Chapter 3

Introduction to Stochastic Models and Markov Chains

The main topic of this thesis is the investigation of particle transport in various types of fluidized bed reactors. We propose to study transport phenomena with the help of mathematical models for the motion of individual particles. This motion is complex and depends on the interaction with a large number of other particles in the reactor, so that a deterministic model is not feasible. In this situation we advocate the use of stochastic models. This chapter gives an overview of the mathematical models used in chemical engineering and a discussion of the stochastic modeling approach used in this thesis.

3.1 Mathematical Modeling

The power of present-day computers and the progress in the area of numerical algorithms is such that the formerly long and tedious procedure of numerical calculation now takes only seconds. This has led to an increasing popularity of mathematical modeling of processes in engineering science. The great advantage of mathematical modeling lies in its potential to partially replace the costly and time consuming process of the actual construction of a test system. Normally a process model of a real system on a smaller scale is built so the scientist can observe the behavior in a replica for which she/he can control all the variables. The mathematical model that results from these small-scale experiments is then a reasonable and quantitative description of these observations, and in some cases a mathematical model can replace the small-scale physical model itself.

In general a mathematical model is not able to represent all the details of an actual process but is able to describe the most important and relevant observable features of the system, which constitutes sufficient knowledge. Also, modeling is an adaptive procedure in which information gained from each experimental trial is used to improve the next generation model. This flexibility is one of the features of the modeling process. The modeling process usually involves several stages,

- select a mathematical model,
- match the model to data,
- remodel the model, and
possibly additional experiments,

until the scientist is satisfied that she/he understands the system.

3.2 Deterministic vs. Stochastic Modeling

Mathematical models can roughly be divided into two categories, namely deterministic and stochastic models. Deterministic models are usually expressed in terms of differential equations that, together with initial and boundary conditions, exactly predict the development of a system. Stochastic models are given by random variables whose outcomes are uncertain and where we can only compute the probabilities of possible outcomes. Though useful, the division line between deterministic and stochastic modeling approaches is far from being sharp. Research in the area of nonlinear dynamical systems has shown that deterministic models can exhibit seemingly stochastic behavior. On the other hand, stochastic models can have deterministic behavior when viewed at the right scales or when computing aggregate quantities.

Mathematical models are tools to describe physical reality, and on this level the distinction between stochastic and deterministic phenomena is again not as sharp as it might look at first sight. The way it is often taught in elementary courses, that random experiments have an unpredictable outcome whereas classical deterministic experiments have a predictable outcome, is an oversimplification. For illustration, we consider two examples. Throwing fair dice is an experiment that is usually regarded as a prime example of a random experiment, to be modeled by a random variable that takes values 1, …, 6 with equal probability. On the other hand, a die is a solid body whose motion follows Newton’s laws of mechanics and which can be described by a system of differential equations. The outcome of a toss can be perfectly predicted if we know the initial conditions and the external forces that act on the dice. That the latter are almost always impossible to determine with a sufficient degree of accuracy, is the main reason for the prevalence of stochastic models for tossing of dice. As a second example, we look at the motion of molecules in a fluid. The motion of an individual molecule follows classical mechanical laws. There are however so many forces acting on any molecule, as a result of collisions with the other molecules, that it is practically impossible to predict the path of an individual particle. Thus a stochastic model, in this case Brownian motion, is the only reasonable approach. When studying particle motion at a macroscopic level, we find again deterministic phenomena such as diffusion transport that can very well be modeled by differential equations. For this case we argue that a full understanding of diffusion transport can only be obtained from models for the random motion of individual particles. This link has
been made in Albert Einstein 1905 paper on Brownian motion, which was one of two papers that were cited in the laudation for Einstein’s Nobel prize award.

When modeling particle transport in chemical reactors, the choice between a deterministic and a stochastic approach amounts to a choice between a macroscopic and a microscopic approach. In a microscopic approach, the object of study is the random path of a single particle through the reactor. In contrast, in a macroscopic approach we investigate the evolution of densities of a large number of marked particles. Traditional models in chemical engineering are macroscopic models, given by particle differential equations such as the diffusion equation

\[ \frac{\partial}{\partial t} p(t,x) = D \frac{\partial^2}{\partial x^2} p(t,x) \]

which governs the transport of marked particles in a liquid at rest.

Whether a deterministic or a stochastic model is preferable, depends among other things on the nature of the process to be modeled. Some indicators for the use of a stochastic modeling strategy can be given along the following lines:

- **The process has a strong element of random motion of material particles.** For such processes, stochastic modeling is intuitively appealing, and consistent with the nature of the process. A stochastic model also gives more information about statistical uncertainties involved in the process than a deterministic model.
- **The process is complex, and involves many discrete events.** Formulating deterministic models demands that the process is converted to one which is continuous in time and space. This is often not possible, or it involves sweeping assumptions. For a stochastic model this is not necessary, and the model can be formulated directly. The processes discussed in this thesis provide examples.

Fortunately, there are many works and examples that can guide model selection. Different approaches can lead to the same solution via different ways. All of them have the same goal, which is to simulate and to understand the real system as well as possible. The final analysis of which model should be selected involves considering which is most likely to lead to the aspect in which one is interested, and how much time and money is available for a description of the system.

