

**Identifiability and sensitivity analysis of heterogeneous
cell population models**

Master Thesis

Shen Zeng

May 31, 2013

Betreuer:

Jun.-Prof. Dr.-Ing. Steffen Waldherr

Prof. Dr.-Ing. Frank Allgöwer

Institut für Systemtheorie und Regelungstechnik
Universität Stuttgart
Prof. Dr.-Ing. Frank Allgöwer

Contents

Abstract	5
Deutsche Zusammenfassung	7
1. Introduction	9
1.1. Motivation	9
1.2. Focus of this thesis	10
1.3. Outline and contributions	11
1.3.1. Modeling of heterogeneous cell populations	11
1.3.2. Parameter estimation and identifiability	12
1.3.3. Sensitivity analysis	12
I. Preliminaries	15
2. Models for heterogeneous cell populations and the direct problems	17
2.1. Monte Carlo-based modeling	18
2.2. Density-based modeling	21
2.2.1. The advection equation	22
2.2.2. Method of characteristics	24
2.3. Discussion	29
3. Frobenius-Perron operators and the advection equation	31
3.1. An overview of properties of the advection equation	31
3.2. Frobenius-Perron operators	32
3.3. Frobenius-Perron operators and the advection equation	36
4. Parameter estimation	39
4.1. Cell populations with output	39
4.2. The parameter estimation problem	41

II. Structural identifiability and sensitivity analysis	43
5. Structural identifiability	45
5.1. Motivation	45
5.2. Criteria for structural identifiability	46
5.2.1. An illustrative example	46
5.2.2. Necessary condition	49
5.2.3. Discussion on sufficient condition	50
6. Sensitivity analysis	53
6.1. Sensitivity analysis and the Frechet derivative	53
6.2. Sensitivity analysis for heterogeneous cell populations models	55
6.3. Output sensitivity	60
6.4. Output sensitivity and structural identifiability	62
7. Conclusions	65
7.1. Summary and discussion	65
7.2. Outlook	66
A. Mathematical background	69
A.1. Dynamical systems	69
A.2. Probability theory	72
A.3. Multivariable calculus	74
B. Matlab functions	77
B.1. Example Monte Carlo method	77
B.2. Example method of characteristics	78
B.3. Example sensitivity analysis	81
Bibliography	83

Abstract

In this thesis, we introduce novel concepts to the modeling and analysis of heterogeneous cell populations. Heterogeneous cell populations can be interpreted as large populations of structurally identical cells with *heterogeneous* parameters and initial conditions. They appear in biological systems such as tissues of higher organisms or colonies of microorganisms [1].

A well-known approach for the modeling of heterogeneous cell populations is the so called density-based approach, in which the state of a heterogeneous cell population is given by the probability density of the cell states. The evolution of the probability densities is in this approach given in terms of a partial differential equation. We extend this approach via a measure theoretical consideration, which exploits the probabilistic nature of the problem. The result of this novel ansatz is a framework in which the evolution of densities is described by operators.

One of the key tasks in the analysis of heterogeneous cell population models is parameter estimation. For heterogeneous cell populations we want to estimate the probability density of parameters and initial conditions. However, to be able to perform parameter estimation, one always needs specific identifiability properties of a system. We formulate for the first time the concept of structural identifiability of a heterogeneous cell population model. It is revealed that this concept is closely related to observability of the corresponding single cell model. The connection between both concepts is studied and illuminated in a concrete example.

The second emphasis of this thesis is the implementation of sensitivity analysis to the class of heterogeneous cell population models. Here we study sensitivity with respect to variations or misspecifications in the probability density of parameters and initial conditions.

Deutsche Zusammenfassung

In dieser Arbeit führen wir neue Konzepte zur Modellierung und Analyse von heterogenen Zellpopulationen ein. Heterogene Zellpopulationen können als große Populationen von strukturell identischen Zellen mit heterogenen Parametern und Anfangswerten interpretiert werden. Solche Populationen findet man in biologischen Systemen, wie etwa Gewebe von höheren Organismen oder Kolonien von Mikroorganismen [1].

Ein bekannter Ansatz für die Modellierung von heterogenen Zellpopulationen ist der sogenannte dichte-basierte Ansatz, in welchem der Zustand einer heterogenen Zellpopulation gegeben ist durch die Wahrscheinlichkeitsdichte der Zellzustände. Die zeitliche Evolution der Wahrscheinlichkeitsdichten wird in diesem Ansatz durch eine partielle Differentialgleichung beschrieben. Wir erweitern die bisherigen Ergebnisse mit einer maßtheoretischen Betrachtung, welche die probabilistische Natur des Problems ausnutzt. Das Ergebnis dieses neuen Ansatzes ist ein Rahmen, in dem die Evolution der Wahrscheinlichkeitsdichte durch Operatoren beschrieben sind.

Eine Hauptaufgabe in der Analyse von heterogenen Zellpopulationen ist die Parameterschätzung. Hierbei ist es das Ziel die Wahrscheinlichkeitsdichte von Parametern und Anfangswerten zu schätzen. Als Voraussetzung von Parameterschätzung sind jedoch immer bestimmte Identifizierbarkeitseigenschaften eines Systems notwendig. Wir formulieren erstmalig das Konzept der strukturellen Identifizierbarkeit von heterogenen Zellpopulationsmodellen, welches eng verbunden ist mit der Beobachtbarkeit des zugehörigen Einzelzellmodells. Wir studieren die Verbindung zwischen den beiden Konzepten und zeigen diese in einem anschaulichen Beispiel auf.

Der zweite Schwerpunkt dieser Arbeit ist die Implementierung der Sensitivitätsanalyse auf die Klasse der heterogenen Zellpopulationsmodellen. Wir studieren hierbei die Sensitivität bezüglich Variationen, bzw. Misspezifikationen in der Wahrscheinlichkeitsdichte von Parametern und Anfangswerten.

1. Introduction

1.1. Motivation

One of the major goals in the field of systems biology is an understanding of biological processes on a single cell level. To establish this goal, the idea is to model processes in a cell via differential equations. One then studies these models with system theoretical tools to obtain both a quantitative and a qualitative understanding of such processes. For the past decades, a lot of effort has been put into the modeling and analysis of single cells and this approach can be nowadays considered well established [1].

We briefly recall that single cell models describe the dynamics of signaling molecule concentrations or activities within a cell (see e.g. [1], [2]). The dynamics are obtained from first-principles, e.g. via biochemical reactions, and are given by an ordinary differential equation

$$\dot{z}(t) = f(z(t), \theta), \quad z(0) = z_0. \quad (1.1)$$

Here $z(t) \in \mathbb{R}^n$ is a vector of protein concentrations, gene activities, etc. at the time instance t , and $\theta \in \mathbb{R}^q$ is a vector of parameter values that describe reaction kinetics, enzyme concentrations and other cellular properties which remain constant [1].

Having arrived at such understanding in terms of a single cell model, the next level to naturally consider is the population level. The study of cell populations is for example required to understand physiological dynamics in living tissues or metabolic processes in a bioreactor from a cellular perspective [1]. As we are now looking at populations of cells, we naturally have to assume the presence of a certain variability between individuals in the population. For example, even two cells from a population of genetically identical cells still might differ slightly due to phenotypic differences among cells, such as differences in initial protein abundance or in gene expression. In view of our single cell model (1.1) this means that we have to assume that the initial conditions and parameters are distributed within a population.

In some cell populations, this distribution is not essential in the sense that the cells behave very similarly despite the distribution. Those populations can be described by an average or typical single cell model (1.1) without loss of biological information (cf. [3]). For some other cell populations, however, the fact that initial conditions and parameters are distributed, cannot be neglected. They show a heterogeneous behavior and the application of average models would lead to biologically meaningless results.

Gradually, researchers became aware of such *heterogeneous cell populations*. Many cell populations were discovered which showed heterogenous behavior (cf. [3] and references therein) and where an application of average models lead to wrong results. We shall mention an example from the cell death pathway. There, in the process of programmed cell death upon an external stimulus, heterogeneity of a cell population manifests itself in the observation that, only some cells die whereas other cells survive [1].

In this thesis, we study heterogeneous cell population models that consist of structurally identical single cell models, given by (1.1), and where heterogeneity is due to a distribution in initial conditions and parameters (cf. [1], [5]). We believe that the study of such populations is the first step towards a mechanistic understanding of important biological functions driven by heterogeneity.

1.2. Focus of this thesis

While the focus of previous work on heterogeneous cell population models was much on parameter estimation (see e.g. [11], [21], [22], [23]), equally important tasks such as the study of identifiability and sensitivity analysis were left open.

In this thesis, we make the first step towards a study of identifiability, and as a second focus we implement sensitivity analysis for the class of heterogeneous cell populations. In the following section, we briefly summarize key tasks in the modeling and analysis of heterogeneous cell populations, as well as our contributions to these.

1.3. Outline and contributions

1.3.1. Modeling of heterogeneous cell populations

Modeling of heterogeneous cell populations aims at extrapolating the single cell model (1.1) to the cell population level. For this, there exist two frameworks in the literature [1]. In the first modeling framework, we model a cell population literally as a large number of cells, each with individual initial conditions and parameter values. The individual values are realized as samples from a given distribution. Such a model is called an individual-based population model (IBPM) and results in Monte Carlo-based simulation models (see [6]).

In an alternative approach, we describe a heterogeneous cell population by the probability density of its cell states. In contrast to the IBPM approach, information about individual cells is omitted here. We will later reveal that the evolution of this cell population density is governed by a partial differential equation. Models that describe the density of a population via partial differential equations are called population balance models (PBM) in the engineering literature (see [7], [8], [9]).

In **Chapter 2**, we survey the two mentioned frameworks, as well as the solutions to their *direct problems*. The direct problem here is to describe the evolution of the cell population when given a single cell model and an initial probability density. In the IBPM approach, we sample initial conditions from the given probability density and given these, solve the differential equation for the single cell model. For the PBM approach we derive the mentioned partial differential equation for the evolution of densities. We reveal that the PDE is an advection equation, for which the well-known method of characteristics can be applied as a solution technique.

Since the PBM approach has a tremendous advantage over the IBPM approach in terms of a theoretical framework, we will further pursue this approach. In **Chapter 3**, we extend this approach by exploiting the probabilistic nature of the problem and thereby develop a formalism based on operators. This operator-based formalism will be the base for all subsequent problems described in the following.

1.3.2. Parameter estimation and identifiability

More sophisticated than direct problems are *inverse problems*. For the inverse problem for heterogeneous cell population models, we start out with a given single cell model and want to estimate the initial probability density based on measurements of a heterogeneous cell population.

In **Chapter 4**, we introduce the type of measurements in heterogeneous cell populations and also briefly survey the state of the art of parameter estimation methods. Our focus, however, is more on the conceptually important question of identifiability of heterogeneous cell population models, which is necessary for parameter estimation.

In **Chapter 5**, our contribution is the introduction the novel concept of structural identifiability for heterogeneous cell population models. A heterogeneous cell population is said to be structurally identifiable, if it is theoretically possible to reconstruct the initial probability density from the knowledge of the probability density of an output at all times. This definition is thus somewhat analogous to the concept of observability for linear finite-dimensional systems. We are in fact able to illuminate the connection between the structural identifiability of a heterogeneous cell population model and the observability of the underlying single cell model in the linear case.

1.3.3. Sensitivity analysis

In the analysis of single cell models, the study of sensitivity with respect to perturbations in the initial condition or perturbations of parameters is a very important topic. Therefore we try to establish a similar sensitivity analysis for the class of heterogeneous cell population models. For heterogeneous cell population models we study sensitivity with respect to perturbations in the probability density of initial conditions and parameters.

In **Chapter 6**, we introduce local sensitivity analysis in a general framework using the notion of Frechet derivatives. This notion is the natural generalization of the classical derivative to functions that are defined between two normed, possibly infinite-dimensional vector spaces. The introduction of this general framework is necessary because perturbations in the probability density of initial conditions and parameters are elements of the space of integrable functions, which is an infinite-dimensional vector space.

Based on the machinery of multivariate calculus, and our operator formalism developed in Chapter 3, we compute the sensitivity operator and show that it is well-behaved and does neither depend on the single cell dynamics, nor on the particular perturbation. Lastly, we introduce sensitivity analysis also for heterogeneous cell population models with output measurements. Our result shows that in this case the sensitivity operator is still well-behaved, but now does depend on the structural identifiability of the heterogeneous cell population.

Part I.

Preliminaries

2. Models for heterogeneous cell populations and the direct problems

While single cell models are of undeniable importance, single cell models alone in general fail as models for cell populations. This is due to the existence of heterogeneity in a lot of cell populations. In this thesis, we consider cell populations with structurally identical single cells and where heterogeneity is due to differences in initial conditions and parameters in (1.1). Thus, the implicit assumptions are the absence of interactions among cells and secondly, that the dynamics of the single-cells is deterministic.

Since cell populations consist of millions of individual cells, we can treat these differences in a probabilistic framework, i.e. we assume that initial conditions and parameters are described by a probability distribution. In this chapter we present two existing modeling approaches for heterogeneous cell populations, namely individual-based population models (IBPM) and population balance models (PBM).

Before we proceed, let us note that our single cell model in parametric form (1.1) can be put into the more convenient form

$$\dot{x}(t) = F(x(t)), \quad x(0) = x_0. \quad (2.1)$$

We achieve this by introducing the extended state

$$x := (z_1, \dots, z_n, \theta_1, \dots, \theta_q)$$

and the extended vector field

$$F := (f_1, \dots, f_n, 0, \dots, 0).$$

The dimension of the extended state shall be denoted $d := n + q$. From now on we will exclusively refer to system (2.1) for the sake of a simpler notation.

2.1. Monte Carlo-based modeling

In IBPM, or Monte Carlo-based modeling (see [6]), the quite natural idea is to view a heterogeneous cell population as an ensemble of individual cells. We implement this by simulating the single cell model a large number of times with different initial conditions which are drawn as samples from a given probability distribution. To put it more mathematically, we fix a sufficiently large number of cells N and consider an ensemble of differential equations

$$\begin{aligned}\dot{x}^{(1)} &= F(x^{(1)}), \\ &\vdots \\ \dot{x}^{(N)} &= F(x^{(N)}).\end{aligned}$$

The associated initial conditions

$$x_0^{(1)}, \dots, x_0^{(N)}$$

shall be realizations of i.i.d. random variables $X_0^{(1)}, \dots, X_0^{(N)}$ with a given probability distribution. We can easily formulate a heuristic solution to the direct problem in this framework:

- (i) Take N samples from the given distribution,
- (ii) for each sample we solve (2.1) and stop at some specific time T .

