8

Discrete time processes embedded in continuous time systems

In this chapter, our goal is to introduce a way in which discrete time processes may be embedded in continuous time systems without altering the phase space. To do this, we adopt a strictly probabilistic point of view, not embedding the deterministic system $S: X \to X$ in a continuous time process, but rather embedding its Frobenius–Perron operator $P: L^1(X) \to L^1(X)$ that acts on L^1 functions. The result of this embedding is an abstract form of the Boltzmann equation. This chapter requires some elementary definitions from probability theory and a knowledge of Poisson processes, which are introduced following the preliminary remarks of the next section.

8.1 The relation between discrete and continuous time processes

For a semidynamical system $\{S_t\}_{t\geq 0}$ on a phase space X, if we fix the time t at some value t_0, then by property (b) of Definition 7.2.3,

$$S_{nt_0}(x) = S_{t_0} \circ \overset{n}{\cdots} \circ S_{t_0}(x) = S^n_{t_0}(x) \qquad \text{for all } x \in X.$$

Thus, in this fashion, a discrete time system may be generated from any continuous time (semidynamical) system. It is possible that a study of the discrete time system may yield some partial information concerning the continuous time system from which it was derived.

Another way of obtaining a discrete time system from a continuous one is as follows. Again, suppose we are given a semidynamical system $\{S_t\}_{t\geq 0}$. Also assume that we can find a closed set $A \subset X$ such that, if $x \in A$, then, for $t > 0$ sufficiently small $S_t(x) \notin A$, that is, each trajectory leaves A immediately (see Figure 8.1.1). Further, if every trajectory that starts in A eventually returns to A, that is, for every $x \in A$ there is a $t' > 0$ such that $S_{t'}(x) \in A$, then we may define a new mapping, the **first return map.** This is given by

$$\tilde{S}(x) = S_{t'}(x),$$

where t' is the smallest time $t' > 0$ such that $S_{t'}(x) \in A$. Again, by studying $\tilde{S}$,

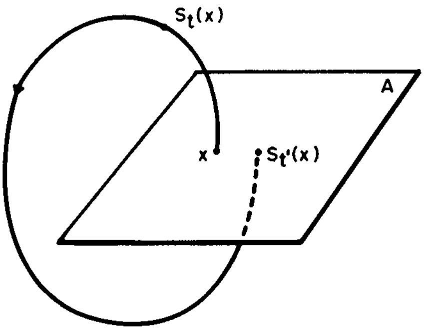

Figure 8.1.1. Determination of the first return (or Poincaré) map for a semidynamical system $\{S_t\}_{t\geq 0}$.

we may gain some insight into the properties of $\{S_t\}_{t\geq 0}$. This method was introduced by Poincaré, and the first return map is often called the **Poincaré map.**

Thus it is relatively straightforward to devise ways to study continuous time processes by a reduction to a discrete time system. However, given a discrete time system $S: X \to X$, it is much more difficult to embed it in a continuous time system and, indeed, such embedding is, in general, impossible. That is, given $\tilde{S}: X \to X$, generally there does not exist a $\{S_t\}_{t\geq 0}$ such that $\tilde{S}(x) = S_{t_0}(x)$ for some $t_0 \geq 0$ [see Zdun, 1977]. For example, in previous chapters we considered the parabolic transformation $S(x) = 4x(1 - x)$, $x \in [0, 1]$. It can be proved that there does not exist a semidynamical system $\{S_t\}_{t\geq 0}$ on $[0, 1]$ such that $S_{t_0}(x) = 4x(1 - x)$. Of course, it is always possible to embed a discrete time process into a continuous time system by altering the phase space in an appropriate way.

8.2 Probability theory and Poisson processes

Up to this point we have almost never used the word probabilistic even though we have dealt, from the outset, with normalized measures that are also measures on a probability space. In this short section, we review all of the material necessary for an understanding of Poisson processes.

The fundamental notion of probability theory is that of a **probability space** $(\Omega, \mathcal{F}, \text{prob})$, where Ω is a nonempty set called the **space of** all possible **elementary events,** $\mathcal{F}$ is a σ-algebra of subsets of Ω, which are called **events**, and "prob" is a normalized measure on $\mathcal{F}$. The equality

$$\text{prob}(A) = p, \qquad A \in \mathcal{F}$$

means that the probability of event A is p. From the fact that prob is a measure, it immediately follows that

$$\mathrm{prob}\left(\bigcup_i A_i\right) = \sum_i \mathrm{prob}(A_i)\,, \tag{8.2.1}$$

where the $A_i \in \mathcal{F}$ are mutually disjoint, that is, $A_i \cap A_j = \emptyset$ for all $i \neq j$. To introduce the concept of independence, we define it as follows.

Definition 8.2.1. In a sequence of events $A_1, A_2, \ldots$ (finite or not), the events are called **independent** if, for any increasing sequence of integers $k_1 < k_2 < \cdots < k_n$,

$$\mathrm{prob}(A_{k_1} \cap A_{k_2} \cap \cdots \cap A_{k_n}) = \mathrm{prob}(A_{k_1}) \cdots \mathrm{prob}(A_{k_n})\,. \tag{8.2.2}$$

Equation (8.2.2) just means that the probability of all the events A_{k_i} occurring is the product of the probabilities that each will occur separately.

Random variables are defined next.

Definition 8.2.2. A **random variable** ξ is a measurable transformation from Ω into R. More precisely, ξ: $\Omega \to R$ is a random variable if, for any Borel set $B \subset R$,

$$\xi^{-1}(B) = \{\omega \in \Omega: \xi(\omega) \in B\} \in \mathcal{F}\,.$$

This set is customarily written in the more compact notation $\{\xi \in B\}$. Thus, for any Borel set $B \subset R$, $\mathrm{prob}\{\xi \in B\}$ is well defined.

A function $f \in D(R)$ is called the **density** of the random variable ξ if

$$\mathrm{prob}\{\xi \in B\} = \int_B f(x)\,dx \tag{8.2.3}$$

for any Borel set $B \subset R$.

Let $\xi_1, \xi_2, \ldots$ be a sequence of random variables. We say the ξ_i are **independent** if, for any sequence of Borel sets $B_1, B_2, \ldots$, the events

$$\{\xi_1 \in B_1\}, \{\xi_2 \in B_2\}, \ldots$$

are independent. Thus a finite sequence of independent random variables satisfies

$$\mathrm{prob}\{\xi_1 \in B_1, \ldots, \xi_n \in B_n\} = \mathrm{prob}\{\xi_1 \in B_1\} \cdots \mathrm{prob}\{\xi_n \in B_n\}\,, \tag{8.2.4}$$

and the probability that all events $\{\xi_i \in B_i\}$ will occur is simply given by the product of the probabilities that each will occur separately.

We are now in a position to make the concept of a stochastic process precise with the following definition.

Definition 8.2.3. A **stochastic process** $\{\xi_t\}$ is a family of random variables that depends on a parameter t, usually called time. If t assumes only integer values,

$t = 1, 2, \ldots$, then the stochastic process reduces to a sequence $\{\xi_n\}$ of random variables called a **discrete time stochastic process.** However, if t belongs to an interval (bounded or not) of R, then the stochastic process is called a **continuous time stochastic process.**

By its very definition, a stochastic process $\{\xi_t\}$ is a function of two variables, namely, time t and event ω, but this is seldom made explicit by writing $\{\xi_t(\omega)\}$. If the time is fixed, then ξ_t is simply a random variable. However, if ω is fixed, then the mapping $t \to \xi_t(\omega)$ is called the **sample path** of the stochastic process.

Two important properties that stochastic processes may have are given in the following definition.

Definition 8.2.4. A continuous time stochastic process $\{\xi_t\}_{t\geq 0}$ has **independent increments** if, for any sequence of times $t_0 < t_1 < \cdots < t_n$, the random variables

$$\xi_{t_1} - \xi_{t_0}, \xi_{t_2} - \xi_{t_1}, \ldots, \xi_{t_n} - \xi_{t_{n-1}}$$

are independent. Further, if for any t_1 and t_2 and Borel set $B \subset R$,

$$\operatorname{prob}\{\xi_{t_2+t'} - \xi_{t_1+t'} \in B\} \tag{8.2.5}$$

does not depend on t', then the continuous time stochastic process $\{\xi_t\}$ has **stationary independent increments.**

Before giving the definition of a Poisson process, we note that a stochastic process $\{\xi_t\}$ is called a **counting process** if its sample paths are nondecreasing functions of time with integer values. Counting processes will be denoted by $\{N_t\}_{t\geq 0}$.

