Probability and Its Applications

Emma Horton
Andreas E. Kyprianou

Stochastic Neutron Transport

And Non-Local Branching Markov Processes

Birkhäuser

Birkhäuser

Probability and Its Applications

Probability and Its Applications is designed for monographs on all aspects of probability theory and stochastic processes, as well as their connections with and applications to other areas such as mathematical statistics and statistical physics.

Emma Horton • Andreas E. Kyprianou

Stochastic Neutron Transport

And Non-Local Branching Markov Processes

Birkhäuser

Emma Horton
Department of Statistics
University of Warwick
Coventry, UK

Andreas E. Kyprianou
Department of Statistics
University of Warwick
Coventry, UK

ISSN 2297-0371 ISSN 2297-0398 (electronic)
Probability and Its Applications
ISBN 978-3-031-39545-1 ISBN 978-3-031-39546-8 (eBook)
https://doi.org/10.1007/978-3-031-39546-8

Mathematics Subject Classification: 60-xx, 60Jxx, 60J80

This book is published under the imprint Birkhäuser, www.birkhauser-science.com by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

Dedicated to fusion after fission

Preface

The neutron transport equation (NTE) is a fundamental equation which describes the flux of neutron particles in a fissile material. Although nothing more than a special case of the Boltzmann equation for kinetic particle systems, the mathematical significance of the NTE found prominence during WWII as part of the Manhattan project. Despite the sad associations with its beginnings, its real significance lay with the civil end of the Manhattan project that grew into the global development of nuclear energy. The development of nuclear power concurrently demanded a general understanding of how to handle neutron driven radioactivity. In the modern age, neutron transport modelling and, more generally, radiation transport modelling play out among an evenly balanced community of academics and industrialists across several sectors, including both fission and fusion reactor design, safety management and decommissioning, nuclear medical technology for both treatment and diagnosis, the food hygiene irradiation industry, and finally, the rapidly growing number of private and state organisations dedicated to interplanetary space exploration and associated extraterrestrial operations.

Having already captured and held the attention of nuclear physicists and engineers for decades, neutron transport modelling has somehow been absent from the attention of the mathematical community for around half a century. This is even more notable when one takes account of the fact that the desire to solve the NTE is precisely what drove polymaths Ulam and von Neumann, among others in the Manhattan project, to develop the concept of Monte Carlo simulation. Moreover, there had been numerous contributions from mathematicians in the post-war surge of interest in radiation transport modelling leading to the late 1960s. During this time, the relationship between fission modelling to the then evolving theory of branching processes was not missed. Many of the earliest works in this field are quite clear about the connection between the two and how one can read the behaviour of a certain type of branching process out of the NTE. In this text, we refer to the aforesaid as a neutron branching process (NBP); cf [7, 70, 94, 99, 100, 102, 103, 108, 110]. With time, an understanding of the mathematical nature of this relationship largely settled into more practical modelling questions. As alluded to above, the opportunity to study a much deeper relationship has essentially not been exploited much beyond the literature available up to the early 1970s.

In contrast, the theory of branching processes has remained of central interest to probability theory, all the way through the development of post-WWII mathematics to the present day. The classical and most basic setting of the Bienyamé–Galton–Watson (BGW) process has received significant attention since the 1940s [84] with the majority of foundational results emerging through the 1960s, 1970s, and 1980s; c.f. [5, 6, 70]. In spite of some very early work dating back to the 1940s–1960s era in both Western and Soviet literature (e.g. [61, 75–77, 84, 101, 120, 126, 129]), the fertile ground of spatial analogues of the BGW process really only began to gain momentum from around the mid-1970s. Multi-type branching processes [5], Crump–Mode–Jagers processes [80], branching random walks [16, 82, 123], branching diffusions and superprocesses, [53, 55, 90], and more recently fragmentation and growth fragmentation processes [12, 13] are all examples of processes which build out of BGW processes by including the notion of spatial movement and/or allow for fractional mass, infinite activity, and non-local mass creation.

In many ways, the aforementioned connection made between the NTE and an underlying spatial branching process in the 1940s[1] was ahead of its time. It is nonetheless surprising that the inclusion of NBPs as a fundamental example in the general theory of branching processes appears to have left by the wayside by probabilists.

One of the motivations for this book has been to correct this loss of visibility within the probabilistic community. In particular, this is especially so since the application areas of general radiation transport have become much more integrated into modern living, e.g. through healthcare. Moreover, in doing so, we will take the opportunity to present a summary of the sudden burst of material that has been published in the last three years by the authors of this book together with various collaborators; a large segment of which formed the PhD thesis of the first author [73], as well some parts coming from the PhD thesis of Isaac Gonzalez [66]. We are therefore writing for a more mathematically literate audience.

Precisely where in the modern-day zoology of branching processes, NBPs land is quite interesting. In short, an NBP is nothing more than a branching Markov process in a bounded domain. That said, the underlying Markov process is that of a particle that moves in straight lines and scatters (changes direction) at with rates that are spatially dependent. This is what we refer to in this text as a neutron random walk (NRW). Moreover, the rate at which fission (branching) occurs is also spatially dependent and occurs in a way that is non-local. The state space of the underlying Markov process is the tensor product of position and velocity. Offspring are positioned non-locally relative to last position-velocity location of the parent particle through a change in velocity rather than their position. In this sense, and

[1] This connection was certainly known to those in Los Alamos as part of the Manhattan project, and had certainly reached formal scientific literature by the 1950s; cf. [61].

as one would expect, neutrons fire off from the position that fission occurs albeit in different directions with different speeds.

It is the combination of an irregular underlying Markov process (i.e. an NRW), non-local branching, and spatial inhomogeneity that makes NBPs difficult and therefore interesting to analyse. In some respects, to make headway into the theory of NBPs is to make progress on the theory of general branching Markov processes (BMPs). In that case, one may argue that it suffices to write a book on purely the BMP setting. Nonetheless, there are some graces that come from the specificity of the neutron model.

Taking the pros and cons of the specific setting of NBPs versus general setting of BMPs into account, the book is divided into two parts. The first part runs through a modern take on the relationship between the NBP and the NTE, as well as some of the core relevant results concerning the growth and spread of mass of the NBP. For the most part, we only offer proofs that are specific to the NTE and NBP in this part of the book. As such, the reader gains quicker insight into how deeply the theory of branching processes penetrates into neutronics.

In Part II, we look at the generalisation of some of the mathematical theory announced for NBPs in Part I, albeit, now, offering their proofs in the broader context. This allows us the opportunity to provide an understanding of why the NBP is as malleable as it appears to be. In essence, Part II of the book picks up a story that began with a small cluster of literature produced by Asmussen and Herring [3–5] in the 1970s and 1980s. Their work was based on the general abstract framework for branching processes introduced in the three foundational papers of Ikeda, Nagasawa, and Watanabe [75–77], which allowed for arbitrary Markovian motion of particles and non-local branching mechanisms. Asmussen and Herring showed that, with the addition of an assumed spectral decomposition of an associated mean semigroup, a number of significant tractable results could be attained.

In principle, each of the two parts of this book could form separate texts. Clearly, the application of branching processes to neutron transport stands on its own merits given the real-world applications. Moreover, a study of the Asmussen–Hering class is deserving of attention within its own right if one considers how computationally robust it is. Nonetheless, probability is a field of mathematics which is cultured by the rigour of pure mathematics whilst, at the same time, keeping application well within view. Hence, in keeping with this tradition, we have decided to keep the two parts within the same text.

Coventry, UK
2023

Emma Horton
Andreas E. Kyprianou

Acknowledgements

The journey as probabilists to nuclear enlightenment has been complemented by the support of numerous scientists and mathematicians who have received us as cross-disciplinarians with open arms.

These include Paul Smith and his colleagues Max Shepherd, Geoff Dobson, and Simon Richards from the ANSWERS nuclear software modelling group. Through Paul Smith, we formed a special relationship with the Nuclear Engineering Monte Carlo group at Cambridge University under the leadership of Eugene Shwageraus, from which we have learned an enormous amount.

We have enjoyed ongoing collaboration with Sarah Osman and Colin Baker from University College London Hospital, Ana Lourenço and Russell Thomas at the National Physical Laboratory, and Andrea Zoia and Eric Dumonteil from the Commissariat à l'énergie atomique et aux énergies alternatives (CEA).

Numerical analysts Ivan Graham, Rob Scheichl, and Tristan Pryer at the University of Bath have worked alongside us providing alternative mathematical perspectives. We would like to thank all of these individuals for their support, intellectual stimulation, and encouragement.

A great deal of the material presented in this text book comes from a series of recently published papers with Simon Harris, Alex Cox, Denis Villemonais, and Minmin Wang (cf. [30–32, 69, 74]), with whom we have greatly enjoyed collaborating. Moreover, the motivation to go back and look at neutron branching processes and, more generally, branching Markov processes was very much fuelled by two significant EPSRC grants (EP/P009220/1 and EP/W026899/1) and accompanying industrial funding from the ANSWERS group thanks to Paul Smith.

Eugene Shwageraus provided us with the images for Figs. 1.3 and 1.2 is an image provided to us by Paul Smith, and Matt Roberts wrote the software that produced Fig. 11.1 and helped us with the simulations. Our thanks to all three for giving us permission to include these (simulation) images. In alphabetical order, James Ayer, Bastien Mallein, Juan Carlos Pardo, Gianmarco del Pino, Ellen Powell, Terence Tsui, Minmin Wang, and Alex Watson all provided feedback on an earlier draft of this book and we would like to thank them for their input.

Contents

Part I
Stochastic Neutron Transport

Chapter 1
Classical Neutron Transport Theory

The neutron transport equation (NTE) describes the flux of neutrons through an inhomogeneous fissile medium. It serves as a core example of a wider family of radiation transport equations, all of which are variants of a general category of Boltzmann transport equations. Our objective in this book is to assemble some of the main mathematical ideas around neutron transport and their relationship with the modern theory of branching Markov processes. In this first chapter, we will introduce the underlying physical processes of nuclear fission, how this gives rise to the NTE, and moreover, we discuss the classical context in which the NTE can be rigorously understood as a well-defined mathematical equation.

1.1 Basic Neutronics

Shortly, we will formally introduce the NTE as describing neutron flux. That is, the average total length travelled by all free neutrons per unit time and volume. In order to understand its structure, it is first necessary to understand the basic physical phenomena that govern the movement of neutrons in fissile environments. Below we give a brief summary of the physical processes at play. We make no apology for skipping some rather sophisticated physics in favour of a dynamical description of what one might consider a macroscopic view on the scale of the human observer.

Configuration Space We start by introducing what we call the *configuration* of a particle. Each neutron can be identified by:

- Its position $r \in D$, where $D \subset \mathbb{R}^3$ is open, bounded, and smooth enough that for every $r \in \partial D$, a unique normal vector $\mathbf{n}_r$ can be identified
- Its velocity $\upsilon \in V$, where $V = \{\upsilon \in \mathbb{R}^3 : \upsilon_{\texttt{min}} \leq |\upsilon| \leq \upsilon_{\texttt{max}}\}$, where $0 < \upsilon_{\texttt{min}} \leq \upsilon_{\texttt{max}} < \infty$

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8_1

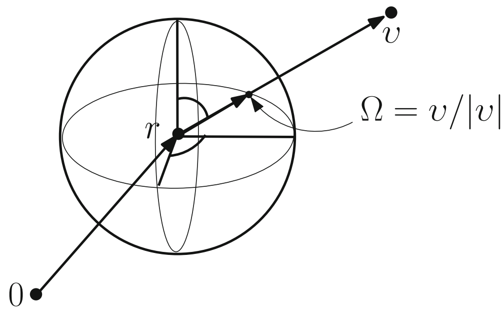

Fig. 1.1 The configuration variables (r, υ) offer the minimal mathematical information for each neutron particle

We refer to (r, υ) as the particle's configuration, and the set $D \times V$ as the configuration space. All quantities that appear in the NTE will be functionals on this space (Fig. 1.1).

In the nuclear physics and engineering literature, it is more common to work instead with physical position $r \in D$, unit direction $\Omega \in \mathbb{S}^2$ (the unit sphere), and energy $E \in (0, \infty)$. We make the global assumption that the energy of particles is such that one need not worry about quantum or relativistic effects. As such, particle energy is related to its speed, v, via the simple Newtonian kinematic equation $E = mv^2/2$, where m is the mass of a neutron. The velocity of a particle is then simply given by $\upsilon = v\Omega$. Our notation therefore deviates from these norms because, from a mathematical and probabilistic perspective, it makes more sense to work with the minimal number of necessary variables.

Advection Transport If at time $t = 0$, a neutron is found with configuration $(r, \upsilon) \in D \times V$, it will move along the trajectory $(r + \upsilon s, \upsilon)$, $s \geq 0$, until it either hits the boundary of the domain D, at which point we will consider it as absorbed and no longer tracked (equivalently, no longer in existence), or one of three other possibilities occurs. These are: scattering, fission, or neutron capture, each of which we describe next.

Scattering A scattering event occurs when a neutron and an atomic nucleus interact. As alluded to above, we shy away from the delicate physics of such subatomic interactions. It suffices to note that, for our purposes, the interaction causes what is tantamount to an instantaneous change in velocity. A more subtle form of modelling may distinguish between elastic scattering (conservation of energy) and inelastic scattering (dissipation of energy); however, for our introductory approach, we will consider only elastic scattering.

Fission Roughly speaking, nuclear fission occurs when a neutron collides with a nucleus, which then causes the nucleus to split into several smaller parts, thus releasing more neutrons. We again skirt over the subatomic subtleties and think of a fission event as an instantaneous removal of an incident neutron, to be replaced by a number of new neutrons at the same physical location but with differing velocities.

In reality, the incident neutron is absorbed by the nucleus, which results in the nucleus becoming “excited” or unstable. In turn, the nucleus splits open and releases surplus neutrons. However, not all neutrons are released from the nucleus at the same time, which leads us to two categories of fission neutrons. The so-called *prompt neutrons* are released immediately (on a time scale of about 10^{-14} seconds) from a fission event. Most neutron yield (about 99%) from fission events is made up of prompt neutrons. When an excited nucleus splits open and releases prompt neutrons, the remaining fragments from the nucleus may still be in an excited state. After some time, these fragments may release further neutrons, called *delayed neutrons*, in order to become stable again. Delayed neutrons constitute a much lower proportion of the total yield from a fission event. In our models, we will only consider prompt neutrons. Nonetheless, we will return to a brief discussion on distinguishing the two cases at the end of this chapter.

Neutron Capture Neutron capture occurs when a neutron is absorbed by a nucleus during a neutron–nucleus interaction, and as the name suggests, there is no further neutron release as a consequence. This is all we really care about for the purposes of our modelling, but in reality, the situation is inevitably more complex with the release of gamma rays, for example, from the nucleus as a consequence. In principle, neutron capture is highly dependent on the incident energy, as well as the nucleus involved. In our presentation of the NTE, we will combine nuclear fission and neutron capture by considering the latter as a fission event where no neutrons are released.

1.2 Neutron Transport Equation (NTE)

Neutron flux and neutron density are two fundamental physical measurements that describe a neutron configuration. Let us start by defining the latter. Henceforth denoted by $\Psi_t(r, \upsilon)$, neutron density is the expected number of neutrons present at time $t \geq 0$ at position $r \in D$ and with velocity $\upsilon \in V$, per unit volume, per unit velocity. In other words, the expected number of neutrons to be found in a domain $D_0 \subseteq D$ with velocities in $V_0 \subseteq V$ is nothing more than $\int_{D_0}\int_{V_0} \Psi_t(r, \upsilon)\mathrm{d}r\mathrm{d}\upsilon$.

In contrast, neutron flux, henceforth denoted by $\Gamma_t(r, \upsilon)$, is the density in $D \times V$ of the average track length covered by neutrons per unit time. Flux is fundamentally related to neutron density via the equation

$$\Gamma_t(r, \upsilon) = |\upsilon|\Psi_t(r, \upsilon), \qquad r \in D,\ \upsilon \in V,\ t \geq 0. \tag{1.1}$$

Flux is a preferable quantity for determining the rate of interaction of neutrons with other atoms because interaction rates (or cross sections) are usually expressed as “per unit track length”. This is another point of departure from standard modelling in the nuclear physics and engineering literature that we enforce here. As this text is fundamentally about probability theory, our preference is to work with rates as a "per unit time" rather than "per unit track length" quantity.

The following key quantities give the rates in question as well as some additional kernels that further describe interaction events:

$\sigma_{\mathtt{s}}(r, \upsilon')$:	The density in $D \times V$ of the rate per unit time at which scattering occurs from incoming velocity υ'
$\sigma_{\mathtt{f}}(r, \upsilon')$:	The density in $D \times V$ of the rate per unit time at which fission occurs from incoming velocity υ'
$\sigma(r, \upsilon')$:	The sum of the rates $\sigma_{\mathtt{f}} + \sigma_{\mathtt{s}}$, also known as the *collision rate*
$\pi_{\mathtt{s}}(r, \upsilon', \upsilon)$:	The probability density that an outgoing velocity due to a scatter event is υ from an incoming velocity υ', given that scattering occurs, necessarily satisfying $\int_V \pi_{\mathtt{s}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon' = 1$
$\pi_{\mathtt{f}}(r, \upsilon', \upsilon)$:	The average number of neutrons emerging at velocity υ from fission with incoming velocity υ', given that a fission or capture event occurs, satisfying $\int_V \pi_{\mathtt{f}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon' < \infty$
$\mathcal{Q}_t(r, \upsilon)$:	A non-negative source term giving the density in $D \times V$ of the rate at which neutrons are generated from a radioactive source.

For the quantities σ, $\sigma_{\mathtt{s}}$, and $\sigma_{\mathtt{f}}$, one can easily convert them to rates per unit track length by multiplying by $|\upsilon|$. In the nuclear physics and engineering literature, the quantities

$$\varsigma(r, \upsilon) := |\upsilon|^{-1}\sigma(r, \upsilon), \;\; \varsigma_{\mathtt{s}}(r, \upsilon) := |\upsilon|^{-1}\sigma_{\mathtt{s}}(r, \upsilon) \;\text{ and }\; \varsigma_{\mathtt{f}}(r, \upsilon) := |\upsilon|^{-1}\sigma_{\mathtt{f}}(r, \upsilon) \tag{1.2}$$

on $D \times V$ are known as *cross sections*. Because of the simple relationship between rates, we will somewhat abuse this notation and also refer to σ, $\sigma_{\mathtt{s}}$, and $\sigma_{\mathtt{f}}$ as cross sections. It is usual and indeed consistent with the physics of neutron transport, to impose the following important global assumption:

(H1) $\boldsymbol{\sigma_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{s}}}$ **and** $\boldsymbol{\pi_{\mathtt{f}}}$ **are uniformly bounded away from infinity.**

Now let us turn our attention to deriving the neutron transport equation using a semi-rigorous approach based on the basic principles of mean particle behaviour. Take $D_0 \subset D$, and consider the change in neutron density in $D_0 \times \{\upsilon\}$ in between time t and Δt. First note that this quantity can be expressed as

$$\int_{D_0} [\Psi_{t+\Delta t}(r, \upsilon) - \Psi_t(r, \upsilon)]\mathrm{d}r. \tag{1.3}$$

This quantity can be also derived by considering the different ways the neutron density can change according to the dynamics of the particles. We emphasise that in what follows we only consider prompt neutrons and that neutron capture is seen as fission with no neutron output.

We first consider the "gains" that lead to a neutron having a configuration (r, υ) at time $t + \Delta t$. There are three possibilities: scattering, fission, and an external source. If it has undergone a scattering, the contribution to (1.3) is given by

$$\Delta t \int_{D_0} \int_V \sigma_{\mathtt{s}}(r, \upsilon')\pi_{\mathtt{s}}(r, \upsilon', \upsilon)\Psi_t(r, \upsilon')\mathrm{d}\upsilon'\mathrm{d}r.$$

Similarly, the contribution due to fission is given by

$$\Delta t \int_{D_0} \int_V \sigma_{\mathtt{f}}(r, \upsilon')\pi_{\mathtt{f}}(r, \upsilon', \upsilon)\Psi_t(r, \upsilon')\mathrm{d}\upsilon'\mathrm{d}r.$$

Finally, the contribution due to the external source is given by

$$\Delta t \int_{D_0} \mathscr{Q}_t(r, \upsilon)\mathrm{d}r.$$

We now consider the "losses", which comes from the particles "leaving" the configuration (r, υ) due to a collision. Since this is characterised by the rate σ, the total loss in neutron density due to collisions is given by

$$\Delta t \int_{D_0} \sigma(r, \upsilon)\Psi_t(r, \upsilon)\mathrm{d}r.$$

Finally, particles may be "lost" or "gained" via the boundary of D_0, denoted ∂D_0. Recall that particles are travelling at speed $|\upsilon|$, thereby covering a distance $|\upsilon|\Delta t$ in the time interval $[t, t + \Delta t]$, and are moving in direction $\Omega = |\upsilon|^{-1}\upsilon$. Then the net number of particles that pass through $D_0 \times \{\upsilon\}$ via a small patch $\mathrm{d}S$ of the surface ∂D_0 is given by $|\upsilon|\Delta t \times \Omega \cdot \mathbf{n}_r \Psi_t(r, \upsilon)\mathrm{d}S = \Delta t \upsilon \cdot \mathbf{n}_r \Psi_t(r, \upsilon)\mathrm{d}S$. Hence, the total change in the number of particles in $D_0 \times \{\upsilon\}$ via ∂D_0 is given by

$$\Delta t \oint_{\partial D_0} \upsilon \cdot \mathbf{n}_r \Psi_t(r, \upsilon)\mathrm{d}S = \Delta t \int_{D_0} \upsilon \cdot \nabla_r \Psi_t(r, \upsilon)\mathrm{d}r,$$

where the equality follows from the divergence theorem and ∇_r is the gradient differential operator with respect to r.

Combining these gains and losses, and equating them with (1.3), we obtain

$$\begin{aligned}
&\int_{D_0} [\Psi_{t+\Delta t}(r, \upsilon) - \Psi_t(r, \upsilon)]\mathrm{d}r \\
&= \Delta t \int_{D_0} \mathscr{Q}_t(r, \upsilon)\mathrm{d}r - \Delta t \int_{D_0} \upsilon \cdot \nabla_r \Psi_t(r, \upsilon)\mathrm{d}r - \Delta t \int_{D_0} \sigma(r, \upsilon)\Psi_t(r, \upsilon)\mathrm{d}r \\
&\qquad + \Delta t \int_{D_0} \int_V \sigma_{\mathtt{s}}(r, \upsilon')\pi_{\mathtt{s}}(r, \upsilon', \upsilon)\Psi_t(r, \upsilon')\mathrm{d}\upsilon'\mathrm{d}r \\
&\qquad\qquad + \Delta t \int_{D_0} \int_V \sigma_{\mathtt{f}}(r, \upsilon')\pi_{\mathtt{f}}(r, \upsilon', \upsilon)\Psi_t(r, \upsilon')\mathrm{d}\upsilon'\mathrm{d}r.
\end{aligned}$$

Since D_0 was arbitrary, it follows that

$$
\begin{aligned}
\Psi_{t+\Delta t}(r,\upsilon) - \Psi_t(r,\upsilon) = &\Delta t \mathscr{Q}_t(r,\upsilon) - \Delta t \upsilon\cdot\nabla_r\Psi_t(r,\upsilon) - \Delta t\sigma(r,\upsilon)\Psi_t(r,\upsilon)\\
&+ \Delta t\int_V \Psi_t(r,\upsilon')\mathrm{d}\upsilon'\sigma_{\mathtt{s}}(r,\upsilon')\pi_{\mathtt{s}}(r,\upsilon',\upsilon)\\
&+ \Delta t\int_V \Psi_t(r,\upsilon')\mathrm{d}\upsilon'\sigma_{\mathtt{f}}(r,\upsilon')\pi_{\mathtt{f}}(r,\upsilon',\upsilon)\mathrm{d}\upsilon'.
\end{aligned}
$$

Dividing both sides by Δt and letting $\Delta t \to 0$, we obtain the so-called *forward neutron transport equation*:

$$
\begin{aligned}
\frac{\partial}{\partial t}\Psi_t(r,\upsilon) = &\mathscr{Q}_t(r,\upsilon) - \upsilon\cdot\nabla_r\Psi_t(r,\upsilon) - \sigma(r,\upsilon)\Psi_t(r,\upsilon)\\
&+ \int_V \Psi_t(r,\upsilon')\sigma_{\mathtt{s}}(r,\upsilon')\pi_{\mathtt{s}}(r,\upsilon',\upsilon)\mathrm{d}\upsilon'\\
&+ \int_V \Psi_t(r,\upsilon')\sigma_{\mathtt{f}}(r,\upsilon')\pi_{\mathtt{f}}(r,\upsilon',\upsilon)\mathrm{d}\upsilon'.
\end{aligned}
\tag{1.4}
$$

We will also need the following initial and boundary conditions:

$$
\begin{cases}
\Psi_0(r,\upsilon) = g(r,\upsilon) & \text{for } r\in D,\ \upsilon\in V,\\
\Psi_t(r,\upsilon) = 0 & \text{for } t\ge 0 \text{ and } r\in\partial D \text{ if } \upsilon\cdot\mathbf{n}_r < 0,
\end{cases}
\tag{1.5}
$$

where $\mathbf{n}_r$ is the outward facing unit normal at $r \in \partial D$. The second of these two conditions ensures that particles just outside the domain D travelling towards it cannot enter it.

1.3 Flux Versus Neutron Density

Earlier we noted that our preferred configuration variables (r,υ) are chosen in place of the triplet (r,Ω,E) and that we have defined our cross sections as rates per unit time rather than rates per unit track length, cf. (1.2). If we consider the system (1.4) and (1.5) in terms of flux rather than in terms of neutron density, then we see that (1.4) and (1.5) become

$$
\begin{aligned}
&\frac{1}{|\upsilon|}\frac{\partial}{\partial t}\Gamma_t(r,\Omega,E) + \Omega\cdot\nabla_r\Gamma_t(r,\Omega,E) + \varsigma(r,\Omega,E)\Gamma_t(r,\Omega,E)\\
&\qquad = \int_{\mathbb{S}^2\times(0,\infty)} \Gamma_t(r,\Omega',E')\varsigma_{\mathtt{s}}(r,\Omega',E')\pi_{\mathtt{s}}(r,\Omega',E',\Omega,E)\mathrm{d}\Omega'\mathrm{d}E'
\end{aligned}
$$

$$+\int_{\mathbb{S}^2\times(0,\infty)} \Gamma_t(r,\Omega',E')\varsigma_{\mathtt{f}}(r,\Omega',E')\pi_{\mathtt{f}}(r,\Omega',E',\Omega,E)\mathrm{d}\Omega'\mathrm{d}E'$$
$$+\mathscr{Q}_t(r,\Omega,E), \tag{1.6}$$

$$\begin{cases} \Gamma_0(r,\Omega,E) = g(r,\Omega,E) & \text{for } r\in D, \upsilon\in V, \\ \\ \Gamma_t(r,\Omega,E)=0 & \text{for } t\geq 0 \text{ and } r\in\partial D \text{ if } \upsilon\cdot\mathbf{n}_r<0. \end{cases} \tag{1.7}$$

Here, we have used the previously discussed fact that functional dependency on (r,υ) is equivalent to functional dependency on (r,Ω,E) when we have the relations $\Omega = \upsilon/|\upsilon|$ and (1.2). Moreover, recalling from (1.1) that $|\upsilon| = \sqrt{2E/m}$, we have used the fact that the probability density $\pi_{\mathtt{s}}(r,\upsilon',\upsilon)\mathrm{d}\upsilon'$ over V is alternatively treated as the density $\pi_{\mathtt{s}}(r,\Omega',E',\Omega,E)\mathrm{d}\Omega'\mathrm{d}E'$ over $\mathbb{S}^2\times(0,\infty)$ under the alternative configuration parameters (i.e., an adjustment in the definition of $\pi_{\mathtt{s}}$ is needed in terms of the new configuration variables), with a similar statement for $\pi_{\mathtt{f}}$ holding. Technically speaking, in (1.7), we should write $\Gamma_0(r,\Omega,E) = |\upsilon|g(r,\Omega,E)$; however, we can simply define this as a new function g representing the initial flux profile.

The system (1.6) and (1.7) is the classical form in which the NTE is presented in the physics and engineering literature. As alluded to previously, we have chosen to deviate from these classical norms as our principal interest in this text is to examine the underpinning mathematical structure of the NTE in relation to stochastic processes. As such, it is more convenient to keep the NTE in as compact a form as possible.

As the reader is probably already aware, there is a technical problem with (1.4) and (1.5). Our calculations do not adhere to the degree of rigour that is needed to confirm that the limits that define derivatives are well defined. For example, it is not clear that we can interpret $\upsilon\cdot\nabla_r\Psi_t$ in a pointwise sense as Ψ_t may not be smooth enough. It turns out that the right way to see the above calculations is in the setting of an appropriate functional space.

1.4 Classical Solutions as an Abstract Cauchy Problem (ACP)

Let us henceforth continue our discussion without the source term for convenience. In other words, we will make the blanket assumption

$$\mathscr{Q}_t\equiv 0, \qquad t\geq 0.$$

The usual way to find solutions to the NTE is to pose it as an *abstract Cauchy problem* (ACP). More formally, we need to move our discussion into the language of functional spaces. To this end, let us introduce $L^2(D \times V)$ as the Hilbert space of square integrable functions on $D \times V$, defined with the usual inner product

$$\langle f, g\rangle = \int_{D\times V} f(r, \upsilon)g(r, \upsilon)\mathrm{d}r\mathrm{d}\upsilon, \qquad f, g \in L^2(D \times V).$$

Before introducing the notion of the NTE in the form of an ACP on $L^2(D \times V)$, we first need to introduce three linear operators whose domains are embedded in $L^2(D \times V)$ (we will be more precise about these later on) and whose actions on $g \in L^2(D \times V)$ are given by

$$\begin{cases} \mathscr{T} g(r, \upsilon) := -\upsilon \cdot \nabla_r g(r, \upsilon) - \sigma(r, \upsilon)g(r, \upsilon) \\ \mathscr{S} g(r, \upsilon) := \int_V g(r, \upsilon')\sigma_{\mathtt{s}}(r, \upsilon')\pi_{\mathtt{s}}(r, \upsilon', \upsilon)\mathrm{d}\upsilon' \\ \mathscr{F} g(r, \upsilon) := \int_V g(r, \upsilon')\sigma_{\mathtt{f}}(r, \upsilon')\pi_{\mathtt{f}}(r, \upsilon', \upsilon)\mathrm{d}\upsilon'. \end{cases} \tag{1.8}$$

We name them, respectively, the transport, scattering, and fission operators.

Then the NTE (1.4) can be posed as the following ACP:

$$\begin{cases} \dfrac{\partial}{\partial t}\Psi_t = (\mathscr{T} + \mathscr{S} + \mathscr{F})\Psi_t \\ \Psi_0 = g, \end{cases} \tag{1.9}$$

where $\Psi_t \in L^2(D \times V)$. Specifically, this means that $(\Psi_t, t \geq 0)$ is continuously differentiable in this space. In other words, there exists a $\upsilon_t \in L^2(D \times V)$, which is time-continuous in $L^2(D \times V)$ with respect to $\|\cdot\|_2 = \langle\cdot, \cdot\rangle^{1/2}$, such that $\lim_{h\to 0}\|h^{-1}(\Psi_{t+h} - \Psi_t) - \upsilon_t\|_2 = 0$ for all $t \geq 0$. Moreover, we refer to ν_t as $\partial\Psi_t/\partial t$, $t \geq 0$.

The theory of c_0-semigroups gives us a straightforward approach to describing (what turns out to be) the unique solution to (1.9). Recall that a c_0-semigroup also goes by the name of a strongly continuous semigroup and, in the present context, this means a family of time-indexed operators, $(V_t, t \geq 0)$, on $L^2(D \times V)$, such that:

(i) $V_0 = \mathrm{Id}$.
(ii) $V_{t+s}g = V_t V_s g$, for all $s, t \geq 0$, $g \in L^2(D \times V)$.
(iii) For all $g \in L^2(D \times V)$, $\lim_{h\to 0}\|V_h g - g\|_2 = 0$.

To see how c_0-semigroups relate to (1.9), let us pre-emptively define $(\Psi_t g, t \geq 0)$ to be the semigroup generated by $\mathscr{G}$ via the orbit

$$\Psi_t g := \exp(t\mathscr{G})g, \qquad g \in L^2(D \times V), \tag{1.10}$$

where we interpret the exponential of the operator $t\mathscr{G}$ in the obvious way,

$$\exp(t\mathscr{G})g = \sum_{k=0}^{\infty} \frac{t^k}{k!} \underbrace{\mathscr{G}[\mathscr{G}[\cdots \mathscr{G}}_{k\text{-fold}} g]].$$

Standard c_0-semigroup theory tells us that $\Psi_t g \in \mathtt{Dom}(\mathscr{G})$ for all $t \geq 0$, providing $g \in \mathtt{Dom}(\mathscr{G})$, where

$$\mathtt{Dom}(\mathscr{G}) := \left\{ g \in L^2(D \times V) : \lim_{h \to 0} h^{-1}(\Psi_h g - g) \text{ exists in } L^2(D \times V) \right\}. \tag{1.11}$$

Understanding the structure of $\mathtt{Dom}(\mathscr{G})$ in more detail now gives us the opportunity to bring (1.5) into the discussion.

Both the scattering and fission operators, $\mathscr{S}$ and $\mathscr{F}$, are $\|\cdot\|_2$-continuous mappings from $L^2(D \times V)$ into itself. Showing this is a straightforward exercise that uses the Cauchy–Schwarz inequality and (H1). The domain of $\mathscr{T}$ is a little more complicated. To this end, let us define

$$\partial(D \times V)^- = \{(r, \upsilon) \in D \times V \text{ such that } r \in \partial D \text{ and } \upsilon \cdot \mathbf{n}_r < 0\}.$$

The domain of $\mathscr{T}$ is known to satisfy

$$\mathtt{Dom}(\mathscr{T}) = \{g \in L^2(D \times V) \text{ such that } \upsilon \cdot \nabla_r g \in L^2(D \times V) \text{ and } g|_{\partial(D\times V)^-} = 0\}. \tag{1.12}$$

As such, it follows that $\mathtt{Dom}(\mathscr{G}) = \mathtt{Dom}(\mathscr{T})$.

Now returning to (1.9), we can conclude the following.

Theorem 1.1 *Defining* $(\Psi_t, t \geq 0)$ *via* (1.10)*, we have that it is the unique classical solution of* (1.9) *in* $L^2(D \times V)$.

1.5 Principal Eigendecomposition and Criticality

One of the main tools used for modelling with the NTE is its leading asymptotic behaviour. Roughly speaking, this means looking for an associated triple of a "leading" eigenvalue $\lambda_* \in \mathbb{R}$ and non-negative left and right eigenfunctions, ϕ and $\tilde{\phi}$, respectively, such that

$$\lambda_* \langle g, \phi \rangle = \langle g, \mathscr{G}\phi \rangle \quad \text{and} \quad \lambda_* \langle \tilde{\phi}, f \rangle = \langle \tilde{\phi}, \mathscr{G} f \rangle,$$

for $g \in L^2(D \times V)$ and $f \in \mathtt{Dom}(\mathscr{G})$. The sense in which λ_* is "leading" turns out to mean that it is the eigenvalue with the largest real component. The accompanying eigenfunction ϕ is often referred to as the "ground state".

Depending on the nature of the underlying generators, a natural behaviour for equations of the type (1.9) is that its solution asymptotically projects on to its ground state in the sense that

$$\Psi_t = \mathrm{e}^{\lambda_* t}\langle \tilde{\phi}, g\rangle \phi + o(\mathrm{e}^{\lambda_* t}) \quad \text{as} \quad t \to \infty. \tag{1.13}$$

The approximation (1.13) can be seen as a functional version of the Perron–Frobenius theorem.[1] Using spectral theory for c_0-semigroups, the following result is one of many different variations of its kind, which provides sufficient conditions for the existence of a ground state for which (1.13) holds.

Theorem 1.2 *Let D be convex and non-empty. We assume the following irreducibility conditions. Assume that the cross sections*

$$\sigma_{\mathtt{f}}(r, \upsilon)\pi_{\mathtt{f}}(r, \upsilon, \upsilon') \text{ and } \sigma_{\mathtt{s}}(r, \upsilon)\pi_{\mathtt{s}}(r, \upsilon, \upsilon')$$

are piecewise continuous[2] *on $\bar{D} \times V \times V$ and*

$$\sigma_{\mathtt{f}}(r, \upsilon)\pi_{\mathtt{f}}(r, \upsilon, \upsilon') > 0 \quad \text{on} \quad D \times V \times V. \tag{1.14}$$

Then:

(i) The neutron transport operator $\mathscr{G}$ has a simple and isolated eigenvalue $\lambda_* > -\infty$, which is leading in the sense that

$$\lambda_* = \sup\{\mathrm{Re}(\lambda) : \lambda \text{ is an eigenvalue of } \mathscr{G}\},$$

and which has corresponding non-negative right eigenfunction, ϕ, in $L^2(D \times V)$ and left eigenfunction $\tilde{\phi}$ in $L^2(D \times V)$.

(ii) There exists an $\varepsilon > 0$ such that, as $t \to \infty$,

$$\|\mathrm{e}^{-\lambda_* t}\Psi_t f - \langle \tilde{\phi}, f\rangle \phi\|_2 = O(\mathrm{e}^{-\varepsilon t}), \tag{1.15}$$

for all $f \in L^2(D \times V)$, where $(\Psi_t, t \geq 0)$ is defined in (1.10).

---

[1] For the reader unfamiliar with the classical Perron–Frobenius theorem, it will be discussed in more detail in the next chapter.

[2] A function is piecewise continuous if its domain can be divided into an exhaustive finite partition (for example, polytopes) such that there is continuity in each element of the partition. This is precisely how cross sections are stored in numerical libraries for modelling of nuclear reactor cores.

The importance of the asymptotic decomposition (1.15) is that it allows us to introduce the notion of *criticality*. The setting $\lambda_* > 0$ indicates an exponential growth in the neutron density, $\lambda_* < 0$ indicates exponential decay in neutron density, and $\lambda_* = 0$ corresponds to stabilisation of the neutron density. When considering the NTE in the context of nuclear reactor modelling, it is clear that the most desirable scenario is $\lambda_* = 0$, which would suggest the production of neutrons through fission is perfectly balanced, on average, by nuclear capture and absorption in control rods and the surrounding shielding materials.

1.6 Remarks on Diffusive Approximations

We may consider the NTE in the form (1.4) as irregular. The difference and advection operators involved in the NTE force us to work with less smooth functions than one would typically work with, for example, in a diffusive setting. As mathematicians, when confronted with such an irregular system, we sometimes have the urge to scale away irregularities via a diffusive approximation. To this end, let us define

$$\mu_t(r, E) := \int_{\mathbb{S}^2} \Psi_t(r, \upsilon\Omega)\mathrm{d}\Omega, \tag{1.16}$$

$$J_t(r, E) := \int_{\mathbb{S}^2} \upsilon\Omega\Psi_t(r, \upsilon\Omega)\mathrm{d}\Omega, \tag{1.17}$$

where we recall that $\upsilon = \sqrt{2E/m}$, $\mathbb{S}^2$ denotes the unit sphere in $\mathbb{R}^3$. Typically, cross sections only depend on velocity through energy, which comes from the fact that the media through which neutrons diffuse are piecewise homogeneous. For convenience, let us write $\varsigma_{\mathtt{s}}(r, E)$ and $\varsigma_{\mathtt{f}}(r, E)$in place of $\sigma_{\mathtt{s}}(r, \upsilon)$ and $\sigma_{\mathtt{f}}(r, \upsilon)$, respectively, and $\varsigma(r, \upsilon) = \varsigma_{\mathtt{f}}(r, \upsilon) + \varsigma_{\mathtt{s}}(r, \upsilon)$. Under the further assumption that scattering and fission are isotropic, it follows that

$$\begin{aligned}\frac{\partial \mu_t}{\partial t}(r, E) &= -\nabla_r \cdot J_t(r, E) - \varsigma(r, E)\mu_t(r, E)\\ &\quad + \int_{E_{min}}^{E_{max}} \mu_t(r, E')\varsigma_{\mathtt{s}}(r, E')\mathrm{d}E'\\ &\quad + \int_{E_{min}}^{E_{max}} \mu_t(r, E')\varsigma_{\mathtt{f}}(r, E')\nu(r, E', E)\mathrm{d}E',\end{aligned} \tag{1.18}$$

where $\nu(r, E', E)$ is the average number of neutrons emerging with energy E' from a fission event caused by a neutron with energy E at position r and E_{min} (respectively, E_{max}) is the energy corresponding to the velocity υ_{min} (respectively, υ_{max}).

Now, under certain assumptions, which will be discussed shortly, Fick's law states that

$$J_t(r, E) = -D\nabla_r \mu_t(r, E), \tag{1.19}$$

where D is the diffusion coefficient. Substituting this into (1.18) yields

$$\begin{aligned}\frac{\partial \mu_t}{\partial t}(r, E) = D\Delta\mu_t(r, E) - \varsigma(r, E)\mu_t(r, E) \\ + \int_{E_{min}}^{E_{max}} \mu_t(r, E')\varsigma_{\mathtt{s}}(r, E')\mathrm{d}E' \\ + \int_{E_{min}}^{E_{max}} \mu_t(r, E')\varsigma_{\mathtt{f}}(r, E')\nu(r, E', E)\mathrm{d}E'.\end{aligned} \tag{1.20}$$

But does such an approximation make sense? Let us return to the approximation given in (1.19). In order to make this substitution, several assumptions are required. First, in order to derive Fick's law, one must assume that absorption occurs at a much slower rate than scattering and so the approximation is technically not valid in highly absorbing regions or vacuums. We also made the assumption of isotropic scattering, which is only really valid at low temperatures. Finally, this approximation is only valid when neutrons are sufficiently far from a neutron source or the surface of a material, since this leads to abrupt changes in the scattering and absorption cross sections. Consider for example Fig. 1.2 that gives us a slice of what the internal geometry of a reactor might look like.[3]

Admittedly, the diffusion approximation works well in many practical settings; however, our objective in this monograph is to understand how to work in the more accurate setting of the NTE in its native form.

1.7 Primer to the Stochastic Approach

Having rejected working with a diffusive approximation, we are still confronted with the inhomogeneous nature of the NTE. This makes the prospect of an analytically closed-form solution to (1.4) highly unlikely. While symmetries of nuclear reactors can be exploited, as alluded to above, each subsection of the reactor may contain many different materials, which, in turn, have different properties. For example, the scattering and fission cross sections, $\sigma_{\mathtt{s}}$, $\sigma_{\mathtt{f}}$, $\pi_{\mathtt{s}}$, and $\pi_{\mathtt{f}}$, depend on both the position and velocity of the incident neutron. It seems inevitable that numerical solutions are the best we can hope for.

[3] We would like to thank researchers from the ANSWERS software group at Jacobs who kindly provided this image.

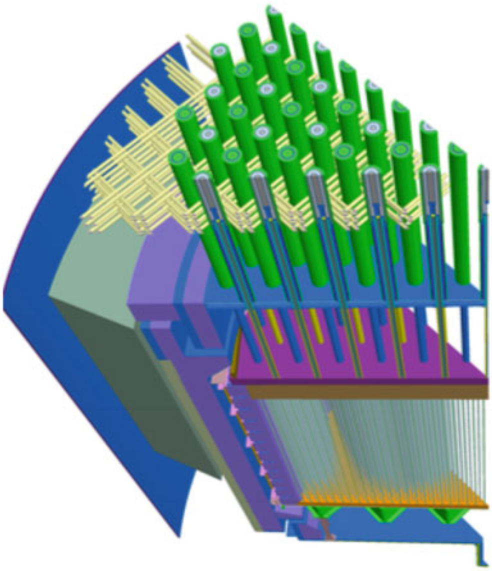

Fig. 1.2 The geometry of a nuclear reactor core representing a physical domain D, on to which the different cross-sectional values of $\sigma_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{s}}, \pi_{\mathtt{f}}$ are mapped, also as a function of neutron velocity

After many years of experimentation, the values of cross sections are well understood for different materials, as well as for different neutron energies (which are in turn a function of velocity). Figure 1.3 gives two examples of cross sections, which are typical in terms of their complexity.[4] Their values are stored in extensive numerical libraries that are held in national archives[5] and are referred to as *nuclear data*. Any numerical model will require extensive reference to nuclear data, and this alone presents an additional source of computational complexity.

Fundamentally, there are two types of numerical solutions to the NTE that one can imagine working with: deterministic and stochastic. Roughly speaking, deterministic numerical models rely on a classical method of meshing the underlying variables of time, space, and velocity and turning the NTE (1.4) into a complex system of algebraic linear equations that one can aim to solve via inversion.

Our objective in this text is to explore the mathematical principles that underpin the stochastic approach. Fundamentally, this entails representing solutions to the NTE as path averages of certain underlying stochastic processes. In other words, we are interested here in the so-called *Feynman–Kac* representation of solutions to the NTE. An understanding of Feynman–Kac representation is what opens the door to Monte Carlo simulation, which is another approach to numerically generating solutions.

---

[4] We would like to thank Prof. Eugene Shwageraus from the Cambridge Nuclear Energy Centre for allowing us to reproduce these images.

[5] Different national nuclear organisations have each generated different libraries.

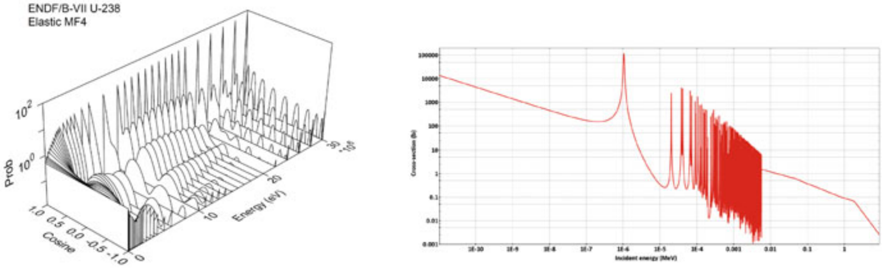

Fig. 1.3 Left: An example of the probability of elastic scatter as a function of energy and polar angle for U-238 (with symmetry in the azimuthal angle), prepared using [27]. Right: An example of microscopic fission cross section for Pu-240 as a function of the incident neutron energy, prepared using [118]. The images were prepared using ENDF-7 nuclear data cf. [20]

In the coming chapters, we will explain how solutions to the NTE can be identified via an entirely different approach that embraces underlying stochastic representations. It will turn out that there are fundamentally two different types of Markov processes whose paths provide us with information about the solutions to the NTE. The first of these is what we call a neutron random walk (NRW), and the second is what we will call a neutron branching process (NBP). Our focus for the remainder of Part I of this book will be to pursue how such processes can be defined, how they fit into an alternative notion of "solution" to the NTE, and how their underlying properties give us insight into the behaviour of these solutions. We also gain some insight into what we should expect of the physical process of neutron fission itself.

1.8 Comments

It is a little difficult to locate the original derivation of the NTE. But one would speculate with confidence that a more formalised mathematical approach was undoubtedly rooted in the *Manhattan Project* and known to the likes of Ulam and von Neumann. In the 1960s, numerous papers and books emerged in which the derivation of the neutron transport equation can be found. Examples include classical texts such as Davison and Sykes [36], with more modern derivations found in [8, 94, 95], for example. A layman's introduction to the role of neutron transport and, more generally, how nuclear reactors work can be found in the outstanding text of Tucker [128], which makes for light-hearted but highly insightful reading.

Already by the 1950s, there was an understanding of how to treat the NTE in special geometries and by imposing isotropic scattering and fission, see, for example, Lehner [91] and Lehner and Wing [92, 93]. It was also understood quite early on that the natural way to state the NTE is via the linear differential transport equation associated to a suitably defined operator on a Banach space (i.e., an L^p space).

Moreover, it was understood that in this formulation, a spectral decomposition should play a key role in representing solutions, as well as identifying the notion of criticality; see, for example, Jörgens [81], Pazy and Rabinowitz [109].

The description of the NTE in the functional setting of an L^p space was promoted by the work of R. Dautray and collaborators, who showed how c_0-semigroups form a natural framework within which one may analyse the existence and uniqueness of solutions to the NTE; see [34] and [33]. A similar approach has also been pioneered by Mokhtar-Kharroubi [98]. The theory of c_0-semigroups is a relatively standard tool, for which a solid account can be found in [52]. For the proof of the precise nature of the domains of $\mathscr{T}$, $\mathscr{S}$, and $\mathscr{F}$, see, for example, Lemma XXI.1 of [33]. As such, Theorem 1.1 is essentially restating Proposition II.6.2 of [52]. Mokhtar-Kharroubi [98] also gives a very precise account of the spectral analysis of the combined generators $\mathscr{T} + \mathscr{S} + \mathscr{F}$.

On a more humorous note, Tucker [128] points out that the significance of criticality is often misunderstood in fictional media. At moments of impending nuclear reactor catastrophe in dramatic storylines, an underling will often be heard reporting to their superior with great anxiety *"The reactor has gone critical!"*. In fact, this is a moment when everyone should go back about their business, content that their power source is operating as intended.

Let us return briefly to some of our earlier remarks on delayed neutrons that have formally entered into the nuclear literature as early as 1947 in Courant and Wallace [28]; see also the 1958 publication of Weinberg and Wigner [130].

As previously mentioned, when a fission event occurs, the nucleus is split into unstable fragments (called isotopes), which later release more neutrons in order to return to a stable energetic state. Thus, modelling neutron density in this case requires one to keep track of the number of prompt neutrons as well as the number of isotopes. We therefore need to enhance the configuration space to distinguish between these two objects.

We will denote type 1 particles to be neutrons and type 2 particles to be isotopes. We write $\Psi^{(2)}(r, \upsilon)$ for the particle density of isotopes at $r \in D$, which carry the *label* $\upsilon \in V$. Although isotopes are stationary in space, for technical reasons, we assign them with a velocity, which is inherited from the incident fission neutron. Accordingly, the configuration space is now $D \times V \times \{1, 2\}$. In order to write down the relevant equations, we introduce the following additional cross sections:

ϑ:	Rate at which neutrons are released from isotopes
$m_i(r, \upsilon)$:	The average number of isotopes resulting from a collision at $r \in D$ with incoming velocity[6] $\upsilon \in V$.
$m_o(r, \upsilon', \upsilon)$:	The average number of delayed neutrons released with outgoing velocity $\upsilon \in V$ from an isotope at $r \in D$ with label υ'.

[6] Note that since isotopes inherit velocity from the incident fission neutron, this means the *label* of the isotope is also υ.

Then, the (forward) NTE with delayed neutrons can be written as follows:

$$
\begin{aligned}
\frac{\partial}{\partial t}\Psi_t^{(1)}(r,\upsilon) = \mathscr{Q}_t(r,\upsilon) - \upsilon\cdot\nabla_r\Psi_t^{(1)}(r,\upsilon) - \sigma(r,\upsilon)\Psi_t^{(1)}(r,\upsilon) \\
+ \int_V \sigma_{\mathtt{f}}(r,\upsilon')\Psi_t^{(1)}(r,\upsilon')\pi_{\mathtt{f}}(r,\upsilon',\upsilon)\mathrm{d}\upsilon' \\
+ \vartheta\int_V \Psi_t^{(2)}(r,\upsilon')m_o(r,\upsilon',\upsilon)\mathrm{d}\upsilon' \\
+ \int_V \sigma_{\mathtt{s}}(r,\upsilon')\Psi_t^{(1)}(r,\upsilon')\pi_{\mathtt{s}}(r,\upsilon',\upsilon)\mathrm{d}\upsilon',
\end{aligned}
\tag{1.21}
$$

with

$$
\frac{\partial}{\partial t}\Psi_t^{(2)}(r,\upsilon) = -\,\vartheta\Psi_t^{(2)}(r,\upsilon) + \sigma_{\mathtt{f}}(r,\upsilon)m_i(r,\upsilon)\Psi_t^{(1)}(r,\upsilon). \tag{1.22}
$$

The above pair of equations illustrate how to include both prompt and delayed neutrons. Of course, we could develop a more precise version of (1.21)–(1.22), since we could also keep track of the different types of isotopes and require that $\pi_{\mathtt{f}}$ also depends on the type of the particle.

In this more general setting, Theorem 1.2 has been stated and proved in greater generality in [30], allowing for the inclusion of other types of nuclear emissions such as alpha, beta, and gamma radiation, as well as delayed neutrons.

An important fact about delayed neutrons is that they are a crucial element of why nuclear reactors can be controlled; cf. discussion in Tucker [128]. The time taken to release these neutrons can be anything from a few seconds to minutes. This has the implication that any control actions that are put into place do not result in instantaneous change in neutron density, but rather a gradual response over seconds and minutes.

Chapter 2
Some Background Markov Process Theory

Before we embark on our journey to explore the NTE in a stochastic context, we need to lay out some core theory of Markov processes that will appear repeatedly in our calculations. After a brief reminder of some basics around the Markov property, we will focus our attention on what we will call expectation semigroups. These are the tools that we will use to identify neutron density and provide an alternative representation of solutions to the NTE that will be of greater interest to us. As such, in this chapter, we include a discussion concerning the asymptotic behaviour of expectation semigroups in terms of a leading eigentriple.

2.1 Markov Processes

We are interested in modelling randomly evolving particles, whose state space is denoted by E. The space E can be a physical space, but it can also be something more abstract. For many of the general results in this text, we will allow E to be a locally compact Hausdorff space with a countable base. The reader may nonetheless be content with taking E to be an open bounded subset of $\mathbb{R}^d$, or in the context of neutron transport, $D \times V$. We will use $\mathscr{B}(E)$ to denote the Borel sets on E.

To E, we append the "cemetery" state $\dagger \notin E$. One may think of it as an auxiliary point where the process is sent if it is "killed". We will take $\mathscr{E}_\dagger$ to be the Borel σ-algebra on $E_\dagger := E \cup \{\dagger\}$ and $E_\dagger^{[0,\infty]}$ to be the space of paths $\omega : [0, \infty] \to E_\dagger$ such that $\omega(\infty) = \dagger$, and if $\omega(t) = \dagger$, then $\omega(s) = \dagger$ for all $s \geq t$.

The sequence of functions $\xi := (\xi_t, t \geq 0)$ represents the family of coordinate maps $\xi_t : E_\dagger^{[0,\infty]} \to E_\dagger$ given by $\xi_t(\omega) = \omega_t$ for all $t \in [0, \infty]$ and $\omega \in E_\dagger^{[0,\infty]}$. Denote by $\mathfrak{G}_t = \sigma(\xi_s, 0 \leq s \leq t)$, and $\mathfrak{G} = \sigma(\xi_s, s \in [0, \infty])$, the canonical filtrations of ξ. We may also think of $\mathfrak{G}_t$ as the space of possible events that occur up to time t and $\mathfrak{G}$ as the space of all possible events.

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8_2

Thus, we may think of ξ as nothing more than an infinite dimensional random variable defined as a measurable function of the event space $E_\dagger^{[0,\infty]}$ via the identity map. What makes ξ a Markov process is the way in which we assign probabilities to events in $\mathfrak{G}$. To this end, let us introduce the family of probability measures on $(E_\dagger^{[0,\infty]}, \mathfrak{G})$, denoted by $\mathbf{P} := (\mathbf{P}_x, x \in E_\dagger)$, which satisfy a number of conditions.

Definition 2.1 (Markov Process) The process $(\xi, \mathbf{P})$ is called a *Markov process* on state space E, with cemetery state $\dagger$ and lifetime ζ (also called the killing time) if the following conditions hold:

(Regularity) For each $B \in \mathfrak{G}$, the map $x \mapsto \mathbf{P}_x(B)$ is $\mathscr{E}_\dagger$-measurable.
(Normality) $\mathbf{P}_x(\xi_0 = x) = 1$ for all $x \in E_\dagger$.
(càdlàg[1] paths) For all $x \in E$, the path functions $t \to \xi_t$ are $\mathbf{P}_x$-almost surely right continuous on $[0, \infty)$ and have left limits on $[0, \zeta)$ where the random time

$$\zeta = \inf\{t > 0 : \xi_t = \dagger\} \tag{2.1}$$

is called the *lifetime* of ξ.
(Markov property) For all $x \in E_\dagger$, $s, t \geq 0$, and bounded measurable functions f on $E_\dagger$ such that $f(\dagger) = 0$, we have on $\{t < \zeta\}$

$$\mathbf{E}_x[f(\xi_{t+s})|\mathfrak{G}_t] = \mathbf{E}_y[f(\xi_s)]_{y=\xi_t}$$

$\mathbf{P}_x$-almost surely.

In the setting where $\mathbf{P}_x(\zeta < \infty) = 0$ for all $x \in E$, we call $(\xi, \mathbf{P})$ a *conservative Markov process*. However, throughout this text, we will often encounter *non-conservative Markov processes.*

In Definition 2.1, we can also ask for a slightly stronger notion of the Markov property, namely the *strong Markov property*. To this end, let us recall that τ is a *stopping time* for the filtration $(\mathfrak{G}_t, t \geq 0)$ if

$$\{\tau \leq t\} \in \mathfrak{G}_t, \qquad t \geq 0.$$

Moreover, we can identify the natural sigma algebra $\mathfrak{G}_\tau$ associated to any stopping time τ via the definition

$$A \in \mathfrak{G}_\tau \text{ if and only if } A \cap \{\tau \leq t\} \in \mathfrak{G}_t \text{ for all } t > 0.$$

The most common examples of stopping times that we will work with are those that take the form

$$\tau^D = \inf\{t > 0 : \xi_t \in D\},$$

---

[1] Continue à droite, limite à gauche.

where D is a Borel-measurable domain in E. We also note from (2.1) that $\zeta = \tau^{\{\dagger\}}$, and the lifetime of ξ is also a stopping time when we consider the Markov process to be defined on the extended space with the cemetery point included.

Definition 2.2 (Strong Markov Process) In addition to being a Markov process, $(\xi, \mathbf{P})$ is a *strong Markov process* if the Markov property is replaced by the following condition.
(Strong Markov Property) For all $x \in E_\dagger$, $s \geq 0$, stopping times τ, and bounded measurable functions f on $E_\dagger$ such that $f(\dagger) = 0$, on the event $\{\tau < \zeta\}$, we have

$$\mathbf{E}_x[f(\xi_{\tau+s})|\mathfrak{G}_\tau] = \mathbf{E}_y[f(\xi_s)]_{y=\xi_\tau}$$

$\mathbf{P}_x$-almost surely.

2.2 Expectation Semigroups and Evolution Equations

Let $B(E)$ be the space of bounded, measurable functions on E with the additional requirement that $f \in B(E)$ is forced to satisfy $f(\dagger) = 0$. The subset $B^+(E)$ of $B(E)$ consists only of non-negative functions. If we define $\|\cdot\|$ to be the supremum norm, then the pair $(B(E), \|\cdot\|)$ forms a Banach space.

Still denoting $(\xi, \mathbf{P})$ as a Markov process, for $s, t \geq 0$, $g \in B(E)$, and $x \in E$, define

$$\mathtt{Q}_t[g](x) = \mathbf{E}_x[g(\xi_t)]. \tag{2.2}$$

We refer to the family of operators $\mathtt{Q} = (\mathtt{Q}_t, t \geq 0)$ on $B(E)$ as the *expectation semigroup* associated to $(\xi, \mathbf{P})$. Technically speaking, we can think of $\mathtt{Q}$ as a family of linear operators on $(B(E), \|\cdot\|)$. It is also not unusual to restrict the definition in (2.2) to just $g \in B^+(E)$.

Expectation semigroups are natural objects that characterise the mean evolution of a Markov process. The notion of the expectation semigroup can also be taken in a slightly broader sense. Suppose that $\gamma : E \mapsto \mathbb{R}$. Then we can also define

$$\mathtt{Q}_t^\gamma[g](x) = \mathbf{E}_x\left[\mathrm{e}^{\int_0^t \gamma(\xi_s)\mathrm{d}s} g(\xi_t)\right], \qquad x \in E,\ t \geq 0. \tag{2.3}$$

It may of course be the case that, without further restriction in γ, the above definition may take infinite values, or not be well defined for $g \in B(E)$. As such, we shall henceforth assume that $\gamma \in B(E)$. This allows us to deduce that $\mathtt{Q}_t^\gamma[g] \in B(E)$. Moreover, this also gives us simple control over the growth of $(\mathtt{Q}_t^\gamma, t \geq 0)$ in the sense that

$$\mathtt{Q}_t^\gamma[g](x) \leq \mathrm{e}^{\|\gamma\|t}\|g\|, \qquad t \geq 0, x \in E, \gamma, g \in B(E). \tag{2.4}$$

The term $\exp\left(\int_0^t \gamma(\xi_s)\mathrm{d}s\right)$, $t \geq 0$, is often referred to as a *multiplicative potential*. In essence, $\mathtt{Q}^\gamma$ is a semigroup in the sense of (2.2) for the bivariate Markov process

$$\left(\xi_t\,,\ \int_0^t \gamma(\xi_s)\mathrm{d}s\right), \qquad t \geq 0, \tag{2.5}$$

albeit that we are forcing a choice of function of the above pair to take the form $g(x)\mathrm{e}^y$, with x playing the role of ξ_t and y playing the role of $\int_0^t \gamma(\xi_s)\mathrm{d}s$. To see why the pair (2.5) forms a Markov process, let us write

$$I_t = I + \int_0^t \gamma(\xi_s)\mathrm{d}s, \qquad t \geq 0,$$

and we can abuse our existing notation and write $\mathbf{P}_{(x,I)}$ for the law of the pair (ξ_t, I_t), $t \geq 0$, when issued from (x, I). Moreover, using $\dagger$ as the cemetery state for $(I_t, t \geq 0)$, with the lifetime of the pair being that of ξ, note that, for bounded, measurable $f : E \times \mathbb{R} \mapsto [0, \infty)$ such that $f(\dagger, \dagger) = 0$ and $s, t \geq 0$, on $\{t < \zeta\}$,

$$\begin{aligned}
&\mathbf{E}_{(x,I)}\left[f(\xi_{t+s}, I_{t+s})|\mathfrak{G}_t\right] \\
&= \mathbf{E}_{(x,I)}\left[f\left(\xi_{t+s}, I + \int_0^t \gamma(\xi_u)\mathrm{d}u + \int_t^{t+s} \gamma(\xi_u)\mathrm{d}u\right)\Big|\,\mathfrak{G}_t\right] \\
&= \mathbf{E}_{(y,J)}\left[f\,(\xi_s, I_s)\right]_{y=\xi_t,\, J=I_t}.
\end{aligned}$$

Technically speaking, we can drop the aforesaid qualifier that $\{t < \zeta\}$ because our convention would insist that $\mathbf{E}_{(y,J)}[f(\xi_s, I_s)] = 0$ for $(y, J) = (\dagger, \dagger)$.

The following lemma gives us a sense of why (2.2) and (2.3) are deserving of the name semigroup.

Lemma 2.1 *Suppose that γ is uniformly bounded from above. The expectation semigroup $\mathtt{Q}^\gamma$ satisfies the semigroup property. That is, for all $g \in B(E)$,*

$$\mathtt{Q}_t^\gamma[\mathtt{Q}_s^\gamma[g]] = \mathtt{Q}_{t+s}^\gamma[g] \text{ on } E,$$

with the understanding that $\mathtt{Q}_0^\gamma[g] = g$.

Proof This is a simple consequence of the Markov property. Indeed, for all $s, t \geq 0$ and $x \in E$,

$$\begin{aligned}
\mathtt{Q}_{t+s}^\gamma[g](x) &= \mathbf{E}_x\left[\mathrm{e}^{\int_0^{t+s}\gamma(\xi_s)\mathrm{d}s}g(\xi_{t+s})\right] \\
&= \mathbf{E}_x\left[\mathrm{e}^{\int_0^{t}\gamma(\xi_s)\mathrm{d}s}\mathbf{E}_y\left[\mathrm{e}^{\int_0^{s}\gamma(\xi_s)\mathrm{d}s}g(\xi_s)\right]_{y=\xi_t}\right]
\end{aligned}$$

$$= \mathtt{Q}_t^\gamma[\mathtt{Q}_s^\gamma[g]](x),$$

as required.

Remark 2.1 In the setting that $\gamma \le 0$, we can alternatively see $\mathtt{Q}^\gamma$ as the expectation semigroup of an auxiliary Markov process that agrees with $(\xi, \mathbf{P})$ but experiences an additional form of killing at rate γ, i.e., when at $x \in E$, the process is killed at rate $\gamma(x)$. One may think of the killing as a clock that rings, sending ξ to the cemetery state $\dagger$ in a way that depends on the path of ξ (in addition to the possibility that ξ has a finite lifetime). Indeed, if we denote by T the time at which the clock rings, then, given $(\xi_s, s \le t)$, remembering that we are specifically in the setting that $\gamma \le 0$, the probability that the clock has not rung by time $t \ge 0$ on the event $\{t < \zeta\}$ is given by

$$\Pr(T > t|\mathfrak{G}_t) = \mathrm{e}^{\int_0^t \gamma(\xi_s)\mathrm{d}s},$$

where we recall that $\mathfrak{G}_t = \sigma(\xi_s, s \le t)$. As such, even if $\mathtt{Q}$ is conservative, $\mathtt{Q}^\gamma$ can always be seen as a non-conservative Markov process. $\diamond$

Often, the expectation semigroup $\mathtt{Q}^\gamma = (\mathtt{Q}_t^\gamma, t \ge 0)$ forms the basis of an evolution

$$\chi_t(x) := \mathtt{Q}_t^\gamma[g](x) + \int_0^t \mathtt{Q}_s^\gamma[h_{t-s}](x)\mathrm{d}s, \qquad t \ge 0, x \in E, \tag{2.6}$$

where $g \in B^+(E)$ and $h : [0,\infty) \times E \to [0,\infty)$ is such that $\sup_{s \le t} |h_s| \in B^+(E)$ for all $t \ge 0$. In such equations, $(\mathtt{Q}_t^\gamma, t \ge 0)$ is called the *driving semigroup*. Note that the assumptions on γ, g, and $(h_t, t \ge 0)$ imply that $\sup_{s \le t} |\chi_s| \in B^+(E)$ for all $t \ge 0$.

Before considering an alternative form of (2.6), let us first consider an example of such an evolution equation. Consider a Brownian motion, $(B_t, t \ge 0)$, in $\mathbb{R}$ that, at rate $\alpha \in B^+(\mathbb{R})$, jumps to a new position in $\mathbb{R}$ according to the law μ. Hence, if T is the time of the first such jump, then

$$\Pr(T > t|\sigma(B_s : s \le t)) = \mathrm{e}^{-\int_0^t \alpha(B_s)\mathrm{d}s}, \qquad t \ge 0.$$

Let $(X, \mathbb{P}_x)$ denote this process when initiated from a single particle at $x \in \mathbb{R}$ and set $\chi_t(x) := \mathbb{E}_x[f(X_t)]$, for $f \in B^+(\mathbb{R})$. Then, by conditioning on the time of the first jump of the process, we have

$$\chi_t(x) = \mathbf{E}_x[\mathrm{e}^{-\int_0^t \alpha(B_s)\mathrm{d}s} f(B_t)] \\ + \int_0^t \mathbf{E}_x\left[\alpha(B_s)\mathrm{e}^{-\int_0^s \alpha(B_u)\mathrm{d}u} \int_{\mathbb{R}} \mathbb{E}_y[f(X_{t-s})]\mu(\mathrm{d}y)\right]\mathrm{d}s$$

$$= \mathtt{Q}_t^{-\alpha}[f](x) + \int_0^t \mathtt{Q}_s^{-\alpha}[h_{t-s}]\mathrm{d}s, \tag{2.7}$$

where $h_{t-s}(x) = \alpha(x) \int_{\mathbb{R}} \chi_{t-s}(y)\mu(\mathrm{d}y)$.

Let us now turn to an alternative form of (2.6), which will turn out to be of repeated use later in this text and will allow us to work effectively with such evolution equations.

Theorem 2.1 *Suppose that $|\gamma| \in B^+(E)$, $g \in B^+(E)$, and $\sup_{s\le t} |h_s| \in B^+(E)$, for all $t \geq 0$. If $(\chi_t, t \geq 0)$ is represented by (2.6), then it also solves*

$$\chi_t(x) = \mathtt{Q}_t[g](x) + \int_0^t \mathtt{Q}_s[h_{t-s} + \gamma\chi_{t-s}](x)\mathrm{d}s, \qquad t \geq 0,\ x \in E. \tag{2.8}$$

The converse statement is also true if $(\chi_t, t \geq 0)$ solves (2.8) with $\sup_{s\le t} |\chi_s| \in B^+(E)$, for all $t \geq 0$.

One should understand Theorem 2.1 as a form of calculus by which the *multiplicative potential* in $\mathtt{Q}^\gamma$ is removed and appears instead as an *additive potential* in the integral term. We also note that within the class of solutions $(\chi_t, t \geq 0)$ for which $\sup_{s\le t} |\chi_s| \in B^+(E)$, for all $t \geq 0$, both (2.6) and (2.8) have unique solutions. This is easy to see via Grönwall's Lemma.

For example, in the case of (2.8), suppose $\chi_t^{(i)}(x)$, $t \geq 0$, $x \in E$, $i = 1, 2$, represent two solutions and, accordingly, $\tilde{\chi}_t = \sup_{x\in E} |\chi_t^{(1)}(x) - \chi_t^{(2)}(x)|$. It is easy to see with the help of (2.4) that $\sup_{s\le t} \tilde{\chi}_s < \infty$, for $s \leq t$. With the assumption $|\gamma| \in B^+(E)$, we easily deduce that

$$\tilde{\chi}_t \leq \int_0^t \tilde{\chi}_{t-s}\mathtt{Q}_s[|\gamma|](x)\mathrm{d}s \leq C \int_0^t \tilde{\chi}_{t-s}\mathrm{d}s, \qquad t \geq 0, x \in E,$$

for some constant $C > 0$. Grönwall's Lemma implies that $\tilde{\chi}_t = 0$, $t \geq 0$, and this gives us uniqueness of (2.8) as required. A similar argument can be produced for (2.6) to prove the claimed uniqueness in this case.

Proof of Theorem 2.1 We start by noting that if we write

$$\Gamma_t = \mathrm{e}^{\int_0^t \gamma(\xi_s)\mathrm{d}s}, \qquad t \geq 0,$$

then straightforward calculus tells us that on $\{t < \zeta\}$,

$$\mathrm{d}\left(\frac{1}{\Gamma_t}\right) = -\frac{\gamma(\xi_t)}{\Gamma_t}\mathrm{d}t.$$

As such, for $t \geq 0$,

$$\frac{1}{\Gamma_t} = 1 - \int_0^t \frac{\gamma(\xi_u)}{\Gamma_u}\mathrm{d}u. \tag{2.9}$$

Now assume that (2.6) holds. We have, with the help of the Markov property, the simple computation

$$\begin{aligned}
\mathtt{Q}_s[\gamma \mathtt{Q}^{\gamma}_{t-s}[g]](x) &= \mathbf{E}_x\left[\gamma(\xi_s)\mathbf{E}_y\left[\Gamma_{t-s} g(\xi_{t-s})\right]_{y=\xi_s}\right] \\
&= \mathbf{E}_x\left[\gamma(\xi_s)\frac{\Gamma_t}{\Gamma_s} g(\xi_t)\right]
\end{aligned}$$

and, hence using (2.6),

$$\begin{aligned}
\mathtt{Q}_s[\gamma\chi_{t-s}](x) &= \mathtt{Q}_s[\gamma \mathtt{Q}^{\gamma}_{t-s}[g]](x) + \int_0^{t-s} \mathtt{Q}_s\left[\gamma \mathtt{Q}^{\gamma}_u[h_{t-s-u}]\right](x)\mathrm{d}u \\
&= \mathbf{E}_x\left[\gamma(\xi_s)\frac{\Gamma_t}{\Gamma_s}g(\xi_t)\right] + \int_0^{t-s}\mathbf{E}_x\left[\gamma(\xi_s)\frac{\Gamma_{s+u}}{\Gamma_s}h_{t-s-u}(\xi_{s+u})\right]\mathrm{d}u \\
&= \mathbf{E}_x\left[\gamma(\xi_s)\frac{\Gamma_t}{\Gamma_s}g(\xi_t)\right] + \int_s^{t}\mathbf{E}_x\left[\gamma(\xi_s)\frac{\Gamma_{u}}{\Gamma_s}h_{t-u}(\xi_{u})\right]\mathrm{d}u.
\end{aligned}$$

Integrating again, applying Fubini's theorem, and using (2.9), we get

$$\begin{aligned}
\int_0^t \mathtt{Q}_s[\gamma\chi_{t-s}](x)\mathrm{d}s &= \mathbf{E}_x\left[\Gamma_t g(\xi_t)\int_0^t \frac{\gamma(\xi_s)}{\Gamma_s}\mathrm{d}s\right] \\
&\qquad + \int_0^t\int_s^t \mathbf{E}_x\left[\gamma(\xi_s)\frac{\Gamma_u}{\Gamma_s}h_{t-u}(\xi_u)\right]\mathrm{d}u\,\mathrm{d}s \\
&= \mathbf{E}_x\left[\Gamma_t g(\xi_t)\int_0^t \frac{\gamma(\xi_s)}{\Gamma_s}\mathrm{d}s\right] \\
&\qquad + \int_0^t \mathbf{E}_x\left[\Gamma_u h_{t-u}(\xi_u)\int_0^u \frac{\gamma(\xi_s)}{\Gamma_s}\mathrm{d}s\right]\mathrm{d}u \\
&= \mathtt{Q}^{\gamma}_t[g](x) - \mathtt{Q}_t[g](x) + \int_0^t \left(\mathtt{Q}^{\gamma}_u[h_{t-u}](x) - \mathtt{Q}_u[h_{t-u}](x)\right)\mathrm{d}u.
\end{aligned}$$

Rearranging, this tells us that

$$\mathtt{Q}_t[g](x) + \int_0^t \mathtt{Q}_s[h_{t-s} + \gamma\chi_{t-s}](x)\mathrm{d}s = \mathtt{Q}^{\gamma}_t[g](x) + \int_0^t \mathtt{Q}^{\gamma}_u[h_{t-u}](x)\mathrm{d}u.$$

Said another way, (2.6) implies (2.8).

Reversing the arguments above, with the assumption that $\sup_{s\le t}|\chi_s| \in B^+(E)$, for all $t \ge 0$, we also see that (2.8) solves (2.6).

2.3 The Heuristics of Infinitesimal Generators

In the previous section, we alluded to the fact that expectation semigroups are natural objects that characterise the law of a Markov process. It turns out that this can be brought out in sharper focus by considering the rate of change of an expectation semigroup, which brings forward the notion of an infinitesimal generator. In turn, this connects us back to the medium of evolution equations. The discussion we will present here will be entirely heuristic, although in later chapters we will see examples of the general presentation here in careful detail.

Under the right assumptions, it turns out that, for functions g in an appropriate subset of $B(E)$,

$$\mathtt{L}[g](x) := \lim_{t\downarrow 0} \frac{\mathrm{d}}{\mathrm{d}t}\mathtt{Q}_t[g](x), \qquad x \in E, \tag{2.10}$$

is well defined, where $\mathtt{L}$ is an operator whose image is a subspace of functions that map E into $\mathbb{R}$. For the sake of convenience, and to be consistent with what is presented in Chap. 1, let us refer to the set of functions $g \in B(E)$ for which (2.10) holds as $\mathtt{Dom(L)}$.

We may think of the operator $\mathtt{L}$ as telling us everything we need to know about the expectation semigroup and hence about the law of the Markov process. Therefore, if it can be identified, it is a natural "mathematical package" with which to characterise the Markov process.

As a first observation in this respect, we can easily note that, if we write $\mathtt{L}^\gamma$ for the generator associated to $\mathtt{Q}^\gamma$, then, providing γ is bounded from above,

$$\mathtt{L}^\gamma = \mathtt{L} + \gamma. \tag{2.11}$$

Indeed, for all $g, \gamma \in B(E)$ such that $\mathtt{L}g$ is well defined,

$$\begin{aligned}
&\lim_{t\downarrow 0} \frac{\mathtt{Q}^\gamma_t[g](x) - g(x)}{t} \\
&= \lim_{t\downarrow 0} \frac{\mathbf{E}_x[g(\xi_t)] - g(x)}{t} + \lim_{t\downarrow 0} \mathbf{E}_x\left[\frac{(\mathrm{e}^{\int_0^t \gamma(\xi_s)\mathrm{d}s} - 1)}{t} g(\xi_t)\right] \\
&= \mathtt{L}g(x) + \gamma g(x), \qquad x \in E,
\end{aligned} \tag{2.12}$$

where the second term on the right-hand side follows from bounded convergence (cf. (2.4)).

As a second observation, let us consider the setting that $(\xi, \mathbf{P})$ is a continuous-time Markov chain on $E = \{1, \ldots, n\}$. In that case, our expectation semigroup captures nothing more than the transition matrix. Indeed, suppose $(p_t(i, j), i, j \in \{1, \cdots, n\})$ are the transition probabilities of our chain. Any function $g \in B(E)$ can be represented as a vector, and hence, for $i \in \{1, \cdots, n\}$,

$$\mathtt{Q}_t[g](i) = \mathbf{E}_i[g(\xi_t)] = \sum_{j=1}^{n} p_t(i,j)g(j), \qquad t \geq 0, \tag{2.13}$$

so that $\mathtt{Q}_t[g](i)$ is nothing more than the i-th row of the matrix multiplication of $\mathtt{Q}_t$ and the vector g.

In this setting, we also know that we can write $\mathtt{Q}_t = \exp(t\mathtt{L})$, $t \geq 0$, where the $n \times n$ matrix $\mathtt{L}$ is the so-called the *intensity matrix* of ξ, whose off-diagonal entries give the rates of transition from one state to another, and whose row sums are identically zero. Moreover, $\mathtt{Dom}(\mathtt{L})$ is the n-dimensional Euclidian vector space, $\mathbb{R}^n$.

For more complex Markov processes with an uncountable state space, the operator $\mathtt{L}$ does not necessarily take as simple a form as a matrix. In the Markov chain example, we understand

$$\exp(t\mathtt{L}) = \sum_{n=0}^{\infty} \frac{t^n}{n!}\mathtt{L}^n \tag{2.14}$$

to be a series of matrices and hence a matrix itself. The meaning of (2.14) for a more general setting remains a formality, in a similar spirit to the discussion in Sect. 1.9. Nonetheless, in this book, we will encounter explicit analytical representations of the operator $\mathtt{L}$ for those Markov processes that will be of concern to us, which will prove to be informative at both heuristic and rigorous levels.

Tautologically speaking, the definition of the infinitesimal generator (2.10) makes sense whenever the limit is well defined. In essence, one may think of the existence of the limit in (2.10) as a restriction on the class of functions on which the operator $\mathtt{L}$ acts. In order to develop a theory that characterises this class of functions, one generally needs to ask more of the semigroup $\mathtt{Q}$. This brings us to the notion of a *Feller semigroup*.

We say that $(\mathtt{Q}_t, t \geq 0)$ is a Feller semigroup if (i) for $t \geq 0$, $\mathtt{Q}_t$ maps, $C^+(E)$, the space of non-negative, bounded continuous functions on E, to itself, and (ii) for each $f \in C^+(E)$ and $x \in E$, $\lim_{t\to 0} \mathtt{Q}_t[f](x) = f(x)$. It turns out that for Feller semigroups, the generator as defined in (2.10) is well defined providing $f \in C_0^+(E)$, the space of functions in $C^+(E)$ that converge to zero at any infinite boundary points.[2] Stochastic processes with Feller semigroups are, naturally, called *Feller processes* and have additional convenient properties. This includes, for example, the ability to define versions of the process with right-continuous paths with left limits as well as possessing the strong Markov property.

In this text, we generally avoid assuming that our semigroups are Feller, with the exception of some examples. The reason for this is that the stochastic processes we will see in the setting of neutron transport will turn out not to have Feller

---

[2] This means $f(x) \to 0$ as $\inf_{y\in\partial E}\|x - y\| \to 0$, providing the latter limit is possible within E.

semigroups. This also largely explains why we avoid working with generators, preferring instead to work directly with semigroups.

2.4 Feynman–Kac Heuristics

Recall that we write $g \in \mathtt{Dom}(\mathtt{L})$ as the subset of $B(E)$ for which (2.10) holds. If $g \in \mathtt{Dom}(\mathtt{L})$, then, heuristically speaking, it should also be the case that $\mathtt{Q}_t[g] \in \mathtt{Dom}(\mathtt{L})$, for all $t \geq 0$. Indeed, the semigroup property allows us to deduce that, for $g \in \mathtt{Dom}(\mathtt{L})$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathtt{Q}_t[g] = \lim_{h\to 0} \frac{\mathtt{Q}_{t+h}[g] - \mathtt{Q}_t[g]}{h} = \lim_{h\to 0} \frac{\mathtt{Q}_h\mathtt{Q}_t[g] - \mathtt{Q}_t[g]}{h} = \mathtt{L}\mathtt{Q}_t[g]. \tag{2.15}$$

Similarly, providing we can pass the limit and derivative through the semigroup, we also have that, for $g \in \mathtt{Dom}(\mathtt{L})$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathtt{Q}_t[g] = \lim_{h\to 0} \mathtt{Q}_t\left[\frac{\mathtt{Q}_h[g] - g}{h}\right] = \mathtt{Q}_t\left[\lim_{h\to 0} \frac{\mathtt{Q}_h[g] - g}{h}\right] = \mathtt{Q}_t[\mathtt{L}g]. \tag{2.16}$$

Clearly, some careful checking and perhaps additional assumptions may be needed to turn these heuristics into rigour.

An alternative way of reading (2.15) and (2.16) is in a milder form, via the integral equations

$$\mathtt{Q}_t[g] = \mathtt{Q}_s[g] + \int_s^t \mathtt{L}\mathtt{Q}_u[g]\mathrm{d}u = \mathtt{Q}_s[g] + \int_s^t \mathtt{Q}_u[\mathtt{L}g]\mathrm{d}u, \qquad t \geq s \geq 0.$$

This is a very basic form of integral equation in the spirit of (2.6) and (2.8). In later chapters of this text, we will often refer to equations of this form as *mild* equations, describing expectation semigroup evolutions.

Generally speaking, the relationship between the expectation semigroup $(\mathtt{Q}_t, t \geq 0)$ and the differential equation (2.15) is referred to as a *Feynman–Kac* representation. A more formalised Feynman–Kac theory will concern itself with developing appropriate mathematical conditions under which the solution to the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}u_t(x) = \mathtt{L}u_t(x), \qquad x \in E,\ t \geq 0, \tag{2.17}$$

with initial condition $u_0 = g$, can be uniquely identified as $(\mathtt{Q}_t[g],\ t \geq 0)$ on $(B(E), \|\cdot\|)$.

A more general version of (2.17) that carries an association with the stochastic process $(\xi, \mathbf{P})$ via the use of expectation semigroups is the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}u_t(x) = (\mathtt{L} + \gamma)u_t(x) + h, \qquad x \in E,\ t \geq 0, \tag{2.18}$$

where $h \in B^+(E)$, with initial condition $u_0 = g$. For (2.18), Feynman–Kac theory will provide sufficient conditions for which its unique solution is given by the classical *Feynman–Kac formula*

$$u_t(x) = \mathbf{E}_x\left[\mathrm{e}^{\int_0^t \gamma(\xi_s)\mathrm{d}s} g(\xi_t)\right] + \mathbf{E}_x\left[\int_0^t \mathrm{e}^{\int_0^u \gamma(\xi_s)\mathrm{d}s} h(\xi_u)\mathrm{d}u\right], \qquad x \in E,\ t \geq 0. \tag{2.19}$$

The Feynman–Kac solution (2.19) is a predictable formula if we reflect on the earlier demonstrated fact that the pair $(\xi_t\,,\ \int_0^t \gamma(\xi_s)\mathrm{d}s)$, $t \geq 0$, is Markovian. In a similar spirit, we can show that the triplet

$$(\xi_t, I_t, J_t) := \left(\xi_t\,,\ \int_0^t \gamma(\xi_s)\mathrm{d}s,\ \int_0^t \mathrm{e}^{\int_0^u \gamma(\xi_s)\mathrm{d}s} h(\xi_u)\mathrm{d}u\right), \qquad t \geq 0$$

is Markovian. As such, it is implicit that in (2.19), one should think of the cemetery state as being $(\dagger, \dagger, \dagger)$ for the triplet above and that functionals such as $F(J_t) = J_t$ carry the additional convention that $F(\dagger) = 0$.

If we define $\mathtt{Q}_t^{\gamma,h}[g](x)$ to be the right-hand side of (2.19), then the aforesaid Markov property can be used to show that

$$\mathtt{Q}_{t+s}^{\gamma,h}[g](x) = \mathtt{Q}_t^{\gamma,h}[\mathtt{Q}_s^{\gamma,h}[g](x)](x) + \mathbf{E}_x\left[\int_0^t \mathrm{e}^{\int_0^u \gamma(\xi_s)\mathrm{d}s} h(\xi_u)\mathrm{d}u\right], \tag{2.20}$$

for $s, t \geq 0$. Unfortunately, this means that $(\mathtt{Q}_t^{\gamma,h}, t \geq 0)$ is not a semigroup as we have defined it. Nonetheless, we can follow our heuristic notion of infinitesimal generator and calculate

$$\mathtt{L}^{\gamma,h}[g](x) := \lim_{t\downarrow 0}\frac{\mathrm{d}}{\mathrm{d}t}\mathtt{Q}_t^{\gamma,h}[g](x), \qquad x \in E. \tag{2.21}$$

In the spirit of (2.12), we can easily see that

$$\lim_{t\downarrow 0}\frac{\mathrm{d}}{\mathrm{d}t}\mathtt{Q}_t^{\gamma,h}[g](x) = (\mathtt{L} + \gamma)g + h.$$

This clearly brings us back to the right-hand side of (2.18), at least heuristically, taking account of (2.20).

2.5 Perron–Frobenius-Type Behaviour

Recalling the discussion in the previous section, let us continue with the example that $(\xi, \mathbf{P})$ is a Markov chain and $E = \{1, \ldots, n\}$. In this setting, the infinitesimal generator $\mathtt{L}$ corresponds to the *intensity matrix* of our Markov chain, that is, $\mathtt{L} = (\mathtt{L}(i, j), i, j \in \{1, \ldots, n\})$, where

$$\mathtt{L}(i, j) = \frac{\mathrm{d}}{\mathrm{d}t} p_t(i, j)|_{t=0}$$

and the transitions $p_t(i, j)$ were introduced just before (2.13). We assume that $\mathtt{L}(i, j) > 0$ for all $i, j \in E$ with $i \neq j$. However, we do not exclude the possibility of an additional cemetery state † and hence of $(\mathtt{L}\mathbf{1})_i \leq 0$ for $i = 1, \ldots, n$, where $\mathbf{1} = (1, \ldots, 1)$. That is to say, the row sums of $\mathtt{L}$ are at most zero indicating the possibility that the Markov chain may jump to † from at least one of the states in E.

Perron–Frobenius theory tells us that, subject to the requirement that E is irreducible,[3] there exists a leading eigenvalue $\lambda_{\mathtt{c}} \leq 0$ such that

$$p_t(i, j) \sim \mathrm{e}^{\lambda_{\mathtt{c}} t} \varphi(i)\tilde{\varphi}(j) + o(\mathrm{e}^{\lambda_{\mathtt{c}} t}) \text{ as } t \to \infty, \tag{2.22}$$

where $\varphi = (\varphi(1), \ldots, \varphi(n))$, resp. $\tilde{\varphi} = (\tilde{\varphi}(1), \ldots, \tilde{\varphi}(n))$, are the (unique up to a multiplicative constant) right, resp. left, eigenvectors of $\mathtt{L}$ with eigenvalue $\lambda_{\mathtt{c}}$. In relation to $\mathtt{L}$, φ and $\tilde{\varphi}$ satisfy

$$\sum_{j=1}^{n} L(i, j)\varphi(j) = \lambda_c \varphi(i) \quad \text{and} \quad \sum_{k=1}^{n} \tilde{\varphi}(k) L(k, i) = \lambda_c \tilde{\varphi}(i). \tag{2.23}$$

This implies that

$$\langle g, \mathtt{L}\varphi \rangle = \lambda_{\mathtt{c}} \langle g, \varphi \rangle \text{ and } \langle \tilde{\varphi}, \mathtt{L}g \rangle = \lambda_{\mathtt{c}} \langle \tilde{\varphi}, g \rangle, \tag{2.24}$$

for all $g \in B(E)$, where we extend existing notation and write $\mathtt{L}g$ for matrix multiplication of g by $\mathtt{L}$ and $\langle \cdot, \cdot \rangle$ is the usual Euclidian inner dot product.[4] Equivalently, we may identify φ and $\tilde{\varphi}$ by the stability equations

$$\langle g, \mathtt{Q}_t[\varphi] \rangle = \mathrm{e}^{\lambda_{\mathtt{c}} t} \langle g, \varphi \rangle \text{ and } \langle \tilde{\varphi}, \mathtt{Q}_t[g] \rangle = \mathrm{e}^{\lambda_{\mathtt{c}} t} \langle \tilde{\varphi}, g \rangle, \qquad t \geq 0. \tag{2.25}$$

The reason why the eigenvalue λ_c is non-positive here is that the asymptotic (2.22) captures the rate at which probability is lost from the Markov chain

[3] Recall that a Markov chain is irreducible if, for each pair (i, j) in the state space, there is a non-zero probability that, starting in state i, one will eventually visit state j.

[4] Elsewhere in this text we use the notation $\langle \cdot, \cdot \rangle$ to denote inner products on other Hilbert spaces.

because of absorption to the cemetery state. If the Markov chain is conservative, i.e., $(\mathtt{L}\mathbf{1})_i = 0$, for $i = 1, \ldots, n$, then we see that $\lambda_c = 0$ and $\varphi = \mathbf{1}$. Recalling the chain is irreducible, this also allows us to deduce that $\tilde{\varphi}$ is the *stationary distribution* of the chain. Indeed, from the second inequality in (2.25), we see that

$$\langle \tilde{\varphi}, \mathtt{Q}_t[g]\rangle = \langle \tilde{\varphi}, g\rangle, \qquad t \geq 0.$$

On the other hand, if the chain is non-conservative, then $\lambda_{\mathtt{c}} < 0$. For this setting, the eigenvector $\tilde{\varphi}$ is called the *quasi-stationary distribution*. To see the meaning of this choice of terminology, note from the first term in (2.22) that

$$\mathrm{e}^{-\lambda_{\mathtt{c}} t}\mathbf{P}_i(t < \zeta) \sim \varphi(i), \qquad i \in E.$$

Combining it again with (2.22), we see that

$$\lim_{t\to\infty} \mathbf{P}_i(\xi_t = j | t < \zeta) = \tilde{\varphi}(j) \qquad i, j \in E. \tag{2.26}$$

Given the Perron–Frobenius asymptotic representation of the operator family $(\mathtt{Q}_t, t \geq 0)$ in this rather special setting, it is natural to ask if a similar result is true in general. The class of Markov processes is vast, and hence, one should expect a number of relatively stringent conditions in order to replicate the setting of finite-state Markov chains.