For this specific time $T > 0$ the state of the population under scrutiny is given by all the single cell states $x_1(T), \dots, x_N(T)$. However, to get a better understanding of the population as a whole, it is customary to look at the histogram of these single cell states.

Example 2.1 (A simulation example, [10]). We present an implementation of the Monte Carlo-based simulation. For this, consider the model given by the ordinary differential equation

$$\dot{x} = F(x) = \frac{V_{\max} \cdot x^\beta}{K^\beta + x^\beta} - k_d \cdot x. \quad (2.2)$$

This model describes the concentration x of a protein X . As we can guess from the first term, X activates its own expression. The second term models dilution of X due to cell-growth with a rate constant $k_d > 0$.

In this example we assume that x_0 is a random variable with initial density p_0 , which may represent differences in the *abundance of protein X* in individuals of the cell population. For parameters we choose (cf. [10])

$$V_{\max} = 1, \quad K = 2, \quad \beta = 4, \quad k_d = 0.01.$$

The dynamics of system (2.2) with our specific choice of parameters can be deduced from Figure 2.1. There we plotted sections of graph(F) = $\{(x, F(x))\}$. From the graph we see that trajectories initialized left of the steady state $x \approx 0.55$ converge to the origin, while trajectories initialized right of $x \approx 0.55$ converge to a steady state far right.

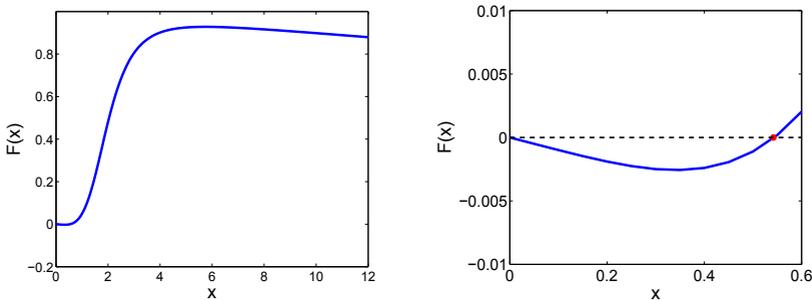


Figure 2.1.: On the left we depicted the graph of the vector field F within the interval $[0, 12]$. On the right we depicted a zoom that focuses on the values of F within the interval $[0, 0.6]$. The red dot indicates a steady state.

Let us now simulate the effect of the considered heterogeneity. The following code takes samples from a normal distribution with mean $\mu = 1$ and variance $\sigma^2 = 0.07$ and uses these as initial conditions for the differential equation (2.2).

Listing 2.1: Monte Carlo simulation

```
sample_size = 5000;
n = normrnd(1, sqrt(0.07), [1 sample_size]);
hist(n, sample_size/100);

for j = 1:sample_size
    [t, x] = ode45('vector_field', [0 50], n(j));
    p(j) = x(end);
end
hist(p, sample_size/100);
```

Figure 2.2 shows on the left the histogram of the samples from the initial distribution and on the right the histogram of the solutions to the ordinary differential equation (2.2) with the samples as initial conditions at time $T = 50$. We notice that the histogram on the right is concentrated in the interval $[35, 40]$ and that there is also a peak (red color) for very small values of $x(T)$. In other words, the initial probability distribution evolved into a bimodal distribution. What this means concretely for the cell population is that for one bulk of cells, the concentration of protein X increases as time increases, while for the other bulk of cells, the concentration tends to zero (red color).

The mathematical reason for this can be seen by inspecting the graph of F in Figure 2.1. As we already pointed out, trajectories that are initialized left of the steady state $x \approx 0.55$ converge to the origin, while trajectories initialized right of $x \approx 0.55$ converge to a steady state far right. In Figure 2.2 we further see on the left, that while a large portion of cells does have an initial protein abundance $x_0 > 0.55$, for the other majority we have $x_0 < 0.55$ (red color). We shall note that in practice one starts with a single cell model and the observation of such heterogeneous behavior (and not with the initial distribution) and then tries to explain the observed phenomena, which is of course much harder than the presented direct problem.

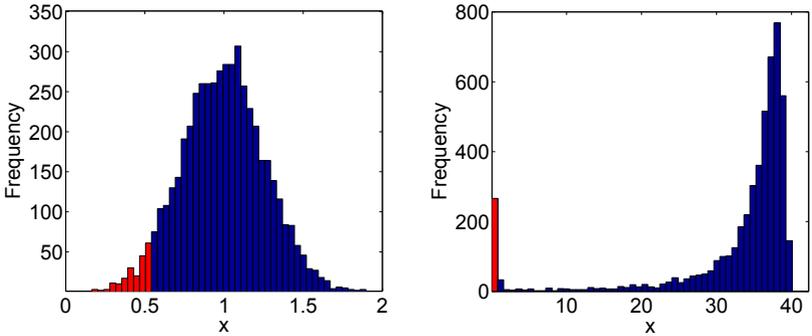


Figure 2.2.: On the left we depict the histogram of samples drawn from the initial distribution. The red colored bars hint the existence of cells with an initial protein abundance $x_0 < 0.55$. On the right we depict the histogram of $x_1(T), \dots, x_N(T)$, which is seen to be a bimodal histogram.

To conclude, the Monte Carlo method offers an easy heuristics to solve the direct problem in IBPM. However, there are two main disadvantages in IBPM. First, Monte Carlo methods in general require a large sample size, which leads to a high computational effort. Secondly, in this approach we lack a theoretical framework in which inverse problems such as the estimation of the initial density from an output density, or problems like sensitivity analysis with respect to perturbations in the initial condition can be treated. Thus, for the remainder of this thesis we will put our effort into developing such framework and tackling the mentioned problems therein.

2.2. Density-based modeling

In PBM we think of a population truly as a whole instead of a set of individuals as in the previous approach. From a mathematical point of view, this simply means that we omit the detour using realizations of random variables and directly model heterogeneous cell populations through the (probability) density of the cell states $x(t)$. It might be familiar for the reader from other areas such as physics, that such density-based approaches typically yield partial differential equations (PDE) for the evolution of densities. In fact in our specific framework our partial differential equation arises out of a random initial value problem. A random initial value problem is given by an ordinary differential equation

$$\dot{x}(t) = F(x(t)),$$

and a random variable X_0 with a given probability density for the initial conditions. Because the initial condition is a random variable, so are the states $x(t) = X_t$. The solution to a random initial value problems should thus be given by the probability density of the states.

We begin by first introducing our probabilistic framework. Our probability space shall consist of the sample space \mathbb{R}^d , the corresponding Borel algebra \mathcal{B} on \mathbb{R}^d and a given probability measure \mathbb{P}_0 . Additionally we assume that the probability measure \mathbb{P}_0 has a probability density p_0 , i.e.

$$\mathbb{P}_0(B) = \int_B p_0 \, d\mu \quad \text{for all } B \in \mathcal{B}.$$

Here we mean Lebesgue integration with respect to the Lebesgue measure μ .

Since heterogeneity of a cell population is in this framework encoded in the probability density of the cell states $x(t)$, the direct problem for PBM can be formulated as:

Given a randomness in the initial condition x_0 in terms of a probability density, what is the resulting probability density of $x(t)$?

Remark 2.2. This problem was already addressed in 1981 (see [13]) in the sensitivity analysis of chemical reactants. There it has been termed “stochastic sensitivity analysis” and is used as a technique for global sensitivity analysis. The interpretation and application of this question is treated very differently there, but from a mathematical point of view it is the same problem as ours.

2.2.1. The advection equation

Let us denote $p(t, \cdot)$ the probability density of the state $x(t)$ for some fixed time $t \geq 0$. Then, to say in advance, we will show that the function

$$\begin{aligned} p : \mathbb{R}_+ \times \mathbb{R}^d &\rightarrow \mathbb{R}, \\ (t, x) &\mapsto p(t, x) \end{aligned}$$

is governed by a partial differential equation of the form

$$\begin{aligned} \frac{\partial}{\partial t} p(t, x) + \operatorname{div}(p(t, x)F(x)) &= 0, \\ p(0, x) &= p_0(x), \end{aligned} \tag{2.3}$$

for all $t > 0$ and all $x \in \mathbb{R}^d$.

In physics, this equation is a so-called *advection equation*, which itself is a special case of a *continuity equation*. In a fluid flow, “advection” refers to the process of something being carried along passively (autonomously) by a fluid, such as a dye [12]. To be more precise, we can think of a substance with some initial density p_0 being placed in a vector field $F : x \mapsto F(x)$. Again the substance is carried along passively by the vector field. Then, if $(t, x) \mapsto p(t, x)$ is a solution to the advection equation (2.3), then the function

$$p(T, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}$$

precisely describes the density of the substance at some fixed time $T > 0$.

With a similar reasoning to the above we justify that the advection equation applies also to advection of a random point, which was our initial motivation.

We now present a simple example that hopefully provides the reader already with an intuitive understanding of the density-based approach.

Example 2.3 (Constant vector field). Let us consider the case where the vector field of the differential equation is *constant*, i.e. is given as

$$F : x \mapsto c = (c_1, \dots, c_d),$$

and the initial conditions are randomly distributed according to a probability density p_0 . Very loosely speaking, one would expect the shape of the distribution to stay the same, but moving along the constant vector field. One quickly comes up with the solution

$$p(t, x) = p_0(x - ct).$$

It is not hard to verify that this function solves the advection equation. Figure 2.3 illustrates this for $d = 2$ and a Gaussian-like initial density that is transported along a constant vector field.

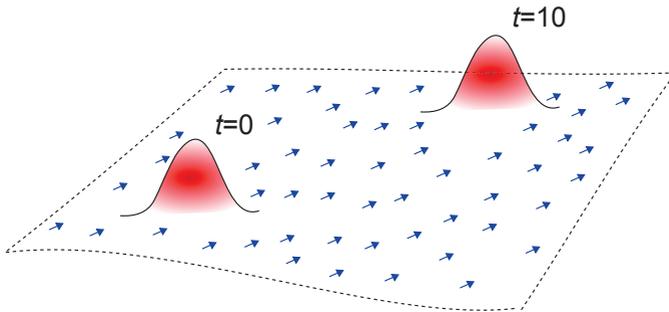


Figure 2.3.: A density gets transported along a constant vector field.

A standard derivation of the advection equation can be found in [12]. For the sake of completeness this derivation is presented here.

Suppose that a vector field F is advecting a quantity and let $p(t, x)$ be the density of that quantity at time t and at location x . We assume that the quantity being advected is neither created nor destroyed, and that

$$H_V(t) = \int_V p(t, x) \, d\mu(x)$$

is the total amount inside the control volume V at time t . As is customary, our ansatz is to look at

$$\frac{dH_V}{dt} = \int_V \partial_t p(t, x) \, d\mu(x) = - \int_{\partial V} j(t, x) \cdot n(x) \, dA(x).$$

From here on it is left to determine the flux $j(t, x)$ and to apply the divergence theorem. The rate at which material is transported across a small piece of surface is proportional to the velocity component normal to the surface, the area of the small piece of surface and the density of the material being advected, thus

$$j(t, x) \cdot n(x) \, dA(x) = p(t, x)(F(x) \cdot n(x)) \, dA(x).$$

Hence the advective flux is $j(t, x) = p(t, x)F(x)$ and applying the divergence theorem

$$0 = \int_V \partial_t p(t, x) \, d\mu(x) + \int_V \operatorname{div}(p(t, x)F(x)) \, d\mu(x).$$

This yields the claimed partial differential equation (2.3) since the above equality holds for any control volume V .

2.2.2. Method of characteristics

In this section, we present a well-known technique for solving the advection equation (2.3), called the method of characteristics. For this we first consider $x : t \mapsto x(t)$, the solution to the initial value problem (2.1). In the context of partial differential equations such as the advection equation, such solution curve is called a *characteristic*. We will show that we can determine the value of the PDE solution along these characteristics by solving ordinary differential equations. This, in turn, allows us to solve the PDE, as we will show afterwards.

Let us denote

$$p(t, x(t)) =: M(t)$$

the PDE solution along a characteristic. Then we have on the one hand

$$M(0) = p(0, x(0)) = p_0(x_0),$$

and on the other hand, differentiating yields

$$\frac{d}{dt}M(t) = \frac{d}{dt}p(t, x(t)) = \frac{\partial}{\partial t}p + \nabla p \cdot \dot{x} = -(\operatorname{div} F)(x(t))p(t, x(t)).$$

The last equality follows from the chain rule

$$\operatorname{div}(Fp) = \operatorname{div}(F)p + \nabla p \cdot F = \operatorname{div}(F)p + \nabla p \cdot \dot{x},$$

together with the advection equation

$$\frac{\partial}{\partial t}p + \operatorname{div}(Fp) = 0.$$

The solution of the advection equation along the characteristic can thus be obtained by solving the initial value problem

$$\begin{aligned} \frac{d}{dt}M(t) &= -(\operatorname{div} F)(x(t))M(t), \\ M(0) &= p_0(x_0). \end{aligned} \tag{2.4}$$

Let us now show how we can determine the value $p(T, x_T)$ for arbitrary time T and arbitrary point x_T in state space (and thus solve the PDE) using the previous result. First, we need to find x_0 such that

$$\Phi_T x_0 = x_T,$$

where $\Phi : (t, x) \mapsto \Phi_t x$ denotes the flow generated by the vector field F . We can obtain this x_0 by solving the differential equation $\dot{x} = F(x)$ backwards in time. Taking x_0 and the associated characteristic $x : t \mapsto x(t) = \Phi_t x_0$, we then solve the system in (2.4) to obtain M . Now observe that

$$M(T) = p(T, x(T)) = p(T, x_T),$$

by construction.

Example 2.4 (Simulation example). In this example we apply the method of characteristics on the heterogeneous cell population model from Example 2.1. For this we choose the same initial probability density as before, i.e. the probability density of

$$N(\mu = 1, \sigma^2 = 0.07^2).$$

This time however, we shall define the end time $T = 15$ as the figures will turn out more clearly. We denote the probability density of the states $x(T)$ as “the end density” in the following.