Definition 8.2.5. A **Poisson process** is a counting process $\{N_t\}_{t\geq 0}$ with stationary independent increments satisfying:

(a) $N_0 = 0;$ (8.2.6a)

(b) $\lim\limits_{t\to 0}(1/t)\operatorname{prob}\{N_t \geq 2\} = 0;$ (8.2.6b)

(c) The limit

$$\lambda = \lim_{t\to 0}(1/t)\operatorname{prob}\{N_t = 1\} \tag{8.2.6c}$$

exists and is positive; and

(d) $\operatorname{prob}\{N_t = k\}$ as functions of t are continuous.

A classic example of a Poisson process is illustrated by a radioactive substance placed in a chamber equipped with a device for detecting and counting the total number of atomic disintegrations N_t that have occurred up to a time t. The amount of the substance must be sufficiently large such that during the time of observation there is no significant decrease in the mass. This ensures that the probability (8.2.5) is independent of t'. It is an experimental observation that the number of disintegrations that occur during any given interval of time is independent of the number occurring during any other disjoint interval, thus giving stationary independent increments. Conditions (a)–(c) in Definition 8.2.5 have the following interpretations within this example: $N_0 = 0$ simply means that we start to count disintegrations from time $t = 0$. Condition (b) states that two or more disintegrations are unlikely in a short time, whereas (c) simply means that during a short time t the probability of one disintegration is proportional to t.

Also, the classical derivations of the Boltzmann equation implicitly assume that molecular collisions are a Poisson process. This fact will turn out to be important later.

It is interesting that from the properties of the Poisson process we may derive a complete description of the way the process depends on time. Thus we may derive an explicit formula for

$$p_k(t) = \text{prob}\{N_t = k\}. \tag{8.2.7}$$

This is carried out in two steps. First we derive an ordinary differential equation for $p_k(t)$, and then we solve it. In our construction it will be useful to rewrite equations (8.2.6a) through (8.2.6c) using the notation of (8.2.7):

$$p_0(0) = 1, \tag{8.2.8a}$$

$$\lim_{t\to 0} \frac{1}{t} \sum_{i=2}^{\infty} p_i(t) = 0, \tag{8.2.8b}$$

and

$$\lambda = \lim_{t\to 0}(1/t)p_1(t). \tag{8.2.8c}$$

To obtain the differential equation for $p_k(t)$, we first start with $p_0(t)$, noting that $p_0(t+h)$ may be written as

$$p_0(t+h) = \text{prob}\{N_{t+h} = 0\} = \text{prob}\{N_{t+h} - N_t + N_t - N_0 = 0\}.$$

Since N_t is not decreasing, hence $(N_{t+h} - N_t) + (N_t - N_0) = 0$ if and only if $(N_{t+h} - N_t) = 0$ and $(N_t - N_0) = 0$. Thus,

$$\begin{aligned} p_0(t+h) &= \text{prob}\{(N_{t+h} - N_t) = 0 \text{ and } (N_t - N_0) = 0\} \\ &= \text{prob}\{N_{t+h} - N_t = 0\}\,\text{prob}\{N_t - N_0 = 0\} \\ &= \text{prob}\{N_h - N_0 = 0\}\,\text{prob}\{N_t - N_0 = 0\} \\ &= p_0(h)\,p_0(t), \end{aligned} \tag{8.2.9}$$

where we have used the property of stationary independent increments. From (8.2.9) we may write

$$\frac{p_0(t+h)-p_0(t)}{h}=\frac{p_0(h)-1}{h}p_0(t). \tag{8.2.10}$$

Since $\sum_{i=0}^{\infty} p_i(t)=1$, we have

$$\frac{p_0(h)-1}{h}=-\frac{p_1(h)}{h}-\frac{1}{h}\sum_{i=2}^{\infty} p_i(h),$$

and, thus, by taking the limit of both sides of (8.2.10) as $h\to 0$, we obtain

$$\frac{dp_0(t)}{dt}=-\lambda p_0(t). \tag{8.2.11}$$

The derivation of the differential equation for $p_k(t)$ proceeds in a similar fashion. Thus

$$\begin{aligned}
p_k(t+h) &= \text{prob}\{N_{t+h}=k\}\\
&= \text{prob}\{N_{t+h}-N_t+N_t-N_0=k\}\\
&= \text{prob}\{N_t-N_0=k \text{ and } N_{t+h}-N_t=0\}\\
&\quad+\text{prob}\{N_t-N_0=k-1 \text{ and } N_{t+h}-N_t=1\}\\
&\quad+\sum_{i=2}^{k}\text{prob}\{N_t-N_0=k-i \text{ and } N_{t+h}-N_t=i\}\\
&= p_k(t)p_0(h)+p_{k-1}(t)p_1(h)+\sum_{i=2}^{k}p_{k-i}(t)p_i(h).
\end{aligned}$$

As before, we have

$$\frac{p_k(t+h)-p_k(t)}{h}=\frac{p_0(h)-1}{h}p_k(t)+\frac{p_1(h)}{h}p_{k-1}(t)+\frac{1}{h}\sum_{i=2}^{k}p_{k-i}(t)p_i(h),$$

and, by taking the limit as $h\to 0$, we obtain

$$\frac{dp_k(t)}{dt}=-\lambda p_k(t)+\lambda p_{k-1}(t). \tag{8.2.12}$$

The initial conditions for $p_0(t)$ and $p_k(t)$, $k\geqslant 1$, are just $p_0(0)=1$ (by definition), and this immediately gives $p_k(0)=0$ for all $k\geqslant 1$. Thus, from (8.2.11), we have

$$p_0(t)=e^{-\lambda t}. \tag{8.2.13}$$

Substituting this into (8.2.12) when $k=1$ gives

$$\frac{dp_1(t)}{dt}=-\lambda p_1(t)+\lambda e^{-\lambda t}$$

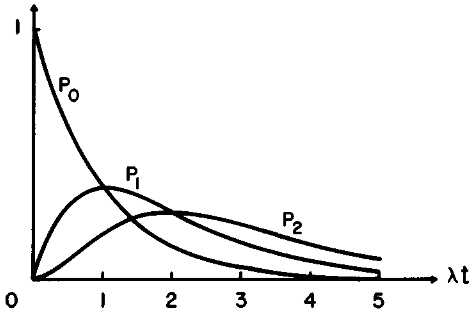

Figure 8.2.1. Probabilities $p_0(t)$, $p_1(t)$, $p_2(t)$ versus λt for a Poisson process.

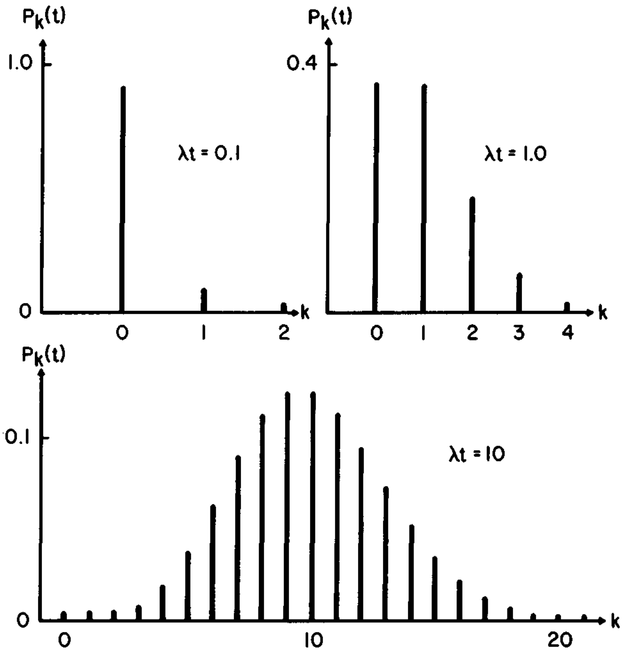

Figure 8.2.2. Plots of $p_k(t)$ versus k for a Poisson process with $\lambda t = 0.1$, 1.0, or 10.

whose solution is

$$p_1(t) = \lambda t e^{-\lambda t}.$$

Repeating this procedure for $k = 2, \ldots$ we find, by induction, that

$$p_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \tag{8.2.14}$$

The behavior of $p_k(t)$ as a function of t is shown in Figure 8.2.1 for $k = 0$, 1, and 2. Figure 8.2.2 shows $p_k(t)$ versus k for several values of λt.