First let us introduce a little more notation. In the same sense as (2.13), we can think of $\mathtt{Q}_t$ as an integral with respect to a transition measure.[5] Indeed,

$$\mathtt{Q}_t[g](x) = \int_E g(x) p_t(x, \mathrm{d}y), \qquad g \in B(E),$$

where

$$p_t(x, \mathrm{d}y) = \mathbb{P}_x(\xi_t \in \mathrm{d}y, t < \zeta), \qquad x, y \in E.$$

To give a natural analytical home for $(\mathtt{Q}_s, s \geq 0)$ and other measures that we will encounter, we define the space of finite measures, $\mathcal{M}_f(E)$ say, so that for $g \in B(E)$ and $\mu \in \mathcal{M}_f(E)$,

$$\mu[f] = \int_E f(y)\mu(\mathrm{d}y). \tag{2.27}$$

We may now comfortably identify $\mathtt{Q}_s$ as an element of $\mathcal{M}_f(E)$, noting that it has additional dependencies on $x \in E$ and $s \geq 0$.

[5] Technically speaking, $p_t(x, \mathrm{d}y)$ is called a kernel rather than a measure because of its dependency on $x \in E$.

A measure $\mu \in \mathcal{M}_f(E)$ has a density, say m, if

$$\mu[f] = \langle m, f\rangle,$$

where, now, we identify the inner product $\langle \cdot, \cdot\rangle$ in the usual way for an $L^2(E)$ Hilbert space, i.e.,

$$\langle f, g\rangle = \int_E f(x)g(x)\mathrm{d}x, \qquad f, g \in L^2(E). \tag{2.28}$$

In other words, μ is a continuous linear functional on $L^2(E)$.

In the next theorem, we assume that the Markov process $(\xi, \mathbf{P})$ is non-conservative. A typical example to think of is a Markov process that is killed when it leaves a physical domain.

Theorem 2.2 (Perron–Frobenius Semigroup Decomposition) *Suppose that $(\xi, \mathbf{P})$ is a non-conservative Markov process for which*

$$\mathbf{P}_x(t < \zeta) > 0 \textit{ and } \mathbf{P}_x(\zeta < \infty) = 1,$$

for all $t \geq 0$ and $x \in E$. In addition, suppose that there exists a probability measure ν on E such that:

(A1) **There exist** $t_0, c_1 > 0$ **such that for each** $x \in E$,

$$\mathbf{P}_x(\xi_{t_0} \in \cdot\ | t_0 < \zeta) \geq c_1 \nu(\cdot).$$

(A2) **There exists a constant** $c_2 > 0$ **such that for each** $x \in E$ **and for every** $t \geq 0$,

$$\mathbf{P}_\nu(t < \zeta) \geq c_2 \mathbf{P}_x(t < \zeta).$$

Then, there exist $\lambda_c < 0$, a probability measure, η, on E and a function $\varphi \in B^+(E)$ such that η, resp. φ, is an eigenmeasure, resp. eigenfunction, of $(\mathtt{Q}_t, t \geq 0)$ with eigenvalue $\exp(\lambda_c t)$. That is, for all $g \in B(E)$,

$$\eta[\mathtt{Q}_t[g]] = \mathrm{e}^{\lambda_c t}\eta[g] \quad \textit{and} \quad \mathtt{Q}_t[\varphi] = \mathrm{e}^{\lambda_c t}\varphi, \quad t \geq 0. \tag{2.29}$$

Moreover, there exist $C, \varepsilon > 0$ such that

$$\left\| \mathrm{e}^{-\lambda_c t}\mathbf{P}_\cdot(t < \zeta) - \varphi \right\| \leq C\mathrm{e}^{-\varepsilon t}, \qquad t \geq 0, \tag{2.30}$$

where $\|\cdot\|$ is the supremum norm over E, and

$$\sup_{g \in B(E): \|g\| \leq 1} \left\| \mathrm{e}^{-\lambda_c t}\varphi^{-1}\mathtt{Q}_t[g] - \eta[g] \right\| \leq C\mathrm{e}^{-\varepsilon t}, \qquad t \geq 0. \tag{2.31}$$

Remark 2.2 Let us make a number of observations concerning Theorem 2.2:

(1) We can think of the assumption (A1) as a mixing condition. It ensures that transition probabilities of the Markov process, on the event of survival, can be uniformly controlled from below.
(2) The assumption (A2) means that the highest non-absorption probability among all initial points in E has the same order of magnitude as the non-absorption probability starting from ν.
(3) The eigenvalue λ_c can be thought of as the generic rate of loss of probability as it becomes increasingly unlikely for ξ to survive with time.
(4) If one can show that $(\xi_t, t < \zeta)$ admits a bounded density with respect to Lebesgue measure, then the same can be said for the eigenmeasure η.

Note that φ and η in the above theorem generalise the notions of the eigenfunctions φ and $\tilde{\varphi}$, respectively, discussed at the start of the section. In the following two sections, we discuss some deeper implications of Theorem 2.2, in particular, giving a precise interpretation of the eigenfunction φ and the eigenmeasure η. The latter turns out to be what is known as a *quasi-stationary distribution*, and the former turns out to have a harmonic property leading to what is known as a Doob h-transform that conditions the Markov process $(\xi, \mathbf{P})$ to behave in an exceptional way.

2.6 Quasi-Stationarity

Let us remain under the assumptions of Theorem 2.2. Following the reasoning that leads to statement (2.26), albeit now in a general setting, we see that for $g \in B(E)$,

$$\lim_{t\to\infty} \mathbf{E}_x[g(\xi_t)|t < \zeta] = \eta[g]. \tag{2.32}$$

By dominated convergence, (2.32) implies that, for any probability distribution μ supported on E, which can be thought of as randomising the point of issue of $(\xi, \mathbf{P})$,

$$\lim_{t\to\infty} \mathbf{E}_\mu[g(\xi_t)|t < \zeta] = \eta[g], \tag{2.33}$$

where $\mathbf{P}_\mu = \int_E \mu(\mathrm{d}x)\mathbf{P}_x$. In that respect, by appropriately normalising η such that $\eta[1] = 1$, we can think of η as a *quasi-limiting distribution*.

Suppose we take $g(x) = \mathbf{P}_x(s < \zeta)$ in (2.33), for some fixed $s > 0$, then, with the help of the Markov property, starting from the right-hand side of (2.33), it reads

$$\mathbf{P}_\eta(s < \zeta) = \lim_{t\to\infty} \frac{\mathbf{E}_\mu[\mathbf{P}_x(s < \zeta)_{x=\xi_t}]}{\mathbf{P}_\mu(t < \zeta)} = \lim_{t\to\infty} \frac{\mathbf{P}_\mu(t + s < \zeta)}{\mathbf{P}_\mu(t < \zeta)}. \tag{2.34}$$

On the other hand, if we take $g(x) = \mathbf{E}_x[f(\xi_s)]$, for some $f \in B(E)$ and fix $s > 0$, then a similar calculation shows us that

$$
\begin{aligned}
\mathbf{E}_\eta[f(\xi_s)] &= \lim_{t\to\infty} \frac{\mathbf{E}_\mu[\mathbf{E}_y[f(\xi_s)]_{y=\xi_t}]}{\mathbf{P}_\mu(t<\zeta)} \\
&= \lim_{t\to\infty} \frac{\mathbf{E}_\mu[f(\xi_{t+s})]}{\mathbf{P}_\mu(t<\zeta)} \\
&= \lim_{t\to\infty} \frac{\mathbf{E}_\mu[f(\xi_{t+s})]}{\mathbf{P}_\mu(t+s<\zeta)} \frac{\mathbf{P}_\mu(t+s<\zeta)}{\mathbf{P}_\mu(t<\zeta)} \\
&= \lim_{t\to\infty} \mathbf{E}_\mu[f(\xi_{t+s})|t+s<\zeta] \frac{\mathbf{P}_\mu(t+s<\zeta)}{\mathbf{P}_\mu(t<\zeta)} \\
&= \eta[f]\mathbf{P}_\eta(s<\zeta),
\end{aligned}
$$

where we have used (2.34). In conclusion, this tells us that our quasi-limiting distribution is also a *quasi-stationary distribution*, that is,

$$
\mathbf{E}_\eta[f(\xi_s)|s<\zeta] = \eta[f], \qquad s \geq 0.
$$

Said another way, when the process is issued from a randomised state using η, then conditional on survival, at any fixed time later, its position is still distributed according to η. From this perspective, it is not surprising that, when issuing the Markov process from its quasi-stationary distribution, its lifetime is exactly exponentially distributed, that is,

$$
\mathbf{P}_\eta(\zeta > t) = \mathrm{e}^{\lambda_c t}.
$$

This is a fact that is easily derived from (2.29).

While it is easy to see that a quasi-stationary distribution satisfies (2.33) with $\mu = \eta$, we note that the concepts of quasi-stationary distributions and quasi-limiting distributions are actually equivalent in the context of Theorem 2.2.

On a final note, we mention that a variant of (2.33) is the situation for which there exists a time-dependent sequence $a(t)$ and a probability measure ν such that

$$
\lim_{t\to\infty} \mathbf{P}_\mu[g(\xi_t/a(t))|t<\zeta] = \nu[g],
$$

for all $g \in B(E)$ and $\mu \in \mathcal{M}_f(E)$. When this happens, the measure ν is called the *Yaglom* limit. We will encounter this concept later in this book as well.

2.7 Martingales, Doob h-Transforms, and Conditioning

In the previous section, we have examined the probabilistic meaning and functionality of the left eigenmeasure η. Here we do the same for the right eigenfunction. Before doing so, we remind the reader of the notion of a martingale.

Definition 2.3 A stochastic process $(M_t, t \geq 0)$ is a martingale with respect to the filtration $(\mathfrak{G}_t, t \geq 0)$ if:

(i) For each $t \geq 0$, M_t is $\mathfrak{G}_t$-measurable.
(ii) For all $t \geq 0$, $\mathbf{E}[|M_t|] < \infty$.
(iii) For all $s, t \geq 0$, $\mathbf{E}[M_{t+s}|\mathfrak{G}_s] = M_t$.

Lemma 2.2 *Suppose the conclusion of Theorem 2.2 holds. Then the process*

$$M_t^{\mathrm{c}} := \mathrm{e}^{-\lambda_{\mathrm{c}} t}\varphi(\xi_t), \qquad t \geq 0,$$

is a martingale with respect to $\mathfrak{G}$ *for* $(\xi, \mathbf{P}_x)$.

Proof Clearly, condition (i) of Definition 2.3 is satisfied. Next note that $\mathbf{E}_x|M_t^{\mathrm{c}}| < \infty$ for $t \geq 0$, $x \in E$ (which is obvious from the boundedness of φ) and, for $s, t \geq 0$, the Markov property and the second equality in (2.29) tell us that

$$\begin{aligned}
\mathbf{E}_x[M_{t+s}^{\mathrm{c}}|\mathfrak{G}_t] &= \mathrm{e}^{-\lambda_{\mathrm{c}}(t+s)}\mathbf{E}_x\,[\varphi(\xi_{t+s})|\mathfrak{G}_t] \\
&= \mathrm{e}^{-\lambda_{\mathrm{c}}(t+s)}\mathbf{E}_y\,[\varphi(\xi_s)]_{y=\xi_t} \\
&= \mathrm{e}^{-\lambda_{\mathrm{c}} t}\mathrm{e}^{-\lambda_{\mathrm{c}} s}\mathtt{Q}_s[\varphi](\xi_t) \\
&= \mathrm{e}^{-\lambda_{\mathrm{c}} t}\varphi(\xi_t) \\
&= M_t^{\mathrm{c}},
\end{aligned}$$

as required.

One of the fundamental properties of martingales is that they maintain constant expectation. Specifically,

$$\mathbf{E}_x[M_{t+s}^{\mathrm{c}}] = \mathbf{E}_x[\mathbf{E}_x[M_{t+s}^{\mathrm{c}}|\mathfrak{G}_t]] = \mathbf{E}_x[M_t^{\mathrm{c}}], \qquad s, t \geq 0,$$

and hence, since $M_0^{\mathrm{c}} = \varphi(x)$, it follows that

$$\mathbf{E}_x[M_t^{\mathrm{c}}] = \varphi(x), \qquad x \in E,\ t \geq 0.$$

While martingales have many uses, non-negative martingales that are defined as functionals of a Markov process find a natural home in defining the so-called *Doob h-transforms*. This is a way of tilting the law of a Markov process $(\xi, \mathbf{P})$ in a way that keeps the resulting object still within the class of Markov processes.

We can define a new family of measures $\mathbf{P}^{\mathrm{c}} = (\mathbf{P}_x^{\mathrm{c}}, x \in E)$ on $(E_\dagger^{[0,\infty]}, \mathfrak{G})$ by the relation

$$\mathbf{P}_x^{\mathrm{c}}(A) = \mathbf{E}_x\left[\mathbf{1}_A \mathrm{e}^{-\lambda_c t}\frac{\varphi(\xi_t)}{\varphi(x)}\right], \qquad x \in E,\ t \geq 0,\ A \in \mathfrak{G}_t. \tag{2.35}$$

A shorthand way of writing this change of measure is

$$\left.\frac{\mathrm{d}\mathbf{P}_x^{\mathrm{C}}}{\mathrm{d}\mathbf{P}_x}\right|_{\mathfrak{G}_t} = M_t^{\mathrm{C}}, \qquad t \geq 0,\ x \in E. \tag{2.36}$$

It is worth noting that any event $A \in \mathfrak{G}_t$ also belongs to $\mathfrak{G}_{t+s}$ for any $s \geq 0$. As such, the definitions (2.35) and (2.36) may appear to be ambiguous. This is where the martingale property plays an important role to circumvent this potential problem since, using standard properties of expectations, we have

$$\mathbf{E}_x\left[\mathbf{1}_A M_{t+s}^{\mathrm{C}}\right] = \mathbf{E}_x\left[\mathbf{1}_A \mathbf{E}_x\left[\left. M_{t+s}^{\mathrm{C}}\right| \mathfrak{G}_t\right]\right] = \mathbf{E}_x\left[\mathbf{1}_A M_t^{\mathrm{C}}\right]. \tag{2.37}$$

Lemma 2.3 *The process* $(\xi, \mathbf{P}^{\mathrm{C}})$ *is a Markov process. Moreover, it is a strong Markov process if* $(\xi, \mathbf{P})$ *is.*

Proof The Markov property is easily verified as, for $f \in B(E)$ and $s, t \geq 0$,

$$\begin{aligned}
\mathbf{E}_x^{\mathrm{C}}\left[f(\xi_{t+s})|\mathfrak{G}_t\right] &= \mathbf{E}_x\left[\left. f(\xi_{t+s})\frac{\varphi(\xi_{t+s})}{\varphi(x)}\right| \mathfrak{G}_t\right] \\
&= \mathbf{E}_y\left[f(\xi_s)\frac{\varphi(\xi_s)}{\varphi(x)}\right]_{y=\xi_s} \\
&= \mathbf{E}_y^{\mathrm{C}}[f(\xi_s)]_{y=\xi_s}.
\end{aligned}$$

What is also clear from this calculation is that it passes through verbatim in the setting that t is a stopping time, providing the strong Markov property is available for the process $(\xi, \mathbf{P})$.

The new Markov process $(\xi, \mathbf{P}^{\mathrm{C}})$ is not just an abstract phenomenon. It has a genuine meaning connected to conditioning, albeit in a different way to quasi-stationarity.

Assuming it exists, consider the limit

$$\mathtt{P}_t^{\mathrm{C}}[g](x) := \lim_{s\to\infty} \mathbf{E}_x[g(\xi_t)|\zeta > t+s], \qquad x \in E,\ g \in B(E),\ t \geq 0. \tag{2.38}$$

The limit (2.38) is tantamount to conditioning the Markov process $(\xi, \mathbf{P})$ at each finite time to survive for an arbitrary amount of time in the future. The next theorem shows the connection between this conditioned process and $(\xi, \mathbf{P}^{\mathrm{C}})$, defined by the martingale change of measure (2.36).

Lemma 2.4 *Suppose the conclusion of Theorem 2.2 holds. The family of operators* $(\mathtt{P}_t^{\mathrm{C}}, t \geq 0)$ *is well defined and equal to the expectation semigroup of* $(\xi, \mathbf{P}^{\mathrm{C}})$. *That is to say,*

$$\mathtt{P}_t^{\mathrm{C}}[g](x) = \mathbf{E}_x^{\mathrm{C}}[g(\xi_t)], \qquad x \in E, t \geq 0, g \in B^+(E).$$

Proof We can use Bayes' expression for conditional expectation and write, for $s, t \geq 0$, $g \in B(E)$, $x \in E$,

$$\begin{aligned}\mathbf{E}_x[g(\xi_t)|\zeta > t+s] &= \frac{\mathbf{E}_x[g(\xi_t)]}{\mathbf{P}_x(\zeta > t+s)} \\ &= \mathbf{E}_x\left[g(\xi_t)\frac{\mathbf{P}_y(\zeta > s)|_{y=\xi_t}}{\mathbf{P}_x(\zeta > t+s)}\right],\end{aligned} \tag{2.39}$$

where we recall that g returns the value 0 on $\{\dagger\}$. It follows from Theorem 2.2 and Fatou's Lemma that, for $g \in B^+(E)$, then

$$\liminf_{s\to\infty} \mathbf{E}_x[g(\xi_t)|\zeta > t+s] \geq \mathbf{E}_x\left[g(\xi_t)\mathrm{e}^{-\lambda_c t}\frac{\varphi(\xi_t)}{\varphi(x)}\right] = \mathbf{E}^c_x[g(\xi_t)]. \tag{2.40}$$

On the other hand, suppose we write $h(x) = \bar{g} - g(x)$, for $x \in E$, where $\bar{g} = \sup_{y\in E} g(y)$. Recalling Lemma 2.2, by applying (2.40) to h, which is necessarily in $B^+(E)$, we discover that

$$\begin{aligned}\liminf_{s\to\infty}\{\bar{g} - \mathbf{E}_x[g(\xi_t)|\zeta > t+s]\} &\geq \mathbf{E}_x\left[\{\bar{g} - g(\xi_t)\}\mathrm{e}^{-\lambda_c t}\frac{\varphi(\xi_t)}{\varphi(x)}\right] \\ &= \bar{g} - \mathbf{E}^c_x[g(\xi_t)].\end{aligned}$$

Rearranging the above gives us

$$\limsup_{s\to\infty} \mathbf{E}_x[g(\xi_t)|\zeta > t+s] \leq \mathbf{E}^c_x[g(\xi_t)]. \tag{2.41}$$

Putting (2.40) and (2.41) together gives us the result.

To conclude this section, let us make a few remarks concerning the generator of the resulting process $(\xi, \mathbf{P}^c)$. Suppose to this end, we write $\mathtt{L}^c$ as the generator of $(\xi, \mathbf{P}^c)$. In keeping with the definition (2.10) and the heuristic style of reasoning, we have that, for a suitable class of $f \in B(E)$,

$$\mathtt{L}^c f(x) = \lim_{t\to 0}\frac{\mathrm{d}}{\mathrm{d}t}\mathtt{P}^c_t[f](x) = \lim_{h\to 0}\frac{\varphi^{-1}\mathrm{e}^{-\lambda_c h}\mathtt{P}_h[\varphi f](x) - f(x)}{h}, \qquad x \in E. \tag{2.42}$$

As $\mathtt{L}$ is a linear operator and φ is an eigenfunction of $\mathtt{L}$, this reduces more simply to

$$\begin{aligned}\mathtt{L}^c f(x) &= \lim_{h\to 0}\frac{\varphi^{-1}\mathtt{P}_h[\varphi f](x) - f(x)}{h} + \lim_{h\to 0}\frac{\varphi^{-1}(\mathrm{e}^{-\lambda_c h} - 1)\mathtt{P}_h[\varphi f](x)}{h} \\ &= \frac{\mathtt{L}[\varphi f](x)}{\varphi(x)} - \lambda_c f(x), \qquad x \in E.\end{aligned} \tag{2.43}$$

This observation will be particularly useful in the next section when we discuss two specific examples.

2.8 Two Instructive Examples of Doob h-Transforms

We offer two examples in this section. One that deals with the setting of a finite-state space, and hence necessarily $(\xi, \mathbf{P})$, is a Markov chain, and the second an uncountable state space, for which we have chosen a diffusive process.

Markov Chains with Absorption Let us return to the setting of a Markov chain on $E = \{1, \ldots, n\}$ with an additional cemetery state $\{\dagger\}$, as described in Sect. 2.5. Assuming irreducibility on E and that $\{\dagger\}$ is accessible from at least one of the states in E, we recall that (2.22) holds with $\lambda_c < 0$ and the intensity matrix $\mathtt{L}$ satisfies $(\mathtt{L}\mathbf{1})_i < 0$ for at least one $i \in E$.

Let us now consider the effect of the change of measure (2.36). We can try to understand the effect of this change of measure via the generator $\mathtt{L}^{\mathrm{c}}$. From (2.43), we note that for any vector $f \in B(\{1, \cdots, n\})$,

$$
\begin{aligned}
L^{\mathrm{c}}[f](i) &= \frac{\sum_{j=1}^{n} \mathtt{L}(i,j)\varphi(j)f(j)}{\varphi(i)} - \lambda_c f(i) \\
&= \frac{\sum_{j=1}^{n} \varphi(j)\big(\mathtt{L}(i,j) - \lambda_c \delta(i,j)\big) f(j)}{\varphi(i)},
\end{aligned}
\tag{2.44}
$$

where $\delta(i, j)$ is the identity matrix. In short, the new transition matrix of $(\xi, \mathbf{P}^{\mathrm{c}})$ is given by the $n \times n$ matrix

$$
\mathtt{L}^{\mathrm{c}}(i,j) = \frac{\varphi(j)}{\varphi(i)}\big(\mathtt{L}(i,j) - \lambda_c \delta(i,j)\big), \qquad i, j \in \{1, \cdots, n\}.
$$

We can thus see that the effect of the Doob h-transform on the transition semigroup passes through to a natural analogue of the Doob h-transform on the intensity matrix. Writing $\mathtt{1} = (1, \cdots, 1)$ for the vector of ones, it is also worthy of note that, appealing to the fact that φ is an eigenvector,

$$
\mathtt{L}\mathtt{1} = \frac{\sum_{j=1}^{n} \varphi(j)\big(\mathtt{L}(i,j) - \lambda_c \delta(i,j)\big)}{\varphi(i)} = \frac{\mathtt{L}\varphi(i) - \lambda_c \varphi(i)}{\varphi(i)} = 0.
$$

This means that $(\xi, \mathbf{P}^{\mathrm{c}})$ is a conservative process; in other words, there is no loss of probability from transitions to $\{\dagger\}$. Said another way, $(\xi, \mathbf{P}^{\mathrm{c}})$ is almost surely forbidden from entering the state $\{\dagger\}$ and, in this sense, we may think of it as the original Markov chain conditioned to avoid absorption in $\{\dagger\}$.

Brownian Motion in a Compact Domain Suppose that D is a bounded open connected[6] set in $\mathbb{R}^d$ that has positive d-dimensional Lebesgue measure (volume). Let us write ∂D for the boundary of D and we will assume the following smoothness on ∂D. For each $y \in \partial D$, there exists an *exterior radius* $r_y^{(e)}$ and *interior radius* $r_y^{(i)}$ such that a ball, say $\mathbb{B}_y^{(e)}$, of radius $r_y^{(e)}$ can be placed exterior to D and a ball, say $\mathbb{B}_y^{(i)}$, of radius $r_y^{(i)}$ can be placed interior to D so that $\mathbb{B}_y^{(e)} \cap \mathbb{B}_y^{(i)} \cap \partial D = \{y\}$.

We are interested in the setting that ξ is a Brownian motion until it first exits D. That is to say, our Markov process has the infinitesimal generator

$$\mathtt{L}f(x) = \frac{1}{2}\nabla^2 f(x), \qquad x \in D, \tag{2.45}$$

and, for the sake of clarity, its semigroup is given by

$$\mathtt{P}_t[f](x) = \mathbf{E}_x[f(\xi_t)\mathbf{1}_{(t<\tau^D)}], \qquad x \in D,\ t \geq 0,$$

where $\tau^D = \inf\{t > 0 : \xi_t \notin D\}$. We may think of ∂D as the cemetery state $\{\dagger\}$ and τ^D as the lifetime ζ of the process.

In this setting, it turns out that all the eigenvalues, in the sense of (2.24), are real, satisfying

$$\cdots \leq \lambda_4 \leq \lambda_3 \leq \lambda_2 < \lambda_1 =: \lambda_{\mathtt{c}} < 0,$$

such that $\lim_{k\to\infty} \lambda_k = -\infty$. The associated right eigenfunctions φ_k, $k \geq 0$, are such that, up to an appropriate normalisation, they form an orthonormal basis in $L^2(D)$ (with respect to Lebesgue measure). Moreover, $\varphi := \varphi_1$ is strictly positive in D and continuous up to and including the boundary where it is zero-valued. It turns out that associated to each eigenvalue is a left eigenfunction that, for each $k = 1, 2, \ldots$, is identically equal to φ_k.

As a consequence of these facts, it turns out that

$$\mathtt{P}_t[f](x) = \sum_{k\geq 1} \mathrm{e}^{-\lambda_k t}\varphi_k(x)\langle f, \varphi_k\rangle, \qquad f \in B(D). \tag{2.46}$$

In addition, using means other than verifying (A1) and (A2), classical theory tells us that (2.31) and (2.30) automatically hold. In that case, it is at least heuristically obvious from (2.46) that the rate of decay ε in (2.31) and (2.30) can be precisely identified as equal to $\lambda_{\mathtt{c}} - \lambda_2$.

Since the so-called ground state, φ, is zero-valued on ∂D and can be shown to be a continuous function, we can see that the change of measure (2.35) is such that those paths that stray too close to the boundary are penalised. We can also see this

[6] In this setting, a connected set simply means that, if $x, y \in D$, then there is a path joining x to y that remains entirely in D.

via (2.43). Remembering that φ is an eigenfunction of (2.45), for twice continuously differentiable functions on D, we have

$$
\begin{aligned}
\mathtt{L}^{\mathrm{c}} f(x) &= \frac{1}{2\varphi(x)}\nabla^2\big(\varphi(x)f(x)\big) - \lambda_{\mathrm{c}} f(x)\\
&= \frac{1}{2}\nabla^2 f(x) + \left(\frac{\nabla\varphi}{\varphi}\right)\cdot\nabla f(x) + \frac{f(x)}{2\varphi(x)}\nabla^2\varphi(x) - \lambda_{\mathrm{c}} f(x)\\
&= \frac{1}{2}\nabla^2 f(x) + \left(\frac{\nabla\varphi}{\varphi}\right)\cdot\nabla f(x), \qquad x\in E.
\end{aligned}
$$

The generator $\mathtt{L}^{\mathrm{c}}$ is now identifiable as that of a Brownian motion with a drift given by $\varphi^{-1}\nabla\varphi$. More precisely, we can identify the Markov process $(\xi, \mathbf{P}^{\mathrm{c}})$ as equal in law to the solution to the SDE

$$
\xi_t^{\mathrm{c}} = X_t + \int_0^t \frac{\nabla\varphi(\xi_s^{\mathrm{c}})}{\varphi(\xi_s^{\mathrm{c}})}\mathrm{d}s, \qquad t\geq 0.
$$

With the above representation, we now see the ground state property that $\varphi = 0$ on ∂D means that the speed of drift is dominated by φ^{-1}, which becomes arbitrarily large in value as the process ξ^{c} approaches the boundary. Moreover, the direction of the drift is given by $\nabla\varphi$, which is always pointing inwards to the interior of D (because φ decreases continuously to zero on ∂D). In other words, the process $(\xi, \mathbf{P}^{\mathrm{c}})$ feels an increasingly large repulsive force from the boundary the closer it gets. Indeed, one can show that, under the change of measure (2.36), it never touches the boundary. We thus see that the effect of the Doob h-transform is to condition the original process to remain in the domain D eternally.

2.9 Comments

Starting with Markov himself in his initial work at the turn of the twentieth century, the general theory of Markov processes has a long history but owes a lot of its measure-theoretic formality to Kolmogorov, Feller, and Dynkin. Key improvements and summaries of the formalised theory of Markov processes were given in the highly influential texts of Dynkin [50], Blumenthal and Getoor [8], Dellacherie and Meyer [37–40], Rogers and Williams [116, 117], Chung and Walsh [25], among others. Most of what appears in this chapter can be found in these texts.

In Sect. 2.1, for a very general setting, it is not uncommon for the event space $E_\dagger^{[0,\infty]}$ to be replaced by the subspace $\mathbb{D}(E)$ of coordinate maps that are right continuous with left limits in E. In that case, the filtration $(\mathfrak{G}_t, t\geq 0)$ to be generated by open subsets of $\mathbb{D}(E)$. In order to specify these open sets, a topology is needed. The generally accepted fit-for-purpose topology in this respect is the Skorokhod J_1-topology; cf. Chapter VI of [79]. Theorem 2.1 is a general result that can be found in a slightly less general form in Lemma 1.2, Chapter 4 of Dynkin [49]. The heuristic discussion of generators in Sect. 2.3 can be replaced with a more formal treatment (as alluded to at the end of that section). The reader is referred to Ethier and Kurtz [57] for one of many references. Perron–Frobenius behaviour for Markov chains discussed in Sect. 2.5 is nicely summarised in the book of Seneta [119]. Quasi-stationarity, discussed in Sect. 2.6, is a theme that has run through probability theory as a folklore rather than a formalised theory for decades. The book of Collet et al. [26] presents one of the few treatments in the form of a monograph. We also refer the reader to the works of Champagnat and Villemonais [21, 22] where the assumptions (A1) and (A2) and their implications were introduced. The general theory of Doob h-transforms alluded to in Sect. 2.7 is another folklore that does not necessarily enjoy a complete single point of reference. One may refer to Chapter X of Doob [42] or Chung and Walsh [25], for example, for further insight. See also Bliedtner and Hansen [18]. The examples of Doob h-transforms in Sect. 2.8 are classical; see, for example, Powell [112] for a recent exposition on the Brownian setting and, more specifically, [35, Theorem 1.6.8] for spectral properties of the Laplacian.

Chapter 3
Stochastic Representation of the Neutron Transport Equation

In this chapter, we break away from the classical view of the NTE described in Chap. 1 and begin our journey into stochastic representation of its solutions. The main objective of this chapter is to look at alternative interpretations of solutions to the NTE in terms of averaging over paths of neutrons. More precisely, we look at the connection to two families of stochastic processes that underly different Feynman–Kac representations of NTE solutions. This sets the scene for the remainder of the first part of this book that delves into a detailed analysis of the path properties of these stochastic processes and how this embodies the physical process of fission as much as it encapsulates the behaviour of solutions.

3.1 Duality and the Backward NTE

Recall the standard setup from Chap. 1, where we defined $(\Psi_t, t \geq 0)$ as the time evolution of neutron density, which, under the assumption (H1), was described by the NTE (with no source term) as an ACP in $L^2(D \times V)$,

$$\begin{aligned}\frac{\partial}{\partial t}\Psi_t(r, \upsilon) = &-\upsilon \cdot \nabla_r \Psi_t(r, \upsilon) - \sigma(r, \upsilon)\Psi_t(r, \upsilon)\\ &+ \int_V \Psi_t(r, \upsilon')\sigma_{\mathtt{s}}(r, \upsilon')\pi_{\mathtt{s}}(r, \upsilon', \upsilon)\mathrm{d}\upsilon'\\ &+ \int_V \Psi_t(r, \upsilon')\sigma_{\mathtt{f}}(r, \upsilon')\pi_{\mathtt{f}}(r, \upsilon', \upsilon)\mathrm{d}\upsilon'. \end{aligned} \tag{3.1}$$

Moreover, we have the additional boundary conditions

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8_3

$$\begin{cases} \Psi_0(r, \upsilon) = f(r, \upsilon) & \text{for } r \in D, \upsilon \in V, \\ \Psi_t(r, \upsilon) = 0 & \text{for } t \geq 0 \text{ and } r \in \partial D \text{ if } \upsilon \cdot \mathbf{n}_r < 0, \end{cases} \tag{3.2}$$

where $\mathbf{n}_r$ is the outward facing normal at $r \in \partial D$ and $f \in L^2(D \times V)$ that belongs to the domain of $\mathscr{G} := \mathscr{T} + \mathscr{S} + \mathscr{F}$. From (1.10), we also understand the solution in $\mathtt{Dom}(\mathscr{G})$ as an orbit of the linear operator $\exp(t\mathscr{G})$. It is natural to wonder what the meaning of the "dual" of this solution looks like on $L^2(D \times V)$.

As a first step in this direction, let us first examine what we mean by the dual of the operator $\mathscr{G}$. Suppose that $\mathscr{O}$ is an operator mapping $L^2(D \times V)$ to itself. We would like to know if there is an operator $\hat{\mathscr{O}}$ such that, for all $g \in \mathtt{Dom}(\hat{\mathscr{O}})$ and $f \in \mathtt{Dom}(\mathscr{O})$,

$$\langle f, \mathscr{O} g\rangle = \langle \hat{\mathscr{O}} f, g\rangle. \tag{3.3}$$

We will first study what the duals of the individual operators $\mathscr{T}$, $\mathscr{S}$, and $\mathscr{F}$ look like.

To this end, let us begin with the operator $\mathscr{T} = -\upsilon \cdot \nabla_r - \sigma$. Suppose momentarily that $f, g \in \mathtt{Dom}(\mathscr{T})$, that is $f, g \in L^2(D \times V)$ such that both $\upsilon \cdot \nabla_r f$ and $\upsilon \cdot \nabla_r g$ are well defined as distributional derivatives in $L^2(D \times V)$. We can verify by a simple integration by parts that, for $\upsilon \in V$,

$$\langle f, \upsilon \cdot \nabla_r g\rangle = \int_{\partial(D\times V)} (\upsilon \cdot \upsilon') f(r, \upsilon') g(r, \upsilon') \mathrm{d}r \mathrm{d}\upsilon' - \langle \upsilon \cdot \nabla_r f, g\rangle. \tag{3.4}$$

Hence, if g respects the second of the boundary conditions in (3.2), and we additionally insist that f respects the condition

$$f(r, \upsilon) = 0 \text{ for } r \in \partial D \text{ if } \upsilon \cdot \mathbf{n}_r > 0, \tag{3.5}$$

then (3.4) reduces to the nicer duality relation

$$\langle f, \upsilon \cdot \nabla_r g\rangle = -\langle \upsilon \cdot \nabla_r f, g\rangle.$$

In short, recalling the definitions in (1.8), the transport operator $\mathscr{T}$ is dual to $\mathtt{T} - \sigma$ where

$$\mathtt{T} f(r, \upsilon) := \upsilon \cdot \nabla_r f(r, \upsilon), \tag{3.6}$$

for

$$f \in \mathtt{Dom}(\mathtt{T}) := \{f \in L^2(D \times V) \text{ such that } \upsilon \cdot f \in L^2(D \times V) \text{ and } f|_{\partial(D\times V)^+=0}\},$$

with

$$\partial(D \times V)^+ := \{(r, \upsilon) \in D \times V \text{ such that } r \in \partial D \text{ and } \upsilon \cdot \mathbf{n}_r > 0\}.$$

We can similarly consider the duals of the operators $\mathscr{S}$ and $\mathscr{F}$, which were given in (1.8). To this end, note that Fubini's theorem tells us that, for $f, g \in L^2(D \times V)$,

$$\begin{aligned}
&\langle f, \int_V g(\cdot, \upsilon')\sigma_{\mathtt{s}}(\cdot, \upsilon')\pi_{\mathtt{s}}(\cdot, \upsilon', \cdot)\mathrm{d}\upsilon'\rangle \\
&\qquad = \int_{D\times V\times V} f(r, \upsilon)\sigma_{\mathtt{s}}(r, \upsilon')g(r, \upsilon')\pi_{\mathtt{s}}(r, \upsilon', \upsilon)\mathrm{d}\upsilon'\mathrm{d}r\mathrm{d}\upsilon \\
&\qquad = \int_{D\times V} \sigma_{\mathtt{s}}(r, \upsilon') \int_V f(r, \upsilon)\pi_{\mathtt{s}}(r, \upsilon', \upsilon)\mathrm{d}\upsilon\, g(r, \upsilon')\mathrm{d}r\mathrm{d}\upsilon' \\
&\qquad = \langle \sigma_{\mathtt{s}}(\cdot, \cdot) \int_V f(\cdot, \upsilon)\pi_{\mathtt{s}}(\cdot, \cdot, \upsilon)\mathrm{d}\upsilon, g\rangle.
\end{aligned}$$

Replacing $\sigma_{\mathtt{s}}$ and $\pi_{\mathtt{s}}$ by $\sigma_{\mathtt{f}}$ and $\pi_{\mathtt{f}}$, respectively, yields the duality relation for $\mathscr{F}$.

Let us now summarise the *backward* transport, scattering, and fission operators on $L^2(D \times V)$ together as

$$\begin{aligned}
&\mathtt{T}f(r, \upsilon) := \upsilon \cdot \nabla_r f(r, \upsilon) \\
&\mathtt{S}f(r, \upsilon) := \sigma_{\mathtt{s}}(r, \upsilon) \textstyle\int_V f(r, \upsilon')\pi_{\mathtt{s}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon' - \sigma_{\mathtt{s}}(r, \upsilon) f(r, \upsilon) \\
&\mathtt{F}f(r, \upsilon) := \sigma_{\mathtt{f}}(r, \upsilon) \textstyle\int_V f(r, \upsilon')\pi_{\mathtt{f}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon' - \sigma_{\mathtt{f}}(r, \upsilon) f(r, \upsilon).
\end{aligned} \tag{3.7}$$

As in Chap. 1, we have $\mathtt{Dom}(\mathtt{T} + \mathtt{S} + \mathtt{F}) = \mathtt{Dom}(\mathtt{T})$. The reader will immediately note that although the sum $\mathtt{T} + \mathtt{S} + \mathtt{F}$ is the dual of the sum $\mathscr{T} + \mathscr{S} + \mathscr{F}$, the same cannot be said for the individual operators "$\mathtt{T}$", "$\mathtt{S}$", and "$\mathtt{F}$". That is to say, the way we have grouped the terms does not allow us to say that $\mathtt{T}$ is the adjoint operator to $\mathscr{T}$ and so on. The reason for this is because of the way that we will shortly make the association of the operators $\mathtt{T}$, $\mathtt{S}$, $\mathtt{F}$ to certain stochastic processes. Nonetheless, we have the following duality relation between $\mathtt{T} + \mathtt{S} + \mathtt{F}$ and $\mathscr{T} + \mathscr{S} + \mathscr{F}$, which follows from the calculations above.

Theorem 3.1 *Assume* $f \in \mathtt{Dom}(\mathtt{T} + \mathtt{S} + \mathtt{F})$, $g \in \mathtt{Dom}(\mathscr{T} + \mathscr{S} + \mathscr{F})$. *Then we have*

$$\langle f, (\mathscr{T} + \mathscr{S} + \mathscr{F})g\rangle = \langle(\mathtt{T} + \mathtt{S} + \mathtt{F})f, g\rangle. \tag{3.8}$$

Theorem 3.1 also alludes to the possibility that the dual of the operator $\exp(t\mathscr{G})$ takes the form $\exp(t\hat{\mathscr{G}})$, where $\hat{\mathscr{G}} := (\mathtt{T} + \mathtt{S} + \mathtt{F})$. Moreover, since $(\Psi_t, t \geq 0)$

solves a linear equation (3.1) in $L^2(D \times V)$, then there is also a dual—also known as the backward—NTE on $L^2(D \times V)$ taking the form

$$\frac{\partial}{\partial t}\hat{\Psi}_t(r, \upsilon) = \upsilon \cdot \nabla_r \hat{\Psi}_t(r, \upsilon) + \sigma_{\mathtt{s}}(r, \upsilon) \int_V \{\hat{\Psi}_t(r, \upsilon')\pi_{\mathtt{s}}(r, \upsilon, \upsilon') - \hat{\Psi}_t(r, \upsilon)\}\mathrm{d}\upsilon'$$
$$+ \sigma_{\mathtt{f}}(r, \upsilon) \int_V \hat{\Psi}_t(r, \upsilon')\pi_{\mathtt{f}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon' - \sigma_{\mathtt{f}}(r, \upsilon)\hat{\Psi}_t(r, \upsilon), \tag{3.9}$$

with physical boundary conditions

$$\begin{cases} \hat{\Psi}_0(r, \upsilon) = g(r, \upsilon) & \text{for } r \in D, \upsilon \in V, \\ \hat{\Psi}_t(r, \upsilon) = 0 & \text{for } t \geq 0 \text{ and } r \in \partial D \text{ if } \upsilon \cdot \mathbf{n}_r > 0. \end{cases} \tag{3.10}$$

Indeed, in a similar style to Chap. 1, we should think of (3.9)–(3.10) as an ACP, which has a unique solution when considered on $\mathtt{Dom}(\mathtt{T} + \mathtt{S} + \mathtt{F}) = \mathtt{Dom}(\mathtt{T})$, identified as the orbit

$$\hat{\Psi}_t = \exp(t(\mathtt{T} + \mathtt{S} + \mathtt{F}))g, \qquad t \geq 0.$$

The following result formalises the above discussion.

Corollary 3.1 *The solution to* (3.1) *with boundary conditions* (3.2)*, specifically with* $\Psi_0 = f \in \mathtt{Dom}(\mathscr{T} + \mathscr{S} + \mathscr{F})$*, and the solution to* (3.9) *with boundary conditions* (3.10)*, specifically with* $\hat{\Psi}_0 = g \in \mathtt{Dom}(\mathtt{T} + \mathtt{S} + \mathtt{F})$*, are dual in the sense that*

$$\langle g, \Psi_t \rangle = \langle \hat{\Psi}_t, f \rangle, \qquad t \geq 0. \tag{3.11}$$

Proof Since Ψ_t is a solution to (3.1) on $L^2(D \times V)$, we have

$$\frac{\partial}{\partial t}\langle g, \Psi_t \rangle = \langle g, (\mathscr{T} + \mathscr{S} + \mathscr{F})\Psi_t \rangle.$$

Similarly, we have

$$\frac{\partial}{\partial t}\langle \hat{\Psi}_t, f \rangle = \langle (\mathtt{T} + \mathtt{S} + \mathtt{F})\hat{\Psi}_t, f \rangle.$$

Since $f \in \mathtt{Dom}(\mathtt{T} + \mathtt{S} + \mathtt{F})$ and $g \in \mathtt{Dom}(\mathscr{T} + \mathscr{S} + \mathscr{F})$, it follows that $\hat{\Psi}_t \in \mathtt{Dom}(\mathtt{T} + \mathtt{S} + \mathtt{F})$ and $\Psi \in \mathtt{Dom}(\mathscr{T} + \mathscr{S} + \mathscr{F})$. Thus, Theorem 3.1 can be applied, and we have, for $t \geq 0$,

$$\frac{\partial}{\partial t}\langle g, \Psi_t \rangle = \langle g, (\mathscr{T} + \mathscr{S} + \mathscr{F})\Psi_t \rangle = \langle (\mathtt{T} + \mathtt{S} + \mathtt{F})\hat{\Psi}_t, f \rangle = \frac{\partial}{\partial t}\langle \hat{\Psi}_t, f \rangle.$$

In other words, for $t \geq 0$,

$$\langle g, \Psi_t \rangle = \langle g, f \rangle + \int_0^t \frac{\partial}{\partial s} \langle g, \Psi_s \rangle \mathrm{d}s = \langle g, f \rangle + \int_0^t \frac{\partial}{\partial s} \langle \hat{\Psi}_s, f \rangle \mathrm{d}s = \langle \hat{\Psi}_t, f \rangle.$$

The result now follows. □

Remark 3.1 In much of the nuclear physics and engineering literature, Eq. (3.9) is referred to as the *adjoint* NTE.

It is not yet clear what added value the backward formulation of the NTE on $L^2(D \times V)$ gives us. Our next objective is to show that solutions to the backward NTE can be identified with a category of integral equations that we saw in Chap. 2, see in particular (2.6), which are also known as *mild equations*, or *Duhamel equations* in the PDE literature.

Before introducing the mild form of the backward NTE, we need to familiarise ourselves with expectation semigroups of some Markov processes that are of relevance. To this end, we will start by defining the so-called advection semigroup, which provides an alternative way to describe particle transport, as opposed to the operator $\mathtt{T}$. We will then construct two families of stochastic processes, which will form the basis of the solutions to the aforementioned mild equations. Finally, we will formally introduce the mild version of the NTE and discuss its relation to the classical integro-differential version presented until now.

3.2 Advection Transport

In free space, neutrons move in straight lines. If a particle has a velocity $\upsilon \in V$, then we can describe its motion in space by composing its position with a test function say $g \in B^+(D \times V)$, the space of non-negative and uniformly bounded functions on $D \times V$, with the additional constraint that g has value zero on a cemetery state. We will discuss the latter in more detail shortly.

If the initial position is $r \in D$, then evolution over time is represented by the *advection semigroup*, that is, the family of operators

$$\mathtt{U}_t : B^+(D \times V) \to B^+(D \times V) \qquad t \geq 0,$$

defined for $r \in D$, $\upsilon \in V$ and $g \in B^+(D \times V)$ by

$$\mathtt{U}_t[g](r, \upsilon) = \begin{cases} g(r + \upsilon t, \upsilon), & t < \kappa^D_{r,\upsilon}, \\ 0, & \text{otherwise}, \end{cases} \tag{3.12}$$

where

$$\kappa^D_{r,\upsilon} := \inf\{t > 0 : r + \upsilon t \notin D\} \tag{3.13}$$

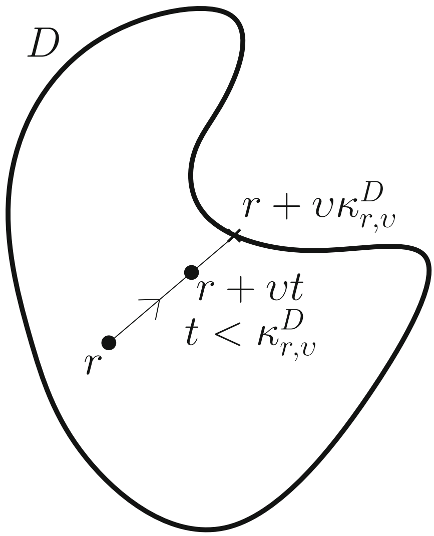

Fig. 3.1 The deterministic time $\kappa^D_{r,\upsilon}$ is the time at which a particle issued from r and travelling with velocity υ exits the domain D

is the deterministic time that a neutron released from r with velocity $\upsilon \in V$ leaves the physical domain D (Fig. 3.1). Note that the definition of $\mathtt{U}$ in (3.12) is equivalent to setting g to be zero if the neutron's configuration enters the set $\{r \in \partial D, \upsilon \in V : \mathbf{n}_r \cdot \upsilon > 0\}$, where $\mathbf{n}_r$ is the outward facing normal at $r \in \partial D$. As alluded to above, this prescribes functions in $B^+(D \times V)$ to be zero on a natural cemetery state.

We also note that we may extend the definition of $\mathtt{U}$ to $r \in \partial D$ with $\mathbf{n}_r \cdot \upsilon < 0$ through the right-hand side of (3.12); however, we will work on the open set D and the case where particles are on ∂D with incoming velocity will never be treated. We can also extend the absorbing set to include velocities satisfying $\mathbf{n}_r \cdot \upsilon = 0$, when $r \in \partial D$; however, this is again somewhat esoteric for the same reasons.

The following result identifies $\mathtt{U} := (\mathtt{U}_t, t \geq 0)$ as having the semigroup property.

Lemma 3.1 *The operator family* $\mathtt{U} = (\mathtt{U}_t, t \geq 0)$ *has the semigroup property, that is,* $\mathtt{U}_{t+s} = \mathtt{U}_t\mathtt{U}_s$ *for all* $s, t \geq 0$.

Proof Let us start by noting that we may more compactly write

$$\mathtt{U}_t[g](r, \upsilon) = g(r + \upsilon t, \upsilon)\mathbf{1}_{(t<\kappa^D_{r,\upsilon})}, \quad t \geq 0.$$

Then note that for $s \geq 0$

$$\kappa^D_{r+\upsilon s,\upsilon} = \inf\{t > 0 : r + \upsilon(t + s) \notin D\} = (\kappa^D_{r,\upsilon} - s) \vee 0,$$

so that, for $t \geq 0$,

$$t < \kappa^D_{r+\upsilon s,\upsilon} \text{ if and only if } t + s < \kappa^D_{r,\upsilon}.$$

Hence, for $g \in B(D \times V)$ and $s, t \geq 0$,

$$\begin{aligned}\mathtt{U}_s[\mathtt{U}_t[g]](r, \upsilon) &= \mathtt{U}_t[g](r + \upsilon s, \upsilon)\mathbf{1}_{(s<\kappa^D_{r,\upsilon})}\\ &= g(r + \upsilon(t + s), \upsilon)\mathbf{1}_{(t<\kappa^D_{r+\upsilon s,\upsilon})}\mathbf{1}_{(s<\kappa^D_{r,\upsilon})}\\ &= \mathtt{U}_{t+s}[g](r, \upsilon),\end{aligned}$$

as required. □

The semigroup $(\mathtt{U}_t, t \geq 0)$ is a well-defined object that we are going work with *in place of* the operator $\mathtt{T}$. To see the connection with $\mathtt{T}$, fix $\upsilon \in V$ and suppose that the function $g(\cdot, \upsilon)$ is regular enough, i.e., $g(\cdot, \upsilon) \in \mathrm{Dom}(\mathtt{T})$ for each $\upsilon \in V$, where

$$\mathrm{Dom}(\mathtt{T}) := \{f \in L^2(D \times V) : \upsilon \cdot \nabla_r f \in L^2(D \times V)\}.$$

Then for this class of functions and for $0 \leq t < \kappa^D_{r,\upsilon}$, $r \in D$, $\upsilon \in V$, we have on $L^2(D \times V)$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathtt{U}_t[g](r, \upsilon) = \frac{\mathrm{d}}{\mathrm{d}t}g(r + \upsilon t, \upsilon) = \upsilon \cdot \nabla_r g(r + \upsilon t, \upsilon) = \mathtt{T}\mathtt{U}_t[g](r, \upsilon), \tag{3.14}$$

with boundary condition and initial condition given by

$$\begin{cases}\mathtt{U}_t[g](r, \upsilon) = 0, & r \in \partial D, \upsilon \in V : \mathbf{n}_r \cdot \upsilon > 0\\ \mathtt{U}_0[g](r, \upsilon) = g(r, \upsilon).\end{cases} \tag{3.15}$$

Note that the conditions given in (3.15) are very much reminiscent of those given in (3.10) for the backward NTE. As we have seen in Chap. 1, for each $\upsilon \in V$, we may formally and uniquely identify solutions to (3.14) on $L^2(D \times V)$ via the orbit

$$\mathtt{U}_t[g](r, \upsilon) = \mathrm{e}^{t\mathtt{T}}g(r, \upsilon), \qquad t \geq 0, r \in D, \upsilon \in V. \tag{3.16}$$

We will shortly see that, for the purposes of probabilistic analysis, the pointwise definition (3.12) will be preferable to the more abstract functional sense in which (3.16) is understood.

3.3 Neutron Random Walk

A neutron random walk (NRW) on $D \times V$ is defined by the two characteristics:

$\alpha(r, \upsilon)$: The rate at which scattering occurs from incoming velocity υ

$\pi(r, \upsilon, \upsilon')\mathrm{d}\upsilon'$: The probability that an incoming velocity υ scatters to an outgoing velocity υ'

As a probability density, π necessarily satisfies $\int_V \pi(r, \upsilon, \upsilon')\mathrm{d}\upsilon' = 1$, for all $\upsilon \in V$. For convenience, we can assume that both are uniformly bounded from above and that, pointwise, $\alpha\pi$ is bounded away from 0. The stochastic evolution of a NRW follows the following rules:

- When issued from $r \in D$ with velocity $\upsilon \in V$, the NRW propagates linearly with that velocity until either it exits the domain D, in which case it is killed, or at the random time T a scattering occurs, where

$$\Pr(T > t) = \exp\{-\textstyle\int_0^t \alpha(r + \upsilon\ell, \upsilon)\mathrm{d}\ell\}, \quad t \geq 0.$$

- When the scattering event occurs in position-velocity configuration (r, υ), a new velocity υ' is selected with probability $\pi(r, \upsilon, \upsilon')\mathrm{d}\upsilon'$.

We denote the associated suite of laws by $\mathbf{P} = (\mathbf{P}_{(r,\upsilon)}, r \in D, \upsilon \in V)$. If we write $(R, \Upsilon) = ((R_t, \Upsilon_t), t \geq 0)$ for the position-velocity of the resulting continuous-time random walk on $D \times V$, then it is easy to show that $((R, \Upsilon), \mathbf{P})$ is a Markov process (Fig. 3.2). Fundamentally, this is because, for $s, t \geq 0$, we have

$$\Pr(T > t + s | T > s) = \exp\{-\textstyle\int_s^{t+s} \alpha(r + \upsilon\ell, \upsilon)\mathrm{d}\ell\} = \exp\{-\textstyle\int_0^t \alpha(r_s + \upsilon\ell, \upsilon)\mathrm{d}\ell\},$$

where $r_s = r + \upsilon s$. Moreover, this carries the implication that, if T_s is the time to the next scatter event after time $s \geq 0$, then the residual lifetime satisfies

$$\mathbf{P}_{(r,\upsilon)}(T_s > t | (R_u, \Upsilon_u), 0 \leq u \leq s) = \exp\{-\textstyle\int_0^t \alpha(R_s + \Upsilon_s\ell, \Upsilon_s)\mathrm{d}\ell\},$$

for $t \geq 0, r \in D, \upsilon \in V$. Thereafter, the scattering occurs still with probability density π, and the process continues according to the stochastic rules in the bullet points above. Note that if we assume that $\alpha\pi$ is pointwise bounded away from zero, then scattering is possible everywhere in $D \times V$. It is worthy of note that, despite the fact that the pair (R, Υ) is Markovian, neither R nor Υ is Markovian as individual processes.

We call the process (R, Υ) an $\alpha\pi$-NRW. It suffices to identify the NRW by the product $\alpha\pi$. Indeed, when $\alpha\pi$ is given as a single rate function, the density π, and hence the rate α, can easily be separated out by normalisation of the product by its total mass to make it a probability distribution. Occasionally, we will have recourse to be specific about the parameters α and π in the law $\mathbf{P}$. In that case, we will work with the probabilities $(\mathbf{P}^{\alpha\pi}_{(r,\upsilon)}, r \in D, \upsilon \in V)$.

Suppose now we denote by

$$\tau^D = \inf\{t > 0 : R_t \in \partial D \text{ and } \mathbf{n}_{R_t} \cdot \Upsilon_t > 0\}, \tag{3.17}$$

the time of first exit from D of the $\alpha\pi$-NRW. We can thus introduce the expectation semigroup of the process (R, Υ) killed on exiting D by

$$\phi_t[g](r, \upsilon) = \mathbf{E}_{(r,\upsilon)}[g(R_t, \Upsilon_t)\mathbf{1}_{(t<\tau^D)}], \qquad t \geq 0,\ r \in D,\ \upsilon \in V,\ g \in B^+(D\times V). \tag{3.18}$$

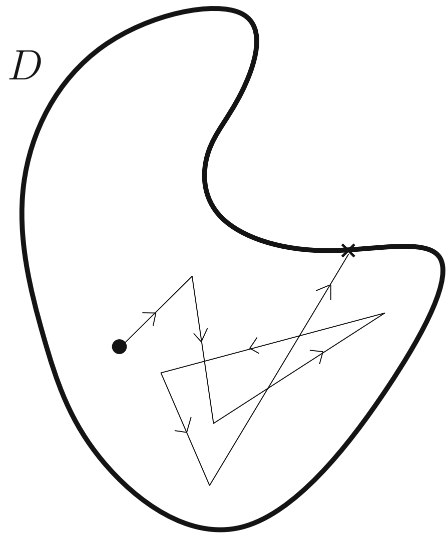

Fig. 3.2 A realisation of a path of a neutron random walk (R, Υ) until exiting its domain D

We technically do not need to include the indicator in our definition, as g should necessarily take the value zero on the cemetery state, where the NRW is sent when it is either absorbed into a nucleus in D or it hits the set $\{r \in D, \upsilon \in V : \mathbf{n}_r \cdot \upsilon > 0\}$. However, we prefer to stress the role of the boundary. Again, although we choose not to, as with $\mathtt{U}$, we may extend the definition of ϕ to allow $r \in \partial D$ such that $\mathbf{n}_r \cdot \upsilon < 0$, that is, starting on the boundary of D with an inward facing velocity, by simply evaluating the expectation in (3.18) that is non-trivial.

Let us also define the operator

$$\tilde{\mathtt{S}} f(r, \upsilon) = \alpha(r, \upsilon) \int_V \{f(r, \upsilon') - f(r, \upsilon)\}\pi(r, \upsilon, \upsilon')\mathrm{d}\upsilon', \quad r \in D, \upsilon \in V, \tag{3.19}$$

for $f \in B^+(D \times V)$. We are now in a position to write down the evolution equation associated to the NRW, which describes the evolution of the semigroup $\phi_t[g]$. Recalling the advection semigroup $(\mathtt{U}_t, t \geq 0)$ defined in the previous section, we have the following lemma.

Lemma 3.2 *For each* $g \in B^+(D \times V)$, $(\phi_t[g], t \geq 0)$ *is a solution in* $B^+(D \times V)$ *to*

$$\phi_t[g](r, \upsilon) = \mathtt{U}_t[g](r, \upsilon) + \int_0^t \mathtt{U}_s[\tilde{\mathtt{S}}\phi_{t-s}[g]](r, \upsilon)\mathrm{d}s, \qquad t \geq 0. \tag{3.20}$$

Proof Starting with the expression in (3.18), we can condition on the first scattering event to obtain

$$\phi_t[g](r,\upsilon) = g(r+\upsilon t,\upsilon)\mathrm{e}^{-\int_0^t \alpha(r+\upsilon s,\upsilon)\mathrm{d}s}\mathbf{1}_{(t<\kappa^D_{r,\upsilon})}$$
$$+\int_0^{t\wedge\kappa^D_{r,\upsilon}} \alpha(r+\upsilon s,\upsilon)\mathrm{e}^{-\int_0^{s\wedge\kappa^D_{r,s}}\alpha(r+\upsilon\ell,\upsilon)\mathrm{d}\ell}$$
$$\left(\int_V \phi_{t-s}(r+s\upsilon,\upsilon')\pi(r+s\upsilon,\upsilon,\upsilon')\mathrm{d}\upsilon'\right)\mathrm{d}s.$$

We may now appeal to Theorem 2.1 and deduce that (3.20) holds. □

The evolution (3.20) gives us a clear analytical representation of the behaviour of (R,Υ), i.e., linear movement interlaced by scattering until the physical boundary D is reached. It is important to note that the evolution Eq. (3.20) is now well defined pointwise as well as on $L^2(D\times V)$. In the $L^2(D\times V)$ setting, we would have expected to see the operator $\mathtt{T}+\tilde{\mathtt{S}}$ describing the evolution of (R,Υ). However, with our description now being pointwise, we see instead in (3.20) a mixture of the well-behaved difference operator $\tilde{\mathtt{S}}$ and the semigroup $(\mathtt{U}_t, t\geq 0)$, whose role replaces that of $\mathtt{T}$.

3.4 Neutron Branching Process

In the previous two sections, we addressed the Markov structure of pure advection and then the more complex setting of advection with scattering. In this section, we introduce a further level of complexity and look at the Markov structure of a system of particles that experience independent movement as neutron random walks, but which also undergo fission. We call such a process a *neutron branching process* (NBP), and as a stochastic model, it is the mathematical object that most closely describes the real-world phenomenon of nuclear fission. Moreover, its evolution can be described in mean using the cross sections $(\sigma_{\mathtt{s}},\sigma_{\mathtt{f}},\pi_{\mathtt{s}},\pi_{\mathtt{f}})$ that appear in the NTE. Earlier we assumed that (H1) is in force, and we will do so here as well for convenience.

Consider a randomly evolving configuration of particles that are specified at time $t\geq 0$ via their physical location and velocity in $D\times V$. We will write this configuration as

$$\big((r_i(t),\upsilon_i(t)) : i=1,\ldots,N_t\big),$$

where N_t is the number of particles alive at time $t\geq 0$. In order to describe the stochastic evolution of our collection of particles, it turns out to be more convenient to represent them as a process in the space of finite atomic measures. To this end, let us define the counting measure

$$X_t(A) = \sum_{i=1}^{N_t}\delta_{(r_i(t),\upsilon_i(t))}(A), \qquad A\in\mathscr{B}(D\times V),\ t\geq 0, \tag{3.21}$$

where δ is the Dirac measure, defined on $\mathscr{B}(D \times V)$, the Borel subsets of $D \times V$. The process $(X_t, t \geq 0)$ is valued in the space of finite counting measures

$$\mathscr{M}_c(D \times V) := \{\textstyle\sum_{i=1}^n \delta_{(r_i,\upsilon_i)} : n \in \mathbb{N}, (r_i, \upsilon_i) \in D \times V, \ i = 1, \cdots, n\} \tag{3.22}$$

and evolves randomly according to the following rules:

- A particle positioned at r with velocity υ will continue to move along the trajectory $r + \upsilon t$, $t \geq 0$, until one of the following things happens.
- The particle leaves the physical domain D, in which case it is instantaneously killed.
- Independently of all other particles in the system, a scattering event occurs and makes an instantaneous change of velocity. For a particle in the system with position and velocity (r, υ), if we write $T_{\mathtt{s}}$ for the random time that scattering may occur, then

  $$\Pr(T_{\mathtt{s}} > t) = \exp\{-\textstyle\int_0^t \sigma_{\mathtt{s}}(r + \upsilon s, \upsilon)\mathrm{d}s\}, \qquad t \geq 0.$$

  When scattering occurs at configuration (r, υ), the new velocity is chosen to be $\upsilon' \in V$ independently with probability $\pi_{\mathtt{s}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon'$.
- Independently of all other particles in the system, a fission event occurs causing the creation of several new particles, each of which will acquire a new velocity. For any particle with configuration (r, υ), if we write $T_{\mathtt{f}}$ for the random time that fission may occur, then

  $$\Pr(T_{\mathtt{f}} > t) = \exp\{-\textstyle\int_0^t \sigma_{\mathtt{f}}(r + \upsilon s, \upsilon)\mathrm{d}s\}, \qquad t \geq 0.$$

  When fission occurs, the particle undergoing fission instantaneously ceases to exist and a random number of new particles, say $N \geq 0$, are created at the same point in space but with randomly distributed, and possibly correlated, velocities, say $(\upsilon_i : i = 1, \cdots, N)$. The outgoing velocities are described by the atomic random measure

  $$\mathsf{Z}(A) := \sum_{i=1}^{N} \delta_{\upsilon_i}(A), \qquad A \in \mathscr{B}(V). \tag{3.23}$$

  When fission occurs at location $r \in D$ from a particle with incoming velocity $\upsilon \in V$, we denote by $\mathcal{P}_{(r,\upsilon)}$ the law of Z. It is worthy of note that $\mathcal{P}_{(r,\upsilon)}(N = 0) > 0$, which corresponds to neutron capture (that is, where a neutron collides with a nucleus but the collision does not result in fission).

We will write $\mathbb{P}_{\delta_{(r,\upsilon)}}$ for the law of X when issued from a single particle with space-velocity configuration $(r, \upsilon) \in D \times V$. More generally, for $\mu \in \mathscr{M}_c(D \times V)$, we understand

$$\mathbb{P}_\mu := \mathbb{P}_{\delta_{(r_1,\upsilon_1)}} \otimes \cdots \otimes \mathbb{P}_{\delta_{(r_n,\upsilon_n)}} \text{ when } \mu = \textstyle\sum_{i=1}^n \delta_{(r_i,\upsilon_i)}.$$

In other words, the process X when issued from initial configuration μ is equivalent to issuing n independent copies of X, each with configuration (r_i, υ_i), $i = 1, \cdots, n$. We will frequently write $(X, \mathbb{P})$, where $\mathbb{P} = (\mathbb{P}_\mu : \mu \in \mathcal{M}_c(D \times V))$ to refer to the NBP we have defined above.

In essence, the NBP is parameterised by the quantities $\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}$ and the family of measures $\mathcal{P} = (\mathcal{P}_{(r,\upsilon)}, r \in D, \upsilon \in V)$, and accordingly, we also refer to $(X, \mathbb{P})$ as a $(\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \mathcal{P})$-NBP. Clearly, our use of the quantities that appear as cross sections in the NTE is pre-emptive, and one of our objectives is to show the connection between the NBP and the NTE.

An important point to make here is that the "data" we need to define our NBP, i.e., $(\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \mathcal{P})$, are not equivalent to the data $(\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{f}})$. Indeed, the latter is less data than the former. The connection between the cross section $\pi_{\mathtt{f}}$ and $\mathcal{P}$ in our NBP is through the relation

$$\int_V g(\upsilon')\pi_{\mathtt{f}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon' = \mathcal{E}_{(r,\upsilon)}\left[\int_V g(\upsilon')\mathsf{Z}(\mathrm{d}\upsilon')\right] = \mathcal{E}_{(r,\upsilon)}[\mathsf{Z}[g]], \tag{3.24}$$

for all $\upsilon \in V$ and $g \in B(V)$, where we are treating Z as a random measure in $\mathcal{M}_c(V)$, the space of finite atomic measures on V.

This begs the question as to whether, given a quadruple $(\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{f}})$, at least one $(\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \mathcal{P})$-NBP exists such that (3.24) holds. In order to construct an example of such a $\mathcal{P}$, we first introduce some further assumptions on the model parameters:

(H2) **We have** $\sigma_{\mathtt{s}}\pi_{\mathtt{s}} + \sigma_{\mathtt{f}}\pi_{\mathtt{f}} > 0$ **on** $D \times V \times V$.

(H3) **There is an open ball** B **compactly embedded in** D **such that** $\sigma_{\mathtt{f}}\pi_{\mathtt{f}} > 0$ **on** $B \times V \times V$.

Assumption (H2) ensures that at least some activity occurs, whether it be scattering or fission. Together with (H3), it ensures that there is at least some fission as well as scattering. Assumption (H2) also avoids the degenerate scenario that once in certain space-velocity configurations, the evolution of the system is entirely deterministic.

We also introduce another assumption on the number of fission offspring.

(H4) **Fission offspring are bounded in number by the constant** $n_{\mathtt{max}} > 1$.

Assumption (H4) is automatically satisfied for the physical processes we intend to use our NBP to model since the maximum number of neutrons that can be emitted during a fission event with positive probability is finite. For example, in an environment where the heaviest nucleus is *Uranium-235*, there are at most 143 neutrons that can be released in a fission event, albeit, in reality, it is more likely that 2 or 3 are released. In particular, this means that

$$\sup_{r\in D, \upsilon\in V} \int_V \pi_{\mathtt{f}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon' \le n_{\mathtt{max}}.$$

Now let us suppose (H1) and (H4) hold. Then, for a given $\pi_{\mathtt{f}}$, define

$$n_{\mathtt{max}} = \min\{k \ge 1 : \sup_{(r,\upsilon)\in D\times V} \int_V \pi_{\mathtt{f}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon' \le k\}.$$

Define the ensemble $(\upsilon_i, i = 1, \cdots, N)$ such that:

(i) $N \in \{0, n_{\max}\}$.
(ii) For each $(r, \upsilon) \in D \times V$, set

$$\mathscr{P}_{(r,\upsilon)}(N = n_{\max}) = \frac{\int_V \pi_{\mathtt{f}}(r, \upsilon, \upsilon'')\mathrm{d}\upsilon''}{n_{\max}}.$$

(iii) On the event $\{N = n_{\max}\}$, each of the $n_{\max}$ neutrons are released with the same velocities $\upsilon_1 = \cdots = \upsilon_{n_{\max}}$, the distribution of this common velocity is given by

$$\mathscr{P}_{(r,\upsilon)}(\upsilon_i \in \mathrm{d}\upsilon' | N = n_{\max}) = \frac{\pi_{\mathtt{f}}(r, \upsilon, \upsilon')}{\int_V \pi_{\mathtt{f}}(r, \upsilon, \upsilon'')\mathrm{d}\upsilon''}\mathrm{d}\upsilon', \qquad i = 1, \ldots, n_{\max}.$$

With the construction (i)–(iii) for $\mathscr{P}_{(r,\upsilon)}$, we can now easily calculate for bounded and measurable $g : V \to [0, \infty)$,

$$\begin{aligned}
&\int_V g(\upsilon')\pi_{\mathtt{f}}(r, v, \upsilon')\mathrm{d}\upsilon' \\
&\quad = 0 \times \left(1 - \mathscr{P}_{(r,\upsilon)}(N = n_{\max})\right) \\
&\qquad\qquad + \mathscr{P}_{(r,\upsilon)}(N = n_{\max})n_{\max} \int_V g(\upsilon')\mathscr{P}_{(r,\upsilon)}(\upsilon_i \in \mathrm{d}\upsilon' | N = n_{\max}) \\
&\quad = \frac{\int_V \pi_{\mathtt{f}}(r, \upsilon, \upsilon'')\mathrm{d}\upsilon''}{n_{\max}} n_{\max} \int_V g(\upsilon') \frac{\pi_{\mathtt{f}}(r, \upsilon, \upsilon')}{\int_V \pi_{\mathtt{f}}(r, \upsilon, \upsilon'')\mathrm{d}\upsilon''}\mathrm{d}\upsilon' \\
&\quad = \int_V g(\upsilon')\pi_{\mathtt{f}}(r, v, \upsilon')\mathrm{d}\upsilon',
\end{aligned}$$

thus matching (3.24), as required.

Like all spatial branching Markov processes, the NBP $(X, \mathbb{P})$ respects the *Markov branching property*. In order to demonstrate this, let us introduce a little more notation. Define the natural filtration of $(X, \mathbb{P})$ by

$$\mathfrak{F}_t := \sigma((r_i(s), \upsilon_i(s)) : i = 1, \cdots, N_s, s \leq t), \qquad t \geq 0. \tag{3.25}$$

Recall that $B^+(D \times V)$ is the space of non-negative, bounded, measurable functions on $D \times V$ and, treating X_t as a random measure in $\mathscr{M}_c(D \times V)$, defined in (3.22), by appealing to the notation in (2.27), we can write

$$X_t[f] = \sum_{i=1}^{N_t} f(r_i(t), \upsilon_i(t)), \qquad t \geq 0,\ f \in B^+(D \times V).$$

Since the function $f \in B^+(D \times V)$ is forced, by definition, to score zero on the cemetery state, when particles reach the boundary of D or are absorbed into a nucleus and are removed from the system, they no longer contribute to the sum.

Lemma 3.3 (Markov Branching Property) *For all $g \in B^+(D \times V)$ and $\mu \in \mathcal{M}_c(D \times V)$ written $\mu = \sum_{i=1}^n \delta_{(r_i, \upsilon_i)}$, we have*

$$\mathbb{E}_\mu \left[\mathrm{e}^{-X_t[g]}\right] = \prod_{i=1}^n \mathrm{v}_t[g](r_i, \upsilon_i), \qquad t \geq 0, \tag{3.26}$$

where

$$\mathrm{v}_t[g](r, \upsilon) := \mathbb{E}_{\delta_{(r,\upsilon)}} \left[\mathrm{e}^{-X_t[g]}\right], \qquad r \in D,\ \upsilon \in V. \tag{3.27}$$

In this respect, $(X, \mathbb{P})$ is a Markov process with state space $\mathcal{M}_c(D \times V)$. In particular,[1]

$$\mathbb{E}\left[\mathrm{e}^{-X_{t+s}[g]} \middle| \mathfrak{F}_t\right] = \prod_{j=1}^{N_t} \mathrm{v}_s[g](r_j(t), \upsilon_j(t)), \tag{3.28}$$

where $((r_j(t), \upsilon_j(t)),\ j = 1, \ldots, N_t)$ is the collection of particles alive at time t.

Proof By construction, once particles come into existence, they do not interact with one another. Hence, given $\mathfrak{F}_t$, for any fixed $t \geq 0$, the expectation of a product of the NBP population at time t will result in expectations over the individual trees that grow from each particle at that time.

Suppose a particle, labelled i, has space-velocity configuration (r, υ) in $D \times V$ at time 0 and exists over a time horizon $[0, t]$. Writing $T_{\mathtt{f}}^{(i)}$ for its fission time, the chance that fission has not occurred by time $t + s$ given that it has not yet occurred by time t is given by

$$\mathbb{P}(T_{\mathtt{f}}^{(i)} > t + s | T_{\mathtt{f}}^i > t) = \frac{\mathbb{P}(T_{\mathtt{f}}^{(i)} > t + s)}{\mathbb{P}(T_{\mathtt{f}}^{(i)} > t)} = \mathrm{e}^{-\int_t^{t+s} \sigma_{\mathtt{f}}(r_i(u), \upsilon_i(u))\mathrm{d}u},$$

where $(r_i(\cdot), \upsilon_i(\cdot))$ is the $\sigma_{\mathtt{s}}\pi_{\mathtt{s}}$-NRW that describes the path of particle i that is issued from a configuration (r, υ). With a straightforward change of time, we can write

$$\int_t^{t+s} \sigma_{\mathtt{f}}(r_i(u), \upsilon_i(u))\mathrm{d}u = \int_0^s \sigma_{\mathtt{f}}(r_i'(\ell), \upsilon_i'(\ell))\mathrm{d}\ell,$$

where $(r_i'(\cdot), \upsilon_i'(\cdot))$ is also a $\sigma_{\mathtt{s}}\pi_{\mathtt{s}}$-NRW, albeit that it is issued from the configuration $(r_i(t), \upsilon_i(t))$.

[1] We always treat the product over an empty set as equal to unity.

Putting these pieces together, we note that when we consider the process X conditional on $\mathfrak{F}_t$, for some fixed $t \geq 0$, the evolution of the trees that grow out of each of the particles with respective configurations $(r_i(t), \upsilon_i(t))$, $i = 1, \cdots, N_t$, is equal in law to an independent collection, $(X, \mathbb{P}_{\delta_{(r_i(t),\upsilon_i(t))}})$, $i = 1, \cdots, N_t$, of NBPs.

More precisely, suppose $N_s^{(i)}$ is the number of descendants in the subtree of the NBP that is initiated from a single mass at space-velocity configuration $(r_i(t), \upsilon_i(t))$ and run over the time horizon $[t, s+t]$, and $(r_j^{(i)}(s), \upsilon_j^{(i)}(s))$, $j = 1, \ldots, N_s^{(i)}$ are the positions in the aforesaid subtree after the relative s units of time; see Fig. 3.3. Then we can decompose the population of particles at time $t+s$ as groups of descendants from particles of the NBP at time t and write

$$
\begin{aligned}
&\mathbb{E}\left[\exp\left(-\sum_{i=1}^{N_{t+s}} g(r_i(t+s), \upsilon_i(t+s))\middle| \mathfrak{F}_t\right)\right] \\
&= \mathbb{E}\left[\exp\left(-\sum_{i=1}^{N_t}\sum_{j=1}^{N_s^{(i)}} g(r_j^{(i)}(s), \upsilon_j^{(i)}(s))\right)\middle| \mathfrak{F}_t\right] \\
&= \prod_{i=1}^{N_t} \mathbb{E}\left[\exp\left(-\sum_{j=1}^{N_s^{(i)}} g(r_j^{(i)}(s), \upsilon_j^{(i)}(s))\right)\middle| \mathfrak{F}_t\right] \\
&= \prod_{i=1}^{N_t} \mathrm{v}_s[g](r_i(t), \upsilon_i(t)).
\end{aligned} \tag{3.29}
$$

In short, (3.28) and hence (3.26) hold. □

The Markov branching property also leads us to the expectation semigroup of the NBP, which is key to understanding the relationship between the NBP and the NTE. With pre-emptive notation, we are interested in

$$
\psi_t[g](r, \upsilon) := \mathbb{E}_{\delta_{(r,\upsilon)}}[X_t[g]], \qquad t \geq 0, r \in D, \upsilon \in V, \tag{3.30}
$$

for $g \in B^+(D \times V)$. As usual we extend the definition of this semigroup (as well as the value of g) to take the value zero when the process enters the cemetery state, that is, where particles are absorbed in D into a nucleus or hit a boundary point $\{r \in D, \upsilon \in V : \mathbf{n}_r \cdot \upsilon > 0\}$. It is not immediately clear that (3.30) constitutes an expectation semigroup; however, the lemma below affirms this fact.

Lemma 3.4 *Assume (H1) holds. The family of operators $(\psi_t, t \geq 0)$ on $B^+(D\times V)$ is an expectation semigroup.*

Proof Similar to the calculation (3.29), the Markov branching property tells us that for $s, t \geq 0$ and $g \in B^+(D \times V)$,

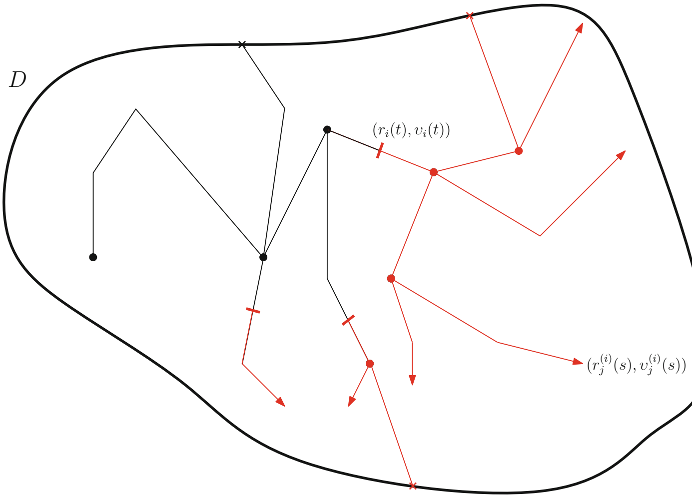

Fig. 3.3 The evolution of a NBP up to time $t + s$ from a single particle. The path in black denotes the evolution over the time interval $[0, t)$, and the path in red denotes the evolution over the time interval $[t, t + s]$. With arbitrary labelling, a particle with position-velocity configuration $(r_i(t), \upsilon_i(t))$ is labelled at time t and s units of time later, again with arbitrary labelling, a descendent particle with position-velocity $(r_j^{(i)}(s), \upsilon_j^{(i)}(s))$ is also labelled

$$\begin{aligned}\mathbb{E}[X_{t+s}[g]|\mathfrak{F}_t] &= \mathbb{E}\left[\sum_{i=1}^{N_{t+s}} g(r_i(t+s), \upsilon_i(t+s))\middle|\mathfrak{F}_t\right]\\ &= \mathbb{E}\left[\sum_{i=1}^{N_t}\sum_{j=1}^{N_s^{(i)}} g(r_j^{(i)}(s), \upsilon_j^{(i)}(s))\middle|\mathfrak{F}_t\right],\end{aligned} \tag{3.31}$$

where $((r_j^{(i)}(s), \upsilon_j^{(i)}(s)), j = 1\cdots, N_s^{(i)})$ is the collection of descendants in the subtree of the NBP that is initiated from a single particle at space-velocity configuration $(r_i(t), \upsilon_i(t))$, run over the time horizon $[t, s + t]$; see Fig. 3.3. Now taking advantage of the Markov branching property, we can continue the computation in (3.31) with

$$\begin{aligned}\mathbb{E}[X_{t+s}[g]|\mathfrak{F}_t] &= \sum_{i=1}^{N_t}\mathbb{E}\left[\sum_{j=1}^{N_s^{(i)}} g(r_j^{(i)}(s), \upsilon_j^{(i)}(s))\middle|\mathfrak{F}_t\right]\\ &= \sum_{i=1}^{N_t}\mathbb{E}_{\delta_{(r_i(t),\upsilon_i(t))}}[X_s[g]]\\ &= X_t[\psi_s[g]].\end{aligned}$$

Hence, taking expectations again across these equalities,

$$\psi_{t+s}[g](r,\upsilon) = \psi_t[\psi_s[g]](r,\upsilon), \qquad s,t \geq 0, r \in D, \upsilon \in V, g \in B^+(D\times V),$$

which constitutes the expectation semigroup property. □

As with Lemma 3.2 for the NRW, the semigroup $(\psi_t, t \geq 0)$ solves a mild evolution equation (in the pointwise sense) in $B^+(D\times V)$, which we now investigate. To this end, recall the definition of the operators $\mathtt{S}$ and $\mathtt{F}$ defined in (3.7).

Lemma 3.5 *Under (H1), for $g \in B^+(D\times V)$, there exist constants $C_1, C_2 > 0$ such that $\psi_t[g]$, as given in (3.30), is uniformly bounded by $C_1 \exp(C_2 t)$, for all $t \geq 0$. Moreover, $(\psi_t[g], t \geq 0)$ is the unique solution in $B^+(D\times V)$ to the mild equation:*

$$\psi_t[g] = \mathtt{U}_t[g] + \int_0^t \mathtt{U}_s[(\mathtt{S}+\mathtt{F})\psi_{t-s}[g]]\mathrm{d}s, \qquad t \geq 0, \tag{3.32}$$

for which (3.5) holds.

Proof To show that (3.30) solves (3.32) is a simple matter of conditioning the expression in (3.30) on the first fission or scatter event (whichever occurs first). In that case, we observe, for $t \geq 0$, $r \in D$, $\upsilon \in V$, $g \in B^+(D\times V)$,

$$\begin{aligned}
&\psi_t[g](r,\upsilon)\\
&= \mathrm{e}^{-\int_0^t \sigma(r+\upsilon\ell,\upsilon)\mathrm{d}\ell} g(r+\upsilon t,\upsilon)\mathbf{1}_{(t<\kappa^D_{r,\upsilon})}\\
&\qquad + \int_0^t \mathbf{1}_{(s<\kappa^D_{r,\upsilon})} \frac{\sigma_{\mathtt{f}}(r+\upsilon s,\upsilon)}{\sigma(r+\upsilon s,\upsilon)} \sigma(r+\upsilon s,\upsilon)\mathrm{e}^{-\int_0^s \sigma(r+\upsilon\ell,\upsilon)\mathrm{d}\ell}\\
&\qquad\qquad\qquad \int_V \psi_{t-s}(r+\upsilon s,\upsilon')\pi_{\mathtt{f}}(r,\upsilon,\upsilon')\mathrm{d}\upsilon'\mathrm{d}s\\
&\qquad + \int_0^t \mathbf{1}_{(s<\kappa^D_{r,\upsilon})} \frac{\sigma_{\mathtt{s}}(r+\upsilon s,\upsilon)}{\sigma(r+\upsilon s,\upsilon)} \sigma(r+\upsilon s,\upsilon)\mathrm{e}^{-\int_0^s \sigma(r+\upsilon\ell,\upsilon)\mathrm{d}\ell}\\
&\qquad\qquad\qquad \int_V \psi_{t-s}(r+\upsilon s,\upsilon')\pi_{\mathtt{s}}(r,\upsilon,\upsilon')\mathrm{d}\upsilon'\mathrm{d}s\\
&= \mathrm{e}^{-\int_0^t \sigma(r+\upsilon\ell,\upsilon)\mathrm{d}\ell} g(r+\upsilon t,\upsilon)\mathbf{1}_{(t<\kappa^D_{r,\upsilon})}\\
&\qquad + \int_0^t \mathbf{1}_{(s<\kappa^D_{r,\upsilon})} \mathrm{e}^{-\int_0^s \sigma(r+\upsilon\ell,\upsilon)\mathrm{d}\ell} (\mathtt{S}+\mathtt{F}+\sigma)\psi_{t-s}(r+\upsilon s,\upsilon)\mathrm{d}s,
\end{aligned}$$

where we recall that $\sigma = \sigma_{\mathtt{f}} + \sigma_{\mathtt{s}}$ is the joint rate at which either a fission or scatter occurs and, on the event that either of these two events occur, the chance that it is a scatter event is $\sigma_{\mathtt{s}}/\sigma$ and the chance that it is a fission event is $\sigma_{\mathtt{f}}/\sigma$. Next, applying Theorem 2.1, we can remove the multiplicative potential, and recalling the definition of $(\mathtt{U}_s, s \geq 0)$ in (3.12), we recover the Eq. (3.32).

To show uniqueness of (3.32), suppose that $\psi^{(1)}$ and $\psi^{(2)}$ are two bounded non-negative solutions to (3.32) and, for each $g \in B^+(D \times V)$, $r \in D$, $\upsilon \in V$, set

$$\tilde{\psi}_t[g] = \sup_{r \in D, \upsilon \in V} |\psi_t^{(1)}[g](r, \upsilon) - \psi^{(2)}[g]_t(r, \upsilon)|.$$

Necessarily, $\tilde{\psi}_t[g] \geq 0$ has zero initial condition and

$$\begin{aligned}\tilde{\psi}_t[g] &= \sup_{r \in D, \upsilon \in V} \left| \int_0^t \mathtt{U}_s[(\mathtt{S} + \mathtt{F})(\psi_{t-s}^{(1)}[g] - \psi_{t-s}^{(2)}[g])](r, \upsilon) \mathrm{d}s \right| \\ &\leq (2\bar{\sigma}_{\mathtt{s}} + (n_{\mathtt{max}} + 1)\bar{\sigma}_{\mathtt{f}}) \mathtt{Vol}(V) \int_0^t \tilde{\psi}_{t-s}[g] \mathrm{d}s \qquad t \geq 0,\end{aligned}$$

where we have used the definitions of $\mathtt{S}$ and $\mathtt{F}$ in (3.7), where

$$\bar{\sigma}_{\mathtt{s}} := \sup_{r \in D, \upsilon \in V} \sigma_{\mathtt{s}}(r, \upsilon),$$

with a similar definition for $\bar{\sigma}_{\mathtt{f}}$, and $\mathtt{Vol}(V)$ is the Lebesgue volume of V. We may now appeal directly to Grönwall's lemma to deduce that $\tilde{\psi}_t[g] = 0$, $t \geq 0$.

Finally, to show domination by exponential growth, we may return to the stochastic definition of X_t and note that $N_t := X_t[1]$, $t \geq 0$, is the number of particles in the NBP at time t. Our objective is to show that, on the same probability space, we can pathwise upper bound the counting process $(N_t, t \geq 0)$ by $(\tilde{N}_t, t \geq 0)$, i.e., $N_t \leq \tilde{N}_t$ almost surely, $t \geq 0$, where $\tilde{N}$ is a Bienyamé–Galton–Watson branching process in continuous time. The construction is such that $\tilde{N}$ produces precisely $n_{\mathtt{max}}$ offspring at each branching event (i.e., a net difference of $n_{\mathtt{max}} - 1$), and whenever N increases due to a branching event, so does $\tilde{N}$. In addition, $\tilde{N}$ permits each individual to additionally branch at rate $\bar{\sigma}_{\mathtt{f}} - \sigma_{\mathtt{f}}$, again producing precisely $n_{\mathtt{max}}$ offspring. The stochastic domination of N by $\tilde{N}$ is clear, and also the choice of the additional branching rate is such that $\tilde{N}$ is a Bienyamé–Galton–Watson process with branching rate $\bar{\sigma}_{\mathtt{f}}$ and offspring distribution that is concentrated on $n_{\mathtt{max}}$ with probability one. For any $g \in B^+(D \times V)$, $r \in D$, and $\upsilon \in V$, we may now use the fact that the Bienyamé–Galton–Watson processes grow exponentially at a rate given by its branching rate multiplied by the mean number of offspring minus one to obtain

$$\mathbb{E}_{\delta_{(r,\upsilon)}}[X_t[g]] \leq \|g\| \mathbb{E}_{\delta_{(r,\upsilon)}}[X_t[1]] \leq \|g\| \mathrm{e}^{\bar{\sigma}_{\mathtt{f}}(n_{\mathtt{max}} - 1)t},$$

where $\|\cdot\|$ denotes the supremum norm on $B^+(D \times V)$. The proof is complete. $\square$

3.5 Mild NTE vs. Backward NTE

Despite the fact that (3.32) gives us an interpretation of the NTE on $B^+(D \times V)$, it is still unclear how our NBP relates to the backward NTE in (3.9) on the space $L^2(D \times V)$. The following theorem shows that when we treat $\psi_t : B^+(D \times V) \mapsto B^+(D \times V)$, $t \geq 0$, as a family of functions in $L^2(D \times V)$, then it agrees with the unique solution to the backward NTE in (3.9).

Theorem 3.2 *Under the assumptions of Lemma 3.5, if $g \in B^+(D \times V)$, then the solution to* (3.32) *solves* (3.9) *in $L^2(D \times V)$.*

Proof Consider the adjusted ACP with inhomogeneity given by

$$\begin{cases} \dfrac{\partial u_t}{\partial t} = \mathtt{T}u_t + (\mathtt{S} + \mathtt{F})\hat{\Psi}_t \\ u_0 = g \end{cases}, \tag{3.33}$$

where $(\hat{\Psi}_t, t \geq 0)$ is taken as the solution to (3.9) with boundary condition (3.10). By taking the difference of two solutions and invoking the uniqueness of the homogenous ACP

$$\begin{cases} \dfrac{\partial u_t}{\partial t} = \mathtt{T}u_t \\ u_0 = 0 \end{cases} \tag{3.34}$$

in $L^2(D \times V)$, we note that the solution to (3.33) is unique in $L^2(D \times V)$. However, on the one hand, straightforward differentiation gives us that, providing $g \in \mathtt{Dom}(\mathtt{T} + \mathtt{S} + \mathtt{F}) = \mathtt{Dom}(\mathtt{T})$,

$$u_t := \mathrm{e}^{t\mathtt{T}} g + \int_0^t \mathrm{e}^{(t-s)\mathtt{T}} (\mathtt{S} + \mathtt{F})\hat{\Psi}_s \mathrm{d}s, \qquad t \geq 0, \tag{3.35}$$

solves (3.33). On the other hand, taking account of the fact that $(\hat{\Psi}_t, t \geq 0)$ solves (3.9), it is also clear by inspection in (3.33) that

$$u_t = \hat{\Psi}_t, \qquad t \geq 0,$$

solves (3.33). Uniqueness thus tells us that, on $L^2(D \times V)$,

$$\hat{\Psi}_t = \mathtt{U}_t[g] + \int_0^t \mathtt{U}_s[(\mathtt{S} + \mathtt{F})\hat{\Psi}_{t-s}]\mathrm{d}s, \qquad t \geq 0,$$

where we have appealed to the identification of $\exp(t\mathtt{T})$ as equal to $\mathtt{U}_t$ on $L^2(D \times V)$, cf. (3.16), and we have also reversed the direction of integration in (3.35). In conclusion, whereas $(\psi_t[g], t \geq 0)$ solves (3.32) in the pointwise sense on $B^+(D \times V)$, $(\hat{\Psi}_t, t \geq 0)$ solves it in $L^2(D \times V)$.