We begin by gridding the space on which the initial probability density is defined on. We choose on the interval $[0, 5]$ a uniform grid with grid size 0.1 and obtain the grid points

$$x_0^1, \dots, x_0^{50}.$$

These grid points we raise as the initial conditions for $\dot{x} = F(x)$. Taking x_0^i and $x^i : t \mapsto x^i(t)$, the solution obtained by choosing the initial condition x_0^i , we solve (2.4) to obtain the values of the PDE solution along the characteristics.

In Figure 2.4 we plot the curves

$$t \mapsto (t, x^i(t), p(t, x^i(t)))$$

to illustrate the result. We also separately plotted the characteristics and the end density in Figure 2.5. We see that choosing a uniform grid for the initial density yields a not so satisfying approximation for the end density.

Nevertheless we now know the support of the end density so that we can determine the value $p(T, x)$ for every $x \in \text{supp } p(T, \cdot)$ using the previously discussed method. The previously discussed method is implemented as follows:

- (i) Choose a desired grid of the state space for the probability density at some fixed end time $T > 0$. Then, for any point x_T in that grid determine $\Phi_{-T}x_T =: x_0$.
- (ii) Given x_0 and the characteristic, we compute, by solving (2.4), the value of the probability density along the characteristic and thus $p(T, x_T)$ for all grid points.

Applying this two-step procedure (see also [10]) we obtain the characteristics shown in the left plot in Figure 2.6. The right-hand side shows the plot of the end density that we obtain by using the two-step procedure. The result is a very good approximation of the actual density. The code for the two-step procedure is provided in the appendix.

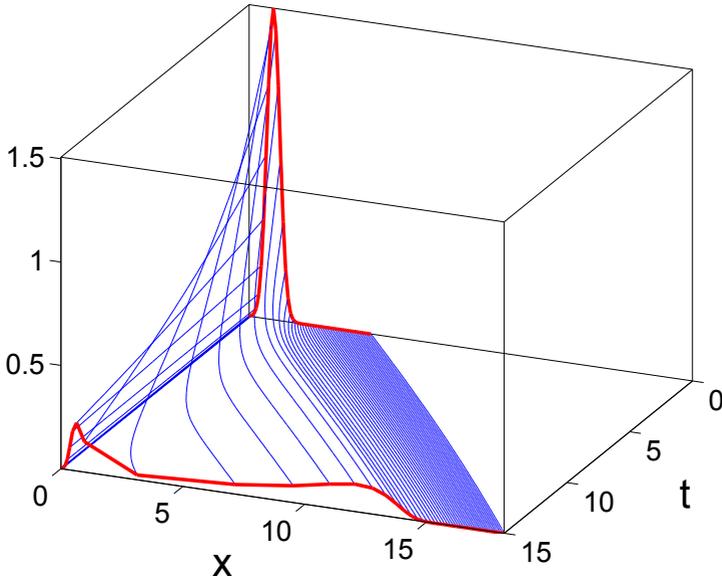


Figure 2.4.: The thick red lines show the initial and the end density. The initial density is the density of the normal distribution $N(\mu = 1, \sigma^2 = 0.07^2)$. The thin blue lines depict the traces $\{(t, x^i(t), p(t, x^i(t))) : 0 \leq t \leq 15\}$ that we obtain through the method of characteristics.

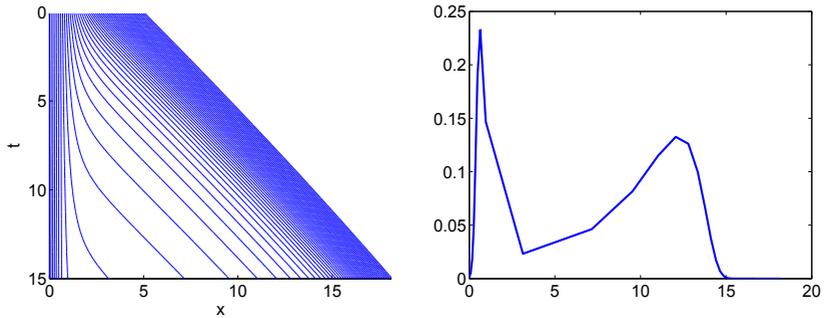


Figure 2.5.: Left: Characteristics obtained by gridding the initial density. Right: A rough approximation of the end density as in Figure 2.4.

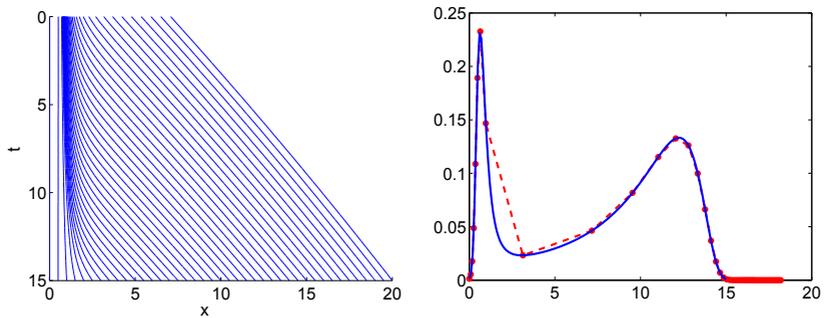


Figure 2.6.: Left: Characteristics obtained via the two-step procedure with a uniform grid with grid size 0.5 on the end density. Right: An approximation of the end density obtained via the two-step procedure with a uniform grid with grid size 0.1 for the end density (blue). For comparison we have also plotted the previous approximation from Figure 2.5 (dashed and red).

2.3. Discussion

We have presented two modeling frameworks for heterogeneous cell populations. In the first framework (IBPM) we model a heterogeneous cell population as an ensemble of structurally identical single cell models with individual initial conditions. In the PBM framework we model a heterogeneous cell population via the density of cell states. For these two frameworks we presented techniques to solve the direct problems, namely a Monte Carlo method and the method of characteristics. To conclude this chapter, we briefly discuss the advantages and disadvantages of the modeling frameworks, as well as the applicability of the respective methods (see also [10]).

Let us start with the applicability as simulation methods. The method of characteristics allows us to evaluate at any time $t > 0$ the value of the evolved probability density at an arbitrary point by using the two-step method. In contrast to the Monte Carlo method, in which we only obtain an approximation in terms of histograms or kernel density estimates, this is a big advantage. We can use this two-step procedure to compute the probability of arbitrary regions in the state space, by gridding only the region and then applying the two-step method. Using this method, we also have no problems to determine the probability of regions with low probability, which is a major problem in the Monte Carlo method.

However, the accuracy of the computed end density depends strongly on the grid on the end density. We typically would choose a uniform grid on the end density. For higher-dimensional problems we are clearly confronted with the “curse of dimensionality”. The computational effort becomes even worse if we consider output mappings. If we want to determine the density of an output $y(t) = Hx(t)$, then using the method of characteristics, we would have to first compute the density of the states and then compute the density of the output (denoted p_H) through marginalization

$$p_H(t, y) = \int_{H^{-1}(\{y\})} p(t, x) dS.$$

Here we denote by dS integration of a surface and we note that we exclude the case that the matrix H is nonsingular. This is in slight anticipation to Section 4.1, where output mappings are properly introduced.

As for the Monte Carlo method, the problem of choosing a grid does not exist. The Monte Carlo method does not depend on the dimension or the end time.

Furthermore, outputs can easily be considered without any additional effort. However, the fact that Monte Carlo methods typically require large sample sizes remains a key disadvantage.

From a theoretical point of view, the density-based approach is clearly much more appealing than the individual-based approach. For the remainder of thesis we will therefore focus on the density-based approach. In the next chapter we extend the PDE approach by introducing Frobenius-Perron operators, that will yield us a very useful tool kit for all subsequent analysis.

3. Frobenius-Perron operators and the advection equation

In the previous chapter we have shown that in the density-based approach a heterogeneous cell population is modeled by the advection equation. This approach has a tremendous advantage over the IBPM approach, as it offers a framework in which direct and inverse problems can be properly studied.

This chapter is devoted to extend this density-based approach. We introduce a measure theoretical approach, which is novel in the context of heterogeneous cell population models. The approach is based on the notion of pushforward measures and results in the theory of Frobenius-Perron operators. Following the introduction, we illuminate the connection between Frobenius-Perron operators and the advection equation. Lastly we prove important qualitative properties of the advection equation using the novel framework.

3.1. An overview of properties of the advection equation

To start with, we shortly state the mentioned important qualitative properties of the advection equation. Since we have already discussed that there is a very picturesque interpretation of the advection equation, it is clear that we would expect several properties from the solution of the equation. The first property that we expect is the conservation of total mass, i.e.

$$\int_{\mathbb{R}^d} p(t, x) d\mu \equiv 1.$$

Another property is that for $p_0 \geq 0$ we also have $p(t, x) \geq 0$ for all $t > 0$ and $x \in \mathbb{R}^d$. We shall restate these two properties as:

- (i) $p(t, \cdot) \geq 0$, for $p_0 \geq 0$,
- (ii) $\|p(t, \cdot)\| = \|p_0\|$, for $p_0 \geq 0$.

3. Frobenius-Perron operators and the advection equation

Let us also talk about the equation itself rather than its solutions. We recall that the initial value problem was to find a function $p = p(t, x)$ that satisfies

$$\frac{\partial}{\partial t} p(t, x) + \operatorname{div}(p(t, x)F(x)) = 0,$$

with the initial condition $p(0, x) = p_0$. This initial value problem can be shown to be well-posed in the sense of Hadamard, i.e. for continuously differentiable p_0 ,

- (i) a solution $p(t, \cdot)$ exists for all $t \geq 0$,
- (ii) this solution is unique,
- (iii) this solution depends continuously on p_0 .

Furthermore the equation is reversible in time, which is for example not the case for the heat equation.

3.2. Frobenius-Perron operators

This section is devoted to giving a short review of the more general theory of Frobenius-Perron operators (sometimes also called transfer operators). Those play an important role in the study of dynamical systems, as they concern the behavior of densities evolving under the influence of deterministic systems. Our presentation is following the textbook “Chaos, Fractals, and Noise” by Lasota and Mackey [14]. For the mathematical background see Appendix A.1 and A.2.

We start with introducing the evolution of probability densities in a measure theoretical framework. The idea is conceivably natural and simple and depicted in Figure 3.1. Suppose we have a flow $\Phi : (t, x) \mapsto \Phi_t x$ corresponding to the differential equation $\dot{x}(t) = F(x(t))$. Fix some time $t > 0$ and consider the mapping $\Phi_t : \mathbb{R}^d \rightarrow \mathbb{R}^d$ from the left probability space (which models the randomness of the initial conditions) to the right space. On the right space we define the canonical probability measure there, namely the pushforward measure with respect to Φ_t , i.e. for all $B \in \mathcal{B}$ we have

$$\mathbb{P}(t)(B) := \mathbb{P}_0(\Phi_t^{-1}(B)). \tag{3.1}$$

This pushforward measure is the probability distribution of the states $x(t)$.

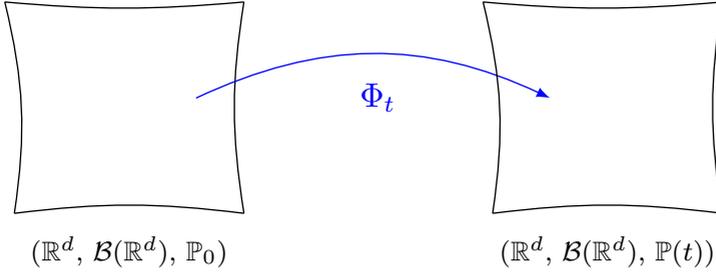


Figure 3.1.: On the left-hand side is the probability space that governs the randomness in the initial condition and on the right-hand side is the probability space that governs the randomness in the states at some fixed time $t > 0$. We can interpret this with the mapping Φ_t between both spaces.

From (3.1) we immediately see that an operator

$$P_t : p_0 \mapsto P_t p_0,$$

mapping a probability density p_0 to the probability density $P_t p_0$ of the push-forward measure under Φ_t , would have to satisfy

$$\int_{\Phi_t^{-1}(B)} p_0 \, d\mu = \int_B P_t p_0 \, d\mu \quad \text{for } B \in \mathcal{B}. \quad (3.2)$$

This is because

$$\mathbb{P}_0(\Phi_t^{-1}(B)) = \int_{\Phi_t^{-1}(B)} p_0 \, d\mu \quad \text{and} \quad \mathbb{P}(t)(B) = \int_B P_t p_0 \, d\mu.$$

Under the assumption that all transformations Φ_t are nonsingular with respect to the Lebesgue measure μ , i.e.

$$\mu(\Phi_t^{-1}(B)) = 0 \quad \text{for each } B \in \mathcal{B} \text{ such that } \mu(B) = 0,$$

we can follow that equation (3.2) for each fixed $t \geq 0$ uniquely defines an operator

$$p_0 \mapsto P_t p_0.$$

This follows from the following version of the Radon-Nikodym theorem.

3. Frobenius-Perron operators and the advection equation

Theorem 3.1 (Radon-Nikodym, Corollary 2.2.1 in [14]). *Let $(\Omega, \mathcal{B}, \mu)$ be a σ -finite measure space and let ν be another σ -finite measure on (X, \mathcal{B}) such that $\nu(B) = 0$ whenever $\mu(B) = 0$. Then there exists a unique $f \in L^1$ such that*

$$\nu(B) = \int_B f \, d\mu$$

for any measurable set B .

Here our (probability) measure is given by $\nu(B) = \int_{\Phi_t^{-1}(B)} p_0 \, d\mu$, which satisfies the assumptions of the Radon-Nikodym theorem.

Recall that our initial motivation was to study the mapping, that applied to the initial density function gives the density of the pushforward measure. Let us now generalize this concept and consider for arbitrary $h \in L^1$ the equation

$$\int_{\Phi_t^{-1}(B)} h \, d\mu = \int_B P_t h \, d\mu \quad \text{for } B \in \mathcal{B}. \quad (3.3)$$

In [14], Section 3.2, the following theorem is proven based on the previous result for probability densities.