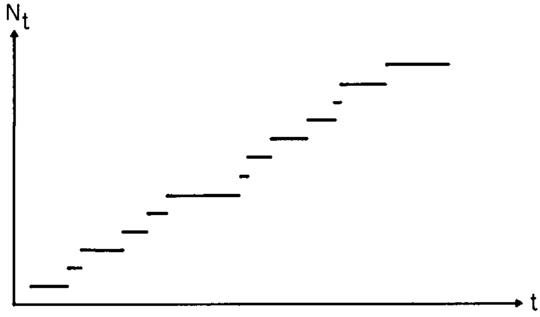

Figure 8.3.1. A sample path for a Poisson process.

Remark 8.2.1. Note that in our derivation of equation (8.2.12) we have only used $h > 0$ and, therefore, the derivative p_k' on the left-hand side of (8.2.12) is, in fact, the right-hand derivative of p_k. However, it is known [Szarski, 1967] that, if the right-hand derivative p_k' exists and the p_k are continuous [as they are here by assumption (d) of Definition 8.2.5], then there is a unique solution to (8.2.12). Thus the functions (8.2.14) give the unique solution to the problem. □

Although the way we have introduced Poisson processes and derived the expressions for $p_k(t)$ is the most common, there are other ways in which this may be accomplished. However, all these derivations, as indicated by properties (a)–(c) of Definition 8.2.5, show that a Poisson process results if the events counted by N_t are caused by a large number of independent factors, each of which has a small probability of incrementing N_t.

8.3 Discrete time systems governed by Poisson processes

A particular sample path for a Poisson process might look like the one shown in Figure 8.3.1. In this section we develop some ideas and tools that will allow us to study the behavior of a deterministic discrete time process given by a nonsingular transformation $S: X \to X$ on a measure space $(X, \mathcal{A}, \mu)$ coupled with a Poisson process $\{N_t\}_{t\geqslant 0}$. The coupling is such that, even though the dynamics are deterministic, the *times* at which the transformation S operates are determined by the Poisson process. Thus we consider the situation in which each point $x \in X$ is transformed into $S^{N_t}(x)$. This may be written symbolically as

$$x \to S^{N_t}(x)$$

for times in the interval $[0, \infty)$. Specifically we consider the following problem. Given an initial distribution of points $x \in X$, with density f, how does this distribution evolve in time? We denote the time-dependent density by $u(t, x)$ and set $u(0, x) = f(x)$.

The solution of this problem starts with a calculation of the probability that

$$S^{N_t}(x) \in A \tag{8.3.1}$$

for a given set $A \subset X$ and time $t > 0$. This probability depends on two factors: the initial density f and the counting process $\{N_t\}$.

To be more precise, we need to calculate the measure of the set

$$\{(\omega, x): S^{N_t(\omega)}(x) \in A\}. \tag{8.3.2}$$

This, in turn, requires some assumptions concerning the **product space** $\Omega \times X$ given by

$$\Omega \times X = \{(\omega, x): \omega \in \Omega, x \in X\}$$

that contains all sets of the form (8.3.2). In the space $\Omega \times X$ we define (see Theorem 2.2.2) a **product measure** that, for the sets $C \times A$, $C \in \mathcal{F}$, $A \in \mathcal{A}$, is given by $\text{prob}(C)\mu_f(A)$, and we denote it by $\text{Prob}(C \times A)$ or

$$\text{Prob}(C \times A) = \text{prob}(C)\mu_f(A), \tag{8.3.3}$$

where, as usual,

$$\mu_f(A) = \int_A f(x)\mu(dx).$$

This measure is denoted by "Prob" since it is a probability measure. Equation (8.3.3) intuitively corresponds to the assumption that the initial position x and the stochastic process $\{N_t\}_{t\geq 0}$ are independent.

Now we may proceed to calculate the measure of the set (8.3.2). This set may be rewritten as the union of disjoint subsets in the following way:

$$\begin{aligned}\{(\omega, x): S^{N_t(\omega)}(x) \in A\} &= \bigcup_{k=0}^{\infty} \{N_t(\omega) = k, S^k(x) \in A\}. \\ &= \bigcup_{k=0}^{\infty} \{N_t(\omega) = k\} \times \{S^k(x) \in A\}.\end{aligned}$$

Thus the Prob of this set is

$$\begin{aligned}\text{Prob}\{S^{N_t} \in A\} &= \sum_{k=0}^{\infty} \text{Prob}\{N_t(\omega) = k, S^k(x) \in A\} \\ &= \sum_{k=0}^{\infty} \text{prob}\{N_t = k\}\mu_f(x \in S^{-k}(A)) \\ &= \sum_{k=0}^{\infty} p_k(t) \int_{S^{-k}(A)} f(x)\mu(dx) \\ &= \sum_{k=0}^{\infty} p_k(t) \int_A P^k f(x)\mu(dx)\end{aligned} \tag{8.3.4}$$

so that

$$\text{Prob}\{S^{N_t} \in A\} = \int_A \sum_{k=0}^{\infty} p_k(t) P^k f(x) \mu(dx) \qquad \text{for } A \in \mathcal{A}, \tag{8.3.5}$$

where, as before, P denotes the Frobenius–Perron operator associated with S, and we have assumed that $S: X \to X$ is nonsingular.

The integrand on the right-hand side of (8.3.5) is just the desired density, $u(t, x)$:

$$u(t,x) = \sum_{k=0}^{\infty} p_k(t) P^k f(x). \tag{8.3.6}$$

[Note that the change in order of integration and summation in arriving at (8.3.5) is correct since $\|P^k f\| = 1$ and $\sum_{k=0}^{\infty} p_k(t) \equiv 1$. Thus the sequence on the right-hand side of (8.3.6) is strongly convergent in L^1.]

Differentiating (8.3.6) with respect to t and using (8.2.12), we have

$$\begin{aligned}\frac{\partial u(t,x)}{\partial t} &= \sum_{k=0}^{\infty} \frac{dp_k(t)}{dt} P^k f(x) \\ &= -\lambda \sum_{k=0}^{\infty} p_k(t) P^k f(x) + \lambda \sum_{k=1}^{\infty} p_{k-1}(t) P^k f(x).\end{aligned}$$

Since the last two series are strongly convergent in L^1, the initial differentiation was proper. Thus we have

$$\begin{aligned}\frac{\partial u(t,x)}{\partial t} &= -\lambda u(t,x) + \lambda \sum_{k=0}^{\infty} p_k(t) P^{k+1} f(x) \\ &= -\lambda u(t,x) + \lambda P \sum_{k=0}^{\infty} p_k(t) P^k f(x) \\ &= -\lambda u(t,x) + \lambda P u(t,x).\end{aligned}$$

Therefore $u(t, x)$ satisfies the differential equation

$$\frac{\partial u(t,x)}{\partial t} = -\lambda u(t,x) + \lambda P u(t,x) \tag{8.3.7}$$

with, from (8.3.6), the initial condition

$$u(0,x) = f(x).$$

We may always change the time scale in (8.3.7) to give

$$\frac{\partial u(t,x)}{\partial t} = -u(t,x) + P u(t,x). \tag{8.3.8}$$

From a formal point of view, equation (8.3.7) is a generalization of the system of differential equations (8.2.11) and (8.2.12) derived for the Poisson process. Consider the special case where X is the set of nonnegative integers $\{0, 1, \ldots\}$, μ is a counting measure, and $S(x) = x + 1$. For a single point $n \geqslant 1$,

$$Pf(n) = f(n - 1)$$

and when $n = 0$, $Pf(0) = 0$. Thus, from (8.3.7), we have

$$\frac{\partial u(t,n)}{\partial t} = -\lambda u(t,n) + \lambda u(t, n - 1) \qquad n \geqslant 1$$

and

$$\frac{\partial u(t,0)}{\partial t} = -\lambda u(t,0),$$

which are identical with equations (8.2.12) and (8.2.11), respectively, except that the initial condition is more general than for the Poisson process since $u(0, n) = f(n)$.

8.4 The linear Boltzmann equation: an intuitive point of view

Our derivation in the preceding section of equation (8.3.8) for the density $u(t, x)$ was quite long as we wished to be precise and show the connection with Poisson processes. In this section we present a more intuitive derivation of the same result, using arguments similar to those often employed in statistical mechanics.