### 3.3 Historical Remarks

Many mathematical models have been proposed for describing transport of particles in fluidized beds.
Early studies of solids mixing in fluidized beds were related to axial solid dispersion, i.e. May (1959), Rowe and Sutherland (1964), Wollard and Potter (1968) and Schügerl (1976) or radial solid dispersion, i.e. Cranfield (1978), Shi et al. (1985) and Tomeczek et al. (1992). The vast majority of the literature mentioned above is concerned with so-called diffusion or dispersion models by which the diffusion or dispersion coefficient have been estimated by fitting a dispersion equation to the data obtained experimentally.

Sitnai (1981) proposed a convection model for vertical solids mixing in one dimensional fluidized beds, with properties of bubbles, with the convection coefficient being constant. Sitnai’s convection model bears close similarities with the countercurrent backing model of Freyer and Potter (1978) and Kunii and Levenspiel (1969). Throughout years of development, most of fluidized bed models are based on the convection-diffusion model’s concept.

For mixing and segregation fluidized beds, the best known model was proposed by Gibilaro and Rowe (1974). The model was extended by Naimer et al. (1982) and several groups of researchers (Chiba et al. 1976, 1979; Chen, 1980; Tanimoto et al., 1981; Hoffmann et al., 1990, 1993 etc.). More details of the Gibilaro and Rowe model can be found in Chapter 5.

The literature concerned with the residence time distribution (RTD) of particles in continuous fluidized beds is also extensive. A number of researchers have tried to identify and model the physical phenomena governing the transport of particles in a fluidized bed, e.g. Berruti et al. (1988), Haines and King (1972) and Morris et al. (1964). Rowe and Partridge (1962) proposed the idea of solids mixing in fluidized beds. Hoffmann and Paarhuis (1990) extended Rowe and Partridge idea to the problem of the RTD of the solids in continuous beds. There is extensive literature on each specific type of fluidized beds, e.g. slugging fluidized beds, batch freely bubbling fluidized beds. The literature concerning each specific type of fluidized beds studied in this thesis can be found within the relevant chapter.

Most of the mentioned models are deterministic models based on conservation equations. For a complex system such as the transport of particles in fluidized beds, the application of the approach, however, often does not lead directly to a soluble model. We believe the stochastic model for the transport of an individual particle in a fluidized bed should be further exploited.

Our work builds on earlier work by Dehling et al. (1999) and Hoffmann et al. (1998) who used stochastic models of transport in continuous fluidized beds to compute residence time distributions. In a related context, stochastic processes
have been used to model horizontal transport of particles in sediment beds. A fundamental contribution in this field is H.A. Einstein’s 1936 thesis. Here a stochastic approach was motivated by the observation that, in repeated experiments, identical particles in the sediment of a river undergo very different displacements, whereas the overall distribution of a larger number of particles each time is very similar. The problems investigated and the type of model studied in that are, however, very different from those in this thesis.

Stochastic models focus on a single particle and model its path through the reactor as a result of random effects (including, for example, upward convection in wakes and segregation effects etc.), whereas deterministic models focus on densities, and thus on large numbers of particles. If a very large number of particles are involved, the law of large numbers guarantees a deterministic behavior of, for example, particle densities and this limit thus draws a bridge between microscopic and macroscopic models. In this case both models eventually yield the same results. If the number of particles chosen is small, a microscopic model provides more information concerning the fluctuations about the mean behavior. In many cases, a microscopic model is conceptually easier to formulate as it concentrates on a simple and physically intuitive object, namely the path of a single particle.

Stochastic processes model the development of a random system over the course of time. They obtain their name from the Greek word for guessing, ‘stokhastikos’. A stochastic process is a family \((X_t)_{t \geq 0}\) of random variables, indexed by the parameter \(t\). In our applications, \(t\) denotes time, and \(X_t\) describes the state of the system, or same aspect of it, at time \(t\). The set of possible values of \(X_t\) is called state space, and will be denoted by \(S\). The state can e.g., be the location of a given particle in the reactors in which case \(S\) is some bounded subset of \(\mathbb{R}^3\).

Stochastic processes can be categorized according to the possible values of \(t\) and \(X_t\), as follows:

- Discrete state, discrete time stochastic processes
- Discrete state, continuous time stochastic processes
- Continuous state, discrete time stochastic processes
- Continuous state, continuous time stochastic processes

The nature of the physical processes considered in this thesis calls for models with continuous state space and in continuous time. Our initial models will, however, always be discrete state, discrete time models. Such models are easier to formulate and moreover they allow direct numerical calculations and simulations. We can
obtain continuous models by letting both space and time discretization converge to zero.

Generally, there are two types of stochastic processes, independent and dependent processes. Independence would be a good model for such systems as repeated experiments, in which future states of the system are independent of past and present states. In complex systems, such as chemical engineering processes including fluidization, generally past or/and present states of the system affect the future states.

