7}$$

aka the (Kolmogorov) backward equation.

So starting from Q we derived, under certain assumptions but still in a rather formal manner, the diffusion equation (3.7). Recall the observation made on page 1.5 : In fact we often proceed in the other direction. That is, we use modelling considerations to come up with a and b and next solve (3.7) in order to define and investigate Q.

The variable x in (3.7) corresponds to the initial position at time zero and not to (the distribution of) the current position at time t. But, as we now show, information can be transferred by duality, i.e., by taking adjoints. To streamline the presentation, assume that Q has a density, i.e., assume that (2.9) holds. This allows us to rewrite (3.4) in the form

$$v(t,x) = \int_{\mathcal{R}} u(x,t,y)\, \psi(y)\, dy \tag{3.8}$$

Let ϕ be any integrable function defined on $\mathcal{R}$. Define a pairing between integrable functions and continuous functions by

$$\langle \phi, \psi \rangle = \int_{\mathcal{R}} \phi(z)\, \psi(z)\, dz \tag{3.9}$$

Remark True duality ~~is between L_1 and~~ leads from L_1 to L_∞ and from continuous functions to measures. The smoothing properties of the diffusion equation allow us to focus on elements of L_∞ that are continuous

and to focus on measures that are absolutely continuous (i.e., have a density), without losing any information.

End of Remark

Then

$$\langle \phi, v(t,\cdot)\rangle = \int_{\Omega} \phi(z) \int_{\Omega} u(z,t,y)\,\psi(y)\,dy\,dz$$

$$= \int_{\Omega} \left(\int_{\Omega} \phi(z) u(z,t,y)\,dz \right) \psi(y)\,dy$$

$$= \int_{\Omega} w(t,y)\,\psi(y)\,dy = \langle w(t,\cdot), \psi\rangle$$

where

$$w(t,y) = \int_{\Omega} \phi(z)\,u(z,t,y)\,dz \tag{3.10}$$

Recalling (2.11) we write (3.10) as $w(t,\cdot) = T(t)\phi$. If we similarly write (3.8) as $v(t,\cdot) = T^*(t)\psi$, then the identity above assumes a form that might be familiar from functional analysis

$$\langle \phi, T^*(t)\psi\rangle = \langle T(t)\phi, \psi\rangle \tag{3.11}$$

The obvious next question is: how does this identity translate to the level of the infinitesimal generator?

The right hand side of (3.7) defines in general only partly the unbounded operator that generates the semigroup of operators $T^*(t)$, because for unbounded operators one needs to specify the domain of definition in addition to (or, in fact, before) specifying the action.

And an important part of the specification of the domain is formed by boundary conditions. All of this was clarified by W. Feller in the famous paper "The parabolic differential equations and the associated semi-groups of transformations", Annals of Mathematics (1952) 55 (3): 468-519. An extensive treatment of the multi-dimensional case is presented in the recent book: K. Taira, Semigroups, Boundary Value Problems and Markov Processes, Springer 2014

If we multiply (3.7) by ϕ, integrate, the right hand side by parts, and assume boundary terms vanish, we arrive at, after renaming the variable (in order) to stay in line with (3.10)

$$\frac{\partial w}{\partial t}(t,y) = \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(a(y)w(t,y)\right) - \frac{\partial}{\partial y}\left(b(y)w(t,y)\right) \tag{3.12}$$

which is called the (Kolmogorov) forward equation (but also the Fokker-Planck equation). The vanishing of the boundary terms may either be related to the behaviour of the coefficients a and b at the boundary or to conditions imposed on the solutions of (3.7) and (3.12). If we put $b=0$ and $a=2$, the equations are in fact identical and, assuming $\mathcal{R} = [-1,+1]$, the boundary terms are given by

$$\left[\phi\psi' - \phi'\psi\right]_{-1}^{+1}$$

If we impose that the functions themselves are zero at the boundary, we call the boundary an absorbing barrier and interpret this as the death of the individual

upon reaching the boundary. As a result $Q_t(x, \mathbb{R}) < 1$ for $t > 0$, i.e., the probability distribution is defective. If we impose that the derivatives are zero at the boundary, we call the boundary a reflecting barrier and interpret this as the bouncing back of an individual upon hitting the boundary. In this case $Q_t(x, \mathbb{R}) = 1$ for $t > 0$. Of course it may happen that $x = 1$ is absorbing while $x = -1$ is reflecting (or the other way around).

Summary If individuals move, independently of each other, randomly in space and some conditions are satisfied, the generator takes the form of a diffusion operator, with boundary conditions incorporated in the domain of definition of the operator. In fact there are two forms, forward and backward, related to each other by duality. Often we motivate the diffusion operator, in particular the coefficients and the boundary conditions, directly in terms of modelling considerations and next constructively define solutions for the partial differential equation with boundary conditions. As long as we stay in the linear realm, we can either interpret the solution as describing the (possibly defective) probability distribution of the position of one individual or as ~~the~~ a description of how a large population of individuals is distributed in space.