Assume that we have a hypothetical system consisting of N particles enclosed in a container, where N is a large number. Each particle may change its velocity $x = (v_1, v_2, v_2)$ from x to $S(x)$ only by colliding with the walls of the container. Our problem is to determine how the velocity distribution of particles evolves with time. Thus we must determine the function $u(t, x)$ such that

$$N \int_A u(t,x)\,dx$$

is the number of particles having, at time t, velocities in the set A.

The change in the number of particles, whose velocity is in A, between t and $t + \Delta t$ is given by

$$N \int_A u(t + \Delta t, x)\,dx - N \int_A u(t,x)\,dx. \tag{8.4.1}$$

From our assumption, such a change can only take place through collisions with the walls of the container. Take Δt to be sufficiently small so that a negligible number of particles make two or more collisions with a wall during Δt. Thus, the number of particles striking the wall during a time Δt with velocity in A before the collision [and, therefore, having velocities in $S(A)$ after the collision] is

$$N\lambda\,\Delta t\int_A u(t,x)\,dx, \tag{8.4.2}$$

where λN is the number of particles striking the walls per unit time. In this idealized, abstract example we neglect the quite important physical fact that the faster particles are striking the walls of the container more frequently than are the slower particles.

Conversely, to find the number of particles whose velocity is in A after the collision, we must calculate the number having velocities in the set $S^{-1}(A)$ before the collision. Again, assuming Δt to be sufficiently small to make the number of double collisions by single particles negligible, we have

$$N\lambda\,\Delta t\int_{S^{-1}(A)} u(t,x)\,dx. \tag{8.4.3}$$

Hence the total change in the number of particles with velocity in the set A over a short time Δt is given by the difference between (8.4.3) and (8.4.2):

$$N\lambda\,\Delta t\int_{S^{-1}(A)} u(t,x)\,dx - N\lambda\,\Delta t\int_A u(t,x)\,dx. \tag{8.4.4}$$

By combining equation (8.4.1) with equation (8.4.4), we have

$$N\int_A [u(t+\Delta t,x) - u(t,x)]\,dx = \lambda N\,\Delta t\left\{\int_{S^{-1}(A)} u(t,x)\,dx - \int_A u(t,x)\,dx\right\},$$

and, since

$$\int_{S^{-1}(A)} u(t,x)\,dx = \int_A Pu(t,x)\,dx,$$

where P is the Frobenius–Perron operator associated with S, we have

$$\begin{aligned} N\int_A &[u(t+\Delta t,x) - u(t,x)]\,dx \\ &= \lambda N\,\Delta t\int_A [-u(t,x) + Pu(t,x)]\,dx. \end{aligned} \tag{8.4.5}$$

Equation (8.4.5) is exact to within an error that is small compared to Δt.

By dividing through in (8.4.5) by Δt and passing to the limit $\Delta t \to 0$, we obtain

$$\int_A \frac{\partial u(t,x)}{\partial t}\,dx = \lambda\int_A [-u(t,x) + Pu(t,x)]\,dx,$$

which gives

$$\frac{\partial u(t,x)}{\partial t} = -\lambda u(t,x) + \lambda P u(t,x).$$

Thus we have again arrived at equation (8.3.7).

In this derivation we assumed that the particle, upon striking the wall, changed its velocity from x to $S(x)$, where $S: X \to X$ is a point-to-point transformation. An alternative physical assumption, which is more general from a mathematical point of view, would be to assume that the change in velocity is not uniquely determined but is a probabilistic event. In other words, we might assume that collision with the walls of the container alters the distribution of particle velocities. Thus, if before the collision the particles have a velocity distribution with density g, then after collision they have a distribution with density Pg, where $P: L^1(X) \to L^1(X)$ is a Markov operator.

So, assume as before that $u(t,x)$ is the density of the distribution of particles having velocity x at time t, so

$$N \int_A u(t,x)\,dx$$

is the number of particles with velocities in A. Once again,

$$\lambda N\,\Delta t \int_A u(t,x)\,dx$$

is the number of particles with velocity in A that will collide with the walls in a time Δt, whereas

$$\lambda N\,\Delta t \int_A Pu(t,x)\,dx$$

is the number of particles whose velocities go into A because of collisions over a time Δt. Thus,

$$-\lambda N\,\Delta t \int_A u(t,x)\,dx + \lambda N\,\Delta t \int_A Pu(t,x)\,dx$$

is the net change, due to collisions over a time Δt in the number of particles whose velocities are in A.

Combining this result with (8.4.1), we immediately obtain the balance equation (8.4.5), which leads once again to (8.3.7). The only difference is that P is no longer a Frobenius–Perron operator corresponding to a given one-to-one deterministic transformation S, but it is an arbitrary Markov operator.

Since in our intuitive derivations of (8.3.7) presented in this section, we used arguments that are employed to derive the Boltzmann equation, we will call equation (8.3.7) a **linear abstract Boltzmann equation** corresponding to a collision (Markov) operator P. To avoid confusion with the usual Boltzmann equation, bear in mind that x corresponds to the particle velocity and not to position. Indeed, it is because we assume that the only source of change for particle velocity is collisions with the wall, that drift and external force terms do not appear in (8.3.7).

Our next goal will not be to apply equation (8.3.7) to specific physical systems. Rather, we will demonstrate the interdependence between the properties of discrete time deterministic processes, governed by $S: X \to X$ or a Markov operator, and the continuous time process, determined by (8.3.7). The next four sections are devoted to an examination of the most important properties of (8.3.7), and then in the last section we demonstrate that the Tjon–Wu representation of the Boltzmann equation is a special case of (8.3.7).

8.5 Elementary properties of the solutions of the linear Boltzmann equation

To facilitate our study of the linear Boltzmann equation (8.3.7), we will consider the solution $u(t,x)$ as a function from the positive real numbers, R^+, into L^1

$$u: R^+ \to L^1.$$

Thus, by writing (8.3.8) in the form

$$\frac{du}{dt} = (P - I)u, \tag{8.5.1}$$

where P is a given Markov operator and I is the identity operator, we may apply the Hille–Yosida theorem 7.8.1 to the study of equation (8.3.8).

All three assumptions (a)–(c) of the Hille–Yosida theorem are easily shown to be satisfied by the operator $(P - I)$ of equation (8.5.1). First, since $A = P - I$ is defined on the whole space, L^1, $\mathfrak{D}(A) = L^1$ and property (a) is thus trivially satisfied.

To check property (b), rewrite the resolvent equation $\lambda f - Af = g$ using $A = P - I$ to give

$$(\lambda + 1)f - Pf = g. \tag{8.5.2}$$

Equation (8.5.2) may be easily solved by the method of successive approximations. Starting from an arbitrary f_0, we define f_n by

$$(\lambda + 1)f_n - Pf_{n-1} = g,$$

so, as a consequence,

$$f_n = \frac{1}{(\lambda + 1)^n} P^n f_0 + \sum_{k=1}^{n} \frac{1}{(\lambda + 1)^k} P^{k-1} g. \tag{8.5.3}$$

Since $\|P^k g\| \leqslant \|g\|$, the series in (8.5.3) is, therefore, convergent, and the unique solution f of the resolvent equation (8.5.2) is

$$f \equiv R_\lambda g = \lim_{n\to\infty} f_n = \sum_{k=1}^{\infty} \frac{1}{(\lambda + 1)^k} P^{k-1} g. \tag{8.5.4}$$

Remark 8.5.1. The method of successive approximations applied to an equation such as (8.5.2) will always result in a solution (8.5.3) that converges to a unique limit, as $n \to \infty$, when $\|P\| \leqslant \lambda + 1$. The limiting solution given by (8.5.4) is called a **von Neumann series.** □

To check that the linear Boltzmann equation satisfies property (c) of the Hille–Yosida theorem, integrate (8.5.4) over the entire space X to give

$$\begin{aligned}
\int_X R_\lambda g(x)\mu(dx) &= \sum_{k=1}^{\infty} \frac{1}{(\lambda + 1)^k} \int_X P^{k-1} g(x)\mu(dx) \\
&= \sum_{k=1}^{\infty} \frac{1}{(\lambda + 1)^k} \int_X g(x)\mu(dx) \\
&= \frac{1}{\lambda} \int_X g(x)\mu(dx) = \frac{1}{\lambda},
\end{aligned}$$

where we used the integral-preserving property of Markov operators in passing from the first to the second line. Thus,

$$\int_X \lambda R_\lambda g(x)\mu(dx) = 1,$$

and, since λR_λ is linear, nonnegative, and also preserves the integral, it is a Markov operator. Thus condition (c) is automatically satisfied (see Corollary 7.8.1).