### 3.4 Markov Processes

The simplest type of models that exhibit dependence are Markov processes, where the future development of the process only depends on the present value, but not on the past. In many systems, indeed, the past of the state does not influence the future of the system and its future state only depends on the present. This characteristic, called the Markov property, can be expressed by

\[
P(X_{t+1} \in A \mid X_0 = x_0, \ldots, X_n = x_n) = P(X_{t+1} \in A \mid X_n = x_n),
\]

(3.1)

for all times \(0 = t_0 \leq t_1 \leq \ldots \leq t_n \leq t_{n+1}\) and all \(x_0, \ldots, x_n\) in the state space. \(p_{s,t}(x, A) := P(X_t \in A \mid X_s = x)\) is called the transition kernel of the process. It gives the probability that the process takes at time \(t\) a value in \(A\) given that at time \(s\) its value was \(x\).

A Markov process is called time invariant if \(p_{s,t}(x, A)\) only depends on the difference \(t - s\). In this case we thus have a one-parameter family of transition kernels \(p_t(x, A) := P(X_{s+t} \in A \mid X_s = x)\). In what follows we will restrict attention to time-invariant processes.

A Markov process is completely specified by its transition kernel and the initial distribution \(\pi\), i.e. the distribution of \(X_0\). We can then compute arbitrary probabilities concerning the process \((X_t)\), e.g., we have

\[
P(X_t \in A) = \int_A p_t(x, A) \, d\pi(x).
\]

Computations are much simpler for Markov chains with discrete time and discrete state space. In this case, condition (3.1) can be replaced by

\[
P(X_{n+1} = x_{n+1} \mid X_0 = x_0, \ldots, X_n = x_n) = P(X_{n+1} = x_{n+1} \mid X_n = x_n).
\]

(3.2)
Again assuming time invariance, the transitions of the process are completely governed by the one-step transition probabilities

\[ p(x, y) := P(X_{n+1} = y | X_n = x). \]

The associated matrix \[ P = (p(x, y))_{x, y \in S} \] is called the transition matrix. We often write the entries of this matrix as \( p_{xy} \) rather than \( p(x, y) \). If the initial state of the system is chosen according to the probability distribution \( P(X_0 = x) \), we can compute the distribution at time 1 using the law of total probability as follows:

\[
P(X_1 = x) = \sum_{x_0 \in S} P(X_0 = x_0) \cdot P(X_1 = x | X_0 = x_0)
\]

(3.3)

Introducing the probability vector \( p(n) := (p(n, x))_{x \in S} \), where \( p(n, x) := P(X_n = x) \), we can rewrite (3.3) in vector-matrix notation as \( p(1) = p(0) \cdot P \).

In the same way we obtain the general formula \( p(n) = p(n-1) \cdot P \) from where we get by induction \( p(n) = p(0) \cdot P^n \).

The distribution vector of the state of the system at time \( n \) can thus be computed by multiplying the vector of initial distributions \( n \) times from the right by the transition matrix.

We will adopt Markov chains as our main vehicle to achieve the model. The reasons are, first, that they have a rich theory, much of which can be presented at an elementary level. Secondly, there are large number of systems arising in practice that have already been modeled by Markov chains, so that this model has many useful applications and gives plenty of information.

The next section exhibits how the microscopic model can be used to derive macroscopic properties, and how the stochastic model can lead to more traditional deterministic models of macroscopic behavior.

### 3.5 From Stochastic to Deterministic Modeling

In this section we will present some generalities about stochastic processes. Our main goal is to provide a link between stochastic and deterministic modeling which we will illustrate in detail for the one-dimensional diffusion of particles in a stationary liquid. We will show how a stochastic model for the transport of individual particles leads to the same partial differential equation for the particle
density that can be obtained by macroscopic considerations. We stress the fact that beyond the mean behavior modeled by the deterministic approach, a stochastic modeling approach also captures the stochastic fluctuations about that mean. In a stochastic approach, we model the motion of a single particle in the reactor by a stochastic process \((X_t)_{t \geq 0}\), where \(X_t\) gives the location of the particle at time \(t\). We will always assume that the process is Markovian, i.e., that for times \(0 \leq t_1 \leq \ldots \leq t_n \leq t\) we have
\[
P(X_{t_1} \in A \mid X_{t_0} = x_1, \ldots, X_{t_n} = x_n) = P(X_t \in A \mid X_t = x_n)
\]
where \(P(X_t \in A \mid X_t = x_n)\) denotes the conditional probability that \(X_t\) is in \(A\), given that its value at time \(t_n\) was \(x_n\). This assumption states that the probability distribution of the particle’s location in the future only depends on the present position \(X_t\) and not on the past. The process is fully specified by the transition probabilities \(P(X_{s+t} \in A \mid X_s = \xi)\) for \(s, t \geq 0\) together with the initial distribution \(P(X_0 \in \bullet)\). If the process is also time-homogeneous, the transition probabilities are independent of \(s\) and thus we can define \(Q_t(\xi, A) = P(X_{s+t} \in A \mid X_s = \xi)\).

Let now \(\pi\) be the probability density of the particle’s initial position. Then the particle distribution at time \(t\) is given by \(\mu_t(A) = \int Q_t(\xi, A) \pi(\xi) d\xi\).