Since $D \times V$ is bounded, we have $B^+(D \times V) \subset L^2(D \times V)$, and hence, we can consider

$$\|\psi_t[g] - \hat{\Psi}_t\|_2 = \left\| \int_0^t \mathtt{U}_s[(\mathtt{S} + \mathtt{F})\{\psi_{t-s}[g] - \hat{\Psi}_{t-s}\}]\mathrm{d}s \right\|_2, \qquad t \geq 0.$$

To this end, let us note that, for $T > 0$, and $\omega_t \in L^2(D \times V)$, $t \leq T$, we have

$$\begin{aligned} \left\| \int_0^t \omega_s \mathrm{d}s \right\|_2^2 &= \int_{D\times V} \left(t \int_0^t \omega_s(r, \upsilon) \frac{\mathrm{d}s}{t} \right)^2 \mathrm{d}r\mathrm{d}\upsilon \\ &\leq \int_{D\times V} t^2 \left(\int_0^t \omega_s(r, \upsilon)^2 \frac{\mathrm{d}s}{t} \right) \mathrm{d}r\mathrm{d}\upsilon \\ &\leq T \int_0^t \|\omega_s\|_2^2 \mathrm{d}s, \qquad t \leq T, \end{aligned} \tag{3.36}$$

where in the first inequality we have used Jensen's inequality. Moreover, for $g \in L^2(D \times V)$,

$$\begin{aligned} \|\mathtt{U}_s[g]\|_2^2 &= \int_{D\times V} \mathbf{1}_{(s<\kappa^D_{r,\upsilon})} g(r + \upsilon s, \upsilon)^2 \mathrm{d}r\mathrm{d}\upsilon \\ &\leq \int_{D\times V} g(r', \upsilon)^2 \mathrm{d}r'\mathrm{d}\upsilon \\ &= \|g\|_2^2, \end{aligned} \tag{3.37}$$

where the inequality follows as a consequence that, for each υ, the integral of $r \mapsto \mathbf{1}_{(s<\kappa^D_{r,\upsilon})} g^2(r + \upsilon s, \upsilon)$ integrates over a subdomain of D in the first argument as r varies over D. Also, we have for the operator $\mathtt{S}$ (and similarly for $\mathtt{F}$) that, for $g \in L^2(D \times V)$,

$$\begin{aligned} \|(\mathtt{S} + \sigma_{\mathtt{S}})g\|_2 &= \left(\int_{D\times V} \left(\int_V g(r, \upsilon')\sigma_{\mathtt{S}}(r, \upsilon)\pi_{\mathtt{S}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon' \right)^2 \mathrm{d}r\mathrm{d}\upsilon \right)^{1/2} \\ &\leq C \left(\int_{D\times V} \left(\int_V g(r, \upsilon') \times 1 \,\mathrm{d}\upsilon' \right)^2 \mathrm{d}r\mathrm{d}\upsilon \right)^{1/2} \\ &\leq C \left(\mathtt{Vol}(V) \int_{D\times V} \int_V g(r, \upsilon')^2 \mathrm{d}\upsilon'\mathrm{d}r \right)^{1/2} \\ &\leq C\mathtt{Vol}(V)\|g\|_2, \end{aligned} \tag{3.38}$$

where the constant C appears by using (H1) and upper estimating the uniformly bounded cross sections, and in the second inequality, we have used the Cauchy–Schwarz inequality.

It thus follows from (3.36), (3.37), and (3.38) that, for $t \le T$, writing $\omega_t = \psi_t[g] - \hat{\Psi}_t$, $t \ge 0$,

$$
\begin{aligned}
\|\omega_t\|_2^2 &= \left\| \int_0^t \mathtt{U}_s[(\mathtt{S}+\mathtt{F})\omega_{t-s}]\mathrm{d}s \right\|_2^2 \\
&\le T \int_0^t \|\mathtt{U}_s[(\mathtt{S}+\mathtt{F})\omega_{t-s}]\|_2^2 \mathrm{d}s \\
&\le T \int_0^t \|(\mathtt{S}+\mathtt{F})\omega_{t-s}\|_2^2 \mathrm{d}s \\
&\le T \int_0^t (\|(\mathtt{S}+\sigma_{\mathtt{s}})\omega_s\|_2 + \|(\mathtt{F}+\sigma_{\mathtt{f}})\omega_s\|_2 + \|\sigma\omega_s\|_2)^2 \,\mathrm{d}s \\
&\le C' \int_0^t \|\omega_s\|_2^2 \mathrm{d}s, \qquad t \le T,
\end{aligned}
\tag{3.39}
$$

where the constant C' comes from the fact that σ is uniformly bounded, cf. (H1).

The final inequality in (3.39) together with Grönwall's Lemma now tells us that $\|\omega_t\|_2 = 0$, for all $t \le T$. Since T is chosen arbitrarily, it follows that $(\psi_t[g], t \ge 0)$ and $(\hat{\Psi}_t, t \ge 0)$ are indistinguishable in $L^2(D \times V)$. □

3.6 Re-Oriented Mild NTE

The representation (3.32) for the evolution of the expectation semigroup (3.30) is not the only way to describe the system in the form of a recursion. The mild NTE (3.32) clearly tells us that a particle drifts with pure advection until either it undergoes scatter or fission, which accounts for the appearance of the integral term in which the sum of the two generators $\mathtt{S}+\mathtt{F}$ appears. Moreover, we recall that this representation came about by conditioning the expectation that defines ψ_t on the first fission or scatter event. In this spirit, it is easy to imagine that by first conditioning on just the first fission event, we will get a different recursion.

Until the first fission event, a neutron will evolve simply as a $\sigma_{\mathtt{s}}\pi_{\mathtt{s}}$-NRW killed when it exits D. Suppose we denote by $(\mathtt{P}_t, t \ge 0)$ the expectation semigroup of the latter, as discussed in (3.18). We will need a slightly stronger assumption than (H2).

(H2)* **We have** $\inf_{r\in D, \upsilon,\upsilon'\in V} \alpha(r,\upsilon)\pi(r,\upsilon,\upsilon') > 0$.

Lemma 3.6 *Assume (H1) holds. For each* $g \in B^+(D \times V)$*, the expectation semigroup* $(\psi_t[g], t \geq 0)$ *is also the unique solution in* $B^+(D \times V)$ *to*

$$\psi_t[g] = \mathtt{P}_t[g] + \int_0^t \mathtt{P}_s[\mathtt{F}\psi_{t-s}[g]]\mathrm{d}s, \qquad t \geq 0. \tag{3.40}$$

The heuristic we used to write down this lemma is in fact sufficient to develop a rigorous proof in the spirit of the proof of Lemma 3.5. As such we leave the details to the reader.

Again we need to stress that the Eq. (3.40) is understood in the pointwise sense. That said, if we were again to turn to the reasoning of Theorem 3.2, we can also express (3.40) on $L^2(D \times V)$ in the form

$$\psi_t[g] = \mathrm{e}^{t(\mathtt{T}+\mathtt{S})}g + \int_0^t \mathrm{e}^{s(\mathtt{T}+\mathtt{S})}[\mathtt{F}\psi_{t-s}[g]]\mathrm{d}s, \qquad t \geq 0. \tag{3.41}$$

Now referring to the conclusion of Theorem 2.1, which, technically speaking, is for expectation semigroups on $B^+(D \times V)$ rather than c_0-semigroups on $L^2(D \times V)$, we may think of transferring the multiplicative potential $(\mathrm{e}^{t\mathtt{S}}g, t \geq 0)$ to an additive potential. If we were at our liberty to use Theorem 2.1 in this context, then this would give us equivalence of with solutions to (3.41) and bring us back to

$$\psi_t[g] = \mathrm{e}^{t\mathtt{T}}g + \int_0^t \mathrm{e}^{s\mathtt{T}}[(\mathtt{S}+\mathtt{F})\psi_{t-s}[g]]\mathrm{d}s, \qquad t \geq 0,$$

which is the $L^2(D \times V)$ analogue of (3.32). Of course, as soon as we re-prove Theorem 2.1 for equations on L^2 spaces, and this can indeed be done, then the discussion above can be made rigorous. This is an unnecessary distraction for the course of this text, and we leave this missing result for the reader to prove as an exercise.

3.7 Comments

The probabilistic interpretation of the NTE was appreciated from its very first mathematical handling; see for example Davison and Sykes [36], Bell [7] and Williams [131] and references therein to name but a few textbooks. Indeed, the physical description of nuclear fission, when governed by basic principles, allowing for additional randomness, is nothing more than a branching Markov process. Numerous derivations of the NTE from this perspective can be found in the literature to various degrees of rigour; see, e.g., Bell [7], Mori et al. [100], Pazy and Rabinowitz [110], Lewins [94], and Pázsit and Pál. [107].

A more modern treatment of the probabilistic representation through Feynman–Kac expectation semigroups and the connection to the theory of Markov diffusions is found in Dautray et al. [34]. A purely probabilistic approach can be found in Lapeyre et al. [89]. The perspective we have illustrated in this chapter, in particular via mild equations on $(B(D \times V), \|\cdot\|)$ and the indistinguishability between solutions in $(B(D \times V), \|\cdot\|)$ and $L^2(D \times V)$, largely comes from the recent works of Cox et al. [30, 31] and Harris et al. [69, 74].

Chapter 4
Many-to-One, Perron–Frobenius and Criticality

Now that the precise mathematical relationship between the NTE and the NBP is clear, we now look at how we can profit from this. The first port of call in this respect is to understand how to provide a rigorous analogue of the spectral asymptotic behaviour given in Theorem 1.2 for the NTE as an $L^2(D \times V)$ solution but now for the setting of $B^+(D \times V)$ solutions that emerge from our mild NTE formulation (3.32). The way we will do this is to draw the general Perron–Frobenius result for Markov processes in Theorem 2.2 into our current setting. This requires us to develop a second representation of the solution to the mild NTE (3.32) in terms of a single particle Markov process.

4.1 Many-to-One Representation

As alluded to above, there is a second representation of the expectation semigroup $(\psi_t, t \geq 0)$. To describe the second stochastic representation of (3.32), we define

$$\beta(r, \upsilon) = \sigma_{\mathtt{f}}(r, \upsilon) \left(\int_V \pi_{\mathtt{f}}(r, \upsilon, \upsilon') \mathrm{d}\upsilon' - 1 \right) \geq - \sup_{r \in D, \upsilon \in V} \sigma_{\mathtt{f}}(r, \upsilon) > -\infty, \tag{4.1}$$

where the lower bound is due to assumption (H1). Let us also introduce a specific $\alpha\pi$-NRW (recall the NRW was introduced in Sect. 3.3) and define, for $r \in D$, $\upsilon, \upsilon' \in V$,

$$\alpha(r, \upsilon)\pi(r, \upsilon, \upsilon') = \sigma_{\mathtt{s}}(r, \upsilon)\pi_{\mathtt{s}}(r, \upsilon, \upsilon') + \sigma_{\mathtt{f}}(r, \upsilon)\pi_{\mathtt{f}}(r, \upsilon, \upsilon'), \tag{4.2}$$

where π is taken to be a probability density. The latter assumption forces the definition

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8_4

$$\alpha(r,\upsilon) = \sigma_{\mathtt{s}}(r,\upsilon) + \sigma_{\mathtt{f}}(r,\upsilon)\int_V \pi_{\mathtt{f}}(r,\upsilon,\upsilon')\mathrm{d}\upsilon'. \tag{4.3}$$

Lemma 4.1 (Many-to-One) *Under the assumptions of Lemma 3.5, we have the second representation*

$$\psi_t[g](r,\upsilon) = \mathbf{E}_{(r,\upsilon)}\left[\mathrm{e}^{\int_0^t \beta(R_s,\Upsilon_s)\mathrm{d}s} g(R_t,\Upsilon_t)\mathbf{1}_{\{t<\tau^D\}}\right], \tag{4.4}$$

for $t \geq 0$, $r \in D$, $\upsilon \in V$, and $g \in B^+(D\times V)$, where

$$\tau^D = \inf\{t > 0 : R_t \in \partial D \textit{ and } \mathbf{n}_{R_t}\cdot\Upsilon_t > 0\}$$

and $\mathbf{P}_{(r,\upsilon)}$ is the law of the $\alpha\pi$-NRW starting from a single neutron with configuration (r,υ).

Proof The proof is relatively straightforward, appealing to a proof similar to Lemma 3.5. We start by noting that the expression (4.4) has the expectation semigroup property thanks to Lemma 2.1. By conditioning the right-hand side of (4.4) on the first scatter event, with the help of Theorem 2.1, we obtain

$$\begin{aligned}
\psi_t[g](r,\upsilon) &= \mathbf{1}_{(t<\kappa^D_{r,\upsilon})}\mathrm{e}^{-\int_0^t \alpha(r+\upsilon\ell,\upsilon)\mathrm{d}\ell}\mathrm{e}^{\int_0^t \beta(r+\upsilon\ell,\upsilon)\mathrm{d}\ell} g(r+\upsilon t,\upsilon)\\
&\quad + \int_0^t \mathbf{1}_{(s<\kappa^D_{r,\upsilon})}\alpha(r+\upsilon s,\upsilon)\mathrm{e}^{-\int_0^s \alpha(r+\upsilon\ell,\upsilon)\mathrm{d}\ell}\\
&\qquad \mathrm{e}^{\int_0^s \beta(r+\upsilon\ell,\upsilon)\mathrm{d}\ell}\int_V \psi_{t-s}[g](r+\upsilon s,\upsilon')\pi(r+\upsilon s,\upsilon,\upsilon')\mathrm{d}\upsilon'\mathrm{d}s\\
&= \mathtt{U}_t[g](r,\upsilon) + \int_0^t \mathtt{U}_s[(\tilde{\mathtt{S}}+\beta)\psi_{t-s}[g]](r,\upsilon)\mathrm{d}s,
\end{aligned} \tag{4.5}$$

where $\tilde{\mathtt{S}}$ was defined in (3.19). It is straightforward algebra to show, with the help of (4.2), that, for $f \in B^+(D\times V)$,

$$(\tilde{\mathtt{S}}+\alpha)f = (\mathtt{S}+\sigma_{\mathtt{s}})f + (\mathtt{F}+\sigma_{\mathtt{f}})f,$$

and hence, recalling (4.3) and noting that $\beta - \alpha = -(\sigma_{\mathtt{s}}+\sigma_{\mathtt{f}})$,

$$(\tilde{\mathtt{S}}+\beta)f = (\mathtt{S}+\mathtt{F})f.$$

Consequently, we see that the right-hand side of (4.4) also gives an alternative representation of the unique solution to (3.32). □

Remark 4.1 The proof of the many-to-one Lemma 4.1 gives us yet another representation of the mild equation. Indeed, (4.5) tells us that, in $B^+(D\times V)$,

$$\psi_t[g] = \mathtt{U}_t[g] + \int_0^t \mathtt{U}_s[(\tilde{\mathtt{S}} + \beta)\psi_{t-s}[g]]\mathrm{d}s, \qquad t \geq 0. \tag{4.6}$$

Recalling the discussion in Sect. 3.6, it is easy to see how one may heuristically transform between the Eqs. (4.6), (3.40), and (3.32).

Define

$$\bar{\beta} := \sup_{r \in D, \upsilon \in V} \beta(r, \upsilon),$$

which is finite thanks to (H1). Let us now introduce the expectation semigroup $\mathtt{P}^\dagger := (\mathtt{P}^\dagger_t, t \geq 0)$ via

$$\begin{aligned}
\mathtt{P}^\dagger_t[g](r, \upsilon) &:= \mathrm{e}^{-\bar{\beta}t}\psi_t[g](r, \upsilon) \\
&= \mathbf{E}_{(r,\upsilon)}\left[\mathrm{e}^{\int_0^t (\beta(R_s, \Upsilon_s) - \bar{\beta})\mathrm{d}s} g(R_t, \Upsilon_t)\mathbf{1}_{\{t < \tau^D\}}\right] \\
&= \mathbf{E}_{(r,\upsilon)}\left[g(R_t, \Upsilon_t)\mathbf{1}_{\{t < \mathtt{k}\}}\right] \\
&=: \mathbf{E}^\dagger_{(r,\upsilon)}\left[g(R_t, \Upsilon_t)\right], \qquad t \geq 0, r \in D, \upsilon \in V, g \in B^+(D \times V),
\end{aligned} \tag{4.7}$$

where

$$\mathtt{k} = \inf\{t > 0 : \int_0^t (\bar{\beta} - \beta(R_s, \Upsilon_s))\mathrm{d}s > \mathbf{e}\} \wedge \tau^D, \tag{4.8}$$

and $\mathbf{e}$ is an independent exponentially distributed random variable with mean 1. Another way to describe $\mathtt{P}^\dagger := (\mathtt{P}^\dagger_t, t \geq 0)$ is the expectation semigroup of the $\alpha\pi$-neutron random walk killed at rate $(\bar{\beta} - \beta)$. To see where the definition in (4.8) comes from, we note that for any $\gamma \in B^+(D \times V)$,

$$\mathbf{P}_{(r,\upsilon)}\left(\mathbf{e} > \int_0^t \gamma(R_s, \Upsilon_s)\mathrm{d}s \,\middle|\, ((R_u, \Upsilon_u), u \leq t)\right) = \mathrm{e}^{-\int_0^t \gamma(R_s, \Upsilon_s)\mathrm{d}s}.$$

4.2 Perron–Frobenius Asymptotic

We will naturally write $\mathbf{P}^\dagger_{(r,\upsilon)}$ for the (sub)probability measure associated to $\mathbf{E}^\dagger_{(r,\upsilon)}$, $r \in D, \upsilon \in V$. The family $\mathbf{P}^\dagger := (\mathbf{P}^\dagger_{(r,\upsilon)}, r \in D, \upsilon \in V)$ now defines a Markov family of probability measures on the path space of the neutron random walk with cemetery state $\{\dagger\}$, which is where the path is sent when hitting the boundary ∂D with an outgoing velocity, when the clock associated to the killing rate $\bar{\beta} - \beta$ rings or when there is a fission event with zero offspring in the underlying branching process. We remind the reader for future calculations that we can extend the domain of functions on $D \times V$ to accommodate taking a value on $\{\dagger\}$ by insisting that this value is always 0.

This gives us a variation on the theme of Lemma 4.1, namely that the uniqueness of bounded non-negative solutions to (3.32) gives us, for $g \in B^+(D \times V)$

$$\mathbb{E}_{\delta_{(r,\upsilon)}}[X_t[g]] = \mathrm{e}^{\bar{\beta}t}\mathbf{E}^{\dagger}_{(r,\upsilon)}[g(R_t, \Upsilon_t)], \qquad t \geq 0, r \in D, \upsilon \in V.$$

With this identity in hand, we can revisit some of the known theories available to us for Markov processes to look for behaviour of the type described in Theorem 2.2. Specifically, by verifying hypotheses (A1) and (A2) there, we can translate the conclusion for the Markov process $((R, \Upsilon), \mathbf{P}^{\dagger})$ to the setting of the semigroup $(\psi_t, t \geq 0)$ and obtain the following result.

Theorem 4.1 (Perron–Frobenius Asymptotics of the Mild Solution) *Suppose that D is convex and (H1) and (H2*) hold. Then, for the semigroup $(\psi_t, t \geq 0)$ identified by (3.32), there exist a constant $\lambda_* \in \mathbb{R}$, a positive*[1] *right eigenfunction $\varphi \in B^+(D \times V)$, and a left eigenmeasure that is absolutely continuous with respect to Lebesgue measure on $D \times V$ with density $\tilde{\varphi} \in B^+(D \times V)$, both having associated eigenvalue $\mathrm{e}^{\lambda_* t}$, and such that φ (resp., $\tilde{\varphi}$) is uniformly (resp., a.e. uniformly) bounded away from zero on each compactly embedded subset of $D \times V$. In particular for all $g \in B^+(D \times V)$*

$$\langle \tilde{\varphi}, \psi_t[g]\rangle = \mathrm{e}^{\lambda_* t}\langle \tilde{\varphi}, g\rangle \quad \textit{(resp. } \psi_t[\varphi] = \mathrm{e}^{\lambda_* t}\varphi\textit{)} \quad t \geq 0. \tag{4.9}$$

Moreover, there exist constants $C, \varepsilon > 0$ such that

$$\sup_{g \in B_1^+(D \times V)} \left\| \mathrm{e}^{-\lambda_* t}\varphi^{-1}\psi_t[g] - \langle \tilde{\varphi}, g\rangle \right\| \leq C\mathrm{e}^{-\varepsilon t}, \quad t \geq 0. \tag{4.10}$$

The proof of the Theorem 4.1 is extremely technical and therefore pushed to the end of this chapter (and can be skipped if the reader prefers to move forward in the book). As alluded to above, it consists of verifying the two assumptions (A1) and (A2) of Theorem 2.2. For want of a better way to summarise the approach behind this, it is done by brute force, which is not surprising given the significantly inhomogeneous nature of the problem. As such, while the requirement that D is convex is a sufficient condition to help us verify (A1) and (A2), it is certainly not necessary.

In addition to the above paragraph of remarks, there are a number of others that we should also make concerning this theorem. First let us make a remark on notation

[1] To be precise, by a positive eigenfunction, we mean a mapping from $D \times V \to (0, \infty)$. This does not prevent it being valued zero on ∂D, as D is open and bounded.

involving $\tilde{\varphi}$. In the statement of Theorem 4.1, and repeatedly in the forthcoming text, the reader will observe that where we write, for example, $\langle f, \tilde{\varphi}\rangle$, we could also write $\eta[f]$ with η being a measure that is absolutely continuous with respect to $\mathrm{d}r\mathrm{d}\upsilon$ on $D \times V$ with density $\tilde{\varphi}$. In the second part of this book, we will look at a general version of (4.10) in which we will treat $\tilde{\varphi}$ as a measure rather than a function.

Next, it is customary that the eigenfunctions $\tilde{\varphi}$ and φ are normalised so that $\langle \varphi, \tilde{\varphi}\rangle = 1$. As we will see in Chap. 6, it will turn out that we can also interpret the product $\varphi\tilde{\varphi}$ as the stationary density of an auxiliary Markov process.

Finally, we remark that Theorem 4.1 clearly mirrors the situation for the forward NTE on $L^2(D \times V)$ given in Theorem 1.2. Clearly, the eigentriple λ_*, φ, $\tilde{\varphi}$ of Theorem 4.1 and of Theorem 1.2 must agree. In particular, on $L^2(D \times V)$, we must have that $\varphi = \tilde{\phi}$ and $\tilde{\varphi} = \phi$. Indeed, this can be shown rigorously in the spirit of the proof of Theorem 3.2.

4.3 The Notion of Criticality

Before turning to the proof of Theorem 4.1, let us enter into a discussion concerning the notion of *criticality* for the NBP, which is now afforded to us thanks to Theorem 4.1. Roughly speaking, since

$$\mathbb{E}_{\delta_{(r,\upsilon)}}[X_t[1]] = \psi_t[1](r,\upsilon) \sim \mathrm{e}^{\lambda_* t}\langle \tilde{\varphi}, 1\rangle \varphi(r,\upsilon), \qquad r \in D, \upsilon \in D,$$

as $t \to \infty$, we can think of λ_* as the rate of growth of the total number of particles in the NBP at time t. Moreover, we can think of this total mass as being distributed across the space-velocity domain $D \times V$ in a way that is proportional to $\tilde{\varphi}$. Indeed, the fact that $\tilde{\varphi}$ satisfies

$$\int_{D\times V} \mathbb{E}_{\delta_{(r,\upsilon)}}[X_t[g]]\,\tilde{\varphi}(r,\upsilon)\,\mathrm{d}r\,\mathrm{d}\upsilon = \langle \tilde{\varphi}, \psi_t[g]\rangle = \mathrm{e}^{\lambda_* t}\langle \tilde{\varphi}, g\rangle,$$
$$t \geq 0,\ g \in B^+(V \times D),$$

tells us that starting the NBP from a single particle that is distributed according to a normalised version of the density $\tilde{\varphi}$ will result in an average number of particles $\mathrm{e}^{\lambda_* t}$ that are scattered in space according to $\tilde{\varphi}$.

It seems quite clear that λ_* now parameterises the overall growth rate of the NBP and, in particular, the sign of λ_* gives us a notion of criticality. In particular, we say that the NBP is

$$\text{subcritical/critical/supercritical} \quad \text{when} \quad \lambda_* < 0/\lambda_* = 0/\lambda_* > 0.$$

Below we give some insights into the significance of the path behaviour of the NBP, as opposed to its mean behaviour, which we will demonstrate in forthcoming chapters:

Subcritical: The average number of particles in our NBP decays exponentially to zero at rate λ_*; however, we will see later in Chap. 6 that the actual number of particles becomes zero almost surely.

Critical: The average number of particles in our NBP remains constant; however, we will see later in Chap. 6 that, just as in the subcritical case, the actual number of particles becomes zero almost surely.

Supercritical: The average number of particles in our NBP grows exponentially at rate λ_* on the event of survival that occurs with probability in $(0, 1)$. We will see later in Chap. 12 that the *actual* number of particles almost surely scales exponentially at rate λ_*.

In the supercritical setting, it is clear that survival occurs with probability strictly less than one because, with positive probability, each neutron in any initial configuration may simply move to the boundary and be killed before ever having the chance to scatter or undergo fission.

In terms of reactor physics, the above stochastic behaviour provides an elementary view of how nuclear fission behaves in reactors. The ideal scenario is to keep a nuclear reactor held in a state of balance between fission and absorption as otherwise the reaction fizzles out (subcritical) or grows out of control (supercritical). Interestingly, nuclear reactor operators know extremely well that it is never possible to hold a reactor precisely in a critical state, as the activity of the fission will begin to decay. Instead, it is necessary to put the reactor into a state of slight supercriticality and, with the slow exponential growth of fission, periodically insert control rods to bring fission activity down to acceptable levels.[2]

4.4 Proof of the Perron–Frobenius Asymptotic

This section is dedicated to the proof of Theorem 4.1, and thus we assume the assumptions of Theorem 4.1 are in force. As alluded to earlier, our approach to proving Theorem 4.1 will be to extract the existence of the eigentriple λ_*, φ, and $\tilde{\varphi}$ for the expectation semigroup $(\psi_t, t \geq 0)$ from the existence of a similar triple for the semigroup $(\mathtt{P}^{\dagger}_t, t \geq 0)$, defined in (4.7). Indeed, from (4.7), it is clear that the eigenvalue of the former differs from the eigenvalue of the latter only by the constant $\overline{\beta}$. Moreover, the core of our proof relies on Theorem 2.2.

[2] In fact, this is still an over-simplification of what is really happening. There are many other considerations that pertain to the presence of fast and slow neutrons, the influence of fission by-products, and thermal hydraulics among several other influencing physical processes.

The proof is essentially broken into three large blocks of calculations. The first two address the respective proofs that assumptions (A1) and (A2) are satisfied, so that the conclusions of Theorem 2.2 hold. For the third block of calculations, we will prove the stated regularity properties of φ and $\tilde{\varphi}$; namely that φ is uniformly bounded away from 0 on each compactly embedded subset of $D \times V$ and that η admits a positive bounded density with respect to the Lebesgue measure on $D \times V$.

In order to pursue this agenda, we start by introducing two alternative assumptions to (A1) and (A2):

There exists an $\varepsilon > 0$ such that:

(B1) **The set** $D_\varepsilon := \{r \in D : \inf_{y \in \partial D} |r - y| > \varepsilon \mathtt{v}_{\max}\}$ **is non-empty and connected.**

(B2) **There exist** $0 < s_\varepsilon < t_\varepsilon$ **and** $\gamma > 0$ **such that, for all** $r \in D \backslash D_\varepsilon$, **there is a measurable set** $K_r \subset V$ **satisfying** $\mathrm{Vol}(K_r) \geq \gamma > 0$ **and, for all** $\upsilon \in K_r$, $r + \upsilon s \in D_\varepsilon$ **for every** $s \in [s_\varepsilon, t_\varepsilon]$ **and** $r + \upsilon s \notin \partial D$ **for all** $s \in [0, s_\varepsilon]$.

It is easy to verify that (B1) and (B2) are implied when we assume that D is a non-empty and convex, for example. They are also satisfied if the boundary of D is a smooth, connected, compact manifold and ε is sufficiently small. In geometrical terms, (B2) means that each of the sets

$$L_r := \left\{ z \in \mathbb{R}^3 : \frac{\|z - r\|}{\|\upsilon\|} \in [s_\varepsilon, t_\varepsilon], \upsilon \in K_r \right\}, \qquad r \in D \backslash D_\varepsilon \tag{4.11}$$

is included in D_ε and has Lebesgue measure at least $\gamma(t_\varepsilon^2 - s_\varepsilon^2)/2$. Roughly speaking, for each $r \in D$ that is within $\varepsilon \mathtt{v}_{\max}$ of the boundary ∂D, L_r is the set of points from which one can issue a neutron with a velocity chosen from $\upsilon \in K_r$ such that (ignoring scattering and fission) we can ensure that it passes through $D \backslash D_\varepsilon$ during the time interval $[s_\varepsilon, t_\varepsilon]$.

Our proof of Theorem 4.1 thus consists of proving that assumptions (B1) and (B2) imply assumptions (A1) and (A2) of Theorem 2.2.

Verification of (A1)

We begin by considering several technical lemmas. The first is a straightforward consequence of D being a bounded subset of $\mathbb{R}^3$ and so does not deserve a proof.

Lemma 4.2 *Let $B(r, \upsilon)$ be the ball in $\mathbb{R}^3$ centred at r with radius υ:*

(i) There exist an integer $\mathtt{n} \geq 1$ and $r_1, \ldots, r_n \in D_\varepsilon$ such that

$$D_\varepsilon \subset \bigcup_{i=1}^{\mathtt{n}} B(r_i, \mathtt{v}_{\max}\varepsilon/32)$$

and $D_\varepsilon \cap B(r_i, \mathtt{v}_{\mathtt{max}}\varepsilon/32) \neq \emptyset$ *for each* $i \in \{1, \dots, \mathfrak{n}\}$.

(ii) *For all* $r, r' \in D_\varepsilon$, *there exist* $m \leq n$ *and* $i_1, \dots, i_m$ *distinct in* $\{1, \dots, \mathfrak{n}\}$ *such that* $r \in B(r_{i_1}, \mathtt{v}_{\mathtt{max}}\varepsilon/32)$, $r' \in B(r_{i_m}, \mathtt{v}_{\mathtt{max}}\varepsilon/32)$, *and for all* $1 \leq j \leq m-1$, $B(r_{i_j}, \mathtt{v}_{\mathtt{max}}\varepsilon/32) \cap B(r_{i_{j+1}}, \mathtt{v}_{\mathtt{max}}\varepsilon/32) \neq \emptyset$.

Heuristically speaking, the above lemma ensures that there is a universal covering of D_ε by the balls $B(r_i, \mathtt{v}_{\mathtt{max}}\varepsilon/32)$, $1 \leq i \leq \mathfrak{n}$, such that, between any two points r and r' in D_ε, there is a sequence of overlapping balls

$$B(r_{i_1}, \mathtt{v}_{\mathtt{max}}\varepsilon/32), \cdots, B(r_{i_m}, \mathtt{v}_{\mathtt{max}}\varepsilon/32)$$

that one may pass through in order to get from r to r'.

The next lemma provides a minorisation of the law of (R_t, Υ_t) under $\mathbf{P}^\dagger$. In the statement of the lemma, we use $\mathtt{dist}(r, \partial D)$ for the distance of r from the boundary ∂D.

Define $\underline{\alpha} = \inf_{r\in D, \upsilon\in V} \alpha(r, \upsilon) > 0$ and $\underline{\pi} = \inf_{r\in D, \upsilon, \upsilon'\in V} \pi(r, \upsilon, \upsilon')$. We will also similarly write $\overline{\alpha}$ and $\overline{\pi}$ with obvious meanings. We note that due to the assumption (H1) we have $\overline{\alpha} < \infty$ and $\overline{\pi} < \infty$, and hence, combining this with the fact that we assumed $\inf_{r,\in D, \upsilon, \upsilon'\in V} \alpha(r, \upsilon)\pi(r, \upsilon, \upsilon') > 0$, it follows that

$$\underline{\alpha} = \frac{1}{\overline{\pi}} \inf_{r\in D, \upsilon\in V} \alpha(r, \upsilon)\overline{\pi} \geq \frac{1}{\overline{\pi}} \inf_{r\in D, \upsilon, \upsilon'\in V} \alpha(r, \upsilon)\pi(r, \upsilon, \upsilon') > 0,$$

and a similar calculation shows that $\underline{\pi} > 0$.

Lemma 4.3 *For all* $r \in D$, $\upsilon \in V$, *and* $t > 0$ *such that* $\mathtt{v}_{\mathtt{max}}t < \mathtt{dist}(r, \partial D)$, *the law of* (R_t, Υ_t) *under* $\mathbf{P}^\dagger_{(r,\upsilon)}$, *defined in* (4.7), *satisfies*

$$\begin{aligned}
&\mathbf{P}^\dagger_{(r,\upsilon)}(R_t \in \mathrm{d}z, \Upsilon_t \in \mathrm{d}\upsilon) \\
&\geq \frac{C\mathrm{e}^{-\overline{\alpha}t}}{t^2}\left[\frac{t}{2}\left(\mathtt{v}_{\mathtt{max}}^2 - \left(\mathtt{v}_{\mathtt{min}} \vee \frac{|z-r|}{t}\right)^2\right)\right. \\
&\qquad \left. - |z-r|\left(\mathtt{v}_{\mathtt{max}} - \mathtt{v}_{\mathtt{min}} \vee \frac{|z-r|}{t}\right)\right]\mathbf{1}_{\{z\in B(r, \mathtt{v}_{\mathtt{max}}t)\}}\, \mathrm{d}z\, \mathrm{d}\upsilon,
\end{aligned} \tag{4.12}$$

where $C > 0$ *is a positive constant.*

Proof Fix $r_0 \in D$. Let J_k denote the kth jump time of (R_t, Υ_t) under $\mathbf{P}^\dagger_{(r,\upsilon)}$, and let Υ_0 be uniformly distributed on V. Assuming that $\mathtt{v}_{\mathtt{max}}t < \mathtt{dist}(r_0, \partial D)$, we first give a minorisation of the density of (R_t, Υ_t), with initial configuration (r_0, Υ_0), on the event $\{J_1 \leq t < J_2\}$. Note that, on this event, we have

$$R_t = r_0 + J_1\Upsilon_0 + (t - J_1)\Upsilon_{J_1},$$

where Υ_{J_1} is the velocity of the process after the first jump. Then

$$
\begin{aligned}
&\mathbf{E}^{\dagger}_{(r_0,\Upsilon_0)}[f(R_t,\Upsilon_t)\mathbf{1}_{\{J_1\le t<J_2\}}] \\
&= \int_0^t \mathrm{d}s \int_V \mathrm{d}\upsilon_0 \int_V \mathrm{d}\upsilon_1 \alpha(r_0+\upsilon_0 s,\upsilon_0)\mathrm{e}^{-\int_0^s \alpha(r_0+\upsilon_0 u,\upsilon_0)\mathrm{d}u}\mathrm{e}^{-\int_0^{t-s}\alpha(r_0+\upsilon_0 s+\upsilon_1 u,\upsilon_1)\mathrm{d}u} \\
&\qquad\qquad \times \pi(r_0+\upsilon_0 s,\upsilon_0,\upsilon_1) f(r_0+\upsilon_0 s+(t-s)\upsilon_1,\upsilon_1) \\
&\ge \underline{\alpha}\mathrm{e}^{-\overline{\alpha}t}\underline{\pi}\int_V \mathrm{d}\upsilon_1 \int_0^t \mathrm{d}s \int_V \mathrm{d}\upsilon_0 f(r_0+s\upsilon_0+(t-s)\upsilon_1,\upsilon_1),
\end{aligned} \tag{4.13}
$$

where we have used the bounds on α and π. We now make the change of variables $\upsilon_0 \mapsto (\rho_0,\theta_0,\varphi_0)$ and $\upsilon_1 \mapsto (\rho_1,\theta_1,\varphi_1)$ so that (4.13) becomes

$$
\begin{aligned}
&\mathbf{E}^{\dagger}_{(r_0,\Upsilon_0)}[f(R_t,\Upsilon_t)\mathbf{1}_{\{J_1\le t<J_2\}}] \\
&\ge C_1\underline{\alpha}\mathrm{e}^{-\overline{\alpha}t}\underline{\pi}\int_0^t \mathrm{d}s \int_{\upsilon_{\min}}^{\upsilon_{\max}} \mathrm{d}\rho_1 \int_0^{\pi}\mathrm{d}\varphi_1 \int_0^{2\pi}\mathrm{d}\theta_1 \int_{\upsilon_{\min}}^{\upsilon_{\max}}\mathrm{d}\rho_0 \int_0^{\pi}\mathrm{d}\varphi_0\int_0^{2\pi}\mathrm{d}\theta_0 \\
&\qquad f(r_0+\Theta_{\rho_0,\rho_1,\theta_1,\varphi_1}(s,\theta_0,\varphi_0),\widetilde{\Theta}(\rho_1,\theta_1,\varphi_1))\delta(\rho_0,\theta_0,\varphi_0)\delta(\rho_1,\theta_1,\varphi_1),
\end{aligned} \tag{4.14}
$$

where

$$
\Theta_{\rho_0,\rho_1,\theta_1,\varphi_1}(s,\theta_0,\varphi_0) = \begin{bmatrix} s\rho_0\sin\varphi_0\cos\theta_0 + (t-s)\rho_1\sin\varphi_1\cos\theta_1 \\ s\rho_0\sin\varphi_0\sin\theta_0 + (t-s)\rho_1\sin\varphi_1\sin\theta_1 \\ s\rho_0\cos\varphi_0 + (t-s)\rho_1\cos\varphi_1 \end{bmatrix} \tag{4.15}
$$

represents the spatial variable $s\upsilon_0 + (t-s)\upsilon_1$ in polar coordinates,

$$
\widetilde{\Theta}(\rho_1,\theta_1,\varphi_1) = \begin{bmatrix} \rho_1\sin\varphi_1\cos\theta_1 \\ \rho_1\sin\varphi_1\sin\theta_1 \\ \rho_1\cos\varphi_1 \end{bmatrix} \tag{4.16}
$$

represents υ_1 in polar coordinates,

$$
\delta(\rho,\theta,\varphi) = \rho^2\sin\varphi, \tag{4.17}
$$

is the determinant of the Jacobian matrix for the change of variables from Cartesian to polar coordinates, and C_1 is an unimportant normalising constant.

For fixed ρ_0, ρ_1, θ_1, and φ_1, we first consider the part of (4.14) given by

$$(s,\theta_0,\varphi_0)\\ \mapsto \int_0^t \mathrm{d}s \int_0^\pi \mathrm{d}\varphi_0 \int_0^{2\pi} \mathrm{d}\theta_0\\ f(r_0+\Theta_{\rho_0,\rho_1,\theta_1,\varphi_1}(s,\theta_0,\varphi_0), \widetilde{\Theta}(\rho_1,\theta_1,\varphi_1))\delta(\rho_0,\theta_0,\varphi_0). \tag{4.18}$$

The Jacobian of $\Theta_{\rho_0,\rho_1,\theta_1,\varphi_1}$, as a function of (s,θ_0,φ_0), is given by

$$\begin{bmatrix} \rho_0\cos\theta_0\sin\varphi_0-\rho_1\cos\theta_1\sin\varphi_1 & -s\rho_0\sin\theta_0\sin\varphi_0 & s\rho_0\cos\varphi_0\cos\theta_0 \\ \rho_0\sin\theta_0\sin\varphi_0-\rho_1\sin\theta_1\sin\varphi_1 & s\rho_0\cos\theta_0\sin\varphi_0 & s\rho_0\cos\varphi_0\sin\theta_0 \\ \rho_0\cos\varphi_0-\rho_1\cos\varphi_1 & 0 & -s\rho_0\sin\varphi_0 \end{bmatrix},$$

whose determinant $\mathtt{det}(D_{\rho_0,\rho_1,\theta_1,\varphi_1}(s,\theta_0,\varphi_0))$ satisfies

$$\frac{\delta(\rho_0,\theta_0,\varphi_0)}{\mathtt{det}(D_{\rho_0,\rho_1,\theta_1,\varphi_1}(s,\theta_0,\varphi_0))} \geq \frac{1}{4s^2\mathtt{v}_{\mathtt{max}}^3} \geq \frac{1}{4t^2\mathtt{v}_{\mathtt{max}}^3}, \qquad s\leq t.$$

We thus have the following lower bound for (4.18)

$$\frac{1}{4t^2\mathtt{v}_{\mathtt{max}}^3}\int_0^t \mathrm{d}s\int_0^\pi \mathrm{d}\varphi_0\int_0^{2\pi}\mathrm{d}\theta_0 f(r_0+\Theta_{\rho_0,\rho_1,\theta_1,\varphi_1}(s,\theta_0,\varphi_0),\widetilde{\Theta}(\rho_1,\theta_1,\varphi_1)) \tag{4.19}$$
$$\times\,\mathtt{det}(D_{\rho_0,\rho_1,\theta_1,\varphi_1}(s,\theta_0,\varphi_0)).$$

Making another change of variables $(s,\theta_0,\varphi_0)\mapsto r\in\mathbb{R}^3$ and using the fact that, regardless of the values of ρ_1, θ_1, and φ_1, $\Theta_{\rho_0,\rho_1,\theta_1,\varphi_1}$ maps $(0,t)\times(0,\pi)\times(0,2\pi)$ surjectively onto a set that contains $B(\rho_0 t)$, where $B(r)$ is the ball in $\mathbb{R}^3$ of radius r centred at the origin, (4.19), and hence, (4.18), is bounded below by

$$\frac{1}{4t^2\upsilon_{\mathtt{max}}^3}\int_{B(\rho_0 t)} f(r,\widetilde{\Theta}(\rho_1,\theta_1,\varphi_1))\mathrm{d}r. \tag{4.20}$$

Substituting this equation back into (4.14) and changing $(\rho_1,\theta_1,\varphi_1)$ back to Cartesian coordinates, we have

$$\mathbf{E}^{\dagger}_{(r_0,\Upsilon_0)}[f(R_t,\Upsilon_t)\mathbf{1}_{\{J_1\leq t<J_2\}}] \geq \frac{C_2\mathrm{e}^{-\overline{\alpha}t}}{t^2}\int_{\mathtt{v}_{\mathtt{min}}}^{\mathtt{v}_{\mathtt{max}}}\mathrm{d}\rho_0\int_{B(\rho_0 t)}\mathrm{d}r\int_V \mathrm{d}\upsilon_1 f(r,\upsilon_1), \tag{4.21}$$

where $C_2=\underline{\alpha\pi}C_1/(4\mathtt{v}_{\mathtt{max}}^3)$.

Now suppose we fix an initial configuration $(r_0,\upsilon_0)\in D\times V$, with $t\mathtt{v}_{\mathtt{max}}<\mathtt{dist}(r_0,\partial D)$. By considering the event $\{J_2\leq t<J_3\}$ and noting that the scattering kernel is bounded below by $\underline{\pi}$, we may apply the Markov property together

with (4.21) to the process at time J_1 before choosing the new velocity. Using the bounds on α and π as before, and recalling that Υ_0 is uniformly distributed, we have

$$\begin{aligned}
&\mathbf{E}^{\dagger}_{(r_0,\upsilon_0)}[f(R_t,\Upsilon_t)\mathbf{1}_{\{J_2\le t<J_3\}}] \\
&\quad\ge \int_0^t \mathrm{d}s\,\underline{\alpha}\mathrm{e}^{-\overline{\alpha}s}\underline{\pi}\mathbf{E}^{\dagger}_{(r_0+s\upsilon_0,\Upsilon_0)}[f(R_{t-s},\Upsilon_{t-s})\mathbf{1}_{\{J_1\le t-s<J_2\}}] \\
&\quad\ge \int_0^t \mathrm{d}s\,\underline{\alpha}\mathrm{e}^{-\overline{\alpha}s}\underline{\pi}\frac{C_2\mathrm{e}^{-\overline{\alpha}(t-s)}}{(t-s)^2}\int_V \mathrm{d}\upsilon_1\int_{\mathrm{v_{min}}}^{\mathrm{v_{max}}}\mathrm{d}\rho_0\int_{\rho_0(t-s)B}\mathrm{d}r f(r_0+s\upsilon_0+r,\upsilon_1) \\
&\quad\ge \frac{C_3\mathrm{e}^{-\overline{\alpha}t}}{t^2}\int_0^t \mathrm{d}s\int_V \mathrm{d}\upsilon_1\int_{\mathrm{v_{min}}}^{\mathrm{v_{max}}}\mathrm{d}\rho_0\int_{\rho_0(t-s)B}\mathrm{d}r f(r_0+s\upsilon_0+r,\upsilon_1) \\
&\quad= \frac{C_3\mathrm{e}^{-\overline{\alpha}t}}{t^2}\int_0^t \mathrm{d}s\int_V \mathrm{d}\upsilon_1\int_{\mathrm{v_{min}}}^{\mathrm{v_{max}}}\mathrm{d}\rho_0\int_{r_0+s\upsilon_0+\rho_0(t-s)B}\mathrm{d}y f(y,\upsilon_1),
\end{aligned} \tag{4.22}$$

where we have used the substitution $y = r_0 + s\upsilon_0 + r$ to obtain the final line and C_3 is another constant in $(0,\infty)$. Now note that, for $s \le \rho_0 t/(\rho_0 + \mathrm{v_{max}})$, we have $r_0 + B(\rho_0 t - (\rho_0 + \mathrm{v_{max}})s) \subset r_0 + s\upsilon_0 + B(\rho_0(t-s))$. Combining this with (4.22) and using Fubini's theorem, we have

$$\begin{aligned}
&\mathbf{E}^{\dagger}_{(r_0,\upsilon_0)}[f(R_t,\Upsilon_t)\mathbf{1}_{\{J_2\le t<J_3\}}] \\
&\quad\ge \frac{C_3\mathrm{e}^{-\overline{\alpha}t}}{t^2}\int_V \mathrm{d}\upsilon_1\int_{\mathrm{v_{min}}}^{\mathrm{v_{max}}}\mathrm{d}\rho_0 \\
&\qquad\qquad \int_{\mathbb{R}}\mathbf{1}_{\left\{0\le s\le\frac{\rho_0}{\rho_0+\mathrm{v_{max}}}t\right\}}\mathrm{d}s\int_{\mathbb{R}^3}\mathrm{d}y\mathbf{1}_{\{|y-r_0|\le\rho_0 t-(\rho_0+\mathrm{v_{max}})s\}}f(y,\upsilon_1) \\
&\quad= \frac{C_3\mathrm{e}^{-\overline{\alpha}t}}{t^2}\int_V \mathrm{d}\upsilon_1\int_{\mathrm{v_{min}}}^{\mathrm{v_{max}}}\mathrm{d}\rho_0\int_{\mathbb{R}}\mathrm{d}s\int_{\mathbb{R}^3}\mathrm{d}y\mathbf{1}_{\left\{0\le s\le\frac{\rho_0 t-|y-r_0|}{\rho_0+\mathrm{v_{max}}}\right\}}f(y,\upsilon_1) \\
&\quad= \frac{C_3\mathrm{e}^{-\overline{\alpha}t}}{t^2}\int_V \mathrm{d}\upsilon_1\int_{\mathrm{v_{min}}}^{\mathrm{v_{max}}}\mathrm{d}\rho_0\int_{\mathbb{R}^3}\mathrm{d}y\mathbf{1}_{\{|y-r_0|\le\rho_0 t\}}\left(\frac{\rho_0 t-|y-r_0|}{\rho_0+\mathrm{v_{max}}}\right)f(y,\upsilon_1).
\end{aligned} \tag{4.23}$$

We finally compute the integral with respect to $\rho_0 \in (\mathrm{v_{min}}, \mathrm{v_{max}})$. In order to do so, we first note that since $\rho_0 < \mathrm{v_{max}}$, the integrand in (4.23) is bounded below by

$$\frac{\rho_0 t - |y - r_0|}{2\mathrm{v_{max}}}.$$

Absorbing $1/2\mathrm{v_{max}}$ into the constant C_3, applying Fubini, and computing the ρ_0 integral yield

$$
\begin{aligned}
&\mathbf{E}^{\dagger}_{(r_0,\upsilon_0)}[f(R_t,\Upsilon_t)]\\
&\geq \frac{C_3\mathrm{e}^{-\overline{\alpha}t}}{t^2}\int_V \mathrm{d}\upsilon_1\int_{\mathbb{R}^3}\mathrm{d}y\Bigg[\frac{t}{2}\left(\mathtt{v}_{\mathtt{max}}^2-\left(\mathtt{v}_{\mathtt{min}}\vee\frac{|y-r|}{t}\right)^2\right)\\
&\qquad -|y-r|\left(\mathtt{v}_{\mathtt{max}}-\mathtt{v}_{\mathtt{min}}\vee\frac{|y-r|}{t}\right)\Bigg]\mathbf{1}_{\{|y-r_0|\leq \mathtt{v}_{\mathtt{max}}t\}}f(y,\upsilon_1),
\end{aligned}
\tag{4.24}
$$

as required. □

We now turn to the proof of (A1) under the assumptions of (B1) and (B2).

Proof *(That (A1) Holds)* We start by proving (A1) for initial configurations in $D_\varepsilon \times V$. To this end, fix $(r,\upsilon)\in D_\varepsilon\times V$. From Lemma 4.2, there exists an $i\in\{1,\dots,\mathfrak{n}\}$ such that $r\in B(r_i,\mathtt{v}_{\mathtt{max}}\varepsilon/32)\cap D_\varepsilon$. Then, for each $t\in[\varepsilon/2,\varepsilon)$, Lemma 4.3 yields

$$
\begin{aligned}
&\mathbf{P}^{\dagger}_{(r,\upsilon)}(R_t\in\mathrm{d}z,\Upsilon_t\in\mathrm{d}w)\\
&\qquad\geq \frac{C\mathrm{e}^{-\overline{\alpha}t}}{t^2}\Bigg[\frac{t}{2}\left(\mathtt{v}_{\mathtt{max}}^2-\left(\mathtt{v}_{\mathtt{min}}\vee\frac{|z-r|}{t}\right)^2\right)\\
&\qquad\qquad -|z-r|\left(\mathtt{v}_{\mathtt{max}}-\mathtt{v}_{\mathtt{min}}\vee\frac{|z-r|}{t}\right)\Bigg]\mathbf{1}_{\{z\in B(r,\mathtt{v}_{\mathtt{max}}t)\}}\,\mathrm{d}z\,\mathrm{d}w.
\end{aligned}
\tag{4.25}
$$

Now, if $j\in\{1,\dots\mathfrak{n}\}$ is such that $B(r_i,\mathtt{v}_{\mathtt{max}}\varepsilon/32)\cap B(r_j,\mathtt{v}_{\mathtt{max}}\varepsilon/32)\neq\emptyset$, the triangle inequality implies that $D_\varepsilon\cap(B(r_i,\mathtt{v}_{\mathtt{max}}\varepsilon/32)\cup B(r_j,\mathtt{v}_{\mathtt{max}}\varepsilon/32))\subset B(r,\mathtt{v}_{\mathtt{max}}\varepsilon/8)\subset B(r,\mathtt{v}_{\mathtt{max}}t)$, with the latter inclusion following from the fact that $t\in[\varepsilon/2,\varepsilon)$.

Hence, for $z\in B(r_i,\mathtt{v}_{\mathtt{max}}\varepsilon/32)\cup B(r_j,\mathtt{v}_{\mathtt{max}}\varepsilon/32)$ and $t\in[\varepsilon/2,\varepsilon)$, the density on the right-hand side of (4.25) is bounded below by a constant $C_\varepsilon>0$, which is independent of r, υ, i, and j. Hence,

$$
\begin{aligned}
&\mathbf{P}^{\dagger}_{(r,\upsilon)}(R_t\in\mathrm{d}z,\Upsilon_t\in\mathrm{d}w)\\
&\qquad\geq C_\varepsilon\mathbf{1}_{\{z\in D_\varepsilon\cap(B(r_i,\varepsilon/32)\cup B(r_j,\varepsilon/32))\}}\,\mathrm{d}z\,\mathrm{d}w,\qquad z\in D, w\in V.
\end{aligned}
\tag{4.26}
$$

Now let $t\geq(\mathfrak{n}+1)\varepsilon/2$. By writing $t=k\varepsilon/2+t'$, for some $k\geq\mathfrak{n}$ and $t'\in[\varepsilon/2,\varepsilon)$. We will demonstrate that a repeated application of (4.26) will lead to the inequality

$$
\mathbf{P}^{\dagger}_{(r,\upsilon)}(R_t\in\mathrm{d}z,\Upsilon_t\in\mathrm{d}w)\geq C_\varepsilon c_\varepsilon^k\mathbf{1}_{\{z\in D_\varepsilon\}}\mathrm{d}z\mathrm{d}w,\qquad z\in D, w\in V,
\tag{4.27}
$$

for $(r,\upsilon)\in D_\varepsilon\times V$, where $c_\varepsilon>0$ is another unimportant constant that depends only on ε and is defined in the following analysis.

To this end, we start by noting that, since $r\in D_\varepsilon$ and $\upsilon\in V$, there exist $i_0,i_1\in\{1,\dots,\mathfrak{n}\}$ such that $r\in B(r_{i_0},\mathtt{v}_{\mathtt{max}}\varepsilon/32)$ and $B(r_{i_0},\mathtt{v}_{\mathtt{max}}\varepsilon/32)\cap$

$B(r_{i_1}, \mathtt{v}_{\mathtt{max}}\varepsilon/32) \cap D_\varepsilon \neq \emptyset$. Applying (4.26) at time t' (recall that we have identified $t = k\varepsilon/2 + t'$ for some $k \geq \mathfrak{n}$), we obtain

$$
\begin{aligned}
&\mathbf{P}^{\dagger}_{(r,\upsilon)}(R_t \in \mathrm{d}z, \Upsilon_t \in \mathrm{d}w) \\
&= \mathbf{P}^{\dagger}_{(r,\upsilon)}(R_{t'+k\varepsilon/2} \in \mathrm{d}z, \Upsilon_{t'+k\varepsilon/2} \in \mathrm{d}w) \\
&\geq \mathbf{E}^{\dagger}_{(r,\upsilon)}\left[\mathbf{1}_{\{R_{t'} \in B(r_{i_1}, \mathtt{v}_{\mathtt{max}}\varepsilon/32)\cap D_\varepsilon, \Upsilon_{t'} \in V\}} \mathbf{P}^{\dagger}_{(R_{t'}, \Upsilon_{t'})}(R_{k\varepsilon/2} \in \mathrm{d}z, \Upsilon_{k\varepsilon/2} \in \mathrm{d}w)\right] \\
&= \int_{B(r_{i_1}, \mathtt{v}_{\mathtt{max}}\varepsilon/32)\cap D_\varepsilon} \int_V \mathbf{P}^{\dagger}_{(r',\upsilon')}(R_{k\varepsilon/2} \in \mathrm{d}z, \Upsilon_{k\varepsilon/2} \in \mathrm{d}w) \mathbf{P}^{\dagger}_{(r,\upsilon)}(R_{t'} \in \mathrm{d}r', \Upsilon_{t'} \in \mathrm{d}\upsilon') \\
&\geq C_\varepsilon \int_{B(r_{i_1}, \mathtt{v}_{\mathtt{max}}\varepsilon/32)\cap D_\varepsilon} \int_V \mathbf{P}^{\dagger}_{(r',\upsilon')}(R_{k\varepsilon/2} \in \mathrm{d}z, \Upsilon_{k\varepsilon/2} \in \mathrm{d}w) \\
&\qquad\qquad \times \mathbf{1}_{\{r' \in (B(r_{i_0}, \mathtt{v}_{\mathtt{max}}\varepsilon/32) \cup B(r_{i_1}, \mathtt{v}_{\mathtt{max}}\varepsilon/32))\cap D_\varepsilon\}} \mathrm{d}r'\mathrm{d}\upsilon' \\
&= C_\varepsilon \int_{B(r_{i_1}, \mathtt{v}_{\mathtt{max}}\varepsilon/32)\cap D_\varepsilon} \int_V \mathbf{P}^{\dagger}_{(r',\upsilon')}(R_{k\varepsilon/2} \in \mathrm{d}z, \Upsilon_{k\varepsilon/2} \in \mathrm{d}w)\mathrm{d}r'\mathrm{d}\upsilon'.
\end{aligned}
\tag{4.28}
$$

We now turn our attention to $\mathbf{P}^{\dagger}_{(r',\upsilon')}(R_{k\varepsilon/2} \in \mathrm{d}z, \Upsilon_{k\varepsilon/2} \in \mathrm{d}w)$, for $(r', \upsilon') \in (B(r_{i_1}, \mathtt{v}_{\mathtt{max}}\varepsilon/32) \cap D_\varepsilon) \times V$ and $k \geq \mathfrak{n}$. Thanks to Lemma 4.2, for all $i_{k+1} \in \{1, \dots, \mathfrak{n}\}$, there exist $i_2, \dots, i_k \in \{1, \dots, \mathfrak{n}\}$ such that $B(r_{i_j}, \varepsilon/32) \cap B(r_{i_{j+1}}, \varepsilon/32) \neq \emptyset$ for every $j \in \{1, \dots, k\}$. Note, here we see the importance of choosing $k \geq \mathfrak{n}$, to ensure the validity of the previous statement.

Applying (4.26) and following the same steps that lead to (4.28), we obtain

$$
\begin{aligned}
&\mathbf{P}^{\dagger}_{(r',\upsilon')}(R_{k\varepsilon/2} \in \mathrm{d}z, \Upsilon_{k\varepsilon/2} \in \mathrm{d}w) \\
&\geq C_\varepsilon \int_{B(r_{i_2}, \varepsilon/32)\cap D_\varepsilon} \int_V \mathbf{P}^{\dagger}_{(r'',\upsilon'')}(R_{(k-1)\varepsilon/2} \in \mathrm{d}z, \Upsilon_{(k-1)\varepsilon/2} \in \mathrm{d}w)\mathrm{d}r''\mathrm{d}\upsilon''.
\end{aligned}
\tag{4.29}
$$

Iterating this step a further $k-2$ times, we obtain

$$
\begin{aligned}
&\mathbf{P}^{\dagger}_{(r',\upsilon')}(R_{k\varepsilon/2} \in \mathrm{d}z, \Upsilon_{k\varepsilon/2} \in \mathrm{d}w) \\
&\geq C_\varepsilon c_\varepsilon^{k-2} \int_{B(r_{i_k}, \mathtt{v}_{\mathtt{max}}\varepsilon/32)\cap D_\varepsilon} \int_V \mathbf{P}^{\dagger}_{(r'',\upsilon'')}(R_{\varepsilon/2} \in \mathrm{d}z, \Upsilon_{\varepsilon/2} \in \mathrm{d}w)\mathrm{d}r''\mathrm{d}\upsilon'',
\end{aligned}
\tag{4.30}
$$

where $c_\varepsilon = C_\varepsilon \mathrm{Vol}(V) \min_{i=1,\dots,n} \mathrm{Vol}(B(r_i, \mathtt{v}_{\mathtt{max}}\varepsilon/32) \cap D_\varepsilon)$. Using this inequality to bound the right-hand side of (4.28) yields

$$\mathbf{P}^{\dagger}_{(r,\upsilon)}(R_t \in \mathrm{d}z, \Upsilon_t \in \mathrm{d}w)$$
$$\geq C_\varepsilon c_\varepsilon^{k-1} \int_{B(r_{i_k},\varepsilon/32)\cap D_\varepsilon} \int_V \mathbf{P}^{\dagger}_{(r',\upsilon')}(R_{\varepsilon/2} \in \mathrm{d}z, \Upsilon_{\varepsilon/2} \in \mathrm{d}w)\mathrm{d}r'\mathrm{d}\upsilon'. \tag{4.31}$$

We now apply (4.26) a final time at time $\varepsilon/2$ to obtain

$$\mathbf{P}^{\dagger}_{(r,\upsilon)}(R_t \in \mathrm{d}z, \Upsilon_t \in \mathrm{d}w) \geq C_\varepsilon c_\varepsilon^k \mathbf{1}_{\{z\in B(r_{i_{k+1}},\varepsilon/2)\cap D_\varepsilon\}}\,\mathrm{d}z\,\mathrm{d}w. \tag{4.32}$$

Since this inequality holds for every $i_{k+1} \in \{1, \dots, \mathfrak{n}\}$, it also follows that

$$\mathbf{P}^{\dagger}_{(r,\upsilon)}(R_t \in \mathrm{d}z, \Upsilon_t \in \mathrm{d}w) \geq C_\varepsilon c_\varepsilon^k \sup_{i_{k+1}\in\{1,\dots,\mathfrak{n}\}} \mathbf{1}_{\{z\in B(r_{i_{k+1}},\varepsilon/2)\cap D_\varepsilon\}}\,\mathrm{d}z\,\mathrm{d}w$$
$$\geq C_\varepsilon c_\varepsilon^k\,\mathbf{1}_{\{z\in D_\varepsilon\}}\,\mathrm{d}z\,\mathrm{d}w,$$

where the final line follows from Lemma 4.2 since $k+1 > \mathfrak{n}$. This is the lower bound claimed in (4.27).

Finally, noting that for any two events A, B, $\Pr(A|B) = \Pr(A \cap B)/\Pr(B) \geq \Pr(A \cap B)$, we have that for initial conditions $(r, \upsilon) \in D_\varepsilon \times V$, any $t_0 \geq (\mathfrak{n}+1)\varepsilon/2$ and ν equal to Lebesgue measure on $D_\varepsilon \times V$, there exists a constant $c_1 \in (0, \infty)$ such that

$$\mathbf{P}_{(r,\upsilon)}((R_{t_0}, \Upsilon_{t_0}) \in \cdot\ |t_0 < \mathtt{k}) \geq c_1\nu(\cdot),$$

as required by (A1).

We now prove (A1) for initial conditions in $(D\backslash D_\varepsilon) \times V$. Once again, we recall that assumptions (B1) and (B2) are in force. Choose $r \in D\backslash D_\varepsilon$, $\upsilon \in V$, and define the (deterministic) time

$$\kappa_{r,\upsilon}^{D\backslash D_\varepsilon} := \inf\{t > 0 : r + t\upsilon \notin D_\varepsilon\},$$

which is the time it would take a neutron released at r with velocity υ to hit the boundary of $D\backslash D_\varepsilon$ if no scatter or fission took place. Importantly, the boundary of $D\backslash D_\varepsilon$ is made up of the union $\partial D \cup \partial D_\varepsilon$. Note in particular that $\kappa_{r,\upsilon}^{D\backslash D_\varepsilon}$ is not a random time but entirely deterministic. We first consider the case $r + \kappa_{r,\upsilon}^{D\backslash D_\varepsilon}\upsilon \in \partial D_\varepsilon$, for which we have

$$\mathbf{P}^{\dagger}_{(r,\upsilon)}(R_{\kappa_{r,\upsilon}^{D\backslash D_\varepsilon}} \in \partial D_\varepsilon) \geq \mathrm{e}^{-\bar{\alpha}\kappa_{r,\upsilon}^{D\backslash D_\varepsilon}} \geq \mathrm{e}^{-\bar{\alpha}\mathrm{diam}(D)/\mathtt{v}_{\min}}. \tag{4.33}$$

Combining this with (4.27) and the Markov property, for all $t \geq (\mathfrak{n}+1)\varepsilon/2$

$$\mathbf{P}_{(r,\upsilon)}(R_{\kappa_{r,\upsilon}^{D\backslash D_\varepsilon}+t} \in \mathrm{d}z, \Upsilon_{\kappa_{r,\upsilon}^{D\backslash D_\varepsilon}+t} \in \mathrm{d}w | \kappa_{r,\upsilon}^{D\backslash D_\varepsilon} + t < \mathtt{k})$$

$$
\begin{aligned}
&\geq \mathbf{P}^{\dagger}_{(r,\upsilon)}(R_{\kappa_{r,\upsilon}^{D\setminus D_\varepsilon}+t} \in \mathrm{d}z, \Upsilon_{\kappa_{r,\upsilon}^{D\setminus D_\varepsilon}+t} \in \mathrm{d}w) \\
&\geq \mathrm{e}^{-\bar{\alpha}\mathrm{diam}(D)/\mathtt{v}_{\mathtt{min}}} C_\varepsilon c_\varepsilon^k \mathbf{1}_{\{z\in D_\varepsilon\}}\,\mathrm{d}z\,\mathrm{d}w,
\end{aligned} \tag{4.34}
$$

where $k \geq \mathfrak{n}$ is such that $t = k\varepsilon/2 + t'$ for some $t' \in [\varepsilon/2, \varepsilon)$.

On the other hand, suppose $r + \kappa_{r,\upsilon}^{D\setminus D_\varepsilon}\upsilon \in \partial D$. Then, recalling the assumptions (B1) and (B2), it follows that $\{J_1 < \kappa_{r,\upsilon}^{D\setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon), \Upsilon_{J_1} \in K_{r+\upsilon J_1}, J_2 > t_\varepsilon\} \subset \{R_{t_\varepsilon} \in D_\varepsilon, t_\varepsilon < \mathtt{k}\}$. Heuristically speaking, this is because if the first jump occurs before time $\kappa_{r,\upsilon}^{D\setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon)$, then the process has not hit the boundary, and there are still (at least) s_ε units of time left until t_ε. By then choosing the new velocity, Υ_{J_1}, from $K_{r+\upsilon J_1}$, thanks to the assumption (B1) and the remarks around (4.11), this implies that the process will remain in $D\setminus D_\varepsilon$ for s_ε units of time, at some point in time after which, it will move into D_ε, providing the process does not jump again before entering D_ε. Combining this with the usual bounds on α, and recalling from (B2) that $\mathrm{Vol}(K_r) > \gamma > 0$ for all $r \in D\setminus D_\varepsilon$ and $\upsilon \in V$, we have

$$
\begin{aligned}
\mathbf{P}_{(r,\upsilon)}(R_{t_\varepsilon} \in D_\varepsilon, t_\varepsilon < \mathtt{k}) &\geq \mathbf{P}^{\dagger}_{(r,\upsilon)}(J_1 < \kappa_{r,\upsilon}^{D\setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon), \Upsilon_{J_1} \in K_{r+\upsilon J_1}, J_2 > t_\varepsilon) \\
&\geq \underline{\pi}\gamma \mathrm{e}^{-\overline{\alpha}t_\varepsilon} \mathbf{P}^{\dagger}_{(r,\upsilon)}(J_1 < \kappa_{r,\upsilon}^{D\setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon)).
\end{aligned} \tag{4.35}
$$

Along with (4.27), this implies that, for all $r \in D\setminus D_\varepsilon$, $\upsilon \in V$ and $t \geq (\mathfrak{n}+1)\varepsilon/2$ such that $t + t_\varepsilon \geq \kappa_{r,\upsilon}^{D\setminus D_\varepsilon}$

$$
\begin{aligned}
&\mathbf{P}_{(r,\upsilon)}(R_{t+t_\varepsilon} \in \mathrm{d}z, \Upsilon_{t+t_\varepsilon} \in \mathrm{d}w | t + t_\varepsilon < \mathtt{k}) \\
&\geq \frac{\mathbf{P}_{(r,\upsilon)}(R_{t_\varepsilon} \in D_\varepsilon, t_\varepsilon < \mathtt{k}, R_{t+t_\varepsilon} \in \mathrm{d}z, \Upsilon_{t+t_\varepsilon} \in \mathrm{d}w)}{\mathbf{P}_{(r,\upsilon)}(t + t_\varepsilon < \mathtt{k})} \\
&= \frac{\mathbf{P}^{\dagger}_{(r,\upsilon)}(R_{t+t_\varepsilon} \in \mathrm{d}z; \Upsilon_{t+t_\varepsilon} \in \mathrm{d}w | R_{t_\varepsilon} \in D_\varepsilon, t_\varepsilon < \mathtt{k})\mathbf{P}_{(r,\upsilon)}(R_{t_\varepsilon} \in D_\varepsilon, t_\varepsilon < \mathtt{k})}{\mathbf{P}_{(r,\upsilon)}(t + t_\varepsilon < \mathtt{k})} \\
&\geq \inf_{r\in D_\varepsilon, \upsilon\in V} \mathbf{P}^{\dagger}_{(r,\upsilon)}(R_t \in \mathrm{d}z; \Upsilon_t \in \mathrm{d}w) \\
&\qquad\qquad \times \frac{\mathbf{P}^{\dagger}_{(r,\upsilon)}(J_1 < \kappa_{r,\upsilon}^{D\setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon))}{\mathbf{P}_{(r,\upsilon)}(t + t_\varepsilon < \mathtt{k})} \underline{\pi}\gamma \mathrm{e}^{-\overline{\alpha}t_\varepsilon} c_\varepsilon^k \mathbf{1}_{\{z\in D_\varepsilon\}}\,\mathrm{d}z\,\mathrm{d}w \\
&\geq \frac{\mathbf{P}^{\dagger}_{(r,\upsilon)}(J_1 < \kappa_{r,\upsilon}^{D\setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon))}{\mathbf{P}_{(r,\upsilon)}(t + t_\varepsilon < \mathtt{k})} \underline{\pi}\gamma \mathrm{e}^{-\overline{\alpha}t_\varepsilon} C_\varepsilon c_\varepsilon^k \mathbf{1}_{\{z\in D_\varepsilon\}}\,\mathrm{d}z\,\mathrm{d}w.
\end{aligned} \tag{4.36}
$$

Now, since we are considering the case $r + \kappa_{r,\upsilon}^{D\setminus D_\varepsilon}\upsilon \in \partial D$ and $t + t_\varepsilon \geq \kappa_{r,\upsilon}^{D\setminus D_\varepsilon}$, it follows that $\{t + t_\varepsilon < \mathtt{k}\} \subset \{J_1 < \kappa_{r,\upsilon}^{D\setminus D_\varepsilon}\}$. Then,

$$\frac{\mathbf{P}^{\dagger}_{(r,\upsilon)}(J_1 < \kappa^{D\backslash D_\varepsilon}_{r,\upsilon} \wedge (t_\varepsilon - s_\varepsilon))}{\mathbf{P}_{(r,\upsilon)}(t + t_\varepsilon < \mathtt{k})} \geq \frac{\mathbf{P}^{\dagger}_{(r,\upsilon)}(J_1 < \kappa^{D\backslash D_\varepsilon}_{r,\upsilon} \wedge (t_\varepsilon - s_\varepsilon))}{\mathbf{P}_{(r,\upsilon)}(J_1 < \kappa^{D\backslash D_\varepsilon}_{r,\upsilon})}$$
$$\geq \frac{1 - \mathrm{e}^{-\underline{\alpha}(\kappa^{D\backslash D_\varepsilon}_{r,\upsilon} \wedge (t_\varepsilon - s_\varepsilon))}}{1 - \mathrm{e}^{-\overline{\alpha}\kappa^{D\backslash D_\varepsilon}_{r,\upsilon}}}, \tag{4.37}$$

with the bound on the right-hand side above being itself bounded below by a constant that does not depend on (r, υ). Substituting this back into (4.36), this proves (A1) with ν taken as Lebesgue measure on $D_\varepsilon \times V$ as before, t_0 can be sufficiently taken as $(\mathfrak{n} + 1)\varepsilon/2 + \mathrm{diam}(D)/\mathtt{v}_{\mathtt{min}}$, and we may start with any initial configurations in $D\backslash D_\varepsilon \times V$. □

Verification of (A2)

In order to prove (A2), we require the following lemma.

Lemma 4.4 *For all $r \in D$ and $\upsilon \in V$, recalling that J_k denotes the k-th jump time of the process (R, Υ), we have*

$$\mathbf{P}^{\dagger}_{(r,\upsilon)}(J_7 < \mathtt{k}, R_{J_7} \in \mathrm{d}z) \leq C\mathbf{1}_{\{z\in D\}}\,\mathrm{d}z, \tag{4.38}$$

for some constant $C > 0$, and

$$\mathbf{P}^{\dagger}_{\nu}(J_1 < \mathtt{k}, R_{J_1} \in \mathrm{d}z) \geq c\mathbf{1}_{\{z\in D\}}\,\mathrm{d}z, \tag{4.39}$$

for another constant $c > 0$, where ν, from the proof of (A1), is Lebesgue measure on $D_\varepsilon \times V$.

Proof Let us first prove (4.38). We couple the neutron transport random walk in D with one on the whole of $\mathbb{R}^3$. Denote by $(\hat{R}_t, \hat{\Upsilon}_t)$ the neutron random walk in $\hat{D} = \mathbb{R}^3$, coupled with (R, Υ) such that $\hat{R}_t = R_t$ and $\hat{\Upsilon}_t = \Upsilon_t$ for all $t < \mathtt{k}$ and $(R_0, \Upsilon_0) = (\hat{R}_0, \hat{\Upsilon}_0) = (r, \upsilon)$, for $r \in D$, $\upsilon \in V$. Denote by $\hat{J}_1 < \hat{J}_2 < \ldots$ the jump times of $\hat{\Upsilon}_t$. Then for each $k \geq 1$ such that $J_k < \mathtt{k}$, we have $\hat{J}_k = J_k$. Due to the inequality

$$\mathbf{E}^{\dagger}_{(r,\upsilon)}[f(R_{J_7}); J_7 < \mathtt{k}] \leq \mathbf{E}_{(r,\upsilon)}[f(\hat{R}_{\hat{J}_7})], \qquad r \in D, \upsilon \in V, \tag{4.40}$$

we will consider the distribution of $\hat{R}_{\hat{J}_i}$ for $i \geq 2$. We first look at the case when $i = 2$. For $(r, \upsilon) \in D \times V$ and non-negative, bounded, measurable functions f,

$$\mathbf{E}_{(r,\upsilon)}[f(\hat{R}_{\hat{J}_2})] = \mathbf{E}_{(r,\upsilon)}[f(r + \upsilon\hat{J}_1 + \hat{\Upsilon}_{\hat{J}_1}(\hat{J}_2 - \hat{J}_1))]$$

$$\leq \bar{\alpha}^2 \bar{\pi} \int_0^\infty \mathrm{d}j_1 \int_V \mathrm{d}\upsilon_1 \int_0^\infty \mathrm{d}j_2 \mathrm{e}^{-\underline{\alpha}(j_1+j_2)} f(r + \upsilon j_1 + \upsilon_1 j_2). \tag{4.41}$$

For j_1 fixed, we consider the integrals over υ_1 and j_2 in (4.41). Making the change of variables $\upsilon_1 \mapsto (\rho, \varphi, \theta)$, we have

$$\begin{aligned}
&\int_V \mathrm{d}\upsilon_1 \int_0^\infty \mathrm{d}j_2 \mathrm{e}^{-\underline{\alpha} j_2} f(r + \upsilon j_1 + \upsilon_1 j_2) \\
&\quad \leq \int_{\mathtt{v}_{\min}}^1 \mathrm{d}\rho \int_0^{2\pi} \mathrm{d}\theta \int_0^\pi \mathrm{d}\varphi \int_0^\infty \mathrm{d}j_2 \mathrm{e}^{-\underline{\alpha} j_2} f\left(r + \upsilon j_1 + \widetilde{\Theta}(\rho j_2, \theta, \varphi)\right) \rho^2 \sin\varphi,
\end{aligned} \tag{4.42}$$

where $\widetilde{\Theta}$ was defined in (4.16). Now making the substitution $u = \rho j_2$ in (4.42),

$$\begin{aligned}
&\int_V \mathrm{d}\upsilon_1 \int_0^\infty \mathrm{d}j_2 \mathrm{e}^{-\underline{\alpha} j_2} f(r + \upsilon j_1 + \upsilon_1 j_2) \\
&\quad \leq \int_{\mathtt{v}_{\min}}^{\mathtt{v}_{\max}} \mathrm{d}\rho \int_0^{2\pi} \mathrm{d}\theta \int_0^\pi \mathrm{d}\varphi \int_0^\infty \mathrm{d}u \mathrm{e}^{-\underline{\alpha} u/\rho} f\left(r + \upsilon j_1 + \widetilde{\Theta}(u, \theta, \varphi)\right) \rho \sin\varphi \\
&\quad \leq C \int_0^{2\pi} \mathrm{d}\theta \int_0^\pi \mathrm{d}\varphi \int_0^\infty \mathrm{d}u \mathrm{e}^{-\underline{\alpha} u/\mathtt{v}_{\max}} f\left(r + \upsilon j_1 + \widetilde{\Theta}(u, \theta, \varphi)\right) \sin\varphi,
\end{aligned} \tag{4.43}$$

where $C = \mathtt{v}_{\max}(\mathtt{v}_{\max} - \mathtt{v}_{\min})$. Making a final change of variables $(u, \theta, \varphi) \mapsto x \in \mathbb{R}^3$, we have

$$\int_V \mathrm{d}\upsilon_1 \int_0^\infty \mathrm{d}j_2 \mathrm{e}^{-\underline{\alpha} j_2} f(r + \upsilon j_1 + \upsilon_1 j_2) \leq C \int_{\mathbb{R}^3} \mathrm{d}x\, f(r + \upsilon j_1 + x) \frac{\mathrm{e}^{-\underline{\alpha}|x|/\mathtt{v}_{\max}}}{|x|^2}. \tag{4.44}$$

Substituting this back into (4.41) yields

$$\mathbf{E}_{(r,\upsilon)}[f(\hat{R}_{\hat{J}_2})] \leq \bar{\alpha} K \int_0^\infty \mathrm{d}j_1 \mathrm{e}^{-\underline{\alpha} j_1} \int_{\mathbb{R}^3} \mathrm{d}x f(r + \upsilon j_1 + x) \frac{\mathrm{e}^{-\underline{\alpha}|x|/\mathtt{v}_{\max}}}{|x|^2}, \tag{4.45}$$

where $K = \bar{\alpha}\bar{\pi} C$. Iterating this process over the next five jumps of the process gives

$$\begin{aligned}
&\mathbf{E}_{(r,\upsilon)}[f(\hat{R}_{\hat{J}_7})] \\
&\leq \bar{\alpha} K^6 \int_0^\infty \mathrm{d}j_1 \mathrm{e}^{-\underline{\alpha} j_1} \int_{\mathbb{R}^3} \mathrm{d}x_1 \ldots \int_{\mathbb{R}^3} \mathrm{d}x_6 f(r + \upsilon j_1 + x_1 + \cdots + x_6) g(x_1) \ldots g(x_6),
\end{aligned} \tag{4.46}$$

where $g(x) = \mathrm{e}^{-\underline{\alpha}|x|/\mathtt{v}_{\max}}/|x|^2$, $x \in \mathbb{R}^3$. Now, $g \in L^p(\mathbb{R}^3)$ for each $p < 3/2$ so that, in particular, $g \in L^{6/5}(\mathbb{R}^3)$. Hence, repeatedly applying Young's inequality implies that the six-fold convolution $*^6 g \in L^\infty(\mathbb{R}^3)$. (The reader will note that this is the fundamental reason we have focused our calculations around the 7th jump time J_7, rather than it being an arbitrary choice.) Making the substitution $x = x_1 + \cdots + x_6$,

$$\mathbf{E}^\dagger_{(r,\upsilon)}[f(\hat{R}_{\hat{J}_7})] \leq \bar{\alpha} K^6 \| *^6 g\|_\infty \int_0^\infty \mathrm{d}j_1 \mathrm{e}^{-\underline{\alpha} j_1} \int_{\mathbb{R}^3} \mathrm{d}x_1 \ldots \int_{\mathbb{R}^3} \mathrm{d}x_6 f(r + \upsilon j_1 + x). \tag{4.47}$$

Finally, setting $z = r + \upsilon j_1 + x$ yields

$$\mathbf{E}^\dagger_{(r,\upsilon)}[f(R_{J_7});\, J_7 < \mathtt{k}] \leq \mathbf{E}_{(r,\upsilon)}[f(\hat{R}_{\hat{J}_7})] \leq C' \int_{\mathbb{R}^3} f(z)\mathrm{d}z, \tag{4.48}$$

where $C' = \bar{\alpha} K^6 \| *^6 g\|_\infty$, which completes the proof of (4.38).

We now prove (4.39). For $r, r' \in \mathbb{R}^3$, let $[r, r']$ denote the line segment between r and r'. For all $f \in \mathscr{B}(\mathbb{R}^3)$, recalling the definition of ν from the proof of (A1) and using the usual bounds on α,

$$\begin{aligned}&\mathbf{E}_\nu[f(R_{J_1});\, J_1 < \mathtt{k}] \\ &\quad \geq \int_{D_\varepsilon} \frac{\mathrm{d}r}{\mathrm{Vol}(D_\varepsilon)} \int_V \frac{\mathrm{d}\upsilon}{\mathrm{Vol}(V)} \int_0^\infty \mathrm{d}s\, \mathbf{1}_{\{[r,r+s\upsilon]\subset D\}}\, \underline{\alpha} \mathrm{e}^{-\bar{\alpha}s} f(r + s\upsilon),\end{aligned} \tag{4.49}$$

where $\mathrm{Vol}(D_\varepsilon) = \int_{D_\varepsilon} \mathrm{d}r$ and $\mathrm{Vol}(V) = \int_V \mathrm{d}\upsilon$. Following a similar method to those employed in the proof of Lemma 4.3 and (4.38) and changing first to polar coordinates via $\upsilon \mapsto (\rho, \theta, \varphi)$, followed by the substitution $u = s\rho$, and finally changing back to Cartesian coordinates via $(u, \theta, \varphi) \mapsto x$, the right-hand side of (4.49) is bounded below by

$$C \int_{D_\varepsilon} \mathrm{d}r \int_{\mathbb{R}^3} \mathrm{d}x\, \mathbf{1}_{\{[r,r+x]\subset D\}} \frac{\underline{\alpha} \mathrm{e}^{-\bar{\alpha}s/\mathtt{v}_{\min}}}{|x|^2} f(r + x), \tag{4.50}$$

where $C > 0$ is a constant. Making a final substitution of $x = z - r$ yields

$$\begin{aligned}\mathbf{E}_\nu[f(R_{J_1});\, J_1 < \mathtt{k}] &\geq C \int_{D_\varepsilon} \mathrm{d}r \int_D \mathrm{d}z \mathbf{1}_{\{[r,z]\subset D\}} \frac{\underline{\alpha} \mathrm{e}^{-\bar{\alpha}|z-r|/\mathtt{v}_{\min}}}{|z - r|^2} f(z) \\ &\geq C \frac{\mathtt{v}_{\min}^2 \underline{\alpha} \mathrm{e}^{-\bar{\alpha}\mathrm{diam}(D)/\upsilon_{\min}^2}}{(\mathrm{diam}(D))^2} \int_{D_\varepsilon} \mathrm{d}r \int_D \mathrm{d}z \mathbf{1}_{\{[r,z]\subset D\}} f(z).\end{aligned} \tag{4.51}$$

For all $z \in D \backslash D_\varepsilon$, (B1) and the discussion thereafter now imply that

$$\int_{D_\varepsilon} \mathbf{1}_{\{[r,z]\subset D\}} \mathrm{d}r \geq \mathrm{Vol}(L_z) \geq \frac{\gamma}{2}(t_\varepsilon^2 - s_\varepsilon^2), \tag{4.52}$$

where s_ε and t_ε are defined in (B2), and L_z is defined in (4.11). On the other hand, for all $z \in D_\varepsilon$,

$$\int_{D_\varepsilon} \mathbf{1}_{\{[r,z]\subset D\}} \mathrm{d}r \geq \mathrm{Vol}(D_\varepsilon \cap B(r, \varepsilon)). \tag{4.53}$$

Since the map $z \mapsto \mathrm{Vol}(D_\varepsilon \cap B(z, \varepsilon))$ is continuous and positive on the compact set $\bar{D}_\varepsilon$, the latter equation is uniformly bounded below by a strictly positive constant. It then follows that for every $z \in D$, the integral $\int_{D_\varepsilon} \mathrm{d}r \mathbf{1}_{\{[r,z]\subset D\}}$ is bounded below by a positive constant. Using this to bound the right-hand side of (4.51) gives us the desired result. □

Proof *(That (A2) Holds)* Let $t \geq 7\mathrm{diam}(D)/\mathtt{v}_{\mathtt{min}}$ and note that on the event $\{\mathtt{k} > t\}$, we have $J_7 \leq 7\mathrm{diam}(D)/\mathtt{v}_{\mathtt{min}}$ almost surely. This inequality along with the strong Markov property implies that

$$\begin{aligned}\mathbf{P}_{(r,\upsilon)}(t < \mathtt{k}) &\leq \mathbf{E}^{\dagger}_{(r,\upsilon)}\left[\mathbf{1}_{\{J_7<t\}}\mathbf{P}_{(R_{J_7},\Upsilon_{J_7})}\left(t - s < \mathtt{k}\right)_{s=J_7}\right] \\ &\leq \mathbf{E}^{\dagger}_{(r,\upsilon)}\left[\mathbf{P}_{(R_{J_7},\Upsilon_{J_7})}\left(t - \frac{7\mathrm{diam}(D)}{\mathtt{v}_{\mathtt{min}}} < \mathtt{k}\right)\right].\end{aligned} \tag{4.54}$$

Since π is uniformly bounded above, conditional on $\{J_7 < \infty, R_{J_7} \in \mathrm{d}z\}$, the density of Υ_{J_7} is bounded above by $\overline{\pi}$ multiplied by Lebesgue measure on V. Combining this with (4.38) and (4.54), we obtain

$$\mathbf{P}_{(r,\upsilon)}(t < \mathtt{k}) \leq C' \int_D \int_V \mathbf{P}_{(z,w)}\left(t - \frac{7\mathrm{diam}(D)}{\mathtt{v}_{\mathtt{min}}} < \mathtt{k}\right) \mathrm{d}w\, \mathrm{d}z, \tag{4.55}$$

for some $C' \in (0, \infty)$. Similarly, for $t \geq \mathrm{diam}(D)/\mathtt{v}_{\mathtt{min}}$, Eq. (4.39), the fact that the inclusion $\{t < \mathtt{k}\} \subset \{J_1 \leq \mathrm{diam}(D)/\mathtt{v}_{\mathtt{min}}\}$, the strong Markov property, and the fact that π is uniformly bounded below entail that

$$\begin{aligned}\mathbf{P}_\nu(t < \mathtt{k}) &= \mathbf{E}^{\dagger}_\nu\left[\mathbf{1}_{\{J_1\leq \mathtt{k}\}}\mathbf{P}_{(R_{J_1},\Upsilon_{J_1})}\left(t - s < \mathtt{k}\right)_{s=J_1}\right] \\ &\geq \mathbf{E}^{\dagger}_\nu\left[\mathbf{1}_{\{J_1\leq \mathtt{k}\}}\mathbf{P}_{(R_{J_1},\Upsilon_{J_1})}\left(t < \mathtt{k}\right)\right] \\ &\geq c' \int_D \int_V \mathbf{P}_{(z,w)}(t < \mathtt{k})\, \mathrm{d}w\, \mathrm{d}z,\end{aligned}$$

for some $c' \in (0, \infty)$, where ν is Lebesgue measure on $D_\varepsilon \times V$. Putting (4.54) and (4.55) together, for all $t \geq 8\mathrm{diam}(D)/\mathtt{v}_{\mathtt{min}}$, we have

$$\mathbf{P}_{(r,\upsilon)}(t < \mathtt{k}) \le \frac{C'}{c'}\mathbf{P}_\nu\left(t - \frac{7\mathrm{diam}(D)}{\mathtt{v}_{\mathtt{min}}} < \mathtt{k}\right). \tag{4.56}$$

Now, recalling t_0 and ν from the proof of (A1), it follows from (A1) that

$$\mathbf{P}^\dagger_\nu((R_{t_0}, \Upsilon_{t_0}) \in \cdot) \ge c_1\mathbf{P}_\nu(t_0 < \mathtt{k})\nu(\cdot). \tag{4.57}$$

The event $\{t < \mathtt{k}\}$ occurs if the particle has either been killed on the boundary of D or if it has been absorbed by fissile material, which occurs at rate $\bar{\beta} - \beta$. Since t_0 and ν are fixed, and $\overline{\beta} - \beta \le \overline{\beta} + 1 < \infty$ by assumption, $\mathbf{P}_\nu(t_0 < \mathtt{k}) \ge K$ for some constant $K > 0$. Thus, keeping $t \ge 8\mathrm{diam}(D)/\mathtt{v}_{\mathtt{min}}$, using (4.57),

$$\begin{aligned}\mathbf{P}_\nu\left(t - \frac{7\mathrm{diam}(D)}{\mathtt{v}_{\mathtt{min}}} + t_0 < \mathtt{k}\right) &= \mathbf{E}_\nu\left[\mathbf{1}_{\{t_0<\mathtt{k}\}}\mathbf{P}^\dagger_{(R_{t_0},\Upsilon_{t_0})}\left(t - \frac{7\mathrm{diam}(D)}{\mathtt{v}_{\mathtt{min}}} < \mathtt{k}\right)\right]\\ &\ge \tilde{c}_1\mathbf{P}_\nu\left(t - \frac{7\mathrm{diam}(D)}{\mathtt{v}_{\mathtt{min}}} < \mathtt{k}\right),\end{aligned} \tag{4.58}$$

where $\tilde{c}_1 = Kc_1$.

Now define $N = \lceil 7\mathrm{diam}(D)/(\mathtt{v}_{\mathtt{min}}t_0)\rceil$. Then, for any $t > 0$, $t - 7\mathrm{diam}(D)/\mathtt{v}_{\mathtt{min}} + Nt_0 \ge t$ so that, trivially,

$$\mathbf{P}_\nu(t < \mathtt{k}) \ge \mathbf{P}_\nu\left(t - \frac{7\mathrm{diam}(D)}{\mathtt{v}_{\mathtt{min}}} + Nt_0 < \mathtt{k}\right). \tag{4.59}$$

Applying (4.58) N times implies that

$$\mathbf{P}_\nu(t < \mathtt{k}) \ge \tilde{c}_1^N\mathbf{P}_\nu\left(t - \frac{4\mathrm{diam}(D)}{\mathtt{v}_{\mathtt{min}}} < \mathtt{k}\right). \tag{4.60}$$

Combining this with (4.56) completes the proof of (A2). □

Regularity of φ and $\tilde{\varphi}$

We thus far proved that the conclusions of Theorem 2.2 are valid under our assumptions. In order to conclude that Theorem 4.1 holds true, it remains to prove that φ is uniformly bounded away from 0 on each compactly embedded subset of $D \times V$ and the existence of a positive bounded density for the left eigenmeasure η.