Theorem 3.2. *Let $(\Omega, \mathcal{B}, \mu)$ be a measure space and let $\Phi_t : \Omega \rightarrow \Omega$ be a non-singular transformation with respect to μ . Then, for each fixed $t \geq 0$, equation (3.3) defines a unique operator $P_t : L^1(\Omega) \rightarrow L^1(\Omega)$.*

We denote $P_t : L^1 \rightarrow L^1$ the Frobenius-Perron operator corresponding to Φ_t . Exploiting equation (3.2) it is an easy exercise to verify the following properties:

- (i) P_t is linear,
- (ii) $P_t h \geq 0$, if $h \geq 0$,
- (iii) $\int_{\Omega} P_t h \, d\mu = \int_{\Omega} h \, d\mu$.

Although the definition of the Frobenius-Perron operator is given by a quite abstract mathematical theorem of Radon-Nikodym, it should be realized that it precisely describes the evolution of a density p_0 by a transformation Φ_t . Furthermore, properties (i) to (iii) of the transformed density $P_t p_0$ are *exactly* what one would expect on intuitive grounds [14].

Along side the properties (i)-(iii), Frobenius-Perron operators have another important property. They are contractions, as shown in the next theorem.

Theorem 3.3 (Contraction property). *Let P_t be a Frobenius-Perron operator and let $h \in L^1$. Then $\|P_t h\| \leq \|h\|$, i.e. the Frobenius-Perron operator is a contraction.*

Proof. A simple calculation yields the claim

$$\begin{aligned} \|P_t h\| &= \|P_t(h^+ - h^-)\| = \|P_t(h^+) - P_t(h^-)\| \\ &\leq \|P_t(h^+)\| + \|P_t(h^-)\| = \|h^+\| + \|h^-\| = \|h\|. \end{aligned}$$

Here we have used the decomposition $h = h^+ - h^-$, linearity of P_t , triangle inequality of the L^1 -norm and the fact that $\|P_t h\| = \|h\|$ for $h \geq 0$. \square

One might wonder, when equality holds. The following theorem gives us an answer in terms of a sufficient condition.

Theorem 3.4. *Let P_t be a Frobenius-Perron operator and let $h \in L^1$. Then equality $\|P_t h\| = \|h\|$ holds if*

$$(P_t(h^+))(x) = 0 \quad \text{or} \quad (P_t(h^-))(x) = 0 \tag{3.4}$$

is true almost everywhere.

Proof. From the previous proof we see that for proving $\|P_t h\| = \|h\|$, it suffices to show the equality

$$\|P_t(h^+) - P_t(h^-)\| = \|P_t(h^+)\| + \|P_t(h^-)\|. \tag{3.5}$$

Since (3.4) is true almost everywhere, we can conclude that

$$|(P(h^+))(x) - (P(h^-))(x)| = |(P(h^+))(x)| + |(P(h^-))(x)|$$

is true almost everywhere. Integration over Ω yields the claim $\|P_t h\| = \|h\|$. \square

Loosely speaking, if the functions $P_t(h^+)$ and $P_t(h^-)$ have disjoint supports, then no cancellation is possible yielding the equality (3.5). We will come back to this property in the sensitivity analysis in Chapter 6.

3.3. Frobenius-Perron operators and the advection equation

So far it should have become intuitively clear that Frobenius-Perron operators describe the evolution or flow of a density under a differential equation. At the same time, as we discussed earlier, the evolution of densities under a differential equation is described by the advection equation. In this section we unify both results by illuminating the precise connection between Frobenius-Perron operators and the advection equation.

We first show that the family of Frobenius-Perron operators inherits the properties of the flow $\Phi : (t, x) \mapsto \Phi_t x$ (cf. Section 7.4 in [14] and Appendix A.1). To see that

$$P_0 h = h$$

we recall that $\Phi_0^{-1}(B) = B$ and, consequently,

$$\int_B P_0 h \, d\mu = \int_{\Phi_0^{-1}(B)} h \, d\mu = \int_B h \, d\mu.$$

To see that

$$P_{t+t'} h = P_t(P_{t'} h)$$

we recall that $\Phi_{t+t'}^{-1}(B) = \Phi_{-t'}(\Phi_{-t}(B))$ and, thus,

$$\int_B P_{t+t'} h \, d\mu = \int_{\Phi_{-t'}(\Phi_{-t}(B))} h \, d\mu = \int_{\Phi_t^{-1}(B)} P_{t'} h \, d\mu = \int_B P_t(P_{t'} h) \, d\mu.$$

Thus we have rigorously shown that what we have been studying is in fact the flow of densities in a well-defined sense (cf. Appendix A.1, definition of a flow).

To illuminate the connection between the Frobenius-Perron operator and the advection equation, we need to determine the *infinitesimal generator* for the group $\{P_t\}_{t \geq 0}$. To be more precise, we would have to show that for an arbitrary continuously differentiable function $h : \mathbb{R}^d \rightarrow \mathbb{R}$ we have

$$\lim_{t \rightarrow 0} \frac{P_t h - h}{t} = -\operatorname{div}(h(t, x)F(x)). \quad (3.6)$$

This in turn shows that for an arbitrary continuously differentiable function $h : \mathbb{R}^d \rightarrow \mathbb{R}$, the function defined by

$$p(t, x) = P_t h(x),$$

satisfies the advection equation

$$\begin{aligned} \frac{\partial}{\partial t} p(t, x) + \operatorname{div}(p(t, x)F(x)) &= 0, \\ p(0, x) &= h. \end{aligned}$$

Equation (3.6) is of course true, but we omit presenting the derivation and discussing the details, since it can be found in Section 7.6 of [14] in full detail.

To conclude, the fact that Frobenius-Perron operators describe the flow of densities and the fact that the infinitesimal generator for the group of Frobenius-Perron operators is given by the advection equation fit perfectly.

After having established this connection, we see that the properties

- (i) $p(t, \cdot) \geq 0$, for $p_0 \geq 0$,
- (ii) $\|p(t, \cdot)\| = \|p_0\|$, for $p_0 \geq 0$,

that we mentioned in the introduction, follow directly from the properties of the Frobenius-Perron operator. For the well-definedness in the sense of Hadamar, we first see that existence and uniqueness follow from the well-definedness of Frobenius-Perron operators and that continuity follows from the contraction property formulated in Theorem 3.3.

Furthermore, linearity of the Frobenius-Perron operator translates into the fact that the advection equation satisfies the superposition principle. Lastly, the fact that the advection equation is reversible in time follows from the group properties of the family of Frobenius-Perron operators.

4. Parameter estimation

In the previous chapters we have shown that the evolution of the cell state density is given in terms of a partial differential equation. If the initial density is given, we can simulate numerically the evolution of a heterogeneous cell population by solving the partial differential equation.

In this chapter we focus on inverse problems, which consist in reconstructing the initial density from observations of the cell population. In Section 4.1 we introduce for this purpose heterogeneous cell populations with output. Afterwards we formulate the parameter estimation problem and briefly review the state of the art of parameter estimation.

Although parameter estimation is a very important tasks in the analysis of heterogeneous cell population models, it is not within the main focus of this thesis and shall only serve as a motivation for the study of structural identifiability in the next chapter.

4.1. Cell populations with output

An important inverse problem in the study of heterogeneous cell population models is the reconstruction of the initial probability density. Suppose for the moment that we could measure for a fixed time $t > 0$ the cell states $x(t)$ for a large amount of cells. Then we could obtain a good estimate of the density $p(t, \cdot)$ and since the process of advection is reversible in time, we could therefore obtain a good estimate of the initial density. However, the assumption that we can measure the whole state is not very realistic.

As is well-known from practical control problems, one typically cannot measure the whole state $x(t)$ of a system $\dot{x} = F(x)$, but only a specific output $y(t)$. This is mostly because the application of measurements for all state variables $x_i(t)$ is too costly or that there are simply no methods available to measure a certain state variable.

4. Parameter estimation

For cell populations this means that for a fixed time $t > 0$ we do not know the probability density of $x(t)$ but only the probability density of the outputs $y(t)$. In this thesis, we focus on the case of linear outputs

$$y(t) = Hx(t),$$

with a singular matrix $H \in \mathbb{R}^{m \times n}$.

A treatment of this situation in a probabilistic framework is straightforward. For this consider the illustration of this situation given in Figure 4.1.

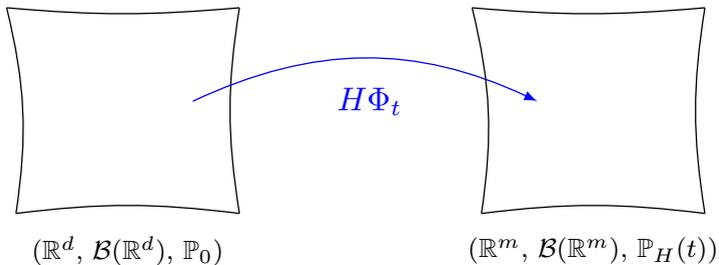


Figure 4.1.: On the left we have the probability space describing the distribution of the initial conditions x_0 . Between the left and the right space is defined the mapping $x_0 \mapsto H\Phi_t x_0 = y(t)$ and thus the right space shall describe the distribution of the outputs, denoted $\mathbb{P}_H(t)$.

In the spirit of Section 3.2, we define the distribution of the outputs $\mathbb{P}_H(t)$ as the pushforward measure of \mathbb{P}_0 under $H\Phi_t$. In other words, for an arbitrary $B_y \in \mathcal{B}(\mathbb{R}^m)$ we set

$$\mathbb{P}_H(t)(B_y) := \mathbb{P}_0((H\Phi_t)^{-1}(B_y)).$$

Given the density of the states $p = p(t, x)$ we can determine the density of the distribution $\mathbb{P}_H(t)$ through marginalization

$$p_H(t, y) = \int_{H^{-1}(\{y\})} p(t, x) \, dS. \quad (4.1)$$

Here we denote by dS integration of a surface and we note that we exclude the case that the matrix H is nonsingular. To see that (4.1) is true, observe that for all $B_y \in \mathcal{B}(\mathbb{R}^m)$,

$$\int_{B_y} p_H(t, y) dy = \int_{B_y} \int_{H^{-1}(\{y\})} p(t, x) dS dy = \int_{H^{-1}(B_y)} p(t, x) d\mu = \mathbb{P}_H(t)(B_y).$$

We are now ready to formulate the parameter estimation problem.

4.2. The parameter estimation problem

A crucial task in the study of heterogeneous cell populations models is the estimation of the unknown initial probability distribution from measurement data. Again, this is because in our framework, heterogeneity of a cell population is encoded in the initial probability distribution (cf. Example 2.1). The available measurement data typically consists of realizations of an ensemble of i.i.d. random variables

$$X \sim \mathbb{P}_H(t_i),$$

where t_i are fixed time instances (cf. [21]). This type of measurement data is available through experiments such as high-throughput fluorescence flow cytometry or microscopy [21].¹ It is to be stressed, that the particular difficulties of this setup arise from the fact that only the distribution of the outputs is known for specific sampling times (or all times, respectively). In particular, output trajectories of individuals, are not given which can be regarded the fundamental problem here.

For parameter estimation based on measurement data, several methods have been proposed by scientists at the Institute for Systems Theory and Automatic Control, which we would like to briefly survey. The studied methods reach from a sampling-based approach (see [5]) and ℓ_2 -norm minimization (see [11]) to maximum-likelihood estimation (see [22]) and a Bayesian approach (see [23]). Although the first method performed well for noise-free data, the assumption of noise-free data was quickly seen to be unpractical. Therefore, all subsequent methods account for measurement noise.

¹As an alternative problem, one could also consider determining the unknown initial probability distribution from the function $p_H : (t, y) \mapsto p_H(t, y)$. This is discussed in the outlook in Section 7.2.

The methods have been shown to perform well for artificial examples, but an open problem is to rigorously justify the consistency of the estimators. To prove consistency for concrete applications, it is clear that we need to study identifiability properties of the systems under scrutiny. However, identifiability properties of heterogeneous cell population models with output measurements have not been studied so far. In the next chapter, we introduce for the first time the concept of structural identifiability for heterogeneous cell population models with output measurements.

Part II.

Structural identifiability and sensitivity analysis

5. Structural identifiability

In order to do parameter estimation, we always need some kind of identifiability property of a model or system. In this chapter we define structural identifiability for heterogeneous cell population models. This definition is motivated by the observability property in linear finite-dimensional systems. In Section 5.2 we focus on the case that the single cell models of a heterogeneous cell population model are linear. We establish connections between the structural identifiability of the heterogeneous cell population model and the observability of the single cell model.

5.1. Motivation

Structural identifiability for heterogeneous cell population models is motivated by *observability* for linear finite-dimensional systems. Thus, to start with, let us recall that one of the many equivalent formulations for observability of linear finite-dimensional systems is given by (cf. [15], Definition 4.2 or [16])

$$(\forall t \geq 0 : y(t; x'_0) = y(t; x''_0)) \Rightarrow x'_0 = x''_0,$$

where $x'_0, x''_0 \in \mathbb{R}^n$ are arbitrary initial conditions. In heterogeneous cell population models with output measurements the output distribution

$$\mathbb{P}_H(t)(B_y) := \mathbb{P}_0((H\Phi_t)^{-1}(B_y))$$

takes the role as “the output” and the initial distribution that takes the role as the “initial condition”. Therefore it is somewhat evident to define *structural identifiability of heterogeneous cell population models* as the property

$$(\forall t \geq 0 : \mathbb{P}_H(t; \mathbb{P}'_0) = \mathbb{P}_H(t; \mathbb{P}''_0)) \Rightarrow \mathbb{P}'_0 = \mathbb{P}''_0,$$

where \mathbb{P}'_0 and \mathbb{P}''_0 are arbitrary probability distributions. Due to the definition of the output distribution, this is equivalent to

$$(\forall t \geq 0 \ \forall B_y \in \mathcal{B}(\mathbb{R}^m) : \mathbb{P}'_0((H\Phi_t)^{-1}(B_y)) = \mathbb{P}''_0((H\Phi_t)^{-1}(B_y))) \Rightarrow \mathbb{P}'_0 = \mathbb{P}''_0.$$

5. Structural identifiability

And lastly, since we assume that the probability measures \mathbb{P}'_0 and \mathbb{P}''_0 have probability densities, we can rewrite structural identifiability as the property that the following statement

$$\forall t \geq 0 \quad \forall B_y \in \mathcal{B}(\mathbb{R}^m) : \int_{(H\Phi_t)^{-1}(B_y)} p'_0 \, d\mu = \int_{(H\Phi_t)^{-1}(B_y)} p''_0 \, d\mu$$

implies $p'_0 = p''_0$ (almost everywhere). In the following, we study criteria for this implication to hold or to not hold.