Therefore, by the Hille–Yosida theorem, the linear Boltzmann equation (8.3.8) generates a continuous semigroup of Markov operators, $\{\hat{P}_t\}_{t\geqslant 0}$.

To determine an explicit formula for $\hat{P}_t$, we first write

$$\begin{aligned}
A_\lambda f = \lambda A R_\lambda f &= \lambda(P - I)R_\lambda f \\
&= \lambda(P - I)\sum_{k=1}^{\infty} \frac{1}{(\lambda + 1)^k} P^{k-1} f \\
&= \lambda \sum_{k=1}^{\infty} \frac{1}{(\lambda + 1)^k} P^k f - \lambda \sum_{k=1}^{\infty} \frac{1}{(\lambda + 1)^k} P^{k-1} f,
\end{aligned}$$

so

$$\lim_{\lambda\to\infty} A_\lambda f = Pf - f.$$

Thus, by the Hille–Yosida theorem and equation (7.8.3), the unique semigroup corresponding to $A = P - I$ is given by

$$\hat{P}_t f = e^{t(P-I)} f, \tag{8.5.5}$$

and the unique solution to equation (8.3.8) with the initial condition $u(0,x) = f(x)$ is

$$u(t,x) = e^{t(P-I)} f(x). \tag{8.5.6}$$

Although we have determined the solution of (8.3.8) using the Hille–Yosida theorem, precisely the same result could have been obtained by applying the method of successive approximations to equation (8.5.1). However, our derivation once again illustrates the techniques involved in using the Hille–Yosida theorem and establishes that (8.3.8) generates a continuous semigroup of Markov operators. Finally, we note that if P in equation (8.3.8) is a Frobenius–Perron operator corresponding to a nonsingular transformation S, the solution can be obtained by substituting equation (8.2.14) into equation (8.3.6).

In addition to the existence and uniqueness of the solution to (8.3.8), other properties of $\hat{P}_t$ may be demonstrated.

Property 1. From inequality (7.4.7) we know that, given $f_1, f_2 \in L^1$, the norm

$$\|\hat{P}_t f_1 - \hat{P}_t f_2\| \tag{8.5.7}$$

is a nonincreasing function of time t.

Property 2. If for some $f \in L^1$ the limit

$$f_* = \lim_{t\to\infty} \hat{P}_t f \tag{8.5.8}$$

exists, then, for the same f,

$$\lim_{t\to\infty} \hat{P}_t (Pf) = f_*. \tag{8.5.9}$$

To show this, we prove even more, namely that

$$\lim_{t\to\infty} \hat{P}_t (f - Pf) = 0 \tag{8.5.10}$$

for all $f \in L^1$. Now,

$$\hat{P}_t f = e^{t(P-I)} f = e^{-t} e^{tP} f = e^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!} P^n f, \tag{8.5.11}$$

and

$$\hat{P}_t (Pf) = e^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!} P^{n+1} f = e^{-t} \sum_{n=1}^{\infty} \frac{t^{n-1}}{(n-1)!} P^n f.$$

Taking the norm of $\hat{P}_t f - \hat{P}_t(Pf)$, we have

$$\|\hat{P}_t f - \hat{P}_t(Pf)\| \leq e^{-t}\left\| \sum_{n=1}^{\infty} \left[\frac{t^n}{n!} - \frac{t^{n-1}}{(n-1)!}\right] P^n f\right\| + e^{-t}\|f\|$$
$$\leq e^{-t} \sum_{n=1}^{\infty} \left|\frac{t^n}{n!} - \frac{t^{n-1}}{(n-1)!}\right| \|f\| + e^{-t}\|f\|.$$

If $t = m$, an integer, then

$$e^{-t} \sum_{n=1}^{\infty} \left|\frac{t^n}{n!} - \frac{t^{n-1}}{(n-1)!}\right| = 2e^{-m}\left(\frac{m^m}{m!} - \frac{1}{2}\right)$$

since almost all of the terms in the series cancel. However, by Stirling's formula, $m! = m^m e^{-m}\sqrt{2\pi m}\,\theta_m$, where $\theta_m \to 1$ as $m \to \infty$. Thus for integer t,

$$\|\hat{P}_t f - \hat{P}_t(Pf)\|$$

converges to zero as $t \to \infty$. Since, by property 1 this quantity is a nonincreasing function, then (8.5.10) is demonstrated for all $t \to \infty$. Finally, inserting (8.5.8) into (8.5.10) gives the desired result, (8.5.9).

Remark 8.5.2. Note that the sum of the coefficients of $\hat{P}_t f$ given in equation (8.5.11) is identically 1, and thus the solutions of the linear Boltzmann equation $u(t,x) = \hat{P}_t f(x)$ bear a strong correspondence to the averages $A_n f$ studied earlier in Chapter 5, with n and t playing analogous roles. □

Property 3. The operators P and $\hat{P}_t$ commute, that is, $P\hat{P}_t f = \hat{P}_t P f$ for all $f \in L^1$. This is easily demonstrated by applying P to (8.5.11):

$$P(\hat{P}_t f) = e^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!} P^{n+1} f$$
$$= e^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!} P^n(Pf) = \hat{P}_t(Pf).$$

Property 4. If for some $f \in L^1$ the limit (8.5.8), $f_* = \lim_{t\to\infty} \hat{P}_t f$, exists, then f_* is a fixed point of the Markov operator P, that is,

$$Pf_* = f_*.$$

To show this, note that if

$$f_* = \lim_{t\to\infty} \hat{P}_t f,$$

then, by (8.5.9),

$$Pf_* = \lim_{t\to\infty} P(\hat{P}_t f) = \lim_{t\to\infty} \hat{P}_t(Pf) = f_*,$$

which gives the desired result. Further, the same argument shows that, if $f_* = \lim_{n\to\infty} \hat{P}_{t_n} f$ exists for some subsequence $\{t_n\}$, then $Pf_* = f_*$.

Property 5. If $Pf_* = f_*$ for some $f_* \in L^1$, then also $\hat{P}_t f_* = f_*$.

This is also easy to show. Write $Pf_* = f_*$ as

$$(P - I)f_* = 0.$$

Since $(P - I) = A$ is an infinitesimal operator, and every solution of $Af = 0$ is a fixed point of the semigroup (see Section 7.8), we have immediately that $\hat{P}_t f_* = f_*$.

8.6 Further properties of the linear Boltzmann equation

As shown in the preceding section, the solutions of the linear Boltzmann equation are rather regular in their behavior, that is, the distance between any two solutions never increases. Now we will show that, under a rather mild condition, $\hat{P}_t f$ always converges to a limit.

Recall our definition of precompactness (Section 5.1) and observe that every sequence $\{f_n\}$ that is weakly precompact contains a subsequence that is weakly convergent. Analogously, if for a given f, the trajectory $\{\hat{P}_t f\}$ is weakly precompact, then there exists a sequence $\{t_n\}$ such that $\{\hat{P}_{t_n} f\}$ is weakly convergent as $t_n \to \infty$. To see this, take an arbitrary sequence of numbers $t_n' \to \infty$ and, then, applying the definition of precompactness to $\{\hat{P}_{t_n'} f\}$, choose a weakly convergent subsequence $\{\hat{P}_{t_n} f\}$ of $\{\hat{P}_{t_n'} f\}$.

Theorem 8.6.1. If the trajectory $\{\hat{P}_t f\}$ is weakly precompact, then there exists a fixed point of P.

Proof: If $\{\hat{P}_t f\}$ is weakly precompact, then there exists a sequence $\{t_n\}$ such that

$$\lim_{n\to\infty} \hat{P}_{t_n} f = f_* \qquad \text{weakly} \tag{8.6.1}$$

exists. This implies the weak convergence of

$$\lim_{n\to\infty} \hat{P}_{t_n}(Pf) = \lim_{n\to\infty} P(\hat{P}_{t_n} f) = Pf_*. \tag{8.6.2}$$

However, from (8.5.10), we have

$$\lim_{n\to\infty} \hat{P}_{t_n}(f - Pf) = 0,$$

and, thus, from equations (8.6.1) and (8.6.2), we have

$$Pf_* = f_*,$$

which establishes the claim. Note also from property 5 of $\hat{P}_t$ (Section 8.5) that this implies $\hat{P}_t f_* = f_*$. ■

Theorem 8.6.2. For a given $f \in L^1$, if the trajectory $\{\hat{P}_t f\}$ is weakly precompact, then $\hat{P}_t f$ strongly converges to a limit.