If the transition probability distribution has a density \(q_t(\xi, x)\), i.e.,
\[
P(X_{s+t} \in A \mid X_s = \xi) = \int q_t(\xi, x) dx ,
\]
the particle distribution has density
\[
p(t, x) = \int q_t(\xi, x) \pi(\xi) d\xi.
\]
Under very general assumptions the probability density satisfies a forward differential equation \(\frac{\partial}{\partial t} p(t, x) = L_x p(t, x)\) where \(L_x\) is a differential operator acting on the space variable \(x\). Roughly speaking the forward, or Fokker-Planck, equation states that the density change \(\frac{\partial}{\partial t} p(t, x)\) is a result of gains and losses expressed by \(L_x\). The best known example is the convection-diffusion equation,
\[
\frac{\partial}{\partial t} p(t, x) = D \frac{\partial^2}{\partial x^2} p(t, x) - V \frac{\partial}{\partial x} p(t, x).
\]
In our experience, it is often easier to formulate a discrete model first and to approximate a continuous model by taking a sufficiently fine discretization.
Subdividing space into a finite number of cells \( C_i, i \in I \), and modeling only transitions at times \( n \cdot \varepsilon, n = 0, 1, 2, \ldots \), we obtain a Markov chain \( (X_{nc})_{n \geq 0} \) with state space \( I \). The Markov chain is specified by its transition probabilities \( p_{ij} = P(X_{(n+1)c} = j \mid X_{nc} = i) \). As an analogue of the forward differential equation, we now obtain a difference equation for the probability function,

\[
P(X_{(n+1)c} = j) - P(X_{nc} = j) = \sum_{i \neq j} P(X_{nc} = i) p_{ij} - (1 - p_{jj}) P(X_{nc} = j).
\]

In the limit, with discretization steps converging to 0, this difference equation becomes the Fokker-Planck equation.

A great advantage of discretization is that it is often easier to find the transition probabilities \( p_{ij} \) of the discrete Markov chain than the transition densities \( q_r(\xi, x) \).

In this thesis we actually consider only discrete models, as these give sufficiently good results for small discretization steps.

For the purpose of illustration, we will explain in some detail both deterministic and stochastic approaches for a one-dimensional diffusion of marked particles in a liquid, e.g., one might think of the diffusion of ink in a horizontal tube of water. A macroscopic model will focus on modeling \( c(t, x) \), the density of marked particles at time \( t \) in location \( x \). The essential modeling assumption is that a density gradient leads to a flux \( j \) of marked particles against the direction of the gradient. Moreover, Fick's first law asserts that:

\[
j(t, x) = -D \frac{\partial}{\partial x} c(t, x),
\]

which is shown diagrammatically in Fig. 3.1.

![Figure 3.1 Particle flow from a high density to low density region (deterministic).](image)

As a result of this particle flux, the particle density (concentration) changes: during the time interval \( [t, t+\Delta t] \), the amount \( \Delta t j(t, x) \) will enter the interval \( [x, x+\Delta x] \) from...
the left and the amount $\Delta t \cdot j(t,x+\Delta x)$ will leave the same interval to the right. As net effect, we get an increase of the number of marked particles in $[x,x+\Delta x]$ by $\Delta t \cdot [j(t,x) - j(t,x+\Delta x)]$ and thus,

$$(c(t + \Delta t, x) - c(t, x)) \cdot \Delta x = (j(t,x) - j(t,x+\Delta x)) \cdot \Delta t$$

Dividing both sides by $\Delta x \cdot \Delta t$, and letting $\Delta x, \Delta t \to 0$, we obtain the partial differential equation

$$\frac{\partial}{\partial t} c(t,x) = -\frac{\partial}{\partial x} j(t,x). \tag{3.5}$$

In contrast with (3.1), this equation does not contain a modeling assumption, but is simply a consequence of mass-balance considerations.

Together, (3.1) and (3.2) yield the diffusion equation

$$\frac{\partial}{\partial t} c(t,x) = D \frac{\partial^2}{\partial x^2} c(t,x) \tag{3.6}$$

also known as Fick’s second law.

![Figure 3.2 Particle flow from a high density to low density region (stochastic).](image)

The microscopic approach attempts to model the motion of a single particle by a Markov process. To keep the presentation simple, we will focus on a discrete model, discretizing both time and space. We split the tube into cells of width $\Delta$ with centers at $i \Delta$, $i \in I$, and discretize time into steps of size $\varepsilon$ (see Fig.3.2). Our main model assumption specifies that between time $n \varepsilon$ and $(n + 1) \varepsilon$, a particle can move by one cell forwards or backwards, with probabilities $\nu_i \cdot \varepsilon$ and $\mu_i \cdot \varepsilon$ respectively, or remain in the same cell, with probability $(1 - \nu_i \cdot \varepsilon - \mu_i \cdot \varepsilon)$. If we denote by $X_{n\varepsilon}$ the position of the particle at time $n \varepsilon$ and by
\[ p(n \varepsilon, i \Delta) = P(X_{n \varepsilon} = i \Delta) \] its probability function, we obtain the recursion equation
\[
p((n + 1) \varepsilon, i \Delta) = p(n \varepsilon, i \Delta) + p(n \varepsilon, (i - 1) \Delta) \nu_{i-1} \varepsilon + p(n \varepsilon, (i + 1) \Delta) \mu_{i+1} \varepsilon - p(n \varepsilon, i \Delta) \lambda_{i} \varepsilon
- p(n \varepsilon, i \Delta) \mu_{i} \varepsilon
= p(n \varepsilon, (i - 1) \Delta) \nu_{i-1} \varepsilon + p(n \varepsilon, (i + 1) \Delta) \mu_{i+1} \varepsilon + p(n \varepsilon, i \Delta) \left[ 1 - \nu_{i} \varepsilon - \mu_{i} \varepsilon \right]
\]