5.2. Criteria for structural identifiability

In this section we try to deduce criteria for structural identifiability. We focus on the case that the vector field F of the single cell model is linear, i.e. $\Phi_t x = e^{At} x$. Before we proceed with formulating and proving theorems, let us first consider a simple example. This example shall give us an intuitive understanding of the connection between observability of (A, H) and the structural identifiability of the corresponding cell population.

5.2.1. An illustrative example

Let us consider the single cell model $\dot{z} = \theta - z$. We can think of z as the concentration of a protein in a gene regulatory network and θ as a constant expression rate. The term $-z$ is due to degradation of the protein. This model can be rewritten into the form $\dot{x} = Ax$, by choosing

$$x = \begin{pmatrix} z \\ \theta \end{pmatrix}, \quad A = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}.$$

We sketched a phase portrait of $\dot{x} = Ax$ in Figure 5.1. Consider additionally the two different output matrices

$$H' = \begin{pmatrix} 0 & 1 \end{pmatrix} \quad \text{and} \quad H'' = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

The first output matrix could be interpreted as a measurement of RNA abundance in a cell, while the second output matrix corresponds to measurements of the protein abundance. We quickly see that (A, H') is not observable, while (A, H'') is observable.

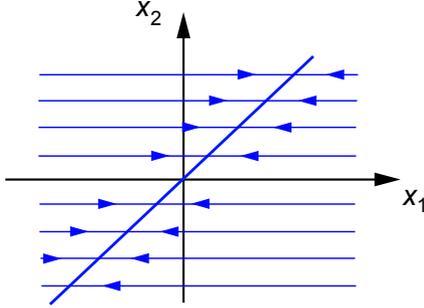


Figure 5.1.: Phase portrait of $\dot{x} = Ax$.

Recall that a cell population with linear single cell dynamics is structurally identifiable by definition if

$$\forall t \geq 0 \quad \forall B_y \in \mathcal{B}(\mathbb{R}^m) : \int_{(He^{At})^{-1}(B_y)} p'_0 \, d\mu = \int_{(He^{At})^{-1}(B_y)} p''_0 \, d\mu \quad (5.1)$$

implies that $p'_0 = p''_0$.

In the following we try to illustrate how this implication is related to the observability of (A, H) . First of all we have in general that

$$\ker He^{At} = e^{-At}(H^{-1}(\{0\})).$$

For the first output matrix we have

$$H'^{-1}(\{0\}) = \mathbb{R} \times \{0\}.$$

This subspace is invariant under the flow of $\dot{x} = Ax$ and therefore it holds that

$$\ker H'e^{At} = \mathbb{R} \times \{0\} \text{ for all } t \geq 0.$$

Thus for arbitrary $t \geq 0$ and $B_y \in \mathcal{B}(\mathbb{R}^m)$, the set $(H_1e^{At})^{-1}(B_y)$ is some combination of strips as depicted on the left of Figure 5.2.

The fact that (5.1) holds, however, does not imply that $p'_0 = p''_0$. Consider for example the case depicted on the right of Figure 5.2. There, integration along those strips yields the same value, while the densities are not identical.

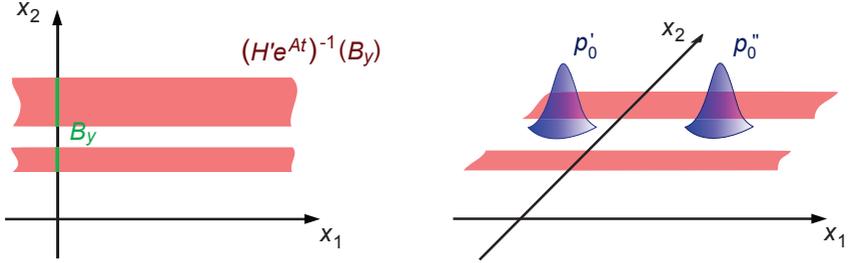


Figure 5.2.: Left: The green set shall depict some $B_y \in \mathcal{B}(\mathbb{R}^m)$. The set $(H'e^{At})^{-1}(B_y)$ is then given by the red strips. Right: A choice of $p'_0 \neq p''_0$ such that integration along all possible “strips” always yield the same value.

Let us see what happens, if we choose H'' which yielded (A, H'') observable. Geometrically viewed, the kernel $H''^{-1}(\{0\})$ gets transported by the flow in the way depicted in Figure 5.3. This is in accordance with the observability of (A, H'') , since we can see that the intersection is trivial, i.e.

$$\bigcap_{t \geq 0} \ker H'' e^{At} = \{0\}.$$

Loosely speaking, observability enforces the subspaces $\ker H'' e^{At}$ to move, as t is changing. For this special case this results that the integral along all the strips depicted on the right of Figure 5.3 must be zero. It is somehow clear that this fact together with the condition

$$\forall t \geq 0 \quad \forall B_y \in \mathcal{B}(\mathbb{R}^m) : \int_{(He^{At})^{-1}(B_y)} p'_0 - p''_0 \, d\mu = 0$$

should imply that $p'_0 - p''_0 = 0$ almost everywhere.

To conclude, this example gives us an intuitive idea about the connection of the observability of (A, H) and the structural identifiability of the corresponding heterogeneous cell population model.

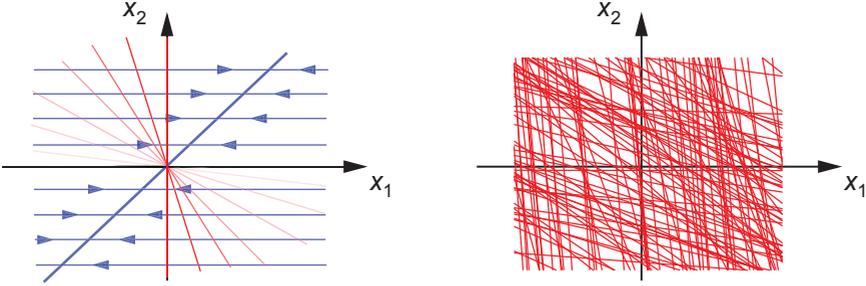


Figure 5.3.: Left: The “transportation” of the kernel $H''^{-1}(\{0\})$ by the flow is hinted. The transparency of $\ker H''e^{At}$ decreases as the time t increases. Right: Since (A, H'') is observable, we can choose B_y and $t \geq 0$ such that $(H''e^{At})^{-1}(B_y)$ is any line depicted. Since the integral along all the strips has to be zero, it seems intuitively clear, that this implies $p'_0 - p''_0 = 0$.

5.2.2. Necessary condition

In the previous section we have seen in a concrete example why a non-observable single cell model (A, H) renders the corresponding heterogeneous cell population not structurally identifiable. In this section we generalize the idea from the example to obtain the following theorem.

Theorem 5.1 (Necessary condition for structural identifiability). *Suppose a given heterogeneous cell population is structurally identifiable, i.e. (5.1) does imply the equality $p'_0 = p''_0$. Then (A, H) has to be observable.*

Proof. Our proof strategy is to show that under the assumption that (A, H) is not observable, there exist probability densities $p'_0 \neq p''_0$ for which equation (5.1) is true. First we fix an arbitrary probability density p'_0 .

It is well-known that (A, H) not being observable is equivalent to the fact that the observability map

$$He^{A(\cdot)} : x_0 \mapsto He^{A(\cdot)}x_0 = y(\cdot),$$

5. Structural identifiability

is not injective, or equivalently, that

$$\ker He^{A(\cdot)} = \bigcap_{t \geq 0} \ker He^{At} \text{ is non-trivial.}$$

Therefore we can pick a non-zero vector $v \in \bigcap_{t \geq 0} \ker He^{At}$, and given that define our second probability density p_0'' by shifting our first probability density p_0' along v ,

$$p_0''(x) := p_0'(x + v).$$

Now obviously we have $p_0' \neq p_0''$, while for all $t \geq 0$ and $B_y \in \mathcal{B}(\mathbb{R}^m)$ we have

$$\begin{aligned} \int_{(He^{At})^{-1}(B_y)} p_0''(x) \, d\mu &= \int_{(He^{At})^{-1}(B_y)} p_0'(x + v) \, d\mu \\ &= \int_{v + (He^{At})^{-1}(B_y)} p_0' \, d\mu. \end{aligned}$$

Lastly, we observe that

$$v + (He^{At})^{-1}(B_y) = (He^{At})^{-1}(B_y),$$

since $v \in \ker He^{At}$ for all $t \geq 0$. Thus we have shown that equation (5.1) is true, while $p_0' \neq p_0''$. This concludes the proof. \square

5.2.3. Discussion on sufficient condition

To prove that (A, H) being observable is a sufficient condition for structural identifiability, we would have to show that, given the observability of (A, H) , the condition

$$\forall t \geq 0 \quad \forall B_y \in \mathcal{B}(\mathbb{R}^m) : \int_{(He^{At})^{-1}(B_y)} p_0' - p_0'' \, d\mu = 0$$

is strong enough to enforce $p_0' - p_0'' = 0$. This however is still an open problem.

In the previous example this fact seemed plausible, but even there we have not found a way to rigorously prove this. Should we be able to prove it for the concrete example, we could try to generalize the idea of the proof as we did for the necessary condition.

One approach could be to try to characterize the space

$$\mathcal{H}_t := \left\{ h \in L^1 \mid \forall B_y \in \mathcal{B}(\mathbb{R}^m) : \int_{(He^{At})^{-1}(B_y)} h \, d\mu = 0 \right\}.$$

This is because it is easily seen that structural identifiability is equivalent to the fact that

$$\bigcap_{t \geq 0} \mathcal{H}_t = \{0\}. \quad (5.2)$$

Thus, once we found a “useful” characterization of the space \mathcal{H}_t , we could study whether observability of (A, H) does yield a trivial intersection (5.2).

However, the only thing that we know so far about \mathcal{H}_t is that the set

$$\left\{ h = p_0 - p_0(\cdot + v) \in L^1 \mid p_0 \text{ a probability density and } v \in \ker He^{At} \right\}$$

is a subset of \mathcal{H}_t . By considering the linear span of this set we obtain a subspace of \mathcal{H}_t . However, it is not immediately clear, if this subspace is the whole space \mathcal{H}_t or not, which is a major issue. We leave the characterization of the kernels as an open problem.

6. Sensitivity analysis

Generally speaking, sensitivity analysis investigates how “sensitive” a system model is to variations in its parameter values. In other words we investigate quantitatively the effects of parameter variations on the behavior of a system. This makes sensitivity analysis an important tool for tasks such as finding particularly influential or identifying insignificant parameters for model reduction [17].

For sensitivity analysis of systems that are described by parametric ordinary differential equations, such as single cell models, there exists already a large amount of literature (see e.g. [18], [19]). The goal of this chapter is to implement sensitivity analysis for the class of heterogeneous cell population models that we introduced.

We start this chapter by giving a short introduction on the general theme of sensitivity analysis. Afterwards we introduce sensitivity analysis for heterogeneous cell populations with respect to perturbations in the initial density. This kind of sensitivity analysis is particularly important to us because due to the estimation of the initial density, we naturally have to assume that there is a certain mismatch. Following the introduction we show that in contrast to single cell models, in heterogeneous cell population models (without outputs) there are no particularly influential or insignificant variations. Loosely speaking, we could say that “all variations are the equal”. Lastly, we introduce sensitivity analysis for heterogeneous cell populations with output.

6.1. Sensitivity analysis and the Frechet derivative

The most prominent sensitivity analysis is mostly local and mostly linear. Although the methods in local and linear sensitivity analysis all share the same mathematical idea, sensitivity in more practical areas however is still often handled rather vividly. Let us therefore review the idea of local and linear sensitivity on a more general basis.

6. Sensitivity analysis

In general, we have two normed vector spaces $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ and a function f that is defined on an open subset of E , taking values in F , i.e.

$$f : U \subset E \rightarrow F.$$

The function f shall describe a given system and the normed vector space E here takes the role of our parameter space. Since f is in general a nonlinear function defined on possibly infinite-dimensional spaces, studying f directly is quite hard. Therefore, we wish to approximate for a given $x_0 \in E$ the behavior of f in a neighborhood of x_0 by a linear and bounded operator. Those are of course much easier to study and to handle. This directly motivates the following definition.

Definition 6.1 (Frechet derivative). *$f : U \subset E \rightarrow F$ is called differentiable in $x_0 \in E$ if there exists a linear and bounded operator $T_{x_0} \in L(E, F)$, such that we have the following decomposition*

$$f(x_0 + h) = f(x_0) + T_{x_0}h + o(\|h\|), \quad h \rightarrow 0.$$

Then we denote $T_{x_0} =: f'(x_0)$ the Frechet derivative of the function f in the point x_0 .

The study of local and linear sensitivity thus accounts to the study of properties of the Frechet derivative. For example, if one studies a differentiable function

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

then it is well-known that the Frechet derivative is given by

$$f'(x_0) : h \mapsto J_f(x_0)h,$$

where $J_f(x_0)$ is the Jacobian of f at point x_0 . By definition we have

$$f(x_0 + h) = f(x_0) + J_f(x_0)h + o(\|h\|), \quad h \rightarrow 0.$$

We can now find those (sufficiently small) variations that lead to the largest changes by studying the properties of the Jacobian.

6.2. Sensitivity analysis for heterogeneous cell populations models

Let us start with the case in which output mappings are neglected. In this case a heterogeneous cell population can be modeled by the advection equation

$$\begin{aligned}\frac{\partial}{\partial t}p(t, x) + \operatorname{div}(p(t, x)F(x)) &= 0, \\ p(0, x) &= p_0.\end{aligned}$$

For the question of sensitivity of heterogeneous cell populations we consider perturbations $h \in L^1$ such that $p_0 + h$ is also a probability density.¹ We then consider the solution to the PDE with perturbed initial density.

$$\begin{aligned}\frac{\partial}{\partial t}p(t, x) + \operatorname{div}(p(t, x)F(x)) &= 0, \\ p(0, x) &= p_0 + h.\end{aligned}$$

By doing so we are considering the results of misspecification of the initial probability density on the state density for some given time $t > 0$.