Lemma 4.5 *The right eigenfunction φ is uniformly bounded away from* 0 *on each compactly embedded subset of $D \times V$, and the probability measure η admits a positive density with respect to the Lebesgue measure on $D \times V$, which corresponds to the quantity $\tilde{\varphi}$ and which is uniformly bounded from above and a.e. uniformly bounded from below on each compactly embedded subset of $D \times V$.*

Proof For all $\varepsilon > 0$, we deduce from the eigenfunction property of φ (cf. Theorem 2.2) and from (4.27) that there exist a time $t_\varepsilon > 0$ and a constant $\tilde{C}_\varepsilon > 0$ such that

$$\varphi(r,\upsilon) = \mathrm{e}^{-\lambda_c t_\varepsilon}\mathtt{P}_{t_\varepsilon}[\varphi](r,\upsilon) \geq \mathrm{e}^{-\lambda t_\varepsilon}\tilde{C}_\varepsilon \int_{D_\varepsilon \times V} \varphi(z,w)\mathrm{d}z\mathrm{d}w > 0,$$

for all $(r,\upsilon) \in D_\varepsilon \times V$. It follows that φ is uniformly bounded away from 0 on each compactly embedded domain of $D \times V$.

Using the same notations as in the proof of Lemma 4.4, we consider the neutron transport random walk $(\hat{R}_t, \hat{\Upsilon}_t)$ in $\hat{D} = \mathbb{R}^3$, coupled with (R,Υ) such that $\hat{R}_t = R_t$ and $\hat{\Upsilon}_t = \Upsilon_t$ for all $t < \mathtt{k}$. We also denote by $\hat{J}_1 < \hat{J}_2 < \dots$ the jump times of $(\hat{\Upsilon}_t)_{t\geq 0}$. Let $T \geq 0$ be a random time independent of $(\hat{R},\hat{\Upsilon})$ with uniform law on $[\underline{T}, \bar{T}]$, where $\underline{T} < \bar{T}$ are fixed and $\underline{T} \geq 7\mathrm{diam}(D)/\mathtt{v}_{\mathtt{min}}$. We first prove that the law of $(\hat{R}_T, \hat{\Upsilon}_T)$ after the 7th jump admits a uniformly bounded density with respect to the Lebesgue measure. We conclude by using the coupling with (R,Υ) and the quasi-stationary property of η in (7.21).

For all $k \geq 7$ and for any positive, bounded, and measurable function f vanishing outside of $D \times V$, we have

$$\begin{aligned}
&\mathbf{E}[f(\hat{R}_T,\hat{\Upsilon}_T)\mathbf{1}_{\{\hat{J}_k \leq T < \hat{J}_{k+1}\}} \mid \hat{R}_0, \hat{\Upsilon}_0, T] \\
&= \mathbf{E}[f(\hat{R}_0 + \hat{J}_1\hat{\Upsilon}_0 + \dots + \hat{J}_k\hat{\Upsilon}_{k-1} + (T - \hat{J}_1 - \dots - \hat{J}_k)\hat{\Upsilon}_k, \hat{\Upsilon}_k) \\
&\quad \times \mathbf{1}_{\{\hat{J}_k \leq T < \hat{J}_{k+1}\}} \mid \hat{R}_0, \hat{\Upsilon}_0, T] \\
&= \int_0^T \mathrm{d}s_1\, \alpha(\hat{R}_0 + \upsilon_0 s_1, \upsilon_0)\mathrm{e}^{-\int_0^{s_1}\alpha(\hat{R}_0+\upsilon_0 u,\upsilon_0)\mathrm{d}u} \\
&\quad \times \int_V \mathrm{d}\upsilon_1 \pi(r_0 + \upsilon_0 s_1, \upsilon_0, \upsilon_1) \\
&\quad \times \int_0^{T-s_1} \mathrm{d}s_2\, \alpha(\hat{R}_0 + \upsilon_0 s_1 + \upsilon_1 s_2, \upsilon_1)\mathrm{e}^{-\int_0^{s_2}\alpha(\hat{R}_0+\upsilon_0 s_1+\upsilon_1 u,\upsilon_1)\mathrm{d}u} \\
&\quad \times \cdots \\
&\quad \times \int_V \mathrm{d}\upsilon_{k-1}\pi(\hat{R}_0 + \upsilon_0 s_1 + \dots + \upsilon_{k-2}s_{k-1}, \upsilon_{k-2}, \upsilon_{k-1}) \\
&\quad \times \int_0^{T-s_1-\dots-s_{k-1}} \mathrm{d}s_k\, \alpha(\hat{R}_0 + \upsilon_0 s_1 + \dots + \upsilon_{k-1}s_k, \upsilon_{k-1}) \\
&\qquad\qquad \mathrm{e}^{-\int_0^{s_k}\alpha(\hat{R}_0+\upsilon_0 s_1+\dots+\upsilon_{k-2}s_{k-1}+\upsilon_{k-1}u,\upsilon_{k-1})\mathrm{d}u} \\
&\quad \times \int_V \mathrm{d}\upsilon_k \pi(\hat{R}_0 + \upsilon_0 s_1 + \dots + \upsilon_{k-1}s_k, \upsilon_{k-1}, \upsilon_k)
\end{aligned}$$

$$\mathrm{e}^{-\int_0^{T-s_1-\cdots-s_k} \alpha(\hat{R}_0+\upsilon_0 s_1+\cdots+\upsilon_{k-1}s_k+\upsilon_k u, \upsilon_k)\mathrm{d}u}$$
$$\times f(\hat{R}_0 + \upsilon_0 s_1 + \cdots + \upsilon_{k-1}s_k + \upsilon_k(t - s_1 - \cdots - s_k), \upsilon_k).$$

Hence,

$$\begin{aligned}
&\mathbf{E}[f(\hat{R}_T, \hat{\Upsilon}_T)\mathbf{1}_{\{\hat{J}_k \le T < \hat{J}_{k+1}\}} \mid \hat{R}_0, \hat{\Upsilon}_0, T] \\
&\le \bar{\alpha}^k \bar{\pi}^k \mathrm{e}^{-T\underline{\alpha}} \int_0^T \mathrm{d}s_1 \int_V \mathrm{d}\upsilon_1 \cdots \int_0^{T-s_1-\cdots-s_{k-1}} \mathrm{d}s_k \int_V \mathrm{d}\upsilon_k \\
&\qquad f(\hat{R}_0 + \upsilon_0 s_1 + \cdots + \upsilon_{k-1}s_k + \upsilon_k(T - s_1 - \cdots - s_k), \upsilon_k).
\end{aligned}$$

Taking the expectation with respect to T, we obtain

$$\begin{aligned}
&\mathbf{E}[f(\hat{R}_T, \hat{\Upsilon}_T)\mathbf{1}_{\{\hat{J}_k \le T < \hat{J}_{k+1}\}} \mid \hat{R}_0, \hat{\Upsilon}_0] \\
&\le \frac{\bar{\alpha}^k \bar{\pi}^k}{\bar{T}} \int_0^{\bar{T}} \mathrm{d}t \int_0^t \mathrm{d}s_1 \int_V \mathrm{d}\upsilon_1 \cdots \int_0^{t-s_1-\cdots-s_{k-1}} \mathrm{d}s_k \int_V \mathrm{d}\upsilon_k \\
&\qquad \times f(\hat{R}_0 + \upsilon_0 s_1 + \cdots + \upsilon_{k-1}s_k + \upsilon_k(t - s_1 - \cdots - s_k), \upsilon_k).
\end{aligned}$$

Using the change of variable $(u_1, \ldots, u_k, u_{k+1}) = (s_1, \ldots, s_k, t - s_1 - \cdots - s_k)$ yields

$$\begin{aligned}
&\mathbf{E}[f(\hat{R}_T, \hat{\Upsilon}_T)\mathbf{1}_{\{\hat{J}_k \le T < \hat{J}_{k+1}\}} \mid \hat{R}_0, \hat{\Upsilon}_0] \\
&\le \frac{\bar{\alpha}^k \bar{\pi}^k}{\bar{T}} \int_{[0,\bar{T}]^{k+1}} \mathrm{d}u\, \mathbf{1}_{0 \le u_1+\cdots+u_{k+1} \le \bar{T}} \int_{V^k} \mathrm{d}\upsilon \\
&\qquad \times f(\hat{R}_0 + \upsilon_0 u_1 + \cdots + \upsilon_{k-1}u_k + \upsilon_k u_{k+1}, \upsilon_k).
\end{aligned}$$

The same approach as in Lemma 4.4 shows that there exists a constant $C > 0$ (which does not depend on $\hat{R}_0$ nor on $\hat{\Upsilon}_0$) such that, for all measurable functions $g : \mathbb{R}^3 \to [0, \infty)$,

$$\int_{[0,\bar{T}]^7} \mathrm{d}u \int_{V^6} \mathrm{d}\upsilon\, g(\hat{R}_0 + \hat{\Upsilon}_0 u_1 + \cdots + \upsilon_6 u_7) \le C \int_{\mathbb{R}^d} \mathrm{d}x g(x).$$

Hence,

$$\begin{aligned}
&\mathbf{E}[f(\hat{R}_T, \hat{\Upsilon}_T)\mathbf{1}_{\{\hat{J}_k \le T < \hat{J}_{k+1}\}} \mid \hat{R}_0, \hat{\Upsilon}_0] \\
&\qquad \le \frac{C\bar{\alpha}^k \bar{\pi}^k}{\bar{T}} \int_{[0,\bar{T}]^{k+1-7}} \mathrm{d}u\, \mathbf{1}_{0 \le u_8+\cdots+u_{k+1} \le \bar{T}} \int_{V^{k-6}} \mathrm{d}\upsilon \\
&\qquad\qquad \times \int_{\mathbb{R}^3} \mathrm{d}x f(x + \upsilon_7 u_8 + \cdots + \upsilon_k u_{k+1}, \upsilon_k)
\end{aligned}$$

$$= \frac{C\bar{\alpha}^k\bar{\pi}^k}{\bar{T}} \int_{[0,\bar{T}]^{k+1-7}} \mathrm{d}u\, \mathbf{1}_{0\le u_8+\cdots+u_{k+1}\le \bar{T}} \int_{V^{k-6}} \mathrm{d}\upsilon \int_{\mathbb{R}^3} \mathrm{d}y f(y,\upsilon_k)$$

$$= C\bar{\alpha}^k\bar{\pi}^k \mathrm{Vol}(V)^{k-8} \frac{\bar{T}^{k+1-8}}{(k+1-7)!} \int_D \mathrm{d}y \int_V \mathrm{d}\upsilon_k\, f(y,\upsilon_k),$$

where we used the change of variable $y = x + \upsilon_7 u_8 + \cdots + \upsilon_k u_{k+1}$ and the fact that f vanishes outside $D \times V$. Summing over $k \ge 7$, we deduce that there exists a constant $C' > 0$ (which only depends on $C, \bar{\alpha}, \bar{\pi}$ and $\bar{T}$) such that

$$\mathbf{E}[f(\hat{R}_T, \hat{\Upsilon}_T)\mathbf{1}_{\{\hat{J}_7\le T\}} \mid \hat{R}_0, \hat{\Upsilon}_0] \le C' \int_D \mathrm{d}y \int_V \mathrm{d}\upsilon\, f(y,\upsilon).$$

Similarly, as in the proof of (A2), we chose $\underline{T} \ge 7\mathrm{diam}(D)/\mathtt{v}_{\mathtt{min}}$, so that, on the event $\{\mathtt{k} > T\}$, we have $J_7 \le 7\mathrm{diam}(D)/\mathtt{v}_{\mathtt{min}} \le T$ almost surely. Hence, we obtain that, for any $(r_0, \upsilon_0) \in D \times V$,

$$\begin{aligned}
\mathbf{E}^{\dagger}_{(r_0,\upsilon_0)}[f(R_T,\Upsilon_T); T < \mathtt{k}] &= \mathbf{E}^{\dagger}_{(r_0,\upsilon_0)}[f(R_T,\Upsilon_T); T < \mathtt{k}, J_7 \le T] \\
&\le \mathbf{E}_{(r_0,\upsilon_0)}[f(\hat{R}_T,\hat{\Upsilon}_T); \hat{J}_7 \le T] \\
&\le C' \int_D \mathrm{d}y \int_V \mathrm{d}\upsilon\, f(y,\upsilon).
\end{aligned}$$

Integrating with respect to η and using the quasi-stationary property (7.21) and Fubini's theorem (recall that T and the process (R, Υ) are independent), we obtain

$$\begin{aligned}
\frac{1}{\bar{T}-\underline{T}} \int_{\underline{T}}^{\bar{T}} \mathrm{d}t\, e^{\lambda_c t}\eta[f] &= \frac{1}{\bar{T}-\underline{T}} \int_{\underline{T}}^{\bar{T}} \mathrm{d}t\, \mathbf{E}^{\dagger}_{\eta}[f(R_t,\Upsilon_t); t < \mathtt{k}] \\
&= \mathbf{E}^{\dagger}_{\eta}[f(R_T,\Upsilon_T); T < \mathtt{k}] \\
&\le C' \int_D \mathrm{d}y \int_V \mathrm{d}\upsilon\, f(y,\upsilon).
\end{aligned} \tag{4.61}$$

Since f was chosen arbitrarily, this proves that η admits a uniformly bounded density (from above) with respect to the Lebesgue measure on $D \times V$.

Finally, using the quasi-stationarity of η (7.21) and integrating inequality (4.26) with respect to η imply that (here the time t and the constants $k, C_\varepsilon, c_\varepsilon$ depend on ε as in inequality (4.27)), for all bounded measurable functions f on $D \times E$,

$$\begin{aligned}
e^{\lambda_c t} \int_{D\times V} f(x)\eta(\mathrm{d}x) &= \mathbf{E}^{\dagger}_{\eta}[f(R_t,\Upsilon_t); t < \mathtt{k}] \\
&\ge \eta(D_\varepsilon \times V)\, C_\varepsilon c_\varepsilon^k \int_{D_\varepsilon\times V} f(z,w)\,\mathrm{d}z\,\mathrm{d}w.
\end{aligned}$$

This implies that $\tilde{\varphi}$ is a.e. lower bounded by $e^{-\lambda_c t}\eta(D_\varepsilon \times V)\, C_\varepsilon c_\varepsilon^k$ on $D_\varepsilon \times V$. Since this inequality can be proved for any $\varepsilon > 0$ small enough, one deduces that, on any subset $D_\varepsilon \times V$ with $\varepsilon > 0$ and hence on any compactly embedded subset of $D \times V$, $\tilde{\varphi}$ is a.e. uniformly bounded away from zero. $\square$

4.5 Comments

The majority of this chapter is based on the ideas and calculations presented in Cox et al. [30] and Harris et al. [73]. The material presented in this chapter demarcates classical neutron transport theory from a more modern stochastic perspective that forms the basis of this entire book. As alluded to several times earlier, an important difference with pursuing stochastic representation of solutions on $(B(D \times E), \|\cdot\|)$ as opposed to $L^2(D \times V)$ is the ability to identify solutions both pointwise and via Feynman–Kac representations. As we will see in the forthcoming chapters, this lends itself well to stochastic analysis of the underlying Markov process, which, in some cases, is equivalent to pathwise statements about the underlying NBP.

Chapter 5
Pál–Bell Equation and Moment Growth

The previous chapter largely dealt with the relationship between the NTE and the NBP. The NTE is a linear equation and so there is limited information we can glean about the NBP from the NTE. Recall that the NBP is fundamentally our physical model of fission in an inhomogeneous material and so many questions will go beyond what linear equations can tell us. In this respect, our starting point is the Pál–Bell equation, a non-linear equation which captures a more complete picture of the stochastic behaviour of the NBP.

5.1 Pál–Bell Equation (PBE)

The so-called Pál–Bell equation is a special example of a general non-linear equation that is quite commonly used in the theory of spatial branching processes (as indeed we shall see in the second part of this book). In order to state the Pál–Bell equation, let us recall some basic facts of the NBP and introduce some more notation.

Recall that the way we described the NBP is via the point process

$$X_t(\cdot) := \sum_{i=1}^{N_t} \delta_{(r_i(t),\upsilon_i(t))}(\cdot), \qquad t \geq 0,$$

where

$$\{(r_i(t), \upsilon_i(t)), i = 1, \cdots, N_t\}, \qquad t \geq 0,$$

describes the configuration and number of particles in the system at time $t \geq 0$. In addition, recalling the notation

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8_5

$$X_t[f] := \int_{D\times V} f(r, \upsilon)X_t(\mathrm{d}r, \mathrm{d}\upsilon) = \sum_{i=1}^{N_t} f(r_i(t), \upsilon_i(t)), \quad t \geq 0,\ f \in B^+(D\times V),$$

we previously introduced

$$\mathrm{v}_t[f](r, \upsilon) := \mathbb{E}_{\delta_{(r,\upsilon)}}\left[\mathrm{e}^{-X_t[f]}\right], \qquad t \geq 0,\ f \in B^+(D \times E), \tag{5.1}$$

in (3.27), which was used to verify the branching Markov property of $(X_t, t \geq 0)$. When seen as operators on $B^+(D \times V)$, it is easy to see from the Markov branching property (3.28) that the family $(\mathrm{v}_t, t \geq 0)$ has the semigroup property.

Lemma 5.1 *Assume (H1) holds. For all $s, t \geq 0$ and $g \in B^+(D \times V)$, then* $\mathrm{v}_{t+s}[g] = \mathrm{v}_t[\mathrm{v}_s[g]]$.

Proof Just as with (2.3), this is a simple consequence of the Markov property. Indeed, recalling the decomposition given in (3.28), we can take expectations again and the result follows. □

The structure of the expectation that defines the family $(\mathrm{v}_t, t \geq 0)$ does not appear to be of the form (2.3), which we recall gave us our definition of an expectation semigroup for a general Markov process. Nonetheless, it is in fact consistent with (2.3) when one takes account of the fact that the NBP is a Markov process in $\mathcal{M}_c(D \times V)$. We can think of $\mathrm{v}_t[g]$ as taking the form $\mathbb{E}_{\delta_{(r,\upsilon)}}[G(X_t)]$, for the special class of functionals which are expressed as the negative exponential of an inner product with respect to X_t. As such, whilst Lemma 5.1 does not offer us the expectation semigroup property for all bounded functionals of X_t, it does offer us the expectation semigroup property for a dense family of functionals in the aforementioned class.

It is important to understand why the semigroup $(\mathrm{v}_t, t \geq 0)$ carries more information about the law of the NBP than the semigroup $(\psi_t, t \geq 0)$. From (5.1), we see that, for $g \in B^+(D \times V)$, $r \in D$, $\upsilon \in V$, and $t \geq 0$,

$$\psi_t[g](r, \upsilon) = -\frac{\mathrm{d}}{\mathrm{d}\theta}\mathbb{E}_{\delta_{(r,\upsilon)}}\left[\mathrm{e}^{-\theta X_t[g]}\right]\bigg|_{\theta=0} = -\frac{\mathrm{d}}{\mathrm{d}\theta}\mathrm{v}_t[\theta g](r, \upsilon)\bigg|_{\theta=0}. \tag{5.2}$$

Hence, at the very least, the analytical information contained in $(\psi_t, t \geq 0)$ is equally accessible from $(\mathrm{v}_t, t \geq 0)$.

Just as we have derived an evolution equation for $(\psi_t, t \geq 0)$, which was none other than the mild NTE (3.32), we would like to derive an evolution equation for $(\mathrm{v}_t, t \geq 0)$. Our approach will be similar to the derivation of (3.32), and however, we must be a bit more careful in one aspect. Unlike the additive functional $X_t[g]$ which is equal to zero on the extinction event, the multiplicative functional $\exp(-X_t[g])$ is equal to unity on the extinction event. For this reason, we need to slightly adjust how we work with $(\mathtt{U}_t, t \geq 0)$. We introduce $(\hat{\mathtt{U}}_t, t \geq 0)$ in place of $(\mathtt{U}_t, t \geq 0)$ with the understanding that, for $g \in B^+(D \times V)$,

$$\hat{\mathtt{U}}_t[g](r, \upsilon) = \begin{cases} g(r + \upsilon t, \upsilon) & \text{for } t < \kappa^D_{(r,\upsilon)} = \inf\{t > 0 : r + \upsilon t \notin D\}, \\ 1 & \text{otherwise.} \end{cases} \tag{5.3}$$

Similarly, $\hat{\mathtt{P}}_t$ is a slight adjustment of $\mathtt{P}_t$ which returns a value of 1 on the event of killing. We also need to introduce the non-linear operator, acting on

$$B^+_1(D \times V) = \{f \in B^+(D \times V) : \|f\| \leq 1\}, \tag{5.4}$$

given by

$$\mathtt{G}[f](r, \upsilon) = \sigma_{\mathtt{f}}(r, \upsilon)\mathcal{E}_{(r,\upsilon)}\left[\prod_{i=1}^{N} f(r, \upsilon_i) - f(r, \upsilon)\right]. \tag{5.5}$$

The operator $\mathtt{G}$ is called the *fission mechanism*. In a more general context of branching Markov processes, it is also known as the *branching mechanism*.

Lemma 5.2 (Pál–Bell Equation) *Assume (H1) holds. For $g \in B^+(D \times V)$, we have on $D \times V$*

$$\mathtt{v}_t[g] = \hat{\mathtt{U}}_t[\mathrm{e}^{-g}] + \int_0^t \mathtt{U}_s\left[\mathtt{S}\mathtt{v}_{t-s}[g] + \mathtt{G}[\mathtt{v}_{t-s}[g]]\right]\mathrm{d}s, \qquad t \geq 0. \tag{5.6}$$

Equivalently we have a second representation

$$\mathtt{v}_t[g] = \hat{\mathtt{P}}_t[\mathrm{e}^{-g}] + \int_0^t \mathtt{P}_s\left[\mathtt{G}[\mathtt{v}_{t-s}[g]]\right]\mathrm{d}s, \qquad t \geq 0. \tag{5.7}$$

Solutions to both (5.6) *and* (5.7)*, which are valued in* $[0, 1]$*, are unique.*

Proof The fundamental idea of the proof of the two equations (5.6) and (5.7) is to break the expectation in the definition (5.1) of $\mathtt{v}_t$ either on the first event (scattering or fission) to obtain (5.6) or just on the first fission event to obtain (5.7). We prove (5.6) and leave the derivation of (5.7) in the same fashion as an exercise.

Conditioning on the first event, we get

$$\begin{aligned} &\mathtt{v}_t[g](r, \upsilon) \\ &= \mathrm{e}^{-\int_0^{t\wedge\kappa^D_{(r,\upsilon)}} \sigma(r+\upsilon\ell,\upsilon)\mathrm{d}\ell}[\mathrm{e}^{-g(r+\upsilon t,\upsilon)}\mathbf{1}_{(t<\kappa^D_{(r,\upsilon)})} + \mathbf{1}_{(t\geq\kappa^D_{(r,\upsilon)})}] \\ &+ \int_0^t \mathbf{1}_{(s<\kappa^D_{(r,\upsilon)})}\sigma(r + \upsilon s, \upsilon)\mathrm{e}^{-\int_0^t \sigma(r+\upsilon\ell,\upsilon)\mathrm{d}\ell} \end{aligned}$$

$$\left(\frac{\sigma_{\mathtt{s}}(r+\upsilon s,\upsilon)}{\sigma(r+\upsilon s,\upsilon)}\int_V \mathrm{v}_{t-s}(r+\upsilon s,\upsilon')\pi_{\mathtt{s}}(r+\upsilon s,\upsilon,\upsilon')\mathrm{d}\upsilon' \right.$$
$$\left. + \frac{\sigma_{\mathtt{f}}(r+\upsilon s,\upsilon)}{\sigma(r+\upsilon s,\upsilon)}\mathcal{E}_{(r+\upsilon s,\upsilon)}\left[\prod_{i=1}^{N}\mathrm{v}_{t-s}(r+\upsilon s,\upsilon_i)\right]\right)\mathrm{d}s,$$

where we recall $\kappa^D_{(r,\upsilon)}$ was defined in (3.13) and $\sigma = \sigma_{\mathtt{f}} + \sigma_{\mathtt{s}}$. Appealing to Theorem 2.1, we can transform the multiplicative potential with rate σ into an additive potential and, together with some easy algebra, this gives us straight away (5.6).

The proof is completed as soon as we establish uniqueness. This is again a matter of an application of Grönwall's lemma in the spirit of the proof of uniqueness argument in Lemma 3.5. We leave the details to the reader. □

5.2 Many-to-Two Representation and Variance Evolution

The NTE tells us about the growth of the first moment and the PBE tells us about the law of the NBP. But the question remains how we can extract more specific information out of the latter to complement the former. In this section, we will look at the second moments of the NBP or equivalently the two-point correlation function for particles alive at time $t \geq 0$.

In order to do so, we define the operator

$$\mathscr{V}[f,g](r,\upsilon) = \mathcal{E}_{(r,\upsilon)}\left[\sum_{\substack{i,j=1\\ i\neq j}}^{N} f(r,\upsilon_i)g(r,\upsilon_j)\right], \tag{5.8}$$

for $f, g \in B^+(D\times V)$, $r \in D$, and $\upsilon \in V$. We will often abuse notation and write $\mathscr{V}[g](r,\upsilon)$ instead of $\mathscr{V}[g,g](r,\upsilon)$ for $g \in B^+(D\times V)$ and $r \in D$, $\upsilon \in V$. Note in particular that

$$\mathscr{V}[g](r,\upsilon) = \mathcal{E}_{(r,\upsilon)}\left[\mathsf{Z}[g(r,\cdot)]^2 - \mathsf{Z}[g^2(r,\cdot)]\right]. \tag{5.9}$$

Generally, for (5.9) to be finite, we need a second moment assumption on the number of neutrons produced at each fission event. This is automatically satisfied since we have at most $n_{\mathtt{max}}$ particles produced at fission, cf. (H4).

Lemma 5.3 (Many-to-Two) *Assume (H1) and (H4). Suppose that* $f, g \in B^+(D\times V)$. *Then*

$$\mathbb{E}_{\delta(r,\upsilon)}\Big[X_t[f]X_t[g]\Big] = \psi_t[fg](r,\upsilon) + \int_0^t \psi_s\Big[\sigma_{\mathtt{f}}\mathscr{V}[\psi_{t-s}[f], \psi_{t-s}[g]]\Big](r,\upsilon)\mathrm{d}s, \tag{5.10}$$

for $r \in D$, $\upsilon \in V$. In particular,

$$\mathbb{E}_{\delta(r,\upsilon)}\Big[X_t[f]^2\Big] = \psi_t[f^2](r,\upsilon) + \int_0^t \psi_s\Big[\sigma_{\mathtt{f}}\mathscr{V}[\psi_{t-s}[f]]\Big](r,\upsilon)\mathrm{d}s. \tag{5.11}$$

Proof Suppose that $h \in B^+(D\times V)$ and that $H_s(r,\upsilon)$ on $\mathbb{R}_+ \times D \times V$ is non-negative, continuous, and bounded. Then we claim that

$$\omega_t(r,\upsilon) := \psi_t[h](r,\upsilon) + \int_0^t \psi_s[H_s](r,\upsilon)\mathrm{d}s, \tag{5.12}$$

for $t \geq 0$, $r \in D$, and $\upsilon \in V$, uniquely solves the integral equation

$$\omega_t(r,\upsilon) = \mathtt{U}_t[h](r,\upsilon) + \int_0^t \mathtt{U}_s\Big[H_s + (\mathtt{S}+\mathtt{F})\omega_{t-s}\Big](r,\upsilon)\mathrm{d}s. \tag{5.13}$$

We only sketch the proof as the methods used should now be quite familiar to the reader. First note that we can otherwise write (5.12) as

$$\omega_t(r,\upsilon) := \mathbf{E}_{(r,\upsilon)}\left[\mathrm{e}^{\int_0^t \beta(R_s,\Upsilon_s)\mathrm{d}s} h(R_t,\Upsilon_t)\mathbf{1}_{(t<\tau^D)}\right] + \mathbf{E}_{(r,\upsilon)}\left[\int_0^{t\wedge\tau^D} H_s(R_s,\Upsilon_s)\mathrm{e}^{\int_0^s \beta(R_u,\Upsilon_u)\mathrm{d}u}\mathrm{d}s\right]. \tag{5.14}$$

Next condition the right-hand side of (5.14) on the first fission or scattering event, whichever comes first, thus generating a recursion for $(\omega_t, t\geq 0)$. An application of Theorem 2.1 then gives us (5.13). Finally, the uniqueness of the latter follows by a standard argument appealing to Grönwall's lemma.

To complete the proof of (5.10), it suffices to consider the case $f = g$, as the general form will follow from the polarisation identity

$$2\mathbb{E}_{\delta(r,\upsilon)}[X_t[f]X_t[g]] = \mathbb{E}_{\delta(r,\upsilon)}[X_t[f+g]^2] - \mathbb{E}_{\delta(r,\upsilon)}[X_t[f]^2] - \mathbb{E}_{\delta(r,\upsilon)}[X_t[g]^2].$$

To this end, denote $w_t(r,\upsilon) = \mathbb{E}_{\delta(r,\upsilon)}[X_t[g]^2]$, for $t \geq 0$, $r \in D$, $\upsilon \in V$. If we split $X_t[g]^2$ according to the NBPs that grow out of each of the offspring at the first fission event, say $X^{(i)}$, $i = 1, \cdots, N$, then, conditional on fission occurring at time $s \leq t$, we can write

$$X_t[g]^2 = \left(\sum_{i=1}^{N} X^{(i)}_{t-s}[g]\right)^2 = \sum_{\substack{i,j=1\\ i\neq j}}^{N} X^{(i)}_{t-s}[g] X^{(j)}_{t-s}[g] + \sum_{i=1}^{N} X^{(i)}_{t-s}[g]^2. \tag{5.15}$$

For convenience, write $D_t(r, \upsilon) = \exp(-\int_0^t \sigma(r + \ell\upsilon, \upsilon)\mathrm{d}\ell)$, for $r \in D$, $\upsilon \in V$. Using (5.15), we can now formally split the expectation of $X_t[g]^2$ on the first event, scattering or fission, and then apply Theorem 2.1. This gives us

$$\begin{aligned}
&w_t(r, \upsilon)\\
&= D_t(r, \upsilon)\mathtt{U}_t[g^2](r, \upsilon) + \int_0^t \mathtt{U}_s\left[\sigma_{\mathtt{s}} D_s \int_V w_{t-s}(r, \upsilon')\pi_{\mathtt{s}}(\cdot, \cdot, \upsilon')\mathrm{d}\upsilon'\right](r, \upsilon)\mathrm{d}s\\
&+ \int_0^t \mathtt{U}_s\left[\sigma_{\mathtt{f}} D_s \mathcal{E}_{(\cdot,\cdot)}\left[\sum_{\substack{i,j=1\\ i\neq j}}^{N} \psi_{t-s}[g](\cdot, \upsilon_i)\psi_{t-s}[g](\cdot, \upsilon_j) + \sum_{i=1}^{N} w_{t-s}(\cdot, \upsilon_i)\right]\right](r, \upsilon)\mathrm{d}s\\
&= \mathtt{U}_t[g^2](r, \upsilon) + \int_0^t \mathtt{U}_s[(\mathtt{S} + \mathtt{F})w_{t-s}](r, \upsilon)\mathrm{d}s + \int_0^t \mathtt{U}_s\left[\sigma_{\mathtt{f}} \mathscr{V}[\psi_{t-s}[g]]\right](r, \upsilon)\mathrm{d}s.
\end{aligned}$$

Using the representation of the solution to (5.13) but with $h = g^2$ and $H(s, \cdot, \cdot) = \sigma_{\mathtt{f}} \mathscr{V}[\psi_{t-s}[g]]$, we get (5.11), which we recall is sufficient to establish (5.10). □

A consequence of Theorem 5.3 is that we gain insight into the asymptotic evolution of the variance of the underlying NBP. More precisely, we have the following result which shows that in the supercritical case the second moment behaves like the square of the first moment, in the subcritical case, the second moment behaves like the first moment, and in the critical case, the second moment grows linearly.

Theorem 5.1 *Suppose the assumptions of Lemma 5.3 are met. Then, for any $g \in B^+(D \times V)$, we have the following asymptotic behaviour for the second moment in the supercritical, subcritical, and critical cases:*

(i) If $\lambda_* > 0$,

$$\lim_{t\to\infty} \mathrm{e}^{-2\lambda_* t}\mathbb{E}_{\delta_{(r,\upsilon)}}[X_t[g]^2] = \langle\tilde{\varphi}, g\rangle^2 \int_0^\infty \mathrm{e}^{-2\lambda_* s}\psi_s[\sigma_{\mathtt{f}} \mathscr{V}[\varphi]](r, \upsilon)\mathrm{d}s. \tag{5.16}$$

(ii) If $\lambda_* < 0$,

$$\lim_{t\to\infty} \mathrm{e}^{-\lambda_* t}\mathbb{E}_{\delta_{(r,\upsilon)}}[X_t[g]^2] = \left(\langle\tilde{\varphi}, g^2\rangle + \int_0^\infty \mathrm{e}^{-\lambda_* s}\langle\tilde{\varphi}, \sigma_{\mathtt{f}} \mathscr{V}\big[\psi_s[g]\big]\rangle\mathrm{d}s\right)\varphi(r, \upsilon). \tag{5.17}$$

(iii) *If* $\lambda_* = 0$,

$$\lim_{t\to\infty} \frac{1}{t}\mathbb{E}_{\delta_{(r,\upsilon)}}[X_t[g]^2] = \langle\tilde{\varphi}, g\rangle^2\langle\tilde{\varphi}, \sigma_{\mathtt{f}}\mathscr{V}[\varphi]\rangle\varphi(r,\upsilon). \tag{5.18}$$

Proof Appealing to the Perron–Frobenius Theorem 4.1, we have, for any $f \in B^+(D \times V)$ and $\delta > 0$, that there exist some constant $K \in (0,\infty)$ and some $t_0 = t_0(\delta)$ such that

$$\sup_{t\geq 0}\|\mathrm{e}^{-\lambda_* t}\psi_t[f]\| \leq K\|f\| \text{ and } \|\mathrm{e}^{-\lambda_* t}\psi_t[f]-\langle\tilde{\varphi}, f\rangle\varphi\| \leq \delta\|f\|, \quad \text{for all } t \geq t_0. \tag{5.19}$$

On the other hand, it follows from (5.9) that $\sigma_{\mathtt{f}}\mathscr{V}$ is a symmetric bilinear form, and hence

$$\begin{aligned}|\sigma_{\mathtt{f}}\mathscr{V}[f](r,\upsilon) - \sigma_{\mathtt{f}}\mathscr{V}[h](r,\upsilon)| =& |\sigma_{\mathtt{f}}\mathscr{V}[f-h, f+h](r,\upsilon)| \\ &\leq C\|f-h\|(\|f\| + \|h\|),\end{aligned} \tag{5.20}$$

where, in the last inequality, we have used the boundedness of $\sigma_{\mathtt{f}}$ from (H1) and, thanks to (H4), $C := \|\sigma_{\mathtt{f}}\|\, n_{\mathtt{max}} < \infty$. In particular, taking $h = 0$ yields

$$\|\sigma_{\mathtt{f}}\mathscr{V}[f]\| \leq C\|f\|^2. \tag{5.21}$$

Also, we clearly have

$$f, g \in B^+(D\times V),\ f \leq g \quad \Rightarrow \quad \psi_t[f] \leq \psi_t[g], \quad t \geq 0. \tag{5.22}$$

Let us look at the supercritical case, i.e., $\lambda_* > 0$. Note that the leading term $\psi_t[f^2]$ on the right-hand side of (5.11) is overscaled by $\mathrm{e}^{2\lambda_* t}$ and hence limits away to zero. This leaves us with considering the integral on the right-hand side of (5.10) scaled by $\mathrm{e}^{2\lambda_* t}$. Our objective is to extract the dominant growth rate from (5.11) as $t \to \infty$. To that end, we first split the integral in (5.11) into two parts,

$$\begin{aligned}\int_0^t \psi_s\Big[\sigma_{\mathtt{f}}\mathscr{V}[\psi_{t-s}[g]]\Big](r,\upsilon)\mathrm{d}s =& \int_{t-t_0}^t \psi_s\Big[\sigma_{\mathtt{f}}\mathscr{V}[\psi_{t-s}[g]]\Big](r,\upsilon)\mathrm{d}s \\ &+ \int_0^{t-t_0} \psi_s\Big[\sigma_{\mathtt{f}}\mathscr{V}[\psi_{t-s}[g]]\Big](r,\upsilon)\mathrm{d}s.\end{aligned} \tag{5.23}$$

Note that the first term on the right-hand side is of order $o(\mathrm{e}^{2\lambda_* t})$. Indeed,

$$\int_{t-t_0}^t \psi_s\Big[\sigma_{\mathtt{f}}\mathscr{V}[\psi_{t-s}[g]]\Big](r,\upsilon)\mathrm{d}s \leq \int_{t-t_0}^t \psi_s\Big[\sigma_{\mathtt{f}}\mathscr{V}\Big[K\|g\|\mathrm{e}^{\lambda_*(t-s)}\Big]\Big](r,\upsilon)\mathrm{d}s$$

$$
\begin{aligned}
&\leq \int_{t-t_0}^{t} \psi_s\Big[CK^2\|g\|^2 \mathrm{e}^{2\lambda_*(t-s)}\Big](r,\upsilon)\mathrm{d}s \\
&\leq CK^3\|g\|^2\mathrm{e}^{2\lambda_* t}\int_{t-t_0}^{t} \mathrm{e}^{-\lambda_* s}\mathrm{d}s \\
&= o(\mathrm{e}^{2\lambda_* t}), \qquad \text{as } t\to\infty,
\end{aligned}
$$

where in the first inequality we used (5.19), in the second we used (5.21), and in the final inequality we used (5.22) and (5.19). On the other hand, for the second term in (5.23), we have

$$
\begin{aligned}
&\left|\int_0^{t-t_0} \psi_s[\sigma_{\mathtt{f}}\mathscr{V}\Big[\psi_{t-s}[g]]\Big](r,\upsilon)\mathrm{d}s - \langle\tilde{\varphi}, g\rangle^2\int_0^{t-t_0} \mathrm{e}^{2\lambda_*(t-s)}\psi_s\Big[\sigma_{\mathtt{f}}\mathscr{V}[\varphi]\Big](r,\upsilon)\mathrm{d}s\right| \\
&\quad\leq \int_0^{t-t_0} \psi_s\Big[\Big|\sigma_{\mathtt{f}}\mathscr{V}\big[\psi_{t-s}[g]\big] - \sigma_{\mathtt{f}}\mathscr{V}\big[\langle\tilde{\varphi}, g\rangle \mathrm{e}^{\lambda_*(t-s)}\varphi\big]\Big|\Big](r,\upsilon)\mathrm{d}s \\
&\quad\leq \int_0^{t-t_0} \psi_s\Big[\mathrm{e}^{2\lambda_*(t-s)}\delta CK\|g\|(\|g\| + |\langle\tilde{\varphi}, g\rangle|\|\varphi\|)\Big]\mathrm{d}s \\
&\quad\leq \delta CK\|g\|(\delta\|g\| + 2|\langle\tilde{\varphi}, g\rangle|\|\varphi\|)\mathrm{e}^{2\lambda_* t}\int_0^{t-t_0} \mathrm{e}^{-\lambda_* s}\mathrm{d}s \\
&\quad= O(\delta \mathrm{e}^{2\lambda_* t}),
\end{aligned}
$$

as $t\to\infty$, where the second inequality is due to (5.19) and (5.20) and the third inequality follows from (5.19). Since δ is arbitrary, when combined with the fact that the integral

$$
\int_0^{\infty} \mathrm{e}^{-2\lambda_* s}\psi_s[\sigma_{\mathtt{f}}\mathscr{V}[\varphi]](r,\upsilon)\mathrm{d}s \leq CK\|\varphi\|^2\int_0^{\infty} \mathrm{e}^{-\lambda_* s}\mathrm{d}s < \infty,
$$

the above implies

$$
\int_0^{t-t_0} \psi_s\Big[\sigma_{\mathtt{f}}\mathscr{V}[\psi_{t-s}[g]]\Big](r,\upsilon)\mathrm{d}s \sim \langle\tilde{\varphi}, g\rangle^2\mathrm{e}^{2\lambda_* t}\int_0^{\infty} \mathrm{e}^{-2\lambda_* s}\psi_s[\sigma_{\mathtt{f}}\mathscr{V}[\varphi]](r,\upsilon)\mathrm{d}s,
$$

as $t\to\infty$. The asymptotics in the supercritical case then easily follows.

Next, we consider the subcritical case, i.e., $\lambda_* < 0$. In this setting, the leading term on the right-hand side of (5.11) does not scale away but scales as precisely $\mathrm{e}^{\lambda_* t}$ giving the limit $\langle\tilde{\varphi}, f\rangle\varphi$. For the integral term on the right-hand side of (5.10), we start with a change of variable:

$$
\int_0^{t} \psi_s\Big[\sigma_{\mathtt{f}}\mathscr{V}\big[\psi_{t-s}[g]\big]\Big](r,\upsilon)\mathrm{d}s = \int_0^{t} \psi_{t-s}\Big[\sigma_{\mathtt{f}}\mathscr{V}\big[\psi_s[g]\big]\Big](r,\upsilon)\mathrm{d}s.
$$

Note that by (5.19) and (5.21),

$$\int_{t-t_0}^{t} \psi_{t-s}\Big[\sigma_{\mathrm{f}}\mathscr{V}\big[\psi_s[g]\big]\Big](r,\upsilon)\mathrm{d}s \le CK^2\|g\|^2\int_{t-t_0}^{t} \mathrm{e}^{2\lambda_* s+\lambda_*(t-s)}\mathrm{d}s = o(\mathrm{e}^{\lambda_* t}),$$

where the equality follows as $t\to\infty$, since $\lambda_*<0$. On the other hand, applying (5.19) to ψ_{t-s} and $\sigma_{\mathrm{f}}\mathscr{V}\big[\psi_s[g]\big]$, noting the latter is bounded by $CK^2\mathrm{e}^{2\lambda_* s}\|g\|^2$ as a consequence of (5.21) and (5.22), we get

$$\begin{aligned}
&\left|\int_0^{t-t_0} \psi_{t-s}\Big[\sigma_{\mathrm{f}}\mathscr{V}\big[\psi_s[g]\big]\Big](r,\upsilon)\mathrm{d}s - \mathrm{e}^{\lambda_* t}\varphi(r,\upsilon)\int_0^{t-t_0} \mathrm{e}^{-\lambda_* s}\langle\tilde\varphi,\sigma_{\mathrm{f}}\mathscr{V}\big[\psi_s[g]\big]\rangle\mathrm{d}s\right| \\
&\qquad\le \int_0^{t-t_0} \delta\mathrm{e}^{\lambda_*(t-s)}\Big\|\sigma_{\mathrm{f}}\mathscr{V}\big[\psi_s[g]\big]\Big\|\mathrm{d}s \\
&\qquad\le \int_0^{t-t_0} \delta\mathrm{e}^{\lambda_*(t+s)}CK^2\|g\|^2\mathrm{d}s \\
&\qquad= O(\delta\mathrm{e}^{\lambda_* t}),
\end{aligned}$$

as $t\to\infty$. Arguing as in the supercritical case, we conclude that

$$\int_0^{t} \psi_{t-s}\Big[\sigma_{\mathrm{f}}\mathscr{V}\big[\psi_s[g]\big]\Big](r,\upsilon)\mathrm{d}s \sim \mathrm{e}^{\lambda_* t}\varphi(r,\upsilon)\int_0^{\infty} \mathrm{e}^{-\lambda_* s}\langle\tilde\varphi,\sigma_{\mathrm{f}}\mathscr{V}\big[\psi_s[g]\big]\rangle\mathrm{d}s,$$

as $t\to\infty$, which in turn implies the result in the subcritical case.

The proof in the critical case follows similar arguments. First note that the leading term on the right-hand side of (5.10) will scale away to zero. For the integral term on the right-hand side of (5.10), we can write with a change of variables

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t \psi_s\Big[\sigma_{\mathrm{f}}\mathscr{V}[\psi_{t-s}[g]]\Big](r,\upsilon)\mathrm{d}s = \lim_{t\to\infty}\int_0^1 \psi_{ut}\Big[\sigma_{\mathrm{f}}\mathscr{V}[\psi_{t(1-u)}[g]]\Big](r,\upsilon)\mathrm{d}u.$$

Appealing to (5.19) and dominated convergence, we can pull the limit through the integral to obtain

$$\int_0^1 \langle\tilde\varphi,\sigma_{\mathrm{f}}\mathscr{V}[\langle\tilde\varphi,g\rangle]\varphi\rangle\varphi(r,\upsilon)\mathrm{d}u = \langle\tilde\varphi,g\rangle^2\langle\tilde\varphi,\sigma_{\mathrm{f}}\mathscr{V}[\varphi]\rangle\varphi(r,\upsilon),$$

as required. □

5.3 Moment Growth

The analysis in the previous sections seems hard to extend to higher moments. However there is another approach which brings us back to (5.6) and (5.7). Under the assumptions of (H1) and (H4), we can access all moments of the NBP from $(\mathrm{v}_t, t \geq 0)$ by differentiating as in (5.2) more than once, i.e.,

$$\psi_t^{(k)}[f](r,\upsilon) := \mathbb{E}_{\delta_{(r,\upsilon)}}\left[X_t[f]^k\right] = (-1)^k \frac{\mathrm{d}^k}{\mathrm{d}\theta^k} \mathrm{v}_t[\theta f](r,\upsilon)\bigg|_{\theta=0}, \tag{5.24}$$

for $r \in D$, $\upsilon \in V$, $t \geq 0$, and $k \geq 2$. Note that there is no need to define $\psi^{(1)}$ as we have the special notation ψ. Unlike $(\psi_t, t \geq 0)$, the moment operators $(\psi_t^{(k)}, t \geq 0)$ are not semigroups. They do however satisfy recursion equations which take the form

$$\begin{aligned}\psi_t^{(k)}[f](r,\upsilon) =& \psi_t[f^k](r,\upsilon)\\ &+ \int_0^t \psi_s\left[F_k(\psi_{t-s}^{(k-1)}[f], \cdots, \psi_{t-s}^{(2)}[f], \psi_{t-s}[f])\right](r,\upsilon)\,\mathrm{d}s,\end{aligned} \tag{5.25}$$

where F_k takes a rather complex form. These recursions can be used to develop the asymptotic behaviour in time for the moments in each of the three criticality regimes via an inductive approach. Setting aside the form of F_k, the asymptotic behaviour of $(\psi_t, t \geq 0)$ in Theorem 4.1 together with the scaling limits of $\psi_{t-s}^{(k-1)}[f], \cdots, \psi_{t-s}^{(2)}[f]$ will give us the scaling limit of $\psi_t^{(k)}[f]$. Not surprisingly, the analysis is quite involved and, as it turns out, is not specific to the PBE but works equally well for more general spatial branching processes. For this reason, the proofs of the three main results below are left to Part II of this book where they are restated in the aforesaid general setting.

Theorem 5.2 (Supercritical, $\lambda_* > 0$) *Suppose that (H1), (H2*), and (H4) hold. Fix an integer $k \geq 1$, then*

$$\lim_{t\to\infty} \sup_{\substack{r\in D,\upsilon\in V\\ f\in B_1^+(D\times V)}} \left| \mathrm{e}^{-\lambda_* kt} \frac{\psi_t^{(k)}[f](r,\upsilon)}{\varphi(r,\upsilon)} - k!\langle\tilde{\varphi}, f\rangle^k L_k(r,\upsilon)\right| = 0,$$

where $L_1 = 1$ and we define, iteratively for $k \geq 2$,

$$L_k(r,\upsilon) = \int_0^\infty \mathrm{e}^{-\lambda_* ks}\varphi(r,\upsilon)^{-1}\psi_s\Bigg[\sigma_{\mathrm{f}}\mathcal{E}_\cdot\Bigg[\sum_{[k_1,\ldots,k_N]_k^2} \prod_{\substack{j=1\\ j:k_j>0}}^N \varphi(r,\upsilon_j)L_{k_j}(r,\upsilon_j)\Bigg]\Bigg](r,\upsilon)\mathrm{d}s,$$

and $[k_1,\ldots,k_N]_k^2$ *is the set of all non-negative* N*-tuples* $(k_1,\ldots,k_N)$ *satisfying* $\sum_{i=1}^N k_i = k$ *and at least two of the* k_i *are strictly positive.*

Theorem 5.3 (Subcritical, $\lambda_* < 0$**)** *Suppose that (H1), (H2*), and (H4) hold. Fix an integer* $k \geq 1$*, then*

$$\lim_{t\to\infty} \sup_{\substack{r\in D,\upsilon\in V\\ f\in B_1^+(D\times V)}} \left| \mathrm{e}^{-\lambda_* t} \frac{\psi_t^{(k)}[f](r,\upsilon)}{\varphi(r,\upsilon)} - L_k \right| = 0,$$

where we define iteratively $L_1 = 1$ *and, for* $k \geq 2$*,*

$$L_k = \langle\tilde{\varphi}, f^k\rangle + \int_0^\infty \mathrm{e}^{-\lambda_* s} \left\langle \tilde{\varphi}, \sigma_{\mathtt{f}} \mathscr{E}_\cdot \left[\sum_{[k_1,\ldots,k_N]_k^2} \binom{k}{k_1,\ldots,k_N} \prod_{\substack{j=1\\ j:k_j>0}}^N \psi_s^{(k_j)}[f](\cdot,\upsilon_j) \right] \right\rangle \mathrm{d}s,$$

and $[k_1,\ldots,k_N]_k^n$ *is the set of all non-negative* N*-tuples* $(k_1,\ldots,k_N)$ *such that* $\sum_{i=1}^N k_i = k$ *and exactly* $2 \leq n \leq k$ *of the* k_i *are strictly positive.*

Theorem 5.4 (Critical, $\lambda_* = 0$**)** *Suppose that (H1), (H2*), and (H4) hold. Fix an integer* $k \geq 1$*, then*

$$\lim_{t\to 0} \sup_{\substack{r\in D,\upsilon\in V\\ f\in B_1^+(D\times V)}} \left| t^{-(k-1)} \frac{\psi_t^{(k)}[f](r,\upsilon)}{\varphi(r,\upsilon)} - 2^{-(k-1)} k!\, \langle\tilde{\varphi}, f\rangle^k \langle\tilde{\varphi}, \sigma_{\mathtt{f}} \mathscr{V}[\varphi]\rangle^{k-1} \right| = 0.$$

Let us compare the above results with the results in Theorem 5.1, which corresponds to the setting $k = 2$. It is easy to see that the critical case for $k = 2$ in Theorems 5.4 and 5.1 (iii) directly agree. In the supercritical and subcritical settings, the summations in the definition of L_k degenerate simply to sums over pairs $k_i = 1$ and $k_j = 2 - k_i = 1$ for some $i, j = 1, \cdots, N$ such that $i \neq j$. Both sums attract a factor of 2, either from the $k! = 2$ term which multiplies L_k in the supercritical setting or from the multinomial coefficient of the sum in the subcritical setting. Either way, the summations of products with the respective factor of 2 reduce to the integrals in the definition of L_2 to precisely those of Theorem 5.1 (i) and (ii).

5.4 Running Occupation Moment Growth

Before moving on to an application of the critical moment growth in Theorem 5.4, let us also present an additional suite of results that describe the growth of the moments of the running occupation measure $\int_0^t X_s[g]\mathrm{d}s$, $t \geq 0$. To this end, define, for $r \in D$, $\upsilon \in V$, $g \in B^+(D \times V)$, $k \geq 1$, $t \geq 0$,

$$I_t^{(k)}[g](r, \upsilon) := \mathbb{E}_{\delta_{(r,\upsilon)}}\left[\left(\int_0^t X_s[g]\mathrm{d}s\right)^k\right]. \tag{5.26}$$

For convenience, we will write $I_t[g]$ in place of $I_t^{(1)}[g]$, mirroring similar notation used for the first moment semigroup.

In a similar spirit to (5.24), we can access the moments $I_t^{(k)}[g](r, \upsilon)$ by noting that if we define, for $r \in D$, $\upsilon \in V$ and $g \in B^+(D \times V)$,

$$\mathtt{w}_t[g](r, \upsilon) := \mathbb{E}_{\delta_{(r,\upsilon)}}\left[\exp\left(-\int_0^t X_s[g]\mathrm{d}s\right)\right], \tag{5.27}$$

then

$$I_t^{(k)}[g](r, \upsilon) = (-1)^k \frac{\mathrm{d}^k}{\mathrm{d}\theta^k}\mathtt{w}_t[\theta g]\bigg|_{\theta=0}. \tag{5.28}$$

Moreover, in a similar spirit to the derivation of (5.7), we have that for $g \in B^+(D \times V)$, $t \geq 0$, $r \in D$, and $\upsilon \in V$, $\mathtt{w}_t$ solves

$$\mathtt{w}_t[g](r, \upsilon) = \mathtt{P}_t[1](r, \upsilon) + \int_0^t \mathtt{P}_s\left[\mathtt{G}[\mathtt{w}_{t-s}[g]] - g\mathtt{w}_{t-s}[g]\right](r, \upsilon)\mathrm{d}s. \tag{5.29}$$

A similar recursion to (5.25) then ensues from which an inductive argument can be developed to build up the moment asymptotics.

The complete picture is dealt with in Chap. 9 of Part II of this text. There, one will find the proof of the following results, which mirror Theorems 5.2, 5.3, and 5.4, respectively, albeit in the setting of a general branching Markov process.

Theorem 5.5 (Supercritical, $\lambda_* > 0$) *Suppose that (H1), (H2*), and (H4) hold. Fix an integer $k \geq 1$, then*

$$\lim_{t\to\infty} \sup_{\substack{r\in D,\upsilon\in V\\ f\in B^+(D\times V)}} \left|\varphi(r, \upsilon)^{-1}\mathrm{e}^{-\lambda_* k t} I_t^{(k)}[g](r, \upsilon) - k!\langle\tilde{\varphi}, g\rangle^k L_k\right|,$$

where L_k was defined in Theorem 5.2, albeit with $L_1 = 1/\lambda_*$.

Theorem 5.6 (Subcritical, $\lambda_* < 0$) *Suppose that (H1), (H2*), and (H4) hold. Fix an integer $k \geq 1$, then*

$$\lim_{t\to\infty} \sup_{\substack{r\in D,\upsilon\in V\\ f\in B^+(D\times V)}} \left|\varphi(r, \upsilon)^{-1} I_t^{(k)}[g](r, \upsilon) - k!\langle\tilde{\varphi}, g\rangle^k L_k(r, \upsilon)\right| = 0,$$

where $L_1(r,\upsilon) = \int_0^\infty \varphi(r,\upsilon)^{-1}\psi_s[g](r,\upsilon)\mathrm{d}s$*, and for* $k \geq 2$*, the constants* $L_k(r,\upsilon)$ *are defined recursively via*

$$
\begin{aligned}
L_k(r,\upsilon) = \int_0^\infty & \varphi(r,\upsilon)^{-1}\psi_s \\
& \times \left[\sigma_{\mathrm{f}}\mathcal{E}_\cdot\left[\sum_{[k_1,\dots,k_N]_k^2}\binom{k}{k_1,\dots,k_N}\prod_{\substack{j=1\\ j:k_j>0}}^N \varphi(r,\upsilon_j)L_{k_j}(r,\upsilon_j)\right]\right](x)\,\mathrm{d}s \\
& - k\int_0^\infty \varphi(r,\upsilon)^{-1}\psi_s\Big[g\varphi L_{k-1}\Big](x)\,\mathrm{d}s.
\end{aligned}
$$

Theorem 5.7 (Critical, $\lambda_* = 0$**)** *Suppose that (H1), (H2*), and (H4) hold. Fix* $k \geq 1$*, then*

$$
\lim_{t\to\infty}\sup_{\substack{r\in D,\upsilon\in V\\ f\in B^+(D\times V)}}\left|t^{-(2k-1)}\varphi(r,\upsilon)^{-1}I_t^{(k)}[g](r,\upsilon) - \frac{k!}{2^{k-1}}\langle\tilde{\varphi},g\rangle^k\langle\tilde{\varphi},\sigma_{\mathrm{f}}\mathscr{V}[\varphi]\rangle^{k-1}L_k\right|
$$
$$
= 0,
$$

where $L_1 = 1$ *and* L_k *is defined through the recursion* $L_k = (\sum_{i=1}^{k-1} L_i L_{k-i})/(2k-1)$.

The results in Theorems 5.5, 5.6, and 5.7 are slightly less predictable than Theorems 5.2, 5.3, and 5.4. In the supercritical setting of Theorem 5.5, the exponential growth of the process is still dominant resulting in a growth rate $\mathrm{e}^{k\lambda_* t}$. In the subcritical setting of Theorem 5.6, we will see in forthcoming calculations in the later chapters that $\zeta < \infty$ almost surely, where $\zeta = \inf\{t > 0 : X_t[1] = 0\}$. This tells us that the total occupation $\int_0^\zeta X_s[g]\mathrm{d}s$ is almost surely finite, behaving roughly like an average spatial distribution of mass, i.e., $\langle\tilde{\varphi}, g\rangle$, multiplied by ζ, meaning that no normalisation is required to control the "growth" of the running occupation moments in this case.

Finally, the critical case is somewhat harder to explain heuristically until we have some additional results, which we will address in the next section. We therefore defer our reasoning until the end of the next section.

5.5 Yaglom Limits at Criticality

Another point of interest when it comes to the Pál–Bell equation (5.7) occurs when we set $g = \infty$. Remembering that an empty product is defined to be unity, we see for this special case that

$$1 - \mathrm{v}_t[\infty](r, \upsilon) = \mathbb{P}_{(r,\upsilon)}(\zeta > t), \qquad t \geq 0,$$

where we understand $\mathrm{v}_t[\infty](r, \upsilon) = \lim_{\theta\to\infty} \mathrm{v}_t[\theta](r, \upsilon)$ and

$$\zeta = \inf\{t > 0 : \langle 1, X_t\rangle = 0\}$$

is the extinction time of the NBP.

The first main result of this section (given below) gives the asymptotic decay of the above survival probability in the critical setting, i.e., when $\lambda_* = 0$. It can be stated in a much more general setting than the NBP, and therefore its proof is left to Part II of this book. In order to state it, we need to introduce a new assumption.

(H5) **There exists a constant** $C > 0$ **such that for all** $g \in B^+(D \times V)$,

$$\langle \tilde{\varphi}, \sigma_{\mathtt{f}} \mathscr{V}[g]\rangle \geq C\langle \tilde{\varphi}, \hat{g}^2\rangle, \tag{5.30}$$

where $\hat{g} : D \to [0, \infty) : r \mapsto \int_V g(r, \upsilon')\mathrm{d}\upsilon'$.

Assumption (H5) can be thought of as an irreducibility type condition on the fission operator. It ensures that if there is a fission event (in the stationary distribution), the process should have at least a comparable chance of survival relative to producing an independent number of particles with isotropic velocities at a constant rate.

Indeed, suppose $\sigma_{\mathtt{f}}$ is a constant, and the branching mechanism places an independent random number of offspring, each with independent and uniformly selected velocity in V. In that case, recalling the definition (5.9), for $g \in B^+(D\times V)$, we have

$$\begin{aligned}\sigma_{\mathtt{f}}\mathscr{V}[g] &= \sigma_{\mathtt{f}}\mathscr{E}[N(N-1)]\mathscr{E}[g(U_{\mathbb{S}^2})]^2\\ &= C'\sigma_{\mathtt{f}}\mathscr{E}[N(N-1)]\hat{g}^2\end{aligned}$$

where we have dropped dependency on (r, υ) as this is no longer the case due to the uniformity of the branching mechanism, $U_{\mathbb{S}^2}$ is a random variable that is uniformly distributed on $\mathbb{S}^2$, and C' is a normalisation constant. For an appropriate interpretation of the constant C, the right-hand side of (5.30) is thus equal to the left-hand side of (5.30) for the setting of independent isotropic fission at a constant rate.

The following result is classically known as Kolmogorov's asymptotic for the survival probability in the setting of Bienyamé–Galton–Watson branching processes.

Theorem 5.8 *Suppose (H1), (H2*), (H3), (H4), and (H5) hold and* $\lambda_* = 0$. *Then,*

$$\lim_{t\to\infty} t\mathbb{P}_{\delta_{(r,\upsilon)}}(\zeta > t) = \frac{2\varphi(r, \upsilon)}{\Sigma}, \qquad r \in D, \upsilon \in V,$$

where $\Sigma := \langle \tilde{\varphi}, \sigma_{\mathtt{f}}\mathscr{V}[\varphi]\rangle$.

Combining this result with the conclusion of Theorem 5.4, we discover the following Yaglom limit result for our NBP which echoes a similar result originally proved for Bienyamé–Galton–Watson processes.

Theorem 5.9 *Assume the same conditions as Theorem 5.8. For any $g \in B^+(D\times V)$ and integer $k \geq 1$, we have that*

$$\lim_{t\to\infty} \mathbb{E}_{\delta_{(r,\upsilon)}}\left[\left(\frac{X_t[g]}{t}\right)^k \middle| \zeta > t\right] = 2^{-k}k!\,\langle\tilde{\varphi}, g\rangle^k \Sigma^k$$

or equivalently that

$$\mathtt{Law}\left(\frac{X_t[g]}{t}\middle| \zeta > t\right) \to \mathtt{Law}\left(\mathbf{e}_p\right) \tag{5.31}$$

as $t \to \infty$, where $\mathbf{e}_p$ is an exponential random variable with rate $p := 2/\langle\tilde{\varphi}, g\rangle\Sigma$.

We remark that an alternative way of stating (5.31) is the following:

$$\lim_{t\to\infty} \mathbb{E}_{\delta_{(r,\upsilon)}}\left[\mathrm{e}^{-\theta\frac{X_t[g]}{t}}\middle| \zeta > t\right] = \frac{p}{p+\theta}, \qquad \theta \geq 0.$$

When considering the mean neutron density for a reactor at criticality, when observing the reactor after a large amount of time $t > 0$, our Yaglom limit gives us the counter intuitive result

$$\mathbb{E}_{\delta_{(r,\upsilon)}}\left[X_t[g]\middle| \zeta > t\right] \approx \frac{t\Sigma}{2}\langle\tilde{\varphi}, g\rangle,$$

as $t \to \infty$. This differs from the behaviour of the classical neutron density limiting result which states that

$$\mathbb{E}_{\delta_{(r,\upsilon)}}\left[X_t[g]\right] \sim \langle\tilde{\varphi}, g\rangle\varphi(r,\upsilon),$$

as $t \to \infty$.

Finally, let us conclude by returning to the heuristic explanation for Theorem 5.7. Appealing to Theorem 5.9, we can roughly write, as $t \to \infty$,

$$\begin{aligned}
\mathbb{E}_{\delta_{(r,\upsilon)}}\left[\left(\int_0^t X_s[g]\mathrm{d}s\right)^k\middle| \zeta > t\right] &= \mathbb{E}_{\delta_{(r,\upsilon)}}\left[\left(\int_0^t s\frac{X_s[g]}{s}\mathrm{d}s\right)^k\middle| \zeta > t\right] \\
&\approx \mathbb{E}_{\delta_{(r,\upsilon)}}\left[\mathbf{e}_p\left(\int_0^t s\mathrm{d}s\right)^k\middle| \zeta > t\right] \\
&\approx O(t^{2k}).
\end{aligned}$$

As a consequence, now taking account of Theorem 5.8,

$$\mathbb{E}_{\delta_{(r,\upsilon)}}\left[\left(\int_0^t X_s[g]\mathrm{d}s\right)^k\right] = \mathbb{E}_{\delta_{(r,\upsilon)}}\left[\left(\int_0^t X_s[g]\mathrm{d}s\right)^k \middle| \zeta > t\right]\mathbb{P}_{\delta_{(r,\upsilon)}}(\zeta > t)$$
$$\approx O(t^{2k-1}).$$

5.6 Comments

The Pál–Bell equation is attributed to the concurrent work of Bell [7] and Pál [105, 106]. The Pál–Bell equations are nothing more than examples of the mild non-linear semigroup evolution equations which have appeared regularly in the theory of spatial branching processes and the related theory of superprocesses. See, for example, the discussion in summary of Ikeda et al. [75–77] for the former setting and of Dynkin [50] for the latter setting.

Moment evolution equations and, similarly, evolution equations for occupation moments for NBPs have been considered, for example, in Pazit and Pál [107] (second moments). More significant calculations can be found in the multiple works of Zoia et al. [133–137], Bénichou et al. [9], and Dumonteil and Mazzolo [43]. The results presented in this chapter are based on the more recent work of Dumonteil et al. [44], Cox et al. [31], Harris et al. [72], and Gonzalez et al. [67]. In these articles, a new recursive structure for moments of the NBP and its running occupation functional are established, providing the moment asymptotics in Lemma 5.3 and Theorems 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, and 5.7. The Yaglom limit in Theorem 5.8 is a spatial version of a classical result that was originally proved in the setting of Bienamé–Galton–Watson processes, see, for example, Yaglom [132]. Recent analogues in the spatial setting have also been developed; see, for example, Powell [112] in the setting of branching Brownian motion in a compact domain or Ren et al. [113] in the setting of superprocesses. The result for NBPs presented here was first proved in Harris et al. [72].

Chapter 6
Martingales and Path Decompositions

In this chapter, we use the Perron–Frobenius decomposition of the NBP to show the existence of an intrinsic family of martingales. These are classical objects as far as the general theory of branching processes is concerned and is known to play a key role in the understanding of how particle density aggregates in the pathwise sense (rather than on average). They also serve as the basis of a change of measure, which introduces exceptional stochastic behaviour along one sequential genealogical line of neutron emissions, a so-called *spine* decomposition. In this setting, the exceptional behaviour of the single genealogical line of neutron emissions appears as a biasing of the scatter and fission cross sections to the extent that this one genealogical line never leaves the domain D on which the NBP is defined. The principal gain from examining this spine decomposition in combination with the behaviour of the martingale is that it gives us a sense of how the paths of a surviving NBP look like.

In a similar spirit, in the supercritical setting, we introduce a second decomposition, a so-called *skeletal decomposition*, which requires no change of measure. Here, we consider the behaviour of a supercritical NBP and show that the paths of its particles can be categorised into those that generate a genealogical line of neutrons that eventually become absorbed and those whose genealogical lines of descent survive in D forever. Again, this gives us an unusual insight into how, inside the spatial branching trees of neutrons, there exists a lower density fission process that is keeping the reaction progressing on the event of survival.[1]

[1] Note that even when the NBP is supercritical, a single initial neutron may be absorbed without inducing fission, so survival is not guaranteed.

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8_6

6.1 Martingales

Recall that a martingale was defined in Definition 6.2. We will first discuss the additive martingale that arises naturally from Theorem 4.1. Under the assumptions of Theorem 4.1, recall that the existence of the eigenfunction φ means that

$$\psi_t[\varphi](r, \upsilon) = \mathrm{e}^{\lambda_* t}\varphi(r, \upsilon), \qquad r \in D, \upsilon \in V, t \geq 0. \tag{6.1}$$

This is sufficient to deduce the following easy result.

Lemma 6.1 *Fix $\mu \in \mathcal{M}_c(D \times V)$, the space of finite counting measures on $D \times V$, and define $W = (W_t, t \geq 0)$ as the process*

$$W_t := \mathrm{e}^{-\lambda_* t}\frac{X_t[\varphi]}{\mu[\varphi]}, \qquad t \geq 0. \tag{6.2}$$

Then W is a $\mathbb{P}_\mu$-martingale with unit mean.

Proof Thanks to the semigroup property of (3.30) together with (6.1), we have, for $s, t \geq 0$,

$$\begin{aligned}
\mathbb{E}[W_{t+s}|\mathfrak{F}_t] &= \frac{\mathrm{e}^{-\lambda_* t}}{\mu[\varphi]}\sum_{i=1}^{N_t}\mathrm{e}^{-\lambda_* s}\mathbb{E}\left[X_s^{(i)}[\varphi]\middle|\mathfrak{F}_t\right] \\
&= \frac{\mathrm{e}^{-\lambda_* t}}{\mu[\varphi]}\sum_{i=1}^{N_t}\mathrm{e}^{-\lambda_* s}\psi_s[\varphi](r_i(t), \upsilon_i(t)) \\
&= W_t,
\end{aligned} \tag{6.3}$$

where $X^{(i)}$ is an independent copy of X under $\mathbb{P}_{\delta_{(r_i(t),\upsilon_i(t))}}$ and $\{(r_i(t), \upsilon_i(t)), i = 1, \cdots, N_t\}$ represent the configuration and the number of particles of X at time $t \geq 0$.

Finally, taking expectations again in (6.3), we observe that

$$\mathbb{E}_\mu[W_{t+s}] = \mathbb{E}_\mu[W_t] = 1, \qquad s, t \geq 0,$$

where the second equality follows by considering $t = 0$. Thanks to positivity, this calculation also verifies the requirement that W has finite mean at each $t \geq 0$. □

Analogues of the martingale (6.2) appear in the setting of spatial branching processes in different guises and are sometimes referred to there as the *additive martingale*. As a non-negative martingale, the almost sure limit

$$W_\infty := \lim_{t\to\infty} W_t$$

of (6.2) is assured, thanks to the classical martingale convergence theorem. The next result tells us precisely when this martingale limit is non-zero. In order to state it, we must introduce another hypothesis.

(H3)* **There exists a ball B compactly embedded in D such that**

$$\inf_{r\in B,\upsilon,\upsilon'\in V}\sigma_{\mathtt{f}}(r,\upsilon)\pi_{\mathtt{f}}(r,\upsilon,\upsilon')>0.$$

Theorem 6.1 *Under the assumptions of (H1) and (H2*), we have the following three cases for the martingale $W=(W_t,t\geq 0)$:*

(i) If $\lambda_*>0$ and (H3) holds, then W is $L^1(\mathbb{P})$ convergent.
(ii) If $\lambda_*<0$ and (H3) holds, then $W_\infty=0$ almost surely.
(iii) If $\lambda_*=0$ and (H3) holds, then $W_\infty=0$ almost surely.*

There are various different versions of this theorem that we could have stated. For example, in the supercritical setting $\lambda_*>0$, if we additionally assume that (H4) is in place, which is the norm, then W is also $L^2(\mathbb{P})$ convergent. Note also that (H3*) is just a little bit stronger than the assumption (H3), and hence, a cleaner version of Theorem 6.1 could equally have just assumed (H1), (H2), and (H3*) across the board. These are nonetheless esoteric issues as far as the nuclear modelling perspective is concerned.

On the event that $W_\infty=0$, there are two possible ways the limit could become zero. Either this is because the mass of the martingale continuously limits to zero, or the martingale value jumps to zero because X has become extinct. The former of these two could, in principle, occur in the supercritical case $\lambda_*>0$. Clearly, we have the inclusion

$$\{\zeta<\infty\}\subseteq\{W_\infty=0\},$$

where we recall that

$$\zeta=\inf\{t>0:X_t[1]=0\}.$$

The following theorem frames Theorem 6.1 more concisely, showing the zero set of the martingale limit agrees with extinction. Accordingly, it gives us a valuable statement concerning the survival of X.

Theorem 6.2 *In each of the three cases of Theorem 6.1, we also have that the events $\{W_\infty=0\}$ and $\{\zeta<\infty\}$ almost surely agree. In particular, there is almost sure extinction if and only if $\lambda_*\leq 0$.*

6.2 Strong Laws of Large Numbers

It is particularly interesting to note that, in the setting of a critical system, $\lambda_* = 0$, which is what one would envisage as the natural state in which to keep a nuclear reactor, the results in the previous section evidence a phenomenon that has long been known by engineers and physicists, namely that the fission process eventually dies out. Nonetheless, the neutron density, that is the solution to the NTE (3.9) or (3.32), both predict stabilisation

$$\lim_{t\to\infty} \psi_t[g](r,\upsilon) = \mathbb{E}_{\delta_{(r,\upsilon)}}\big[X_t[g]\big] = \langle\tilde{\varphi}, g\rangle\varphi(r,\upsilon), \quad r \in D, \upsilon \in V, g \in B^+(E). \tag{6.4}$$

In the subcritical setting, both the neutron density and the pathwise behaviour of the NBP agree, in the sense that they both tend to zero.

In the supercritical setting, there is also a concurrence between the behaviour of neutron density and the pathwise behaviour of the NBP, albeit a little more interesting. In this setting, the convergence of the martingale W to a non-trivial limit tells us that when we weight the i-th particle at time $t \geq 0$ with $\varphi(r_i(t), \upsilon_i(t))$, we get an exact convergence. Moreover, because W_∞ is an $L^1(\mathbb{P})$ limit, this convergence would appear to agree with what the NTE predicts. Indeed, the neutron density analogue of the martingale convergence would correspond to its mean value, which, thanks to Theorem 6.1, follows the asymptotic

$$\mathbb{E}_\mu[W_t] = \mathrm{e}^{-\lambda_* t}\frac{\mu[\psi_t[\varphi]]}{\mu[\varphi]} = \mathrm{e}^{-\lambda_* t}\frac{\mathbb{E}_\mu\,[X_t[\varphi]]}{\mu[\varphi]} \sim \langle\tilde{\varphi}, \varphi\rangle = 1 = \mathbb{E}_\mu[W_\infty], \tag{6.5}$$

as $t \to \infty$. The question thus remains as to whether a stronger correspondence holds in the sense that, when we weight the i-th particle in the NBP not by $\varphi(r_i(t), \upsilon_i(t))$, but a more general weight of the form $g(r_i(t), \upsilon_i(t))$, for some $g \in B^+(D \times V)$, do we also get exact convergence? The obvious candidate result in light of (6.4) and (6.5) should be that, $\mathbb{P}_\mu$-almost surely,

$$\mathrm{e}^{-\lambda_* t}\frac{X_t[g]}{\mu[g]} \sim \langle\tilde{\varphi}, g\rangle W_\infty,$$

as $t \to \infty$.

Theorem 6.3 *Suppose the assumptions (H1), (H2*), and (H3) hold and $\lambda_* > 0$. Suppose $\mu \in \mathcal{M}_c(D \times V)$ and $g \in B^+(D \times V)$ and $g/\varphi \in B^+(D \times V)$. Then,*

$$\lim_{t\to\infty} \mathrm{e}^{-\lambda_* t}\frac{X_t[g]}{\mu[\varphi]} = \langle\tilde{\varphi}, g\rangle W_\infty, \tag{6.6}$$

$\mathbb{P}_\mu$-almost surely. Without the requirement that $g/\varphi \in B^+(D \times V)$, the limit additionally holds in $L^2(\mathbb{P})$.

Remark 6.1 In essence, we do not need to assume supercriticality for the statement of Theorem 6.3. Indeed, otherwise under all the other assumptions, since (up to a multiplicative constant) $\mathrm{e}^{-\lambda_* t} X_t[g] \leq \mathrm{e}^{-\lambda_* t} X_t[\varphi] = \mu[\varphi] W_t$, then (6.6) still holds albeit that the limit is trivially zero.

6.3 Spine Decomposition

As with many spatial branching processes, the most efficient way to technically analyse the stochastic growth of the system, as in the proof of Theorem 6.3, for example, is through the pathwise behaviour of the particle system described by the *spine decomposition*. As alluded to in the introduction, this is the result of performing a change of probability measure induced by the martingale (6.2). Whilst classical in the branching process literature, this is unknown in the setting of neutron transport.

To describe the spine decomposition, we introduce the following change of measure, induced by the martingale W_t,

$$\left.\frac{\mathrm{d}\mathbb{P}^{\varphi}_{\mu}}{\mathrm{d}\mathbb{P}_{\mu}}\right|_{\mathfrak{F}_t} = W_t, \qquad t \geq 0, \tag{6.7}$$

where μ belongs to the space of finite atomic measures $\mathcal{M}_c(D \times V)$. In probabilistic terms, this is shorthand for defining the consistent family of probability measures

$$\mathbb{P}^{\varphi}_{\mu}(A) = \mathbb{P}_{\mu}\left[\mathbf{1}_A W_t\right], \qquad t \geq 0,\ A \in \mathfrak{F}_t.$$

We may think of the change of measure (6.7) as a method of biasing or "twisting" the original law $\mathbb{P}_{\mu}$.

In the next theorem, we will formalise an understanding of this change of measure in terms of another $\mathcal{M}_c(D \times V)$-valued stochastic process which is not an NBP. Let us now define it through an algorithmic construction.

1. From the initial configuration $\mu \in \mathcal{M}_c(D \times V)$, with an arbitrary enumeration of particles, the i-th neutron is selected and marked "*spine*" with empirical probability

$$\frac{\varphi(r_i, \upsilon_i)}{\mu[\varphi]}.$$

2. The neutrons $j \neq i$ in the initial configuration that are not marked "*spine*", each issue independent copies of $(X, \mathbb{P}_{\delta_{(r_j,\upsilon_j)}})$, respectively.
3. For the marked neutron, it evolves from its initial configuration as an NRW characterised by the rate function

$$\sigma_{\mathtt{s}}(r, \upsilon)\frac{\varphi(r, \upsilon')}{\varphi(r, \upsilon)}\pi_{\mathtt{s}}(r, \upsilon, \upsilon'), \qquad r \in D, \upsilon, \upsilon' \in V.$$

4. The marked neutron undergoes fission at the accelerated rate $\varphi(r, \upsilon)^{-1}(\mathtt{F} + \sigma_{\mathtt{f}}\mathtt{I})\varphi(r, \upsilon)$, when in physical configuration $r \in D, \upsilon \in V$, at which point, it scatters a random number of particles according to the random measure on V given by $(\mathsf{Z}, \mathscr{P}^{\varphi}_{(r,\upsilon)})$ where

$$\frac{\mathrm{d}\mathscr{P}^{\varphi}_{(r,\upsilon)}}{\mathrm{d}\mathscr{P}_{(r,\upsilon)}} = \frac{\mathsf{Z}[\varphi]}{\mathscr{E}_{(r,\upsilon)}[\mathsf{Z}[\varphi]]}. \tag{6.8}$$

5. When fission of the marked neutron occurs in physical configuration $r \in D, \upsilon \in V$, set

$$\mu = \sum_{i=1}^{n} \delta_{(r,\upsilon_i)}, \text{ where, in the previous step, } \mathsf{Z} = \sum_{i=1}^{n} \delta_{\upsilon_i},$$

and repeat step 1.

The process $X^{\varphi} = (X^{\varphi}_t, t \geq 0)$ describes the physical configuration (position and velocity) of all the particles in the system at time t, for $t \geq 0$, as per the algorithmic description above (Fig. 6.1). In particular, although it is not clear which of the neutrons that contribute to $X_t \in \mathscr{M}_c(D \times V)$ is marked as the spine, it is included in the population at time t. We are also interested in the configuration of the single genealogical line of descent which has been marked "*spine*". The process that the spine follows in configuration space $D \times V$ will be denoted $(R^{\varphi}, \Upsilon^{\varphi}) := ((R^{\varphi}_t, \Upsilon^{\varphi}_t), t \geq 0)$.

The process X^{φ} alone is not Markovian, as it requires knowledge of the spine process $(R^{\varphi}, \Upsilon^{\varphi})$. However, together, the processes $(X^{\varphi}, (R^{\varphi}, \Upsilon^{\varphi}))$ make a Markov pair, whose probabilities we will denote by $(\tilde{\mathbb{P}}^{\varphi}_{\mu,(r,\upsilon)}, \mu \in \mathscr{M}_c(D \times V), r \in D, \upsilon \in V)$.

We will write

$$\tilde{\mathbb{P}}^{\varphi}_{\mu,\varphi\mu} = \sum_{i=1}^{n} \frac{\varphi(r_i, \upsilon_i)}{\mu[\varphi]}\tilde{\mathbb{P}}^{\varphi}_{\mu,(r_i,\upsilon_i)} = \frac{1}{\mu[\varphi]}\int_{D\times V} \varphi(r, \upsilon)\mu(\mathrm{d}r, \mathrm{d}\upsilon)\tilde{\mathbb{P}}^{\varphi}_{\mu,(r,\upsilon)}$$

when $\mu = \sum_{i=1}^{n} \delta_{(r_i,\upsilon_i)}$. In effect, the law $\tilde{\mathbb{P}}^{\varphi}_{\mu,\varphi\mu}$ corresponds to picking the neutron that will be identified as the spine at time $t = 0$ with a density φ with respect to the initial configuration of neutrons, given by μ. This corresponds to step 1 of the construction of $(X^{\varphi}, (R^{\varphi}, \Upsilon^{\varphi}))$. We write, for convenience, $\tilde{\mathbb{P}}^{\varphi} = (\tilde{\mathbb{P}}^{\varphi}_{\mu,\varphi\mu}, \mu \in \mathscr{M}_c)$.

The next result tells us that if we can ensure that we launch $(X^{\varphi}, (R^{\varphi}, \Upsilon^{\varphi}))$ with a particular initial configuration of neutrons and a randomised allocation for the spine among them according to $\tilde{\mathbb{P}}^{\varphi}_{\mu,\varphi\mu}$, then observing the process X^{φ} without

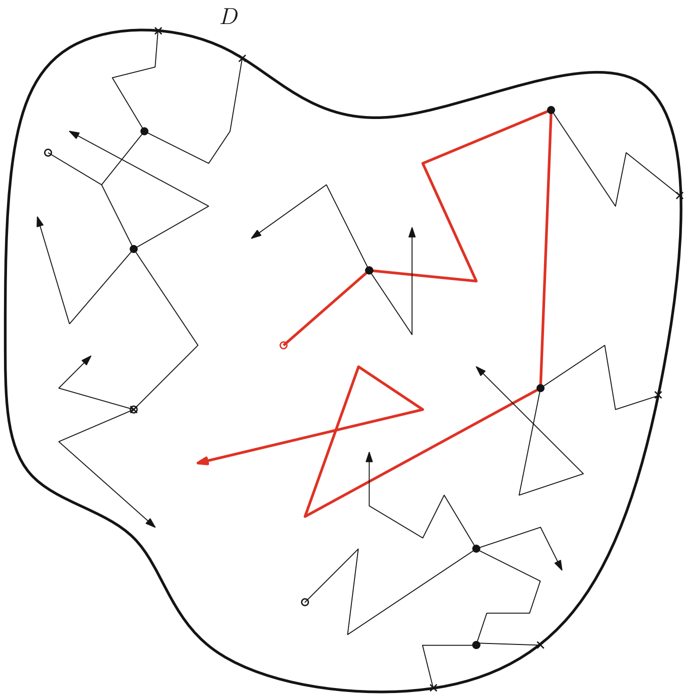

Fig. 6.1 A possible path realisation of the spine decomposition. The red path is that of the spine. The neutrons at time $t = 0$ are depicted by $\circ$, and fission events are depicted by $\bullet$. Particles depicted in black evolve as the bulk of X^φ

knowing $(R^\varphi, \Upsilon^\varphi)$ is equivalent to what one sees of the original NBP under the change of measure (6.7).

Theorem 6.4 *Under assumptions (H1) and (H2*), the process* $(X^\varphi, \tilde{\mathbb{P}}^\varphi)$ *is Markovian and equal in law to* $(X, \mathbb{P}^\varphi)$*, where* $\mathbb{P}^\varphi = (\mathbb{P}^\varphi_\mu, \mu \in \mathscr{M}_c(D \times V))$.

Theorem 6.4 also tells us that the effect of changing probabilities via the mechanism (6.7) results in the NBP taking the shape of what is tantamount to an NRW with immigration. Indeed, aside from other neutrons present at time $t = 0$ which evolve as NBPs, we may think of the spine as a special particle which moves around according to $(R^\varphi, \Upsilon^\varphi)$ and immigrates particles into the bulk of X^φ at a special rate and with a special fission kernel, both depending on $(R^\varphi, \Upsilon^\varphi)$. Once particles immigrate, they evolve as a normal NBP.

We would also like to understand the dynamics of the spine $(R^\varphi, \Upsilon^\varphi)$ as an autonomous process. For convenience, let us denote the family of probabilities of the latter by $\tilde{\mathbf{P}}^\varphi = (\tilde{\mathbf{P}}^\varphi_{(r,\upsilon)}, r \in D, \upsilon \in V)$, in other words, the marginals of $(\tilde{\mathbb{P}}^\varphi_{\mu,(r,\upsilon)}, \mu \in \mathscr{M}_c(D \times V), r \in D, \upsilon \in V)$.

We further define the probabilities $\mathbf{P}^{\varphi} := (\mathbf{P}^{\varphi}_{(r,\upsilon)}, (r, \upsilon) \in D \times V)$ to describe the law of an $\alpha^{\varphi}\pi^{\varphi}$-NRW, where

$$\alpha^{\varphi}(r,\upsilon)\pi^{\varphi}(r,\upsilon,\upsilon') = \frac{\varphi(r,\upsilon')}{\varphi(r,\upsilon)}\left(\sigma_{\mathtt{s}}(r,\upsilon)\pi_{\mathtt{s}}(r,\upsilon,\upsilon') + \sigma_{\mathtt{f}}(r,\upsilon)\pi_{\mathtt{f}}(r,\upsilon,\upsilon')\right), \tag{6.9}$$

for $r \in D$, $\upsilon, \upsilon' \in V$. Recall from (4.1) that

$$\beta(r,\upsilon) = \sigma_{\mathtt{f}}(r,\upsilon)\left(\int_V \pi_{\mathtt{f}}(r,\upsilon,\upsilon')\mathrm{d}\upsilon' - 1\right),$$

and recall that (R, Υ) under $\mathbf{P}$ is the $\alpha\pi$-NRW that appears in the many-to-one Lemma 4.1. We are now ready to identify the spine.

Lemma 6.2 *Under assumptions (H1) and (H2*), the process* $((R^{\varphi}, \Upsilon^{\varphi}), \tilde{\mathbf{P}}^{\varphi})$ *is an NRW equal in law to* $((R, \Upsilon), \mathbf{P}^{\varphi})$ *and, moreover,*

$$\left.\frac{\mathrm{d}\mathbf{P}^{\varphi}_{(r,\upsilon)}}{\mathrm{d}\mathbf{P}_{(r,\upsilon)}}\right|_{\mathfrak{F}_t} = \mathrm{e}^{-\lambda_* t + \int_0^t \beta(R_s,\Upsilon_s)\mathrm{d}s}\frac{\varphi(R_t,\Upsilon_t)}{\varphi(r,\upsilon)}\mathbf{1}_{\{t<\tau^D\}}, \quad t \geq 0, r \in D, \upsilon \in V. \tag{6.10}$$

In addition, $((R, \Upsilon), \mathbf{P}^{\varphi})$ *is conservative with a stationary distribution*

$$\varphi\tilde{\varphi}(r,\upsilon)\mathrm{d}r \text{ on } D \times V.$$

Recalling the discussion in Sect. 2.3, we can develop a heuristic understanding of the motion of the spine via its generator. Taking account of the fact that $((R, \Upsilon), \mathbf{P})$ is an $\alpha\pi$-NRW, in the spirit of the discussion in Sect. 3.3, it is easy to write down the action of its generator as

$$\mathtt{L}f(r,\upsilon) = \upsilon \cdot \nabla_r f(r,\upsilon) + \alpha(r,\upsilon)\int_V \left(f(r,\upsilon') - f(r,\upsilon)\right)\pi(r,\upsilon,\upsilon')\mathrm{d}\upsilon', \tag{6.11}$$

for $f \in B^+(D \times V)$ such that $\nabla_r f$ is well defined. According to Theorem 6.2, the change of measure (6.10) means that the spine under $\tilde{\mathbb{P}}^{\varphi}$ has a generator whose action is given by

$$\mathtt{L}^{\varphi} f := \varphi^{-1}(L + \beta - \lambda_*)(\varphi f),$$

where we have appealed to the spirit of the calculations in (2.11) and (2.43). As such, we have

$$\begin{aligned}\mathtt{L}^{\varphi} f(r,\upsilon) = {}& \upsilon \cdot \nabla_r f(r,\upsilon) + \frac{\upsilon \cdot \nabla_r \varphi(r,\upsilon)}{\varphi(r,\upsilon)} f(r,\upsilon) + \beta f(r,\upsilon) - \lambda_* f(r,\upsilon)\\ &+ \frac{\alpha(r,\upsilon)}{\varphi(r,\upsilon)}\int_V \left(f(r,\upsilon') - f(r,\upsilon)\right)\varphi(r,\upsilon')\pi(r,\upsilon,\upsilon')\mathrm{d}\upsilon'\end{aligned}$$

$$
\begin{aligned}
&\quad + f(r,\upsilon)\frac{\alpha(r,\upsilon)}{\varphi(r,\upsilon)}\int_V \big(\varphi(r,\upsilon') - \varphi(r,\upsilon)\big)\,\pi(r,\upsilon,\upsilon')\mathrm{d}\upsilon' \\
&= \upsilon\cdot\nabla_r f(r,\upsilon) + \frac{(L+\beta-\lambda_*)\varphi(r,\upsilon)}{\varphi(r,\upsilon)} f(r,\upsilon) \\
&\quad + \alpha(r,\upsilon)\int_V \big(f(r,\upsilon') - f(r,\upsilon)\big)\,\frac{\varphi(r,\upsilon')}{\varphi(r,\upsilon)}\pi(r,\upsilon,\upsilon')\mathrm{d}\upsilon'.
\end{aligned}
\tag{6.12}
$$

From Theorem 4.1, we know that φ is an eigenfunction for the semigroup $(\psi_t, t \geq 0)$, which, in turn, by the many-to-one Lemma 4.1, is also the semigroup of $((R, \Upsilon), \mathbf{P})$ with potential β. We would thus expect, at least heuristically, that $(L + \beta - \lambda_*)\varphi = 0$. Using this in (6.12), we conclude that

$$
\begin{aligned}
\mathtt{L}^{\varphi} f(r,\upsilon) = \upsilon\cdot\nabla_r f(r,\upsilon) + \alpha(r,\upsilon)\int_V \big(f(r,\upsilon') - f(r,\upsilon)\big) \\
\times \frac{\varphi(r,\upsilon')}{\varphi(r,\upsilon)}\pi(r,\upsilon,\upsilon')\mathrm{d}\upsilon'.
\end{aligned}
\tag{6.13}
$$

In conclusion, the behaviour of the spine motion under the change of measure is equivalent to a φ-tilting of the motion of the $\alpha\pi$-NRW, favouring the outgoing configurations (r, υ') for which $\varphi(r, \upsilon') > \varphi(r, \upsilon)$, where (r, υ) is the incident configuration, and penalising when the inequality goes in the other direction.