Proof: From Theorem 8.6.1 we know that $\hat{P}_{t_n} f$ converges weakly to an f_* that is a fixed point of P and $\hat{P}_t$. Write $f \in L^1$ in the form

$$f = f - f_* + f_* .$$

Assume that for every $\varepsilon > 0$ the function $f - f_*$ may be written in the form

$$f - f_* = Pg - g + r, \tag{8.6.3}$$

where $g \in L^1$ and $\|r\| \leqslant \varepsilon$. (We will prove in the following that this representation is possible.) By using (8.6.3), we may write

$$\hat{P}_t f = \hat{P}_t(f - f_* + f_*) = \hat{P}_t(Pg - g) + \hat{P}_t f_* + \hat{P}_t r .$$

However $\hat{P}_t f_* = f_*$ and, thus,

$$\|\hat{P}_t f - f_*\| \leqslant \|\hat{P}_t(Pg - g)\| + \|\hat{P}_t r\| .$$

From (8.5.10), the first term on the right-hand side approaches zero as $t \to \infty$, whereas the second term is not greater than ε. Thus

$$\|\hat{P}_t f - f_*\| \leqslant 2\varepsilon$$

for t sufficiently large, and, since ε is arbitrary,

$$\lim_{t\to\infty} \|\hat{P}_t f - f_*\| = 0 ,$$

which completes the proof if (8.6.3) is true. Suppose (8.6.3) is not true, which implies that

$$f - f_* \notin \text{closure}(P - I)L^1(X) .$$

This, in turn, implies by the Hahn–Banach theorem (see Proposition 5.2.3) that there is a $g_0 \in L^\infty$ such that

$$\langle f - f_*, g_0 \rangle \neq 0 \tag{8.6.4}$$

and

$$\langle h, g_0 \rangle = 0$$

for all $h \in \text{closure}(P - I)L^1(X)$. In particular

$$\langle (P - I)P^n f, g_0 \rangle = 0 ,$$

since $(P - I)P^n f \in (P - I)L^1(X)$, so

$$\langle P^{n+1}f, g_0\rangle = \langle P^n f, g_0\rangle$$

for $n = 0, 1, \ldots$. Thus, by induction, we have

$$\langle P^n f, g_0\rangle = \langle f, g_0\rangle . \tag{8.6.5}$$

Furthermore, since $e^{-t}\sum_{n=0}^{\infty} t^n/n! = 1$, we may multiply both sides of (8.6.5) by $e^{-t}t^n/n!$ and sum over n to obtain

$$\left\langle e^{-t}\sum_{n=0}^{\infty}\frac{t^n}{n!}P^n f, g_0\right\rangle = \langle f, g_0\rangle ,$$

or

$$\langle \hat{P}_t f, g_0\rangle = \langle f, g_0\rangle . \tag{8.6.6}$$

Substituting $t = t_n$ and taking the limit as $t \to \infty$ in (8.6.6) gives

$$\langle f_*, g_0\rangle = \langle f, g_0\rangle ,$$

and, thus,

$$\langle f_* - f, g_0\rangle = 0 ,$$

which contradicts equation (8.6.4). Thus (8.6.3) is true. ■

8.7 Effect of the properties of the Markov operator on solutions of the linear Boltzmann equation

From the results of Section 8.6, some striking properties of the solutions of the linear Boltzmann equation emerge. The first of these is stated in the following corollary.

Corollary 8.7.1. If for $f \in L^1$ there exists a $g \in L^1$ such that

$$|\hat{P}_t f| \leqslant g \qquad t \geqslant 0 , \tag{8.7.1}$$

then the (strong) limit

$$\lim_{t\to\infty} \hat{P}_t f \tag{8.7.2}$$

exists. That is, either $\hat{P}_t f$ is not bounded by any integrable function or $\hat{P}_t f$ is strongly convergent.

Proof: Observe that $\{\hat{P}_t f\}$ is weakly precompact by our first criterion of precompactness; see Section 5.1. Thus the limit (8.7.2) exists according to Theorem 8.6.2. ■

With this result available to us, we may go on to state and demonstrate some important corollaries that give information concerning the convergence of solutions $\hat{P}_t f$ of (8.3.8) when the operator P has various properties.

Corollary 8.7.2. If the (Markov) operator P has a positive fixed point f_*, $f_*(x) > 0$ a.e., then the strong limit, $\lim_{t\to\infty} \hat{P}_t f$, exists for all $f \in L^1$.

Proof: First note that when the initial function f satisfies

$$|f| \leqslant cf_* \tag{8.7.3}$$

for some sufficiently large constant $c > 0$, we have

$$|P^n f| \leqslant P^n(cf_*) = cP^n f_* = cf_* .$$

Multiply both sides by $e^{-t}t^n/n!$ and sum the result over n to give

$$\left| e^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!} P^n f \right| \leqslant ce^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!} f_* = cf_* .$$

The left-hand side of this inequality is just $|\hat{P}_t f|$, so that

$$|\hat{P}_t f| \leqslant cf_* ,$$

and, since $\hat{P}_t f$ is bounded, by Corollary 8.7.1 we know that the strong limit $\lim_{t\to\infty} \hat{P}_t f$ exists.

In the more general case when the initial function f does not satisfy (8.7.3), we proceed as follows. Define a new function by

$$f_c(x) = \begin{cases} f(x) & \text{if } |f(x)| \leqslant cf_*(x) \\ 0 & \text{if } |f(x)| > cf_*(x) . \end{cases}$$

It follows from the Lebesgue dominated convergence theorem that

$$\lim_{c\to\infty} \|f_c - f\| = 0 .$$

Thus, by writing $f = f_c + f - f_c$, we have

$$\hat{P}_t f = \hat{P}_t f_c + \hat{P}_t(f - f_c) .$$

Since f_c satisfies

$$|f_c| \leqslant cf_* ,$$

from (8.7.3) we know that $\{\hat{P}_t f_c\}$ converges strongly. Now take $\varepsilon > 0$. Since $\{\hat{P}_t f_c\}$ is strongly convergent, there is a $t_0 > 0$, which in general depends on c, such that

$$\|\hat{P}_{t+t'} f_c - \hat{P}_t f_c\| \leqslant \varepsilon \qquad \text{for } t \geqslant t_0,\ t' \geqslant 0 . \tag{8.7.4}$$

Further,

$$\|\hat{P}_t f - \hat{P}_t f_c\| \leqslant \|f - f_c\| \leqslant \varepsilon \qquad \text{for } t \geqslant 0 \tag{8.7.5}$$

for a fixed but sufficiently large c. From equations (8.7.4) and (8.7.5) it follows that

$$\|\hat{P}_{t+t'} f - \hat{P}_t f\| \leqslant 3\varepsilon \qquad \text{for } t \geqslant t_0,\ t' \geqslant 0\,,$$

which is the Cauchy condition for $\{\hat{P}_t f\}$. Thus $\{\hat{P}_t f\}$ also converges strongly, and the proof is complete. ■

The existence of the strong limit (8.7.2) is interesting, but from the point of view of applications we would like to know what the limit is. In the following corollary we give a sufficient condition for the existence of a unique limit to (8.7.2), noting, of course, that, since (8.7.2) is linear, uniqueness is determined only up to a multiplicative constant.

Corollary 8.7.3. Assume that in the set of all densities $f \in D$ the equation $Pf = f$ has a unique solution f_* and $f_*(x) > 0$ a.e. Then, for any initial density, $f \in D$

$$\lim_{t\to\infty} \hat{P}_t f = f_*\,, \tag{8.7.6}$$

and the convergence is strong.