In the special case where \( \nu_{i} \varepsilon = \mu_{i} \varepsilon = \frac{1}{2} \) for all \( i \) (see Fig. 3.2), we get
\[
p((n + 1) \varepsilon, i \Delta) = \frac{1}{2} p(n \varepsilon, (i - 1) \Delta) + \frac{1}{2} p(n \varepsilon, (i + 1) \Delta).
\]

Subtracting \( p(n \varepsilon, i \Delta) \) from both sides, we obtain a difference equation:
\[
p((n + 1) \varepsilon, i \Delta) - p(n \varepsilon, i \Delta) = \frac{1}{2} p(n \varepsilon, (i - 1) \Delta) - p(n \varepsilon, i \Delta) + \frac{1}{2} p(n \varepsilon, (i + 1) \Delta) \quad (3.7)
\]

Observe the analogy between this difference equation for the probability function and the diffusion equation (3.6): on the left-hand side we have a first order difference in the time variable and on the right-hand side a second order difference in the space variable.

If we let \( \varepsilon, \Delta \to 0 \) in a suitable way, (3.7) formally converges to (3.6). Simple manipulations of (3.7) yield
\[
\frac{p((n + 1) \varepsilon, i \Delta) - p(n \varepsilon, i \Delta)}{\varepsilon} = \frac{\Delta^{2}}{2 \varepsilon} \left( p(n \varepsilon, (i - 1) \Delta) - 2 p(n \varepsilon, i \Delta) + p(n \varepsilon, (i + 1) \Delta) \right)
\]

Observe that by Taylor expansion of any twice differentiable function \( f \) we have
\[
\frac{1}{h^{2}} \left( f(t - h) - 2 f(t) + f(t + h) \right) \to f''(t) \text{ and thus if } \Delta, \varepsilon \to 0 \text{ in such a way that}
\]
\[
\Delta^{2} = 2 D \varepsilon \text{ we get, in the limit: } \frac{\partial}{\partial t} p(t, x) = D \frac{\partial^{2}}{\partial x^{2}} p(t, x), \text{ i.e., the diffusion equation.}
\]

The heuristic approach to the derivation of the diffusion equation, by letting the discretization steps in a discrete random walk model converge to zero, can be justified and made precise. Wiener (1923) established existence of Brownian motion, a continuous analogue of the random walk process. Brownian motion is the appropriate model for the continuous motion of particles without drift. Its density \( c(t, x) \) satisfies the diffusion equation (3.3). The explanation for the fact that we obtain the same PDE for the particle density, a macroscopic quantity and the probability density, a microscopic quantity, is to be found in the law of large numbers. The empirical distribution of a very large number of particles namely follows the probability distribution. The idea that diffusion phenomena
could be explained by a stochastic model for the motion of individual particles, dates back to Einstein’s 1905 paper on Brownian motion.

In analogous ways, stochastic models can be built for more complex processes, and in some cases, though not in all, the limit can be shown to reduce to differential conservation equations.

3.6 A Discrete Stochastic Model for Fluidization

In this section we will present a model for particle transport in continuous fluidized beds that is the basis for the modeling approach followed in this thesis. This model was introduced by Hoffmann and Dehling (1998) and rigorously studied by Dehling, Hoffmann and Stuut (1999). These papers were in turn inspired by the work of Hoffmann and Paarhuis (1990), who were the first to propose a stochastic model for the motion of individual particles in a fluidized bed. Hoffmann and Paarhuis ran Monte-Carlo simulations to investigate the properties of their model, whereas Dehling, Hoffmann and Stuut gave a mathematical analysis of the model. As a result of the analysis, numerical calculations became possible and meaningful parameters of the model could be identified.

Dehling, Hoffmann and Stuut applied their model to the study of RTD curves, which provide information about the distribution of residence times of particles in the reactor. If \( T \) denotes the time a randomly chosen particle spends inside the reactor then \( T \) is a random variable. Its distribution function, \( F(t) = P(T \leq t) \) is the so-called RTD curve. In macroscopic terms, \( F(t) \) gives the fraction of particles that spend at most time \( t \) inside the reactor. The RTD curves of fluidized bed reactors exhibit characteristics that could not be explained by traditional models. It was a great success of the stochastic modeling approach that the model RTD curves gave good predictions of experimental RTD curves.

In this thesis we study batch fluidized beds and thus the model has to be modified accordingly. The concept and the basic ideas are, however, still the same. The difference lies technically in the boundary conditions and in the type of questions considered. Obviously, RTD-curves make no sense for batch models. In turn, the distribution of marked particles in the reactor and issues of mixing and separation of multi-type particles become objects of study.