6.4 Skeletal Decomposition

There is also a second path decomposition that is fundamental to understanding the stochastic behaviour of the NBP in the supercritical setting, namely the *skeleton* decomposition. In rough terms, for the NBP, we can speak of genealogical lines of descent, meaning neutrons that came from a fission event of a neutron that came from a fission event of a neutron, and so on, back to one of the initial neutrons at time $t = 0$. If we focus on an individual genealogical line of descent embedded in the NBP, it has a space-velocity trajectory which takes the form of an NRW whose spatial component may or may not hit the boundary of D. When the NBP survives for all time, which of course is only possible in the supercritical setting, there must necessarily be at least one genealogical line of descent whose spatial trajectories remain in D forever.

It turns out that there are many immortal genealogical lines of descent that survive eternally on the survival set of our NBP. Indeed, together they create an entire subtree which is itself an NBP, albeit with biased stochastic behaviour relative to the original NBP, known as *the skeleton*. The basic idea of the skeletal decomposition is to understand the space-velocity dynamics and the fission process of the embedded neutron branching process of immortal genealogies. For the remaining neutron genealogies that go to the boundary of D or end in neutron capture, the skeletal decomposition identifies them as immigrants that are "dressed" along the path of the skeleton.

The remainder of this section is devoted to a description of the skeletal decomposition for the NBP. As with most of the exposition in this part of the book, the skeletal decomposition can be stated in the much more general setting of non-local branching Markov processes, and we will deal with this in full rigour in Part II of the book.

Let us start by first introducing some more notation. First, define

$$w(r, \upsilon) := \mathbb{P}_{\delta_{(r,\upsilon)}}(\zeta < \infty), \qquad r \in D, \upsilon \in V, \tag{6.14}$$

where we recall $\zeta := \inf\{t \geq 0 : X_t[1] = 0\}$. We extend the definition of w to allow it to take the value 1 on the cemetery state $\dagger$. We will also frequently use the notation

$$p(r, \upsilon) := 1 - w(r, \upsilon), \qquad r \in D, \upsilon \in V,$$

for the survival probability.

Given the configuration $\{(r_i(t), \upsilon_i(t)), i = 1, \cdots, N_t\}$ of our NBP at time $t \geq 0$, it is clear that

$$\{\zeta < \infty\} = \bigcap_{i=1}^{N_t} \{X_s^{(i)}[1] = 0 \text{ for some } s \geq t\},$$

where $(X^{(i)}, i = 1, \cdots, N_t)$ are independent copies of X under the respective probabilities $\mathbb{P}_{\delta_{(r_i(t),\upsilon_i(t))}}$, $i = 1, \cdots, N_t$. It follows that, by conditioning on $\mathfrak{F}_t = \sigma(X_s, s \leq t)$, for $t \geq 0$,

$$w(r, \upsilon) = \mathbb{E}_{\delta_{(r,\upsilon)}}\left[\prod_{i=1}^{N_t} w(r_i(t), \upsilon_i(t))\right]. \tag{6.15}$$

Recall that $\hat{\mathtt{P}}_t$ is a slight adjustment of $\mathtt{P}_t$ which returns a value of 1 on the event of killing. Taking Lemma 5.2 and (6.15) into account, it is easy to deduce that w is an invariant solution to the Pál-Bell equation (5.7). Hence,

$$w(r, \upsilon) = \hat{\mathtt{P}}_t[w](r, \upsilon) + \int_0^t \mathtt{P}_s\left[\mathtt{G}[w]\right](r, \upsilon)\mathrm{d}s, \quad t \geq 0, \; r \in D, \upsilon \in V, \tag{6.16}$$

where we recall that

$$\mathtt{G}[w](r,\upsilon) = \sigma_{\mathtt{f}}(r,\upsilon)\mathscr{E}_{(r,\upsilon)}\left[\prod_{i=1}^{N} w(r,\upsilon_i) - w(r,\upsilon)\right]. \tag{6.17}$$

We also have the following lemma, which will be necessary for the development of the skeletal decomposition.

Lemma 6.3 *Assuming* $\lambda_* > 0$, *we have* $\inf_{r\in D,\upsilon\in V} w(r,\upsilon) > 0$ *and* $w(r,\upsilon) < 1$, *for* $r \in D, \upsilon \in V$.

Proof On account of the inclusion $\{\zeta < \infty\} \subseteq \{W_\infty = 0\}$, we see that $w(r,\upsilon) \le \mathbb{P}_{\delta_{(r,\upsilon)}}(W_\infty = 0)$, $r \in D, \upsilon \in V$. Recalling from Theorem 6.1 that W converges both almost surely and in $L^1(\mathbb{P})$ to its limit, we have that $\mathbb{P}_{\delta_{(r,\upsilon)}}(W_\infty = 0) < 1$ for $r \in D, \upsilon \in V$. This, combined with the fact that every particle may leave the bounded domain D directly, without scattering or undergoing fission, with positive probability, gives us that

$$\mathrm{e}^{-\int_0^{\kappa^D_{r,\upsilon}} \sigma(r+\upsilon s,\upsilon)\mathrm{d}s} < w(r,\upsilon) < 1 \text{ for all } r \in D, \upsilon \in V. \tag{6.18}$$

Note that the lower bound is uniformly bounded away from 0 thanks to the boundedness of D, the minimal velocity $\mathtt{v}_{\mathtt{min}}$ (which together uniformly upper bound $\kappa^D_{r,\upsilon}$), and the uniformly upper bounded rates of fission and scattering. The upper inequality becomes an equality for $r \in \partial D$ and $\upsilon \cdot \mathbf{n}_r > 0$. □

Recalling that $\mathtt{P}$ is the semigroup for the $\sigma_s\pi_s$-NRW killed on exiting D, we may rewrite (6.16) in the form

$$w(r,\upsilon) = \mathbf{E}_{(r,\upsilon)}[w(R_{t\wedge\tau_D}, \Upsilon_{t\wedge\tau_D})] + \mathbf{E}_{(r,\upsilon)}\left[\int_0^{t\wedge\tau_D} w(R_s,\Upsilon_s)\frac{\mathtt{G}[w](R_s,\Upsilon_s)}{w(R_s,\Upsilon_s)}\mathrm{d}s\right],$$

$t \ge 0$, where we recall that τ_D denotes the first time the $\sigma_{\mathtt{S}}\pi_{\mathtt{S}}$-NRW exits D and we have extended the definition of w to take the value 1 on the cemetery state, which, in the current setting, is the boundary of D. Noting that, thanks to the previous lemma,

$$\sup_{r\in D,\upsilon\in V} \mathtt{G}[w](r,\upsilon)/w(r,\upsilon) < \infty,$$

we can appeal to Theorem 2.1 to obtain

$$w(r,\upsilon) = \mathbf{E}_{(r,\upsilon)}\left[w(R_{t\wedge\tau_D}, \Upsilon_{t\wedge\tau_D})\exp\left(\int_0^{t\wedge\tau_D} \frac{\mathtt{G}[w](R_s,\Upsilon)}{w(R_s,\Upsilon_s)}\mathrm{d}s\right)\right], \tag{6.19}$$

for all $r \in D, \upsilon \in V, t \ge 0$. This identity will turn out to be extremely useful in our analysis, and in particular, the equality (6.19) together with the Markov property of (R,Υ) implies that the object in the expectation on the right-hand side of (6.19) is a martingale.

In Theorem 6.5 below, we give the skeletal decomposition in the form of a theorem. In order to state this result, we first need to develop two notions of conditioning. We remind the reader that there is a glossary at the end of the book containing the various notation we introduce and use.

The basic pretext of the skeletal decomposition is that we want to split genealogical lines of descent into those that survive forever and those that are killed. To this end, let $c_i(t)$ denote the label of a particle $i \in \{1, \ldots, N_t\}$. We label a particle "prolific", denoted $c_i(t) = \uparrow$, if it has an infinite genealogical line of descent, and $c_i(t) = \downarrow$, if its line of descent dies out (i.e., "non-prolific"). Ultimately, we want to describe how the spatial genealogical tree of the NBP can be split into a spatial genealogical sub-tree, consisting of $\uparrow$-labelled particles (the skeleton), which is dressed with trees of $\downarrow$-labelled particles.

Let $\mathbb{P}^{\updownarrow} = (\mathbb{P}^{\updownarrow}_{\delta_{(r,\upsilon)}}, r \in D, \upsilon \in V)$ denote the probabilities of the two-labelled process described above. Then, for $t \geq 0$ and $r \in D$, $\upsilon \in V$, we have the following relationship between $\mathbb{P}^{\updownarrow}$ and $\mathbb{P}$:

$$\left.\frac{\mathrm{d}\mathbb{P}^{\updownarrow}_{\delta_{(r,\upsilon)}}}{\mathrm{d}\mathbb{P}_{\delta_{(r,\upsilon)}}}\right|_{\mathfrak{F}_\infty} = \prod_{i=1}^{N_t} \left(\mathbf{1}_{(c_i(t)=\uparrow)} + \mathbf{1}_{(c_i(t)=\downarrow)}\right) = 1, \tag{6.20}$$

where $\mathfrak{F}_\infty = \sigma\left(\cup_{t\geq 0}\mathfrak{F}_t\right)$. Projecting onto $\mathfrak{F}_t$, for $t \geq 0$, we have

$$\begin{aligned}
\left.\frac{\mathrm{d}\mathbb{P}^{\updownarrow}_{\delta_{(r,\upsilon)}}}{\mathrm{d}\mathbb{P}_{\delta_{(r,\upsilon)}}}\right|_{\mathfrak{F}_t} &= \mathbb{E}_{\delta_{(r,\upsilon)}}\left(\prod_{i=1}^{N_t}\left(\mathbf{1}_{(c_i(t)=\uparrow)} + \mathbf{1}_{(c_i(t)=\downarrow)}\right)\Big|\mathfrak{F}_t\right) \\
&= \sum_{I\subseteq\{1,\ldots N_t\}} \prod_{i\in I} \mathbb{P}_{\delta_{(r,\upsilon)}}(c_i(t) = \uparrow |\mathfrak{F}_t) \prod_{i\in\{1,\ldots,N_t\}\setminus I} \mathbb{P}_{\delta_{(r,\upsilon)}}(c_i(t) = \downarrow |\mathfrak{F}_t) \\
&= \sum_{I\subseteq\{1,\ldots N_t\}} \prod_{i\in I} p(r_i(t), \upsilon_i(t)) \prod_{i\in\{1,\ldots,N_t\}\setminus I} w(r_i(t), \upsilon_i(t)),
\end{aligned} \tag{6.21}$$

where we understand the sum to be taken over all subsets of $\{1, \cdots, N_t\}$, each of which is denoted by I.

The decomposition in (6.21) indicates the starting point of how we break up the law of the NBP according to subtrees that are categorised as $\downarrow$ (with probability w) and subtrees that are categorised as $\uparrow$ with $\downarrow$ dressing (with probability p), the so-called *skeletal decomposition*.

In the next two sections, we will examine the notion of the NBP conditioned to die out and conditioned to survive. Thereafter, we will use the characterisation of these conditioned trees to formalise our skeletal decomposition.

↓-*Trees*

Let us start by characterising the law of genealogical trees populated by the marks ↓. Thanks to the branching property, it suffices to consider trees which are issued with a single particle with mark ↓. By the definition of the mark $c_{\emptyset}(0) = \downarrow$, where $\emptyset$ is the initial ancestral particle, this is the same as understanding the law of $(X, \mathbb{P})$ conditioned to become extinct. Indeed, for $A \in \mathfrak{F}_t$,

$$
\begin{aligned}
\mathbb{P}^{\downarrow}_{\delta_{(r,\upsilon)}}(A) &:= \mathbb{P}^{\updownarrow}_{\delta_{(r,\upsilon)}}(A | c_{\emptyset}(0) = \downarrow) \\
&= \frac{\mathbb{P}^{\updownarrow}_{\delta_{(r,\upsilon)}}(A; c_i = \downarrow, \text{ for each } i = 1, \dots, N_t)}{\mathbb{P}^{\updownarrow}_{\delta_{(r,\upsilon)}}(c_{\emptyset}(0) = \downarrow)} \\
&= \frac{\mathbb{E}_{\delta_{(r,\upsilon)}}\left[\mathbf{1}_A \prod_{i=1}^{N_t} w(r_i(t), \upsilon_i(t))\right]}{w(r, \upsilon)}.
\end{aligned}
\tag{6.22}
$$

We are now in a position to characterise the NBP trees which are conditioned to become extinct (equivalently, with genealogical lines of descent which are marked entirely with ↓). Heuristically speaking, the next proposition shows that the conditioning creates a neutron branching process in which particles are prone to die out (whether that be due to being absorbed at the boundary or by suppressing offspring).

Lemma 6.4 (↓ Trees) *For initial configurations of the form $\mu = \sum_{i=1}^{n} \delta_{(r_i, \upsilon_i)}$, for $n \in \mathbb{N}$ and $(r_1, \upsilon_1), \cdots, (r_n, \upsilon_n) \in D \times V$, define the measure $\mathbb{P}^{\downarrow}_{\mu}$ via*

$$
\mathbb{P}^{\downarrow}_{\mu} = \otimes_{i=1}^{n} \mathbb{P}^{\downarrow}_{\delta_{(r_i, \upsilon_i)}},
$$

i.e., starting independent processes at configurations (r_i, υ_i) each under $\mathbb{P}^{\downarrow}_{\delta_{(r_i,\upsilon_i)}}$, for $i = 1, \cdots, n$, where $\mathbb{P}^{\downarrow}_{\delta_{(r_i,\upsilon_i)}}$ was defined in (6.22). Then, under $\mathbb{P}^{\downarrow}_{\mu}$, X is an NBP with motion semigroup $\mathtt{P}^{\downarrow}$ and fission mechanism $\mathtt{G}^{\downarrow}$ defined as follows. The motion semigroup $\mathtt{P}^{\downarrow}$ is that of an NRW with probabilities $(\mathbf{P}^{\downarrow}_{(r,\upsilon)}, r \in D, \upsilon \in V)$, where

$$
\left.\frac{\mathrm{d}\mathbf{P}^{\downarrow}_{(r,\upsilon)}}{\mathrm{d}\mathbf{P}_{(r,\upsilon)}}\right|_{\sigma((R_s,\Upsilon_s), s \le t)} = \frac{w(R_{t\wedge\tau_D}, \Upsilon_{t\wedge\tau_D})}{w(r, \upsilon)} \exp\left(\int_0^{t\wedge\tau_D} \frac{\mathtt{G}[w](R_s, \Upsilon_s)}{w(R_s, \Upsilon_s)} \mathrm{d}s\right), t \ge 0.
\tag{6.23}
$$

For $r \in D$, $\upsilon \in V$, and $f \in B_1^+(D \times V)$, the fission mechanism is given by

$$
\mathtt{G}^{\downarrow}[f] = \frac{1}{w}\left[\mathtt{G}[fw] - f\mathtt{G}[w]\right],
\tag{6.24}
$$

which may otherwise be identified as

$$\mathtt{G}^{\downarrow}[f](r,\upsilon) = \sigma_{\mathtt{f}}^{\downarrow}(r,\upsilon)\mathscr{E}^{\downarrow}_{(r,\upsilon)}\left[\prod_{j=1}^{N} f(r_j,\upsilon_j) - f(r,\upsilon)\right],$$

where

$$\sigma_{\mathtt{f}}^{\downarrow}(r,\upsilon) = \sigma_{\mathtt{f}}(x) + \frac{\mathtt{G}[w](r,\upsilon)}{w(r,\upsilon)} = \frac{\sigma_{\mathtt{f}}(r,\upsilon)}{w(r,\upsilon)}\mathscr{E}_x\left[\prod_{j=1}^{N} w(r_j,\upsilon_j)\right], \qquad r\in D, \upsilon\in V, \tag{6.25}$$

and

$$\left.\frac{\mathrm{d}\mathscr{P}^{\downarrow}_{(r,\upsilon)}}{\mathrm{d}\mathscr{P}_{(r,\upsilon)}}\right|_{\sigma(N,(r_1,\upsilon_1),\ldots,(r_N,\upsilon_N))} = \frac{\prod_{i=1}^{N} w(r_i,\upsilon_i)}{\mathscr{E}_{(r,\upsilon)}\left[\prod_{j=1}^{N} w(r_j,\upsilon_j)\right]}$$
$$= \frac{\sigma_{\mathtt{f}}(r,\upsilon)}{\sigma_{\mathtt{f}}^{\downarrow}(r,\upsilon)w(r,\upsilon)}\prod_{i=1}^{N} w(r_i,\upsilon_i). \tag{6.26}$$

$\updownarrow$*-Trees*

In a similar spirit to the previous section, we can look at the law of our NBP, when issued from a single ancestor, conditioned to have a subtree of prolific individuals. As such, for $A \in \mathfrak{F}_t$, we define

$$\mathbb{P}^{\updownarrow}_{\delta_{(r,\upsilon)}}(A|c_{\emptyset}(0)=\uparrow) = \frac{\mathbb{P}^{\updownarrow}_{\delta_{(r,\upsilon)}}(A;c_i=\uparrow,\ \text{for at least one } i=1,\ldots,N_t)}{\mathbb{P}^{\updownarrow}_{\delta_{(r,\upsilon)}}(c_{\emptyset}(0)=\uparrow)}$$
$$= \frac{\mathbb{E}_{\delta_{(r,\upsilon)}}\left[\mathbf{1}_A\left(1-\prod_{i=1}^{N_t} w(r_i(t),\upsilon_i(t))\right)\right]}{p(r,\upsilon)}. \tag{6.27}$$

In the next proposition, we will describe our NBP under $\mathbb{P}^{\updownarrow}_{\delta_{(r,\upsilon)}}(\cdot|c_{\emptyset}(0)=\uparrow)$. In order to do so, we first need to introduce a type-$\uparrow$-type-$\downarrow$ NBP.

Our type-$\uparrow$-type-$\downarrow$ NBP process, say $X^{\updownarrow} = (X^{\updownarrow}_t, t\geq 0)$, has an ancestor which is of type-$\uparrow$. We will implicitly assume (and suppress from the notation $X^{\updownarrow}$) that $X^{\updownarrow}_0 = \delta_{(r,\upsilon)}$ for $(r,\upsilon)\in D\times V$. Particles in $X^{\updownarrow}$ of type-$\uparrow$ move as a $\mathtt{P}^{\uparrow}$-Markov process, which we will formally define shortly. When a branching event occurs for a type-$\uparrow$ particle, both type-$\uparrow$ and type-$\downarrow$ particles may be produced, but always at least one type-$\uparrow$ is produced. Type-$\uparrow$ particles may be thought of as offspring, and any additional type-$\downarrow$ particles may be thought of as immigrants. Type-$\downarrow$ particles that are created can only subsequently produce type-$\downarrow$ particles in such a way that they give rise to copies of $(X, \mathbb{P}^{\downarrow})$.

The joint branching/immigration rate of type-$\uparrow$ and type-$\downarrow$ particles in $X^{\updownarrow}$ at $r \in D, \upsilon \in V$ is given by

$$\sigma_{\mathtt{f}}^{\updownarrow}(r,\upsilon) = \frac{\sigma_{\mathtt{f}}(r,\upsilon)}{p(r,\upsilon)} \mathcal{E}_{(r,\upsilon)}\left[1 - \prod_{j=1}^{N} w(r_j,\upsilon_j)\right]. \tag{6.28}$$

We can think of the branching rate in (6.28) as the original rate $\sigma_{\mathtt{f}}(r,\upsilon)$ multiplied by the probability (under $\mathcal{P}_{(r,\upsilon)}$) that at least one of the offspring is of type-$\uparrow$, given the branching particle is of type-$\uparrow$

At a branching/immigration event of a type-$\uparrow$ particle, we will write $N^{\uparrow}$ and $((r_i^{\uparrow}, \upsilon_i^{\uparrow}), i = 1, \cdots, N^{\uparrow})$ for the number and positions of type-$\uparrow$ offspring and $N^{\downarrow}$ and $((r_j^{\downarrow}, \upsilon_j^{\downarrow}), j = 1, \cdots, N^{\downarrow})$ for the number and positions of type-$\downarrow$ immigrants. We will write $(\mathcal{P}^{\updownarrow}_{(r,\upsilon)}, r \in D, \upsilon \in V)$ for the joint law of the random variables. Formally speaking, the fission mechanism, $\mathtt{G}^{\updownarrow}$, that describes the offspring/immigrants for a type-$\uparrow$ particle positioned at $r \in D, \upsilon \in V$ is written

$$\mathtt{G}^{\updownarrow}[f,g](r,\upsilon) = \sigma_{\mathtt{f}}^{\updownarrow}(r,\upsilon)\left(\mathcal{E}^{\updownarrow}_{(r,\upsilon)}\left[\prod_{i=1}^{N^{\uparrow}} f(r_i^{\uparrow},\upsilon_i^{\uparrow}) \prod_{j=1}^{N^{\downarrow}} g(r_j^{\downarrow},\upsilon_j^{\downarrow})\right] - f(r,\upsilon)\right), \tag{6.29}$$

for $f, g \in B_1^+(D \times V)$.

For each $r \in D, \upsilon \in V$, the law $\mathcal{P}^{\updownarrow}_{(r,\upsilon)}$ can be defined in terms of an additional random selection from $((r_i,\upsilon_i), i = 1, \cdots, N)$ under $\mathcal{P}_{(r,\upsilon)}$. Write $\mathcal{N}^{\uparrow}$ for the set of indices in $\{1, \cdots, N\}$ that identify the type-$\uparrow$ particles, i.e., $((r_i,\upsilon_i), i \in \mathcal{N}^{\uparrow}) = ((r_j^{\uparrow},\upsilon_j^{\uparrow}), j = 1, \cdots, N^{\uparrow})$. The remaining indices $\{1, \cdots, N\} \setminus \mathcal{N}^{\uparrow}$ then identify the type-$\downarrow$ immigrants from $((r_i,\upsilon_i), i = 1, \cdots, N)$. Thus, to describe $\mathcal{P}^{\updownarrow}_{(r,\upsilon)}$, for any $r \in D, \upsilon \in V$, it suffices to give the law of $(N; (r_1,\upsilon_1), \ldots, (r,_N \upsilon_N); \mathcal{N}^{\uparrow})$. To this end, for $F \in \sigma(N; (r_1,\upsilon_1), \ldots, (r_N,\upsilon_N))$ and $I \subseteq \mathbb{N}$, we will set

$$\begin{aligned}
&\mathcal{P}^{\updownarrow}_{(r,\upsilon)}(F \cap \{\mathcal{N}^{\uparrow} = I\}) \\
&\quad := \mathbf{1}_{\{|I| \geq 1\}} \frac{\sigma_{\mathtt{f}}(r,\upsilon)}{\sigma_{\mathtt{f}}^{\updownarrow}(r,\upsilon) p(r,\upsilon)} \mathcal{E}_{(r,\upsilon)}\Bigg[\mathbf{1}_{F \cap \{I \subseteq \{1,\ldots,N\}\}} \prod_{i \in I} p(r_i,\upsilon_i) \\
&\qquad\qquad \times \prod_{i \in \{1,\ldots,N\} \setminus I} w(r_i,\upsilon_i)\Bigg].
\end{aligned} \tag{6.30}$$

Said another way, for all $I \subseteq \mathbb{N}$,

$$\mathscr{P}^{\updownarrow}_{(r,\upsilon)}(\mathscr{N}^{\uparrow} = I|\sigma(N;(r_1,\upsilon_1),\ldots,(r_N,\upsilon_N)))$$
$$:= \mathbf{1}_{\{|I|\geq 1\}\cap\{I\subseteq\{1,\ldots,N\}\}}\frac{\prod_{i\in I} p(r_i,\upsilon_i)\prod_{i\in\{1,\ldots,N\}\setminus I} w(r_i,\upsilon_i)}{1-\mathscr{E}_{(r,\upsilon)}\left[\prod_{j=1}^N w(r_j,\upsilon_j)\right]}. \tag{6.31}$$

The pairs $((r_i^{\uparrow},\upsilon_i^{\uparrow}), i = 1,\cdots,N^{\uparrow})$ and $((r_j^{\downarrow},\upsilon_j^{\downarrow}), j = 1,\cdots,N^{\downarrow})$ under $(\mathscr{P}^{\updownarrow}_{(r,\upsilon)}, r\in D, \upsilon\in V)$ in (6.29) can thus be seen as equal in law to selecting the type of each particle following an independent sample of the fission offspring $((r_1\upsilon_1),\ldots,(r_N,\upsilon_N))$ under $\mathscr{P}_{(r,\upsilon)}$, where each (r_k,υ_k) is independently assigned either as type-$\uparrow$ with probability $p(r_k,\upsilon_k)$ or as type-$\downarrow$ with probability $w(r_k,\upsilon_k) = 1-p(r_k,\upsilon_k)$, but also conditional on there being at least one type-$\uparrow$.

As such with the definitions above, it is now a straightforward exercise to identify the fission mechanism in (6.29) in terms of $((r_i,\upsilon_i), i = 1,\cdots,N)$ under $(\mathscr{P}_{(r,\upsilon)}, r\in D, \upsilon\in V)$ via the following identity:

$$\begin{aligned}&\mathtt{G}^{\updownarrow}[f,g](r,\upsilon)\\&=\frac{\sigma_{\mathtt{f}}(r,\upsilon)}{p(r,\upsilon)}\mathscr{E}_{(r,\upsilon)}\Bigg[\sum_{\substack{I\subseteq\{1,\ldots,N\}\\|I|\geq 1}}\prod_{i\in I}p(r_i,\upsilon_i)f(r_i,\upsilon_i)\prod_{i\in\{1,\ldots,N\}\setminus I}w(r_i,\upsilon_i)g(r_i,\upsilon_i)\Bigg]\\&\qquad -\sigma_{\mathtt{f}}^{\updownarrow}(r,\upsilon)f(r,\upsilon).\end{aligned} \tag{6.32}$$

Lemma 6.5 (Dressed $\uparrow$-Trees) *For $r\in D, \upsilon\in V$, the process $X^{\updownarrow}$ is equal in law to X under $\mathbb{P}^{\updownarrow}_{\delta_{(r,\upsilon)}}(\cdot|c_{\emptyset}(0)=\uparrow)$. Moreover, both are equal in law to a dressed NBP, say $X^{\uparrow}$, whose motion semigroup and fission mechanism are given by $\mathtt{P}^{\uparrow}$ and $\mathtt{G}^{\uparrow}$, respectively. The motion semigroup $\mathtt{P}^{\uparrow}$ corresponds to the Markov process (R,Υ) on $(D\times V)\cup\{\dagger\}$ with probabilities $(\mathbf{P}^{\uparrow}_{(r,\upsilon)}, r\in D, \upsilon\in V)$ given by (recalling that p is valued 0 on $\dagger$)*

$$\left.\frac{\mathrm{d}\mathbf{P}^{\uparrow}_{(r,\upsilon)}}{\mathrm{d}\mathbf{P}_{(r,\upsilon)}}\right|_{\sigma((R_s,\Upsilon_s),s\leq t)} = \frac{p(R_t,\Upsilon_t)}{p(r,\upsilon)}\exp\left(-\int_0^t\frac{\mathtt{G}[w](R_s,\Upsilon_s)}{p(R_s,\Upsilon_s)}\mathrm{d}s\right), \qquad t\geq 0, \tag{6.33}$$

and the fission mechanism is given by

$$\mathtt{G}^{\uparrow}[f] = \frac{1}{p}\left(\mathtt{G}[pf+w]-(1-f)\mathtt{G}[w]\right), \qquad f\in B_1^+(D\times V). \tag{6.34}$$

The dressing consists of additional particles, which are immigrated non-locally in space at the branch points of $X^{\uparrow}$, with each immigrated particle continuing to evolve as an independent copy of $(X^{\downarrow},\mathbb{P}^{\downarrow})$ from their respective space-point *of immigration, such that the joint branching/immigration mechanism of type-*$\uparrow$ *offspring and type-*$\downarrow$ *immigrants is given by* (6.32).

Theorem 6.5 (Skeletal Decomposition) *Suppose that* $\mu = \sum_{i=1}^{n} \delta_{(r_i,\upsilon_i)}$*, for* $n \in \mathbb{N}$ *and* $(r_1, \upsilon_1), \cdots, (r_n, \upsilon_n) \in D \times V$*. Then* $(X, \mathbb{P}^{\updownarrow}_{\mu})$ *is equal in law to*

$$\sum_{i=1}^{n} \left(\Theta_i X_t^{i,\updownarrow} + (1 - \Theta_i) X_t^{i,\downarrow} \right), \qquad t \geq 0, \tag{6.35}$$

where, for each $i = 1, \ldots, n$*,* Θ_i *is an independent Bernoulli random variable with the probability of success given by*

$$p(r_i, \upsilon_i) := 1 - w(r_i, \upsilon_i), \tag{6.36}$$

and the processes $X^{i,\downarrow}$ *and* $X^{i,\updownarrow}$ *are independent copies of* $(X, \mathbb{P}^{\downarrow}_{\delta_{(r_i,\upsilon_i)}})$ *and* $(X, \mathbb{P}^{\updownarrow}_{\delta_{(r_i,\upsilon_i)}}(\cdot | c_{\emptyset}(0) = \uparrow))$*, respectively.*

As alluded to previously, Theorem 6.5 pertains to a classical decomposition of branching trees in which the process (6.35) describes how the NBP divides into the genealogical lines of descent which are "prolific" (surviving with probability p), in the sense that they create eternal subtrees which never leave the domain and those which are "unsuccessful" (dying with probability w), in the sense that they generate subtrees in which all genealogies die out.

Remark 6.2 It is an easy consequence of Theorem 6.5 that, for $t \geq 0$, the law of $X_t^{\uparrow}$ conditional on $\mathfrak{F}_t = \sigma(X_s, s \leq t)$ is equal to that of a Binomial point process with intensity $p(\cdot)X_t(\cdot)$. The latter, written $\text{BinPP}(pX_t)$, is an atomic random measure given by

$$\text{BinPP}(pX_t) = \sum_{i=1}^{N_t} \Theta_i \delta_{(r_i(t),\upsilon_i(t))},$$

where (we recall) that $X_t = \sum_{i=1}^{N_t} \delta_{(r_i(t),\upsilon_i(t))}$, and Θ_i is a Bernoulli random variable with probability $p(r_i(t), \upsilon_i(t))$, $i = 1, \cdots, N_t$.

Remark 6.3 It is also worth noting that the skeleton process $X^{\uparrow}$, given above, necessarily has at least one type-$\uparrow$ offspring at each branch point and indeed might have exactly one type-$\uparrow$ offspring (although possibly with other simultaneous type-$\downarrow$ immigrants). As such, an alternative way of looking at the type-$\uparrow$ process would be to think of the skeleton of prolific individuals as an NBP with fission mechanism $\mathtt{G}^{\Uparrow}$ that produces *at least two* type-$\uparrow$ offspring at each branch point and with a modified motion $\mathtt{P}^{\Uparrow}$ (in place of $\mathtt{P}^{\uparrow}$) which integrates the event of a single type-$\uparrow$ as an additional scattering in the movement. However, note these additional jumps are special in the sense as they are also potential points of simultaneous immigration of type-$\downarrow$ particles, unlike other jumps corresponding to $\mathtt{P}^{\uparrow}$ where there is no type-$\downarrow$ immigration.

Remark 6.4 As with the spine decomposition, we can understand (heuristically) the motions of $X^{\uparrow}$ and $X^{\downarrow}$ through the action of their generators. By considering only the leading order terms in small time of the process $(X_t)_{t\geq 0}$, the action of the generator can be seen as the result of the limit

$$\mathtt{L}f = \lim_{t\downarrow 0} \frac{1}{t}\left(\mathtt{P}_t[f] - f\right), \tag{6.37}$$

for suitably smooth f. Again, it has been shown (cf., e.g., [30, 33]) that the action of the generator corresponding to $\mathbf{P}$ is given by

$$\mathtt{L}f(r,\upsilon) = \upsilon\cdot\nabla_r f(r,\upsilon) + \int_V \big(f(r,\upsilon') - f(r,\upsilon)\big)\,\sigma_{\mathtt{S}}(r,\upsilon)\pi_{\mathtt{S}}(r,\upsilon,\upsilon')\mathrm{d}\upsilon', \tag{6.38}$$

for $f \in B^+(D\times V)$ such that $\nabla_r f$ is well defined. We emphasise again that, in view of Remark 6.3, this corresponds to motion plus a branching event with one offspring (or scattering).

The change of measure (6.23) induces a generator action given by

$$\begin{aligned}
\mathtt{L}^{\downarrow}f(r,\upsilon) &= \frac{1}{w(r,\upsilon)}\mathtt{L}(wf)(r,\upsilon) + f(r,\upsilon)\frac{\mathtt{G}[w]}{w}(r,\upsilon)\\
&= \upsilon\cdot\nabla_r f(r,\upsilon) + \int_V \big(f(r,\upsilon') - f(r,\upsilon)\big)\,\sigma_{\mathtt{S}}(r,\upsilon)\frac{w(r,\upsilon')}{w(r,\upsilon)}\\
&\quad\times\pi_{\mathtt{S}}(r,\upsilon,\upsilon')\mathrm{d}\upsilon'\\
&\quad+ f(r,\upsilon)\left(\frac{\mathtt{L}w}{w} + \frac{\mathtt{G}[w]}{w}\right)(r,\upsilon)\\
&= \upsilon\cdot\nabla_r f(r,\upsilon) + \int_V \big(f(r,\upsilon') - f(r,\upsilon)\big)\,\sigma_{\mathtt{S}}(r,\upsilon)\frac{w(r,\upsilon')}{w(r,\upsilon)}\\
&\quad\times\pi_{\mathtt{S}}(r,\upsilon,\upsilon')\mathrm{d}\upsilon',
\end{aligned} \tag{6.39}$$

where the fact that the right-hand side of (6.23) is a martingale will lead to $\mathtt{L}w + \mathtt{G}[w] = 0$.

In other words, our heuristic reasoning above shows that the motion on the $\downarrow$-marked tree is tantamount to a w-tilting of the scattering kernel. This tilting favours scattering in a direction where extinction becomes more likely, and as such, $\mathtt{L}^{\downarrow}$ encourages $\downarrow$-marked trees to become extinct "quickly".

Almost identical reasoning shows that the change of measure (6.33) has generator with action

$$
\begin{aligned}
\mathtt{L}^{\uparrow} f(r, \upsilon) &= \frac{1}{p(r, \upsilon)} \mathtt{L}(pf)(r, \upsilon) - f(r, \upsilon) \frac{\mathtt{G}[w]}{p}(r, \upsilon) \\
&= \upsilon \cdot \nabla_r f(r, \upsilon) + \int_V \big(f(r, \upsilon') - f(r, \upsilon)\big) \sigma_{\mathtt{s}}(r, \upsilon) \frac{p(r, \upsilon')}{p(r, \upsilon)} \\
&\quad \times \pi_{\mathtt{s}}(r, \upsilon, \upsilon') \mathrm{d}\upsilon',
\end{aligned} \tag{6.40}
$$

for suitably smooth f, where we have again used $\mathtt{L}w + \mathtt{G}[w] = 0$ and left the calculations that the second equality follows from the first as an exercise for the reader. One sees again a p-tilting of the scattering kernel, and hence $L^{\uparrow}$ rewards scattering in directions that "enable survival". Note that, moreover, for regions of $D \times V$ for which $p(r, \upsilon)$ can be come arbitrarily small (corresponding to a small probability of survival), the scattering rate also becomes very large, and hence $L^{\uparrow}$ "urgently" scatters particles away from such regions.

6.5 Comments

Additive martingales are inherently natural objects that appear in all branching processes. In the setting of Bienaymé–Galton–Watson processes, the additive martingale takes the simple form of the number of individuals alive in the n-th generation, divided by its mean. In more general settings, they rely on the existence of an appropriate eigenvalue and eigenfunction, from which they can be built. Martingale convergence theorems are core to the theory of branching processes in general and Theorem 6.1, proved in [74], is a typical contribution in this respect.

Additive martingales and the spine decomposition that comes with them have found favour as a tool to prove other results for a variety of spatial branching processes. See for example [58, 112, 124] and [53]. Their use for the analysis of neutronics is non-existent prior to the papers of Horton et al. [74] and Cox et al. [32]. The result on the spine decomposition in this chapter is taken from [74].

The strong law of large numbers in Theorem 6.3 is also a result of classical interest for branching processes and a natural extension of martingale convergence. As alluded to earlier in this chapter, whereas the asymptotic behaviour of the neutron transport equation provides average behaviour, Theorem 6.3 and the remark beneath it develop stochastic behaviour, and this explains the difference between stabilisation in mean and yet almost sure extinction at criticality. In the nuclear literature, this phenomenon is called the "critical catastrophe", a phrase coined by Williams [131], which results in "patchyness" of particle clusterings, cf. Dumonteil et al. [45]. As we will see later in Chap. 12, this is a general phenomenon for general branching processes.

Similarly to spine decompositions, skeleton path decompositions are absent from the neutronics literature prior to Harris et al. [69], from where the results in this chapter are taken. More details on and references for the use of spine and skeletal decompositions in branching processes are given in the comments and the end of Chap. 11. In addition, the origin and influences of the proofs of the spine and skeletal decompositions for a more general setting than the one specified in this chapter are also discussed.

Chapter 7
Generational Evolution

The eigenvalue problem in Theorem 4.1 is not the only one that offers insight into the evolution of neutron transport. In this chapter we will consider two different time-independent, or stationary, eigenvalue problems. The one that is of most interest to the nuclear industry is known as the $k_{\mathtt{eff}}$-eigenvalue problem. Roughly speaking, the eigenvalue $k_{\mathtt{eff}}$ has the physical interpretation as being the ratio of neutrons produced during fission events to the number lost due to absorption, either at the boundary or in the reactor due to neutron capture. As such, it characterises a different type of growth to the eigenvalue problem considered in the continuous-time setting in Theorem 4.1. In this chapter, in a similar fashion to the time-dependent setting, we will explore the probabilistic interpretation of k_{eff} and its relation to the classical abstract Cauchy formulation.

7.1 $k_{\mathtt{eff}}$-Eigenvalue Problem

In previous calculations, we worked with the backward transport, scatter, and fission operators defined by the arrangement of the operators in (3.7). In order to state the $k_{\mathtt{eff}}$-eigenvalue problem, we need to consider different arrangements of these operators, again keeping to the backwards setting, as we will largely engage in probabilistic analysis. We need to work with the dual operators, in the $L^2(D \times V)$ sense (see (3.3)), to the forward transport, scatter, and fission operators, $\mathscr{T}$, $\mathscr{S}$, and $\mathscr{F}$. To this end, define

$$\begin{cases} \hat{\mathscr{T}} f(r, \upsilon) := \upsilon \cdot \nabla f(r, \upsilon) - \sigma(r, \upsilon) f(r, \upsilon) \\ \\ \hat{\mathscr{S}} f(r, \upsilon) := \sigma_{\mathtt{s}}(r, \upsilon) \int_V f(r, \upsilon')\pi_{\mathtt{s}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon' \\ \\ \hat{\mathscr{F}} f(r, \upsilon) := \sigma_{\mathtt{f}}(r, \upsilon) \int_V f(r, \upsilon')\pi_{\mathtt{f}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon', \end{cases} \tag{7.1}$$

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8_7

where the cross sections were defined in Chap. 1.

As mentioned above, the $k_{\texttt{eff}}$-eigenvalue problem is a time-independent, or stationary, problem and corresponds to finding a value $k_{\texttt{eff}} > 0$ and a function $0 \le \omega \in L^2(D \times V)$ such that for $(r, \upsilon) \in D \times V$,

$$\hat{\mathscr{T}}\omega(r,\upsilon) + \hat{\mathscr{S}}\omega(r,\upsilon) + \frac{1}{k_{\texttt{eff}}}\hat{\mathscr{F}}\omega(r,\upsilon) = 0. \tag{7.2}$$

As in Chap. 1, we understand the equality to hold on $L^2(D \times V)$.

The parameter $k_{\texttt{eff}}$ can be interpreted as a measure of the typical number of neutrons produced from one generation to the next. From (7.2), we see that, by dividing the fission operator by the eigenvalue $k_{\texttt{eff}}$, this quantity establishes the balance between fission and loss of mass due to neutrons leaving a boundary or neutron capture. Note that, heuristically speaking, apart from losing neutrons to a physical boundary, the transport and scatter operators preserve the total mass.

With this interpretation, the regime $k_{\texttt{eff}} > 1$ corresponds to a supercritical system, $k_{\texttt{eff}} = 1$ corresponds to a critical system, and $k_{\texttt{eff}} < 1$ corresponds to a subcritical system. Moreover, $k_{\texttt{eff}} = 1$ precisely when $\lambda_* = 0$ in the time-dependent problem, and the right eigenfunctions φ and ω are equal as indeed are the left eigenfunctions.

Just as we have seen that the meaning of the time-dependent λ_*-eigenvalue problem can be phrased in terms of a multiplicative invariance with respect to the solution of an ACP (1.9), we can also consider the existence of classical solutions to the $k_{\texttt{eff}}$-eigenvalue problem (7.2).

In order to do this, it is necessary to consider the following ACP on $L^2(D \times V)$:

$$\begin{cases} \dfrac{\partial}{\partial t}u_t = (\hat{\mathscr{T}} + \hat{\mathscr{S}})u_t \\ \quad u_0 = g. \end{cases} \tag{7.3}$$

Then, just as in the spirit of (1.9) and Theorem 1.1, it is not difficult to show that the operator $(\hat{\mathscr{T}} + \hat{\mathscr{S}}, \mathrm{Dom}(\hat{\mathscr{T}} + \hat{\mathscr{S}}))$ generates a unique solution to (7.3) via the c_0-semigroup $(\Lambda_t, t \ge 0)$ given by

$$\Lambda_t g := \exp(t(\hat{\mathscr{T}} + \hat{\mathscr{S}}))g,$$

on $L^2(D \times V)$. Classical semigroup theory applied to the semigroup Λ_t then yields the existence of solutions to (7.2).

Theorem 7.1 *Suppose that (H1) holds, that the cross sections $\sigma_{\texttt{f}}\pi_{\texttt{f}}$ and $\sigma_{\texttt{s}}\pi_{\texttt{s}}$ are piecewise continuous, and that $\sigma_{\texttt{s}}(r,\upsilon)\pi_{\texttt{s}}(r,\upsilon,\upsilon') > 0$ and $\sigma_{\texttt{f}}(r,\upsilon)\pi_{\texttt{f}}(r,\upsilon,\upsilon') > 0$ on $D \times V \times V$. Then there exist a real eigenvalue $k_{\texttt{eff}} > 0$ and associated eigenfunction $0 \le \omega \in L^2(D \times V)$ such that (7.2) holds on $L^2(D \times V)$. Moreover, $k_{\texttt{eff}}$ can be explicitly identified as*

$$k_{\mathtt{eff}} = \sup\left\{k : (\hat{\mathscr{T}} + \hat{\mathscr{S}})\omega + \frac{1}{k}\hat{\mathscr{F}}\omega = 0 \, \textit{for some } \omega \in L^2(D \times V)\right\} \qquad (7.4)$$

and is both algebraically and geometrically simple.

However, our main objective in this chapter is to identify the eigenvalue problem (7.2) in terms of a probabilistic semigroup. For this, we will need to introduce the notion of generational times.

7.2 Generation Time

In order to develop a probabilistic interpretation of the $k_{\mathtt{eff}}$-eigenvalue problem via expectation semigroups, we first need to consider a generational model of the NBP. To this end, let us think of each line of descent in the sequence of neutron creations as a genealogy. In place of $(X_t, t \geq 0)$, we consider the process $(\mathscr{X}_n, n \geq 0)$, where, for $n \geq 1$, $\mathscr{X}_n$ is $\mathscr{M}_c(D \times V)$-valued and can be written

$$\mathscr{X}_n = \sum_{i=1}^{N^{(n)}} \delta_{(r_i^{(n)}, \upsilon_i^{(n)})},$$

where $\{(r_i^{(n)}, \upsilon_i^{(n)}), i = 1, \cdots N^{(n)}\}$ are the position-velocity configurations of the $N^{(n)}$ particles that are in the n-th generation of their genealogical lines of descent. We will use "n-th generation" to mean the collection of neutrons that are produced at the n-th fission event in each genealogical line of descent form the initial ancestor at time zero. In either case, $\mathscr{X}_0$ is consistent with X_0 and is the initial configuration of neutron positions and velocities. As such, for $n \geq 1$, we can think of $\mathscr{X}_n$ as the n-th generation of the system and we refer to $\mathscr{X} = (\mathscr{X}_n, n \geq 1)$ as the neutron generational process (NGP). The reader who is more experienced with the theory of branching processes will know that $\mathscr{X}_n$ is an example of what is called a stopping line.

Appealing to the obvious meaning of $\mathscr{X}_n[g]$, we define the expectation semigroup $(\Phi_n, n \geq 0)$ by

$$\Phi_n[g](r, \upsilon) = \mathbb{E}_{\delta_{(r,\upsilon)}}\left[\mathscr{X}_n[g]\right], \qquad n \geq 0, r \in D, \upsilon \in V, \qquad (7.5)$$

with $\Phi_0[g] := g \in B^+(D \times V)$. To see that this is indeed a semigroup, for $n \geq 0$, let $\mathfrak{K}_n$ denote the sigma algebra generated by the process up to the n-th generation. Fixing $m, n \geq 0$, we have

$$\Phi_{n+m}[g](r, \upsilon) = \mathbb{E}_{\delta_{(r,\upsilon)}}[\mathbb{E}[\mathscr{X}_{n+m}[g] | \mathfrak{K}_n]]$$

$$
\begin{aligned}
&= \mathbb{E}_{\delta_{(r,\upsilon)}}\left[\sum_{i=1}^{N^{(n)}} \mathbb{E}_{\delta_{(r_i^{(n)},\upsilon_i^{(n)})}}[\langle g, \mathcal{X}_m\rangle]\right] \\
&= \Phi_n[\Phi_m[g]](r,\upsilon). \qquad (7.6)
\end{aligned}
$$

We would like to formulate the eigenvalue problem (7.2) in terms of the semigroup $(\Phi_n, n \geq 0)$. We now introduce the problem of finding the (pre-emptively named) pair $k_{\mathtt{eff}} > 0$ and $\omega \in B^+(D \times V)$ such that

$$
\Phi_1[\omega](r,\upsilon) = k_{\mathtt{eff}}\omega(r,\upsilon), \qquad r \in D, \upsilon \in V. \qquad (7.7)
$$

We will shortly show the existence of a solution to (7.7), which plays an important role in the asymptotic behaviour of Φ_n as $n \to \infty$. Before doing so, let us give a heuristic argument as to why (7.7) is another form of the eigenvalue problem (7.2).

First recall that $\sigma_{\mathtt{s}}$ is the instantaneous rate at which scattering occurs and that $\sigma_{\mathtt{f}}(r,\upsilon)\pi_{\mathtt{f}}(r,\upsilon,\upsilon')\mathrm{d}\upsilon'$ is the instantaneous rate at which fission occurs, contributing average flux with velocity υ'. Writing $((R_s, \Upsilon_s), s \geq 0)$ with probabilities $(\mathbf{P}^{\sigma_{\mathtt{s}}\pi_{\mathtt{s}}}_{(r,\upsilon)}, r \in D, \upsilon \in V)$ for the $\sigma_{\mathtt{s}}\pi_{\mathtt{s}}$-NRW, by conditioning on the first fission time, we get, for $r \in D$, $\upsilon \in V$, and $g \in B^+(D \times V)$,

$$
\begin{aligned}
\Phi_n[g](r,\upsilon) &= \mathbf{E}^{\sigma_{\mathtt{s}}\pi_{\mathtt{s}}}_{(r,\upsilon)}\Bigg[\int_0^\infty \mathbf{1}_{(s<\tau_D)}\sigma_{\mathtt{f}}(R_s,\Upsilon_s)\mathrm{e}^{-\int_0^s \sigma_{\mathtt{f}}(R_u,\Upsilon_u)\mathrm{d}u} \\
&\qquad\qquad \times \int_V \pi_{\mathtt{f}}(R_s,\Upsilon_s,\upsilon')\Phi_{n-1}[g](R_s,\upsilon')\mathrm{d}\upsilon'\,\mathrm{d}s\Bigg] \\
&= \mathbf{E}^{\sigma_{\mathtt{s}}\pi_{\mathtt{s}}}_{(r,\upsilon)}\left[\int_0^\infty \mathbf{1}_{(s<\tau_D)}\mathrm{e}^{-\int_0^s \sigma_{\mathtt{f}}(R_u,\Upsilon_u)\mathrm{d}u}\hat{\mathscr{F}}\Phi_{n-1}[g](R_s,\Upsilon_s)\,\mathrm{d}s\right],
\end{aligned}
$$

where we have used (7.1). This tells us that Φ_n solves the mild equation

$$
\Phi_n[g](r,\upsilon) = \int_0^\infty \mathtt{Q}_s\left[\hat{\mathscr{F}}\Phi_{n-1}[g]\right](r,\upsilon)\mathrm{d}s, \qquad (7.8)
$$

where $(\mathtt{Q}_s, s \geq 0)$ is given by

$$
\mathtt{Q}_s[g](r,\upsilon) = \mathbf{E}^{\sigma_{\mathtt{s}}\pi_{\mathtt{s}}}_{(r,\upsilon)}\left[\mathrm{e}^{-\int_0^s \sigma_{\mathtt{f}}(R_u,\Upsilon_u)\mathrm{d}u} g(R_s,\Upsilon_s)\mathbf{1}_{(s<\tau_D)}\right], \qquad (7.9)
$$

and $\tau_D = \inf\{t > 0 : R_t \in \partial D \text{ and } \mathbf{n}_{R_t} \cdot \Upsilon_t > 0\}$. Informally speaking, $(\mathtt{Q}_s, s \geq 0)$ is the expectation semigroup associated with the operator $\hat{\mathscr{T}} + \hat{\mathscr{S}}$. To see why, recall that $\sigma = \sigma_{\mathtt{f}} + \sigma_{\mathtt{s}}$ and hence

$$
\begin{aligned}
&(\hat{\mathscr{T}} + \hat{\mathscr{S}}) f(r, \upsilon) \\
&\quad = \upsilon \cdot \nabla f(r, \upsilon) - \sigma(r, \upsilon) f(r, \upsilon) + \sigma_{\mathtt{s}}(r, \upsilon) \int_V f(r, \upsilon') \pi_{\mathtt{s}}(r, \upsilon, \upsilon') \mathrm{d}\upsilon' \\
&\quad = \upsilon \cdot \nabla f(r, \upsilon) + \sigma_{\mathtt{s}}(r, \upsilon) \int_V [f(r, \upsilon') - f(r, \upsilon)] \pi_{\mathtt{s}}(r, \upsilon, \upsilon') \mathrm{d}\upsilon' - \sigma_{\mathtt{f}}(r, \upsilon),
\end{aligned}
$$

which is the infinitesimal generator of a $\sigma_{\mathtt{s}}\pi_{\mathtt{s}}$-NRW with killing rate $\sigma_{\mathtt{f}}$ as in (7.9); see the discussion around (2.11) for generator heuristics.

If the pair $(k_{\mathtt{eff}}, \omega)$ solves (7.7), an iteration of the semigroup property (7.6) implies that

$$
k_{\mathtt{eff}}^n \omega(r, \upsilon) = \Phi_n[\omega](r, \upsilon), \qquad r \in D, \upsilon \in V.
$$

Substituting this into (7.8) and dividing through by $k_{\mathtt{eff}}^n$ yield

$$
\omega(r, \upsilon) = \int_0^\infty \mathtt{Q}_s \left[\frac{1}{k_{\mathtt{eff}}} \hat{\mathscr{F}} \omega \right] (r, \upsilon) \mathrm{d}s. \tag{7.10}
$$

Now set

$$
V_t[g](r, \upsilon) := \int_0^t \mathtt{Q}_s \, [g] \, (r, \upsilon) \mathrm{d}s. \tag{7.11}
$$

The heuristic Feynman–Kac formula that was discussed for differential equations of the type (2.18) tells us that, since $\mathtt{Q}$ is associated with the generator $\hat{\mathscr{T}} + \hat{\mathscr{S}}$, we should expect that V_t "solves" the equation

$$
\frac{\partial V_t}{\partial t} = (\hat{\mathscr{T}} + \hat{\mathscr{S}}) V_t + g, \tag{7.12}
$$

with $V_0 = 0$. From (7.11), $\partial V_t / \partial t = \mathtt{Q}_t[g]$, which tends to zero as $t \to \infty$ thanks to the transience of (R, Υ); that is to say, (R, Υ) will eventually be killed. Hence, taking $g = k_{\mathtt{eff}}^{-1} \hat{\mathscr{F}} \omega$, letting $t \to \infty$ in (7.12) and (7.11), and using the identity (7.10), we get (7.2).

Conversely, suppose that (k, ω) solves (7.2). Combining this with (7.8) yields

$$
\Phi_1[\omega](r, \upsilon) = \int_0^\infty \mathtt{Q}_s[\hat{\mathscr{F}} \omega](r, \upsilon) \mathrm{d}s = -k \int_0^\infty \mathtt{Q}_s[(\hat{\mathscr{T}} + \hat{\mathscr{S}}) \omega](r, \upsilon) \mathrm{d}s.
$$

Now suppose we can make rigorous the claim that $\mathtt{Q}$ is the semigroup associated with the operator $\hat{\mathscr{T}} + \hat{\mathscr{S}}$. This would (heuristically) imply that $\mathtt{Q}_s = \mathrm{e}^{s(\hat{\mathscr{T}} + \hat{\mathscr{S}})}$, which would in turn yield

$$\Phi_1[\omega](r, \upsilon) = -k \int_0^\infty \mathrm{e}^{s(\hat{\mathscr{T}}+\hat{\mathscr{S}})}[(\hat{\mathscr{T}} + \hat{\mathscr{S}})\omega](r, \upsilon)\mathrm{d}s.$$

Naïvely performing the integration on the right-hand side above and again using the fact that the $\sigma_{\mathtt{s}}\pi_{\mathtt{s}}$-NRW is transient, so that $\lim_{s\to\infty} \mathrm{e}^{s(\hat{\mathscr{T}}+\hat{\mathscr{S}})}\omega = 0$, show that (k, ω) satisfies (7.7).

7.3 Many-to-One Representation

In this section, we construct a many-to-one formula associated with the semigroup $(\Phi_n, n \geq 0)$ in the spirit of the continuous-time counterpart given in Chap. 4. For ease of notation, let

$$m(r, \upsilon) := \int_V \pi_{\mathtt{f}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon' \tag{7.13}$$

denote the mean number of neutrons generated by a fission event at (r, υ) and consider a $\sigma\theta$-NRW, where

$$\theta(r, \upsilon, \upsilon') = \frac{\sigma_{\mathtt{s}}(r, \upsilon)}{\sigma(r, \upsilon)}\pi_{\mathtt{s}}(r, \upsilon, \upsilon') + \frac{\sigma_{\mathtt{f}}(r, \upsilon)}{\sigma(r, \upsilon)}\frac{\pi_{\mathtt{f}}(r, \upsilon, \upsilon')}{m(r, \upsilon)}, \qquad r \in D, \upsilon, \upsilon' \in V. \tag{7.14}$$

We can think of the $\sigma\theta$-NRW as equal in law to the following process. For $k \geq 1$, when the NRW (R, Υ) scatters for the k-th time at (r, υ) (with rate $\sigma(r, \upsilon)$), a coin is tossed and the random variable $\mathtt{I}_k(r, \upsilon)$ takes the value 1 with probability $\sigma_{\mathtt{f}}(r, \upsilon)/\sigma(r, \upsilon)$ and its new velocity is selected according to an independent copy of the random variable $\Theta_k^{\mathtt{f}}(r, \upsilon)$, whose distribution has probability density $\pi_{\mathtt{f}}(r, \upsilon, \upsilon')/m(r, \upsilon)$. On the other hand, with probability $\sigma_{\mathtt{s}}(r, \upsilon)/\sigma(r, \upsilon)$, the random variable $\mathtt{I}_k(r, \upsilon)$ takes the value 0 and its new velocity is selected according to an independent copy of the random variable $\Theta_k^{\mathtt{s}}(r, \upsilon)$, whose distribution has probability density $\pi_{\mathtt{s}}(r, \upsilon, \upsilon')$. As such, the velocity immediately after the k-th scatter of the NRW, given that the position-velocity configuration immediately before is (r, υ), is coded by the random variable

$$\mathtt{I}_k(r, \upsilon)\Theta_k^{\mathtt{f}}(r, \upsilon) + (1 - \mathtt{I}_k(r, \upsilon))\Theta_k^{\mathtt{s}}(r, \upsilon).$$

We can thus identify sequentially $T_0 = 0$ and, for $n \geq 1$,

$$T_n = \inf\{t > T_{n-1} : \Upsilon_t \neq \Upsilon_{t-} \text{ and } \mathtt{I}_{k_t}(R_t, \Upsilon_{t-}) = 1\}, \tag{7.15}$$

where $(k_t, t \geq 0)$ is the process counting the number of scattering events of the $\sigma\theta$-NRW up to time t. We can think of the above description as giving us a marked version of the $\sigma\theta$-NRW, in the spirit of Poisson thinning. As previously mentioned in Sect. 3.3, let us, for convenience, denote the law of this marked $\sigma\theta$-NRW by $\mathbf{P}^{\sigma\theta}_{(r,\upsilon)}$, $r \in D, \upsilon \in V$.

Note, for the above construction of indicators to make sense, we should at least have some region of space for which fission can take place. As such, the assumption (H3) becomes relevant here. Analogously to Lemma 4.1, we have the following many-to-one formula associated with the NGP.

Lemma 7.1 (Generational Many-to-One) *Suppose (H1), (H2), and (H3) hold. The solution to* (7.8) *among the class of expectation semigroups is unique for* $g \in B^+(D \times V)$ *and the semigroup* $(\Phi_n, n \geq 0)$ *may alternatively be represented*[1] *as*

$$\Phi_n[g](r, \upsilon) = \mathbf{E}^{\sigma\theta}_{(r,\upsilon)}\left[\prod_{i=1}^{n} m(R_{T_i}, \Upsilon_{T_i-})g(R_{T_n}, \Upsilon_{T_n})\mathbf{1}_{(T_n<\kappa^D)}\right], \tag{7.16}$$

for $r \in D, \upsilon \in V, n \geq 1$, *(with* $\Phi_0[g] = g$*), where* $(R_t, \Upsilon_t)_{t\geq 0}$ *is the* $\sigma\theta$*-NRW marked at times* $(T_i, i \geq 1)$*,* $m(r, \upsilon)$ *was defined in* (7.13)*, and*

$$\kappa^D := \inf\{t > 0 : R_t \in \partial D \text{ and } \mathbf{n}_{R_t} \cdot \Upsilon_t > 0\}.$$

Proof We first note that the sequence $(\Phi_n, n \geq 0)$ as defined in (7.16) is a semigroup since, due to the strong Markov property, we have

$$\begin{aligned}
&\Phi_{n+m}[g](r, \upsilon)\\
&= \mathbf{E}^{\sigma\theta}_{(r,\upsilon)}\left[\mathbf{E}^{\sigma\theta}\left[\prod_{i=1}^{n+m} m(R_{T_i}, \Upsilon_{T_i-})g(R_{T_{n+m}}, \Upsilon_{T_{n+m}})\mathbf{1}_{(T_{n+m}<\kappa^D)}\middle|\mathfrak{K}_n\right]\right]\\
&= \mathbf{E}^{\sigma\theta}_{(r,\upsilon)}\left[\prod_{i=1}^{n} m(R_{T_i}, \Upsilon_{T_i-})\mathbf{E}^{\sigma\theta}_{(R_{T_n},\Upsilon_{T_n})}\right.\\
&\qquad \left.\times\left[\prod_{i=1}^{m} m(R_{T_i}, \Upsilon_{T_i-})g(R_{T_m}, \Upsilon_{T_m})\mathbf{1}_{(T_m<\kappa^D)}\right]\mathbf{1}_{(T_n<\kappa^D)}\right]\\
&= \Phi_n[\Phi_m[g]](r, \upsilon), \qquad r \in D, \upsilon \in V.
\end{aligned}$$

In order to show that Φ_n defined in (7.16) does indeed solve (7.8), we consider the process at time T_1. Before doing so, we first note that the $\sigma\theta$-NRW has the same dynamics as the $\sigma_{\mathtt{s}}\pi_{\mathtt{s}}$-NRW over the time interval $[0, T_1)$ and, at time T_1, which occurs at rate $\sigma_{\mathtt{f}}$, the new velocity of the $\sigma\theta$-NRW is chosen according to the expectation operator

[1] Here, we define $\prod_{i=1}^{0} \cdot := 1$.

$$\hat{\mathscr{E}}[g](r,\upsilon) := \int_V g(r,\upsilon')\frac{\pi_{\mathtt{f}}(r,\upsilon,\upsilon')}{m(r,\upsilon)}\mathrm{d}\upsilon'.$$

Applying the strong Markov property at time T_1,

$$\begin{aligned}
&\Phi_n[g](r,\upsilon)\\
&= \mathbf{E}^{\sigma\theta}_{(r,\upsilon)}\left[\prod_{i=1}^n m(R_{T_i},\Upsilon_{T_i-})g(R_{T_n},\Upsilon_{T_n})\mathbf{1}_{(T_n<\kappa^D)}\right]\\
&= \mathbf{E}^{\sigma\theta}_{(r,\upsilon)}\left[m(R_{T_1},\Upsilon_{T_1-})\hat{\mathscr{E}}[\Phi_{n-1}[g]](R_{T_1},\Upsilon_{T_1-})\mathbf{1}_{(T_1<\kappa^D)}\right]\\
&= \int_0^\infty \mathbf{E}^{\sigma\theta}_{(r,\upsilon)}\left[\sigma_{\mathtt{f}}(R_s,\Upsilon_s)\mathrm{e}^{-\int_0^s\sigma_{\mathtt{f}}(R_u,\Upsilon_u)\mathrm{d}u}m(R_s,\Upsilon_{s-})\right.\\
&\quad\left.\times\hat{\mathscr{E}}[\Phi_{n-1}[g]](R_s,\Upsilon_{s-})\mathbf{1}_{(s<\kappa^D)}\right]\mathrm{d}s\\
&= \int_0^\infty \mathtt{Q}_s[\hat{\mathscr{F}}\Phi_{n-1}[g]](r,\upsilon)\mathrm{d}s,
\end{aligned}$$

where the final equality follows from the fact that $m\sigma_{\mathtt{f}}\hat{\mathscr{E}} = \hat{\mathscr{F}}$.

It remains to show that (7.8) has a unique solution for $g \in B^+(D\times V)$ among the class of expectation semigroups. Suppose that $(\Phi'_n, n\geq 0)$ is another such solution with $\Phi'_0 = g \in B^+(D\times V)$. Define $\Upsilon_n = \Phi_n - \Phi'_n$, for $n\geq 0$, and note by linearity that $(\Upsilon_n, n\geq 0)$ is another expectation semigroup with $\Upsilon_0 = 0$. Moreover, by linearity, $(\Upsilon_n, n\geq 0)$ also solves (7.8). On account of this, it is straightforward to see by induction that if $\Upsilon_n = 0$ then $\Upsilon_{n+1} = 0$. The uniqueness of (7.8) in the class of expectation semigroups thus follows. □

7.4 Perron–Frobenius Asymptotics

In this section we show that the semigroup $(\Phi_n, n\geq 0)$ exhibits a Perron–Frobenius decomposition in a similar spirit to Chap. 4.

Theorem 7.2 (Generational Perron–Frobenius Asymptotic) *Suppose (H1) and (H4) hold in addition to*

(H3)** **The fission cross section satisfies** $\inf_{r\in D,\upsilon,\upsilon'\in V}\sigma_{\mathtt{f}}(r,\upsilon)\pi_{\mathtt{f}}(r,\upsilon,\upsilon') > 0.$

For the semigroup $(\Phi_n, n\geq 0)$ identified by (7.8), there exist $k_{\mathtt{eff}}\in\mathbb{R}$, a positive[2] *right eigenfunction $\omega\in B^+(D\times V)$, and a left eigenmeasure which is*

[2] To be precise, by a positive eigenfunction, we mean a mapping from $D\times V\to(0,\infty)$. This does not prevent it being valued zero on ∂D, as D is open.

absolutely continuous with respect to the Lebesgue measure on $D \times V$ with positive density $\tilde{\omega} \in B^+(D \times V)$, both having associated eigenvalue $k_{\mathtt{eff}}^n$, and such that ω (respectively, $\tilde{\omega}$) is uniformly (respectively, a.e. uniformly) bounded away from zero on each compactly embedded subset of $D \times V$. Moreover, $k_{\mathtt{eff}}$ is the leading eigenvalue in the sense that, for all $g \in B^+(D \times V)$,

$$\langle \tilde{\omega}, \Phi_n[g] \rangle = k_{\mathtt{eff}}^n \langle \tilde{\omega}, g \rangle \quad \text{and} \quad \Phi_n[\omega] = k_{\mathtt{eff}}^n \omega, \quad n \geq 0, \tag{7.17}$$

and there exists $\gamma > 1$ such that, for all $g \in B^+(D \times V)$,

$$\sup_{g \in B_1^+(D \times V)} \left\| k_{\mathtt{eff}}^{-n} \omega^{-1} \Phi_n[g] - \langle \tilde{\omega}, g \rangle \right\| = O(\gamma^{-n}) \text{ as } n \to +\infty. \tag{7.18}$$

As in the continuous-time setting, to prove the above result, we will work with a variant of the semigroup $(\Phi_n, n \geq 0)$ in order to, then, apply Theorem 2.2. To this end, note that under the assumption (H4), for non-negative functions $g \in B^+(D \times V)$, we have

$$\mathbb{E}_{\delta_{(r,\upsilon)}}[\langle g, \mathscr{X}_1 \rangle] \leq \|g\| \mathbb{E}_{\delta_{(r,\upsilon)}}[\langle 1, \mathscr{X}_1 \rangle] \leq M \|g\|, \tag{7.19}$$

where

$$M = \sup_{r \in D, \upsilon \in V} \int_V \pi_{\mathtt{f}}(r, \upsilon, \upsilon') \mathrm{d}\upsilon' \leq n_{\mathtt{max}}.$$

Dividing both sides of (7.19) by M yields a sub-Markovian semigroup. To see why, let $\Gamma = \min\{n \geq 0 : \mathtt{K}_n(R_{T_n}, \Upsilon_{T_n -}) = 1\}$ where for $n \geq 0$, $r \in D$, and $\upsilon \in V$, the random variable $\mathtt{K}_n(r, \upsilon)$ is an independent indicator random variable which is equal to 0 with probability $m(r, \upsilon)/M$. With this in hand, we can write

$$\begin{aligned}
\Phi_n^{\dagger}[g](r, \upsilon) &:= M^{-n} \Phi_n[g](r, \upsilon) \\
&= \mathbf{E}_{(r,\upsilon)}^{\sigma\theta} \left[\prod_{i=1}^{n} \frac{m(R_{T_i}, \Upsilon_{T_i -})}{M} g(R_{T_n}, \Upsilon_{T_n}) \mathbf{1}_{(T_n < \kappa^D)} \right] \\
&= \mathbf{E}_{(r,\upsilon)}^{\sigma\theta} \left[g(R_{T_n}, \Upsilon_{T_n}) \mathbf{1}_{(T_n < \kappa^D,\, n < \Gamma)} \right] \\
&=: \mathbf{E}_{(r,\upsilon)}^{\sigma\theta,\dagger} \left[g(R_{T_n}, \Upsilon_{T_n}) \right].
\end{aligned} \tag{7.20}$$

In a similar manner to the continuous-time case, we prove Theorem 7.2 by extracting the eigentriple from the semigroup $\Phi^{\dagger}$, defined in (7.20), by using the discrete time counterparts of assumptions (A1) and (A2), which we state here. In order to state them, recalling the notation in (7.20), we define

$$\mathtt{k} = \Gamma \wedge \min\{n \geq 1 : T_n \geq \kappa^D\}.$$

Then, in this setting, (A1) and (A2) read as follows.

(A1)* There exist $n_0, c_1 > 0$ such that for each $(r, \upsilon) \in D \times V$,

$$\mathbf{P}^{\sigma\theta}_{(r,\upsilon)}((R_{T_{n_0}}, \Upsilon_{T_{n_0}}) \in \cdot \,|n_0 < \mathtt{k}) \geq c_1 \nu(\cdot).$$

(A2)* There exists a constant $c_2 > 0$ such that for each $(r, \upsilon) \in D \times V$ and for every $n \geq 0$,

$$\mathbf{P}^{\sigma\theta}_{\nu}(n < \mathtt{k}) \geq c_2 \mathbf{P}^{\sigma\theta}_{(r,\upsilon)}(n < \mathtt{k}).$$

Theorem 7.3 *Assume that (H1), (H3**), and (H4) are in force. Then, there exists an eigenvalue $k_c \in (0, 1)$ accompanied by an eigenmeasure on $D \times V$ with a positive density $\tilde{\omega} \in B^+(D \times V)$ and a positive right eigenfunction $\omega \in B^+(D \times V)$ of $\Phi^\dagger_n$ (defined in (7.20)), i.e., for all $g \in B^+(D \times V)$,*

$$\langle \tilde{\omega}, \Phi^\dagger_n[g] \rangle = k_c^n \langle \omega, g \rangle \quad \textit{and} \quad \Phi^\dagger_n[\omega] = k_c^n \omega, \quad n \geq 0. \tag{7.21}$$

Moreover, there exist $C, \gamma > 0$ such that, for $g \in B^+(D \times V)$ and $n \geq 1$ (independently of g),

$$\left\| k_c^{-n} \omega^{-1} \Phi^\dagger_n[g] - \langle \tilde{\omega}, g \rangle \right\| \leq C\gamma^{-n} \|g\|. \tag{7.22}$$

In particular, setting $g \equiv 1$,

$$\left\| k_c^{-n} \omega^{-1} \mathbf{P}^{\sigma\theta}_{\cdot}(n < \mathtt{k}) - 1 \right\| \leq C\gamma^{-n}, \quad n \geq 1. \tag{7.23}$$

As alluded to above, once Theorem 7.3 is proved, it is straightforward to conclude that ω and $\tilde{\omega}$ are the right and left eigenfunctions corresponding to the eigenvalue $k_{\mathtt{eff}} = k_c M$ for the semigroup Φ_n.

Proof of Theorem 7.3 We will use the notation J_k to denote the k-th scatter event of the random walk (R, Υ) under $\mathbf{P}^{\sigma\theta}$ and recall that T_k denotes the scatter event that corresponds to the k-th fission event in the original NBP. The basis of our proof relies on the fact that, for each $k \geq 1$, $T_k = J_k$ with positive probability.

Another feature of our proof is that we will use a version of Lemma 4.4. Indeed we may take the conclusion of this lemma replacing the law $\mathbf{P}^\dagger$ by $\mathbf{P}^{\sigma\theta,\dagger}$ (the latter defined in (7.20)). We restate Lemma 4.4 in this new format for convenience. □

Lemma 7.2 *Under the assumptions of Theorem 7.3, for all $r \in D$ and $\upsilon \in V$, we have*

$$\mathbf{P}^{\sigma\theta,\dagger}_{(r,\upsilon)}(J_7 < \mathtt{k}, R_{J_7} \in \mathrm{d}z) \leq C \mathbf{1}_{(z \in D)}\, \mathrm{d}z,$$

for some constant $C > 0$, and

$$\mathbf{P}_{\nu}^{\sigma\theta,\dagger}(J_1 < \mathtt{k}, R_{J_1} \in \mathrm{d}z) \geq c\mathbf{1}_{(z \in D)}\, \mathrm{d}z, \tag{7.24}$$

for another constant $c > 0$, where ν is the Lebesgue measure on $D \times V$.

By proving (A1*) and (A2*), we will get the statement of the theorem albeit that the left eigenmeasure does not necessarily have a positive density. However, as in the continuous-time case, that is, Theorem 4.1, it is possible to show that the left eigenmeasure admits a density $\tilde{\omega} \in B^+(D \times V)$, such that, on each compactly embedded subset of $D \times V$, ω (respectively, $\tilde{\omega}$) is positive (respectively, positive almost everywhere). Since the techniques used are similar to those of Lemma 4.5, we leave this part of the proof as an exercise to the reader. Hence, to complete our proof of Theorem 7.3, we focus on proving (A1*) and (A2*).

Proof of (A1)*

Fix $r_0 \in D$ and suppose Υ_0 is uniformly distributed on V. The assumptions (H1) and (H3**) tell us that fission occurs everywhere in the configuration space. With these assumptions in hand, the techniques used in the proof of Theorem 4.1 to prove (4.51) and the discussion thereafter also yield a similar estimate

$$\mathbf{E}_{(r_0,\Upsilon_0)}^{\sigma\theta,\dagger}\left[f(R_{J_1})\mathbf{1}_{(T_1=J_1<\mathtt{k})}\right] \geq C_0 \int_D \mathrm{d}z f(z), \tag{7.25}$$

where $C_0 > 0$ is a constant.

Recall the (deterministic) quantity $\kappa_{r_0,\upsilon_0}^D := \inf\{t > 0 : r_0 + \upsilon_0 t \notin D\}$, for $r_0 \in D$, $\upsilon_0 \in V$, which was introduced in (3.13). From (H3**), the fact that σ is uniformly bounded from above, the strong Markov property, and (7.25), we have

$$\begin{aligned}
&\mathbf{E}_{(r_0,\upsilon_0)}^{\sigma\theta,\dagger}[f(R_{T_2}, \Upsilon_{T_2})\mathbb{1}_{(T_2=J_2<\mathtt{k})}] \\
&\geq C_1 \int_0^{\kappa_{r_0,\upsilon_0}^D} \mathrm{d}s \mathrm{e}^{-\bar{\sigma}s} \underline{\theta} \int_V \mathrm{d}\upsilon_1 \mathbf{E}_{(r_0+\upsilon_0 s,\upsilon_1)}^{\sigma\theta,\dagger}[f(R_{J_1}, \Upsilon_{J_1})\mathbb{1}_{(T_1=J_1<\mathtt{k})}] \\
&\geq C_2 \kappa_{r_0,\upsilon_0}^D \int_D \mathrm{d}r \int_V \mathrm{d}\upsilon f(r,\upsilon),
\end{aligned} \tag{7.26}$$

where $\overline{\sigma} = \sup_{r\in D,\upsilon\in V} \sigma(r,\upsilon)$ and $\underline{\theta} = \inf_{r\in D,\upsilon,\upsilon'\in V} \theta(r,\upsilon,\upsilon')$. The latter is bounded from below because of (H3**). Finally, we note that, due to (H1) and (H3**), $\underline{\sigma} = \inf_{r\in D,\upsilon\in V} \sigma(r,\upsilon) > 0$. Hence, again using (H1), we have

$$\mathbf{P}_{(r_0,\upsilon_0)}^{\sigma\theta}(T_2 < \mathtt{k}) \leq \mathbf{P}^{\sigma\theta}(J_1 < \mathtt{k}) \leq \int_0^{\kappa_{r_0,\upsilon_0}^D} \mathrm{d}s \bar{\sigma} \mathrm{e}^{-\underline{\sigma}s} \leq C_3 \kappa_{r_0,\upsilon_0}^D. \tag{7.27}$$

Combining this with (7.26) yields (A1*) with ν as Lebesgue measure on $D \times V$ and $n_0 = 2$. $\square$

Proof of (A2)*

Again, we use a similar method to the one used in Chap. 4, and however, we state the proof in full to illustrate where the differences occur. Let $n \geq 7$ and note that $T_n - J_7 \geq T_n - T_7$. This and the strong Markov property imply

$$\begin{aligned}\mathbf{P}^{\sigma\theta}_{(r,\upsilon)}(n < \mathtt{k}) &\leq \mathbf{E}^{\sigma\theta,\dagger}_{(r,\upsilon)}\left[\mathbf{P}^{\sigma\theta}_{(R_{J_7},\Upsilon_{J_7})}(n-7 < \mathtt{k})\right] \\ &\leq C' \int_D \int_V \mathbf{P}^{\sigma\theta}_{(z,w)}(n-7 < \mathtt{k})\, \mathrm{d}z\, \mathrm{d}w, \end{aligned} \tag{7.28}$$

where we have used Lemma 7.2 to obtain the final inequality.

Now suppose $n \geq 1$. Recalling the measure ν from (A1*), another application of Lemma 7.2 gives

$$\begin{aligned}\mathbf{P}^{\sigma\theta}_{\nu}(n < \mathtt{k}) &= \mathbf{E}^{\sigma\theta,\dagger}_{\nu}\left[\mathbf{1}_{(J_1<\mathtt{k})}\mathbf{P}^{\sigma\theta}_{(R_{J_1},\Upsilon_{J_1})}(n < \mathtt{k})\right] \\ &\geq c' \int_D \int_V \mathbf{P}^{\sigma\theta}_{(z,w)}(n < \mathtt{k})\, \mathrm{d}z\, \mathrm{d}w. \end{aligned} \tag{7.29}$$

Then, for $n \geq 8$, combining (7.28) and (7.29) yields

$$\mathbf{P}^{\sigma\theta}_{(r,\upsilon)}(n < \mathtt{k}) \leq \frac{C'}{c'}\mathbf{P}^{\sigma\theta}_{\nu}(n-7 < \mathtt{k}). \tag{7.30}$$

It follows from (A1*) that

$$\mathbf{P}^{\sigma\theta,\dagger}_{\nu}((R_{T_{n_0}}, \Upsilon_{T_{n_0}}) \in \cdot) \geq c_1 \mathbf{P}^{\sigma\theta}_{\nu}(n_0 < \mathtt{k})\nu(\cdot). \tag{7.31}$$

Again, due to assumptions (H1) and (H3**),

$$\mathbf{P}^{\sigma\theta}_{\nu}(n_0 < \mathtt{k}) \geq \int_{D\times V} \mathbf{P}^{\sigma\theta}_{(r,\upsilon)}(T_{n_0} = J_{n_0}, n_0 < \mathtt{k})\nu(\mathrm{d}r, \mathrm{d}\upsilon) \geq K, \tag{7.32}$$

for some constant $K > 0$. Then, for $n \geq 8$, due to (7.31) and (7.32),

$$\begin{aligned}\mathbf{P}^{\sigma\theta}_{\nu}(n-7+n_0 < \mathtt{k}) &= \mathbf{E}^{\sigma\theta}_{\nu}\left[\mathbf{1}_{(n_0<\mathtt{k})}\mathbf{P}^{\sigma\theta}_{(R_{T_{n_0}},\Upsilon_{T_{n_0}})}(n-7 < \mathtt{k})\right] \\ &\geq K c_1 \mathbf{P}^{\sigma\theta}_{\nu}(n-7 < \mathtt{k}). \end{aligned} \tag{7.33}$$

Finally, noting that for $n \geq 1$, we have $n - 7 + 4n_0 \geq n$, so that

$$\mathbf{P}^{\sigma\theta}_{\nu}(n < \mathtt{k}) \geq \mathbf{P}^{\sigma\theta}_{\nu}(n - 7 + 4n_0 < \mathtt{k}),$$

and applying (7.33) four times implies

$$\mathbf{P}^{\sigma\theta}_{\nu}(n < \mathtt{k}) \geq (Kc_1)^4 \mathbf{P}^{\sigma\theta}_{\nu}(n - 7 < \mathtt{k}). \tag{7.34}$$

Combining this with (7.30) yields the result. □

7.5 Moment Growth

Just as in the continuous-time setting, it is possible to examine the growth of moments for the different regimes, i.e., we can find analogues of Theorems 5.2, 5.3, and 5.4 for generational time. Before proceeding, we need a little more notation. To this end, we need to introduce the law of the point process of velocities when fission occurs but relative to the configuration of the incident neutron when it came into creation, rather than immediately before undergoing fission. The latter point process was previously denoted by Z, with probabilities $(\mathscr{P}_{(r,\upsilon)}, r \in D, \upsilon \in V)$.

Let us write $\mathsf{Z}_{(r,\upsilon)}$ as a notational convenience for (Z), when it has law $\mathscr{P}_{(r,\upsilon)}$. We thus want to define the probabilities $(P_{(r,\upsilon)}, r \in D, \upsilon \in V)$ for the point process

$$\mathscr{Z} = \sum_{i=1}^{N} \delta_{\upsilon_i},$$

such that the law of $\mathscr{Z}$ under $P_{(r,\upsilon)}$ agrees precisely with the law of

$$\mathsf{Z}_{(R_{\tau_1}, \Upsilon_{\tau_1})} \text{ under } \mathbf{P}^{\sigma_{\mathrm{s}}\pi_{\mathrm{s}}}_{(r,\upsilon)},$$

where τ_1 is the first time that a fission occurs. Another way of seeing this is that, for any $f \in B^+(V)$,

$$E_{(r,\upsilon)}\left[\mathrm{e}^{-\mathscr{Z}[f]}\right] = \mathbf{E}^{\sigma_{\mathrm{s}}\pi_{\mathrm{s}}}_{(r,\upsilon)}\left[\int_0^\infty \sigma_{\mathtt{f}}(R_s, \Upsilon_s)\mathrm{e}^{-\int_0^s \sigma_{\mathtt{f}}(R_u,\Upsilon_u)\mathrm{d}u}\mathscr{E}_{(R_s,\Upsilon_s)}\left[\mathrm{e}^{-\mathsf{Z}[f]}\right]\mathrm{d}s\right]. \tag{7.35}$$

In the following results, we see that in generational time, the results are almost identical to those in the continuous-time setting, given in Theorems 5.5, 5.6, and 5.7, albeit that one must take care to note that the triple $(k_{\mathtt{eff}}, \omega, \tilde{\omega})$ replaces the role of $(\lambda_*, \varphi, \tilde{\varphi})$ and $(\mathsf{Z}, \mathscr{P})$ is replaced by $(\mathscr{Z}, P)$.

Theorem 7.4 (Supercritical, $k_{\mathtt{eff}} > 1$) *Suppose that (H1), (H3**), and (H4) hold and $k_{\mathtt{eff}} > 1$. Define*

$$\Delta_n^{(\ell)} = \sup_{r\in D, \upsilon\in V, g\in B_1^+(D\times V)} \left| k_{\mathtt{eff}}^{-nk}\Phi_n^{(\ell)}[g](r,\upsilon) - \ell!\,\tilde{\omega}[f]^\ell L_\ell(r,\upsilon)\right|,$$

where $L_1 = \omega(r,\upsilon)$, and for $k \geq 2$, $L_k(r,\upsilon)$ is given by the recursion

$$L_k(r,\upsilon) = \sum_{\ell=0}^{\infty} k_{\mathtt{eff}}^{-\ell(k+1)}\Phi_\ell\left[E_\cdot\Big[\sum_{[k_1,\ldots,k_N]_k^{2+}}\prod_{\substack{j=1\\k_j>0}}^{N} L_{k_j}(\cdot,\upsilon_j)\Big]\right](r,\upsilon), \qquad (7.36)$$

with $[k_1,\ldots,k_N]_k^{2+}$ defining the set of non-negative tuples $(k_1,\ldots,k_N)$, such that $\sum_{j=1}^N k_j = N$ and at least two of the k_j are strictly positive.

Then, for all $\ell \leq k$,

$$\sup_{n\geq 0}\Delta_n^{(\ell)} < \infty \text{ and } \lim_{n\to\infty}\Delta_n^{(\ell)} = 0.$$

Theorem 7.5 (Subcritical, $k_{\mathtt{eff}} < 1$) *Suppose that (H1), (H3**), and (H4) hold and $k_{\mathtt{eff}} < 1$. Define*

$$\Delta_n^{(\ell)} = \sup_{r\in D, \upsilon\in V, g\in B_1^+(D\times V)} \left| k_{\mathtt{eff}}^{-n}\omega(r,\upsilon)^{-1}\Phi_n^{(\ell)}[g](r,\upsilon) - L_\ell\right|,$$

where $L_1 = 1$, and for $\ell \geq 2$, L_ℓ is given by the recursion

$$L_\ell = \langle\tilde{\omega}, f^\ell\rangle + \sum_{n=0}^{\infty} k_{\mathtt{eff}}^{-(n+1)}\left\langle\tilde{\omega}, E_\cdot\Big[\sum_{[k_1,\ldots,k_N]_k^{2+}}\binom{k}{k_1,\cdots,k_N}\prod_{\substack{j=1\\k_j>0}}^{N}\Phi_n^{(k_j)}(\cdot,\upsilon_j)\Big]\right\rangle.$$

Then, for all $\ell \leq k$,

$$\sup_{n\geq 0}\Delta_n^{(\ell)} < \infty \text{ and } \lim_{n\to\infty}\Delta_n^{(\ell)} = 0.$$

Theorem 7.6 (Critical, $k_{\mathtt{eff}} = 1$) *Suppose that (H1), (H3**), and (H4) hold and $k_{\mathtt{eff}} = 1$. Define*

$$\Delta_n^{(\ell)} = \sup_{r\in D, \upsilon\in V, g\in B_1^+(D\times V)} \Big| n^{-(\ell-1)}\omega(r,\upsilon)^{-1}\Phi_n^{(\ell)}[g](r,\upsilon)$$
$$-2^{-(\ell-1)}\ell!\,\langle\tilde{\omega}, f\rangle^\ell\langle\tilde{\omega}, \mathscr{V}[\omega]\rangle^{\ell-1}\Big|,$$

where, again with an abuse of notation in the continuous-time setting,

$$\mathscr{V}[\omega](r,\upsilon) = E_{(r,\upsilon)}\left[\mathscr{Z}[\omega]^2 - \mathscr{Z}[\omega^2]\right] = E_{(r,\upsilon)}\left[\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\omega(z_i)\omega(z_j)\right].$$

Then, for all $\ell \leq k$,

$$\sup_{n\geq 0}\Delta_n^{(\ell)} < \infty \text{ and } \lim_{n\to\infty}\Delta_n^{(\ell)} = 0. \tag{7.37}$$

We will not provide proofs for these results, as in the continuous-time case, as we will handle them again in a more general setting in Part II of this book.

7.6 c-Eigenvalue Problem

There is a third eigenvalue problem, which also appears in some of the nuclear literature, called the c-eigenvalue problem. It is similar to the $k_{\texttt{eff}}$-eigenvalue problem in the sense that it is a time-independent problem, and however there is a subtle difference that arises from requiring the eigenvalue, c, to weight the scattering operator, as well as the fission operator. More precisely, the aim is to find $c > 0$ and a corresponding eigenfunction $\chi \in B^+(D\times V)$ such that

$$\hat{\mathscr{T}}\chi(r,\upsilon) + \frac{1}{c}\left(\hat{\mathscr{S}} + \hat{\mathscr{F}}\right)\chi(r,\upsilon) = 0, \quad r\in D, \upsilon\in V, \tag{7.38}$$

where the operators $\hat{\mathscr{T}}$, $\hat{\mathscr{S}}$. and $\hat{\mathscr{F}}$ were defined in (7.1).

Similarly to the interpretation given for $k_{\texttt{eff}}$, the scalar c can be seen as the effective number of neutrons produced per "collision", where a collision is either a scatter or a fission. Criticality is also categorised in the same way as the $k_{\texttt{eff}}$-eigenvalue problem, and it can be shown that the critical regime also coincides with the critical regimes for the $k_{\texttt{eff}}$- and λ_*-eigenvalue problems, with equality of the left and right eigenfunctions. The c-eigenvalue problem has received the least attention in the literature due to its similarities with the $k_{\texttt{eff}}$-eigenvalue problem. Indeed, many of the methods used for the latter can be adapted to analyse the former.

As for the $k_{\texttt{eff}}$-eigenvalue problem, we can also think of the discrete-time index for the c-eigenvalue problem as generational time, where, now, for $n \geq 1$, generation n is interpreted as the collection of neutrons immediately after the moment that they engage in the n-th collision along their genealogical line of descent. By conditioning on the first collision event in the expectation given by (7.5), arguments similar to those leading to (7.8) show that Φ_n solves the mild equation

$$\Phi_n[g](r,\upsilon)=\int_0^\infty \mathtt{U}_s^\sigma\left[(\hat{\mathscr{F}}+\hat{\mathscr{S}})\Phi_{n-1}[g]\right](r,\upsilon)\mathrm{d}s, \tag{7.39}$$

for $r\in D$, $\upsilon\in V$, and $g\in B^+(D\times V)$, where $(\mathtt{U}_s^\sigma, s\geq 0)$ is given by

$$\begin{aligned}\mathtt{U}_s^\sigma[g](r,\upsilon)&=\mathrm{e}^{-\int_0^s\sigma(r+\upsilon u,\upsilon)\mathrm{d}u}g(r+\upsilon s,\upsilon)\mathbf{1}_{(s<\kappa^D_{r,\upsilon})}\\&=\mathrm{e}^{-\int_0^s\sigma(r+\upsilon u,\upsilon)\mathrm{d}u}\mathtt{U}_s[g](r,\upsilon).\end{aligned} \tag{7.40}$$

Then, if (c,χ) satisfies

$$\Phi_1[\chi](r,\upsilon)=c\chi(r,\upsilon),\qquad r\in D,\upsilon\in V, \tag{7.41}$$

one can use (7.40) and similar arguments to those given in the second half of Sect. 7.2 to show that (c,χ) also satisfies

$$0=\hat{\mathscr{T}}\chi+\frac{1}{c}(\hat{\mathscr{S}}+\hat{\mathscr{F}})\chi.$$

We now consider a $\sigma\pi$-NRW, where π was defined in (4.2). Letting $(\mathbf{P}^{\sigma\pi}_{(r,\upsilon)}, r\in D,\upsilon\in V)$ denote the law of the $\sigma\pi$-NRW and defining the weight

$$w(r,\upsilon)=\frac{\sigma_{\mathtt{s}}(r,\upsilon)}{\sigma(r,\upsilon)}+\frac{\sigma_{\mathtt{f}}(r,\upsilon)}{\sigma(r,\upsilon)}m(r,\upsilon)=\frac{\alpha(r,\upsilon)}{\sigma(r,\upsilon)},$$

where α was defined in (4.2), we have the following many-to-one representation:

$$\Phi_n[g](r,\upsilon)=\mathbf{E}^{\sigma\pi}_{(r,\upsilon)}\left[\prod_{i=1}^n w(R_{\tau_i},\Upsilon_{\tau_i-})g(R_{\tau_n},\Upsilon_{\tau_n})\mathbf{1}_{(\tau_n<\tau_D)}\right], \tag{7.42}$$

where now $(\tau_n, n\geq 1)$ denote the successive collision times of the $\sigma\pi$-NRW and $\tau_D=\inf\{t>0: R_t\notin D\}$. We note that the weight function w can be seen as a sharing of the average mass accumulated from scatter and fission events. Indeed, with probability $\sigma_{\mathtt{s}}/\sigma$, there is a scatter event which contributes mass 1, and with probability $\sigma_{\mathtt{f}}/\sigma$, there is a fission event which contributes mass m.