Proof: The proof is straightforward. From Corollary 8.7.2 the $\lim_{t\to\infty} \hat{P}_t f$ exists and is also a nonnegative normalized function. However, by property 4 of $\hat{P}_t$ (Section 8.5), we know that this limit is a fixed point of the Markov operator P. Since, by our assumption, the fixed point is unique it must be f_*, and the proof is complete. ■

In the special case that P is a Frobenius–Perron operator for a nonsingular transformation $S: X \to X$, the condition $Pf_* = f_*$ is equivalent to the fact that the measure

$$\mu_{f_*}(A) = \int_A f_*(x)\mu(dx)$$

is invariant with respect to S. Thus, in this case, from Corollary 8.7.2 the existence of an invariant measure μ_{f_*} with a density $f_*(x) > 0$ is sufficient for the existence of the strong limit (8.7.2) for the solutions of (8.3.8). Since, for ergodic transformations f_* is unique (cf. Theorem 4.2.2), these results may be summarized in the following corollary.

Corollary 8.7.4. Suppose $S: X \to X$ is a nonsingular transformation and P is the corresponding Frobenius–Perron operator. Then with respect to the trajectories $\{\hat{P}_t f\}$ that generate the solutions of the linear Boltzmann equation (8.3.8):

1. If there exists an absolutely continuous invariant measure μ_{f_*} with a positive density $f_*(x) > 0$ a.e., then for every $f \in L^1$ the strong limit, $\lim_{t\to\infty} \hat{P}_t f$ exists; and
2. If, in addition, the transformation S is ergodic, then

$$\lim_{t\to\infty} \hat{P}_t f = f_* \tag{8.7.7}$$

for all $f \in D$.

Now consider the more special case where $(X, \mathscr{A}, \mu)$ is a finite measure space and $S: X \to X$ is a measure-preserving transformation. Since S is measure preserving, f_* exists and is given by

$$f_*(x) = 1/\mu(X) \qquad \text{for } x \in X.$$

Thus $\lim_{t\to\infty} \hat{P}_t f$ always exists. Furthermore, this limit is unique, that is,

$$\lim_{t\to\infty} P_t f = f_* = 1/\mu(X) \tag{8.7.8}$$

if and only if S is ergodic (cf. Theorem 4.2.2).

In closing this section we would like to recall that, from Definition 4.4.1, a Markov operator $P: L^1 \to L^1$ is exact if and only if the sequence $\{P^n f\}$ has a strong limit that is a constant for every $f \in L^1$. Although the term exactness is never used in talking about the behavior of stochastic semigroups, for the situation where (8.7.8) holds, then, the behavior of the trajectory $\{\hat{P}_t f\}$ is precisely analogous to our original definition of exactness. Figuratively speaking, then, we could say that S is ergodic if and only if $\{\hat{P}_t\}_{t\geq 0}$ is exact.

8.8 Linear Boltzmann equation with a stochastic kernel

In this section we consider the linear Boltzmann equation

$$\frac{\partial u(t,x)}{\partial t} + u(t,x) = Pu$$

where the Markov operator P is given by

$$Pf(x) = \int_X K(x,y) f(y)\, dy \tag{8.8.1}$$

and $K(x,y)$: $X \times X \to R$ is a stochastic kernel, that is,

$$K(x,y) \geqslant 0 \tag{8.8.2}$$

and

$$\int_X K(x,y)\,dx = 1\,. \tag{8.8.3}$$

For this particular formulation of the linear Boltzmann equation, we will show some straightforward applications of the general results presented earlier.

The simplest case occurs when we are able to evaluate the stochastic kernel from below. Thus we assume that for some integer m the function $\inf_y K_m(x,y)$ is not identically zero, so that

$$\int_X \inf_y K_m(x,y)\,dx > 0 \tag{8.8.4}$$

(K_m is the m times iterated kernel K). In this case we will show that the strong limit

$$\lim_{t\to\infty} \hat{P}_t f = f_* \tag{8.8.5}$$

exists for all densities $f \in D$, where f_* is the unique density that is a solution of

$$f(x) = \int_X K(x,y)f(y)\,dy\,. \tag{8.8.6}$$

The proof of this is quite direct. Set

$$h(x) = \inf_y K_m(x,y)\,.$$

By using the explicit formula (8.5.11) for the solution $\hat{P}_t f$, we have

$$\hat{P}_t f = e^{-t}\sum_{n=0}^{\infty}\frac{t^n}{n!}P^n f\,.$$

However, for $n \geqslant m$, we may write

$$P^n f(x) = \int_X K_m(x,y)P^{n-m}f(y)\,dy \geqslant h(x)\,,$$

and thus the explicit solution $\hat{P}_t f$ becomes

$$\begin{aligned}\hat{P}_t f(x) &\geqslant e^{-t}\sum_{n=0}^{m}\frac{t^n}{n!}P^n f(x) + e^{-t}\sum_{n=m+1}^{\infty}\frac{t^n}{n!}h(x)\\ &= e^{-t}\sum_{n=0}^{m}\frac{t^n}{n!}P^n f(x) + h(x)\left[1 - e^{-t}\sum_{n=0}^{m}\frac{t^n}{n!}\right]\\ &\geqslant h(x)\left[1 - e^{-t}\sum_{n=0}^{m}\frac{t^n}{n!}\right].\end{aligned}$$

Thus we have immediately that

$$\hat{P}_t f(x) - h(x) \geq \left(-e^{-t}\sum_{n=0}^{m}\frac{t^n}{n!}\right)h(x),$$

so that

$$(\hat{P}_t f - h)^- \leq \left(e^{-t}\sum_{n=0}^{m}\frac{t^n}{n!}\right)h.$$

Since, however, $e^{-t}t^n \to 0$ as $t \to \infty$, we have

$$\lim_{t\to\infty}\|(\hat{P}_t f - h)^-\| = 0,$$

and, by Theorem 7.4.1, the strong limit f_* of (8.8.5) is unique. Properties 4 and 5 of the solution $\hat{P}_t f$, outlined in Section 8.5, tell us that f_* is the unique solution of $Pf = f$, namely, equation (8.8.6). Thus the proof is complete.

Now we assume, as before, that $K(x, y)$ is a stochastic kernel for which there is an integer m and a $g \in L^1$ such that

$$K_m(x, y) \leq g(x) \qquad \text{for } x, y \in X. \tag{8.8.7}$$

Then the strong limit

$$\lim_{t\to\infty}\hat{P}_t f \tag{8.8.8}$$

exists for all $f \in L^1$.

As before, to prove this we use the explicit series representation of $\hat{P}_t f$, noting first that, because of (8.8.7), we have, for $n \geq m$,

$$|P^n f(x)| = |P^m(P^{n-m}f(x))| \leq \int_X K_m(x, y)\,|P^{n-m}f(y)|\,dy$$

$$\leq g(x)\int_X |P^{n-m}f(y)|\,dy \leq g(x)\,\|f\|.$$

Thus we can evaluate $\hat{P}_t f$ as

$$|\hat{P}_t f| \leq e^{-t}\sum_{n=0}^{m}\frac{t^n}{n!}|P^n f| + \left(e^{-t}\sum_{n=m+1}^{\infty}\frac{t^n}{n!}g\right)\|f\|$$

$$\leq e^{-t}\sum_{n=0}^{m}\frac{t^n}{n!}|P^n f| + g\|f\|.$$

Further, setting

$$r = c\sum_{n=0}^{m}|P^n f|, \qquad c = \sup_{\substack{0\leq t\\ 0\leq n\leq m}} e^{-t}\frac{t^n}{n!},$$

we finally obtain

$$|\hat{P}_t f| \leqslant g\|f\| + r.$$

Evidently, $(g\|f\| + r)$ is an integrable function, and from Corollary 8.7.1 we know that the strong limit (8.8.8) exists.

Under assumption (8.8.7) we have no assurance that the strong limit (8.8.8) is unique. However, some additional properties of $K(x, y)$ may ensure this uniqueness. For example, if X is a bounded interval of the real line or the half-line, (8.8.7) holds, and $K_m(x, y)$ is monotonically increasing or decreasing in x, then

$$\lim_{t\to\infty} \hat{P}_t f = f_* \qquad \text{for all } f \in D, \tag{8.8.9}$$

where f_* is the unique solution of (8.8.6).

To demonstrate this, note that by repeating the proof of Proposition 5.8.1 we may construct an $h(x)$, $h(x) \geqslant 0$, $\|h\| > 0$, such that $K_m(x, y) \geqslant h(x)$. Then the proof follows directly from the assertion following equation (8.8.4).

Analogously, if (8.8.7) holds and $K_m(x, y) > 0$ for $x \in A$, $y \in X$, where A is a set of positive measure, then the limit (8.8.9) exists and is unique.