As we said in the previous section, the study of local and linear sensitivity accounts to the study of properties of the Frechet derivative. In our particular framework the spaces are clearly

$$E = F = L^1(\mathbb{R}^d).$$

For cell populations without output, the Frobenius-Perron operator

$$P_t : L^1 \rightarrow L^1$$

introduced in Chapter 3 takes the role of the function f in Definition 6.1. In view of this definition, we further look at $P_t(p_0 + h)$. Since the Frobenius-Perron operator is linear, we have

$$P_t(p_0 + h) = P_t p_0 + P_t h$$

which satisfies the decomposition in Definition 6.1. Since P_t is furthermore bounded, it actually coincides with its Frechet derivative.

¹A criterium would be that $p_0 + h \geq 0$ and $\int_{\mathbb{R}^d} p_0 + h \, d\mu = 1$, i.e. $p_0 \geq h$ and $\int_{\mathbb{R}^d} h \, d\mu = 0$.

6. Sensitivity analysis

Now we show the identity $\|P_t h\| = \|h\|$ for all $t \geq 0$. For this also recall our measure theoretical discussion of the contraction property of Frobenius-Perron operators in Section 3.2. This time however, we combine the purely measure theoretical approach with the fact that

$$(P_t h)(x) = p(t, x)$$

can be computed via the method of characteristics. First of all it is

$$\|P_t h\| = \int_{\mathbb{R}^d} |P_t h| d\mu = \int_{\mathbb{R}^d} |P_t h^- - P_t h^+| d\mu.$$

Since the positive part h^+ and the negative part h^- are advected separately and because characteristics do not cross, there is no cancellation between $P_t h^+$ and $P_t h^-$. Thus, in view of Theorem 3.3 and 3.4,

$$\begin{aligned} \int_{\mathbb{R}^d} |P_t h^- - P_t h^+| d\mu &= \int_{\mathbb{R}^d} |P_t h^-| d\mu + \int_{\mathbb{R}^d} |P_t h^+| d\mu \\ &= \int_{\mathbb{R}^d} |h^-| d\mu + \int_{\mathbb{R}^d} |h^+| d\mu = \|h\|. \end{aligned}$$

This shows the identity $\|P_t h\| = \|h\|$ for all $t \geq 0$.

Having established this result the question of how misspecifications of the initial density evolve can now be fully answered. Let p_0 be an initial density and consider the misspecified initial density $q_0 = p_0 + h$, where $\|h\| = \delta$. Then for any $t \geq 0$ we have

$$\|P_t q_0 - P_t p_0\| = \|P_t(q_0 - p_0)\| = \|q_0 - p_0\| = \|h\| = \delta.$$

Again, this means that the L^1 -error is preserved for all $t \geq 0$, and this in turn shows that the sensitivity with respect to misspecifications is well-behaved. At this point also recall that in nonparametric statistics the L^1 -norm happens to be the most natural norm to measure errors in estimation between a density and its estimate [20].

Now to somehow illustrate this result, we can draw a diagram as in Figure 6.1. There the result is presented very similarly to the presentations of local sensitivity analysis in the finite-dimensional case.

space of evolved densities at $t > 0$

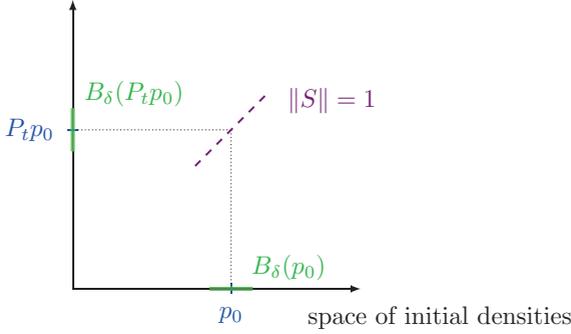


Figure 6.1.: Illustration of the sensitivity result. $B_\delta(p_0)$ and $B_\delta(P_t p_0)$ denote the norm ball with radius δ centered at p_0 and $P_t p_0$, respectively. We have shown that both radii are exactly the same. The dashed violet line shall hint that the sensitivity operator S is linear and has norm one.

Example 6.2. Let us quickly illustrate this sensitivity result with our model $\dot{x} = \frac{Y_{\max} \cdot x^\beta}{K^\beta + x^\beta} - k_d \cdot x$. In the following we fix a “true” probability density, namely that of $N(0.8, 0.01^2)$. Let us consider for example the misspecification $N(0.77, 0.01^2)$. In Figure 6.2 we plotted the misspecified initial density, the true density and the error which is the difference of both. In Figure 6.3 we have plotted the respective propagations.

The L^1 -norm of the difference in the initial densities is computed in Matlab via trapezoidal rule integration (`trapz`) as 0.2385. The L^1 -norm of the difference of the propagations (which is the propagation of the difference by linearity) is computed to be 0.2396. This is clearly almost the same besides a slight difference. The Matlab script can be found in the appendix.

Remark 6.3. The difference stems from the numerical integration, as the probability density at $T = 50$ is given in terms of a grid (Section 2.2.2). To remind the reader of this fact, in Figure 6.4 we depict a zoom of the difference of the two propagations (i.e. the right plot in Figure 6.3).

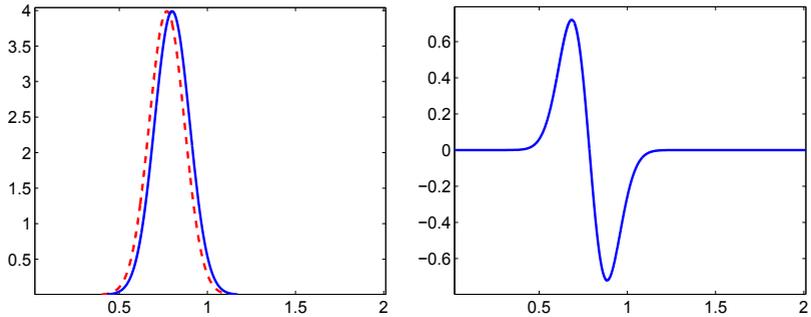


Figure 6.2.: The first figure shows the misspecified initial density (dashed and red), which is the density of $N(0.77, 0.01^2)$ as well as the true initial density (blue) which is the density of $N(0.8, 0.01^2)$. The second figure shows the difference of them.

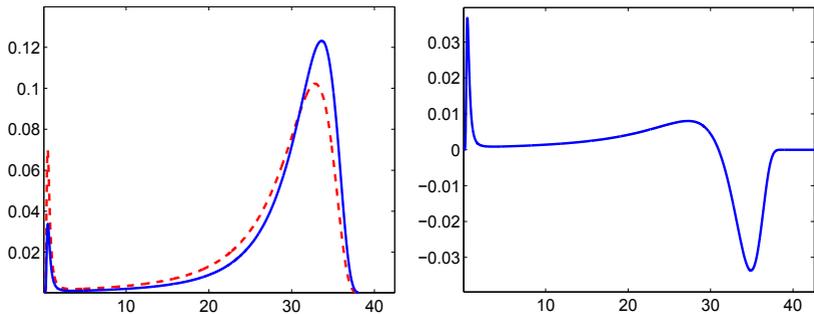


Figure 6.3.: The first figure shows the propagations at $T = 50$ of the misspecified density (dashed and red) and of the true density (blue). The second figure shows the difference of the two propagations.

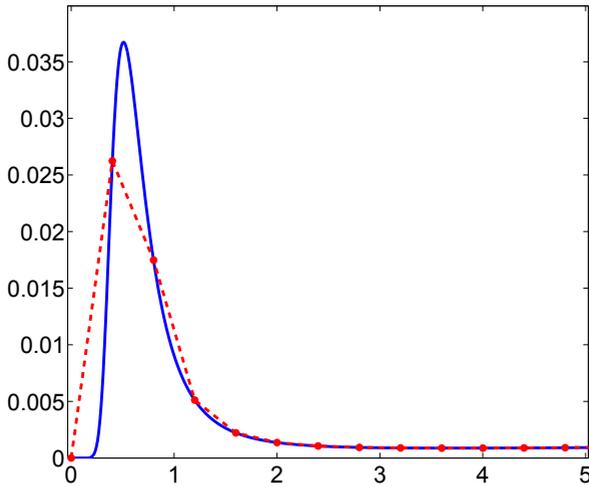


Figure 6.4.: A zoom of the right plot in Figure 6.3 focussing on the values $x \in [0, 5]$. The blue line shows the real difference between the two propagations, while the red dots and lines are hinting the representation via a grid. The grid is chosen inordinately coarse to make the illustration of our point particularly clear. The integral over the grid function in general does not yield the same result as the integral over the real function.

6.3. Output sensitivity

In the previous section we studied the sensitivity operator for heterogeneous cell populations without output. We came to the conclusion, that the sensitivity is an isometry and thus is independent of the “direction” of the variation. The question we address now is whether the same holds true for heterogeneous cell populations with output.

To start with, we recall that the output density is obtained by first solving the advection equation

$$\begin{aligned} \frac{\partial}{\partial t} p(t, x) + \operatorname{div}(p(t, x)F(x)) &= 0, \\ p(0, x) &= p_0, \end{aligned}$$

and then by marginalization

$$p_H(t, y) = \int_{H^{-1}(\{y\})} p(t, x) \, dS.$$

In this section we study the sensitivity of cell population models with output, i.e. the sensitivity of the mapping

$$p_0 \mapsto p_H.$$

For this purpose we first introduce the (well-defined) mapping

$$P_H : p(t, \cdot) \mapsto p_H(t, \cdot),$$

where $p_H(t, \cdot)$ is given by the point-wise almost everywhere definition (4.1). We note that P_H corresponds to the mapping $y = Hx$ in the same way as P_t corresponds to $x(t) = \Phi_t x_0$. This correspondence is illustrated in Figure 6.5.

The mapping $p_0 \mapsto p_H(t, \cdot)$ can be expressed using the operators P_t and P_H (see Figure 6.5) as

$$p_H(t, \cdot) = P_H P_t p_0.$$

The mapping P_H is linear and to see that it is further a bounded operator, we compute

$$\begin{aligned} \int_{\mathbb{R}^m} |p_H(t, y)| \, dy &= \int_{\mathbb{R}^m} \left| \int_{H^{-1}(\{y\})} p(t, x) \, dS \right| \, dy, \\ &\leq \int_{\mathbb{R}^m} \int_{H^{-1}(\{y\})} |p(t, x)| \, dS \, dy = \int_{\mathbb{R}^n} |p(t, x)| \, d\mu. \end{aligned}$$

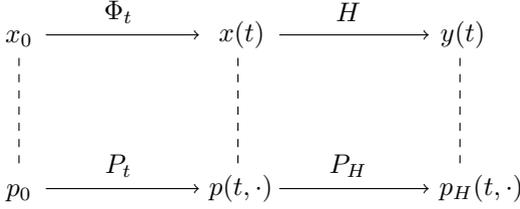


Figure 6.5.: This diagram illustrates the connection between the mappings Φ_t and P_t and the connection between H and P_H . For example, while $x(t)$ is obtained from x_0 through the mapping Φ_t , the corresponding density of $x(t)$ is obtained from the density of x_0 through the mapping P_t .

To sum up, we have shown that $\|P_H\| \leq 1$. Therefore the operator $P_H P_t$ is the composition of two linear and bounded operators, hence also linear and bounded. Its operator norm can further be computed as

$$\|P_H P_t\| \leq \|P_H\| \|P_t\| \leq 1.$$

Again by linearity and boundedness, we have for the output sensitivity operator

$$S_{t,H} = (P_H P_t)' = P_H P_t.$$

However, in contrast to the sensitivity of the states, the output sensitivity is in general not an isometry. To see this, we consider the linear case

$$\dot{x} = Ax, \quad y = Hx.$$

Here we have

$$P_H P_t h = \int_{(He^{At})^{-1}(\{y\})} h \, dS.$$

It is not hard to see that

$$P_H P_t h = 0 \quad \text{if and only if} \quad \forall B_y \in \mathcal{B}(\mathbb{R}^m) : \int_{(He^{At})^{-1}(B_y)} h \, d\mu = 0. \quad (6.1)$$

In the study of structural identifiability we have seen that such h always exists, e.g. pick an arbitrary h from the set

$$\{h = p_0 - p_0(\cdot + v) \in L^1 \mid p_0 \text{ a probability density and } v \in \ker He^{At}\}.$$

The equivalence in (6.1) makes a bridge between the output sensitivity operator and structural identifiability. We study the connection in the next section.

6.4. Output sensitivity and structural identifiability

In the previous section we introduced sensitivity analysis also for heterogeneous cell populations with outputs. By focussing on the *linear case* and using the equivalence (6.1) we have seen that the output sensitivity is not an isometry like the sensitivity operator in the case that outputs are neglected. It is clearly the case that there is a connection between the output sensitivity and structural identifiability. In this section we formulate and prove the precise connection of both concepts.

Theorem 6.4. *Let us be given a heterogeneous cell population model. The following statements are equivalent:*

- (a) *the heterogeneous cell population is structurally identifiable,*
- (b) *for all $h \in L^1, h \neq 0$, there exists $t \geq 0$ such that $S_{t,H}h \neq 0$.*

Proof. By definition of structural identifiability and by virtue of (6.1) a heterogeneous cell population is structurally identifiable if and only if for $h \in L^1$

$$(\forall t \geq 0 P_H P_t h = 0) \Rightarrow h = 0.$$

This is equivalent to

$$\text{for all } h \in L^1, h \neq 0, \text{ there exists } t \geq 0 : P_H P_t h \neq 0.$$

Lastly, we have $S_{t,H} = P_H P_t$ and the claim follows. \square

We conclude our study of the output sensitivity operator with a short summary. Motivated by the question of how a variation in the initial density is propagated to the output density, we derived the output sensitivity operator. This operator was shown to be well-behaved, which is a satisfying result in view of *uncertainty analysis*. Uncertainty analysis in our framework was the study of the resulting uncertainty in the output density, when given an uncertain initial density (e.g. due to misspecification in the parameter estimation process).