We will not include the proof of this many-to-one formula since the arguments are almost identical to those given in the proof of Lemma 7.1.

7.7 Comments

As alluded to earlier in this chapter, the $k_{\mathtt{eff}}$-eigenvalue and c-eigenvalue problems are motivated by different ways of seeing the stability of neutron transport that come from within the nuclear industry itself; cf. [95, 96, 127]. The proof of Theorem 7.1 can be found in [32].

The connection of the $k_{\texttt{eff}}$-eigenvalue problem with generational time has been (at least heuristically) well understood in the nuclear engineering literature, but the relationship that we give in the sketch proof of in Sect. 7.2 comes from [32]. The notion of generational populations (as stopping lines) is a classical concept in the theory of spatial branching processes; see [86].

The generational many-to-one (Lemma 7.1) and the Perron–Frobenius (Theorem 7.2) results come from [21, Theorem 2.1]. In this setting, it is interesting to note that the left eigenfunction $\tilde{\omega}$ does not correspond to the non-negative eigenfunction, say ω^*, of the forward equation $\mathscr{T}\omega^* + \mathscr{S}\omega^* + (k_{\rm eff})^{-1}\mathscr{F}\omega^* = 0$, where we recall that $\mathscr{T}$, $\mathscr{S}$, and $\mathscr{F}$ were defined in (1.8). Indeed, for each $n \geq 0$, up to discounting by $k^n_{\rm eff}$, the eigenfunction ω^* is the stationary distribution of the immediate ancestors of particles in χ_n at the instant preceding their fission. The relationship between $\tilde{\omega}$ and ω^* is given by $\tilde{\omega} = \mathscr{F}\omega^*$. This is a consequence of the fact that the operator $\mathscr{F}$ propagates pre-fission particles to post-fission particles. We refer the reader to [44] for further details. Features of generational moment growth in Theorems 7.4, 7.5, and 7.6 follow from a similar analysis to the real-time moment growth given in Theorems 5.2, 5.3, and 5.4 and are based on [67], although formal proofs are given in Chap. 9 later in Part II of this book.

In a similar fashion to [74], Theorem 7.2 implies that

$$\mathscr{W}_n := k_{\texttt{eff}}^{-n}\frac{\mathscr{X}_n[\omega]}{\mu[\omega]}, \qquad n \geq 0,$$

is a non-negative unit mean martingale under $\mathbb{P}_\mu$, $\mu \in \mathscr{M}_c(D \times V)$. As in the continuous-time limit (cf. Theorem 6.1), one can show that under relatively mild assumptions, $(\mathscr{W}_n, n \geq 0)$ converges in $L^2(\mathbb{P}_\mu)$ in the supercritical case and otherwise has a degenerate limit.

Again, under appropriate assumptions, one may also reconstruct a strong law of large numbers in the spirit of Theorem 6.3,

$$\lim_{n\to\infty} k_{\texttt{eff}}^{-n}\frac{\mathscr{X}_n[g]}{\mu[\omega]} = \langle g, \tilde{\omega}\rangle \mathscr{W}_\infty,$$

where $\mathscr{W}_\infty$ is the limit of the martingale $(\mathscr{W}_n, n \geq 0)$. We leave the details for a more general discussion in Part II of this book.

A challenge that has consumed part of the nuclear engineering literature (cf. [19, 96, 127]) is the problem of how to estimate the quantity $k_{\texttt{eff}}$ for an entire reactor core. As a single number, $k_{\texttt{eff}}$ gives a sense as to whether a reactor core is critical or not; the state of being critical is obviously a desirable trait for the purpose of power generation. The associated eigenvalue ω similarly gives a sense of how neutron density, and hence energy concentration or power generation, occurs within the reactor core. The eigenmeasure $\tilde{\omega}$ on the other hand gives a sense of (quasi)-stationarity given that $\langle \tilde{\omega}, k_{\texttt{eff}}^{-n}\mathscr{X}_n[g]\rangle = \langle \tilde{\omega}, g\rangle$, for $g \in B^+(D \times V)$.

For historical reasons, the nuclear industry has largely preferred to work with the $k_{\mathtt{eff}}$-eigenvalue problem over the λ_*-eigenvalue problem. Accordingly, a culture of Monte Carlo simulation has evolved which implicitly takes advantage of the generational theory of NBPs. We conclude this chapter with a brief discussion of how the analytical foundations of the $k_{\mathtt{eff}}$-eigenvalue problem form the mathematical basis of Monte Carlo algorithms, which ultimately feature in present day industrial software that is used to design and predict the behaviour of nuclear reactor cores.

The definition of $k_{\mathtt{eff}}$ means that

$$k_{\mathtt{eff}} = -\frac{\langle \mathbf{1}, \hat{\mathscr{F}}\omega\rangle}{\langle \mathbf{1}, (\hat{\mathscr{T}} + \hat{\mathscr{S}})\omega\rangle}, \tag{7.43}$$

where $\mathbf{1}$ is the constant function with value one. The identity (7.43) is suggestive of the following conceptual Monte Carlo approach. Start with a set of N neutrons, distributed in $D \times V$ according to some function $\omega^{(0)}$ that serves as an initial guess of ω. In practice, $\omega^{(0)}$ could be taken to be the uniform distribution or based on the solution to an appropriate diffusion approximation of the eigenvalue problem. From this initial configuration, a system of neutrons are simulated according to a NBP that is stochastically consistent with the cross sections $\sigma_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{s}}$, and $\pi_{\mathtt{f}}$ (in the spirit of the discussion preceding Lemma 3.3) until immediately after their first fission events occur (in effect, the first generation $\mathscr{X}_1$). From repeated simulations of $\mathscr{X}_1$, one may construct an updated estimate of ω, say $\omega^{(1)}$, by using the empirical law of $\mathscr{X}_1$ and normalising it to have unit mass. At the same time, the eigenvalue $k_{\mathtt{eff}}$ is approximated by

$$k^{(1)}_{\mathtt{eff}} = -\frac{\langle \mathbf{1}, \hat{\mathscr{F}}\omega^{(1)}\rangle}{\langle \mathbf{1}, (\hat{\mathscr{T}} + \hat{\mathscr{S}})\omega^{(1)}\rangle}.$$

The process is then repeated using $\omega^{(1)}$ as the initial distribution of neutrons, in order to obtain $\omega^{(2)}$ and $k^{(2)}_{\mathtt{eff}}$, and so on.

The starting point of the previously described algorithm is the analytical identity (7.43). A different starting point from which to build another Monte Carlo algorithm would be the observation that

$$k_{\mathtt{eff}} = \lim_{n\to\infty} \frac{1}{n} \log \Phi_n[1](r, \upsilon).$$

Here, as an expectation, $\Phi_n[1]$ can be approximated by Monte Carlo simulation of the generational process $(\mathscr{X}_n, n \geq 0)$, again by simulating a NBP which is stochastically consistent with the data $\sigma_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{s}}$, and $\pi_{\mathtt{f}}$.

In order to calculate the eigenfunction, one can take advantage of the ergodic property proved in Theorem 7.2 to deduce that, under the same assumptions as the aforesaid,

$$\langle\tilde{\omega}, g\rangle\omega(r, \upsilon) = \lim_{n\to\infty} \mathbb{E}_{\delta_{(r,\upsilon)}}\left[\frac{1}{n}\sum_{m=1}^{n} k_{\tt eff}^{-m}\langle\mathscr{X}_m, g\rangle\right].$$

This again suggests that Monte Carlo simulation of the generational process $(\mathscr{X}_n, n \geq 0)$ can be used to extract information concerning both ω and $\tilde{\omega}$. Varying the test function g while keeping (r, υ) fixed allows us to obtain estimates for $\tilde{\omega}$, whereas varying the initial configuration (r, υ) and keeping the test function g fixed allow us to estimate ω.

As one may imagine, there are numerous points in this prescriptive approach that have the potential to introduce bias and additional correlations between the neutrons in successive fission generations. Understanding how to quantify the complexity in such a scheme remains a significant outstanding mathematical challenge. We refer the reader to [19, 31] for a more in-depth analysis of the Monte Carlo algorithms described above, addressing issues such as burn in, bias, and complexity analysis. It is worthy of note that, although the algorithms and efficiency results given in [31] are for time eigenvalues, cf. (7.17), it is straightforward to see how they may be adapted to fit the generational setting (as well as in terms of complexity).

In the setting of generational Monte Carlo, there is a vast literature, too much to list here, with [96] as a core reference. It is worthy of note that, whilst much of the physics and engineering literature engages in various forms of Monte Carlo 'population control' (known in applied probability as interacting particle Monte Carlo) a systematic mathematical analysis that incorporates modern perspectives as laid out in, for example, [29] and [111], has yet to be fully explored.

Part II
Non-local Branching Markov Processes

Chapter 8
A General Family of Branching Markov Processes

Recall that, in Chap. 2, we introduced the notion of a general Markov process on E which is taken to be a locally compact Hausdorff space, to which we can append a cemetery state, †. We used the notation $\mathtt{P} = (\mathtt{P}_t, t \geq 0)$ to denote its associated semigroup and, accordingly, we later referred to it as a $\mathtt{P}$-Markov process. As a generalisation of the neutron branching processes discussed in Chap. 3, we are interested in spatial branching processes that are defined in terms of a $\mathtt{P}$-Markov process and a branching operator. In this chapter and subsequent chapters, we introduce such processes and develop a number of generic results for them. Some of the notations used for neutron branching processes (NBPs) will also be used in this general setting. This is deliberate to give the reader a chance to see how the former is an interesting core example of the latter.

8.1 Branching Markov Processes

Let us define what we mean by a branching Markov process (BMP). As with NBPs, a BMP is a collection of particles that evolves according to certain stochastic rules. Given their point of creation, particles move independently according to a $\mathtt{P}$-Markov process on E in the spirit of Chap. 2. In an event, which we refer to as "branching", particles positioned at x die at rate $\gamma(x)$, where $\gamma \in B^+(E)$, and instantaneously, new particles are created in E according to a point process. We can think of a point process simply as a random variable N, representing the number of offspring, and $(x_1, \cdots, x_N)$ in E representing their locations. A more convenient way of describing the random configuration of these offspring is via the use of random counting measures

$$\mathsf{Z}(A) = \sum_{i=1}^{N} \delta_{x_i}(A), \qquad A \in \mathscr{B}(E), \tag{8.1}$$

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8_8

where $\mathscr{B}(E)$ is the collection of Borel sets in E. As with the branching rate, the law of the aforementioned point process can depend on x, the point of death of the parent, and we denote it by $\mathcal{P}_x$, $x \in E$, with associated expectation operator given by $\mathcal{E}_x$, $x \in E$. Note that this is consistent with the notation in (3.23), where $E = D \times V$ and there is a slight degeneracy in that setting because, whilst in full generality, we should expect x and x_i to be replaced by (r, υ) and (r_i, υ_i), we always have $r_i = r$ in the NBP.

We capture the reproductive distributional information in the so-called branching mechanism

$$\mathtt{G}[f](x) := \gamma(x)\mathcal{E}_x\left[\prod_{i=1}^{N} f(x_i) - f(x)\right], \qquad x \in E, \tag{8.2}$$

where

$$f \in B_1^+(E) := \{f \in B^+(E) : \|f\| \leq 1\},$$

and we recall that $\gamma \in B^+(E)$. Here, we use $\|\cdot\|$ to be the usual supremum norm on $B^+(E)$. Without loss of generality, we can assume that $\mathcal{P}_x(N = 1) = 0$ for all $x \in E$ by viewing a branching event with one offspring as an extra jump in the motion described by $\mathtt{P}$. On the other hand, we do allow for the possibility that $\mathcal{P}_x(N = 0) > 0$ for some or all $x \in E$. In the setting that our BMP is an NBP, allowing $\mathcal{P}_x(N = 0) > 0$ corresponds to the possibility of neutron capture.

Note that the mechanism in (8.2) permits non-local branching in the sense that offspring are not necessarily positioned at the place of their parent's death. In the case of local branching, it reduces to

$$\gamma(x)\left[\sum_{i=1}^{\infty} p_k(x)s^k - s\right], \qquad s \in [0, 1], x \in E,$$

where, for $k \geq 1$ and $x \in E$, $p_k(x)$ denotes the probability that a particle branching at site x produces k offspring. We refer to $(p_k(x), k \geq 0)$ as the offspring distribution at site $x \in E$.

Henceforth we refer to this spatial branching process as a $(\mathtt{P}, \mathtt{G})$-branching Markov process. If the configuration of particles at time t is denoted by

$$\{x_1(t), \ldots, x_{N_t}(t)\},$$

then, on the event that the process has not become extinct or exploded, the branching Markov process can be described as the coordinate process $X = (X_t, t \geq 0)$ in $\mathcal{M}_c(E)$, where

$$X_t(\cdot) = \sum_{i=1}^{N_t} \delta_{x_i(t)}(\cdot), \qquad t \geq 0, \tag{8.3}$$

and we recall that

$$\mathscr{M}_c(E) := \{\textstyle\sum_{i=1}^n \delta_{x_i} : n \in \mathbb{N}, x_i \in E, i = 1, \cdots, n\} \tag{8.4}$$

represents the space of finite counting measures. In particular, X is Markovian in $\mathscr{M}_c(E)$. Its probabilities will be denoted $\mathbb{P} := (\mathbb{P}_\mu, \mu \in \mathscr{M}_c(E))$.

As an obvious extension to the notation we used in Part I of this book, we will work with

$$\mu[f] := \int_E f(x)\mu(\mathrm{d}x), \qquad f \in B^+(E),$$

for all finite measures μ. In particular, when $\mu \in \mathscr{M}_c(E)$, we note that

$$\mu[f] = \sum_{i=1}^n f(x_i) \text{ when } \mu = \sum_{i=1}^n \delta_{x_i}.$$

8.2 Non-linear Semigroup Evolution

As with the NBP, the functional

$$\mathrm{v}_t[f](x) = \mathbb{E}_{\delta_x}\left[\mathrm{e}^{-X_t[f]}\right], \qquad f \in B^+(E),\ t \geq 0, \tag{8.5}$$

is the natural analytical object that gives us a complete understanding of the law of our BMP.

Similarly to the NBP setting, we have the *branching Markov property*. That is, if we define

$$\mathfrak{F}_t = \sigma(x_i(s), i = 1, \cdots, N_s, s \leq t), \qquad t \geq 0,$$

then

$$\mathbb{E}\left[\mathrm{e}^{-X_{t+s}[f]}\middle|\, \mathfrak{F}_t\right] = \prod_{i=1}^{N_t} \mathrm{v}_s[f](x_i(t)). \tag{8.6}$$

From here the semigroup property,

$$\mathrm{v}_{t+s}[f](x) = \mathrm{v}_t[\mathrm{v}_s[f]](x), \qquad s, t \geq 0, x \in E, f \in B^+(E),$$

follows just as in Lemma 5.1. Moreover, for $f \in B^+(E)$ and $x \in E$,

$$\mathtt{v}_t[f](x) = \hat{\mathtt{P}}_t[\mathrm{e}^{-f}](x) + \int_0^t \mathtt{P}_s\left[\mathtt{G}[\mathtt{v}_{t-s}[f]]\right](x)\mathrm{d}s, \qquad t \geq 0, \tag{8.7}$$

where $(\hat{\mathtt{P}}_t, t \geq 0)$ is the adjusted semigroup which returns a value of 1 on the event of killing, i.e., when the particle is absorbed at the boundary, cf. (5.3). More generally, we can consider the joint law of the BMP and its occupation measure $\int_0^t X_s(\cdot)\mathrm{d}s$, given by

$$\mathtt{v}_t[f, g](x) = \mathbb{E}_{\delta_x}\left[\mathrm{e}^{-X_t[f] - \int_0^t X_s[g]\mathrm{d}s}\right], \quad t \geq 0, \ x \in E, \ f, g \in B^+(E), \tag{8.8}$$

which satisfies the evolution equation

$$\mathtt{v}_t[f, g](x) = \hat{\mathtt{P}}_t[\mathrm{e}^{-f}](x) + \int_0^t \mathtt{P}_s\left[\mathtt{G}[\mathtt{v}_{t-s}[f, g]) - g\mathtt{v}_{t-s}[f, g]\right](x)\mathrm{d}s. \tag{8.9}$$

For the proof of both (8.7) and (8.9), we can appeal to reasoning which is essentially the same as for the Pál-Bell equation (PBE) in Lemma 5.2. For example to derive (8.9), it suffices to split the expectation on the first branching event and apply Theorem 2.1. It is also straightforward to see that $\mathtt{v}_t$ is the unique solution to (8.9) when considered as an operator from $B^+(E) \times B^+(E)$ to $B^+(E)$, by appealing to similar reasoning given in the proof of Lemma 3.5.

8.3 Examples of Branching Markov Processes

To give a sense of the generality in the definition of our class of BMP, let us give some concrete examples.

Branching Lévy Processes Particles move according to a Lévy process, $Y = (Y_t, t \geq 0)$ on $E = \mathbb{R}^d$, and branch at constant rate γ. When particles branch, they simultaneously die and reproduce by throwing out offspring according to the point process $(x_i : i = 1, \ldots, N)$ relative to its point of death, with a common law $\mathscr{P}$, which is not spatially dependent. In that case,

$$\mathtt{G}[f](x) = \gamma\mathscr{E}\left[\prod_{i=1}^N f(x + x_i) - f(x)\right], \qquad f \in B_1^+(E), \ x \in E.$$

For BMPs that have an underlying Markov process that is regular, e.g., in this case a Lévy process or another Feller process, it is more usual to write (8.7) in the differential form

$$\frac{\partial}{\partial t}\mathrm{v}_t[f](x) = \mathscr{L}\mathrm{v}_t[f](x) + \mathsf{G}[\mathrm{v}_t[f]](x), \qquad t \geq 0,\ x \in \mathbb{R}^d, \tag{8.10}$$

where $\mathscr{L}$ is the generator of Y. One of the issues with writing the non-linear semigroup evolution as a solution to an integro-differential equation, rather than simply an integral equation, is that it requires a greater degree of smoothness of the solution so that, e.g., the term $\mathscr{L}\mathrm{v}_t[f]$ is meaningful, not to mention the time derivative.

Branching Brownian Motion In the case where the Lévy process Y is taken to be a Brownian motion (in any dimension) on some domain $E = \mathbb{R}^d$ and the branching mechanism is local with no spatial dependence, (8.10) collapses to

$$\frac{\partial}{\partial t}\mathrm{v}_t[f](x) = \frac{1}{2}\Delta\mathrm{v}_t[f](x)+\gamma\left[\sum_{k=1}^{\infty} p_k\mathrm{v}_t[f](x)^k - \mathrm{v}_t[f](x)\right], \quad t \geq 0,\ x \in \mathbb{R}^d,$$

where $(p_k, k \geq 1)$ is the offspring distribution. In particular, in one dimension with dyadic branching, we recover the Fisher–Kolmogorov–Petrsovskii–Piscunov (FKPP) equation

$$\frac{\partial}{\partial t}\mathrm{v}_t = \frac{1}{2}\frac{\partial^2}{\partial x^2}\mathrm{v}_t + \gamma\mathrm{v}_t(\mathrm{v}_t - 1), \qquad t \geq 0, \tag{8.11}$$

where the dependency on space and f is suppressed to allow (8.11) to take a more familiar form. As alluded to in the previous example, for the PDE (8.11) to make sense in the classical sense, we would technically need v_t to be twice continuously differentiable.

Uchiyama Process Another subclass of branching Lévy processes pertains to the setting where particles live in $\mathbb{R}^d$ but have no motion and constant branching rate. When branching occurs, each particle gives birth to offspring according to an independent copy of a point process in $\mathbb{R}^d$, which is centred at the parent's position. Although called a Uchiyama process after the author who introduced it, another way of referring to this process is simply a continuous-time d-dimensional branching random walk. This is because as one follows each genealogical line of descent, one sees a continuous-time random walk. In this setting, the non-linear evolution equation (8.7) is a variant of (8.10) and takes the more specific form

$$\frac{\partial}{\partial t}\mathrm{v}_t[f](x) = \gamma\mathscr{E}\left[\prod_{i=1}^{N}\mathrm{v}_t[f](x + x_i) - \mathrm{v}_t[f](x)\right], \qquad t \geq 0,\ x \in \mathbb{R}^d, \tag{8.12}$$

where $(x_i, i = 1, \cdots, N)$ is the i.i.d. point process of relative offspring positions with probabilities $\mathscr{P}$ that no longer depends on the state space. Once again, we gloss over issue of smoothness in the solution for (8.12) to make sense.

Biggins Process These processes are a generalisation of all of the preceding ones. Particles move according to a Markov process, $M := (M_t, t \geq 0)$ in $\mathbb{R}^d$, which we can assume to be Feller, and branch at a constant rate. Parent particles give birth to a random number of individuals according to the point process on $\mathbb{R}^d \times \mathbb{R}_+$, where the first coordinate describes the spatial displacement of the offspring relative to the parent's birth position, and the second coordinate gives the parent's age at the time of that child's birth. Technically speaking, the Biggins process incorporates time into the state space of particles and lies outside of our definition of BMPs (unless we consider time to be a spatial variable). Nonetheless, a lot of what we will present for BMPs can be extended to the setting of Biggins processes and leave this as a challenge to the ambitious reader.

Branching Markov Additive Process Consider the following Markov additive process (MAP). A particle moves according to one of n possible velocities in $\mathbb{R}^3$, say $\{\mathtt{a}_1, \cdots, \mathtt{a}_n\}$, with the "current" drift being chosen by a continuous-time Markov chain $J = (J_t, t \geq 0)$ on $[n] := \{1, \cdots, n\}$ in the sense that if $J_t = i$, then its drift is $\mathtt{a}_i$. More formally, we can describe this MAP as the joint process $((X_t, J_t), t \geq 0)$, where

$$\mathrm{d}X_t = \mathtt{a}_{J_t}\mathrm{d}t, \qquad t \geq 0.$$

The process X takes values in $\mathbb{R}^3$ and hence $E = \mathbb{R} \times [n]$. In a fully general setting, the branching rate γ depends on the pair (X, J), i.e., $\gamma = \gamma(x, i)$, $x \in \mathbb{R}^3$, $i \in [n]$. However, for convenience, let us simply restrict dependence to $[n]$ and write the respective rates $(\gamma_i, i \in [n])$. When a branching event occurs, the random number of offspring is assigned different velocities so that there are $N^{(i)}$ offspring with velocity $\mathtt{a}_i$, such that $\sum_{i=1}^n N^{(i)} = N$. As with the branching rate, let us assume that the law of this point process depends only on the current drift index in $[n]$. Accordingly, we will write $\mathcal{P}_i$, $i \in [n]$ for its probabilities.

If we are to follow the trend of the previous examples and write (8.7) in differential form, we have

$$\frac{\partial}{\partial t}\mathtt{v}_t[f](x,i) = \mathtt{a}_i \cdot \nabla \mathtt{v}_t[f](x,i) + \sum_{j\in[n]} \mathtt{Q}_{ij}\mathtt{v}_t[f](x,j)$$
$$+ \gamma_i \mathcal{E}_i\left[\prod_{j\in[n]}\prod_{k=1}^{N^{(j)}} \mathtt{v}_t[f](x,k) - \mathtt{v}_t[f](x,i)\right], \quad x \in \mathbb{R}^3, i \in [n], \tag{8.13}$$

where $\mathtt{Q}$ is the transition matrix of J. Again, (8.13) comes with the caution that, written in this differential-difference form, smoothness for the sake of, e.g., working with the operators $\mathtt{a}_i \cdot \nabla$ in a pointwise sense is an issue that requires attention and that (8.7) avoids.

Neutron Branching Process Finally, and most importantly for this text, the NBP is, of course, included in our definition of the general BMP. In this case the branching rate is given by $\gamma = \sigma_{\mathtt{f}}$. The motion semigroup $(\mathtt{P}_t, t \geq 0)$ corresponds to that of an NRW that scatters at rate $\sigma_{\mathtt{s}}$ and chooses its new velocity according to $\pi_{\mathtt{s}}$, with killing when the spatial component exits D; technically speaking, this is when the underlying Markov process (R, Υ) exits $D \times V$, as it will always do so by the process R exiting D. When considering the branching mechanism, one again needs to be careful to remember that the underlying state space is $D \times V$ and so, although fission events are local with respect to particle position, they are non-local in the velocity component. Hence, seen as a branching process on $D \times V$, the branching mechanism is non-local. The non-linear evolution equation (8.7) was given in Theorem 5.2, and due to our deliberate replicated use of notation, it appears identical.

It is worth noting that the NBP is closely related to the previous example in the sense that the neutron velocity space V is analogous to the collection $\{\mathtt{a}_1, \cdots, \mathtt{a}_n\}$. Hence one may think of the dependency of the PBE (5.7) on υ as analogous to the indexing of the solution by i in (8.13). The process of scattering for the NRW in (5.7) is also analogous to the role that the Markov chain J plays in (8.13). The analogy between (5.7) and (8.13) would be complete had we made γ and $\mathscr{P}$ depend fully on $i \in [n]$ and $x \in \mathbb{R}^3$ in (8.13).

8.4 Linear Semigroup Evolution and Many-to-One

In the setting of the NBP, we were very much driven by the desire to characterise solutions to the (mild) NTE in terms of a stochastic process. This resulted in the identification of its solution $(\psi_t, t \geq 0)$ as the mean of linear functionals of the NBP. In the general setting of BMPs, our priorities are the other way around. We are primarily concerned with the evolution of linear functionals of the BMP, which leads us to a linear semigroup evolution equation.

Define the mean semigroup of our BMP as

$$\psi_t[f](x) := \mathbb{E}_{\delta_x}[X_t[f]], \quad x \in E, \ f \in B^+(E), \ t \geq 0. \tag{8.14}$$

Although we are now in a much more general setting, the fact that $(\psi_t, t \geq 0)$ is a semigroup can be shown as in Lemma 3.4. Furthermore, set

$$\mathtt{F}[f](x) = \gamma(x)\mathcal{E}_x\left[\sum_{i=1}^{N} f(x_i) - f(x)\right] =: \gamma(x)(\mathtt{m}[f](x) - f(x)), \quad x \in E. \tag{8.15}$$

Note that both $(\psi_t, t \geq 0)$ and $\mathtt{F}$ are consistent with their definitions for the NBP. Thanks to the calculations we have already undertaken for the NBP, the next lemma requires no proof as it is identical to Lemma 3.6 in the NBP setting. In order to state it, we need to introduce the following assumption:

(G1) $\sup_{x\in E} \mathscr{E}_x[N] < \infty$.

Note that (G1) is equivalent to $\sup_{x\in E} \mathtt{m}[f](x) < \infty$ for $f \in B^+(E)$, which is covered by (H1) in the NBP setting.

Lemma 8.1 *Under (G1), the mean semigroup $(\psi_t, t \geq 0)$ satisfies*

$$\psi_t[f](x) = \mathtt{P}_t[f](x) + \int_0^t \mathtt{P}_s\left[\mathtt{F}\psi_{t-s}[f]\right](x)\mathrm{d}s, \quad t \geq 0,\ x \in E,\ f \in B^+(E). \tag{8.16}$$

As an operator from $B^+(E)$ to itself, $(\psi_t, t \geq 0)$ is uniquely determined by (8.16).

Just as in the NBP setting, the linear semigroup evolution (8.16) deserves a second representation for its unique solution, that is to say, the many-to-one representation. To this end, suppose that $\xi = (\xi_t, t \geq 0)$, with probabilities $\mathbf{P} = (\mathbf{P}_x, x \in E)$, is the Markov process corresponding to the semigroup $\mathtt{P}$. Let us introduce a new Markov process $\hat{\xi} = (\hat{\xi}_t, t \geq 0)$ which evolves as the process ξ, but at rate $\gamma(x)\mathtt{m}[1](x)$ the process is sent to a new position in E, such that for all Borel $A \subset E$, the new position is in A with probability $\mathtt{m}[\mathbf{1}_A](x)/\mathtt{m}[1](x)$. We will refer to the latter as *extra jumps*. Note the law of the extra jumps is well defined thanks to the assumption (G1), which we earlier remarked ensures that $\sup_{x\in E} \mathtt{m}[1](x) < \infty$. We denote the probabilities of $\hat{\xi}$ by $(\hat{\mathbf{P}}_x, x \in E)$. We can now state our many-to-one formula.

Lemma 8.2 (Many-to-One) *Write $\mathtt{B}(x) = \gamma(x)(\mathtt{m}[1](x) - 1)$, $x \in E$. For $f \in B^+(E)$ and $t \geq 0$, under (G1), we have*

$$\psi_t[f](x) = \hat{\mathbf{E}}_x\left[\exp\left(\int_0^t \mathtt{B}(\hat{\xi}_s)\mathrm{d}s\right) f(\hat{\xi}_t)\right]. \tag{8.17}$$

Proof First note that (8.16) is equivalent to

$$\begin{aligned}
\psi_t[f](x) &= \mathtt{P}_t[f] + \int_0^t \mathtt{P}_s\left[\gamma(\mathtt{m}[\psi_{t-s}[f]] - \psi_{t-s}[f])\right](x)\mathrm{d}s \\
&= \mathtt{P}_t[f] + \int_0^t \mathtt{P}_s\left[\gamma\mathtt{m}[1]\left(\frac{\mathtt{m}[\psi_{t-s}[f]]}{\mathtt{m}[1]} - \psi_{t-s}[f]\right)\right](x)\mathrm{d}s \\
&\qquad + \int_0^t \mathtt{P}_s[B\psi_{t-s}[f]]\mathrm{d}s, \qquad t \geq 0, x \in E.
\end{aligned}$$

At the same time, suppose we denote the right-hand side of (8.17) by $\hat{\psi}_t[f](x)$, $t \geq 0$. By conditioning this expectation on the first extra jump, we get, for $f \in B^+(E)$, $x \in E$ and $t \geq 0$,

$$
\begin{aligned}
&\hat{\psi}_t[f](x) \\
&= \mathbf{E}_x\left[\mathrm{e}^{-\int_0^t \gamma(\xi_s)\mathtt{m}[1](\xi_s)\mathrm{d}s}\mathrm{e}^{\int_0^t \mathtt{B}(\xi_s)\mathrm{d}s} f(\xi_t)\right] \\
&\quad + \mathbf{E}_x\left[\int_0^t \mathtt{m}[1](\xi_s)\gamma(\xi_s)\mathrm{e}^{-\int_0^s \gamma(\xi_u)\mathtt{m}[1](\xi_u)\mathrm{d}u}\mathrm{e}^{\int_0^s \mathtt{B}(\xi_u)\mathrm{d}u}\frac{\mathtt{m}[\hat{\psi}_{t-s}[f]](\xi_s)}{\mathtt{m}[1](\xi_s)}\mathrm{d}s\right].
\end{aligned}
\tag{8.18}
$$

Now appealing to Theorem 2.1, we can use the principle of transferring between the multiplicative and additive potential in integral evolution equations to deduce that (8.18) solves (8.16). Uniqueness of solutions to (8.16) now allows us to conclude the statement of the lemma. □

8.5 Asmussen–Hering Class, Criticality, and Ergodicity

In this and the subsequent chapters, we would like to establish results for as general as possible a setting within the class branching Markov processes. It goes without saying that some assumptions will be necessary. One of the main assumptions we will henceforth work with assimilates the properties of the expectation semigroup in Theorem 4.1 that we have proved in the NBP setting.

(G2) **There exist a constant $\lambda_* \in \mathbb{R}$, a function $0 \le \varphi \in B^+(E)$, and a finite measure $\tilde{\varphi}$ such that, for $f \in B^+(E)$,**

$$\psi_t[\varphi] = \mathrm{e}^{\lambda_* t}\varphi \textbf{ and } \tilde{\varphi}\,[\psi_t[f]] = \mathrm{e}^{\lambda_* t}\tilde{\varphi}[f].$$

Furthermore, let us define

$$\varDelta_t = \sup_{x\in E, f\in B_1^+(E)} |\varphi(x)^{-1}\mathrm{e}^{-\lambda_* t}\psi_t[f](x) - \tilde{\varphi}[f]|, \qquad t \ge 0.$$

We assume that

$$\sup_{t\ge 0}\varDelta_t < \infty \textbf{ and } \lim_{t\to\infty}\varDelta_t = 0. \tag{8.19}$$

We refer to branching Markov processes which satisfy (G2) as belonging to the *Asmussen–Hering* class of processes; see the comments section of this chapter for further remarks on why we use this name.

Note that Theorem 4.1 implies that the NBP lies in the Asmussen–Hering class by taking the eigenmeasure in (G2) to be $\tilde{\varphi}(x)\mathrm{d}x$, where $\tilde{\varphi}$ was the left *eigenfunction* given in Theorem 4.1. We note also that (G2) implies (G1) as (8.19) necessitates a finite first moment.

As in the setting of NBPs, as soon as we assume (G2), we can introduce a notion of criticality. Specifically, we say that our BMP is as follows:

Subcritical If $\lambda_* < 0$, in which case the average mass decays to zero over time at rate $|\lambda_*|$.
Critical If $\lambda_* = 0$, in which case the average mass remains essentially constant over time.
Supercritical If $\lambda_* > 0$, in which case we see the average mass in the system growing exponentially over time at rate λ_*.

Our assumption (G2) is relatively strong; however, considering the work that was needed to establish the analogous spectral property for the NBP, this seems a reasonable starting point in order to prove general results.

One advantage of (G2) is that it also provides a strong sense of ergodicity for the semigroup $(\psi_t, t \geq 0)$ as the following theorem below demonstrates, which will be crucial for future calculations. In order to state it, let us introduce a class of functions $\mathscr{C}$ on $B_1^+(E) \times E \times [0, 1] \times [0, \infty)$ such that F belongs to class $\mathscr{C}$ if

$$F[g](x, s) := \lim_{t\to\infty} F[g](x, s, t), \qquad g \in B_1^+(E), x \in E, s \in [0, 1],$$

exists,

$$\sup_{x\in E, s\in[0,1], g\in B_1^+(E)} |\varphi(x) F[g](x, s)| < \infty, \tag{8.20}$$

and

$$\lim_{t\to\infty} \sup_{x\in E, s\in[0,1], g\in B_1^+(E)} \varphi(x)\big|F[g](x, s) - F[g](x, s, t)\big| = 0. \tag{8.21}$$

Note that we have abused notation and used F for both the function that lies in $\mathscr{C}$ and its limit in t, and however, this should not cause any confusion later on.

Theorem 8.1 *Assume (G2) holds, $\lambda = 0$, and that $F \in \mathscr{C}$. Define*

$$\Xi_t = \sup_{x\in E, g\in B_1^+(E)} \left| \frac{1}{\varphi(x)} \int_0^1 \psi_{ut}[\varphi F[g](\cdot, u, t)](x)\mathrm{d}u - \int_0^1 \varphi\tilde{\varphi}\,[F[g](\cdot, u)]\,\mathrm{d}u \right|,$$

$t \geq 0$.

Then

$$\sup_{t\geq 0} \Xi_t < \infty \text{ and } \lim_{t\to\infty} \Xi_t = 0. \tag{8.22}$$

Proof We will show that

$$\lim_{t\to\infty}\sup_{x\in E, g\in B_1^+(E)}\left|\frac{1}{\varphi(x)}\psi_{ut}[\varphi F[g](\cdot,u,t)](x)-\varphi\tilde{\varphi}\,[F[g](\cdot,u)]\right|=0,$$

as, with this in hand, we can deduce that

$$\begin{aligned}
&\lim_{t\to\infty}\sup_{x\in E, g\in B_1^+(E)}\left|\int_0^1\frac{1}{\varphi(x)}\psi_{ut}[\varphi F[g](\cdot,u,t)](x)\mathrm{d}u-\int_0^1\varphi\tilde{\varphi}\,[F[g](\cdot,u)]\,\mathrm{d}u\right|\\
&\quad\le\int_0^1\lim_{t\to\infty}\sup_{x\in E, g\in B_1^+(E)}\left|\frac{1}{\varphi(x)}\psi_{ut}[\varphi F[g](\cdot,u,t)](x)-\varphi\tilde{\varphi}\,[F[g](\cdot,u)]\right|\mathrm{d}u\\
&\quad=0,
\end{aligned}$$

which follows from (8.21). First note that

$$\begin{aligned}
&\left|\frac{1}{\varphi(x)}\psi_{ut}[\varphi F[g](\cdot,u,t)](x)-\varphi\tilde{\varphi}\,[F[g](\cdot,u)]\right|\\
&\le\frac{1}{\varphi(x)}\psi_{ut}[|\varphi F[g](\cdot,u,t)-\varphi F[g](\cdot,u)|](x)\\
&\qquad+\left|\frac{1}{\varphi(x)}\psi_{ut}[\varphi F[g](\cdot,u)](x)-\varphi\tilde{\varphi}\,[F[g](\cdot,u)]\right|.
\end{aligned}$$

Due to assumption (8.21), for t sufficiently large, the first term on the right-hand side above can be controlled by $\varphi^{-1}(x)\psi_{ut}[\varepsilon](x)$. Combining this with the above inequality yields

$$\begin{aligned}
&\sup_{x\in E, g\in B_1^+(E)}\left|\frac{1}{\varphi(x)}\psi_{ut}[\varphi F[g](\cdot,u,t)](x)-\varphi\tilde{\varphi}\,[F[g](\cdot,u)]\right|\\
&\le\sup_{x\in E}\left|\varphi^{-1}(x)\psi_{ut}[\varepsilon](x)-\tilde{\varphi}[\varepsilon](x)\right|+\varepsilon\sup_{x\in E}\tilde{\varphi}[1](x)\\
&\quad+\sup_{x\in E, g\in B_1^+(E)}\left|\frac{1}{\varphi(x)}\psi_{ut}[\varphi F[g](\cdot,u)](x)-\varphi\tilde{\varphi}\,[F[g](\cdot,u)]\right|.
\end{aligned}\tag{8.23}$$

We note that (8.20) and the first (respectively, second) statement of (8.19) in (G2), together with dominated convergence and the fact that ε may be taken arbitrarily small, immediately imply that the first (respectively, second) statement in (8.22) holds. □

8.6 Re-oriented Non-linear Semigroup

As the reader will have observed, the recursion equations of the mild NTE (3.32) and PBE (5.6) enjoy a robustness that allows an interchange between a driving semigroup, e.g., $(\mathtt{U}_t, t \geq 0)$ or $(\mathtt{P}_t, t \geq 0)$, and an action operator in the integral term, e.g., $\mathtt{S}$ or $\mathtt{G}$.

In the general setting of (8.9), we are interested in exercising a similar change in the driving semigroup from $\mathtt{P}$ to ψ for technical reasons pertaining to future calculations around moments. Also, for technical reasons, with ψ as the driving semigroup, it is more convenient to instead look at writing down an evolution equation of

$$\mathtt{u}_t[f, g](x) = \mathbb{E}_{\delta_x}\left[1 - \mathrm{e}^{-X_t[f] - \int_0^t X_s[g]\mathrm{d}s}\right], \qquad t \geq 0,\ x \in E, \tag{8.24}$$

for f, g non-negative and measurable, despite the fact that it is not a semigroup. The main benefit of $(\mathtt{u}_t, t \geq 0)$ in place of $(\mathtt{v}_t, t \geq 0)$ is that the random variable in the expectation on the right-hand side of (8.24) is equal to zero on the extinction set. This removes the need to deal with adjusted semigroups as discussed below (8.7).

Another point of interest when it comes to (8.24) can be seen when we set $f = \theta$ and $g = 0$ and then take the limit as $\theta \to \infty$, in which case we identify $\mathtt{u}_t[\infty, 0] = \lim_{\theta \to \infty} \mathtt{u}_t[\theta, 0]$. Remembering that an empty product is defined to be unity, we see for this special case that

$$\mathtt{u}_t(x) := \mathtt{u}_t[\infty, 0](x) = \mathbb{P}_x(\zeta > t), \qquad t \geq 0, x \in E, \tag{8.25}$$

where

$$\zeta = \inf\{t > 0 : X_t[1] = 0\}$$

is the extinction time of the BMP.

As mentioned above, the change from $\mathtt{P}$ to ψ comes at the cost of changing the operator $\mathtt{G}$. Thus, for $f \in B_1^+(E)$, define

$$\mathtt{A}[f](x) = \gamma(x)\mathcal{E}_x\left[\prod_{i=1}^N (1 - f(x_i)) - 1 + \sum_{i=1}^N f(x_i)\right], \qquad x \in E. \tag{8.26}$$

Theorem 8.2 *Suppose (G1) holds. For all f, g non-negative, measurable functions on E, $x \in E$ and $t \geq 0$, the non-linear semigroup $\mathtt{u}_t[f, g](x)$ satisfies*

$$\mathtt{u}_t[f, g](x) = \psi_t[1 - \mathrm{e}^{-f}](x) - \int_0^t \psi_s\left[\mathtt{A}[\mathtt{u}_{t-s}[f, g]] - g(1 - \mathtt{u}_{t-s}[f, g])\right](x)\mathrm{d}s, \tag{8.27}$$

uniquely in $B_1^+(E)$.

Proof Again, the proof uses standard techniques for integral evolution equations, so we only sketch the proof. Instead of considering $\mathtt{u}_t[f, g]$, we will first work instead with

$$\mathtt{v}_t[f, g] = \mathbb{E}_{\delta_x}\left[\mathrm{e}^{-X_t[f]-\int_0^t X_s[g]\mathrm{d}s}\right], \qquad t \geq 0,\ x \in E,\ f, g \geq 0, \tag{8.28}$$

which will turn out to be more convenient for technical reasons.

By splitting the expectation in (8.28) on the first branching event and appealing to the Markov property, we get, for $f, g \geq 0$, $t \geq 0$, and $x \in E$,

$$\begin{aligned}\mathtt{v}_t[f, g](x) &= \mathbf{E}_x\left[\mathrm{e}^{-\int_0^t \gamma(\xi_s)\mathrm{d}s}\mathrm{e}^{-f(\xi_t)-\int_0^t g(\xi_s)\mathrm{d}s}\right]\\ &\qquad + \mathbf{E}_x\left[\int_0^t \gamma(\xi_s)\mathrm{e}^{-\int_0^s \gamma(\xi_u)+g(\xi_u)\mathrm{d}u}\mathtt{H}[\mathtt{v}_{t-s}[f, g]](\xi_s)\mathrm{d}s\right],\end{aligned}$$

where

$$\mathtt{H}[g](x) = \mathcal{E}_x\left[\prod_{i=1}^N g(x_i)\right], \qquad g \in B_1^+(E), x \in E.$$

Using Theorem 2.1, we can move the multiplicative potential with rate $\gamma + g$ to an additive potential in the above evolution equation to obtain

$$\mathtt{v}_t[f, g](x) = \hat{\mathtt{P}}_t[\mathrm{e}^{-f}](x) + \int_0^t \mathtt{P}_s\left[\mathtt{G}[\mathtt{v}_{t-s}[f, g]) - g\mathtt{v}_{t-s}[f, g]\right](x)\mathrm{d}s, \tag{8.29}$$

which also shows the existence of a non-negative solution to (8.29), bounded by unity.

For $f \in B_1^+(E)$, $x \in E$, define

$$\mathtt{D}[f](x) = \gamma(x)\mathcal{E}_x\left[\prod_{i=1}^N f(x_i) - \sum_{i=1}^N f(x_i)\right] = \gamma(x)\left(\mathtt{H}[f](x) - \mathtt{m}[f](x)\right),$$

and $(\tilde{\mathtt{v}}_t, t \geq 0)$ via

$$\begin{aligned}\tilde{\mathtt{v}}_t[f, g](x) &= \psi_t[\mathrm{e}^{-f}](x) + \int_0^t \psi_s\left[\mathtt{D}\big[\tilde{\mathtt{v}}_{t-s}[f, g]\big] - g\tilde{\mathtt{v}}_{t-s}[f, g]\right](x)\mathrm{d}s\\ &= \hat{\mathbf{E}}_x\left[\mathrm{e}^{\int_0^t \mathtt{B}(\hat{\xi}_s)\mathrm{d}s}\mathrm{e}^{-f(\hat{\xi}_t)}\right]\end{aligned}$$

$$+\hat{\mathbf{E}}_x\left[\int_0^t e^{\int_0^s \mathtt{B}(\hat{\xi}_u)\mathrm{d}u}\Big(\mathtt{D}\big[\tilde{\mathtt{v}}_{t-s}[f,g]\big](\hat{\xi}_s)-g(\hat{\xi}_s)\tilde{\mathtt{v}}_{t-s}[f,g](\hat{\xi}_s)\Big)\,\mathrm{d}s\right], \tag{8.30}$$

for $x\in E$, $t\geq 0$, and $f,g\geq 0$. Note that for the moment we do not claim a solution to (8.30) exists.

For convenience, we will define

$$\mathtt{K}_t[f,g](x)=\hat{\mathbf{E}}_x\left[\int_0^t e^{\int_0^s \mathtt{B}(\hat{\xi}_u)\mathrm{d}u}\Big(\mathtt{D}\big[\tilde{\mathtt{v}}_{t-s}[f,g]\big](\hat{\xi}_s)-g(\hat{\xi}_s)\tilde{\mathtt{v}}_{t-s}[f,g](\hat{\xi}_s)\Big)\,\mathrm{d}s\right],$$

so that $\tilde{\mathtt{v}}_t[f,g](x)=\psi_t[e^{-f}](x)+\mathtt{K}_t[f,g](x)$. By conditioning the right-hand side of (8.30) on the first jump of $\hat{\xi}$ (bearing in mind the dynamics of $\hat{\xi}$ given just before Lemma 8.2), we can use the Markov property (noting that $\mathtt{B}(x)-\gamma\mathtt{m}[1]=\gamma$) to obtain

$$\begin{aligned}
&\tilde{\mathtt{v}}_t[f,g](x)\\
&\quad=\mathbf{E}_x\left[e^{-\int_0^t\gamma(\xi_s)\mathrm{d}s}e^{-f(\xi_t)}\right]\\
&\qquad+\mathbf{E}_x\left[\int_0^t\gamma(\xi_\ell)\mathtt{m}[1](\xi_\ell)e^{-\int_0^\ell\gamma(\xi_s)\mathrm{d}s}\frac{\mathtt{m}[\psi_{t-\ell}[e^{-f}]](\xi_\ell)}{\mathtt{m}[1](\xi_\ell)}\mathrm{d}\ell\right]\\
&\qquad+\mathbf{E}_x\left[e^{-\int_0^t\gamma(\xi_u)\mathtt{m}[1](\xi_u)\mathrm{d}u}\int_0^t e^{\int_0^s \mathtt{B}(\xi_u)\mathrm{d}u}\right.\\
&\qquad\left.\times\big(\mathtt{D}\big[\tilde{\mathtt{v}}_{t-s}[f,g]\big](\xi_s)-g(\xi_s)[\tilde{\mathtt{v}}_{t-s}[f,g]\big)\,\mathrm{d}s\right]\\
&\qquad+\mathbf{E}_x\left[\int_0^t\gamma(\xi_\ell)\mathtt{m}[1](\xi_\ell)e^{-\int_0^\ell\gamma(\xi_u)\mathtt{m}[1](\xi_u)\mathrm{d}u}\right.\\
&\qquad\left(\int_0^\ell e^{\int_0^s \mathtt{B}(\xi_u)\mathrm{d}u}\big(\mathtt{D}\big[\tilde{\mathtt{v}}_{t-s}[f,g]\big](\xi_s)-g(\xi_s)\tilde{\mathtt{v}}_{t-s}[f,g](\xi_s)\big)\,\mathrm{d}s\right.\\
&\qquad\left.\left.+e^{\int_0^\ell \mathtt{B}(\xi_u)\mathrm{d}u}\frac{\mathtt{m}[\mathtt{K}_{t-\ell}[g]](\xi_\ell)}{\mathtt{m}[1](\xi_\ell)}\right)\mathrm{d}\ell\right].
\end{aligned}$$

Gathering terms and exchanging the order of integration in the double integral, this simplifies to

$$\begin{aligned}
&\tilde{\mathtt{v}}_t[f,g](x)\\
&=\mathbf{E}_x\left[e^{-\int_0^t\gamma(\xi_s)\mathrm{d}s}e^{-f(\xi_t)}\right]+\mathbf{E}_x\left[\int_0^t\gamma(\xi_\ell)e^{-\int_0^\ell\gamma(\xi_s)\mathrm{d}s}\mathtt{m}[\tilde{\mathtt{v}}_{t-\ell}[f,g](x)](\xi_\ell)\mathrm{d}\ell\right]\\
&\mathbf{E}_x\left[e^{-\int_0^t\gamma(\xi_u)\mathtt{m}[1](\xi_u)\mathrm{d}u}\int_0^t e^{\int_0^s \mathtt{B}(\xi_u)\mathrm{d}u}\big(\mathtt{D}\big[\tilde{\mathtt{v}}_{t-s}[f,g]\big](\xi_s)-g(\xi_s)[\tilde{\mathtt{v}}_{t-s}[f,g]\big)\,\mathrm{d}s\right]
\end{aligned}$$

$$+\mathbf{E}_x\Bigg[\int_0^t\int_0^t \mathbf{1}_{(s\le\ell)}\gamma(\xi_\ell)\mathtt{m}[1](\xi_\ell)\mathrm{e}^{-\int_0^\ell\gamma(\xi_u)\mathtt{m}[1](\xi_u)\mathrm{d}u}\mathrm{e}^{\int_0^s\mathtt{B}(\xi_u)\mathrm{d}u}$$

$$\Big(\mathtt{D}\big[\tilde{\mathtt{v}}_{t-s}[f,g]\big](\xi_s)-g(\xi_s)\tilde{\mathtt{v}}_{t-s}[f,g](\xi_s))\Big)\,\mathrm{d}\ell\,\mathrm{d}s\Bigg]$$

$$=\mathbf{E}_x\left[\mathrm{e}^{-\int_0^t\gamma(\xi_s)\mathrm{d}s}\mathrm{e}^{-f(\xi_t)}\right]+\mathbf{E}_x\left[\int_0^t\gamma(\xi_\ell)\mathrm{e}^{-\int_0^\ell\gamma(\xi_s)\mathrm{d}s}\mathtt{m}[\tilde{\mathtt{v}}_{t-\ell}[g](x)](\xi_\ell)\mathrm{d}\ell\right]$$

$$+\mathbf{E}_x\left[\int_0^t\mathrm{e}^{-\int_0^s\gamma(\xi_u)\mathrm{d}u}\Big(\mathtt{D}\big[\tilde{\mathtt{v}}_{t-s}[f,g]\big](\xi_s)-g(\xi_s)\tilde{\mathtt{v}}_{t-s}[f,g](\xi_s))\Big)\,\mathrm{d}s\right].$$

Finally, appealing to the change of multiplicative potential to additive potential in Theorem 2.1, we get

$$\tilde{\mathtt{v}}_t[f,g](x)=\hat{\mathtt{P}}_t[\mathrm{e}^{-f}](x)+\int_0^t\mathtt{P}_s\left[\mathtt{G}\big[\tilde{\mathtt{v}}_{t-s}[f,g]\big]-g\tilde{\mathtt{v}}_{t-s}[f,g]\right](x)\mathrm{d}s,$$

and hence $(\tilde{\mathtt{v}}_t, t\ge 0)$ is a solution to (8.29). Reversing these arguments also shows that any solution to (8.29) is also a solution to (8.30) and hence that both equations have domain equal to $B_1^+(E)$. A standard argument using $\gamma\in B^+(E)$, the assumption (G1), and Grönwall's lemma also tells us that (8.29) has a unique solution in $B_1^+(E)$.

To complete the theorem, note that

$$1-\psi_t[\mathrm{e}^{-f}](x)=\psi_t[1-\mathrm{e}^{-f}](x)+1-\psi_t[1](x)$$

and that

$$1-\psi_t[1](x)=\hat{\mathbf{E}}_x\left[\int_0^t\mathtt{B}(\hat{\xi}_s)\mathrm{e}^{\int_0^s\mathtt{B}(\hat{\xi}_u)\mathrm{d}u}\mathrm{d}s\right]=\int_0^t\psi_s[\mathtt{B}](x)\mathrm{d}s.$$

Hence, working from (8.30) and the definitions of $\mathtt{D}$ and $\mathtt{A}$, which are related via

$$\mathtt{D}[1-f](x)=\gamma(x)\mathcal{E}_x\left[\prod_i(1-f(x_i))-\sum_{i=1}^N(1-f(x_i))\right]=\mathtt{A}[f](x)+\mathtt{B}(x),$$

for $x\in E$, $f\ge 0$, we get

$$\begin{aligned}\mathtt{u}_t[f,g](x)&=1-\mathtt{v}_t[f,g](x)\\&=1-\psi_t[\mathrm{e}^{-f}](x)-\int_0^t\psi_s\left[\mathtt{D}\big[1-\mathtt{u}_{t-s}[f,g]\big]\right.\\&\qquad\left.-g(1-\mathtt{u}_{t-s}[f,g])\right](x)\mathrm{d}s\end{aligned}$$

$$= \psi_t[1 - \mathrm{e}^{-f}] - \int_0^t \psi_s \left[\mathtt{A}\left[\mathrm{u}_{t-s}[f, g]\right] - g(1 - \mathrm{u}_{t-s}[f, g]) \right](x)\mathrm{d}s,$$

as required. □ □

8.7 Discrete-Time Branching Markov Processes

In light of the discussion in Chap. 7, we will discuss on occasion discrete-time analogues of our branching Markov process.

We consider a discrete-time spatial branching process $(\mathcal{X}_n, n \geq 0)$. At each unit of time, independently for each individual in the process, a branching event occurs such that, if the parent is at $x \in E$, the new configuration of particles is given by the point process

$$\mathcal{Z} = \sum_{i=1}^{N} \delta_{z_i},$$

with probabilities and $(P_x, x \in E)$. As in the continuous-time setting, we allow for the possibility of absorption, that is, we may have $P_x(N = 0) > 0$.

Letting $N^{(n)}$ denote the number of individuals in the n-th generation, the generational branching process is formally defined via the collection of atomic measures

$$\mathcal{X}_n = \sum_{i=1}^{N^{(n)}} \delta_{x_i^{(n)}}, \qquad n \geq 0,$$

where $\{x_i^{(n)} : i = 1, \ldots, N^{(n)}\}$ denotes the collection of particles in the n-th generation. With a slight conflict of notation with the continuous-time setting, we again work with $(\mathbb{P}_\mu, \mu \in \mathcal{M}_c(E))$ as the law of the branching process defined above.

Remark 8.1 As a special setting, we can easily see a discrete-time Markov branching process embedded in each continuous-time Markov branching process (as defined earlier in this chapter), by sampling it at generational time. That is to say, by considering a continuous-time branching Markov process along each of its genealogical lines of descent at the moment that the n-th generation of offspring appears, we see a discrete-time Markov branching process. In particular, for each $x \in E$, we can link $\mathcal{Z}$ under P_x of the discrete-time setting to Z with probabilities $(\mathcal{P}_x, x \in E)$ of the discrete-time setting via

$$E_x\left[\mathrm{e}^{-\mathcal{Z}[f]}\right] = \mathbf{E}_x\left[\int_0^\infty \gamma(\xi_s)\mathrm{e}^{-\int_0^s \gamma(\xi_u)\mathrm{d}u} \mathcal{E}_{\xi_s}\left[\mathrm{e}^{-\mathsf{Z}[f]}\right]\mathrm{d}s\right], \qquad f \in B^+(E),$$

where we recall $(\xi, \mathbf{P})$ is the stochastic process whose semigroup is given by $(\mathtt{P}_t, t \geq 0)$. This was precisely the scenario we considered for the embedding of the NGP into the NBPs in (7.35).

Returning to the general discrete-time setting, the associated non-linear semigroup is given by

$$\mathtt{V}_n[g](x) := \mathbb{E}_{\delta_x}\left[\prod_{i=1}^{N^{(n)}} g(x_i^{(n)})\right], \qquad n \geq 1,\ x \in E,\ g \in B_1^+(E), \tag{8.31}$$

with $\mathtt{V}_0[g](x) = g(x)$. Analogously to (8.6) in the continuous setting, we have the Markov branching property,

$$\mathbb{E}\left[\prod_{i=1}^{N^{(n+m)}} g(x_i^{(n)}) \middle| \mathfrak{F}_n\right] = \prod_{i=1}^{N^{(n)}} \mathtt{V}_m[g](x_i^{(n)}), \tag{8.32}$$

where $\mathfrak{F}_n = \sigma(x_i^{(\ell)}, i = 1, \cdots, N^{(\ell)}, \ell \leq n)$. Moreover, by taking expectations across (8.32), this similarly justifies the identification of (8.31) as a semigroup via the property $\mathtt{V}_{n+m}[g] = \mathtt{V}_n[\mathtt{V}_m[g]]$, $n, m \geq 0$.

With a slight abuse of notation from the continuous-time setting, we can define the non-linear branching mechanism

$$\mathtt{G}[g](x) = E_x\left[\prod_{i=1}^{N} g(z_i)\right], \qquad x \in E,\ g \in B_1^+(E). \tag{8.33}$$

Noting that $\mathtt{G}[g] = \mathtt{V}_1[g]$, the branching Markov property also gives us the non-linear semigroup evolution equation

$$\mathtt{V}_n[g](x) = \mathtt{G}[\mathtt{V}_{n-1}[g]](x), \qquad n \geq 1. \tag{8.34}$$

Similarly, under the analogue of (G1), we have the associated linear semigroup, defined by

$$\Phi_n[g](x) := \mathbb{E}_{\delta_x}\left[\sum_{i=1}^{N^{(n)}} g(x_i^{(n)})\right], \qquad n \geq 1,\ x \in E,\ g \in B^+(E), \tag{8.35}$$

with $\Phi_0[g](x) = g(x)$, which satisfies the semigroup property $\Phi_{n+m}[g] = \Phi_n[\Phi_m[g]]$; see, for example, the calculation in the setting of the NGP (7.6). Moreover, again abusing notation from the continuous-time setting, the linear branching mechanism is given by

$$\mathtt{F}[g](x) = \mathrm{E}_x\left[\sum_{i=1}^{N} g(z_i)\right], \qquad x \in E,\ g \in B^+(E), \tag{8.36}$$

and noting that $\mathtt{F}[g] = \Phi_1[g]$, we also have the linear evolution equation

$$\Phi_n[g](x) = \mathtt{F}[\Phi_{n-1}[g]](x). \tag{8.37}$$

It is interesting to note that the two evolution equations (8.34) and (8.37) appear to bear little similarity to their analogues in continuous time. In particular, one may wish to ask how to assimilate Theorem 8.2 in the discrete-time setting. The following result may be considered as precisely this assimilation.

Lemma 8.3 *For $n \geq 1$, $x \in E$, $g \in B_1^+(E)$,*

$$\mathtt{V}_n[g](x) = \sum_{\ell=0}^{n-1} \Phi_\ell[(\mathtt{G}-\mathtt{F})[\mathtt{V}_{n-\ell-1}[g]]](x) + \Phi_n[g](x). \tag{8.38}$$

Proof Recalling that $\Phi_1[g](x) = \mathtt{F}[g](x)$, we can add and subtract terms in (8.34) to give us

$$\mathtt{V}_n[g](x) = (\mathtt{G}-\mathtt{F})[\mathtt{V}_{n-1}[g]](x) + \Phi_1[\mathtt{V}_{n-1}[g]](x), \qquad n \geq 1, x \in E, g \in B_1^+(E).$$

Using the same trick with the second term on the right-hand side of the above representation, we have

$$\begin{aligned}
\mathtt{V}_n[g](x) &= (\mathtt{G}-\mathtt{F})[\mathtt{V}_{n-1}[g]](x) + \Phi_1[(\mathtt{G}-\mathtt{F})[\mathtt{V}_{n-2}[g]]](x) \\
&\quad + \Phi_1[\Phi_1[\mathtt{V}_{n-2}[g]]](x) \\
&= (\mathtt{G}-\mathtt{F})[\mathtt{V}_{n-1}[g]](x) + \Phi_1[(\mathtt{G}-\mathtt{F})[\mathtt{V}_{n-2}[g]]](x) + \Phi_2[\mathtt{V}_{n-2}[g]](x),
\end{aligned}$$

where, from (8.37), a simple recursion also tells us that

$$\Phi_n[g](x) = \underbrace{\Phi_1[\Phi_1[\cdots\Phi_1}_{n \text{ times}}[g]]](x).$$

Continuing this recursion, we obtain (8.38). □

8.8 Comments

Branching Markov processes enjoy enormous exposure in the probabilistic literature. So much so that it would be difficult to give a concise overview of the entirety of the literature here. Instead we give some key historical references.

Some of the earliest works on spatial branching processes emerge from the Soviet Union and Japan. Most notably are the works of Sevast'yanov [120–122] and Ikeda et al. [75–77]. Enthusiasm for the, then, fledgling theory of branching processes resulted in a number of classical texts emerging through the late 1960s to mid-1980s, in which there are some spatial treatments. See, for example, Harris [70], Athreya and Ney [6], and Asmussen and Hering [5]. The latter, in particular, summarises a general perspective that we have adopted from this chapter onwards.

The many-to-one formula in Lemma 8.2 is a very general concept that runs throughout classical and modern branching processes, and it would be difficult to give a complete and just historical summary of its evolution through the literature. An early example of where a many-to-one formula is used as a key component to analytical computations is found in the work of Doney [41].

Asmussen and Hering [3–5] observed that the assumption (G2) was a natural way to provide analytical control over the general class of branching Markov processes, from which numerous results can be derived. Some of them are included in the later chapters.

For a lot of branching Markov process literature, it is common to describe the linear and non-linear semigroup in terms of differential equations, such as in the examples of Sect. 8.3. The use of mild equations, analogous to the Pál–Bell equation for neutron branching processes, is less often used for BMPs and is more commonly seen in the superprocesses literature; cf. [49]. The development of their representation in terms of the linear semigroup operator $(\mathtt{T}_t, t \geq 0)$, cf. Theorem 8.2, was recently given in [67]. We will show in the next chapter that, combined with assumption (G2), this representation of the non-linear semigroup offers tractability on asymptotic moment computations.

Chapter 9
Moments

We saw in Sect. 5.3 the claim that the asymptotic moments can be derived by differentiating the non-linear semigroup equation (8.7). This is also the case in the general setting. Indeed, recalling $(\mathtt{v}_t, t \geq 0)$ and $(\mathtt{u}_t, t \geq 0)$ from (8.5) and (8.24), respectively, and noting that $\mathtt{v}_t[\cdot] = \mathtt{u}_t[\cdot, 0]$, $t \geq 0$, it follows that for $f \in B^+(E)$, $x \in E$, $t \geq 0$, and $k \geq 1$, we have

$$\psi_t^{(k)}[f](x) := \mathbb{E}_{\delta_x}\left[X_t[f]^k\right] = (-1)^{k+1}\frac{\mathrm{d}^k}{\mathrm{d}\theta^k}\mathtt{u}_t[\theta f, 0](x)\bigg|_{\theta=0}. \tag{9.1}$$

Working within the Asmussen–Hering class of BMPs, the non-linear equation (8.27) for $(\mathtt{u}_t, t \geq 0)$ offers the advantage that it is written in terms of $(\psi_t, t \geq 0)$. This allows the asymptotic control of $(\psi_t, t \geq 0)$ offered by (G2) (e.g., in Theorem 8.1) to be exploited to yield asymptotic control of the derivatives in θ of $(\mathtt{u}_t[\theta f, 0], t \geq 0)$ and hence $(\psi_t^{(k)}[f], t \geq 0)$ for each $k \geq 2$. Following this logic, we give precise asymptotics of the k-th moments of our class of BMPs in the three criticality regimes. The methods we will use are extremely robust, and thus we may employ the same techniques to the derivatives in θ of $(\mathtt{u}_t[0, \theta f], t \geq 0)$. As such, we are able to analyse the moments of the running occupation functional

$$I_t^{(k)}[g](x) := \mathbb{E}_{\delta_{(x)}}\left[\left(\int_0^t X_s[g]\mathrm{d}s\right)^k\right], \tag{9.2}$$

for $x \in E$, $g \in B^+(E)$, $k \geq 1$, $t \geq 0$.

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8_9

9.1 Evolution Equations for the k-th Moments

Before we can sensibly talk about the k-th moments of our BMP, we need to be sure that they are finite. For this, an assumption is certainly needed on the basic data of the BMP, and somewhat obviously, we need at least to ensure there are k moments of the offspring distribution. We thus fix $k \geq 2$ and introduce the following.

(G3) We have that $\sup_{x\in E} \mathcal{E}_x[\mathsf{Z}[\mathbf{1}]^k] < \infty$.

Before we finally turn our attention to the evolution equation generated by the k-th moment functional $\psi_t^{(k)}$, $t \geq 0$, we will need to remind ourselves of one more classical result, the Leibniz rule for differentiation.

Lemma 9.1 (Leibniz Rule) *Suppose $g_1, \ldots, g_m$ are k-times continuously differentiable functions on $\mathbb{R}$, for $k \geq 1$. Then*

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k}\left(\prod_{i=1}^m g_i(x)\right) = \sum_{k_1+\cdots+k_m=k} \binom{k}{k_1, \ldots, k_m} \prod_{\ell=1}^m g_\ell^{(k_\ell)}(x),$$

such that the sum is taken over all non-negative integers $k_1, \cdots, k_m$ with $\sum_{i=1}^m k_i = k$.

Finally, we are ready to describe an evolution equation for the k-th moment of our MBP, which will lay the foundations for an inductive argument, giving the asymptotic behaviour of moments in each of the three criticality regimes.

Proposition 9.1 *Fix $k \geq 2$. Under the assumptions (G2) and (G2), with the additional assumption that*

$$\sup_{x\in E, s\leq t} \psi_s^{(\ell)}[f](x) < \infty, \qquad \ell \leq k-1,\ f \in B^+(E), t \geq 0, \tag{9.3}$$

it holds that

$$\psi_t^{(k)}[f](x) = \psi_t[f^k](x) + \int_0^t \psi_s\left[\gamma \eta_{t-s}^{(k-1)}[f]\right](x)\,\mathrm{d}s, \qquad t \geq 0, \tag{9.4}$$

where

$$\eta_{t-s}^{(k-1)}[f](x) = \mathcal{E}_x\left[\sum_{[k_1,\ldots,k_N]_k^2} \binom{k}{k_1, \ldots, k_N} \prod_{j=1}^N \psi_{t-s}^{(k_j)}[f](x_j)\right],$$

and $[k_1, \ldots, k_N]_k^2$ is the set of all non-negative N-tuples $(k_1, \ldots, k_N)$ such that $\sum_{i=1}^N k_i = k$ and at least two of the k_i are strictly positive.

Proof Recall from (8.27) that

$$\mathtt{u}_t[\theta f, 0](x) = \psi_t[1 - \mathrm{e}^{-\theta f}](x) - \int_0^t \psi_s\left[\mathbb{A}[\mathtt{u}_{t-s}[\theta f, 0]]\right](x)\mathrm{d}s, \quad t \geq 0. \tag{9.5}$$

It is clear that differentiating the first term k times and setting $\theta = 0$ on the right-hand side of (9.5) yield

$$\left.\frac{\partial^k}{\partial \theta^k}\psi_t[1 - \mathrm{e}^{-\theta f}](x)\right|_{\theta=0} = (-1)^{k+1}\psi_t[f^k](x). \tag{9.6}$$

Thus it remains to differentiate the second term on the right-hand side of (9.5) k times. To this end, without concern for passing derivatives through expectations, using the Leibniz rule in Lemma 9.1, we have

$$\begin{aligned}
&- \left.\frac{\partial^k}{\partial \theta^k}\mathbb{A}[\mathtt{u}_t[\theta f, 0]](x)\right|_{\theta=0} \\
&= \left.\frac{\partial^k}{\partial \theta^k}\gamma(x)\mathcal{E}_x\left[1 - \prod_{i=1}^N \mathbb{E}_{\delta_{x_i}}[\mathrm{e}^{-\theta X_t[f]}] - \sum_{i=1}^N \mathbb{E}_{\delta_{x_i}}[1 - \mathrm{e}^{-\theta X_t[f]}]\right]\right|_{\theta=0} \\
&= -\gamma(x)\mathcal{E}_x\Bigg[\sum_{k_1+\cdots+k_N=k}\binom{k}{k_1, \ldots, k_N}\prod_{j=1}^N (-1)^{k_j}\psi_t^{(k_j)}[f](x_j) \\
&\qquad\qquad\qquad + (-1)^{k+1}\sum_{i=1}^N \psi_t^{(k)}[f](x_i)\Bigg] \\
&= \gamma(x)\mathcal{E}_x\Bigg[(-1)^{k+1}\sum_{k_1+\cdots+k_N=k}\binom{k}{k_1, \ldots, k_N}\prod_{j=1}^N \psi_t^{(k_j)}[f](x_j) \\
&\qquad\qquad\qquad + (-1)^{k}\sum_{i=1}^N \psi_t^{(k)}[f](x_i)\Bigg]
\end{aligned} \tag{9.7}$$

such that the sum is taken over all non-negative integers $k_1, \cdots, k_N$ with $\sum_{i=1}^N k_i = k$.

Next let us look in more detail at the sum/product term on the righthand (9.7). Consider the terms where only one of the k_i in the sum is positive, in which case $k_i = k$ and

$$\binom{k}{k_1, \ldots, k_N} = 1.$$

There are N ways this can happen in the sum–product term and hence

$$\sum_{k_1+\cdots+k_N=k} \binom{k}{k_1,\ldots,k_N} \prod_{j=1}^N \psi_t^{(k_j)}[f](x_j)$$
$$= \sum_{i=1}^N \psi^{(k)}[f](x_i) + \sum_{[k_1,\ldots,k_N]_k^2} \binom{k}{k_1,\ldots,k_N} \prod_{j=1}^N \psi_t^{(k_j)}[f](x_j),$$

where $[k_1,\ldots,k_N]_k^2$ is the set of all non-negative N-tuples $(k_1,\ldots,k_N)$ such that $\sum_{i=1}^N k_i = k$ and at least two of the k_i are strictly positive. Substituting this back into (9.7) yields

$$-\left.\frac{\partial^k}{\partial\theta^k}\mathtt{A}[\mathtt{u}_t[\mathrm{e}^{-\theta f},0]]\right|_{\theta=0}$$
$$= (-1)^{k+1}\gamma(x)\mathcal{E}_x\left[\sum_{[k_1,\ldots,k_N]_k^2} \binom{k}{k_1,\ldots,k_N} \prod_{j=1}^N \psi_t^{(k_j)}[f](x_j)\right].$$

Now let us return to the justification of passing the derivatives through the expectation in the above calculation. We first note that derivatives are limits and so an "epsilon-delta" argument will ultimately require dominated convergence. This is where the assumption (G2) and (9.3) come in. On the right-hand side of (9.7), for a given $f \in B^+(E)$, each of the $\psi_t^{(k_j)}[f](x_j)$ in the sum–product term is uniformly bounded in x_j by assumption (9.3) and the collection $[k_1,\ldots,k_N]_k^2$ means that $0 \le k_j \le k-1$ for each $j = 1,\cdots,N$. Moreover, there can be at most k items in the sum/product. Noting that

$$\sum_{k_1+\cdots+k_N=k} \binom{k}{k_1,\ldots,k_N} = N^k, \tag{9.8}$$

the assumption (G2) allows us to use a domination argument with the k-th order moment to pull the k derivatives through the integral in t as required. □

9.2 Moment Evolution at Criticality

The main result in this section tells us that, under our main assumptions, the k-th moment at time $t \ge 0$ grows like t^{k-1} as $t \to \infty$ at criticality.

Theorem 9.1 (Critical, $\lambda_* = 0$) *Suppose that (G2) holds along with (G2) for some $k \ge 1$. Define*

$$\Delta_t^{(\ell)} = \sup_{x\in E, f\in B_1^+(E)} \left| t^{-(\ell-1)}\varphi(x)^{-1}\psi_t^{(\ell)}[f](x) - 2^{-(\ell-1)}\ell!\,\tilde\varphi[f]^\ell\tilde\varphi\,[\gamma\mathscr{V}[\varphi]]^{\ell-1}\right|,$$

where

$$\mathscr{V}[\varphi](x) = \mathscr{E}_x\left(\mathcal{Z}[\varphi]^2 - \mathcal{Z}[\varphi^2]\right). \tag{9.9}$$

Then, for all $\ell \le k$ *and* $c > 0$*,*

$$\sup_{t\ge c}\Delta_t^{(\ell)} < \infty \text{ and } \lim_{t\to\infty}\Delta_t^{(\ell)} = 0. \tag{9.10}$$

Proof We will prove Theorem 9.1 by induction, starting with the case $k = 1$. In this case, assumption (G2) reads

$$\sup_{t\ge 0}\Delta_t < \infty \text{ and } \lim_{t\to\infty}\Delta_t = 0,$$

which gives us (9.18).

We now assume that the theorem holds true in the branching Markov process setting for some $k \ge 1$ and proceed to show that (9.18) holds for all $\ell \le k + 1$.

To this end, first note that the induction hypothesis implies that (9.3) holds. Hence Proposition 9.1 tells us that

$$\begin{aligned}
&\varphi(x)^{-1}t^{-k}\psi_t^{(k+1)}[f](x)\\
&= \varphi(x)^{-1}t^{-k}\psi_t[f^{(k+1)}](x)\\
&\quad + \varphi(x)^{-1}t^{-k}\int_0^t \psi_s\left[\mathscr{E}_\cdot\left[\sum_{[k_1,\dots,k_N]^2_{k+1}}\binom{k+1}{k_1,\dots,k_N}\prod_{j=1}^N \psi_{t-s}^{(k_j)}[f](x_j)\right]\right](x)\mathrm{d}s\\
&= \varphi(x)^{-1}t^{-k}\psi_t[f^{(k+1)}](x) + \varphi(x)^{-1}t^{-(k-1)}\\
&\quad \times \int_0^1 \psi_{ut}\left[\mathscr{E}_\cdot\left[\sum_{[k_1,\dots,k_N]^2_{k+1}}\binom{k+1}{k_1,\dots,k_N}\prod_{j=1}^N \psi_{t(1-u)}^{(k_j)}[f](x_j)\right]\right](x)\mathrm{d}u,
\end{aligned} \tag{9.11}$$

where we have used the change of variables $s = ut$ in the final equality.

We now make some observations that will simplify the expression on the right-hand side of (9.11) as $t \to \infty$. First note that due to (8.19), the first term on the right-hand side of (9.11) will vanish as $t \to \infty$. Next, note that, if more than two of the k_i in the sum is strictly positive, then renormalising by t^{k-1}

will cause the associated summand to go to zero as well. For example, suppose without loss of generality that k_1 and k_2 are both strictly positive, and we can write $t^{k-1} = t^{(k+1)-2} = t^{k_1-1}t^{k_2-1}t^{k_3}\dots t^{k_N}$. Now the induction hypothesis tells us that the correct normalisation of each of the terms in the product is t^{k_j-1}, which means that the item $\psi^{(k_j)}_{t(1-u)}$ for a third $k_j > 0$ will be "over normalised to zero" in the limit.

To make this heuristic rigorous, we can employ Theorem 8.1. To this end, let us set

$$F[f](x,u,t) := \frac{1}{\varphi(x)t^{k-1}}\mathcal{E}_x\left[\sum_{[k_1,\dots,k_N]^3_{k+1}}\binom{k+1}{k_1,\dots,k_N}\prod_{j=1}^N \psi^{(k_j)}_{t(1-u)}[f](x_j)\right], \tag{9.12}$$

where $[k_1,\dots,k_N]^3_{k+1}$ is the subset of $[k_1,\dots,k_N]^2_{k+1}$, for which at least three of the k_i are strictly positive (which can be an empty set). We will now show that the conditions (8.20) and (8.21) are satisfied.

First note that there are no more than $k+1$ of the k_i that are strictly greater than 1 in the product in (9.12). This follows from the fact that it is not possible to partition the set $\{1,\dots,k+1\}$ into more than $k+1$ non-empty blocks. Next note that

$$\begin{aligned}&\frac{1}{t^{k-1}}\prod_{\substack{j=1\\ j:k_j>0}}^N \psi^{(k_j)}_{t(1-u)}[f](x_j)\\ &\quad= \frac{(t(1-u))^{k+1-\#\{j:k_j>0\}}}{t^{k-1}}\prod_{\substack{j=1\\ j:k_j>0}}^N \varphi(x_j)\cdot\frac{1}{\varphi(x_j)}\frac{\psi^{(k_j)}_{t(1-u)}[f](x_j)}{(t(1-u))^{k_j-1}}.\end{aligned}$$

The product term on the right-hand side is uniformly bounded in x_j and $t(1-u)$ bounded away from 0 due to boundedness of φ and the fact that (9.18) is assumed to hold for all $\ell \le k$ by induction. For $t(1-u)$ in the neighbourhood of zero, it is not necessary to multiply and divide by $(t(1-u))^{k+1-\#\{j:k_j>0\}}$, as the product can be controlled by the factor $t^{-(k-1)}$. Moreover, if $\#\{j : k_j > 0\} \le 2$, the set $[k_1,\dots,k_N]^3_{k+1}$ is empty; otherwise, the term $(t(1-u))^{k+1-\#\{j:k_j>0\}}/t^{k-1}$ is finite for all $t \ge 1$, say. From (9.8) and (G2), we also observe that

$$\sup_{x\in E}\mathcal{E}_x\left[\sum_{[k_1,\dots,k_N]^3_{k+1}}\binom{k+1}{k_1,\dots,k_N}\right] \le \sup_{x\in E}\mathcal{E}_x\left[\mathsf{Z}[\mathbf{1}]^{k+1}\right] < \infty.$$

Taking these facts into account, it is now straightforward to see that the earlier given heuristic can be made rigorous and

$$\sup_{\substack{x\in E,\, f\in B_1^+(E)\\ u\in[0,1]}} \varphi(x)F[f](x,u,t) < \infty \text{ and } \lim_{t\to\infty} \sup_{\substack{x\in E,\, f\in B_1^+(E)\\ u\in[0,1]}} \varphi(x)F[f](x,u,t) = 0 \tag{9.13}$$

hold. In particular, we can use dominated convergence to pass the limit in t through the expectation in (9.12) to achieve the second statement in (9.13).

As F belongs to the class of functions $\mathscr{C}$, defined just before Theorem 8.1, the aforesaid theorem tells us that

$$\lim_{t\to\infty} \sup_{x\in E,\, f\in B_1^+(E)} \left| \frac{1}{\varphi(x)} \int_0^1 \psi_{ut}[\varphi F(\cdot,u,t)](x)\mathrm{d}u \right| = 0. \tag{9.14}$$

Returning to (9.11), since the sum there requires that at least two of the k_i are positive, this means that the only surviving terms in the limit are those that are combinations of two strictly positive terms k_i and k_j such that $i \neq j$ and $k_i + k_j = k + 1$. This can be thought of as choosing $i, j \in \{1, \ldots N\}$ with $i \neq j$, choosing $k_i \in \{1, \ldots, k\}$, and then setting $k_j = k + 1 - k_i$. One should take care of however avoiding double counting each pair (k_i, k_j). Thus, we have

$$\begin{aligned}
&\frac{1}{t^k\varphi(x)}\psi_t^{(k+1)}[f](x)\\
&= \frac{1}{\varphi(x)} \int_0^1 \psi_{ut}\Bigg[\frac{\gamma(\cdot)}{2t^{k-1}}\mathcal{E}_\cdot\Bigg[\sum_{i=1}^N \sum_{\substack{j=1\\ j\neq i}}^N \sum_{k_i=1}^k \binom{k+1}{k_i}\\
&\qquad\qquad \times \psi_{t(1-u)}^{(k_i)}[f](x_i)\psi_{t(1-u)}^{(k+1-k_i)}[f](x_j)\Bigg]\Bigg](x)\mathrm{d}u,
\end{aligned} \tag{9.15}$$

where the factor of $1/2$ appears to compensate for the aforementioned double counting.

In order to show that the right-hand side above delivers the required finiteness and limit (9.18), we again turn to Theorem 8.1. For $x \in E$, $t \geq 0$, and $0 \leq u \leq 1$, in anticipation of using this theorem, we now redefine

$$\begin{aligned}
&F[f](x,u,t)\\
&:= \frac{\gamma(x)}{2\varphi(x)t^{k-1}}\mathcal{E}_x\Bigg[\sum_{i=1}^N \sum_{\substack{j=1\\ j\neq i}}^N \sum_{k_i=1}^k \binom{k+1}{k_i}\psi_{t(1-u)}^{(k_i)}[f](x_i)\psi_{t(1-u)}^{(k+1-k_i)}[f](x_j)\Bigg].
\end{aligned}$$

After some rearrangement, we have

$$F[f](x,u,t)$$
$$= \frac{\gamma(x)(1-u)^{k-1}}{2\varphi(x)}\mathscr{E}_x\Bigg[\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{k_i=1}^{k}\binom{k+1}{k_i}$$
$$\times\, \varphi(x_i)\varphi(x_j)\frac{\psi^{(k_i)}_{t(1-u)}[f](x_i)}{\varphi(x_i)(t(1-u))^{k_i-1}}\frac{\psi^{(k+1-k_i)}_{t(1-u)}[f](x_j)}{\varphi(x_j)(t(1-u))^{k-k_i}}\Bigg]. \tag{9.16}$$

Using similar arguments to those given previously in the proof of (9.14) may, again, combine the induction hypothesis, simple combinatorics, and dominated convergence to pass the limit as $t\to\infty$ through the expectation and show that

$$F[f](x,u) := \lim_{t\to\infty} F[f](x,u,t)$$
$$= (k+1)!(\tilde{\varphi}[\gamma\mathscr{V}[\varphi]]/2)^{k-1}\tilde{\varphi}[f]^{k+1}k\frac{(1-u)^{k-1}}{2\varphi(x)}\gamma(x)\mathscr{V}[\varphi](x), \tag{9.17}$$

for which one uses that

$$(k+1)!(\tilde{\varphi}[\gamma\mathscr{V}[\varphi]]/2)^{k-1}\tilde{\varphi}[f]^{k+1}k\gamma(x)\mathscr{V}[\varphi](x)$$
$$= \mathscr{E}_x\Bigg[\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{k_i=1}^{k}\binom{k+1}{k_i}\varphi(x_i)\varphi(x_j)$$
$$\times\, \frac{k_i!\,\tilde{\varphi}[f]^{k_i}\tilde{\varphi}[\gamma\mathscr{V}[\varphi]]^{k_i-1}}{2^{(k_i-1)}}\frac{(k+1-k_i)!\,\tilde{\varphi}[f]^{k+1-k_i}\tilde{\varphi}[\gamma\mathscr{V}[\varphi]]^{k-k_i}}{2^{(k-k_i)}}\Bigg].$$

Note that, thanks to the assumption (G2), the expression for $F(s,x)$ clearly satisfies (8.20).

Subtracting the right-hand side of (9.17) from the right-hand side of (9.16), again appealing to the induction hypotheses, specifically the second statement in (9.18), it is not difficult to show that, for each $\varepsilon\in(0,1)$,

$$\lim_{t\to\infty}\sup_{x\in E,\, f\in B_1^+(E),\, u\in[0,\varepsilon)}|\varphi(x)F(x,u,t)-\varphi(x)F(x,u)| = 0.$$

On the other hand, again by subtracting the right-hand side of (9.17) from the right-hand side of (9.16), the first statement in the induction hypothesis (9.18) also implies that there exists a constant $C_k>0$ (which depends on k but not ε) such that

$$\lim_{t\to\infty} \sup_{x\in E, f\in B_1^+(E), u\in[\varepsilon,1]} |\varphi(x)F(x,u,t) - \varphi(x)F(x,u)| \leq C_k(1-\varepsilon)^{k-1}.$$

Since we may take ε arbitrarily close to 1, we conclude that (8.21) holds.

In conclusion, since the conditions of Theorem 8.1 are now met, we get the two statements of (9.18) as a consequence. □

Thanks to (G2), it is straightforward to see that $X_t[\varphi]/\varphi(x)$ is a martingale for each $x \in E$ and can thus be used to define the following change of measure:

$$\left.\frac{\mathrm{d}\mathbb{P}_{\delta_x}}{\mathrm{d}\mathbb{Q}_{\delta_x}}\right|_{\mathfrak{F}_t} := \frac{X_t[\varphi]}{\varphi(x)}.$$

Using this and the fact that $\tilde{\varphi}[\varphi] = 1$, we have the following corollary.

Corollary 9.1 *Let $x \in E$ and suppose the conditions of Theorem 9.1 hold. Define*

$$\Delta_t^{(\ell)} = \sup_{x\in E} \left| t^{-\ell}\mathbb{Q}_{\delta_x}[X_t[\varphi]^\ell] - 2^{-\ell}(\ell+1)!\,\tilde{\varphi}\,[\gamma\mathscr{V}[\varphi]]^\ell \right|.$$

Then, for all $\ell \leq k$ and $c > 0$,

$$\sup_{t\geq c} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t\to\infty} \Delta_t^{(\ell)} = 0. \tag{9.18}$$

In particular, if (G2) holds for all $k \geq 1$, under $\mathbb{Q}_{\delta_x}$, $X_t[\varphi]/t$ converges in law to a Gamma random variable with mean $\tilde{\varphi}\,[\gamma\mathscr{V}[\varphi]]^\ell\,/2$.

9.3 Moment Evolution at Non-criticality

Next we turn to the supercritical setting. Whilst in the critical setting, the moments exhibited polynomial growth, the supercritical setting exhibits the phenomenon that the k-th moment grows like the k-th power of the first moment.

Theorem 9.2 (Supercritical, $\lambda_* > 0$) *Suppose that (G2) holds and (G2) holds for some $k \geq 1$. Redefine*

$$\Delta_t^{(\ell)} = \sup_{x\in E, f\in B_1^+(E)} \left| \mathrm{e}^{-\ell\lambda_* t}\varphi(x)^{-1}\psi_t^{(\ell)}[f](x) - \ell!\tilde{\varphi}[f]^\ell L_\ell(x) \right|,$$

where $L_1(x) = 1$ and we define iteratively, for $k \geq 2$,

$$L_k(x) = \int_0^\infty \mathrm{e}^{-\lambda_* ks}\varphi(x)^{-1}\psi_s\Bigg[\gamma\mathscr{E}_\cdot\Bigg[\sum_{[k_1,\ldots,k_N]_k^2}\prod_{\substack{j=1\\ j:k_j>0}}^N \varphi(x_j)L_{k_j}(x_j)\Bigg]\Bigg](x)\mathrm{d}s,$$

where $[k_1,\ldots,k_N]_k^2$ is the set of all non-negative N-tuples $(k_1,\ldots,k_N)$ such that $\sum_{i=1}^N k_i = k$ and at least two of the k_i are strictly positive. Then, for all $\ell \le k$, (9.18) holds.

Proof Suppose for induction that the result is true for all ℓ-th integer moments with $1 \le \ell \le k-1$. From the evolution equation (9.4), noting that $\sum_{j=1}^N k_j = k$, when the limit exists, we have

$$\begin{aligned}
&\lim_{t\to\infty}\mathrm{e}^{-\lambda_* kt}\int_0^t \varphi(x)^{-1}\psi_s\left[\gamma\mathscr{E}_\cdot\left[\sum_{[k_1,\ldots,k_N]_k^2}\binom{k}{k_1,\ldots,k_N}\prod_{j=1}^N \psi_{t-s}^{(k_j)}[f](x_j)\right]\right](x)\mathrm{d}s\\
&= \lim_{t\to\infty} t\int_0^1 \mathrm{e}^{-\lambda_*(k-1)ut}\mathrm{e}^{-\lambda_* ut}\varphi(x)^{-1}\psi_{ut}\Big[H[f](x,u,t)\Big](x)\mathrm{d}u, \qquad (9.19)
\end{aligned}$$

where

$$H[f](x,u,t) := \gamma(x)\mathscr{E}_x\left[\sum_{[k_1,\ldots,k_N]_k^2}\binom{k}{k_1,\ldots,k_N}\prod_{j=1}^N \mathrm{e}^{-\lambda_* k_j t(1-u)}\psi_{t(1-u)}^{(k_j)}[f](x_j)\right].$$

It is easy to see that, pointwise in $x \in E$ and $u \in [0,1)$, using the induction hypothesis and (G2),

$$\begin{aligned}
H[f](x) &:= \lim_{t\to\infty} H[f](x,u,t)\\
&= \gamma(x)\mathscr{E}_x\left[\sum_{[k_1,\ldots,k_N]_k^2}\binom{k}{k_1,\ldots,k_N}\prod_{\substack{j=1\\ j:k_j>0}}^N k_j!\tilde{\varphi}[f]^{k_j}\varphi(x_j)L_{k_j}(x_j)\right]\\
&= k!\tilde{\varphi}[f]^k\gamma(x)\mathscr{E}_x\left[\sum_{[k_1,\ldots,k_N]_k^2}\prod_{\substack{j=1\\ j:k_j>0}}^N \varphi(x_j)L_{k_j}(x_j)\right],
\end{aligned}$$

where we have again used the fact that the k_js sum to k to extract the $\tilde{\varphi}[f]^k$ term.

Using the expressions for $H[f](x,u,t)$ and $H[f](x)$ together with the definition of $L_k(x)$, we have, for any $\epsilon > 0$, as $t \to \infty$,

$$\sup_{x\in E, f\in B_1^+(E)} |\mathrm{e}^{-k\lambda_* t}\varphi(x)^{-1}\psi_t^{(k)}[f](x) - k!\tilde{\varphi}[f]^k L_k(x)|$$

$$\le \sup_{x\in E, f\in B_1^+(E)} \left| t\int_0^1 \mathrm{e}^{-\lambda_*(k-1)ut}\mathrm{e}^{-\lambda_* ut}\varphi(x)^{-1}\psi_{ut}\left[H[f](\cdot,u,t)\right.\right.$$

$$\left.\left.-H[f]\right](x)\mathrm{d}u\right| + \epsilon$$

$$\le t\int_0^1 \mathrm{e}^{-\lambda_*(k-1)ut} \sup_{x\in E, f\in B_1^+(E)} \left|\mathrm{e}^{-\lambda_* ut}\varphi(x)^{-1}\psi_{ut}\left[H[f](\cdot,u,t)\right.\right.$$

$$\left.\left.-H[f]\right](x)\right|\mathrm{d}u + \epsilon, \tag{9.20}$$

where ϵ is an upper estimate for

$$\sup_{x\in E, f\in B_1^+(E)} k!\tilde{\varphi}[f]^k \int_t^\infty \mathrm{e}^{-\lambda_* ks}\varphi(x)^{-1}\psi_s\left[\gamma\mathcal{E}_\cdot\left[\sum_{[k_1,\ldots,k_N]_k^2}\prod_{\substack{j=1\\ j:k_j>0}}^N L_{k_j}(x_j)\right]\right](x)\mathrm{d}s. \tag{9.21}$$

Note that convergence to zero as $t\to\infty$ in (9.21) follows thanks to the induction hypothesis (ensuring that $L_{k_j}(x)$ is uniformly bounded), (9.8), (G2), and the uniform boundedness of γ.