A number of researchers have tried to identify and model the physical phenomena governing the transport of particles in a fluidized bed, e.g., Berruti et al. (1988), Haines and King (1972), Morris et al. (1964) and Rowe and Partridge (1962). The model of Hoffmann and Paarhuis was inspired by the ideas of Rowe and Partridge.
According to Rowe and Partridge particle motion in a fluidized bed can be attributed to three physical processes:

- Upward transport in the wake of rising fluidization bubbles followed by deposition at the top of the bed.
- Downward transport of particles in the particulate (bulk) phase as compensation for the removal of material in the wake phase further down in the reactor.
- Diffusion transport due to the disturbance of particles in the particulate phase by rising bubbles.

The second and the third of these processes are well understood convection and diffusion processes. The first process is a new feature present in fluidized beds. Particle transport in the wake of rising fluidization bubbles is fast in comparison to the other transport processes and adds a discontinuous element to our model. We want to emphasize that not just the particles from the bottom of the reactor are transported upward in the bubble wakes. Due to coalescence of rising bubbles the wake flow increases from bottom to top and thus particles everywhere in the reactor can be caught in the wake and transported upward. Dehling, Hoffmann and Stuut (1999) studied this process in great detail.

Our model for the motion of a single particle in the reactor is Markovian, i.e. we assume that the future of the process depends only on the present position and not on the past. This is a crucial assumption which e.g., excludes the possibility of particles having a velocity. In this context it is again important to realize that any mathematical model is just an image of the real world that tries to portray the main features. We model only the vertical position of a particle, thus disregarding axial motion. In this way we simplify the problem from dimension 3 to dimension 1. This approach makes sense in the situation where the reactor is axially symmetric. Modeling the full three-dimensional motion is still an open problem.

Our modeling approach starts with a discrete model, in which we have partitioned the reactor into a finite number of horizontal cells and where we only record the cell in which the particle is located and not its precise position within the cell. Moreover we study the process in discrete time, at integer multiples of a time unit. In a later stage, we arrive at a continuous model by taking a so-called diffusion limit where both time and space discretizations become finer and finer until they eventually converge to zero. We choose this approach because the discrete model is intuitively appealing and directing based on the above mentioned physical ideas.
Chapter 3: Introduction to Stochastic Models and Markov Chains

Discrete Markov Model

In general, $X_t$ denotes the vertical distance of the particle from the top of the bed at time $t$, $t \geq 0$. Space is discretized by dividing the reactor vertically into $N$ cells of equal width, numbered from top to bottom, with the exterior of the reactor below the distributor plate being denoted as $N+1$. The particles that have entered state $N+1$ can not return to the interior of the reactor, an assumption which is based on the actual transport of the particles in continuous fluidized bed reactors. Time is also discretized, and therefore denoted by $n$, the number of time steps since $t = 0$.

Figure 3.3 gives a schematic representation of a fluidized bed reactor together with the discretized model that is the basis for our Markov process. The bubble size inside the reactor increases with height as does the material transport in the wake. The picture on the right hand side indicates the possible transitions of a particle in one unit of time; the particle can move one cell up or down, stay in the same cell or move all the way to the top of the bed. The first three possibilities model the continuous process of convection and diffusion, whereas the last possibility models the transport in the wake. Our model assumes that wake transport to the top is instantaneous which, though this might seem a rather crude assumption, is still a reasonable approximation.

The vertical distance of the particle from the top of the reactor at time $n$ is given by the cell index $X_n$. When the reactor is operated continuously, particles are constantly added to the top of the bed and removed at the bottom. Once they enter cell $N+1$ they can not return to the reactor forever. A particle that is presently in cell $i$ undergoes during the next time unit one of the following transitions (see Fig. 3.3):

- moves one cell back with probability $\delta_i(1 - \lambda_i)$
- moves one cell ahead with probability $\beta_i(1 - \lambda_i)$
- stays in the same cell with probability $(1 - \beta_i - \delta_i)(1 - \lambda_i)$
- returns to the first cell on top of the bed with probability $\lambda_i$

The probabilities depend on the particle’s location, indicated by the suffix $i$. For example, if a particle is in cell $N$, it has a higher probability $\lambda$ of moving up to cell 1 than elsewhere, according to the theory that bubble wakes are mainly formed at the bottom of the bed. In Dehling and Hoffmann(1998, 1999), the probabilities were computed on the basis of models for the wake flow.
The conditional probability that the particle is in cell \( j \) at time \( n + 1 \), given it was in cell \( i \) at time \( n \), is denoted as \( p_{ij} \):

\[
p_{ij} = P(X_{n+1} = j | X_n = i)
\]  

(3.8)

where \( X_0 = 0 \) and \( X_n \) = location of particle at time \( n \). The model is specified by:

- the probability vector \( p(0) = (p(0,1), \ldots, p(0,N+1)) \) of the particle’s initial position (if a particle starts in the first cell, \( p(0) = (1,0,\ldots,0) \))
- the transition matrix \( P = (p_{ij})_{i,j \leq N+1} \) which gives the transition probabilities from all cells to all cells

The transition probabilities in our model for the interior of the reactor are then given by:

\[
\begin{align*}
p_{i,i} & = \lambda_i \\
p_{i,i-1} & = \delta_i (1 - \lambda_i) \\
p_{i,i+1} & = (1 - \delta_i - \beta_i)(1 - \lambda_i) \\
p_{i,i+1} & = \beta_i (1 - \lambda_i).
\end{align*}
\]  

(3.9)