Our second result in form of Theorem 6.4 is important for the inverse problem of determining the initial density p_0 from the output density at all times, i.e.

$$p_H : (t, y) \mapsto p_H(t, y).$$

By introducing the operator $K : p_0 \mapsto p_H$, we can write the inverse problem as

$$Kp_0 = p_H.$$

Theorem 6.4 now states that the operator K is injective if and only if the heterogeneous cell population is structurally identifiable. Injectivity is typically an important requirement for the *regularization* of inverse problems.

7. Conclusions

7.1. Summary and discussion

In this thesis, we presented a new perspective in the modeling and analysis of heterogeneous cell populations. In the first part of the thesis, we gave a survey on two existing modeling frameworks for general populations. A simple one-dimensional biological example was used to illustrate the approaches. In one of the approaches, i.e. the density-based approach, a heterogeneous cell population is modeled by the probability density of the cell states. The evolution of this probability density is given in terms of a partial differential equation (PDE). The PDE-based formulation provides a powerful theoretical framework in which important questions, regarding heterogeneous cell populations can be studied.

In Chapter 3, we extended these previously known results to a measure theoretical framework which is novel in the study of heterogeneous cell population models. In this framework we showed that the PDE is generating a flow, which describes how densities are evolving under an ordinary differential equation. The flow is given in terms of Frobenius-Perron operators and each transformation of the flow, i.e. each Frobenius-Perron operator has important properties that we exploited in the subsequent studies.

In the second part of this thesis we considered two important concepts for heterogeneous cell population models with outputs, namely structural identifiability and sensitivity analysis. In Chapter 5, we introduced the concept of structural identifiability for heterogeneous cell population models, which is motivated by parameter estimation. We demonstrated in an illustrative example how structural identifiability of a heterogeneous cell population model is related to observability of the underlying single cell model. Based on this example we then proved that observability of the underlying single cell model is necessary for the structural identifiability of its cell population. In addition, the example is used as a plausibility argument for sufficiency.

In Chapter 6, we implemented sensitivity analysis for the class of heterogeneous cell population models. We computed the sensitivity operator based on our knowledge, that the evolution of an initial distribution is given in terms of Frobenius-Perron operators. We showed that the sensitivity is well-behaved, independently of the underlying single cell model and independently of the variation in the initial density. For heterogeneous cell population models with output we showed that the sensitivity is still well-behaved, but now does depend on the underlying single cell model and also on the variation in the initial density. Lastly, we established the connection between output sensitivity and structural identifiability.

To conclude, in this thesis we contributed to the emerging subject of heterogeneous cell population models by introducing a novel approach using Frobenius-Perron operators, by introducing an identifiability property (structural identifiability), and finally, by establishing sensitivity analysis for heterogeneous cell population models. In the following section we discuss open problems and possible research directions for future research.

7.2. Outlook

In Chapter 4, we presented the problem of determining the initial density from output measurements at specific sampling times. There we briefly mentioned an alternative problem of determining the initial density from the knowledge of the function $p_H : (t, y) \mapsto p_H(t, y)$. This problem did not receive any attention so far, possibly because a functional analytical framework was missing. However, with the novel introduction of a formalism based on operators in this thesis, such framework is now available. Thus, one could begin to study whether classical solution techniques for inverse problems (e.g. regularization) can be applied to our framework.

Another problem that we left open in this thesis is a proof the conjectured sufficient condition for structural identifiability. It would be desirable to formalize our plausibility argument for a complete proof. Furthermore, in the study of structural identifiability we were dealing with the linear case only. The use of linear single cell models can be certainly justified in a lot of practically important cases ¹, but it would also be important to establish similar results for the nonlinear case.

¹For example, in early work on modeling heterogeneous cell populations with PDEs, the single cell models were chosen extremely simple.

As a last open point we mention the fact that we did not take division and death events within cell populations into account. When studying the long-time behavior of heterogeneous cell populations, however, these events must be taken into account. Following this inclusion, tasks like parameter estimation, analysis of identifiability properties and sensitivity analysis, need to be adjusted for the future frameworks.

A. Mathematical background

In this chapter we shall briefly review some important mathematical concepts that we make use of in the thesis. Particularly we make extensive use of concepts from the theory of dynamical systems, probability theory and multivariate calculus. The following material however is kept very brief and is not intended to be a first introduction to the subjects.

A.1. Dynamical systems

This section follows the presentation of [14]. Recall that a continuous time process in a phase space Ω is given by a family of mappings

$$\Phi_t : \Omega \rightarrow \Omega, \quad t \geq 0.$$

An illustration is given in Figure A.1. Here the value $\Phi_t(x_0)$ is the position of the system at a time t that started from an initial point $x_0 \in X$ at time $t = 0$. We consider only those processes satisfy the property

$$\Phi_t(\Phi_{t'}(x)) = \Phi_{t+t'}(x).$$

This property essentially says that the dynamics governing the evolution of the system are the same on the intervals $[0, t']$ and $[t, t + t']$.

Example A.1. A well-known example of a continuous time process is given by an autonomous system of differential equations

$$\dot{x} = F(x).$$

Under the assumptions that the vector field F is “sufficiently nice” we can infer existence and uniqueness of solutions. In this case, $\Phi_t(x_0)$ is the solution with initial condition x_0 . Also note that in this example t need not to be restricted to $t \geq 0$, and the system can also be studied for $t \leq 0$.

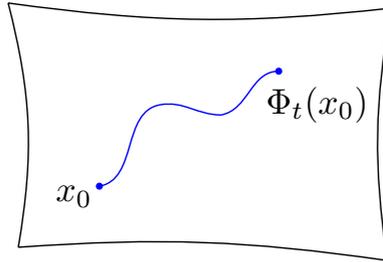


Figure A.1.: A trajectory of a continuous time process in the phase space. At time $t = 0$ the system is at the state x_0 , and at time t it is at the state $\Phi_t(x_0)$.

A flow on the phase space Ω is a mapping

$$\Phi : \mathbb{R} \times \Omega \rightarrow \Omega, \quad (t, x) \mapsto \Phi_t x$$

such that, for all $x \in \Omega$ and all real numbers t and t' it holds that

- (i) $\Phi_0 x = x$,
- (ii) $\Phi_t(\Phi_{t'} x) = \Phi_{t+t'} x$.

Remark A.2. It is clear from the group property in the definition of a flow that

$$\Phi_t(\Phi_{-t}(x)) = x \text{ and } \Phi_{-t}(\Phi_t(x)) = x \text{ for all } t \in \mathbb{R}$$

Thus, for all $t_0 \in \mathbb{R}$, any transformation Φ_{t_0} is invertible.

The system of ordinary differential equations, introduced before, is clearly an example of a dynamical system. For the linear case $F(x) = Ax$ it is well-known that

$$\Phi_t x = e^{At} x.$$

For every fixed $x_0 \in X$, the function $t \mapsto \Phi_t(x_0)$ is called a trajectory of the system. Trajectories of a dynamical system cannot intersect as depicted in Figure A.2.

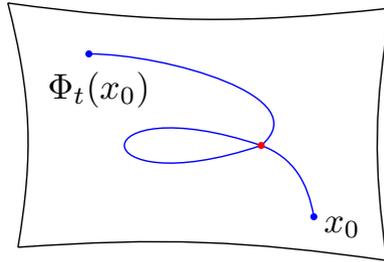


Figure A.2.: The intersecting trajectory (see red dot) shown here is not possible in a dynamical system.

To demonstrate this, assume to the contrary, that for given $x_0 \in \Omega$ we have

$$\Phi_{t_1}(x_0) = \Phi_{t_2}(x_0) \quad t_2 > t_1.$$

By applying Φ_{t-t_1} to both sides of this equation, we have

$$\Phi_{t-t_1}(\Phi_{t_1}(x_0)) = \Phi_{t-t_1}(\Phi_{t_2}(x_0)).$$

By the group property we also have

$$\Phi_{t-t_1}(\Phi_{t_1}(x_0)) = \Phi_t(x_0) \text{ and } \Phi_{t-t_1}(\Phi_{t_2}(x_0)) = \Phi_{t+(t_2-t_1)}(x_0).$$

Hence with $\omega = t_2 - t_1$ this leads to

$$\Phi_t(x_0) = \Phi_{t+\omega}(x_0) \text{ for all } t \in \mathbb{R},$$

implying that the only possible “intersecting” trajectories of a dynamical system are periodic. We shall note that neither can the trajectories $t \mapsto \Phi_t(x'_0)$ and $t \mapsto \Phi_t(x''_0)$ for $x'_0 \neq x''_0$ cross.

There is a connection between the vector field and the dynamical system. One can generate the dynamical system with the vector field. To be more precise, if Φ is differentiable in t , then we can assign it a unique vector field F by defining

$$F(x) = \frac{d}{dt} \Phi_t(x)|_{t=0}, \quad x \in \Omega.$$

The trajectories of the system are solution curves of the vector field. This is because for any t we have

$$\frac{d}{dt} \Phi_t(x)|_t = \frac{d}{ds} \Phi_{s+t}(x)|_{s=0} = \frac{d}{ds} \Phi_s(\Phi_t(x))|_{s=0} = F(\Phi_t(x)).$$

A.2. Probability theory

For a first introduction to probability theory we refer to [29] and the references therein. In probability theory we model randomness through a mathematical structure called probability space. A probability space is a triple $(\Omega, \mathcal{B}, \mathbb{P})$ — a set Ω together with a σ -algebra \mathcal{B} of subsets of Ω and a probability measure \mathbb{P} . Loosely speaking, Ω is the space of samples and the elements of \mathcal{B} are identified as events. A probability measure \mathbb{P} assigns every event B a value $P(B) \in [0, 1]$, called the probability of B . Let us now get rigorous.

A collection \mathcal{B} of subsets of Ω is called a σ -algebra, if

- (i) $B \in \mathcal{B} \implies (\Omega \setminus B) \in \mathcal{B}$
- (ii) for any countable collection $\{B_k\}_{k \in I}$ of $B_k \in \mathcal{B}$ we have $\bigcup_{k \in I} B_k \in \mathcal{B}$
- (iii) $\Omega \in \mathcal{B}$

A real-valued function μ defined on a σ -algebra \mathcal{B} is called a measure, if

- (i) $\mu(\emptyset) = 0$
- (ii) $\mu(B) \geq 0$ for all $B \in \mathcal{B}$
- (iii) for any countable collection $\{B_k\}_{k \in I}$ of pairwise disjoint $B_k \in \mathcal{B}$ we have $\mu(\bigcup_{k \in I} B_k) = \sum_{k \in I} \mu(B_k)$

If \mathcal{B} is a σ -algebra of subsets of Ω and if μ is a measure on \mathcal{B} , then the triple $(\Omega, \mathcal{B}, \mu)$ is called a measure space. Sets belonging to \mathcal{B} are called measurable sets. A measure space $(\Omega, \mathcal{B}, \mu)$ with total measure one, i.e. $\mu(\Omega) = 1$ is called probability space.

For a metric space (M, d) , the Borel σ -algebra $\mathcal{B}(M)$ is defined to be the smallest σ -algebra that contains all open sets of M .

A measure space is called σ -finite if there is a sequence $\{B_k\}_{k \in \mathbb{N}}, B_k \in \mathcal{B}$, satisfying

$$\Omega = \bigcup_{k \in \mathbb{N}} B_k \quad \text{and} \quad \mu(B_k) < \infty \quad \text{for all } k.$$

A certain property involving sets $B \in \mathcal{B}$ is called true almost everywhere, if there exists a set $E \in \mathcal{B}$ with measure zero such that the set for which the property does not hold is contained in E .

A function $h : \Omega \rightarrow \mathbb{R}$ is called measurable if $h^{-1}(\Delta) \in \mathcal{B}$ for all $\Delta \subset \mathbb{R}$, i.e. if for all $\Delta \subset \mathbb{R}$ the pre-image of Δ under h is a measurable set.

Let (Ω', \mathcal{B}') be a measurable space. A measurable function $X : \Omega \rightarrow \Omega'$ is called a random variable. Furthermore

$$P_X : \mathcal{B}' \rightarrow [0, 1], \quad A' \mapsto P(X^{-1}(A'))$$

is a probability measure on (Ω', \mathcal{B}') . It is called pushforward measure of X or also distribution of X .

We skip Lebesgue integration and continue by introducing L^p -spaces. Let $(\Omega, \mathcal{B}, \mu)$ be a measure space and let $f, g : \Omega \rightarrow \mathbb{R}$ be measurable functions. We define the equivalence relation

$$f \sim g \Leftrightarrow f(x) = g(x) \text{ almost everywhere,}$$

which is just following the philosophy that one can ignore sets of measure zero. The equivalence relation yields us equivalence classes \hat{f} . Let p be a real number, $1 \leq p < \infty$. We define

$$\hat{f} \in L^p(\Omega, \mu) \Leftrightarrow \int_{\Omega} |f(x)|^p d\mu < \infty \text{ for some } f \in \hat{f}.$$

Remark A.3. Strictly speaking, elements of $L^p(\Omega, \mu)$ are not functions $f : \Omega \rightarrow \mathbb{R}$ in the set theoretic sense. One can no longer evaluate an element at a single point x (if that point has measure zero). Therefore elements of $L^p(\Omega, \mu)$ are sometimes called generalized functions for convenient reasons that we will not further discuss. If however you find in the thesis an evaluation $f(x)$ of $f \in L^p(\Omega, \mu)$, we mean that a representative of the equivalence class is evaluated. This is a slight abuse of notation, but it is commonly made.

Lastly, we equip this vector space with the norm $\|f\|_{L^p} := \left(\int_{\Omega} |f|^p d\mu \right)^{\frac{1}{p}}$.

A.3. Multivariable calculus

Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed linear vector spaces. Let $D_T \subset E$ be a linear subspace over a field \mathbb{K} . A function $T : D_T \rightarrow F$ is called linear, if

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$$

for all $x, y \in D_T$ and $\alpha, \beta \in \mathbb{K}$.

One of the many equivalent definitions for *continuity* of T is via series, i.e.

$$\forall y_k \in D_T; y_k \xrightarrow{\|\cdot\|_E} y \in D_T : T(y_k) \xrightarrow{\|\cdot\|_F} T(y).$$

A linear operator $T : D_T \subset E \rightarrow F$ is called bounded, if there exists a $C < \infty$ such that

$$\|Tx\|_F \leq C\|x\|_E \text{ for all } x \in D_T.$$

As a matter of fact, a linear operator is continuous if and only if it is bounded.