The induction hypothesis, (9.8), (G2), and dominated convergence ensure that

$$\lim_{t\to\infty}\sup_{x\in E, f\in B_1^+(E), u\in[0,\varepsilon]} |H[f](x,u,t) - H[f](x)| = 0. \tag{9.22}$$

As such, in (9.20), we can split the integral on the right-hand side over $[0,\varepsilon]$ and $(\varepsilon, 1]$, for $\varepsilon\in(0,1)$. Using (9.22), we can ensure that, for any arbitrarily small $\varepsilon' > 0$, making use of the boundedness in (G2), there is a global constant $C > 0$ such that, for all t sufficiently large,

$$t\int_0^\varepsilon \mathrm{e}^{-\lambda_*(k-1)ut}\sup_{x\in E, f\in B_1^+(E)}\left|\mathrm{e}^{-\lambda_* ut}\varphi(x)^{-1}\psi_{ut}\left[H[f](\cdot,u,t) - H[f]\right](x)\right|\mathrm{d}u$$

$$\le \varepsilon' Ct\int_0^\varepsilon \mathrm{e}^{-\lambda_*(k-1)ut}\mathrm{d}u$$

$$= \frac{\varepsilon' C}{\lambda_*(k-1)}(1-\mathrm{e}^{-\lambda_*(k-1)\varepsilon t}). \tag{9.23}$$

On the other hand, we can also control the integral over $(\varepsilon, 1]$, again appealing to (G2), (G2), and (9.8) to ensure that

$$\sup_{x\in E, f\in B_1^+(E), u\in(\varepsilon,1]} \left| \mathrm{e}^{-\lambda_* ut}\varphi(x)^{-1}\psi_{ut}\left[H[f](\cdot,u,t) - H[f]\right]\right| < \infty.$$

We can again work with a (different) global constant $C > 0$ such that

$$\begin{aligned} t\int_\varepsilon^1 &\mathrm{e}^{-\lambda_*(k-1)ut} \sup_{x\in E, f\in B_1^+(E)} \left| \mathrm{e}^{-\lambda_* ut}\varphi(x)^{-1}\psi_{ut}\left[H[f](\cdot,u,t) - H[f]\right]\right| \mathrm{d}u \\ &\leq Ct\int_\varepsilon^1 \mathrm{e}^{-\lambda_*(k-1)ut}\mathrm{d}u \\ &= \frac{C}{\lambda_*(k-1)}(\mathrm{e}^{-\lambda_*(k-1)\varepsilon t} - \mathrm{e}^{-\lambda_*(k-1)t}). \end{aligned} \tag{9.24}$$

In conclusion, using (9.23) and (9.24), we can take limits as $t \to \infty$ in (9.20) and the statement of the theorem follows. □

Finally we state and prove the subcritical case. Our proof will be even briefer given the similarities to the previous two settings. The take-home message is nonetheless different again. Unlike the supercritical setting, this time the k-th moment scales like the first moment.

Theorem 9.3 (Subcritical, $\lambda_* < 0$) *Suppose that (G2) holds and (G2) holds for some $k \geq 1$. Redefine*

$$\Delta_t^{(k)} = \sup_{x\in E, f\in B_1^+(E)} \left|\varphi^{-1}\mathrm{e}^{-\lambda_* t}\psi_t^{(k)}[f](x) - L_k\right|,$$

where $L_1 = \tilde{\varphi}[f]$ and, for $k \geq 2$,

$$L_k = \tilde{\varphi}[f^k] + \int_0^\infty \mathrm{e}^{-\lambda_* s}\tilde{\varphi}\left[\gamma\mathcal{E}_\cdot\left[\sum_{[k_1,\ldots,k_N]_k^2}\binom{k}{k_1,\ldots,k_N}\prod_{\substack{j=1\\ j:k_j>0}}^N \psi_s^{(k_j)}[f](x_j)\right]\right]\mathrm{d}s.$$

Then, for all $\ell \leq k$,

$$\sup_{t\geq 0}\Delta_t^{(\ell)} < \infty \text{ and } \lim_{t\to\infty}\Delta_t^{(\ell)} = 0.$$

Proof First note that since we only compensate by $\mathrm{e}^{-\lambda_* t}$, the term $\psi_t[f^k](x)$ that appears in Eq. (9.4) does not vanish after the normalisation. Due to assumption (G2), we have

$$\lim_{t\to\infty}\varphi^{-1}(x)\mathrm{e}^{-\lambda_* t}\psi_t[f^k](x) = \tilde{\varphi}[f^k].$$

Next we turn to the integral term in (9.4). Define $[k_1, \dots, k_N]_k^{(n)}$, for $2 \le n \le k$ to be the set of tuples $(k_1, \dots, k_N)$ with exactly n positive terms and whose sum is equal to k. Similar calculations to those given above yield

$$\frac{\mathrm{e}^{-\lambda_* t}}{\varphi(x)} \int_0^t \psi_s \left[\gamma \mathcal{E}_x \left[\sum_{[k_1,\dots,k_N]_k^2} \binom{k}{k_1, \dots, k_N} \prod_{j=1}^N \psi_{t-s}^{(k_j)}[f](x_j) \right]\right](x)\mathrm{d}s$$

$$= t \int_0^1 \sum_{n=2}^k \mathrm{e}^{\lambda_*(n-1)ut} \frac{\mathrm{e}^{-\lambda_*(1-u)t}}{\varphi(x)}$$

$$\times \psi_{(1-u)t}\Big[\gamma \mathcal{E}_\cdot\Big[\sum_{[k_1,\dots,k_N]_k^{(n)}} \binom{k}{k_1, \dots, k_N} \prod_{j=1}^N \mathrm{e}^{-\lambda_* ut} \psi_{ut}^{(k_j)}[f](x_j)\Big]\Big](x)\mathrm{d}u. \tag{9.25}$$

Now suppose for induction that the result holds for all ℓ-th integer moments with $1 \le \ell \le k-1$. Roughly speaking, the argument can be completed by noting that the integral in the definition of L_k can be written

$$\int_0^\infty \sum_{n=2}^k \mathrm{e}^{\lambda_*(n-1)s} \tilde{\varphi}\Big[\gamma \mathcal{E}_\cdot\Big[\sum_{[k_1,\dots,k_N]_k^{(n)}} \binom{k}{k_1, \dots, k_N} \prod_{j=1}^N \mathrm{e}^{-\lambda_* s} \psi_s^{(k_j)}[f](x_j)\Big]\Big] \mathrm{d}s, \tag{9.26}$$

which is convergent by appealing to (9.8), hypothesis (G2), the fact that $\gamma \in B^+(E)$, and the induction hypothesis. As a convergent integral, it can be truncated at $t > 0$ and the residual of the integral over (t, ∞) can be made arbitrarily small by taking t sufficiently large. By changing variables in (9.26) when the integral is truncated at arbitrarily large t, so it is of a similar form to that of (9.25), we can subtract it from (9.25) to get

$$t \int_0^1 \sum_{n=2}^k \mathrm{e}^{\lambda_*(n-1)ut} \left(\frac{\mathrm{e}^{-\lambda_*(1-u)t}}{\varphi(x)} \psi_{(1-u)t}[H_{ut}^{(n)}] - \tilde{\varphi}[H_{ut}^{(n)}] \right) \mathrm{d}u,$$

where

$$H_{ut}^{(n)}(x) = \gamma \mathcal{E}_x\Big[\sum_{[k_1,\dots,k_N]_k^{(n)}} \binom{k}{k_1, \dots, k_N} \prod_{j=1}^N \mathrm{e}^{-\lambda_* ut} \psi_{ut}^{(k_j)}[f](x_j)\Big].$$

One proceeds to splitting the integral of the difference over $[0, 1]$ into two integrals, one over $[0, 1-\varepsilon]$ and one over $(1-\varepsilon, 1]$. For the aforesaid integral over $[0, 1-\varepsilon]$, we can control the behaviour of $\varphi^{-1}\mathrm{e}^{-(1-u)t}\psi_{(1-u)t}[H_{ut}^{(n)}] - \tilde{\varphi}[H_{ut}^{(n)}]$ as $t \to$

∞, making it arbitrarily small, by appealing to uniform dominated control of its argument in square brackets thanks to (G2). The integral over $[0, 1-\varepsilon]$ can thus be bounded, as $t \to \infty$, by $t(1 - \mathrm{e}^{\lambda_*(n-1)(1-\varepsilon)})/|\lambda_*|(n-1)$.

For the integral over $(1-\varepsilon, 1]$, we can appeal to the uniformity in (G2) and (G2) to control the entire term $\mathrm{e}^{-(1-u)t}\psi_{(1-u)t}[H^{(n)}_{ut}]$ (over time and its argument in the square brackets) by a global constant. Up to a multiplicative constant, the magnitude of the integral is thus of order

$$t\int_{1-\varepsilon}^{1} \mathrm{e}^{\lambda_*(n-1)ut}\mathrm{d}u = \frac{1}{|\lambda_*|(n-1)}(\mathrm{e}^{\lambda_*(n-1)(1-\varepsilon)t} - \mathrm{e}^{\lambda_*(n-1)t}),$$

which tends to zero as $t \to \infty$. □

9.4 Moments of the Running Occupation at Criticality

As alluded to in the introduction to this chapter, the methods we have used in the previous section are extremely robust and are equally applicable to the setting of the moments of the running occupation. In the critical setting, we have the following main result.

Theorem 9.4 (Critical, $\lambda_* = 0$) *Suppose that (G2) holds along with (G2) for $k \geq 1$. Define*

$$\Delta_t^{(\ell)} = \sup_{x\in E,\, f\in B_1^+(E)} \left| t^{-(2\ell-1)}\varphi(x)^{-1} I_t^{(\ell)}[g](x) - 2^{-(\ell-1)}\ell!\,\tilde{\varphi}[g]^{\ell}\tilde{\varphi}\big[\mathcal{V}[\varphi]\big]^{\ell-1}L_\ell \right|,$$

where $L_1 = 1$ and L_k is defined through the recursion $L_k = (\sum_{i=1}^{k-1} L_i L_{k-i})/(2k-1)$. Then, for all $\ell \leq k$ and $c > 0$,

$$\sup_{t\geq c} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t\to\infty} \Delta_t^{(\ell)} = 0. \tag{9.27}$$

Proof Taking account of (8.27), we see that

$$\mathtt{u}_t[0,\theta g](x) = -\int_0^t \psi_s\left[\mathtt{A}[\mathtt{u}_{t-s}[0,\theta g]] - \theta g(1-\mathtt{u}_{t-s}[0,\theta g])\right](x)\mathrm{d}s. \tag{9.28}$$

Given the proximity of (9.28) to (9.5), it is easy to see that we can apply the same reasoning that we used for $\psi_t^{(k)}[f](x)$ to $I_t^{(k)}[g](x)$ and conclude that, for $k \geq 2$,

$$I_t^{(k)}[g](x) = \int_0^t \psi_s\left[\gamma\hat{\eta}_{t-s}^{(k-1)}[g]\right](x) - k\psi_s[g I_{t-s}^{(k-1)}[g]](x)\mathrm{d}s, \tag{9.29}$$

where $\hat{\eta}^k$ plays the role of η^k albeit replacing the moment operators $\psi^{(j)}$ by the moment operators $I^{(j)}$.

We now proceed to prove Theorem 9.4 by induction. First we consider the setting $k = 1$. In that case,

$$\frac{1}{t}I^{(1)}[g](x) = \frac{1}{t}\mathbb{E}_{\delta_x}\left[\int_0^t X_s[g]\mathrm{d}s\right] = \frac{1}{t}\int_0^t \psi_s[g](x)\mathrm{d}s = \int_0^1 \psi_{ut}[g](x)\mathrm{d}u.$$

Referring now to Theorem 8.1, we can take $F(x, s, t) = g(x)/\varphi(x)$. Since $g \in B^+(E)$, the conditions of the theorem are trivially met and hence

$$\lim_{t\to\infty} \sup_{x\in E, g\in B_1^+(E)} \left|\frac{1}{t}\varphi(x)^{-1}I^{(1)}[g](x) - \tilde{\varphi}[g]\right| = 0.$$

Note that this limit sets the scene for the polynomial growth in $t^{n(k)}$ of the higher moments for some function $n(k)$. If we are to argue by induction, whatever the choice of $n(k)$, it must satisfy $n(1) = 1$.

Next suppose that Theorem 9.4 holds for all integer moments up to and including $k - 1$. We have from (9.29) that

$$\begin{aligned}
&\frac{1}{t^{2k-1}}I_t^{(k)}[g](x)\\
&= \frac{1}{t^{2k-1}}\int_0^t \psi_s\left[\gamma\hat{\eta}_{t-s}^{(k-1)}[g]\right](x)\mathrm{d}s - \frac{1}{t^{2k-1}}\int_0^t k\psi_s[gI_{t-s}^{(k-1)}[g]](x)\mathrm{d}s.
\end{aligned} \tag{9.30}$$

Let us first deal with the rightmost integral in (9.30). It can be written as

$$\begin{aligned}
&\frac{1}{t^{2k-2}}\int_0^1 k\psi_{ut}\left[\varphi F(\cdot, u, t)\right](x)\mathrm{d}u\\
&= \int_0^1 (1-u)^{2k-2}k\psi_{ut}\left[g\frac{1}{(t(1-u))^{2k-2}}I_{t(1-u)}^{(k-1)}[g]\right](x)\mathrm{d}u,
\end{aligned}$$

where F is defined by the equality.

Arguing as in the spirit of the proof of Theorem 9.1, our induction hypothesis ensures that

$$\begin{aligned}
\lim_{t\to\infty} F[g](x, u, t) &= \lim_{t\to\infty} g(1-u)^{2k-2}k\frac{1}{(t(1-u))^{2k-2}}\frac{I_{t(1-u)}^{(k-1)}[g](x)}{\varphi(x)} = 0\\
&=: F(x, u)
\end{aligned}$$

satisfies (8.20) and (8.21). Theorem 8.1 thus tells us that, uniformly in $x \in E$ and $g \in B_1^+(E)$,

$$\lim_{t\to\infty}\frac{1}{t^{2k-1}}\varphi(x)^{-1}\int_0^t k\psi_s[gI^{(k-1)}_{t-s}[g]](x)=0. \tag{9.31}$$

On the other hand, again following the style of the reasoning in the proof of Theorem 9.1, we can pull out the leading order terms, uniformly for $x\in E$ and $g\in B^+(E)$,

$$\begin{aligned}&\lim_{t\to\infty}\frac{1}{t^{2k-1}}\int_0^t \psi_s\left[\gamma\hat{\eta}^{(k-1)}_{t-s}[g]\right](x)\mathrm{d}s\\ &=\lim_{t\to\infty}\int_0^1 \psi_{ut}\Bigg[\frac{\gamma(\cdot)}{2}(1-u)^{2k-2}\mathcal{E}_{\cdot}\Bigg[\sum_{i=1}^N\sum_{\substack{j=1\\ j\neq i}}^N\sum_{k_i=1}^{k-1}\binom{k}{k_i}\varphi(x_i)\varphi(x_j)\\ &\qquad\times\frac{I^{k_i}_{t(1-u)}[g](x_i)}{\varphi(x_i)(t(1-u))^{2k_i-1}}\frac{I^{k-k_i}_{t(1-u)}[g](x_j)}{\varphi(x_j)(t(1-u))^{2k-2k_i-1}}\Bigg]\Bigg](x)\mathrm{d}u.\end{aligned} \tag{9.32}$$

It is again worth noting here that the choice of the polynomial growth in the form $t^{n(k)}$ also constrains the possible linear choices of $n(k)$ to $n(k)=2k-1$ if we are to respect $n(1)=1$ and the correct distribution of the index across (9.32).

Identifying

$$\begin{aligned}&F[g](x,u,t)\\ &=\frac{\gamma(x)}{2\varphi(x)}(1-u)^{2k-2}\mathcal{E}_x\Bigg[\sum_{i=1}^N\sum_{\substack{j=1\\ j\neq i}}^N\sum_{k_i=1}^{k-1}\binom{k}{k_i}\varphi(x_i)\varphi(x_j)\\ &\qquad\times\frac{I^{k_i}_{t(1-u)}[g](x_i)}{\varphi(x_i)(t(1-u))^{2k_i-1}}\frac{I^{k-k_i}_{t(1-u)}[g](x_j)}{\varphi(x_j)(t(1-u))^{2k-2k_i-1}}\Bigg],\end{aligned}$$

our induction hypothesis allows us to conclude that $F[g](x,u):=\lim_{t\to\infty}F[g](x,u,t)$ exists and

$$\varphi(x)F[g](x,u)=(1-u)^{2k-2}k!\frac{\gamma(x)\mathcal{V}[\varphi](x)}{2^{k-1}}\tilde{\varphi}[g]^k\tilde{\varphi}[\mathcal{V}[\varphi]]^{k-1}\sum_{\ell=1}^{k-1}L_\ell L_{k-\ell}.$$

Thanks to our induction hypothesis, we can also easily verify (8.20) and (8.21). Theorem 8.1 now gives us the required uniform (in $x\in E$ and $g\in B^+(E)$) limit

$$\lim_{t\to\infty}\frac{1}{t^{2k-1}}\varphi(x)^{-1}\int_0^t\psi_s\left[\gamma\hat{\eta}^{(k-1)}_{t-s}[g]\right](x)\mathrm{d}s=\frac{k!\tilde{\varphi}[\mathcal{V}[\varphi]]^{k-1}\tilde{\varphi}[g]^k}{2^{k-1}}L_k. \tag{9.33}$$

Putting (9.33) together with (9.31), we get the statement of Theorem 9.4. $\square$

9.5 Moments of the Running Occupation at Non-criticality

We start with the supercritical setting and then move to the subcritical setting. Given the familiarity of the approach at this stage, the proofs we give are very brief, pointing only to the key steps.

Theorem 9.5 (Supercritical, $\lambda_* > 0$) *Suppose that (G2) holds along with (G2) for some $k \geq 1$ and $\lambda_* > 0$. Redefine*

$$\Delta_t^{(\ell)} = \sup_{x\in E,\, f\in B_1^+(E)} \left| \mathrm{e}^{-\ell\lambda_* t}\varphi(x)^{-1} I_t^{(\ell)}[g](x) - \ \ell!\tilde{\varphi}[g]^\ell L_\ell(x) \right|,$$

where $L_k(x)$ was defined in Theorem 9.2 but now with $L_1(x) = 1/\lambda_*$.
Then, for all $\ell \leq k$,

$$\sup_{t\geq 0} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t\to\infty} \Delta_t^{(\ell)} = 0. \tag{9.34}$$

Proof For the case $k = 1$, we have

$$\begin{aligned}
&\left| \mathrm{e}^{-\lambda_* t} \int_0^t \varphi(x)^{-1} \psi_s[g](x)\mathrm{d}s - \frac{\tilde{\varphi}[g]}{\lambda_*} \right| \\
&= \left| \mathrm{e}^{-\lambda_* t} t \int_0^1 \mathrm{e}^{\lambda_* ut} \left(\mathrm{e}^{-\lambda_* ut}\varphi(x)^{-1}\psi_{ut}[g](x) - \tilde{\varphi}[g] \right) \mathrm{d}u - \mathrm{e}^{-\lambda_* t}\frac{\tilde{\varphi}[g]}{\lambda_*} \right| \\
&\leq t \int_0^1 \mathrm{e}^{-\lambda_*(1-u)t} \left| \mathrm{e}^{-\lambda_* ut}\varphi(x)^{-1}\psi_{ut}[g](x) - \tilde{\varphi}[g] \right| \mathrm{d}u + \mathrm{e}^{-\lambda_* t}\frac{\tilde{\varphi}[g]}{\lambda_*}.
\end{aligned} \tag{9.35}$$

Thanks to (G2) and similar arguments to those used in the proof of Theorem 9.2, we may choose t sufficiently large such that the modulus in the integral on the right-hand side of (9.35) is bounded above by an arbitrary small $\varepsilon' > 0$, uniformly in $x \in E$, $g \in B^+(E)$, and $u > \varepsilon \in (0, 1)$. Then, when restricted to $(\varepsilon, 1]$, the aforementioned integral is bounded above by $\varepsilon'(1 - \mathrm{e}^{-\lambda_*\varepsilon t})$. On the other hand, when restricted to $[0, \varepsilon]$, up to a global multiplicative constant, again thanks to (G2), this integral can be bounded by $\mathrm{e}^{-\lambda_*(1-\varepsilon)t} - \mathrm{e}^{-\lambda_* t}$. Since ε' can be taken arbitrarily small and t tends to infinity, this gives the desired result for the integral on the right-hand side of (9.35). The last term in (9.35) is dealt with trivially. The limit in (9.35) also pins down the initial value $L_1(x) = 1/\lambda_*$.

Now assume the result holds for all integers $1 \leq \ell \leq k - 1$. Reflecting on the proof of Theorem 9.2, in the current setting the starting point is (9.29), which is almost the same as (9.4). Our task is thus to evaluate, in the appropriate sense,

$$
\begin{aligned}
&\lim_{t\to\infty} \mathrm{e}^{-\lambda_* kt}\varphi(x)^{-1} I_t^{(k)}[g](x) \\
&= \lim_{t\to\infty} \mathrm{e}^{-\lambda_* kt} \int_0^t \varphi(x)^{-1}\psi_s\left[\gamma\mathcal{E}_\cdot\left[\sum_{[k_1,\ldots,k_N]_k^2} \binom{k}{k_1,\ldots,k_N}\prod_{j=1}^N I_{t-s}^{(k_j)}[g](x_j)\right]\right](x)\mathrm{d}s \\
&\qquad - k \lim_{t\to\infty} \mathrm{e}^{-\lambda_* kt}\int_0^t \varphi(x)^{-1}\psi_s[g I_{t-s}^{(k-1)}[g]](x)\mathrm{d}s. \qquad (9.36)
\end{aligned}
$$

The first term on the right-hand side of (9.36) can be handled in essentially the same way as in the proof of Theorem 9.2. The second term on the right-hand side of (9.36) can easily be dealt with along the lines that we are now familiar with from earlier proofs, using the induction hypothesis. In particular, its limit is zero. Hence, combined with the first term on the right-hand side of (9.36), we recover the same recursion equation for L_k. □

Theorem 9.6 (Subcritical, $\lambda_* < 0$) *Suppose that (G2) holds along with (G2) for some $k \geq 1$. Redefine*

$$
\Delta_t^{(\ell)} = \sup_{x\in E,\, f\in B_1^+(E)} \left|\varphi(x)^{-1} I_t^{(\ell)}[g](x) - L_\ell(x)\right|,
$$

where $L_1(x) = \int_0^\infty \varphi(x)^{-1}\psi_s[g](x)\mathrm{d}s$, and for $k \geq 2$, the constants L_k are defined recursively via

$$
\begin{aligned}
L_k(x) = \int_0^\infty \varphi(x)^{-1}\psi_s\Bigg[\gamma\mathcal{E}_\cdot\Bigg[\sum_{[k_1,\ldots,k_N]_k^2}\binom{k}{k_1,\ldots,k_N}\prod_{\substack{j=1\\ j:k_j>0}}^N \varphi(x_j)L_{k_j}(x_j)\Bigg]\Bigg](x)\,\mathrm{d}s \\
- k\int_0^\infty \varphi(x)^{-1}\psi_s\Big[g\varphi L_{k-1}\Big](x)\,\mathrm{d}s.
\end{aligned}
$$

Then, for all $\ell \leq k$,

$$
\sup_{t\geq 0}\Delta_t^{(\ell)} < \infty \text{ and } \lim_{t\to\infty}\Delta_t^{(\ell)} = 0. \qquad (9.37)
$$

Proof The case $k = 1$ is relatively straightforward, and, again, in the interest of keeping things brief, we point the reader to the fact that, as $t \to \infty$, we have

$$
I_t^{(1)}[g](x) \sim \int_0^\infty \psi_s[g](x)\mathrm{d}s < \infty, \qquad (9.38)
$$

thanks to the exponential decay of $(\psi_t, t \geq 0)$, since $\lambda_* < 0$.

Now suppose the result holds for all integers $1 \leq \ell \leq k-1$. We again refer to (9.29), which means we are interested in handling a limit which is very similar to (9.36), now taking the form

$$\begin{aligned}
&\lim_{t\to\infty} I_t^{(k)}[g](x) \\
&= \lim_{t\to\infty} t \int_0^1 \psi_{ut}\left[\gamma \mathcal{E}_\cdot\left[\sum_{[k_1,\dots,k_N]_k^2} \binom{k}{k_1,\dots,k_N} \prod_{j=1}^N I_{(1-u)t}^{(k_j)}[g](x_j)\right]\right](x)\mathrm{d}u \\
&\qquad - \lim_{t\to\infty} kt \int_0^1 \psi_{ut}\left[g I_{(1-u)t}^{(k-1)}[g]\right](x)\mathrm{d}u. \qquad (9.39)
\end{aligned}$$

Again skipping the details, by treating the integral in (9.39) according to its behaviour over $[0, 1-\varepsilon]$ and $(1-\varepsilon, 1]$, we can quickly see from (9.39) the argument in (9.38), and the induction hypothesis gives us

$$\begin{aligned}
I_t^{(k)}[g](x) \sim \int_0^\infty \psi_s\Big[\gamma \mathcal{E}_\cdot\Big[\sum_{[k_1,\dots,k_N]_k^2} \binom{k}{k_1,\dots,k_N} \prod_{\substack{j=1\\ j:k_j>0}}^N L_{k_j}(x_j)\Big]\Big](x)\,\mathrm{d}s \\
-k\int_0^\infty \psi_s\Big[gL_{k-1}(x)\Big](x)\,\mathrm{d}s, \qquad (9.40)
\end{aligned}$$

which gives us the required recursion for $L_k(x)$. Making these calculations rigorous (left to the reader) as in Theorems 9.2, 9.3, and 9.5 completes the proof. □

9.6 Moments for Discrete-Time Branching Markov Processes

Recall that we introduced the notion of a discrete-time branching Markov process in Sect. 8.7. Just as in the previous sections, under the analogous assumptions to (G2) and (G2), we can expect similar moment growth results.

(G4) **There exist an eigenvalue $\rho_* > 0$, and a corresponding right eigenfunction $0 \leq \omega \in B^+(E)$, and finite eigenmeasure $\tilde{\omega}$ such that, for $f \in B^+(E)$,**

$$\Phi_n[\omega] = \rho_*^n \omega \text{ and } \tilde{\omega}\left[\Phi_n[f]\right] = \rho_*^n \tilde{\omega}[f].$$

Furthermore, let us define

$$\Delta_n = \sup_{x\in E, g\in B_1^+(E)} |\omega(x)^{-1}\rho_*^{-n}\Phi_n[f](x) - \tilde{\omega}[f]|, \qquad n \geq 0. \tag{9.41}$$

We suppose that

$$\sup_{n\geq 0}\Delta_n < \infty \text{ and } \lim_{n\to\infty}\Delta_n = 0. \tag{9.42}$$

(G5) **We have**

$$\sup_{x\in E} E_x(\mathscr{Z}[\mathbf{1}]^k) < \infty. \tag{9.43}$$

Naturally, ρ_*^n plays the role of the lead eigenvalue for the semigroup Φ_n, with corresponding right eigenfunction ω and left eigenmeasure $\tilde{\omega}$. In this setting, ρ_*^n gives the average growth of the number of particles in the system in generation n and thus gives us an analogous notion of criticality to the continuous-time setting. That is, $\rho_* > 1$ corresponds to a *supercritical* system, $\rho_* < 1$ corresponds to a *subcritical* system, and $\rho_* = 1$ corresponds to a *critical* system. The assumption (G5) for a fixed $k \geq 2$ provides control over the moments of the offspring distribution at the first generation. We refer the reader to Remark 8.1 for a brief discussion of the case where the discrete-time BMP is embedded into a continuous-time BMP.

Just as in the continuous-time setting, we obtain perfectly analogous results for the asymptotic moments. We list them below.

Theorem 9.7 (Critical, $\rho_* = 1$) *Suppose that (G4) holds along with (G5) for $k \geq 1$ and $\rho_* = 1$. Define*

$$\Delta_n^{(\ell)} = \sup_{x\in E, g\in B_1^+(E)} \left| n^{-(\ell-1)}\omega(x)^{-1}\Phi_n^{(\ell)}[g](x) - 2^{-(\ell-1)}\ell!\,\tilde{\omega}[f]^\ell\tilde{\omega}[\mathscr{V}[\omega]]^{\ell-1}\right|,$$

where, again abusing notation from the continuous-time setting,

$$\mathscr{V}[\omega](x) = E_x\left[\mathscr{Z}[\omega]^2 - \mathscr{Z}[\omega^2]\right] = E_x\left[\sum_{i=1}^N\sum_{\substack{j=1\\j\neq i}}^N \omega(z_i)\omega(z_j)\right].$$

Then, for all $\ell \leq k$,

$$\sup_{n\geq 1}\Delta_n^{(\ell)} < \infty \textit{ and } \lim_{n\to\infty}\Delta_n^{(\ell)} = 0. \tag{9.44}$$

Theorem 9.8 (Supercritical, $\rho_* > 1$) *Suppose that (G4) holds along with (G5) for $k \geq 1$ and $\rho_* > 1$. Define*

$$\Delta_n^{(\ell)} = \sup_{x\in E, g\in B_1^+(E)} \left|\rho_*^{-nk}\omega(x)^{-1}\Phi_n^{(\ell)}[g](x) - \ell!\,\tilde{\omega}[f]^\ell L_\ell(x)\right|,$$

where $L_1 = 1$, and for $k \geq 2$, $L_k(x)$ is given by the recursion

$$L_k(x) = \omega(x)^{-1}\sum_{\ell=0}^{\infty}\rho_*^{-\ell(k+1)}\Phi_\ell\left[E_\cdot\left[\sum_{[k_1,\ldots,k_N]_k^{2+}}\prod_{\substack{j=1\\k_j>0}}^{N}\omega(x_j)L_{k_j}(x_j)\right]\right](x), \tag{9.45}$$

with $[k_1,\ldots,k_N]_k^{2+}$ defining the set of non-negative tuples $(k_1,\ldots,k_N)$, such that $\sum_{j=1}^N k_j = N$ and at least two of the k_j are strictly positive.

Then, for all $\ell \leq k$,

$$\sup_{n\geq 0}\Delta_n^{(\ell)} < \infty \text{ and } \lim_{n\to\infty}\Delta_n^{(\ell)} = 0.$$

Theorem 9.9 (Subcritical, $\rho_* < 1$) *Suppose that (G4) holds along with (G5) for $k \geq 1$ and $\rho_* < 1$. Define*

$$\Delta_n^{(\ell)} = \sup_{x\in E, g\in B_1^+(E)} \left|\rho_*^{-n}\omega(x)^{-1}\Phi_n^{(\ell)}[g](x) - L_\ell\right|,$$

where $L_1 = 1$, and, for $\ell \geq 2$, L_ℓ is given by the recursion

$$L_\ell = \tilde{\omega}[f^\ell] + \sum_{n=0}^{\infty}\rho_*^{-(n+1)}\tilde{\omega}\left[E_\cdot\left[\sum_{[k_1,\ldots,k_N]_k^{2+}}\binom{k}{k_1,\cdots,k_N}\prod_{\substack{j=1\\k_j>0}}^{N}\Phi_n^{(k_j)}(x_j)\right]\right].$$

Then, for all $\ell \leq k$,

$$\sup_{n\geq 0}\Delta_n^{(\ell)} < \infty \text{ and } \lim_{n\to\infty}\Delta_n^{(\ell)} = 0.$$

The proofs of Theorems 9.7–9.9 are very close to their continuous-time counterparts. Indeed, all three start with the observation that

$$\Phi_n^{(k)}[g](x) = (-1)^k\,\frac{\partial^k}{\partial\theta^k}\mathrm{v}_n[\mathrm{e}^{-\theta g}](x)\Big|_{\theta=0} \qquad n \geq 2, x \in E, f \in B(E),$$

which, in the spirit of Proposition 9.1, leads to

$$\Phi_n^{(k)}[g] = \Phi_n[g^k] + \sum_{\ell=0}^{n-1} \Phi_\ell \left[E_\cdot \left[\sum_{[k_1,\cdots,k_N]_k^2} \binom{k}{k_1,\ldots,k_N} \prod_{j=1}^{N} \Phi_{n-\ell-1}^{(k_j)}[g](z_j) \right]\right], \tag{9.46}$$

on E, for $n \geq 1$ and $g \in B^+(E)$, where $[k_1, \cdots, k_N]_k^2$ is the set of all non-negative N-tuples $(k_1, \ldots, k_N)$ such that $\sum_{i=1}^{N} k_i = k$ and at least two of the k_i are strictly positive.

We now highlight some of the subtle differences in Theorem 9.7 compared to the continuous time.

As with criticality in the continuous setting, it is not difficult to show that the first term on the right-hand side of (9.46) carries no contribution to the asymptotic scaled limit. Moreover, from the sum on the right-hand side of (9.46), the only contribution that matters comes from partitions of the form k_1 and $k_2 = k - k_1$, where $k_1 \in \{1, \cdots, k-1\}$. Our task is thus to show that

$$\Delta_n^{(k),2} = \sup_{x\in E, f\in B_1^+(E)} \left| \frac{1}{n\omega} \sum_{\ell=0}^{n-1} \left(\frac{n-\ell-1}{n}\right)^{k-2} \Phi_\ell\big[\mathrm{H}_{n-\ell-1}^{(k)}[f]\big] \right.$$
$$\left. - 2^{-(k-1)} k!\, \tilde{\omega}\big[\mathscr{V}[\omega]\big]^{k-1} \tilde{\omega}[f]^k \right| \tag{9.47}$$

tends to zero as $n \to \infty$, where (being careful not to double count the non-zero partition $k_1, k_2 : k = k_1 + k_2$)

$$\mathrm{H}_m^{(k)}[f](x) = \frac{1}{2} E_x \left[\sum_{i=1}^{N} \omega(z_i) \sum_{\substack{j=1\\ j\neq i}}^{N} \omega(z_j) \sum_{k_1=1}^{k-1} \binom{k}{k_1} \frac{\Phi_m^{(k_1)}[f](z_i)}{\omega(z_i) m^{k_1-1}} \frac{\Phi_m^{(k-k_1)}[f](z_j)}{\omega(z_j) m^{k-k_1-1}} \right].$$

In a similar way to the analogous part of the proof in the continuous-time setting, we can show that

$$\mathrm{H}_\infty^{(k)}[f](x) := \lim_{m\to\infty} \mathrm{H}_m^{(k)}[f](x) = (k-1)k!\mathscr{V}[\omega](x) 2^{-(k-1)} \tilde{\omega}\big[\mathscr{V}[\omega]\big]^{k-2} \tilde{\omega}[f]^k, \tag{9.48}$$

where, in fact, the convergence can be taken uniformly in both $x \in E$ and $f \in B^+(E)$.

Next, for fixed $\ell \leq n$, define

$$\omega(x) F[g](x, \ell, n) = \left(\frac{n-\ell-1}{n}\right)^{k-2} \mathrm{H}_{n-\ell-1}^{(k)}[g](x),$$

and

$$\omega(x)\hat{F}[g](x,\ell,n) = \left(\frac{n-\ell-1}{n}\right)^{k-2} \mathtt{H}_\infty^{(k)}[g](x). \tag{9.49}$$

Then, the following four facts are easy to verify:

$$\check{F}[g](x) := \lim_{n\to\infty} \frac{1}{n}\sum_{\ell=0}^{n-1} \hat{F}[g](x,\ell,n) = \frac{\mathtt{H}_\infty^{(k)}[f](x)}{k-1} \tag{9.50}$$

exists

$$\lim_{n\to\infty} \sup_{g\in B_1^+(E)} \left|\frac{1}{n}\sum_{\ell=0}^{n-1} \tilde{\omega}\big[\omega\hat{F}[g](\cdot,\ell,n)\big] - \tilde{\omega}\big[\omega\check{F}[g]\big]\right| = 0, \tag{9.51}$$

$$\sup_{x\in E,\ell\le n\in\mathbb{N},g\in B_1^+(E)} |\omega(x)\hat{F}[g](x,\ell,n)| < \infty, \tag{9.52}$$

and

$$\lim_{n\to\infty} \sup_{x\in E,\ell\le n,g\in B_1^+(E)} \omega(x)|F[g](x,\ell,n) - \hat{F}[g](x,\ell,n)| = 0. \tag{9.53}$$

The four properties (9.50)–(9.53) are sufficient to prove an analogue of Theorem 8.1. This states that, under (G4) and (9.50)–(9.53), we have

$$\sup_{n\ge 2} \Xi_n < \infty \text{ and } \lim_{n\to\infty} \Xi_n = 0, \tag{9.54}$$

where

$$\Xi_n = \sup_{x\in E,g\in B_1^+(E)} \left|\frac{1}{n\omega(x)}\sum_{\ell=0}^{n-1} \Phi_\ell[\omega F[g](\cdot,\ell,n)](x)\mathrm{d}u - \tilde{\omega}\big[\omega\check{F}[g]\big]\right|, \qquad t\ge 0.$$

This gives us precisely (9.47), thus proving Theorem 9.7.

The proofs of Theorems 9.8 and 9.9 follow almost verbatim along the same steps in the continuous-time setting albeit, integrals over [0, 1], such as those in (9.20), are played by the role of scaled summations taking the form $n^{-1}\sum_{\ell=0}^{n-1}\cdots$. The details are left to the reader.

9.7 Examples for Specific Branching Processes

We conclude the general technical exposition in this chapter by giving some examples to illustrate our results for specific scenarios, as well as how they relate to neighbouring results.

Continuous-time Bienyamé–Galton–Watson Process We start by considering the simplest branching particle setting where the process is not spatially dependent. In effect, we can take $E = \{0\}$, $\mathtt{P}$ to be the Markov process which remains at $\{0\}$ and a branching mechanism with no spatial dependence. This is the setting of a continuous-time Bienyamé-Galton–Watson process. Its branching rate γ is constant, and the first and second moments of the offspring distribution are given by $m_1 = \mathcal{E}[N]$ and $m_2 = \mathcal{E}[N^2]$, respectively, where N is the number of offspring produced at a branching event. When the process is independent of space, we have $\lambda_* = \gamma(m_1 - 1)$, $\varphi = 1$, $\tilde{\varphi}$ can be taken as $\delta_{\{0\}}$, and (G2) trivially holds. Theorem 9.1 now tells us that, at criticality, i.e., $m_1 = 1$, the limit for the k-th moment of the population size at time $t \geq 0$ satisfies

$$t^{-(k-1)}\mathbb{E}[N_t^k] \sim 2^{-(k-1)}k!\,(\gamma(m_2 - 1))^{k-1}, \qquad \text{as } t \to \infty, \tag{9.55}$$

when $\mathcal{E}[N^k] < \infty$ and $k \geq 1$.

In the supercritical case, i.e., $m_1 > 1$, the limit in Theorem 9.2 simplifies to

$$\mathrm{e}^{-\gamma(m_1-1)kt}\mathbb{E}[N_t^k] \sim k!L_k, \qquad \text{as } t \to \infty, \tag{9.56}$$

where the iteration

$$L_k = \frac{1}{(m_1 - 1)(k - 1)}\mathcal{E}\Bigg[\sum_{[k_1,\ldots,k_N]_k^2} \prod_{\substack{j=1 \\ j:k_j>0}}^{N} L_{k_j}\Bigg], \qquad k \geq 2,$$

holds. Here, although the simplified formula for L_k (on account of no spatial considerations) is still a little complicated, it demonstrates more clearly that the moments in Theorem 9.2 grow according to the leading order terms of the offspring distribution. Indeed, in the case $k = 2$, we have

$$L_2 = \frac{1}{m_1 - 1}\mathcal{E}[\mathrm{card}\{[k_1, \ldots, k_N]_2^2\}] = \frac{1}{m_1 - 1}\frac{\mathcal{E}[N(N-1)]}{2} = \frac{m_2 - m_1}{2(m_1 - 1)}.$$

The constant L_3 can now be computed explicitly in terms of L_2 and $L_1 = 1$, and so on.

The limits in the subcritical case can be detailed similarly and only offer minor simplifications of the constants L_k, $k \geq 1$ presented in the statement of Theorem 9.3. Hence we leave the details for the reader to check.

Branching Brownian Motion in a Bounded Domain In this setting, the semigroup $\mathtt{P}$ corresponds to that of a d-dimensional Brownian motion killed on exiting a C^1 domain $E \subset \mathbb{R}^d$. The branching rate is taken as the constant $\gamma > 0$ and the offspring distribution is not spatially dependent. Moreover, the first and second moments, $m_1 := \mathcal{E}[N]$ and $m_2 = \mathcal{E}[N^2]$, are assumed to be finite. In this setting, the right eigenfunction of the expectation semigroup of the process, φ, exists on E, satisfying Dirichlet boundary conditions, and is accompanied by the left eigenmeasure $\varphi(x)\mathrm{d}x$ on E. The associated eigenvalue is identified explicitly as $\lambda_* = \gamma(m_1 - 1) + \lambda_E$, where λ_E is the ground state eigenvalue of the Laplacian on E. The critical regime thus occurs when $\lambda_E = -\gamma(m_1 - 1)$.

In the spirit of Kolmogorov asymptotic survival probability limit for classical Galton–Watson processes and Theorem 5.8 for NBPs, it is known at criticality that

$$\mathbb{P}_{\delta_x}(\zeta > t) \sim \frac{1}{t}\frac{2(m_1-1)\varphi(x)}{|\lambda_E|(m_2-m_1)\int_E \varphi(x)^3\mathrm{d}x} =: 2\varphi(x)/\Sigma, \qquad x \in E, \tag{9.57}$$

as $t \to \infty$. Moreover, in the spirit of the Yaglom distributional limit for Galton–Watson processes and Theorem 5.9, it is also known that

$$\mathrm{Law}\left(\frac{X_t[f]}{t}\bigg|\, \zeta > t\right) \to \mathrm{Law}(\mathbf{e}_{2/\langle\varphi,f\rangle\Sigma}), \qquad \text{as } t \to \infty, \tag{9.58}$$

where $\mathbf{e}_{\langle\varphi,f\rangle\Sigma/2}$ is an exponentially distributed random variable with rate $\langle\varphi, f\rangle\Sigma/2$. (Note that we understand $\langle f,\varphi\rangle = \int_E \varphi(x)f(x)\mathrm{d}x$ in this context.) In particular, these two results allude to the limit of moments (albeit further moment assumptions would be needed on N), which, in the spirit of (11.28), can be heuristically read as

$$\begin{aligned}\lim_{t\to\infty}\frac{1}{t^{k-1}}\mathbb{E}_{\delta_x}[X_t[f]^k] &= \lim_{t\to\infty} t\mathbb{P}_{\delta_x}(\zeta > t)\mathbb{E}_{\delta_x}\left[\frac{X_t[f]^k}{t^k}\bigg|\, \zeta > t\right] \\ &= k!2^{-(k-1)}\langle f,\varphi\rangle^k\Sigma^{k-1}\varphi(x), \qquad x \in E,\end{aligned} \tag{9.59}$$

for $k \geq 1$. Taking into account the fact that $\tilde{\varphi}(x) = \varphi(x)\mathrm{d}x$ and $\gamma = |\lambda_E|/(m_1-1)$, we see that

$$\tilde{\varphi}[\mathcal{V}[\varphi](x)] = \langle\gamma\varphi^2(m_2-m_1),\varphi\rangle = |\lambda_E|\frac{(m_2-m_1)}{(m_1-1)}\int_E \varphi(x)^3\mathrm{d}x = \Sigma.$$

Hence (9.59) agrees precisely with Theorem 9.1.

Crump–Mode–Jagers (CMJ) Processes Finally, we consider the class of CMJ processes, for which, results are not covered by our class of BMPs. Nonetheless, comparable results to Theorem 9.1 are known, which are worth discussing.

Consider a branching process in which particles live for a random amount of time ζ and during their lifetime give birth to a (possibly random) number of offspring at random times; in essence, the age of a parent at the birth times forms a point process, say $\eta(\mathrm{d}t)$ on $[0, \zeta]$. We denote the law of the latter by $\mathscr{P}$. The offspring reproduce and die as independent copies of the parent particle and the law of the process is denoted by $\mathbb{P}$ when initiated from a single individual. In essence, the CMJ is a non-spatial version of the Biggins process, described in Chap. 8.

Criticality for CMJ processes is usually described in terms of the Malthusian parameter, $\alpha \in \mathbb{R}$, which satisfies $\mathscr{E}[\int_{[0,\infty)} \mathrm{e}^{-\alpha t}\eta(\mathrm{d}t)] = 1$. The critical setting is understood to be the case $\alpha = 0$ (with supercritical $\alpha > 0$ and subcritical $\alpha < 0$). If we write the total number of offspring during a lifetime by $N = \eta[0, \zeta]$, then the critical setting can equivalently be identified by $\mathbb{E}[N] = 1$. Furthermore, let Z_t denote the number of individuals in the population at time $t \geq 0$. Under the moment assumption $\mathbb{E}[N^k] < \infty$ for some $k \geq 1$, it is known that the factorial moments $m_k(t) := \mathbb{E}[Z_t(Z_t - 1)\cdots(Z_t - k + 1)]$ satisfy

$$\lim_{t\to\infty} \frac{m_k(t)}{t^{k-1}} = k!\frac{\mathscr{E}[\zeta]^k}{b^{2k-1}}(m_2 - 1)^{k-1},$$

where $m_2 = \mathscr{E}[N^2]$ and $b = \mathscr{E}[\int_0^\zeta t\eta(\mathrm{d}t)]$. It is an interesting exercise to compare this to the spatially independent example considered above.

9.8 Comments

Despite the fact that understanding the behaviour of moments is a natural and fundamental question to ask for branching Markov processes, until recently, very little appears to be present in the literature beyond second moments. Nonetheless, for higher moments, there are some references which touch upon the topic in a cursory way for the setting of both BMPs and a more exotic class of branching processes known as superprocesses; see, for example, Etheridge [55], Dynkin [48], Harris and Roberts [71], Klenke [83], Fleischman [65], Foutel-Rodier and Schertzer [62] and Powell [112]. This is similarly the case for the occupation measure of branching Markov processes; cf. Dumonteil and Mazzolo [43] and Iscoe [78]. One exception to the lack of treatment of higher moments is Durham [46], whose results for Crump–Mode–Jagers processes in the early 1970s are described in Sect. 9.7. For the general BMP setting, results presented here are reproduced from the recent work of Gonzalez et al. [67].

Chapter 10
Survival at Criticality

We will remain in the setting of the Asmussen–Hering class of BMPs, i.e., assuming (G2), and insist throughout this chapter that we are in the critical setting, that is, $\lambda_* = 0$. Recalling the notation from (8.25), let us define

$$\mathtt{u}_t(x) := \mathbb{P}_{\delta_x}(\zeta > t) \text{ and } a(t) := \tilde{\varphi}[\mathtt{u}_t], \tag{10.1}$$

where $\zeta = \inf\{t > 0 : X_t[1] = 0\}$ is the lifetime of the process. Setting $g = \mathbf{0}$ (the zero function) and $f = \theta$ and letting $\theta \to \infty$ in (8.27), we find that

$$a(t) = \tilde{\varphi}[\psi_t[\mathbf{1}]] - \int_0^t \tilde{\varphi}\big[\psi_s\big[\mathtt{A}[\mathtt{u}_{t-s}]\big]\big]\mathrm{d}s = \tilde{\varphi}[\mathbf{1}] - \int_0^t \tilde{\varphi}\big[\mathtt{A}[\mathtt{u}_s]\big]\mathrm{d}s, \tag{10.2}$$

where $\mathbf{1}$ is the function that is identically unity, we have used the fact that $\tilde{\varphi}$ is a left eigenmeasure with zero eigenvalue at criticality, and we have changed the variable of integration in the second equality. Our aim is to use (10.2) to give us the rate of decay of $a(t)$, and hence $\mathtt{u}_t(x)$, thereby generalising Theorem 5.8 to the setting of BMPs. When coupled with what we know of moment evolution, this will also give us what we need to prove a general Yalgom limit for branching Markov processes in the spirit of Theorem 5.9.

10.1 Yaglom Limit Results for General BMPs

Before stating the main results of this chapter, we need to state a number of additional assumptions. As with much of our notation, we use the same symbols as we have used for the NBP in this more general setting without confusion. For example, we will still work with the variance functional for the branching mechanism defined by

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8_10

$$\mathscr{V}[g](x) = \mathscr{E}_x\left[\mathsf{Z}[g]^2 - \mathsf{Z}[g^2]\right], \qquad x \in E. \tag{10.3}$$

See (5.9) for comparison. We now introduce some further assumptions.

(G6) **The number of offspring produced at a branching event is bounded above by a constant,** $n_{\max}$.

(G7) **There exists a constant** $C > 0$ **such that for all** $g \in B^+(E)$,

$$\tilde{\varphi}\big[\gamma\mathscr{V}[g]\big] \geq C\tilde{\varphi}[g]^2.$$

(G8) **For all** t **sufficiently large,** $\sup_{x\in E}\mathbb{P}_{\delta_x}(t < \zeta) < 1$.

The assumption (G6) simply assumes there is never more than a maximum number of offspring, irrespective of spatial dependency. Assumption (G7) can be thought of as a type of *"spread-out-ness"* requirement that ensures an inherent irreducibility of how the event of branching contributes to mass transportation. We may also think of this as something analogous to a uniform ellipticity condition for diffusive operators. Assumption (G8) ensures that there are no anomalies in our BMP that would allow for guaranteed survival from certain points in space. Note, in particular, that this becomes relevant when E is not bounded.

Theorem 10.1 *Suppose that (G2), (G6), (G7), and (G8) hold. Then,*

$$\lim_{t\to\infty}\sup_{x\in E}\left|\frac{t\mathbb{P}_{\delta_x}(\zeta > t)}{\varphi(x)} - \frac{2}{\Sigma}\right| = 0, \tag{10.4}$$

where

$$\Sigma = \tilde{\varphi}\big[\gamma\mathscr{V}[\varphi]\big]. \tag{10.5}$$

As alluded to above, Theorem 10.1, when combined with Theorem 9.1, implies that, for any $k \geq 1$ and $f \in B^+(E)$,

$$\lim_{t\to\infty}\mathbb{E}_{\delta_x}\left[\left(\frac{X_t[f]}{t}\right)^k \middle| \zeta > t\right] = k!\,\tilde{\varphi}[f]^k\left(\frac{\Sigma}{2}\right)^k. \tag{10.6}$$

The right-hand side above is precisely the k-th moment of an exponential random variable with rate $p := 2/\Sigma\tilde{\varphi}[f]$. In other words, as a generalisation of Theorem 5.9, we have

$$\mathtt{Law}\left(\frac{X_t[f]}{t}\middle| \zeta > t\right) \to \mathtt{Law}\left(\mathbf{e}_p\right),$$

as $t \to \infty$, where $\mathbf{e}_p$ is an exponential random variable with rate p.

We consume the rest of this chapter proving Theorem 10.1. The approach we will take is to first produce coarse lower and upper bounds for the survival probability and then to bootstrap these bounds to give us the precise asymptotic in (10.4). As alluded to above, a key element in our analysis will be understanding the behaviour of $a(t)$ as $t \to \infty$.

10.2 Extinction at Criticality

Let us start by examining extinction, namely the event $\{\zeta < \infty\}$, at criticality. Under our assumptions, as with classical Bienyamé–Galton–Watson branching processes, we find that criticality ensures there is almost sure extinction.

Lemma 10.1 *Assume (G2) with $\lambda_* = 0$ and (G8). For all $x \in E$, we have $\mathbb{P}_{\delta_x}(\zeta < \infty) = 1$.*

Proof We start by proving that for all $x \in E$ and $t_0 > 0$,

$$\lim_{n\to\infty} \mathbb{P}_{X_n}(\zeta \le t_0) = \mathbf{1}_{\{\zeta<\infty\}}, \qquad \mathbb{P}_{\delta_x}\text{-a.s.} \tag{10.7}$$

On the event $\{\zeta < \infty\}$, it is immediate that, for all $x \in E$,

$$\lim_{t\to\infty} \mathbb{P}_{X_t}(\zeta \le t_0) = 1, \tag{10.8}$$

$\mathbb{P}_{\delta_x}$-almost surely. Hence, our proof of (10.7) focuses on what happens on the event of survival.

Let $(T_n, n \in \mathbb{N})$ be any increasing sequence of stopping times. Using the strong Markov property and (10.8), we have that, for all $n \in \mathbb{N}$,

$$\mathbb{P}_{\delta_x}(\zeta < \infty) = \mathbb{E}_{\delta_x}\left[\mathbb{P}_{X_{T_n}}(\zeta < \infty)\right] \ge \mathbb{E}_{\delta_x}\left[\mathbb{P}_{X_{T_n}}(\zeta \le t_0)\right].$$

Using this inequality and Fatou's lemma, we deduce that

$$\begin{aligned}
\mathbb{P}_{\delta_x}(\zeta < \infty) &\ge \liminf_{n\to\infty} \mathbb{E}_{\delta_x}\left[\mathbb{P}_{X_{T_n}}(\zeta \le t_0)\right] \\
&\ge \mathbb{E}_{\delta_x}\left[\liminf_{n\to\infty} \mathbb{P}_{X_{T_n}}(\zeta \le t_0)\right] \\
&\ge \mathbb{E}_{\delta_x}[\mathbf{1}_{\{\zeta<\infty\}}] + \delta \mathbb{P}_{\delta_x}\left(\zeta = \infty \text{ and } \liminf_{n\to\infty} \mathbb{P}_{X_{T_n}}(\zeta \le t_0) \ge \delta\right).
\end{aligned}$$

It follows that, for all $\delta \in (0, 1]$, we have

$$\mathbb{P}_{\delta_x}\left(\zeta = \infty \text{ and } \liminf_{n\to\infty} \mathbb{P}_{X_{T_n}}(\zeta \le t_0) \ge \delta\right) = 0.$$

This implies that, on $\{\zeta = \infty\}$, $\liminf_{n\to\infty} \mathbb{P}_{X_{T_n}}(\zeta \le t_0) = 0$. Since this is true for any sequence of increasing stopping times, in particular $T_n = n$, for $n \in \mathbb{N}$, we deduce that, on $\{\zeta = \infty\}$, $\limsup_{n\to\infty} \mathbb{P}_{X_n}(\zeta \le t_0) = 0$. Together with (10.8), this gives us

$$\lim_{n\to\infty} \mathbb{P}_{X_n}(\zeta \le t_0) = \mathbf{1}_{\{\zeta<\infty\}}$$

$\mathbb{P}_{\delta_x}$-almost surely, as required.

Next we prove that for all $x \in E$, on $\{\zeta = \infty\}$, we have $\mathbb{P}_{\delta_x}$-almost surely that

$$\lim_{n\to\infty} X_n[\varphi] = \infty. \tag{10.9}$$

First note that for any $x \in E$ and $t \ge 0$, we have

$$\mathbb{P}_{\delta_x}(t < \zeta) \le \mathbb{E}_{\delta_x}[N_t] = \psi_t[1](x).$$

Using (G2), we deduce that there exists a $t_0 > 0$ such that, for all $x \in E$,

$$\mathbb{P}_{\delta_x}(t_0 < \zeta) \le \psi_{t_0}[\mathbf{1}](x) \le 2\varphi(x). \tag{10.10}$$

Thanks to the assumption (G8), we can choose t_0 sufficiently large such that there exists a constant $c_0 \in (0, 1)$ such that, uniformly for all $x \in E$,

$$\mathbb{P}_{\delta_x}(t_0 < \zeta) \le c_0 \wedge 2\varphi(x).$$

Using the branching property, we deduce that, for all $\mu = \sum_{i=1}^k \delta_{x_i} \in \mathcal{M}_c(E)$,

$$\mathbb{P}_\mu(\zeta \le t_0) \ge \prod_{i=1}^k (1 - c_0 \wedge 2\varphi(x_i)) \, .$$

Now, using (10.7), we have

$$\mathbf{1}_{\{\zeta<\infty\}} = \limsup_{n\to\infty} \mathbb{P}_{X_n}(\zeta \le t_0) \ge \limsup_{n\to\infty} \prod_{i=1}^{N_n} (1 - c_0 \wedge 2\varphi(x_i(n))) \, , \tag{10.11}$$

where $x_i(n)$ denotes the ith particle alive at time n, $i = 1, \dots, N_n$.

Next, we note that, for each $x_0 \in (0, 1]$, there exists a θ_0 such that $-\log x \le 1 - \theta_0 x$, where $\theta_0 = \theta_0(x_0) > 1$, for x bounded by unity. The limsup in (10.11) tells us that, on $\{\zeta = \infty\}$,

$$\infty = \lim_{n\to\infty} -\log \prod_{i=1}^{N_n} (1 - c_0 \wedge 2\varphi(x_i(n))) = \lim_{n\to\infty} -\sum_{i=1}^{N_n} \log (1 - c_0 \wedge 2\varphi(x_i(n))) \, . \tag{10.12}$$

As $c_0 < 1$, we can find a constant $\theta_0 = \theta_0(1 - c_0) > 1$ such that

$$-\sum_{i=1}^{N_n} \log\left(1 - c_0 \wedge 2\varphi(x_i(n))\right) \le \sum_{i=1}^{N_n}(1-\theta_0) + \theta_0 \sum_{i=1}^{N_n} c_0 \wedge 2\varphi(x_i(n))$$
$$\le 2\theta_0 \sum_{i=1}^{N_n} \varphi(x_i(n)). \qquad (10.13)$$

In conclusion, combining (10.13) with (10.12) gives us (10.9) as promised.

To conclude the proof of the lemma, let us recall that, under the assumption (G2), the eigenfunction φ having zero eigenvalue ensures that $(X_t[\varphi],\ t \ge 0)$ is a martingale. (The proof of this fact is essentially the same as in the NBP setting; see Sect. 6.1.) As $\varphi \in B^+(E)$, $W_t = X_t[\varphi]/\varphi(x) \ge CX_t[\varphi]$. Since W is almost surely convergent, (10.9) implies a contradiction and hence that $\{\zeta < \infty\}$, $\mathbb{P}_{\delta_x}$-a.s. □

10.3 Analytic Properties of the Non-linear Operator $\mathtt{A}$

Before we embark on our pursuit of coarse bounds for the survival probability, it will be important for us to understand the behaviour of the non-linear operator $\mathtt{A}$, particularly if we are to use (10.2) as the basis of our analysis. Recalling the definition (8.26), that is,

$$\mathtt{A}[h](x) = \gamma(x)\mathcal{E}_x\left[\prod_{i=1}^{N}(1 - h(x_i)) - 1 + \sum_{i=1}^{N} h(x_i)\right], \quad x \in E, h \in B_1^+(E), \qquad (10.14)$$

we may think of $\mathtt{A}[h]$ as the branching mechanism $\mathtt{G}[1-h]$ with its linearisation subtracted off. This suggests that the next largest term in the expression for $\mathtt{A}[h]$ should be its quadratic approximation. The next result allows us to control the use of this quadratic approximation in the forthcoming analysis.

Lemma 10.2 *Suppose that (G6) holds. The following statements hold:*

(i) For all $x \in E$ and $h \in B_1^+(E)$, we have

$$0 \le \mathtt{A}[h](x) \le \|\gamma\| n_{\mathtt{max}}.$$

(ii) There exists $C \in (0, \infty)$ such that, for all $h : E \to [0, 1/2]$,

$$\|\mathtt{A}[h](x) - \tfrac{1}{2}\gamma(x)\mathscr{V}[h](x)\| \le C\|h(x)\|^3. \qquad (10.15)$$

(iii) There exists $C \in (0, \infty)$ such that for all $h : E \to [0, 1/2]$,

$$\mathtt{A}[h](x) \geq C\gamma(x)\mathscr{V}[h](x) \geq 0, \qquad x \in E.$$

(iv) There exists $C \in (0, \infty)$ such that for all $h_1, h_2 \in B_1^+(E)$, we have

$$\|\mathscr{V}[h_1] - \mathscr{V}[h_2]\| \leq C\|h_1 - h_2\|.$$

Proof (i) The non-negativity of $\mathtt{A}[h]$ can be demonstrated using an iterative argument. For $n \geq 1$ and $(x_i, 1 \leq i \leq n) \in [0, 1]^n$,

$$f_n(x_1, \ldots, x_n) := \prod_{i=1}^{n}(1 - x_i) - 1 + \sum_{i=1}^{n} x_i \geq 0. \tag{10.16}$$

To see why this is true, we note that $f_1(x_1) = 0$ and, more generally, we have $f_{n+1}(x_1, \ldots, x_n, 0) = f_n(x_1, \ldots, x_n)$ as well as $\partial f_{n+1}(x_1, \ldots, x_{n+1})/\partial x_{n+1} \geq 0$. This yields $\mathtt{A}[h] \geq 0$.

For the second inequality in (i), it suffices to observe that since $h \in B_1^+(E)$,

$$\mathtt{A}[h](x) = \gamma(x)\mathcal{E}_x\left[\prod_i(1 - h(x_i)) - 1 + \sum_i h(x_i)\right] \leq \gamma(x)\mathcal{E}_x\left[\mathsf{Z}[h]\right] \leq \|\gamma\| n_{\mathtt{max}},$$

which is bounded due to (G6) and the global assumption that $\gamma \in B^+(E)$.

(ii) Let us write $\ell(h) = -\mathsf{Z}[\log(1 - h)]$, which is non-negative, and note that

$$\mathtt{A}[h](x) = \mathrm{e}^{-\ell(h)} - 1 + \mathsf{Z}[h].$$

Then, we have

$$\begin{aligned}\left|\mathtt{A}[h](x) - \tfrac{1}{2}\gamma(x)\mathscr{V}[h](x)\right| \leq{}& \gamma(x)\mathcal{E}_x\big[1 - \mathrm{e}^{-\ell(h)} - \ell(h) + \tfrac{1}{2}\ell(h)^2\big] \\ &+ \gamma(x)\mathcal{E}_x\big[\mathsf{Z}[-\log(1 - h) - h - \tfrac{1}{2}h^2]\big] \\ &+ \tfrac{1}{2}\gamma(x)\mathcal{E}_x\big[\big|\mathsf{Z}[\log(1 - h)]^2 - \mathsf{Z}[h]^2\big|\big].\end{aligned} \tag{10.17}$$

Applying the elementary bounds $0 \leq 1 - \mathrm{e}^{-x} - x + \frac{1}{2}x^2 \leq \frac{1}{6}x^3$ for all $x \geq 0$ and $0 \leq -\log(1 - x) - x - \frac{1}{2}x^2 \leq x^3$ for $x \in [0, \frac{1}{2}]$, since both $\gamma \in B^+(E)$ and (G6) hold, we see that the first two terms on the right-hand side of (10.17) are bounded by $C\|h\|^3$. For the third term on the right-hand side of (10.17), writing $\log(1 - h) = -h + R(h)$ with $|R(h)| \leq C\|h\|^2$, we get

$$\left|\mathsf{Z}[\log(1 - h)]^2 - \mathsf{Z}[h]^2\right| \leq 2\mathsf{Z}[h] \cdot \mathsf{Z}\big[|R(h)|\big] + \mathsf{Z}[R(h)]^2 \leq C'\,\|h\|^3,$$

for a constant $C' > 0$, where we have used that $h \in B_1^+(E)$ and (G6) holds. This gives us (ii).

(iii) First note that for all $n \geq 2$ and $(x_i, 1 \leq i \leq n) \in [0, \frac{1}{2}]^n$,

$$\prod_{i=1}^{n}(1 - x_i) - 1 + \sum_{i=1}^{n} x_i \ \geq \ \frac{1}{2^{n-1}} \sum_{1 \leq i, j \leq n : i \neq j} x_i x_j .$$

This can be shown with an induction on n and its proof is therefore omitted. The proof of (iii) now follows by (G6).

(iv) The final claim can be easily checked from the definition of $\mathtt{V}$, combined with (G6). □

10.4 Coarse Bounds for the Survival Probability

We will use probabilistic methods to prove a coarse lower bound for the probability of survival.

Lemma 10.3 *Suppose that (G2) holds. There exists $C \in (0, \infty)$ such that*

$$\mathtt{u}_t(x) \geq \frac{\varphi(x)}{Ct} \qquad \textit{and} \qquad a(t) \geq \frac{1}{Ct},$$

for all $t \geq 1$ and $x \in E$.

Proof Just as in the setting of the NBP (cf. Sect. 6.1), for the Asmussen–Hering class of BMPs at criticality (cf. (G2) with $\lambda_* = 0$), $X_t[\varphi]$, $t \geq 0$, is a martingale. As such, we can introduce the change of measure

$$\left.\frac{\mathrm{d}\mathbb{P}^{\varphi}_{\delta_x}}{\mathrm{d}\mathbb{P}_{\delta_x}}\right|_{\mathfrak{F}_t} = W_t := \frac{X_t[\varphi]}{\varphi(x)}, \qquad t \geq 0,\ x \in E. \tag{10.18}$$

Thanks to the change of measure (10.18) and Theorem 9.1, there exists a $C > 0$ such that

$$\sup_{x \in E} \mathbb{E}^{\varphi}_{\delta_x}[X_t[\varphi]] = \sup_{x \in E} \frac{1}{\varphi(x)} \mathbb{E}_{\delta_x}[X_t[\varphi]^2] \leq Ct,$$

for all $t \geq 1$. By Jensen's inequality, we then get

$$\mathbb{P}_{\delta_x}(\zeta > t) = \mathbb{E}^{\varphi}_{\delta_x}\left[\frac{\varphi(x)}{X_t[\varphi]}\right] \geq \frac{\varphi(x)}{\mathbb{E}^{\varphi}_{\delta_x}[X_t[\varphi]]} \geq \frac{\varphi(x)}{Ct}, \quad t \geq 1. \tag{10.19}$$

The lower bound for $a(t)$ then follows by integrating both sides of (10.19) against $\tilde{\varphi}$ and recalling that we have normalised the left and right eigenfunctions so that $\tilde{\varphi}[\varphi] = 1$. □

Next we turn to a coarse upper bound for the survival probability, which, for the most part, is an analytical proof, making use of Lemma 10.2.

Lemma 10.4 *Under the assumptions of Theorem 10.1, there exists a constant* $\widetilde{C} > 0$ *such that for all* $t \geq t_0$,

$$a(t) \leq \frac{\widetilde{C}}{t} \text{ and } \|\mathtt{u}_t\| \leq \frac{\widetilde{C}}{t}. \tag{10.20}$$

Proof We first show that

$$a(t) \to 0 \text{ and } \|\mathtt{u}_t\| \to 0, \tag{10.21}$$

as $t \to \infty$.

Thanks to Lemma 10.1, we have $\mathbb{P}_{\delta_x}(\zeta < \infty) = 1$, which implies that $\mathtt{u}_t(x) = \mathbb{P}_{\delta_x}(\zeta > t) \to 0$, as $t \to \infty$, for all $x \in E$. Clearly, $a(t) = \tilde{\varphi}[\mathtt{u}_t] \to 0$, as $t \to \infty$, by dominated convergence.

For the uniform convergence of $\mathtt{u}_t$, recalling the definition of $\mathtt{u}_t[f, g](x)$ in (8.24), we note that $\mathtt{u}_{t+s}(x) = \mathtt{u}_t[1 - \mathtt{u}_s, 0](x)$ by the Markov branching property. Recall the non-linear evolution equation in (8.24) tells us that $\mathtt{u}_t[f](x) := \mathtt{u}_t[f, 0](x)$ satisfies

$$\mathtt{u}_t[f](x) = \psi_t[1 - f](x) - \int_0^t \psi_s\left[\mathtt{A}[\mathtt{u}_{t-s}[f]]\right](x)\mathrm{d}s. \tag{10.22}$$

Setting $f = 1 - \mathtt{u}_s$ and using the preceding remarks, we find that

$$0 \leq \mathtt{u}_{t+s}(x) = \psi_t[\mathtt{u}_s](x) - \int_0^t \psi_l\big[\mathtt{A}[\mathtt{u}_{t+s-l}]\big](x)\mathrm{d}l \leq \psi_t[\mathtt{u}_s](x), \tag{10.23}$$

by Lemma 10.2 (i). Combined with the fact that we are working with the Asmussen–Herring class of MBPs, cf. (G2), this yields

$$\|\mathtt{u}_{t+s}\| \leq \|\psi_t[\mathtt{u}_s]\| \leq a(s)\|\varphi\| + O(\mathrm{e}^{-\varepsilon t}). \tag{10.24}$$

Taking first $t \to \infty$ and then $s \to \infty$ gives us that $\|\mathtt{u}_t\| \to 0$ as $t \to \infty$.

With (10.21) in hand, we can now move to the proof of the upper bound on $a(t)$ and $\|\mathtt{u}_t\|$. To this end, fix $t_0 > 0$ such that $\|\mathtt{u}_t\| \leq 1/2$ for all $t \geq t_0$. Note that the integrand in (10.2) is bounded due to our assumptions and Lemma 10.2 (i). It follows that $a(t)$ is differentiable. Differentiating (10.2) for $t \geq t_0$ and then applying the bound in Lemma 10.2 (iii), we obtain for $t \geq t_0$

$$a'(t) = -\tilde{\varphi}\big[\mathtt{A}[\mathtt{u}_t]\big] \le -C_1\tilde{\varphi}\big[\gamma\mathscr{V}[\mathtt{u}_t]\big] \le -C_2\tilde{\varphi}[\mathtt{u}_t]^2 = -C_2 a(t)^2, \tag{10.25}$$

where we have used assumption (G6) in the second inequality.

Integrating from t_0 to t yields

$$a(t) \le \Big(C_2(t-t_0) + a(t_0)^{-1}\Big)^{-1} \le (C_2 t)^{-1},$$

where the last inequality holds for t sufficiently large. The upper bound for $a(t)$ then follows. We may then apply the same techniques as in (10.24) by setting $s = t$ to obtain the uniform bound for $\mathtt{u}_t(x)$. □

10.5 Precise Survival Probability Asymptotics

The next result shows that the long-term behaviour of $\mathtt{u}_t/\varphi$ and $a(t)$ is the same, which will be key to obtaining the correct constants in the bounds obtained in the previous lemma.

Lemma 10.5 *Suppose that there exist $\kappa, \eta \in (0,\infty)$ such that $\|\mathtt{u}_t\| \le \kappa t^{-\eta}$ for all $t > 0$. Then, we can find some constant $C \in (0,\infty)$ which does not depend on κ such that*

$$\sup_{x\in E}\left|\frac{\mathtt{u}_t(x)}{\varphi(x)} - a(t)\right| \le C\kappa^2 t^{-2\eta}, \quad \textit{for all } t > 0.$$

Proof Comparing (10.22) with (10.2), we find that

$$\begin{aligned}\left|\frac{\mathtt{u}_t(x)}{\varphi(x)} - a(t)\right| &\le \left|\frac{\psi_t[\mathbf{1}](x)}{\varphi(x)} - \tilde{\varphi}[\mathbf{1}]\right| + \int_0^t \left|\frac{\psi_{t-s}\big[\mathtt{A}[\mathtt{u}_s]\big](x)}{\varphi(x)} - \tilde{\varphi}\big[\mathtt{A}[\mathtt{u}_s]\big]\right| \mathrm{d}s \\ &\le C_1\mathrm{e}^{-\varepsilon t} + \int_0^t C_1\mathrm{e}^{-\varepsilon(t-s)}\|\mathtt{A}[\mathtt{u}_s]\|\,\mathrm{d}s,\end{aligned}$$

where the constant $\varepsilon > 0$ comes from (G2). Thanks to Lemma 10.4, we can find $t_0 > 0$ such that $\sup_{x\in E}\sup_{t\ge t_0}\mathtt{u}_t(x) \le 1/2$. Take $t \ge 2t_0$, then the integral above can be bounded as follows:

$$\begin{aligned}&\int_0^t C_1\mathrm{e}^{-\varepsilon(t-s)}\|\mathtt{A}[\mathtt{u}_s]\|\,\mathrm{d}s \\ &\qquad = \int_0^{t/2} C_1\mathrm{e}^{-\varepsilon(t-s)}\|\mathtt{A}[\mathtt{u}_s]\|\,\mathrm{d}s + \int_{t/2}^t C_1\mathrm{e}^{-\varepsilon(t-s)}\|\mathtt{A}[\mathtt{u}_s]\|\,\mathrm{d}s\end{aligned}$$

$$\leq \frac{C_1 \|\gamma\| n_{\max}}{\varepsilon} \mathrm{e}^{-\varepsilon t/2} + \int_{t/2}^{t} C_1 \mathrm{e}^{-\varepsilon(t-s)} \|\mathtt{A}[\mathtt{u}_s]\| \,\mathrm{d}s. \tag{10.26}$$

Using the easy observation $\mathscr{V}[\mathtt{u}_s] \leq n_{\max}^2 \|\mathtt{u}_s\|^2$ together with Lemma 10.2 (ii), we deduce for the second term in (10.26) that $\|\mathtt{A}[\mathtt{u}_s]\| \leq \|\mathscr{V}[\mathtt{u}_s]\| + \|\mathtt{u}_s^3\| \leq C_2 \|\mathtt{u}_s\|^2$. The latter is bounded by $C_2 \kappa^2 s^{-2\eta}$ due to the assumption of the lemma. Therefore,

$$\begin{aligned}
&\int_{t/2}^{t} C_1 \mathrm{e}^{-\varepsilon(t-s)} \|\mathtt{A}[\mathtt{u}_s]\| \,\mathrm{d}s \\
&\quad \leq C_3 \kappa^2 \int_{t/2}^{t} \mathrm{e}^{-\varepsilon(t-s)} s^{-2\eta} \mathrm{d}s \\
&\quad = \frac{C_3 \kappa^2}{\varepsilon} \left(t^{-2\eta} - 4\mathrm{e}^{-\varepsilon t/2} t^{-2\eta} \right) + C_4 \int_{t/2}^{t} \mathrm{e}^{-\varepsilon(t-s)} s^{-2\eta-1} \mathrm{d}s = O(t^{-2\eta}).
\end{aligned}$$

Putting the pieces together, we obtain the claimed bound in the lemma. □

***Proof** (of Theorem 10.1)* Applying Lemma 10.5 with $\eta = 1$ and κ being some positive constant, which is permitted thanks to Lemma 10.4, we have

$$\sup_{x \in E} \left| \frac{\mathtt{u}_t(x)}{\varphi(x)} - a(t) \right| = O(t^{-2}), \quad t \to \infty.$$

On the other hand, we have seen in Lemma 10.3 that $a(t)^{-1} = O(t)$. It follows that

$$\sup_{x \in E} \left| \frac{\mathtt{u}_t(x)}{\varphi(x) a(t)} - 1 \right| = O(t^{-1}), \quad t \to \infty. \tag{10.27}$$

Applying Lemma 10.2 (iv), (10.27), and Lemma 10.4, we deduce that

$$\begin{aligned}
\sup_{x \in E} \left| \mathscr{V}[\mathtt{u}_t](x) - a(t)^2 \mathscr{V}[\varphi](x) \right| &= \sup_{x \in E} a(t)^2 \left| \mathscr{V}\left[\frac{\mathtt{u}_t}{a(t)} \right](x) - \mathscr{V}[\varphi](x) \right| \\
&\leq C\, a(t)^2 \sup_{x \in E} \left| \frac{\mathtt{u}_t(x)}{a(t)} - \varphi(x) \right| = O(t^{-3}).
\end{aligned} \tag{10.28}$$

We now see that, for all $t \geq t_0$,

$$\begin{aligned}
a(t) - a(t_0) &= -\int_{t_0}^{t} \tilde{\varphi}[\mathtt{A}[\mathtt{u}_s]] \mathrm{d}s \\
&= -\frac{1}{2} \int_{t_0}^{t} \left(\tilde{\varphi}[\gamma \mathscr{V}[\mathtt{u}_s]] + O(\|\mathtt{u}_s\|^3) \right) \mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{2}\int_{t_0}^{t} \Big(\tilde{\varphi}\big[\gamma\mathscr{V}[\mathrm{u}_s]\big] + O(s^{-3})\Big)\mathrm{d}s \\
&= -\frac{1}{2}\int_{t_0}^{t} \Big(a(s)^2\tilde{\varphi}\big[\gamma\mathscr{V}[\varphi],\big] + O(s^{-3})\Big)\mathrm{d}s \\
&= -\frac{\Sigma}{2}\int_{t_0}^{t} a(s)^2\big(1+o(1)\big)\,\mathrm{d}s,
\end{aligned}
$$

where we have used (10.15) in the second equality, (10.20) in the third, (10.28) in the fourth, and Lemma 10.3 in the final equality. This implies that

$$
a(t) \sim \frac{2}{\Sigma t} \text{ as } t \to \infty.
$$

The desired asymptotic for u_t then follows from (10.27). □

10.6 Remarks on the Neutron Transport Setting

Comparing the statement of Theorem 5.8 with Theorem 10.1, there are seemingly different assumptions at play.

Assumption (H5) vs Assumption (G7) In the setting of neutron transport, the assumption (H5) is weaker than (G7) in the case where E is bounded. To compensate for this, the proof of Theorem 5.8 requires some more work.

The idea in this case is to use the same techniques as in Lemma 10.4 to obtain coarse upper and lower bounds but of order $1/\sqrt{t}$. The key change occurs in (10.25). Specifically, in that setting,

$$
\begin{aligned}
a'(t) = \tilde{\varphi}\big[\mathtt{A}[\mathrm{u}_t]\big] &\le -C_1\tilde{\varphi}\big[\sigma_{\mathtt{f}}\mathscr{V}[\mathrm{u}_t]\big] \\
&\le -C_2\int_{D\times V}\tilde{\varphi}(r,\upsilon)\int_V\int_V \mathrm{u}_t(r,\upsilon_1)\mathrm{u}_t(r,\upsilon_2)\mathrm{d}r\,\mathrm{d}\upsilon\,\mathrm{d}\upsilon_1\mathrm{d}\upsilon_2,
\end{aligned}
\tag{10.29}
$$

where we have used assumption (H5) in the second inequality. Note that Hölder's inequality implies that

$$
\left(\int_D f(r)\mathrm{d}r\right)^3 \le C_3\int_D |f(r)|^3\mathrm{d}r.
$$

Applying this to $r \mapsto \int_V \tilde{\varphi}(r,\upsilon)\mathrm{u}_t(r,\upsilon)\mathrm{d}\upsilon$, we find that

$$
\begin{aligned}
a(t)^3 &= \left(\int_{D\times V} \tilde{\varphi}(r,\upsilon)\mathrm{u}_t(r,\upsilon)\mathrm{d}r\mathrm{d}\upsilon\right)^3 \leq C_3 \int_D \left(\int_V \tilde{\varphi}(r,\upsilon)\mathrm{u}_t(r,\upsilon)\mathrm{d}\upsilon\right)^3 \mathrm{d}r \\
&= C_3 \int_D \mathrm{d}r \int_{V\times V\times V} \tilde{\varphi}(r,\upsilon)\tilde{\varphi}(r,\upsilon_1)\tilde{\varphi}(r,\upsilon_2)\mathrm{u}_t(r,\upsilon)\mathrm{u}_t(r,\upsilon_1)\mathrm{u}_t(r,\upsilon_2) \\
&\quad \times \mathrm{d}\upsilon\,\mathrm{d}\upsilon_1\mathrm{d}\upsilon_2 \\
&\leq C_4\|\mathrm{u}_t\| \int_{D\times V} \tilde{\varphi}(r,\upsilon) \int_V \int_V \mathrm{u}_t(r,\upsilon_1)\mathrm{u}_t(r,\upsilon_2)\mathrm{d}r\,\mathrm{d}\upsilon\,\mathrm{d}\upsilon_1\mathrm{d}\upsilon_2,
\end{aligned}
$$

since $\tilde{\varphi}$ is uniformly bounded. Comparing this with (10.29), we have

$$
a'(t) \leq -\frac{C_5}{\|\mathrm{u}_t\|} a(t)^3, \quad t \geq t_0.
$$

Specifically, in that setting,

$$
a'(t) = -\tilde{\varphi}\big[\mathtt{A}[\mathrm{u}_t]\big] \leq -C_1\tilde{\varphi}\big[\sigma_{\mathtt{f}}\mathscr{V}[\mathrm{u}_t]\big] \leq -C_2\tilde{\varphi}\left[\int_V \mathrm{u}_t(\cdot,\upsilon')\mathrm{d}\upsilon'\right]^2 = -C_2 a(t)^2.
$$

At this point in the argument, we know that $\|\mathrm{u}_t\| \to 0$. Therefore, for any fixed $\epsilon > 0$, there exists $t_0' = t_0'(\epsilon)$ such that

$$
a'(t) \leq -\frac{C_5}{\epsilon^2} a(t)^3, \quad t \geq t_1 := \max(t_0, t_0').
$$

Integrating from t_1 to t yields

$$
a(t) \leq \left(\frac{2C_5}{\epsilon^2}(t-t_1) + a(t_1)^{-2}\right)^{-1/2} \leq \left(\frac{C_5}{\epsilon^2} t\right)^{-1/2},
$$

where the last inequality holds for t sufficiently large. The upper bound for $a(t)$ then follows. We may then apply the same techniques as in (10.24) by setting $s = t$ to obtain the bound for $\|\mathrm{u}_t\|$.

Formally speaking, we have shown that, under the assumptions of Theorem 5.8, there exists a constant $C > 0$ such that, for every $\epsilon > 0$, we can find $t_0 = t_0(\epsilon)$ with

$$
a(t) \leq \frac{\epsilon}{C\sqrt{t}} \text{ and } \|\mathrm{u}_t\| \leq \frac{\epsilon}{C\sqrt{t}}, \tag{10.30}
$$

for all $t \geq t_0$.

Once in the possession of an upper bound for $a(t)$ and $\|\mathrm{u}_t\|$ of $O(t^{-1/2})$, we can bootstrap this further to improve the upper bound to $O(t^{-1})$. Indeed, from Lemma 10.4, we see that the assumption of Lemma 10.5 holds for $\kappa = \epsilon$, where ϵ is given in (10.30), and $\eta = 1/2$. Since ϵ can be taken arbitrarily small, Lemma 10.5 then tells us that $\|(\mathrm{u}_t/\varphi) - a(t)\| = o(t^{-1})$. Combined with Lemma 10.3, this implies that

$$\sup_{r\in D, \upsilon\in V}\left|\frac{\mathtt{u}_t(r,\upsilon)}{\varphi(r,\upsilon)a(t)}-1\right|\to 0,\ \text{as } t\to\infty.$$

Thanks to Lemma 10.2 (iv), we have

$$\|\mathscr{V}[\mathtt{u}_t/a(t)]-\mathscr{V}[\varphi]\|\le C\|(\mathtt{u}_t/a(t))-\varphi\|\to 0. \tag{10.31}$$

On the other hand, as in (10.25), for t sufficiently large, we have

$$a'(t)\le -C_1\tilde{\varphi}\big[\mathscr{V}[\mathtt{u}_t]\big]=-C_1a(t)^2\tilde{\varphi}\big[\mathscr{V}[\mathtt{u}_t/a(t)]\big]\le -C_2a(t)^2,$$

by (10.31). Integrating over t yields the desired bound for $a(t)$. The uniform bound for $\mathtt{u}_t$ follows from the same arguments as previously.

This brings us back to (10.25), and we can continue the reasoning as in the proof of the general case.

The Absence of Assumption (G8) Heuristically, the condition (G8) for NBPs is satisfied because a neutron released from anywhere in $D\times V$ will exit the domain D without undergoing fission or scattering with a minimal probability, implying extinction.

To be more precise, suppose we define $d_0 = 2\inf\{r > 0 : D\subseteq B_a(r) \text{ for some } r \in D\}$, where $B_a(r)$ is a ball of radius a in $\mathbb{R}^3$ centred at r. We can think of d_0 as the "diameter" of the physical domain D. Now note that, for $r\in D$, $\upsilon\in V$,

$$\mathbb{P}_{\delta_{(r,\upsilon)}}(\zeta\le\kappa^D_{r,\upsilon})\ge\exp\left(-\int_0^{\kappa^D_{r,\upsilon}}\sigma(r+\upsilon s,\upsilon)\mathrm{d}s\right)\ge \mathrm{e}^{-\kappa^D_{r,\upsilon}\overline{\sigma}},$$

where we recall that $\sigma=\sigma_{\mathtt{s}}+\sigma_{\mathtt{f}}$ is the sum of the scatter and fission rates and $\overline{\sigma}=\sup_{r\in D,\upsilon\in V}\sigma(r,\upsilon)$ as per (H1), which was assumed in Theorem 5.8. This means that for $r\in D$, $\upsilon\in V$,

$$\sup_{r\in D,\upsilon\in V}\mathbb{P}_{\delta_{(r,\upsilon)}}(\zeta>\kappa^D_{r,\upsilon})\le\sup_{r\in D,\upsilon\in V}\left(1-\mathrm{e}^{-\kappa^D_{r,\upsilon}\overline{\sigma}}\right)\le\left(1-\mathrm{e}^{-d_0\overline{\sigma}/\upsilon_{\mathtt{min}}}\right),$$

where we have used that $\kappa^D_{r,\upsilon}\upsilon_{\mathtt{min}}\le\kappa^D_{r,\upsilon}|\upsilon|\le d_0$. In other words, for all $t>t_0 := d_0/\upsilon_{\mathtt{min}}$,

$$\sup_{r\in D,\upsilon\in V}\mathbb{P}_{\delta_{(r,\upsilon)}}(\zeta>t)\le\sup_{r\in D,\upsilon\in V}\mathbb{P}_{\delta_{(r,\upsilon)}}(\zeta>\kappa^D_{r,\upsilon})\le\left(1-\mathrm{e}^{-d_0\overline{\sigma}/\upsilon_{\mathtt{min}}}\right)<1.$$

10.7 Comments

The asymptotic of the survival probability at criticality for Bienaymé–Galton–Watson processes is a classical result due to Kolmogorov [85], which has since seen many improvements and generalisations to other spatial branching processes. Examples include Crump–Mode–Jagers processes [46], branching Brownian motion [112], and superprocesses [113]. The Kolmogorov survival estimate is an important part of the Yaglom limit, which we presented in the setting of the NBP in Chap. 5 and which was originally formulated for Bienaymé–Galton–Watson processes in [132]. In the setting of isotropic neutron transport, the analogue of the Kolmogorov survival probability asymptotic was first proved in [99]. The main result we present in this chapter, Theorem 10.1, which deals with a more general setting, is reproduced from Harris et al. [72].

We finish this chapter with a short comment regarding the assumption (G6). While this assumption is far from satisfactory, we have been unable to weaken it to a moment assumption on the offspring distribution, for example. However, this assumption is clearly satisfied in the setting of the NBP and was also a necessary assumption in the isotropic case in [99].

Chapter 11
Spines and Skeletons

We have seen in Chap. 6 that a natural way to study the long-term behaviour of the NBP is via spine and skeletal decompositions. As alluded to in Chap. 6, these decompositions can be proved in the setting of the general BMP introduced in Chap. 8. Continuing at this level of generality, we look at the formal proofs of analogues of Lemma 6.2 and Theorem 6.4 for the spine decomposition and Theorem 6.5 for the skeletal decomposition. We also take the opportunity to discuss how the spine decomposition emerges from the skeletal decomposition as a natural consequence of conditioning on survival at criticality, as well as how it explains the particular shape of the limiting moment asymptotics in Theorem 9.1.

11.1 Spine Decomposition

Recall that, for our BMP, $(X, \mathbb{P})$, the underlying Markov process is denoted by $(\xi, \mathbf{P})$ and that $(\hat{\xi}, \hat{\mathbf{P}})$ is the law of $(\xi, \mathbf{P})$ with the additional jumps that appear in the many-to-one formula, cf. Lemma 8.2. We will further assume that assumption (G2), introduced in Sect. 8.5, is in force, which gives us the existence of a lead eigenvalue λ_*, with associated left eigenmeasure $\tilde{\varphi}$ and right eigenfunction φ.

We have, in particular, that

$$W_t := \mathrm{e}^{-\lambda_* t} \frac{X_t[\varphi]}{\mu[\varphi]} \tag{11.1}$$

is a unit mean $\mathbb{P}_\mu$-martingale, where $\mu \in \mathscr{M}_c(E)$. We are interested in the change of measure

$$\left.\frac{\mathrm{d}\mathbb{P}^\varphi_\mu}{\mathrm{d}\mathbb{P}_\mu}\right|_{\mathfrak{F}_t} = W_t, \qquad t \geq 0, \mu \in \mathscr{M}_c(E). \tag{11.2}$$

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8_11

In the next theorem, we will formalise an understanding of this change of measure in terms of another $\mathcal{M}_c(E)$-valued stochastic process $X^\varphi := (X_t^\varphi, t \geq 0)$, which we will now describe through an algorithmic construction:

1. From the initial configuration $\mu \in \mathcal{M}_c(E)$ with an arbitrary enumeration of particles so that $\mu = \sum_{i=1}^n \delta_{x_i}$, the i-th particle is selected and marked "*spine*" with empirical probability
$$\frac{\varphi(x_i)}{\mu[\varphi]}. \tag{11.3}$$
2. Each unmarked particle $j \neq i$ issues an independent copy of $(X, \mathbb{P}_{\delta_{x_j}})$.
3. For the marked particle, issue a copy of the process whose motion is determined by the semigroup
$$\mathtt{P}_t^\varphi[f](x) := \frac{1}{\varphi(x)} \mathbf{E}_x \left[\mathrm{e}^{-\lambda_* t} \mathrm{e}^{\int_0^t \frac{\gamma(\xi_s)}{\varphi(\xi_s)} (\mathtt{m}[\varphi](\xi_s) - \varphi(\xi_s)) \mathrm{d}s} \varphi(\xi_t) f(\xi_t) \right], \tag{11.4}$$
for $x \in E$ and $f \in B^+(E)$.
4. The marked particle undergoes branching at rate
$$\gamma^\varphi(x) := \gamma(x) \frac{\mathtt{m}[\varphi](x)}{\varphi(x)}, \tag{11.5}$$
when at $x \in E$, at which point, it scatters a random number of particles according to the random measure on E given by $(\mathsf{Z}, \mathcal{P}_x^\varphi)$ where
$$\frac{\mathrm{d}\mathcal{P}_x^\varphi}{\mathrm{d}\mathcal{P}_x} = \frac{\mathsf{Z}[\varphi]}{\mathtt{m}[\varphi](x)}. \tag{11.6}$$
Here, we recall $\mathtt{m}[\varphi](x) = \mathcal{E}_x[\mathsf{Z}[\varphi]]$.
5. When the marked particle is at $x \in E$, given the realisation of $(\mathsf{Z}, \mathcal{P}_x^\varphi)$, set $\mu = \mathsf{Z}$ and repeat Step 1.