As boundary conditions we take reflection at the entrance and absorption at the exit which is expressed by \( p_{1,2} = \beta_1 (1 - \lambda_1) \), \( p_{1,i} = 1 - \beta_1 (1 - \lambda_i) \) and \( p_{N+1,N+1} = 1 \).
The transition matrix $P$ is thus given by:

$$
\begin{pmatrix}
1 - \beta_1(1 - \lambda_1) & \beta_1(1 - \lambda_1) & 0 & 0 & \cdots & 0 & 0 \\
\lambda_2 + \delta_2(1 - \lambda_2) & \alpha_2(1 - \lambda_2) & \beta_2(1 - \lambda_2) & 0 & \cdots & 0 & 0 \\
\lambda_3 & \delta_3(1 - \lambda_3) & \alpha_3(1 - \lambda_3) & \beta_3(1 - \lambda_3) & \cdots & 0 & 0 \\
\lambda_4 & 0 & \delta_4(1 - \lambda_4) & \alpha_4(1 - \lambda_4) & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\lambda_N & 0 & 0 & 0 & \cdots & \alpha_N(1 - \lambda_N) & \beta_N(1 - \lambda_N) \\
0 & 0 & 0 & 0 & \cdots & 0 & 1
\end{pmatrix}
$$

This is a so-called stochastic matrix: a square matrix with nonnegative entries giving transition probabilities, and row sums all equal to 1. The probability function of $(X_n)_{n \geq 0}$ gives the probability that the particle is in cell $i$ at time $n$:

$$p(n,i) = P(X_n = i). \quad (3.10)$$

The probability distribution of the particle’s location at time $n$, $p(n)$, can be computed from the probability distribution at time $n-1$ and the transition probabilities:

$$p(n) = p(n-1)P. \quad (3.11)$$

Using the particle’s initial position $p(0)$, $p(n)$ is a row vector obtained by multiplying the conditional probabilities:

$$p(n) = p(0)P^n. \quad (3.12)$$

If $p(0) = (1, 0, \ldots, 0)$ is the initial distribution, $p(n)$ is simply the first row of the $n^{th}$ power of the transition matrix $P$.

An application of (3.8) yields that the probability that a particle is in cell $j$ at time $n+1$ can be computed as

$$p(n+1,j) = \sum_{i=1}^{N+1} p(n,i)p_{i,j}. \quad (3.13)$$
Because of the special structure of our transition matrix, we obtain for \(2 \leq i \leq N\) the following recursion equation

\[
p(n+1,i) = (1-\lambda_{i-1})\beta_{i-1} p(n,i-1) + (1-\lambda_i)\alpha_i p(n,i) + (1-\lambda_{i+1})\delta_{i+1} p(n,i+1)
\]

(3.14)

In the same way, we get at the boundaries i.e. at \(i = 1\) and \(i = N+1\),

\[
p(n+1,1) = \sum_{i=1}^{N} \lambda_i p(n,i) + \delta_2 (1-\lambda_2) p(n,2) + (1-\lambda_i)(1-\beta_i) p(n,1)
\]

and

\[
p(n+1,N+1) = \beta_N (1-\lambda_N) p(n,N) + p(n,N+1).
\]

The above system of recursion equations is the basic concept of stochastic modeling of Dehling and Hoffmann (1999, 2000). For underlying the full details, please consult Dehling et al. (2000).

The continuous model was introduced as limits of discrete Markov chains, obtained by letting space and time discretizations converge to zero.

### 3.7 A Continuous Stochastic Model for Fluidization

In the previous section we have modeled particle transport in fluidized bed reactors by a Markov process in discrete time with a discrete state space. From here, we can obtain a continuous time, continuous state model by letting the discretizations converge to zero in much the same way as this was done in section 3.2. We divide the reactor into \(N\) horizontal segments of height \(\Delta = \frac{h}{N}\) each, where \(h\) denotes the height of the reactor (see Fig. 3.4). We number the cells from top to bottom by \(i = 1, \ldots, N\). If we let \(x\) denote the distance of a particle from the top, then \(x \in ((i-1)\Delta, i\Delta]\) for an element of the \(i\)-th cell. Time is discretized into intervals of length \(\varepsilon\), and we model the particle’s position at times \(n \cdot \varepsilon, n = 0, 1, 2, \ldots\).
In order to obtain a meaningful limit process as $\Delta \to 0$ and $\epsilon \to 0$, the transition probabilities in (3.9) have to depend on $\epsilon$, $\delta$ in a suitable way. Let $\nu: [0, h] \to \mathbb{R}$, $D: [0, h] \to [0, \infty)$ and $\lambda: [0, h] \to [0, \infty)$ be continuous functions, giving velocity, diffusion and rate of return to the top throughout the reactor. We define

$$\beta_i = \frac{\epsilon}{2\Delta^2} D(i\Delta) + \frac{\epsilon}{2\Delta} v(i\Delta)$$

and

$$\delta_i = \frac{\epsilon}{2\Delta^2} D(i\Delta) - \frac{\epsilon}{2\Delta} v(i\Delta)$$

and $\alpha_i = 1 - \delta_i - \beta_i$. A particle that is not returning to the top of the reactor, will move one cell downward or upward with probability $\beta_i$ or $\delta_i$ and stay in the same cell with probability $\alpha_i$. The expected displacement in one time step is thus