To prove this set $\hat{P}_m = \tilde{P}$ and observe that for $f \in D$ the operator $\tilde{P}$ satisfies

$$\tilde{P}f \leqslant g \quad \text{and} \quad \tilde{P}f(x) > 0 \qquad \text{for } x \in A.$$

Thus by Theorem 5.6.1 the limiting function $\lim_{n\to\infty} \tilde{P}^n f$ does not depend on f for $f \in D$. Since $\tilde{P}^n = \hat{P}_{mn}$, the limit (8.8.9) is also independent of f.

It should be noted that the same result holds under even weaker conditions, that is, if (8.8.7) holds and for some integer k

$$\sum_{n=1}^{k} K_n(x, y) > 0 \qquad \text{for } x \in A,\ y \in X.$$

8.9 The linear Tjon–Wu equation

To illustrate the application of the results developed in this chapter we close with an example drawn from the kinetic theory of gases [see Dłotko and Lasota, 1983].

In the theory of dilute gases [Chapman and Cowling, 1960] the Boltzmann equation

$$\frac{DF(t, x, v)}{Dt} = C(F(t, x, v))$$

is studied to obtain information about the particle distribution function F that depends on time (t), position (x), and velocity (v). DF/Dt denotes the total rate of change of F due to spatial gradients and any external forces, whereas the collision operator $C(\cdot)$ determines the way in which particle collisions affect F.

In the case of a spatially homogeneous gas with no external forces the Boltzmann equation reduces to

$$\frac{\partial F(t,v)}{\partial t} = C(F(t,v)). \tag{8.9.1}$$

Bobylev [1976], Krook and Wu [1977], and Tjon and Wu [1979] have shown that in some cases equation (8.9.1) may be transformed into

$$\frac{\partial u(t,x)}{\partial t} = -u(t,x) + \int_x^\infty \frac{dy}{y}\int_0^y u(t,y-z)u(t,z)\,dz, \qquad x > 0, \tag{8.9.2}$$

where $x = (v^2/2)$ (note that x is not a spatial coordinate) and

$$u(t,x) = \text{const}\int_x^\infty \frac{F(t,v)}{\sqrt{v-x}}\,dv.$$

Equation (8.9.2), called the **Tjon–Wu equation** [Barnsley and Cornille, 1981], is nonlinear because of the presence of $u(t,y-z)u(t,z)$ in the integrand on the right-hand side. Thus the considerations of this chapter are of no help in studying the behavior of $u(t,x)$ as $t \to \infty$.

However, note that $\exp(-x)$ is a solution of (8.9.2), a fact that we can use to study a linear problem. Here we will investigate the situation where a small number of particles with an arbitrary velocity distribution f are introduced into a gas, containing many more particles, at equilibrium, so that $u_*(x) = \exp(-x)$. We want to know what the eventual distribution of velocities of the small number of particles tends to.

Thus, on the right-hand side of (8.9.2), we set $u(t,y-z) = u_*(y-z) = \exp[-(y-z)]$, so the resulting **linear Tjon–Wu equation** is of the form

$$\frac{\partial u(t,x)}{\partial t} + u(t,x) = \int_x^\infty \frac{dy}{y}\int_0^y e^{-(y-z)}u(t,z)\,dz, \qquad x > 0. \tag{8.9.3}$$

Equation (8.9.3) is a special case of the linear Boltzmann equation of this chapter with a Markov operator defined by

$$Pf(x) = \int_x^\infty \frac{dy}{y}\int_0^y e^{-(y-z)}f(z)\,dz \tag{8.9.4}$$

for $f \in L^1((0,\infty))$. Using the definition of the exponential integral,

$$-\mathrm{Ei}(-x) \equiv \int_x^\infty (e^{-y}/y)\,dy,$$

equation (8.9.4) may be rewritten as

$$Pf(x) = \int_0^\infty K(x,y)f(y)\,dy, \tag{8.9.5}$$

where

$$K(x,y) = \begin{cases} -e^y\mathrm{Ei}(-y) & 0 < x \leqslant y \\ -e^y\mathrm{Ei}(-x) & 0 < y < x. \end{cases} \tag{8.9.6}$$

To examine the behavior of the solutions $u(t,x)$ of (8.9.3) as $t \to \infty$, we have a number of potential aids available. First, from the preceding section, if $\inf_y K_m(x,y) > 0$ for some m, then we could determine $\lim_{t\to\infty} \hat{P}_t f$. However, $\inf_y K(x,y) = 0$, and further composition of the kernel with itself leads to analytically complex results. Second, if we were able to find a $g(x) \geqslant K_m(x,y)$ for some m, then the results of the preceding section could be applied. However, the maximum of $K(x,y)$ in y occurs at $y = x$ and $-\exp(x)\mathrm{Ei}(-x)$ is not integrable. As before, compositions of $K(x,y)$ become so complicated that it is difficult to work with them.

A third alternative is the following. Note that $f(x) = \exp(-x)$ is a fixed point of (8.9.4). If we can show that $\exp(-x)$ is the unique fixed point of (8.9.4), then we may apply Corollary 8.7.3 to show that

$$\lim_{t\to\infty} u(t,x) = \lim_{t\to\infty} \hat{P}_t f(x) = e^{-x} \tag{8.9.7}$$

for all densities $f \in D((0,\infty))$.

From $Pf = f$ and (8.9.4), we have

$$f(x) = \int_x^\infty \frac{dy}{y} \int_0^y e^{-(y-z)} f(z)\,dz, \tag{8.9.8}$$

which must be solved for f. Since the right-hand side of (8.9.8) is differentiable, f must be differentiable. Its first derivative is

$$\frac{df(x)}{dx} = -\frac{1}{x}\int_0^x e^{-(x-z)} f(z)\,dz.$$

Multiply both sides by $x\exp(x)$ and differentiate again to obtain the nonlinear second-order differential equation

$$x\frac{d^2f}{dx^2} + (x+1)\frac{df}{dx} + f = 0. \tag{8.9.9}$$

We know that one solution of (8.9.9) is $f_1(x) = \exp(-x)$, and a second independent solution may be determined using the d'Alembert reduction method [Kamke, 1959]. This simply consists of substituting $f(x) = g(x)\exp(-x)$ into (8.9.9) and solving the resulting equation for $g(x)$. Once g is determined then the second independent solution of (8.9.8) is $f_2(x) = g(x)\exp(-x)$.

Making this substitution and simplifying gives

$$x\frac{d^2g}{dx^2} + (1-x)\frac{dg}{dx} = 0,$$

which is a first-order equation in dg/dx, easily solved to give

$$\frac{dg}{dx} = \frac{1}{x} e^x.$$

as a particular solution. Thus

$$g(x) = \mathrm{Ei}(x),$$

and the second solution of (8.9.9) is

$$f_2(x) = e^{-x}\mathrm{Ei}(x).$$

Therefore, the general solution of (8.9.9) is

$$f(x) = C_1 e^{-x} + C_2 e^{-x}\mathrm{Ei}(x). \tag{8.9.10}$$

Since we are searching for an $f \in D((0, \infty))$, we must determine C_1 and C_2 such that $f \geqslant 0$ and

$$\int_0^\infty f(x)\, dx = C_1 + C_2 \int_0^\infty e^{-x}\mathrm{Ei}(x)\, dx = 1.$$

However, $\exp(-x)\mathrm{Ei}(x)$ is not integrable, so we must have $C_1 = 1$, $C_2 = 0$, and thus the unique normalized solution of equation (8.9.9) is

$$f_*(x) = e^{-x}. \tag{8.9.11}$$

Hence f_* is also the unique normalized solution of (8.9.8).

Therefore, since $Pf = f$ has a unique nonnegative normalized solution $f_* \in D$ given by (8.9.11), which is also positive, by Corollary 8.7.3, all solutions of the linear Tjon–Wu equation have the limit

$$\lim_{t\to\infty} u(t,x) = e^{-x} \tag{8.9.12}$$

for all initial conditions $u(0,x) = f(x), f \in D((0,\infty))$.

This illustration of applying the tools developed in this chapter to deal with the Tjon–Wu equation is meant to show their power. Given the integro-differential equation (8.9.3), we have been able to show the global convergence of its solutions by examining only the fixed points of the right-hand side. This led to a second-order differential equation that was easily solved, in spite of its nonlinearity. Finally, once the solution was available and shown to satisfy the requirements of Corollary 8.7.3, then the asymptotic behavior of $u(t,x)$, for *all* initial conditions, was also known.