Now let $D_T = E$. Then we define

$$L(E, F) := \{T : E \mapsto F \mid T \text{ is linear and bounded}\},$$

which is a vector space. It is called the space linear and bounded operators from E to F . The canonical norm on $L(E, F)$ is given by

$$\|T\|_{L(E, F)} = \sup_{x \neq 0, x \in E} \frac{\|Tx\|_F}{\|x\|_E}.$$

This norm is called operator norm. We have $\|Tx\|_F \leq \|T\|_{L(E, F)}\|x\|_E$ and $\|T\|_{L(E, F)}$ gives us the smallest constant that satisfies the estimate in the definition of a bounded operator. As a matter of fact, $(L(E, F), \|\cdot\|_{L(E, F)})$ is a normed space and if furthermore $(F, \|\cdot\|_F)$ is a Banach space then so is $(L(E, F), \|\cdot\|_{L(E, F)})$.

For $T \in L(E, F)$ and $S \in L(F, G)$ the composition $ST \in L(E, G)$ satisfied the estimate

$$\|ST\|_{L(E, G)} \leq \|S\|_{L(F, G)} \cdot \|T\|_{L(E, F)}.$$

We are now ready for multivariable differential calculus. Let f be a function that is defined on an open subset U of E , taking values in F , i.e.

$$f : U \subset E \rightarrow F.$$

Since f is in general a nonlinear function defined on possibly infinite-dimensional spaces, studying f directly is quite hard. Therefore, we wish to approximate for a given $x_0 \in E$ the behavior of f in a neighborhood of x_0 by a linear and bounded operator. Those are of course much easier to study and to handle. This directly motivates the following definition.

A function $f : U \subset E \rightarrow F$ is called *differentiable* in $x_0 \in E$ if there exists a linear and bounded operator $T_{x_0} \in L(E, F)$, such that we have the following decomposition

$$f(x_0 + h) = f(x_0) + T_{x_0}h + o(\|h\|), \quad h \rightarrow 0. \quad (\text{A.1})$$

Then we denote $T_{x_0} =: f'(x_0)$ the *Frechet derivative* of the function f in the point x_0 .

Example A.4. As an example consider $f : U \subset \mathbb{R} \rightarrow \mathbb{R}$. In this particular case, linear mappings are either elongations or compressions and thus

$$T_{x_0}h = c(x_0)h,$$

with a constant $c(x_0) \in \mathbb{R}$ that depends on the point $x_0 \in U$. We can rewrite (A.1) as

$$f(x_0 + h) = f(x_0) + c(x_0)h + o(\|h\|), \quad h \rightarrow 0.$$

From this we can directly deduce that

$$c(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

which yields $c(x_0) = \frac{d}{dx}f|_{x=x_0}$, where $\frac{d}{dx}f|_{x=x_0} := \lim_{h \rightarrow 0} \frac{f(x_0+h) - f(x_0)}{h}$.

To sum up, the Frechet derivative of a scalar function is given by the linear and bounded operator $h \mapsto \frac{d}{dx}f|_{x=x_0}h$.

B. Matlab functions

In this chapter we provide for the reader the Matlab functions that we wrote for some of our examples.

B.1. Example Monte Carlo method

For the example of Monte Carlo method, the code is the following:

Listing B.1: Monte Carlo simulation

```
sample_size = 5000;
n = normrnd(1,sqrt(0.07),[1 sample_size]);
hist(n,sample_size/100);

for j = 1:sample_size
    [t,x]=ode45('vector_field', [0 50], n(j));
    p(j) = x(end);
end
hist(p,sample_size/100);
```

Listing B.2: Vector field of differential equation

```
function dxdt = vector_field(t,x)
    Vmax = 1;           % [mol/(volume*time)]
    K = 2;              % [mol/volume]
    b = 4;              % [dimensionless]
    kd = 0.01;         % [1/time]

    dxdt = (Vmax*x^b)/(K^b+x^b) - kd*x;
end
```

B.2. Example method of characteristics

The code for the method of characteristics without two-steps is given by:

Listing B.3: Method of characteristics

```
clear; clc; close all;

% create uniform grid on interval [a,b] with grid size h
a = 0;
b = 5;
h = 0.1;

x0 = a:h:b;

% specify end time
T = 15;

% specify initial density
p0 = mvnpdf(x0', 1, 0.07)';

% solve MOC equation forward in time
for i = 1:size(x0,2)
    [t,u]=ode45('vector_field_moc', [0 T], [x0(i);p0(i)]);
    end_density(:,i) = u(end,:);

    % plot curves t -> (t,x_i(t),p(t,x_i(t)))
    plot3(t,u(:,1),u(:,2));
    hold on;
end

% plot initial density in 3d plot
plot3(zeros(length(x0),1), x0',p0');

% plot end density in 3d plot
plot3(T*ones(length(x0),1),end_density(1,:), end_density(2,:));

% plot end density seperately in 2d plot
figure
plot(end_density(1,:), end_density(2,:));
```

For demonstrating the method of characteristics using the two-step method, the code is the following:

Listing B.4: Method of characteristics (two-step procedure)

```
clear; clc; close all;

% create uniform grid on interval [a,b] with grid size h
a = 0;
b = 15;
h = 0.1;

xT = a:h:b;

% specify end time
T = 15;

% For every point in the end grid xT(i) solve x'=-F(x) with initial
% condition xT(i). Store x(T)=x0(i).
for i = 1:size(xT,2)
    [t,w]=ode45('vector_field', [T 0], xT(i));
    x0(i) = w(end);
end

% specify initial density
p0 = mvnpdf(x0', 1, 0.07)';

% solve MOC equation forward in time with initial conditions
% x0(i) from the previous backwards solving
for i = 1:size(xT,2)
    [t,u]=ode45('vector_field_moc', [0 T], [x0(i);p0(i)]);
    end_density(:,i) = u(end,:);

    % plot curves t -> (t,x_i(t)),p(t,x_i(t))
    plot3(t,u(:,1),u(:,2));
    hold on;
end

% plot initial density in 3d plot
plot3(zeros(length(x0),1), x0',p0');

% plot end density in 3d plot
plot3(T*ones(length(x0),1),end_density(1,:), end_density(2,:));

% plot end density separately in 2d plot
figure
plot(end_density(1,:), end_density(2,:));
```

B. Matlab functions

Listing B.5: Vector field for MOC system

```
function dudt = vector_field_moc(t,u)

    Vmax = 1;           % [mol/(volume*time)]
    K = 2;             % [mol/volume]
    b = 4;             % [dimensionless]
    kd = 0.01;        % [1/time]

    % we define u = [x;M], that is u1 = x and u2=M. Then, when given
    % initial conditions x0, the method of characteristics by solving
    %
    %           dudt = [F(u1); (div F)(u1)*u2]
    %
    dudt(1) = (Vmax*u(1)^b)/(K^b+u(1)^b) - kd*u(1);
    dudt(2) = -(Vmax*b*u(1)^(b-1))/(K^b + u(1)^b) ...
               -kd-(Vmax*b*u(1)^b*u(1)^(b-1))/(K^b + u(1)^b)^2)*u(2);

    dudt = dudt';

end
```

B.3. Example sensitivity analysis

For the sensitivity analysis example (Example 6.2) the code is the following:

Listing B.6: Sensitivity analysis

```
clear; clc; close all;

% create uniform grid on interval [a,b] with grid size h
a = 0;
b = 15;
h = 0.1;

xT = a:h:b;

% specify end time
T = 15;

% For every point in the end grid xT(i) solve x'=-F(x) with initial
% condition xT(i). Store x(T)=x0(i).
for i = 1:size(xT,2)
    [t,w]=ode45('vector_field', [T 0], xT(i));
    x0(i) = w(end);
end

% specify initial density
p0 = mvnpdf(x0', 1, 0.07)';

% solve MOC equation forward in time with initial conditions
% x0(i) from the previous backwards solving
for i = 1:size(xT,2)
    [t,u]=ode45('vector_field_moc', [0 T], [x0(i);p0(i)]);
    end_density(:,i) = u(end,:);

    % plot curves t -> (t,x_i(t),p(t,x_i(t)))
    plot3(t,u(:,1),u(:,2));
    hold on;
end

% plot initial density in 3d plot
plot3(zeros(length(x0),1), x0',p0');

% plot end density in 3d plot
plot3(T*ones(length(x0),1),end_density(1,:), end_density(2,:));

% plot end density seperately in 2d plot
figure
plot(end_density(1,:), end_density(2,:));
```


Bibliography

- [1] S. Waldherr and F. Allgöwer. *Computational modeling, sensitivity analysis and parameter identification for heterogeneous cell populations*. Application for a DFG Research Grant (WA 2965/1-1), 2010.
- [2] B.B. Aldridge, J.M. Burke, D.A. Lauffenborg and P.K. Sorger. *Physicochemical modelling of cell signalling pathways*. Nat. Cell. Biol. 8:1195-1203, 2006.
- [3] S.J. Altschuler and L.F. Wu. *Cellular Heterogeneity: Do Differences Make a Difference?* Cell, Volume 141, Issue 4, Pages 559-563, 2010.
- [4] J. Hasenauer, S. Waldherr, M. Doszczak, P. Scheurich, N. Radde and F. Allgöwer. *Analysis of heterogeneous cell populations: A density-based modeling and identification framework*. Journal of Process Control 21:1417-1425, 2011.
- [5] S. Waldherr, J. Hasenauer and F. Allgöwer. *Estimation of biochemical network parameter distributions in cell populations*. In Proc. of the 15th IFAC Symp. on Syst. Ident., 1265-1270, 2009.
- [6] M.A. Henson, D. Müller, M. Reuss. *Cell population modeling of yeast glycolytic oscillations*. Biochem. J. 368:433-446, 2002.
- [7] A. Fredrickson, D. Ramkrishna and H. Tsuchiya. *Statistics and dynamics of procaryotic cell populations*. Math. Biosci. 1:327-374, 1967.
- [8] A. Fredrickson and N.V. Mantzaris. *A new set of population balance equations for microbial and cell cultures*. Chem. Eng. Sci. 57:2265-2278, 2002.
- [9] M. Stamatakis. *Cell population balance, ensemble and continuum modeling frameworks: Conditional equivalence and hybrid approaches*. Chem. Eng. Sci. 65:1008-1015, 2010.
- [10] A.Y. Weisse, R.H. Middleton and W. Huisinga. *Quantifying uncertainty, variability and likelihood for ordinary differential equation models*. BMC Systems Biology, 2010.

- [11] J. Hasenauer, S. Waldherr, M. Doszczak, P. Scheurich and F. Allgöwer. *Density-based modeling and identification of biochemical networks in cell populations*. In Proc. of the 9th IFAC Symp. Dynam. Contr. Proc. Syst., pp. 306-311. Leuven, Belgium, 2010.
- [12] J. Goodman. *Lecture notes: PDE in Finance, Section 3*. 2008.
- [13] V. Constanza, J.H. Seinfeld. *Stochastic sensitivity analysis in chemical kinetics*. J. Chem. Phys. 74, 3852, 1981.
- [14] A. Lasota and M. C. Mackey. *Chaos, Fractals, and Noise*. Springer, 1994.
- [15] L. Grüne. *Lecture notes: Mathematische Kontrolltheorie I: Lineare Systeme*. 2009.
- [16] H.W. Knobloch, H. Kwakernaak. *Lineare Kontrolltheorie*. Springer-Verlag, 1985.
- [17] A. Charzynska, A. Nalecz, M. Rybinski and A. Gambin. *Sensitivity analysis of mathematical models of signaling pathways*. BioTechnologia, 93:291-308, 2012.
- [18] R. Heinrich and S. Schuster. *The Regulation of Cellular Systems*. Chapman & Hall, New York, 1996.
- [19] H. Rabitz, M. Kramer and D. Dacol. *Sensitivity analysis in chemical kinetics*. Ann. Rev. Phys. Chem. 34:419-461, 1983.
- [20] E. Gine, D. M. Mason and A. Y. Zaitsev. *The L_1 -Norm Density Estimator Process*. The Annals of Probability Vol. 31, No. 2, pp. 719-768, 2003.
- [21] J. Hasenauer. *Modeling and parameter estimation for heterogeneous cell populations*. Dissertation, Universität Stuttgart, 2012.
- [22] J. Hasenauer, S. Waldherr, N. Radde, M. Doszczak, P. Scheurich, and F. Allgöwer. *A maximum likelihood estimator for parameter distributions in heterogeneous cell populations*. In Procedia Computer Science, vol. 1, pp. 1649-1657.
- [23] J. Hasenauer, S. Waldherr, M. Doszczak, N. Radde, P. Scheurich and F. Allgöwer. *Identification of models of heterogeneous cell populations from population snapshot data*. BMC Bioinf., 12(125), 2011.
- [24] T. Luzyanina, D. Roose, T. Schenkel, M. Sester, S. Ehl, A. Meyerhans and G. Bocharov. *Numerical modelling of label-structured cell population growth using CFSE distribution data*. Theor. Biol. Med. Model. 4:26, 2007.

- [25] H.T. Banks, K.L. Sutton, W.C. Thompson, G. Bocharov, D. Roose, T. Schnekel and A. Meyerhans. *Estimation of cell proliferation dynamics using CFSE data*. Bull Math Bio, 2010.
- [26] T. Weidl. *Lecture notes: Analysis II*. 2010.
- [27] T. Weidl. *Lecture notes: Analysis III*. 2010.
- [28] J. Pöschel. *Lecture notes: Dynamische Systeme*. 2009.
- [29] I. Steinwart. *Lecture notes: Wahrscheinlichkeitstheorie*. 2010.
- [30] C. Ebenbauer. *Lecture notes: Einführung in die Regelungstechnik*. 2011.

Eigenständigkeitserklärung

Ich versichere hiermit, dass ich, Shen Zeng, die vorliegende Arbeit selbstständig angefertigt, keine anderen als die angegebenen Hilfsmittel benutzt und sowohl wörtliche, als auch sinngemäß entlehnte Stellen als solche kenntlich gemacht habe. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen.

Ort, Datum

Unterschrift