$$\beta_i \cdot \Delta + \delta_i \cdot (-\Delta) = \epsilon \cdot v(i\Delta),$$

corresponding to a mean velocity $v(i\Delta)$. The mean square displacement in one time step is

$$\beta_i \cdot \Delta^2 + \delta_i \cdot \Delta^2 = \epsilon \cdot D(i\Delta)$$

corresponding to a mean square displacement per time unit $D(i\Delta)$. These considerations motivate our choices of $v$ and $D$, respectively.
We define moreover \( D_0 = \sup_{0 \leq x \leq 1} D(x) \) and require that \( \varepsilon \) and \( \Delta \) be related by the identity

\[
\varepsilon = \frac{\Delta^2}{2D_0}.
\]

With this requirement, all three probabilities, \( \alpha_i \), \( \beta_i \) and \( \delta_i \) are nonnegative for small \( \Delta \), provided that \( v \) and \( D \) are strictly positive functions. In the special case of a constant diffusion, i.e. \( D(x) \equiv D \), we thus get \( \varepsilon = \frac{\Delta^2}{2D} \) and

\[
\begin{align*}
\beta_i &= \frac{1}{4} + \frac{\Delta^2}{4D} v(i\Delta) \\
\delta_i &= \frac{1}{4} - \frac{\Delta^2}{4D} v(i\Delta) \\
\alpha_i &= \frac{1}{2}.
\end{align*}
\]

We model the rate of return to the top of the bed by the rate function \( \lambda(x) \) and thus the probability for a return in a small time interval of length \( \varepsilon \) becomes \( \lambda_i = \varepsilon \cdot \lambda(i\Delta) \).

We model the location of the particle after \( n \) transitions and define the continuous time process

\[
X^n_{\Delta t} = \Delta \cdot \tilde{X}^\Delta_{n[t\Delta]}.
\]

Observe that \( X^n_{\Delta t} \) gives the distance of the lower boundary of the cell that contains our particle, and thus its location up to a discretization error of most \( \Delta \), at time \( t \).

Dabrowski and Dehling (1998) have shown that \( \left( X^n_{\Delta t} \right)_{n \geq 0} \) converges to a limit process \( \left( X_t \right)_{t \geq 0} \) that can be described via a continuous diffusion part and a Poisson process of jumps to the top.

If we enter the above values of \( v, D \) and \( \lambda \) into (3.14), we get the recursion formula
\[ p_{\Delta}((n+1)\varepsilon,i\Delta) = \left\{ \frac{\varepsilon}{2\Delta^2} D((i-1)\Delta) + \frac{\varepsilon}{2\Delta} v((i-1)\Delta) \right\} \left(1 - \varepsilon\lambda((i-1)\Delta) \right) p_{\Delta}(n\varepsilon,(i-1)\Delta) \\
+ \left\{ 1 - \frac{\varepsilon}{\Delta^2} D(i\Delta) \right\} \left(1 - \varepsilon\lambda(i\Delta) \right) p_{\Delta}(n\varepsilon,i\Delta) \\
+ \left\{ \frac{\varepsilon}{2\Delta^2} D((i+1)\Delta) - \frac{\varepsilon}{2\Delta} v((i+1)\Delta) \right\} \left(1 - \varepsilon\lambda((i+1)\Delta) \right) p_{\Delta}(n\varepsilon,(i+1)\Delta) \]

(3.15)

for the distribution of particles at time \( t = n\varepsilon \). If we let \( \Delta,\varepsilon \to 0 \), this difference equation becomes a partial differential equation

\[ \frac{\partial}{\partial t} p(t,x) = \frac{1}{2} \frac{\partial^2}{\partial x^2} \left(D(x)p(t,x)\right) - \frac{\partial}{\partial x} \left(v(x)p(t,x)\right) - \lambda(x)p(t,x) \]

(3.16)

for the density of \( X_t \) at location \( x \). The first two terms on the right hand side are well known from the standard diffusion equation. The third term, \( -\lambda(x)p(t,x) \), is new and describes the removal of particles in the wake of rising fluidization bubbles.

### 3.8 Outlook

In this thesis we investigate a variety of fluidized beds, all different from the one studied by Hoffmann and Dehling et al. All our beds are batch operated, hence there are no particles added to and removed from the reactor. When relating our model to a physical bed, the model neglects the volume occupied by the bubble/wake regions, and the whole model description is based on unit cross-sectional area. We investigate three types of reactors:

- freely bubbling fluidized beds. This type of fluidized bed is commonly found in fluidized bed reactors. More details can be found in Chapter 2 and 4.
- bubbling fluidized beds with internals. This is an innovation putting baffles inside freely bubbling fluidized beds to enhance segregation of particle mixtures. Chapter 5 provides information exclusively about fluidized bed reactors with baffles.
- slugging fluidized beds. This kind of bed was found when the size of bubbles in fluidized beds gets bigger and almost equals the diameter of reactors. This is called a slugging fluidized bed. We modeled slugging fluidized beds by using our stochastic model and the result is shown in Chapter 6.
3.9 References


Chapter 3: Introduction to Stochastic Models and Markov Chains