The process X_t^φ describes the position of all the particles in the system at time $t \geq 0$ (ignoring the marked genealogy). We will also be interested in the configuration of the single genealogical line of descent, which has been marked "*spine*", identified by Steps 1 and 2 above. This process, referred to simply as *the spine*, will be denoted by $\xi^\varphi := (\xi_t^\varphi, t \geq 0)$. Together, the processes (X^φ, ξ^φ) make a Markov pair, whose probabilities we will denote by $(\tilde{\mathbb{P}}_{\mu,x}^\varphi, \mu \in \mathcal{M}_c(E), x \in \mathrm{supp}(\mu))$.

To see the associated Markov property, suppose we are given the pair $(X_t^\varphi, \xi_t^\varphi)$, for $t \geq 0$, then according to steps 2–5 of the algorithm above, to describe the configuration of the pair $(X_{t+s}^\varphi, \xi_{t+s}^\varphi)$, for $s > 0$, it suffices to evolve from each particle in X_t^φ that is not part of the spinal process ξ^φ, an independent copy of $(X, \mathbb{P})$ for s units of time, and to evolve from the initial position ξ_t^φ, an independent copy of ξ^φ, which then follows steps 3–5 above for s units of time.

The way in which X^φ was described algorithmically above, with the spine randomly selected from the initial configuration of particles $\mu \in \mathscr{M}_c(E)$, it has law

$$\tilde{\mathbb{P}}^\varphi_\mu := \sum_{i=1}^n \frac{\varphi(x_i)}{\mu[\varphi]} \tilde{\mathbb{P}}^\varphi_{\mu,x_i} = \frac{1}{\mu[\varphi]} \int_E \varphi(x)\mu(\mathrm{d}x)\tilde{\mathbb{P}}^\varphi_{\mu,x},$$

when $\mu = \sum_{i=1}^n \delta_{x_i}$. Write for convenience $\tilde{\mathbb{P}}^\varphi = (\tilde{\mathbb{P}}^\varphi_\mu, \mu \in \mathscr{M}_c(E))$. The next result gives us the marginal law of the process X^φ under $\tilde{\mathbb{P}}^\varphi$.

Theorem 11.1 *Under assumption (G2), the process $(X^\varphi, \tilde{\mathbb{P}}^\varphi_\mu)$ is Markovian and equal in law to $(X, \mathbb{P}^\varphi_\mu)$, for $\mu \in \mathscr{M}_c(E)$.*

We would also like to understand the dynamics of the spine ξ^φ. For convenience, let us denote the family of probabilities of the latter by $\tilde{\mathbf{P}}^\varphi = (\tilde{\mathbf{P}}^\varphi_x, x \in E)$, in other words, the marginals of $(\tilde{\mathbb{P}}^\varphi_{\mu,x}, \mu \in \mathscr{M}_c(E), x \in E)$.

Lemma 11.1 *Under assumption (G2), the process $(\xi^\varphi, \tilde{\mathbf{P}}^\varphi)$ is equal in law to $(\hat{\xi}, \hat{\mathbf{P}}^\varphi)$, where*

$$\left.\frac{\mathrm{d}\hat{\mathbf{P}}^\varphi_x}{\mathrm{d}\hat{\mathbf{P}}_x}\right|_{\mathfrak{F}_t} = \mathrm{e}^{-\lambda_* t + \int_0^t \mathtt{B}(\hat{\xi}_s)\mathrm{d}s} \frac{\varphi(\hat{\xi}_t)}{\varphi(x)}, \qquad t \geq 0, x \in E, \tag{11.7}$$

and we recall that $\mathtt{B}(x) = \gamma(x)(\mathtt{m}[1](x) - 1)$ and the process $(\hat{\xi}, \hat{\mathbf{P}})$ is described above Lemma 8.2. Equivalently, the process $(\xi^\varphi, \tilde{\mathbf{P}}^\varphi)$ has semigroup $(\tilde{\mathtt{P}}^\varphi_t, t \geq 0)$, which satisfies

$$\tilde{\mathtt{P}}^\varphi_t[g](x) = \hat{\mathbf{P}}_x\left[\mathrm{e}^{-\lambda_* t + \int_0^t \mathtt{B}(\hat{\xi}_s)\mathrm{d}s} \frac{\varphi(\hat{\xi}_t)}{\varphi(x)} g(\hat{\xi}_t)\right], \qquad x \in E, g \in B^+(E), t \geq 0. \tag{11.8}$$

From this conclusion, we deduce that $(\xi^\varphi, \tilde{\mathbf{P}}^\varphi)$ is conservative with a limiting stationary distribution $\varphi(x)\tilde{\varphi}(\mathrm{d}x)$, $x \in E$.

***Proof** (of Theorem 11.1)* There are three main steps to the proof. The first is to characterise the law of transitions of the Markov process $(X, \mathbb{P}^\varphi)$, defined in the change of measure (11.2). The second step is to show that they agree with those of $(X^\varphi, \tilde{\mathbb{P}}^\varphi)$. The third step is to show that $(X^\varphi, \tilde{\mathbb{P}}^\varphi)$ is Markovian. Together these three imply the statement of the theorem.

Step 1 First we look at the multiplicative semigroup that characterises uniquely the transitions of $(X^\varphi, \mathbb{P}^\varphi)$ (cf. [75–77])

$$\mathrm{v}_t^{\varphi}[g](x) := \mathbb{E}_{\delta_x}^{\varphi}\left[\mathrm{e}^{-X_t[g]}\right] = \mathbb{E}_{\delta_x}\left[\mathrm{e}^{-\lambda_* t}\frac{X_t[\varphi]}{\varphi(x)}\mathrm{e}^{-X_t[g]}\right], \tag{11.9}$$

for $t \geq 0$ and $g \in B^+(E)$, where $X_t = \sum_{i=1}^{N_t}\delta_{x_i(t)}$. As we have seen before, we can extend the domain of test functions to include the cemetery state †. For this, we need to insist on the default value $g(\dagger) = 0$ so that $\mathrm{e}^{-g}(\dagger) = 1$.

We start in the usual way by splitting the expectation in the second equality of (11.9) according to whether a branching event has occurred by time t or not.

$$\begin{aligned}\mathrm{v}_t^{\varphi}[g](x) &= \mathbf{E}_x\left[\mathrm{e}^{-g(\xi_t)}\frac{\varphi(\xi_t)}{\varphi(x)}\mathrm{e}^{-\int_0^t \lambda_* + \gamma(\xi_s)\mathrm{d}s}\right] \\ &\quad + \mathbf{E}_x\Bigg[\int_0^t \gamma(\xi_s)\mathrm{e}^{-\int_0^s \lambda_*+\gamma(\xi_u)\mathrm{d}u}\frac{\varphi(\xi_s)}{\varphi(x)} \\ &\qquad \mathcal{E}_{\xi_s}\Big[\sum_{i=1}^N \frac{\varphi(x_i)}{\varphi(\xi_s)}\mathbb{E}\Big[W_{t-s}^i\prod_{j=1}^N \mathrm{e}^{-X_{t-s}^j[g]}\Big|(x_i, i = 1, \cdots, N)\Big]\Big]\mathrm{d}s\Bigg],\end{aligned} \tag{11.10}$$

where, given the offspring positions $(x_i, i = 1, \cdots, N)$, (W^i, X^i) are independent copies of the pair (W, X) under $\mathbb{P}_{\delta_{x_i}}$. We first focus on developing the branching operator in the integral on the right-hand side of (11.10).

Recalling the definition of v_t from (8.7), we have that for all $x \in E$,

$$\begin{aligned}&\gamma(x)\mathcal{E}_x\left[\sum_{i=1}^N \frac{\varphi(x_i)}{\varphi(x)}\mathbb{E}\Big[W_{t-s}^i\prod_{j=1}^N \mathrm{e}^{-X_{t-s}^j[g]}\Big|(x_i, i = 1\cdots, N)\Big]\right] \\ &= \gamma(x)\mathcal{E}_x\left[\frac{\mathsf{Z}[\varphi]}{\varphi(x)}\sum_{i=1}^N\frac{\varphi(x_i)}{\mathsf{Z}[\varphi]}\mathbb{E}_{\delta_{x_i}}\left[W_{t-s}^i\mathrm{e}^{-X_{t-s}^i[g]}\right]\prod_{\substack{j=1\\ i\neq j}}^N \mathbb{E}_{\delta_{x_j}}\left[\mathrm{e}^{-X_{t-s}^j[g]}\right]\right] \\ &= \gamma(x)\frac{\mathtt{m}[\varphi](x)}{\varphi(x)}\mathcal{E}_x^{\varphi}\left[\sum_{i=1}^N\frac{\varphi(x_i)}{\mathsf{Z}[\varphi]}\mathrm{v}_{t-s}^{\varphi}[g](x_i)\prod_{\substack{j=1\\ i\neq j}}^N \mathrm{v}_{t-s}[g](x_j)\right].\end{aligned} \tag{11.11}$$

Now returning to (11.10) with (11.11) in hand, we have

$$\begin{aligned}\mathrm{v}_t^{\varphi}[g](x) &= \mathbf{E}_x\left[\mathrm{e}^{-g(\xi_t)}\frac{\varphi(\xi_t)}{\varphi(x)}\mathrm{e}^{-\int_0^t \lambda_*+\gamma(\xi_s)\mathrm{d}s}\right] \\ &\quad + \mathbf{E}_x\Bigg[\int_0^t \mathrm{e}^{-\int_0^s \lambda_*+\gamma(\xi_u)\mathrm{d}u}\frac{\varphi(\xi_s)}{\varphi(x)}\end{aligned}$$

$$\left.\gamma(\xi_s)\frac{\mathtt{m}[\varphi](\xi_s)}{\varphi(\xi_s)}\mathscr{E}^{\varphi}_{\xi_s}\left[\sum_{i=1}^{N}\frac{\varphi(x_i)}{\mathsf{Z}[\varphi]}\mathtt{v}^{\varphi}_{t-s}[g](x_i)\prod_{\substack{j=1\\ i\neq j}}^{N}\mathtt{v}_{t-s}[g](x_j)\right]\mathrm{d}s\right]$$

$$-\mathbf{E}_x\left[\int_0^t \gamma(\xi_s)\frac{\mathtt{m}[\varphi](\xi_s)}{\varphi(\xi_s)}\mathtt{v}_{t-s}[g](\xi_s)\mathrm{d}s\right]$$

$$+\mathbf{E}_x\left[\int_0^t \gamma(\xi_s)\frac{\mathtt{m}[\varphi](\xi_s)}{\varphi(\xi_s)}\mathtt{v}_{t-s}[g](\xi_s)\mathrm{d}s\right],$$

where we have artificially introduced the final terms. Applying Dynkin's lemma in reverse to change the final term into a multiplicative potential yields

$$\mathtt{v}^{\varphi}_t[g](x)=\mathbf{E}_x\left[\mathrm{e}^{-g(\xi_t)}\frac{\varphi(\xi_t)}{\varphi(x)}\mathrm{e}^{-\int_0^t \lambda_*-\frac{\gamma(\xi_s)}{\varphi(\xi_s)}(\mathtt{m}[\varphi](\xi_s)-\varphi(\xi_s))\mathrm{d}s}\right]$$

$$+\mathbf{E}_x\left[\int_0^t \mathrm{e}^{-\int_0^s \lambda_*-\frac{\gamma(\xi_u)}{\varphi(\xi_u)}(\mathtt{m}[\varphi](\xi_u)-\varphi(\xi_u))\mathrm{d}u}\frac{\varphi(\xi_s)}{\varphi(x)}\right.$$

$$\left.\gamma(\xi_s)\frac{\mathtt{m}[\varphi](\xi_s)}{\varphi(\xi_s)}\mathscr{E}^{\varphi}_{\xi_s}\left[\sum_{i=1}^{N}\frac{\varphi(x_i)}{\mathsf{Z}[\varphi]}\mathtt{v}^{\varphi}_{t-s}[g](x_i)\prod_{\substack{j=1\\ i\neq j}}^{N}\mathtt{v}_{t-s}[g](x_j)\right]\mathrm{d}s\right]$$

$$-\mathbf{E}_x\left[\int_0^t \gamma(\xi_s)\frac{\mathtt{m}[\varphi](\xi_s)}{\varphi(\xi_s)}\mathtt{v}_{t-s}[g](\xi_s)\mathrm{d}s\right]. \tag{11.12}$$

Step 2 Define

$$\tilde{\mathtt{v}}^{\varphi}_t[g](x)=\tilde{\mathbb{E}}^{\varphi}_{\delta_x}\left[\mathrm{e}^{-X^{\varphi}_t[g]}\right],\qquad t\geq 0, \tag{11.13}$$

for $g\in B^+(E)$, where, again, $(x_i(t), i=1\cdots,N_t)$, is the configuration of the population at time $t>0$.

By conditioning $\tilde{\mathtt{v}}^{\varphi}_t$ on the first branch time, it is a straightforward exercise to show that it also solves (11.12). For the sake of brevity, we leave this as an exercise to the reader as the arguments are similar to those appearing earlier. In order to show that (11.12) has a unique solution, we consider $u^{\varphi}_t[g]:=\varphi\mathtt{v}^{\varphi}_t[g]$ and $\tilde{u}^{\varphi}_t[g]:=\varphi\tilde{\mathtt{v}}^{\varphi}_t[g]$. Again, standard arguments show that they both satisfy the same equation. Using boundedness of the branch rate, of the mean offspring, and of φ, by considering the difference $\sup_{x\in E}|u^{\varphi}_t[g](x)-\tilde{u}^{\varphi}_t[g](x)|$, uniqueness of solutions to (11.12) follows from Grönwall's inequality. Again, we leave the details as an exercise to the reader in order to avoid repetition.

Step 3 The joint process $(X^{\varphi},\xi^{\varphi})$ is, by construction, Markovian under $\tilde{\mathbb{P}}^{\varphi}$. We thus need to show that X^{φ} alone demonstrates the Markov property. We do this by showing that for $f\in B^+(E)$ and $\mu\in\mathscr{M}_c(E)$,

$$\tilde{\mathbb{E}}^{\varphi}_{\mu}\left[f(\xi^{\varphi}_t)|X^{\varphi}_t\right] = \frac{X^{\varphi}_t[f\varphi]}{X^{\varphi}_t[\varphi]}, \qquad t \geq 0. \tag{11.14}$$

This says that knowing X^{φ}_t allows one to construct the law of ξ^{φ}_t by using X^{φ}_t as an empirical distribution with an additional density φ. With (11.14) in hand, the desired Markov property follows as, for $g \in B^+(E)$ and $\mu \in \mathscr{M}_c(E)$,

$$\begin{aligned}\tilde{\mathbb{E}}^{\varphi}_{\mu}\left[\mathrm{e}^{-X^{\varphi}_{t+s}[g]}|\mathfrak{F}_t\right] &= \sum_{i=1}^{N_t}\frac{\varphi(x_i(t))}{X^{\varphi}_t[\varphi]}\tilde{\mathbb{E}}^{\varphi}_{\mu',x'}\left[\mathrm{e}^{-X^{\varphi}_s[g]}\right]_{\mu'=X^{\varphi}_t,x'=x_i(t)}\\ &= \tilde{\mathbb{E}}^{\varphi}_{\mu'}\left[\mathrm{e}^{-X^{\varphi}_s[g]}\right]_{\mu'=X^{\varphi}_t},\end{aligned}$$

where we have written $X^{\varphi}_t = \sum_{i=1}^{N_t}\delta_{x_i(t)}$.

We are thus left with proving (11.14) to complete this step. To do so, we note that it suffices to show that for $f \in B^+_1(E)$, $g \in B^+(E)$, $\mu \in \mathscr{M}_c(E)$, and $x \in E$,

$$\tilde{\mathbb{E}}^{\varphi}_{\mu}\left[f(\xi^{\varphi}_t)\mathrm{e}^{-X^{\varphi}_t[g]}\right] = \tilde{\mathbb{E}}^{\varphi}_{\mu}\left[\frac{X^{\varphi}_t[f\varphi]}{X^{\varphi}_t[\varphi]}\mathrm{e}^{-X^{\varphi}_t[g]}\right], \qquad t \geq 0. \tag{11.15}$$

On the left-hand side of (11.15), we have

$$\begin{aligned}&\tilde{\mathbb{E}}^{\varphi}_{\mu}\left[f(\xi^{\varphi}_t)\mathrm{e}^{-X^{\varphi}_t[g]}\right]\\ &= \tilde{\mathbb{E}}^{\varphi}_{\mu}\left[\tilde{\mathbb{E}}^{\varphi}_{\mu}\left[f(\xi^{\varphi}_t)\mathrm{e}^{-X^{\varphi}_t[g]}|\xi^{\varphi}_t\right]\right]\\ &= \sum_{k=1}^{n}\frac{\varphi(x_k)}{\mu[\varphi]}\tilde{\mathbb{E}}^{\varphi}_{\delta_{x_k}}\left[f(\xi^{\varphi}_t)\prod_{i\geq 1:T_i\leq t}\prod_{j=1}^{N_i}\frac{\mathrm{v}_{t-T_i}[g](x_{ij})}{\mathrm{v}_{t-T_i}[g](\xi^{\varphi}_{T_i})}\right]\prod_{\ell\neq k}\mathrm{v}_t[g](x_\ell),\end{aligned}$$

where $\mu = \sum_{k=1}^{n}\delta_{x_i}$ and $(T_i, i \geq 1)$ are the times of fission along the spine at which point, N_i particles are issued at x_{ij}, $j = 1\cdots, N_i$, and we recall $(\mathrm{v}_s, s \geq 0)$ is the non-linear semigroup of $(X, \mathbb{P})$, defined in (8.5).

To deal with the right-hand side of (11.15), we may appeal to Step 1 and Step 2. In particular, the fact that $\tilde{\mathrm{v}}^{\varphi}_t[g](x)$, given by (11.13), is equal to $\mathrm{v}^{\varphi}_t[g](x)$, given by (11.9), tells us that, for each fixed time $t \geq 0$, the laws of $(X^{\varphi}_t, \tilde{\mathbb{P}}^{\varphi}_{\delta_x})$ and $(X_t, \mathbb{P}^{\varphi}_{\delta_x})$ agree. It follows that

$$\begin{aligned}\tilde{\mathbb{E}}^{\varphi}_{\mu}\left[\frac{X^{\varphi}_t[f\varphi]}{X^{\varphi}_t[\varphi]}\mathrm{e}^{-X^{\varphi}_t[g]}\right] &= \mathbb{E}^{\varphi}_{\mu}\left[\frac{X_t[f\varphi]}{X_t[\varphi]}\mathrm{e}^{-X_t[g]}\right]\\ &= \mathrm{e}^{-\lambda_* t}\mathbb{E}_{\mu}\left[\frac{X_t[f\varphi]}{\mu[\varphi]}\mathrm{e}^{-X_t[g]}\right]\end{aligned}$$

$$= \sum_{k=1}^{n} \frac{\varphi(x_k)}{\mu[\varphi]} \mathrm{e}^{-\lambda_* t} \mathbb{E}_{\delta_{x_k}} \left[\frac{X_t[f\varphi]}{\varphi(x_i)} \mathrm{e}^{-X_t[g]} \right] \prod_{\ell \neq k} \mathtt{v}_t[g](x_\ell).$$

The proof of this final step is thus complete as soon as we can show that

$$\tilde{\mathbb{E}}^{\varphi}_{\delta_x} \left[f(\xi^{\varphi}_t) \prod_{i \geq 1: T_i \leq t} \prod_{j=1}^{N_i} \frac{\mathtt{v}_{t-T_i}[g](x_{ij})}{\mathtt{v}_{t-T_i}[g](\xi^{\varphi}_{T_i})} \right] = \mathrm{e}^{-\lambda_* t} \mathbb{E}_{\delta_x} \left[\frac{X_t[f\varphi]}{\varphi(x)} \mathrm{e}^{-X_t[g]} \right], \tag{11.16}$$

for $x \in E$. To this end, we note that splitting the expectation on the right-hand side of (11.16) at a branching event results in a calculation that is almost identical to the one above that concludes with (11.12). More precisely, the expectation on the right-hand side of (11.16) solves (11.12) albeit the role of $g(\xi_t)$ is replaced by $f(\xi_t)g(\xi_t)$. Similarly splitting the expectation on the left-hand side of (11.16) also results in a solution to (11.12) (with the aforementioned adjustment). Uniqueness follows from the same arguments, and hence, the equality in (11.16) now follows, as required. □

With the proof of Theorem 11.1 completed, we can now turn to the proof of Lemma 11.1.

***Proof** (of Lemma 11.1)* The fact that the spine is Markovian is immediate from the definition of ξ^{φ}. Indeed, once its initial configuration is given, it evolves according to the semigroup $\mathtt{P}^{\varphi}_t$ defined in (11.4) and when at position $x \in E$, at rate $\varphi(x)^{-1}\gamma(x)\mathtt{m}[\varphi](x)$, it jumps to a new position y, with distribution

$$\mathcal{E}_x \left[\frac{\mathcal{Z}[\varphi]}{\mathtt{m}[\varphi](x)} \frac{\mathcal{Z}[\varphi \mathbf{1}_{(\cdot \in \mathrm{d}y)}]}{\mathcal{Z}[\varphi]} \right] = \frac{\mathtt{m}[\varphi \mathbf{1}_{(\cdot \in \mathrm{d}y)}](x)}{\mathtt{m}[\varphi](x)},$$

for $y \in E$, where we have used (11.6). Defining

$$\tilde{w}^{\varphi}[g](x) := \tilde{\mathbf{E}}_x[g(\xi^{\varphi}_t)],$$

and splitting on the first jump of the spine, the above description implies that $\tilde{w}^{\varphi}_t$ satisfies

$$\begin{aligned}
&\tilde{w}^{\varphi}[g](x) \\
&= \frac{1}{\varphi(x)} \mathbf{E}_x \left[\mathrm{e}^{-\lambda_* t + \int_0^t \frac{\gamma(\xi^{\varphi}_s)}{\varphi(\xi^{\varphi}_s)} (\mathtt{m}[\varphi](\xi^{\varphi}_s) - \varphi(\xi^{\varphi}_s)) \mathrm{d}s} g(\xi^{\varphi}_t) \varphi(\xi^{\varphi}_t) \mathrm{e}^{-\int_0^t \frac{\gamma(\xi^{\varphi}_s)}{\varphi(\xi^{\varphi}_s)} \mathtt{m}[\varphi](\xi^{\varphi}_s) \mathrm{d}s} \right] \\
&\qquad + \mathbf{E}_x \left[\int_0^t \frac{\gamma(\xi^{\varphi}_s)}{\varphi(\xi^{\varphi}_s)} \mathtt{m}[\varphi](\xi^{\varphi}_s) \mathrm{e}^{-\int_0^s \frac{\gamma(\xi^{\varphi}_u)}{\varphi(\xi^{\varphi}_u)} \mathtt{m}[\varphi](\xi^{\varphi}_u) \mathrm{d}u} \right.
\end{aligned}$$

$$\times \mathrm{e}^{-\lambda_* s+\int_0^s \frac{\gamma(\xi_u^\varphi)}{\varphi(\xi_u^\varphi)}(\mathtt{m}[\varphi](\xi_u^\varphi)-\varphi(\xi_u^\varphi))\mathrm{d}u} \frac{\varphi(\xi_s^\varphi)}{\varphi(x)} \frac{\mathtt{m}[\tilde{w}_{t-s}^\varphi \varphi \mathbf{1}_{(\cdot\in \mathrm{d}y)}](\xi_s^\varphi)}{\mathtt{m}[\varphi](\xi_s^\varphi)}\Bigg]$$

$$= \frac{1}{\varphi(x)}\mathbf{E}_x\left[\mathrm{e}^{-\lambda_* t-\int_0^t \gamma(\xi_s^\varphi)\mathrm{d}s} g(\xi_t^\varphi)\varphi(\xi_t^\varphi)\right]$$
$$+\mathbf{E}_x\left[\int_0^t \frac{\gamma(\xi_s^\varphi)}{\varphi(x)}\mathrm{e}^{-\lambda_* s-\int_0^s \gamma(\xi_u^\varphi)\mathrm{d}u}\mathtt{m}[\tilde{w}_{t-s}^\varphi \varphi \mathbf{1}_{(\cdot\in \mathrm{d}y)}](\xi_s^\varphi)\mathrm{d}s\right]. \tag{11.17}$$

Now recall the change of measure (11.7) and define

$$\hat{w}_t^\varphi[g](x) = \hat{\mathbf{E}}_x\left[\mathrm{e}^{-\lambda_* t+\int_0^t \gamma(\hat{\xi}_s)(\mathtt{m}[1](\hat{\xi}_s)-1)\mathrm{d}s} g(\hat{\xi}_t)\frac{\varphi(\hat{\xi}_t)}{\varphi(x)}\right].$$

Let us now show that $\hat{w}_t^\varphi$ also satisfies (11.17). Recalling that $\hat{\xi}$ jumps at rate $\gamma\mathtt{m}$, we have

$$\hat{w}_t^\varphi[g](x) = \frac{1}{\varphi(x)}\mathbf{E}_x\left[g(\hat{\xi}_t)\varphi(\hat{\xi}_t)\mathrm{e}^{-\lambda_* t+\int_0^t \gamma(\hat{\xi}_s)(\mathtt{m}[1](\hat{\xi}_s)-1)}\mathrm{e}^{-\int_0^t \gamma(\hat{\xi}_s)\mathtt{m}[1](\hat{\xi}_s)\mathrm{d}s}\right]$$
$$+\mathbf{E}_x\Bigg[\int_0^t \gamma(\hat{\xi}_s)\mathtt{m}[1](\hat{\xi}_s)\mathrm{e}^{-\int_0^s \gamma(\hat{\xi}_u)\mathtt{m}[1](\hat{\xi}_u)\mathrm{d}u}$$
$$\times \mathrm{e}^{-\lambda_* s+\int_0^s \gamma(\hat{\xi}_u)(\mathtt{m}[1](\hat{\xi}_u)-1)\mathrm{d}u}\frac{\varphi(\hat{\xi}_s)}{\varphi(x)}\frac{\mathtt{m}[\hat{w}_{t-s}^\varphi \varphi \mathbf{1}_{(\cdot\in \mathrm{d}y)}](\xi_s^\varphi)}{\varphi(\hat{\xi}_s)\mathtt{m}[1](\hat{\xi}_s)}\Bigg]$$
$$= \frac{1}{\varphi(x)}\mathbf{E}_x\left[\mathrm{e}^{-\lambda_* t-\int_0^t \gamma(\xi_s^\varphi)\mathrm{d}s} g(\xi_t^\varphi)\varphi(\xi_t^\varphi)\right]$$
$$+\mathbf{E}_x\left[\int_0^t \frac{\gamma(\xi_s^\varphi)}{\varphi(x)}\mathrm{e}^{-\lambda_* s-\int_0^s \gamma(\xi_u^\varphi)\mathrm{d}u}\mathtt{m}[\hat{w}_{t-s}^\varphi \varphi \mathbf{1}_{(\cdot\in \mathrm{d}y)}](\xi_s^\varphi)\mathrm{d}s\right],$$

as required. Uniqueness follows as in the conclusion of Step 2 in the proof of Theorem 11.1.

Recalling notation from Lemma 8.2,

$$\hat{\mathtt{T}}_t^\varphi[g](x) := \hat{\mathbf{P}}_x^\varphi[g(\hat{\xi}_t)] = \mathrm{e}^{-\lambda_* t}\frac{\psi_t[g\varphi](x)}{\varphi(x)}, \qquad x\in E, \tag{11.18}$$

where $g \in B^+(E)$. Recalling the eigenvalue property of φ, by taking $g \equiv 1$, we see that $\hat{\mathtt{T}}_t^\varphi[1](x) = 1$ and hence $(\hat{\xi}, \hat{\mathbb{P}}^\varphi)$ is conservative. Moreover, $\lim_{t\to\infty}\hat{\mathbf{P}}_x^\varphi[g(\hat{\xi}_t)] = \tilde{\varphi}[\varphi g]$ for all $g\in B^+(E)$. In other words, $\varphi\tilde{\varphi}$ is the density of the stationary distribution of $\hat{\xi}$ under $\hat{\mathbf{P}}^\varphi$. □

11.2 Examples of the Spine Decomposition

In the previous section, we have seen the spine decomposition for our general class of BMPs introduced in Chap. 8. The decomposition is both simple and complicated. It is "simple", in that the change of measure allows one to see the evolution of the BMP as a spine process that is "dressed" with original copies of the original BMP at certain times of "immigration" where the spine acquires siblings. But also "complicated" because the technical details of the aforesaid description are somewhat involved. In particular, the non-locality of our BMP is part of the reason why the technical details are so complex.

To help give a better insight, let us look at how the spine decomposition takes shape for less general but familiar branching processes. We have already seen in Sect. 6.3 how the spine decomposition plays out for the NBP, and it offers little more specificity than in the general case. We therefore focus on two cases where specific aspects of the decomposition can be identified in other familiar detail.

Multi-type Branching Process This is the most basic of non-local branching process. Individuals have no associated motion and have a type belonging to $\{1, \cdots, n\}$. Branching occurs at a constant rate $\gamma > 0$, and particles of type $i \in \{1, \cdots, n\}$ give birth to offspring of all types with law $\mathcal{P}_i$. Assumption (G1) ensures that the mean offspring of each type is finite, irrespective of the type of the parent. Rather than thinking of $(X_t, t \geq 0)$ as a measure-valued process, we can write it as a vector $X_t = (X_t(1), \cdots, X_t(n))$, where $X_t(i)$ simply counts the number of individuals of type i alive at time $t \geq 0$. The mean semigroup can be expressed in terms of a matrix $\psi_t(i, j) = \mathbb{E}_i[X_t(j)]$, $i, j \in \{1, \cdots, n\}$. Assumption (G2) can be understood as the existence of right and left eigenvectors, say $\varphi = (\varphi_1, \cdots, \varphi_n)$ and $\tilde{\varphi} = (\tilde{\varphi}_1, \cdots, \tilde{\varphi}_n)$ with eigenvalue $\lambda_* \in \mathbb{R}$, so that, in the sense of vector–matrix multiplication,

$$\tilde{\varphi}^{\mathrm{T}}\psi_t = \mathrm{e}^{\lambda_* t}\tilde{\varphi}^{\mathrm{T}} \quad \text{and} \quad \psi_t\varphi = \mathrm{e}^{\lambda_* t}\varphi. \tag{11.19}$$

Moreover, the asymptotic behaviour of the semigroup $(\psi_t, t \geq 0)$ is nothing more than the classical Perron–Frobenius asymptotic

$$\lim_{t\to\infty} \mathrm{e}^{-\lambda_* t}\psi_t g = (\tilde{\varphi} \cdot g)\varphi,$$

where $g = (g_1, \cdots, g_n)$ is a non-negative vector in $\mathbb{R}^n$ and $\tilde{\varphi} \cdot g = \tilde{\varphi}^{\mathrm{T}} g$ is the classical Euclidian inner product.

The analogue of the martingale (11.1) is the classical martingale

$$\mathrm{e}^{-\lambda_* t}\frac{X_t \cdot \varphi}{X_0 \cdot \varphi}, \qquad t \geq 0.$$

From Lemma 11.1, in particular, (11.8), the change of measure associated with this martingale induces a spine decomposition in which ξ^φ is a continuous-time Markov chain whose transition semigroup satisfies

$$\tilde{\mathtt{P}}^{\varphi}_t[g](i) = \hat{\mathbf{E}}_i\left[\mathrm{e}^{-\lambda_* t + \int_0^t \gamma(\mathtt{m}_{\hat{\xi}_s}\cdot \mathtt{1} - 1)\mathrm{d}s} \frac{\varphi_{\hat{\xi}_t}}{\varphi_i} g_{\hat{\xi}_t}\right], \qquad t \geq 0, g \in \mathbb{R}^n, i \in \{1, \cdots, n\}, \tag{11.20}$$

where $\mathtt{m}_i = (\mathtt{m}_i(1), \cdots, \mathtt{m}_i(n)) = (\mathscr{E}_i[N(1)], \cdots, \mathscr{E}_n[N(n)])$ and $(N(i), i = 1, \cdots, n)$, are the number of offspring produced of each type at a typical branching event. From the discussion preceding Lemma 8.2, the continuous-time Markov chain $(\hat{\xi}, \hat{\mathbf{P}})$ has intensity matrix $\hat{Q}$ given by

$$\hat{Q}_{ij} = \gamma \mathtt{m}_i \cdot \mathtt{1} \frac{\mathtt{m}_i(j)}{\mathtt{m}_i \cdot \mathtt{1}} = \gamma \mathtt{m}_i(j), \qquad i, j \in \{1, \cdots, n\}, \tag{11.21}$$

where $\mathtt{1}$ is the vector in $\mathbb{R}^n$ whose entries are all unity. Hence, the Q-matrix associated to the spine, say $\tilde{Q}^\varphi$, is easily found by differentiating (11.20) to get

$$\begin{aligned}(\tilde{Q}^\varphi g)_i, &= \frac{\mathrm{d}}{\mathrm{d}t} \mathtt{P}^\varphi_t[g](i)\bigg|_{t=0} \\ &= -\lambda_* g_i + \gamma(\mathtt{m}_i \cdot \mathtt{1} - 1) g_i + \frac{1}{\varphi_i}(\hat{Q}(\varphi \odot g))_i,\end{aligned}$$

where $\varphi \odot g$ is elementwise multiplication of the vectors φ and g. In particular, if $g = \delta_j$, the vector whose elements are all zero albeit an entry of unity for the j-th element, then we see that

$$\tilde{Q}^\varphi_{ij} = \frac{1}{\varphi_i}\left(\hat{Q}_{ij} + \gamma(\mathtt{m}_i \cdot \mathtt{1} - 1)\mathtt{I}_{ij} - \lambda_* \mathtt{I}_{ij}\right)\varphi_j, \qquad i, j \in \{1, \cdots, n\}, \tag{11.22}$$

where $\mathtt{I}$ is the identity matrix. Note however that the eigenvalue λ_* already carries a relationship with $\hat{Q}$ due to Lemma 8.2. Indeed, differentiating the many-to-one formula in (8.17) and taking account of (11.19) yield, for $g \in \mathbb{R}^n$,

$$\lambda_* \varphi_i, = \frac{\mathrm{d}}{\mathrm{d}t}(\psi_t \varphi)_i\bigg|_{t=0} = \gamma(\mathtt{m}_i \cdot \mathtt{1} - 1)\varphi_i + (\hat{Q}\varphi)_i, \qquad i \in \{1, \cdots, n\}. \tag{11.23}$$

Together with (11.22) and (11.21), this tells us that the spine is a continuous-time Markov process with intensity

$$\tilde{Q}^\varphi_{ij} = \gamma \frac{1}{\varphi_i}\left(\mathtt{m}_i(j) - \mathtt{m}_i \cdot \varphi \mathtt{I}_{ij}\right)\varphi_j, \qquad i, j \in \{1, \cdots, n\}.$$

Similarly to (11.23), by using the many-to-one formula in (8.17), we get, for $g \in \mathbb{R}^n$,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\tilde{\varphi}^{\mathtt{T}}\psi_t)_j\bigg|_{t=0} = \gamma\sum_{i=1}^{n}\tilde{\varphi}_i(\mathtt{m}_i\cdot\mathtt{1}-1)\mathtt{I}_{ij} + (\tilde{\varphi}^{\mathtt{T}}\hat{Q})_j$$
$$= \gamma\sum_{i=1}^{n}\tilde{\varphi}_i\Big((\mathtt{m}_i\cdot\mathtt{1}-1)\mathtt{I}_{ij} + \mathtt{m}_i(j)\Big).$$

Hence, differentiating across (11.19) and setting $t = 0$ thus give us

$$\gamma\sum_{i=1}^{n}\tilde{\varphi}_i\Big((\mathtt{m}_i\cdot\mathtt{1}-1)\mathtt{I}_{ij} + \mathtt{m}_i(j)\Big) = \lambda_*\mathtt{I}_{ij}. \tag{11.24}$$

We can use (11.24) to verify that the spine has stationary distribution $(\tilde{\varphi}_i\varphi_i, i = 1, \cdots, n)$, as predicted by Lemma 11.1. Indeed, taking the representation for $\tilde{Q}^{\varphi}$ given in (11.22) and appealing to (11.24),

$$\sum_{i=1}^{n}\tilde{\varphi}_i\varphi_i\tilde{Q}^{\varphi}_{ij} = \gamma\sum_{i=1}^{n}\tilde{\varphi}_i\left(\hat{Q}_{ij} + \gamma(\mathtt{m}_i\cdot\mathtt{1}-1)\mathtt{I}_{ij} - \lambda_*\mathtt{I}_{ij}\right)\varphi_j = 0.$$

Branching Brownian Motion in a Bounded Domain We recall the introduction of this process on p193 in which setting the semigroup $\mathtt{P}$ corresponds to that of a d-dimensional Brownian motion killed on exiting a C^1 domain $E \subset \mathbb{R}^d$. The branching rate is taken as the constant $\gamma > 0$ and the offspring distribution, say $(p_k, k = 0, 1, \cdots)$, is not spatially dependent. Assumption (G1) assumes the mean number of offspring, $m_1 := \mathscr{E}[N]$, is finite and, under (G2), $\lambda_* = \gamma(m_1 - 1) + \lambda_E$, where λ_E is the ground state eigenvalue of the Laplacian on E. Further, if we write $\mathtt{L} = \Delta/2$ on E, as the generator associated to $\mathtt{P}$, the eigenpair (λ_*, φ) is related via the generator equation

$$(\mathtt{L} + \gamma(m_1 - 1))\varphi(x) = \lambda_*\varphi(x), \qquad x \in E,$$

in other words,

$$\mathtt{L}\varphi(x) = \lambda_E\varphi(x), \qquad x \in E.$$

Note that the process $\hat{\xi}$ has no additional jumps and hence also has generator $\mathtt{L}$. From (11.8) in Lemma 11.1, by assuming that $g \in B^+(E)$ is sufficiently smooth, for example, $g \in C^2(E)$, we can apply standard stochastic calculus to deduce that

$$\tilde{\mathtt{L}}^{\varphi}g = \frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathtt{P}}^{\varphi}_t[g](x)\bigg|_{t=0} = -\lambda_*g(x) + \gamma(m_1 - 1)g(x) + \frac{1}{\varphi(x)}\mathtt{L}(\varphi g)(x).$$

We can read out of this calculation that the generator of the spine ξ^φ is given by

$$\tilde{\mathtt{L}}^\varphi g := \lambda_E + \frac{1}{\varphi(x)}\mathtt{L}(\varphi g)(x) = \frac{1}{\varphi(x)}(\mathtt{L} - \lambda_E)(\varphi g)(x), \qquad x \in E,$$

for $g \in C^2(E)$. This is well understood as the Doob h-transform to the generator $\mathtt{L}$, which corresponds to conditioning the Brownian motion to remain in E.

Another interesting feature of this model is what happens at points of immigration along the spine. Because there is no spatial dependency in the offspring distribution nor the branching rate, it is easy to see that $\mathtt{m}[\varphi](x) = \varphi(x)m_1$ and $\mathsf{Z}[\varphi] = \varphi(x)N$, where N is the number of offspring. As such, from (11.5), we see that the branching rate along the spine is simply γm_1, and from (11.6), the offspring distribution is adjusted from $(p_k, k = 0, 1, 2, \cdots)$ to $(kp_k/m_1, k = 0, 1, 2, \cdots)$. Moreover, from (11.3), we also see that the selection of the individual to mark with the spine is uniform among offspring of the last individual in the spine.

The features in the last paragraph are all a consequence of there being a local branching structure with no spatial dependency. Indeed, this is common to spine decompositions of all branching processes that do not exhibit non-locality.

11.3 The Spine Decomposition and Criticality

We would like to focus on the critical setting and discuss how the spine decomposition gives us insight into the behaviour of the process conditional on its survival, as well as of its moment asymptotics, described in Theorem 9.1. In particular, we will go part way to explaining why higher moments can all be written in terms of only the second moment functional $\mathscr{V}[\varphi]$ given in (9.9).

For convenience, let us assume that the conditions of Theorem 10.1 are in force. Recall that this theorem tells us that

$$\lim_{t\to\infty} t\mathbb{P}_{\delta_x}(\zeta > t) = \frac{2\varphi(x)}{\Sigma}, \tag{11.25}$$

where $\Sigma = \tilde{\varphi}\big[\gamma\mathscr{V}[\varphi]\big]$. In particular, this tells us that

$$\lim_{t\to\infty} \frac{\mathbb{P}_{\delta_y}(\zeta > t)}{\mathbb{P}_{\delta_x}(\zeta > t+s)} = \frac{\varphi(y)}{\varphi(x)}, \qquad x, y \in E, s \geq 0. \tag{11.26}$$

An argument involving dominated convergence now allows us to conclude from (11.26) that

$$\lim_{t\to\infty} \mathbb{E}_{\delta_x}\left[\mathrm{e}^{-X_s[g]}\Big| \zeta > t+s\right] = \lim_{t\to\infty} \frac{\mathbb{E}_{\delta_x}\left[\mathrm{e}^{-X_s[g]}\right]}{\mathbb{P}_{\delta_x}(\zeta > t+s)}$$

$$
\begin{aligned}
&= \lim_{t\to\infty} \mathbb{E}_{\delta_x}\left[\mathrm{e}^{-X_s[g]} \frac{\mathbb{P}_{X_s}(\zeta > t)}{\mathbb{P}_{\delta_x}(\zeta > t+s)} \right] \\
&= \mathbb{E}_{\delta_x}\left[\mathrm{e}^{-X_s[g]} \frac{X_s[\varphi]}{\varphi(x)} \right]. \qquad (11.27)
\end{aligned}
$$

The above calculation tells us that the way to ensure that the critical BMP survives is by selecting a single genealogical line of descent, the spine, and ensuring that it never dies out. This is sometimes referred to as *the immortal particle* as, from Lemma 11.1, we see that it has stationary distribution given by $\varphi(x)\tilde{\varphi}(\mathrm{d}x)$, $x \in E$.

In the light of this conditioning, let us now turn our attention to understanding how the spine decomposition heuristically explains why the observed asymptotic moment behaviour in Theorem 9.1 depends only on the second moment functional $\tilde{\varphi}\big[\gamma \mathscr{V}[\varphi]\big]$.

Let us momentarily take for granted (11.25) and the moment asymptotics of Theorem 9.1 and use them to give the promised heuristic explanation of why the asymptotic moments in Theorem 9.1 all end up written in terms of the second moment functional $\mathscr{V}[\varphi]$.

Writing

$$
\mathrm{T}_t^{(k)}[f](x) = \mathbb{E}_{\delta_x}[\langle f, X_t\rangle^k | \zeta > t]\mathbb{P}_{\delta_x}(\zeta > t), \qquad (11.28)
$$

we can consider the behaviour of the conditional moments in (11.28) via the change of measure (11.2). In particular, this means that we can write, for $f \in B^+(E)$,

$$
X_t[f] = f(\xi_t^{\varphi}) + \sum_{i=1}^{n_t} \sum_{j=1}^{N_i} X_{t-T_i}^{ij}[f], \qquad (11.29)
$$

where:

- ξ^{φ} is the spine.
- n_t is the number of branching events along the spine, which arrive at a rate that depends on the motion of ξ.
- N_i is the number of offspring produced such that, at the i-th such branching event, which occurs at time $T_i \le t$.
- $X_{t-T_i}^{ij}$, $j = 1, \ldots, N_i$, are i.i.d. copies of the original branching Markov process initiated from $(x_j, j = 1, \cdots, N_i)$, the positions of the offspring produced at the branching event.

In other words, that under $\mathbb{P}^{\varphi}$, the process X can be decomposed into a single immortal trajectory, of which copies of the original process $(X, \mathbb{P})$ immigrate simultaneously from groups of siblings.

With this in mind, let us consider genealogical lines of descent that contribute to the bulk of the mass of the k-th moment at large times t. For each copy of $(X, \mathbb{P})$ that immigrates onto the spine at time $s > 0$, the probability that the process survives to time $t \geq s$, thus contributing to the bulk of the k-th moment at time t, is $O(1/(t-s)) \approx O(1/t)$, cf. (11.25). If there are multiple offspring at an immigration event at time s, then the chance that at least two of these offspring contribute to the bulk of the k-th moment at time t is $O(1/t^2)$. Moreover, from Lemma 11.1, the semigroup of the spine limits to a stationary distribution $\varphi(x)\tilde{\varphi}(\mathrm{d}x)$, $x \in E$. This stationary behaviour has the effect that the arrival of branching events along the spine begins to look increasingly like a Poisson process as $t \to \infty$. Hence, for large t, $n_t \approx O(t)$.

Putting these pieces together, as $t \to \infty$, there are approximately $O(t)$ branch points along the spine, each of which has the greatest likelihood of a single offspring among immigrating siblings contributing to the bulk of the k-th moment at time t, with probability of order $O(1/t)$. Thus, it is clear that we only expect to see one of each sibling group of immigrants along the spine contributing to the mass of the k-th moment at time t. Now let β^φ denote the spatial rate at which offspring immigrates onto the spine and let $\{x_1, \ldots, x_N\}$ denote their positions at the point of branching *including* the position of the spine at this instance. Let $\mathscr{P}^\varphi$ denote the law of this offspring distribution, and suppose that i^* is the (random) index of the offspring that continues the evolution of the spine. The rate at which a "uniform selection" of a single offspring occurs that is not the spine at a branching event (seen through the function $f \in B^+(E)$) is given by

$$
\begin{aligned}
\gamma^\varphi(x)\mathscr{E}_x^\varphi\left[\sum_{i=1}^N f(x_i)\mathbf{1}_{(i\neq i^*)}\right] &= \gamma(x)\frac{\mathscr{E}_x\left[\mathsf{Z}[\varphi]\right]}{\varphi(x)}\mathscr{E}_x \\
&\quad\times\left[\frac{\mathsf{Z}[\varphi]}{\mathscr{E}_x\left[\mathsf{Z}[\varphi]\right]}\sum_{i=1}^N\frac{\varphi(x_i)}{\mathsf{Z}[\varphi]}\sum_{\substack{i=1\\ j\neq i}}^N f(x_j)\right] \\
&= \frac{\gamma(x)}{\varphi(x)}\mathscr{E}_x\left[\sum_{i=1}^N\varphi(x_i)\sum_{\substack{i=1\\ j\neq i}}^N f(x_j)\right] \\
&= \frac{\gamma(x)}{\varphi(x)}\mathscr{E}_x\left[\mathsf{Z}[f]\mathsf{Z}[\varphi]-\mathsf{Z}[\varphi f]\right],
\end{aligned}
\tag{11.30}
$$

where we have used the features of the spine decomposition given in (11.3), (11.5) and (11.6).

We know from the assumed behaviour of the first moment in (G2) that it is the projection of $\langle f, X_t\rangle$ on to $X_t[\varphi]$, with coefficient $\langle f, \tilde{\varphi}\rangle$, which dominates the mean growth. In this spirit, let us take $f = \varphi$ for simplicity, and we see that in (11.30) we get precisely $\mathscr{V}[\varphi](x)/\varphi(x)$ on the right-hand side.

Hence, finally, we conclude our heuristic by observing that the rate at which immigration off the spine contributes to the bulk of the k-th moment limit of $X_t[\varphi]$ is determined by the second moment functional $\mathscr{V}[\varphi]$; together with (11.28) and the associated remarks above, this goes some way to explaining the appearance of the limit in Theorem 9.1.

11.4 T-Skeletal Decomposition

In this section, we remain in the setting of a general BMP, as described in Chap. 8. Our aim is to provide a more general skeletal decomposition to the one presented in Chap. 6. In particular, under appropriate assumptions, we decompose the branching process into the genealogies that survive up to time $T < \infty$, in which individuals will carry the mark $\uparrow$, and those that die out before time T, in which individuals will carry the mark $\downarrow$. The calculations we present are robust, and the reader is encouraged to consider that they are equally valid for the setting $T = \infty$ (as presented in Chap. 6), when the process survives with positive probability. We will provide further remarks at the end of this section to this end. The reader will note that we have not assumed (G1) or (G2) in our analysis.

As with the skeletal decomposition of the NBP in Sect. 6.4 (albeit now in finite time), we see that our BMP decomposes as equal in law to a "thinner" tree of all $\uparrow$-marked individuals, dressed with immigrating trees of all $\downarrow$-marked individuals.

For now, let us fix $0 \leq t \leq T < \infty$. In order to describe the evolution of the $\{\uparrow, \downarrow\}$-valued marks along the running population, consider the configuration of the BMP at time t, with positions in E given by $\{x_j(t),\ j = 1, \cdots, N_t\}$. Given $\mathfrak{F}_T$, if particle i is such that it has descendants alive at time T, then we define its mark $c_i^T(t) =\uparrow$. On the other hand, if every line of descent from the i-th particle has become extinct by time T, then define its mark $c_i^T(t) =\downarrow$.

Next, set

$$w_T(x) := \mathbb{P}_{\delta_x}(\zeta < T), \qquad 0 \leq T < \infty, x \in E,$$

where we recall that $\zeta = \inf\{t > 0 : X_t[1] = 0\}$ is the extinction time of the process. We will also frequently use the notation

$$p_T(x) := \mathbb{P}_{\delta_x}(\zeta > T) = 1 - w_T(x), \qquad 0 \leq T < \infty,$$

for the survival (up to time T) probability. The quantity w_T will be central to the skeletal decomposition. The extreme cases that $w_T \equiv 0$ and $w_T \equiv 1$ will turn out to be degenerate for our purposes. With this in mind, there are other exclusions we need to be mindful of which lead us to the following assumptions that are in force throughout this section:

(G9) Extinction by time T is uniformly bounded away from zero,

$$\inf_{x \in E} w_T(x) > 0.$$

(G10) Extinction by time T is not a certainty,

$$w_T(x) < 1 \text{ for } x \in E.$$

As alluded to in the introduction to this section, we want to describe how the spatial genealogical tree of the BMP up to time T can be split into a spatial genealogical subtree, consisting of $\uparrow$-labelled particles (the skeleton), which is dressed with trees of $\downarrow$-labelled particles. To this end, let $\mathbb{P}^{\updownarrow,T} = (\mathbb{P}^{\updownarrow,T}_\mu, \mu \in \mathcal{M}_c(E))$ denote the probabilities of the $\{\uparrow,\downarrow\}$-marked BMP, where $\mu \in \mathcal{M}_c(E)$ is the initial spatial configuration of particles. Then, writing as before $\{x_i(t) : i = 1, \ldots, N_t\}$ for the set of particles alive at time t, we have the following relationship between $\mathbb{P}^{\updownarrow,T}_\mu$ and $\mathbb{P}_\mu$:

$$\left.\frac{\mathrm{d}\mathbb{P}^{\updownarrow,T}_\mu}{\mathrm{d}\mathbb{P}_\mu}\right|_{\mathfrak{F}_t} = \prod_{i=1}^{N_t}\left(\mathbf{1}_{(c_i^T(t)=\uparrow)} + \mathbf{1}_{(c_i^T(t)=\downarrow)}\right) = 1. \tag{11.31}$$

Projecting onto $\mathfrak{F}_t$, for $t \geq 0$, we have

$$\begin{aligned}
\left.\frac{\mathrm{d}\mathbf{P}^{\updownarrow,T}_{\delta_x}}{\mathrm{d}\mathbf{P}_{\delta_x}}\right|_{\mathfrak{F}_t} &= \mathbf{E}_{\delta_x}\left(\prod_{i=1}^{N_t}\left(\mathbf{1}_{(c_i(t)=\uparrow)} + \mathbf{1}_{(c_i^T(t)=\downarrow)}\right)\middle|\mathfrak{F}_t\right)\\
&= \sum_{I\subseteq\{1,\ldots N_t\}}\prod_{i\in I}\mathbf{P}_{\delta_x}(c_i^T(t) = \uparrow|\mathfrak{F}_t)\prod_{i\in\{1,\ldots,N_t\}\setminus I}\mathbf{P}_{\delta_x}(c_i^T(t) = \downarrow|\mathfrak{F}_t)\\
&= \sum_{I\subseteq\{1,\ldots N_t\}}\prod_{i\in I}p_{T-t}(x_i(t))\prod_{i\in\{1,\ldots,N_t\}\setminus I}w_{T-t}(x_i(t)),
\end{aligned} \tag{11.32}$$

where we understand the sum to be taken over all subsets of $\{1, \ldots, N_t\}$, each of which is denoted by I, and we have appealed to the branching Markov property (8.6) in the last equality. Technically speaking, the right-hand side of (11.32) is equal to unity on account of the fact that the right-hand side of (11.31) is unity. Despite the rather complex looking identity for (11.32), this is indeed true since, for any $n \in \mathbb{N}$ and $a_1, \ldots, a_n, b_1, \ldots, b_n \geq 0$,

$$\prod_{i=1}^{n}(a_i + b_i) = \sum_{I\subset\{1,\ldots,n\}}\prod_{i\in I}a_i\prod_{j\in\{1,\ldots,n\}\setminus I}b_j, \tag{11.33}$$

where the sum is taken over all subsets I of $\{1, \ldots, n\}$. Hence, for any $t \leq T$,

$$\sum_{I\subset\{1,\ldots,N_t\}}\prod_{i\in I}p_{T-t}(x_i(t))\prod_{j\in\{1,\ldots,N_t\}\setminus I}w_{T-t}(x_j(t))$$

$$= \prod_{i=1}^{N_t} (p_{T-t}(x_i(t)) + w_{T-t}(x_i(t))) = 1.$$

The decomposition in (11.32) is the starting point of how we break up the law of the $(\mathtt{P}, \mathtt{G})$-BMP according to subtrees that are categorised as $\downarrow$ and subtrees that are categorised as $\uparrow$ with $\downarrow$ dressing, the so-called skeletal decomposition.

In what follows, we consider different ways that the extinction probability w_T can be used to condition $(X, \mathbb{P})$. These conditioned versions are important as, when conditionally dependent copies of them are put together in the right way, we can begin to interpret the meaning of the change of measure (11.32).

$\downarrow$-*Trees*

We start by describing the evolution of the trees conditioned to be extinct by time T. Thanks to the branching property, it suffices to consider trees that are issued with a single particle with mark $\downarrow$. By definition of the mark $c_\emptyset^T(0) =\downarrow$, where $\emptyset$ is the label of the initial ancestral particle, this is the same as understanding the law of $(X, \mathbb{P})$ conditioned to become extinct. Indeed, for $A \in \mathfrak{F}_t$

$$\begin{aligned}
\mathbb{P}_{\delta_x}^{\downarrow,T}(A) &:= \mathbb{P}_{\delta_x}^{\updownarrow,T}(A | c_\emptyset^T(0) = \downarrow) \\
&= \frac{\mathbb{P}_{\delta_x}^{\updownarrow,T}(A\,;\ c_i^T(t) = \downarrow \text{ for each } i = 1, \ldots, N_t)}{\mathbb{P}_{\delta_x}^{\updownarrow,T}(c_\emptyset^T(0) = \downarrow)} \\
&= \frac{\mathbb{E}_{\delta_x}[\mathbf{1}_A \prod_{i=1}^{N_t} w_{T-t}(x_i(t))]}{w_T(x)}.
\end{aligned} \tag{11.34}$$

Moreover, conditioning the extinction event $\{\zeta < T\}$ on $\mathfrak{F}_t$, for $0 \le t \le T$, yields

$$w_T(x) = \mathbb{E}_{\delta_x}\left[\prod_{i=1}^{N_t} w_{T-t}(x_i(t))\right]. \tag{11.35}$$

Thus we may use (11.35) to define the following change of measure,

$$\left.\frac{\mathrm{d}\mathbb{P}_\mu^{\downarrow,T}}{\mathrm{d}\mathbb{P}_\mu}\right|_{\mathfrak{F}_t} := \frac{\prod_{i=1}^{N_t} w_{T-t}(x_i(t))}{\prod_{i=1}^{n} w_T(x_i)}, \tag{11.36}$$

for $\mu = \sum_{i=1}^n \delta_{x_i} \in \mathscr{M}_c(E)$. It is also straightforward to see that, using the non-linear evolution equation (8.7) and (11.35), w_T satisfies

$$w_T(x) = \hat{\mathtt{P}}_t[w_{T-t}](x) + \int_0^t \mathtt{P}_s\left[w_{T-s}\frac{\mathtt{G}[w_{T-s}]}{w_{T-s}}\right](x)\,\mathrm{d}s, \quad x \in E, \tag{11.37}$$

where $(\hat{\mathtt{P}}_t, t \geq 0)$ is the adjusted semigroup that returns a value of 1 on the event of killing, and we recall that $\mathtt{G}$ is the branching mechanism given by (8.2). Using Theorem 2.1, it follows that

$$w_T(x) = \mathbf{E}_x\left[w_{T-t}(\xi_{t\wedge \mathtt{k}})\exp\left(\int_0^{t\wedge\mathtt{k}} \frac{\mathtt{G}[w_{T-s}](\xi_s)}{w_{T-s}(\xi_s)}\right)\right], \quad x \in E,\ 0 \leq t \leq T, \tag{11.38}$$

where $\mathtt{k}$ is the lifetime of the process ξ (which we distinguish from ζ, the lifetime of the branching Markov process X). This identity will turn out to be extremely useful in our analysis. In particular, the equality (11.38) together with the Markov property of ξ implies that the object in the expectation on the right-hand side of (11.38) is a martingale.

We are now in a position to characterise the law of the process under (11.36).

Lemma 11.2 (↓-Trees) *Under $\mathbb{P}^{\downarrow,T}_{\delta_x}$, $(X_t, t \leq T)$ is a time-inhomogeneous BMP with motion semigroup $\mathtt{P}^{\downarrow,T}$ and branching mechanism $(\mathtt{G}^{\downarrow,T}_s, s \leq T)$ defined as follows. The motion semigroup $\mathtt{P}^{\downarrow,T}$ is that of the Markov process ξ with probabilities $(\mathbf{P}^{\downarrow,T}_x, x \in E)$, where*

$$\left.\frac{\mathrm{d}\mathbf{P}^{\downarrow,T}_x}{\mathrm{d}\mathbf{P}_x}\right|_{\sigma(\xi_s, s\leq t)} = \frac{w_{T-t}(\xi_{t\wedge\mathtt{k}})}{w_T(x)}\exp\left(\int_0^{t\wedge\mathtt{k}}\frac{\mathtt{G}[w_{T-s}](\xi_s)}{w_{T-s}(\xi_s)}\mathrm{d}s\right), \qquad t \geq 0. \tag{11.39}$$

For $x \in E$ and $f \in B^+_1(E)$, when a branching event occurs at time $s \leq T$, the branching mechanism is given by

$$\mathtt{G}^{\downarrow,T}_s[f](x) = \frac{1}{w_{T-s}(x)}\left[\mathtt{G}[f w_{T-s}] - f\mathtt{G}[w_{T-s}]\right](x). \tag{11.40}$$

Proof First let us show that the change of measure (11.36) results in a particle process that respects the Markov branching property. It is clear from the conditioning in (11.36) that, by construction, every particle in the resulting process under the new measure $\mathbb{P}^{\downarrow,T}_\mu$ must carry the mark ↓, i.e., become extinct by time T.

Let us define, for $g \in B^+_1(E)$, $x \in E$, and $0 \leq t \leq T$,

$$\mathtt{u}^{\downarrow,T}_t[g](x) = \mathbb{E}^{\updownarrow,T}_{\delta_x}\left[\prod_{i=1}^{N_t} g(x_i(t))\middle| c^T_\emptyset(0) = \downarrow\right] = \frac{1}{w_T(x)}u_t[g w_{T-t}](x), \tag{11.41}$$

which describes the evolution of the process X under $\mathbb{P}^{\downarrow,T}$. In particular, for $g \in B_1^+(E)$, $x \in E$ and $s, t \geq 0$ such that $t + s < T$, note that

$$
\begin{aligned}
&\mathbb{E}_{\delta_x}^{\downarrow,T}\left[\prod_{i=1}^{N_{t+s}} g(x_i(t+s)) \middle| \mathfrak{F}_t\right] \\
&= \frac{1}{w_T(x)} \prod_{i=1}^{N_t} w_{T-t}(x_i(t)) \mathbb{E}_{\delta_x}\left[\frac{\prod_{j=1}^{N_s^i} g(x_j^i(s)) w_{T-t-s}(x_j^i(s))}{w_{T-t}(x_i(t))} \middle| \mathfrak{F}_t\right] \\
&= \frac{1}{w_T(x)} \prod_{i=1}^{N_t} w_{T-t}(x_i(t)) u_s^{\downarrow,T-t}[g](x_i(t)), \qquad (11.42)
\end{aligned}
$$

where, given $\mathfrak{F}_t$, $\{(x_j^i(s)), j = 1, \cdots, N_s^i\}$ is the configuration of particles at time $t + s$ that are descendants of the i-th particle alive at time t for $i = 1, \ldots, N_t$. The equality (11.42) is a manifestation of the Markov branching property.

Given the statement of the lemma, it thus suffices for the remainder of the proof to show that, for $g \in B_1^+(E)$,

$$
\mathtt{u}_t^{\downarrow,T}[g](x) = \hat{\mathtt{P}}_t^{\downarrow,T}[g](x) + \int_0^t \mathtt{P}_s^{\downarrow,T}[\mathtt{G}_s^{\downarrow,T}[\mathtt{u}_{t-s}^{\downarrow,T-s}[g]](x)\mathrm{d}s, \quad 0 \leq t \leq T,\ x \in E, \qquad (11.43)
$$

holds, where, similarly to Chap. 5, $\hat{\mathtt{P}}^{\downarrow,T}$ is an adjustment of $\hat{\mathtt{P}}$ that returns the value one on the event of killing. (Recall this is a consequence of the empty product being defined as unity.) The reader will, by now, recognise the above equation as the semigroup evolution equation of a BMP, albeit a time-inhomogeneous one, with Markov motion affiliated to the semigroup $\mathtt{P}^{\downarrow,T}$ and time-dependent branching mechanism $(\mathtt{G}_s^{\downarrow,T}, s \leq T)$.

Using (11.41) and splitting the $(\mathtt{P}, \mathtt{G})$-BMP on the first branching event, it follows that, for $g \in B_1^+(E)$ and $0 \leq t \leq T$,

$$
\begin{aligned}
\mathtt{u}_t^{\downarrow,T}[g] &= \frac{1}{w_T}\hat{\mathtt{P}}_t[g w_{T-t}] + \frac{1}{w_T}\int_0^t \mathtt{P}_s[\mathtt{G}[\mathtt{u}_{t-s}[g w_{T-t}]]]\mathrm{d}s \\
&= \frac{1}{w_T}\hat{\mathtt{P}}_t[g w_{T-t}] + \frac{1}{w_T}\int_0^t \mathtt{P}_s[\mathtt{G}[w_{T-s}\mathtt{u}_{t-s}^{\downarrow,T-s}[g]]]\mathrm{d}s \\
&= \frac{1}{w_T}\hat{\mathtt{P}}_t[g w_{T-t}] + \frac{1}{w_T}\int_0^t \mathtt{P}_s\left[w_{T-s}\frac{\mathtt{G}[w_{T-s}\mathtt{u}_{t-s}^{\downarrow,T-s}[g]]}{w_{T-s}}\right]\mathrm{d}s \\
&\quad + \frac{1}{w_T}\int_0^t \mathtt{P}_s\left[\frac{\mathtt{G}[w_{T-s}]}{w_{T-s}} w_{T-s}\mathtt{u}_{t-s}^{\downarrow,T-s}[g]\right]\mathrm{d}s \\
&\quad - \frac{1}{w_T}\int_0^t \mathtt{P}_s\left[\frac{\mathtt{G}[w_{T-s}]}{w_{T-s}} w_{T-s}\mathtt{u}_{t-s}^{\downarrow,T-s}[g]\right]\mathrm{d}s. \qquad (11.44)
\end{aligned}
$$

Applying Theorem 2.1 to the above, we have

$$
\begin{aligned}
\mathtt{u}_t^{\downarrow,T}[g](x) &= \frac{1}{w_T(x)}\mathbf{E}_x\left[g(\xi_{t\wedge\mathtt{k}})w_{T-t}(\xi_{t\wedge\mathtt{k}})\exp\left(\int_0^{t\wedge\mathtt{k}}\frac{\mathtt{G}[w_{T-u}](\xi_u)}{w_{T-u}(\xi_u)}\mathrm{d}u\right)\right]\\
&\quad+\frac{1}{w_T(x)}\mathbf{E}_x\Bigg[\int_0^{t\wedge\mathtt{k}}\frac{\mathtt{G}[w_{T-s}u_{t-s}^{\downarrow,T-s}[g]](\xi_s)}{w_{T-s}(\xi_s)}w_{T-s}(\xi_s)\\
&\qquad\qquad\exp\left(\int_0^s\frac{\mathtt{G}[w_{T-u}](\xi_u)}{w_{T-u}(\xi_u)}\mathrm{d}u\right)\mathrm{d}s\Bigg]\\
&\quad-\frac{1}{w_T(x)}\mathbf{E}_x\Bigg[\int_0^{t\wedge\mathtt{k}}\frac{\mathtt{G}[w_{T-s}](\xi_s)}{w_{T-s}(\xi_s)}u_{t-s}^{\downarrow,T-s}[g](\xi_s)w_{T-s}(\xi_s)\\
&\qquad\qquad\exp\left(\int_0^s\frac{\mathtt{G}[w_{T-u}](\xi_u)}{w_{T-u}(\xi_u)}\mathrm{d}u\right)\mathrm{d}s\Bigg]\\
&= \hat{\mathtt{P}}_t^{\downarrow,T}[g](x)\\
&\quad+\int_0^t\mathtt{P}_s^{\downarrow,T}\left[\frac{\mathtt{G}_s[w_{T-s}u_{t-s}^{\downarrow,T-s}[g]]}{w_{T-s}}-\frac{\mathtt{G}_s[w_{T-s}]}{w_{T-s}}u_{t-s}^{\downarrow,T-s}[g]\right](x)\mathrm{d}s\\
&= \hat{\mathtt{P}}_t^{\downarrow,T}[g](x)+\int_0^t\mathtt{P}_s^{\downarrow,T}\left[\mathtt{G}_s^{\downarrow,T}[u_{t-s}^{\downarrow,T-s}[g]\right](x)\mathrm{d}s,
\end{aligned}
$$

where we have used the definitions (11.39) and (11.40) to obtain the last two equalities. □

Let us study, in more detail, the structure of the branching mechanism $\mathtt{G}_s^{\downarrow,T}$. Using the definition (11.40), we have, for $f\in B_1^+(E)$ and $x\in E$,

$$
\begin{aligned}
\mathtt{G}_s^{\downarrow,T}[f](x) &= \frac{1}{w_{T-s}(x)}\left[\mathtt{G}[fw_{T-s}]-f\mathtt{G}[w_{T-s}]\right](x)\\
&= \frac{1}{w_{T-s}(x)}\Bigg[\gamma(x)\mathcal{E}_x\left[\prod_{i=1}^N f(x_i)w_{T-s}(x_i)\right]\\
&\quad-\gamma(x)f(x)w_{T-s}(x)-f\mathtt{G}[w_{T-s}](x)\Bigg]\\
&= \frac{\gamma(x)}{w_{T-s}(x)}\mathcal{E}_x\left[\prod_{i=1}^N f(x_i)w_{T-s}(x_i)\right]-\left(\gamma(x)+\frac{\mathtt{G}[w_{T-s}]}{w_{T-s}}(x)\right)f(x)\\
&= \gamma^{\downarrow,T-s}(x)\left(\frac{\gamma(x)}{\gamma^{\downarrow,T-s}(x)w_{T-s}(x)}\mathcal{E}_x\left[\prod_{i=1}^N w_{T-s}(x_i)f(x_i)\right]-f(x)\right),
\end{aligned}
\tag{11.45}
$$

where

$$\gamma^{\downarrow,T-s}(x) := \gamma(x) + \frac{\mathtt{G}[w_{T-s}](x)}{w_{T-s}(x)} = \frac{\gamma(x)}{w_{T-s}(x)}\mathscr{E}_x\left[\prod_{j=1}^{N} w_{T-s}(x_j)\right]. \tag{11.46}$$

Now, returning to (11.45), using the above definition of $\gamma^{\downarrow,T-s}$, it is straightforward to show that

$$\frac{\gamma(x)}{\gamma^{\downarrow,T-s}(x)w_{T-s}(x)}\prod_{i=1}^{N} w_{T-s}(x_i) = \frac{\prod_{i=1}^{N} w_{T-s}(x_i)}{\mathscr{E}_x\left[\prod_{j=1}^{N} w_{T-s}(x_j)\right]},$$

whose expectation under $\mathscr{E}_x$ is unity. Thus we may define the following change of measure

$$\left.\frac{\mathrm{d}\mathscr{P}_x^{\downarrow,T-s}}{\mathrm{d}\mathscr{P}_x}\right|_{\sigma(N,x_1,\ldots,x_N)} = \frac{\gamma(x)}{\gamma^{\downarrow,T-s}(x)w_{T-s}(x)}\prod_{i=1}^{N} w_{T-s}(x_i). \tag{11.47}$$

Combining this with (11.45), we get the conclusion of the following corollary.

Corollary 11.1 *For $f \in B_1^+(E)$ and $s \leq T$,*

$$\mathtt{G}_s^{\downarrow,T}[f](x) = \gamma^{\downarrow,T-s}(x)\mathscr{E}_x^{\downarrow,T-s}\left[\prod_{j=1}^{N} f(x_j) - f(x)\right], \qquad x \in E,$$

where $\gamma^{\downarrow,T-s}(x)$ is defined in (11.46) and $\mathscr{P}_x^{\downarrow,T-s}$ is defined in (11.47).

Dressed ↑-Trees

In a similar spirit to the previous section, we can look at the law of our BMP, when issued from a single ancestor with mark ↑, in other words, conditioned to have a subtree that survives until time T. To this end, for $A \in \mathfrak{F}_t$, $x \in E$ and $t \leq T$, note that

$$\begin{aligned}\mathbb{P}_{\delta_x}^{\updownarrow,T}(A \mid c_\emptyset^T(0) = \uparrow) &= \frac{\mathbb{P}_{\delta_x}^{\updownarrow,T}(A;\ c_i^{T-t}(t) = \uparrow \text{ for at least one } i = 1,\ldots,N_t)}{\mathbb{P}_{\delta_x}^{\updownarrow,T}(c_\emptyset^T(0) = \uparrow)}\\ &= \frac{\mathbb{E}_{\delta_x}[\mathbf{1}_A(1 - \prod_{i=1}^{N_t} w_{T-t}(x_i(t)))]}{p_T(x)}.\end{aligned} \tag{11.48}$$

We want to describe our BMP under $\mathbb{P}_{\delta_x}^{\updownarrow,T}(\cdot \mid c_\emptyset^T(0) = \uparrow)$. In order to do so, we first need to introduce a type-↑-type-↓ BMP. Our type-↑-type-↓ BMP process, say

$X^{\updownarrow,T} = (X_t^{\updownarrow,T}, t \le T)$, has an initial ancestor that is of type $\uparrow$. We will implicitly assume (and suppress from the notation $X^{\updownarrow,T}$) that $X^{\updownarrow,T} = \delta_x$, for $x \in E$. Particles in $X^{\updownarrow,T}$ of type $\uparrow$ move as a $\mathtt{P}^{\uparrow,T}$-Markov process, for some semigroup $\mathtt{P}^{\uparrow,T}$, which we will introduce in the lemma below. When a branching event occurs for a type-$\uparrow$ particle, both type-$\uparrow$ and type-$\downarrow$ particles may be produced, but always at least one type-$\uparrow$ is produced. Type-$\uparrow$ particles may be thought of as offspring, and any additional type-$\downarrow$ particles may be thought of as immigrants. Type-$\downarrow$ particles that are created can only subsequently produce type-$\downarrow$ particles, giving rise to $\downarrow$-trees, as described above.

Lemma 11.3 (Dressed $\uparrow$-Trees) *For $x \in E$, the process $X^{\updownarrow,T}$ is equal in law to X under $\mathbb{P}_{\delta_x}(\cdot|\zeta > T)$. Moreover, both are equal in law to a dressed time-inhomogeneous BMP, say $X^{\uparrow,T}$, where the motion semigroup $\mathtt{P}^{\uparrow,T}$ corresponds to the Markov process ξ on $E\cup\{\dagger\}$ with probabilities $(\mathbf{P}_x^{\uparrow,T}, x \in E)$ given by (recalling that p is valued 0 on $\dagger$)*

$$\left.\frac{\mathrm{d}\mathbf{P}_x^{\uparrow,T}}{\mathrm{d}\mathbf{P}_x}\right|_{\sigma(\xi_s, s\le t)} = \frac{p_{T-t}(\xi_t)}{p_T(x)}\exp\left(-\int_0^t \frac{\mathtt{G}[w_{T-s}](\xi_s)}{p_{T-s}(\xi_s)}\mathrm{d}s\right), \quad t \le T, \tag{11.49}$$

and the branching mechanism at time s is given by

$$\mathtt{G}_s^{\uparrow,T}[f] = \frac{1}{p_{T-s}}\left(\mathtt{G}[p_{T-s}f + w_{T-s}] - (1-f)\mathtt{G}[w_{T-s}]\right), \qquad f \in B_1^+(E). \tag{11.50}$$

The dressing consists of additional particles, which are immigrated non-locally in space at the branch points of $X^{\uparrow,T}$, with each immigrated particle at time $s \in [0, T)$ continuing to evolve as an independent copy of $(X^{\downarrow,T-s}, \mathbb{P}^{\downarrow,T-s})$ from their respective space-time point of immigration, such that the joint branching/immigration mechanism of type-$\uparrow$ offspring and type-$\downarrow$ immigrants at time $s \in [0, T]$ is given by

$$\begin{aligned}\mathtt{G}_s^{\updownarrow,T}&[f, g](x)\\ &:= \frac{\gamma(x)}{p_{T-s}(x)}\mathcal{E}_x\Bigg[\sum_{\substack{I\subseteq\{1,\dots,N\}\\ |I|\ge 1}}\prod_{i\in I} p_{T-s}(x_i)f(x_i)\prod_{j\in\{1,\dots,N\}\setminus I} w_{T-s}(x_j)g(x_j)\Bigg]\\ &\quad - \gamma^{\updownarrow,T-s}(x)f(x)\end{aligned} \tag{11.51}$$

and

$$\gamma^{\updownarrow,T-s}(x) = \frac{\gamma(x)}{p_{T-s}(x)}\mathcal{E}_x\left[1 - \prod_{j=1}^N w_{T-s}(x_j)\right] = \gamma(x) - \frac{\mathtt{G}[w_{T-s}](x)}{p_{T-s}(x)}. \tag{11.52}$$

Proof We may think of $((x_i(t), c_i^{T-t}(t)), i \leq N_t)$, $t \leq T$, under $\mathbb{P}^{\updownarrow, T-t}$ as a two-type branching process. To this end, for $t \leq T$, let us write $N_t^{\uparrow,T} = \sum_{i=1}^{N_t} \mathbf{1}_{(c_i^{T-t}(t)=\uparrow)}$ and $N_t^{\downarrow,T} = N_t - N_t^{\uparrow,T}$. Define, for $f, g \in B_1^+(E)$,

$$\mathrm{u}_t^{\updownarrow,T}[f,g](x) = \mathbb{E}_{\delta_x}^{\updownarrow}\left[\varPi_t^T[f,g]\middle| c_\emptyset^T(0) = \uparrow\right], \qquad t \geq 0, \tag{11.53}$$

where, for $t \geq 0$,

$$\varPi_t^T[f,g] = \prod_{i=1}^{N_t^{\uparrow,T}} p_{T-t}(x_i^{\uparrow}(t)) f(x_i^{\uparrow}(t)) \prod_{j=1}^{N_t^{\downarrow,T}} w_{T-t}(x_j^{\downarrow}(t)) g(x_j^{\downarrow}(t)),$$

where

$$(x_i^{\uparrow}(t), i = 1, \cdots, N_t^{\uparrow,T}) = (x_i(t) \text{ such that } c_i^T(t) = \uparrow,\ i = 1 \cdots N_t)$$

and $(x_i^{\downarrow}(t), i = 1, \cdots, N_t^{\downarrow,T})$ is similarly defined. Clearly, there is an implicit T dependence on each of the $x_i^{\uparrow}$ and $x_j^{\downarrow}$; however, we suppress this dependency in the notation for the sake of simplicity.

We can break the expectation in the definition of $\mathrm{u}_t^{\updownarrow,T}[f,g]$ over the first branching event, noting that, until that moment occurs, the initial ancestor is necessarily of type $\uparrow$. Again remembering $p_T(\dagger) = 0$, we have

$$\begin{aligned}
&\mathrm{u}_t^{\updownarrow,T}[f,g](x) \\
&= \frac{\mathbb{E}_{\delta_x}\left[\varPi_t^T[f,g]\mathbf{1}_{(c_\emptyset^T(0)=\uparrow)}\right]}{\mathbb{P}_{\delta_x}(c_\emptyset^T(0) = \uparrow)} \\
&= \frac{1}{p_T(x)}\mathbf{E}_x\left[p_{T-t}(\xi_t) f(\xi_t) \mathrm{e}^{-\int_0^t \gamma(\xi_u)\mathrm{d}u}\right] \\
&\quad + \frac{1}{p_T(x)}\mathbf{E}_x\Bigg[\int_0^t p_{T-s}(\xi_s) \frac{\gamma(\xi_s)}{p_{T-s}(\xi_s)} \mathrm{e}^{-\int_0^s \gamma(\xi_u)\mathrm{d}u} \\
&\times \mathcal{E}_{\xi_s}\Bigg[\sum_{\substack{I \subseteq \{1,\ldots,N\} \\ |I| \geq 1}} \prod_{i \in I} p_{T-s}(x_i) \mathrm{u}_{t-s}^{\updownarrow,T-s}[f,g](x_i) \\
&\qquad \prod_{j \in \{1,\ldots,N\} \setminus I} w_{T-s}(x_j) \mathrm{u}_{t-s}^{\downarrow,T-s}[g](x_j)\Bigg] \mathrm{d}s\Bigg].
\end{aligned} \tag{11.54}$$

To help the reader interpret (11.54) better, we note that the first term on the right-hand side comes from the event that no branching occurs up to time t, in which case the initial ancestor is positioned at ξ_t. Moreover, we have used the fact that $\mathbf{P}_{\delta_x}(c^T_\emptyset(0) =\uparrow \mid \mathfrak{S}_t) = p_{T-t}(\xi_t)$. The second term is the consequence of a branching event occurring at time $s \in [0,t]$, at which point in time, the initial ancestor is positioned at ξ_s and thus has offspring scattered at $(x_i, i = 1\cdots, N)$ according to $\mathcal{P}_{\xi_s}$. The contribution thereof from time s to t can either be expressed as $\mathtt{u}^{\updownarrow,T-s}_{t-s}[f,g]$, with probability p_{T-s}, if a given offspring is of type-$\uparrow$ (thereby growing a tree of particles marked both $\uparrow$ and $\downarrow$), or be expressed as $u^{\downarrow,T-s}_{t-s}[g]$, with probability w_{T-s}, if a given offspring is of type-$\downarrow$ (thereby growing a tree of particles marked only with $\downarrow$). Hence, projecting the expectation of $\varPi^T_t[f,g]\mathbf{1}_{(c^T_\emptyset(0)=\uparrow)}$ onto the given configuration $(x_i, i = 1\cdots, N)$ at time s, we get the sum inside the expectation with respect to $\mathcal{P}_{\xi_s}$, catering for all the possible markings of the offspring of the initial ancestor, while ensuring that at least one of them is of type $\uparrow$ (which guarantees $c^T_\emptyset(0) =\uparrow$). In both expectations, the event of killing is accommodated for thanks to the fact that $p_T(\dagger) = f(\dagger) = \gamma(\dagger) = 0$.

We may now substitute the definition (11.51) into (11.54) to get

$$
\begin{aligned}
&\mathtt{u}^{\updownarrow,T}_t[f,g](x)\\
&= \frac{1}{p_T(x)}\mathbf{E}_x\left[p_{T-t}(\xi_t)f(\xi_t)\mathrm{e}^{-\int_0^t\gamma(\xi_u)\mathrm{d}u}\right]\\
&\quad+\frac{1}{p_T(x)}\mathbf{E}_x\Bigg[\int_0^t p_{T-s}(\xi_s)\frac{\gamma(\xi_s)}{p_{T-s}(\xi_s)}\mathrm{e}^{-\int_0^s\gamma(\xi_u)\mathrm{d}u}\\
&\quad\times\left[\mathtt{G}^{\updownarrow}_{T-s}[\mathtt{u}^{\updownarrow,T-s}_{t-s}[f,g], u^{\downarrow,T-s}_{t-s}[g]](\xi_s)+\gamma^{\updownarrow,T-s}(\xi_s)\mathtt{u}^{\updownarrow,T-s}_{t-s}[f,g](\xi_s)\right]\mathrm{d}s\Bigg].
\end{aligned}
\tag{11.55}
$$

Next, recall the first equality in (11.52) that for any $t \le T$, $\gamma(x) = \gamma^{\updownarrow,T-t}(x) + \mathtt{G}[w_{T-t}](x)/p_{T-t}(x)$. In each of the terms on the right-hand side of (11.55), we can exchange the exponential potential $\exp(-\int_0^\cdot \gamma(\xi_u)\mathrm{d}u)$ for the exponential potential $\exp(-\int_0^\cdot \mathtt{G}[w](\xi_u)/p(\xi_u)\mathrm{d}u)$ by transferring the difference in the exponent to an additive potential using Theorem 2.1. In this exchange, the term

$$
\gamma^{\updownarrow,T-s}(\xi_s)\mathtt{u}^{\updownarrow,T-s}_{t-s}[f,g](\xi_s)
$$

is removed as an additive potential on the right-hand side of (11.55) and replaced as a multiplicative potential. Then recalling the change of measure (11.49) that defines the semigroup $\mathtt{P}^{\uparrow}$, we get that, on E,

$$
\mathtt{u}^{\updownarrow,T}_t[f,g] = \mathtt{P}^{\uparrow,T-t}_t[f] + \int_0^t \mathtt{P}^{\uparrow,T-s}_s\left[\mathtt{G}^{\updownarrow}_{T-s}[\mathtt{u}^{\updownarrow,T-s}_{t-s}[f,g], u^{\downarrow,T-s}_{t-s}[g]]\right]\mathrm{d}s,\quad t\le T.
\tag{11.56}
$$

For the first term on the right-hand side of (11.56), there is no need to define the adjusted semigroup $\hat{\mathtt{P}}^{\uparrow,T}$ as the semigroup $\mathtt{P}^{\uparrow,T}$ is that of a conservative process on E. Indeed, the change of measure (11.49) defines the law of a conservative Markov process. To see why, note from Eq. (11.37) that can be rearranged to give

$$\begin{aligned}
p_T(x) &= 1 - w_T(x) \\
&= 1 - \hat{\mathtt{P}}_t[w_{T-t}](x) - \int_0^t \mathtt{P}_s\left[p_{T-s}\frac{\mathtt{G}[w_{T-s}]}{p_{T-s}}\right](x)\,\mathrm{d}s \\
&= \mathtt{P}_t[p_{T-t}](x) - \int_0^t \mathtt{P}_s\left[p_{T-s}\frac{\mathtt{G}[w_{T-s}]}{p_{T-s}}\right](x)\,\mathrm{d}s, \quad x \in E, \qquad (11.57)
\end{aligned}$$

where we have multiplied and divided by p_{T-s} in the integral term instead of w_{T-s}. Using Theorem 2.1, we can thus interpret 11.57 as saying

$$p_T(x) = \mathbf{E}_x\left[p_{T-t}(\xi_t)\exp\left(-\int_0^t \frac{\mathtt{G}[w_{T-s}](\xi_s)}{p_{T-s}(\xi_s)}\mathrm{d}s\right)\right], \qquad x \in E, t \in [0,T].$$

Together with the Markov property, the above identity is sufficient to show that the right-hand side of (11.49) is a unit mean martingale, rendering the interpretation of (11.49) as that of a change of measure describing a conservative process.

Returning to (11.56), we see that it is the semigroup evolution equation of a two-type BMP in which $\downarrow$-marked particles immigrate off an $\uparrow$-marked BMP. We have yet to verify, however, that the $\uparrow$-marked BMP is indeed the process described in the statement of the proposition. In order to do this, we need to show that $\mathtt{G}_s^{\uparrow,T}[f] = \mathtt{G}_s^{\updownarrow,T}[f,1]$, for all $f \in B_1^+(E)$ and $0 \le s \le T$, where $\mathtt{G}_s^{\uparrow,T}$ was given in (11.50).

Using (11.33) and the definition of $\gamma^{\updownarrow,s}$ in (11.52), for $x \in E$ and $0 \le s \le T$, we have

$$\begin{aligned}
\mathtt{G}_s^{\uparrow,T}[f](x) &:= \mathtt{G}_s^{\updownarrow,T}[f,1](x) \\
&= \frac{\gamma(x)}{p_{T-s}(x)}\mathscr{E}_x\Bigg[\sum_{\substack{I\subseteq\{1,\dots,N\}\\|I|\ge 1}}\prod_{i\in I}p_{T-s}(x_i)f(x_i)\prod_{i\in\{1,\dots,N\}\setminus I}w_{T-s}(x_i)\Bigg] \\
&\qquad - f(x)\frac{\gamma(x)}{p_{T-s}(x)}\mathscr{E}_x\left[1-\prod_{i=1}^N w_{T-s}(x_i)\right] \\
&= \frac{\gamma(x)}{p_{T-s}(x)}\mathscr{E}_x\left[\prod_{k=1}^N(p_{T-s}(x_i)f(x_i)+w_{T-s}(x_i)) - \prod_{k=1}^N w_{T-s}(x_k)\right] \\
&\qquad - f(x)\frac{1}{p_{T-s}(x)}[\gamma(x) - \mathtt{G}[w_{T-s}](x) - \gamma(x)w_{T-s}(x)]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{p_{T-s}(x)}\,[\mathtt{G}[p_{T-s}f + w_{T-s}] \\
&\qquad + \gamma(x)(p_{T-s}(x)f(x) + w_{T-s}(x)) - \mathtt{G}[w_{T-s}](x) - \gamma(x)w_{T-s}(x)] \\
&\qquad\qquad - f(x)\frac{1}{p_{T-s}(x)}[\gamma(x) - \mathtt{G}[w_{T-s}](x) - \gamma(x)w_{T-s}(x)] \\
&= \frac{1}{p_{T-s}(x)}\left\{\mathtt{G}[p_{T-s}f + w_{T-s}](x) - (1-f)\mathtt{G}[w_{T-s}](x)\right\},
\end{aligned}
$$

since $p_s(x) + w_s(x) = 1$. This completes the proof. □

As with the process $(X, \mathbb{P}^{\downarrow,T})$, we can develop the expression for the branching mechanism $\mathtt{G}_s^{\updownarrow,T}[f, g]$ further and show that it can be associated with a change of measure of the original offspring distribution.

In a similar manner to (11.32), the offspring distribution associated with the two-type tree, say $(\mathscr{P}_x^{\updownarrow,T-t}, 0 \le t \le T)$, can be defined in terms of an additional random selection from $(x_i, i = 1, \ldots, N)$ under $\mathscr{P}_x$. To this end, for $x \in E$ and $0 \le t \le T$, set

$$
\left.\frac{\mathrm{d}\mathscr{P}_x^{\updownarrow,T-t}}{\mathrm{d}\mathscr{P}_x}\right|_{\sigma(N;x_1,\ldots,x_N)}
:= \frac{\gamma(x)}{\gamma^{\updownarrow,s}(x)p_{T-t}(x)}\mathscr{E}_x\left[\sum_{\substack{I\subset\{1,\ldots,N\}\\|I|\ge 1}}\prod_{i\in I}p_{T-t}(x_i)\prod_{i\in\{1,\ldots,N\}\setminus I}w_{T-t}(x_i)\right]. \tag{11.58}
$$

Note that we require $|I| \ge 1$ in the sum since there must be at least one particle whose genealogy survives until time T by definition of $X^{\updownarrow}$. Using the definition of $\gamma^{\updownarrow,T-t}$ in (11.52) and the identity in (11.33), we have

$$
\begin{aligned}
&\frac{\gamma(x)}{\gamma^{\updownarrow,T-t}(x)p_{T-t}(x)}\mathscr{E}_x\left[\sum_{\substack{I\subset\{1,\ldots,N\}\\|I|\ge 1}}\prod_{i\in I}p_{T-t}(x_i)\prod_{i\in\{1,\ldots,N\}\setminus I}w_{T-t}(x_i)\right] \\
&\qquad = \frac{\sum_{I\subset\{1,\ldots,N\}}\mathbf{1}_{|I|\ge 1}\prod_{i\in I}p_{T-t}(x_i)\prod_{i\in\{1,\ldots,N\}\setminus I}w_{T-t}(x_i)}{1-\mathscr{E}_x[\prod_{i=1}^N w_{T-t}(x_i)]} \\
&\qquad = 1,
\end{aligned} \tag{11.59}
$$

so that (11.58) is indeed a change of measure.

Recalling (11.51), the branching mechanism of the type-↑-type-↓ process at time s can hence be identified as per the statement of the following corollary.

Corollary 11.2 *We have for $f, g \in B_1^+(E)$ and $x \in E$,*

$$\mathtt{G}_s^{\updownarrow,T}[f,g](x) = \gamma^{\updownarrow,T-s}(x)\left(\mathscr{E}_x^{\updownarrow,T-s}\left[\prod_{i=1}^{N_s^{\uparrow,T}} f(x_i^{\uparrow})\prod_{j=1}^{N_s^{\downarrow,T}} g(x_i^{\downarrow})\right] - f(x)\right), \tag{11.60}$$

where $\gamma^{\updownarrow,T-s}(x)$ is defined in (11.52), $\mathscr{P}_x^{\updownarrow,T-s}$ is defined in (11.58), and, given a branching event at time s, the set $(x_i^{\uparrow}; i = 1, \ldots, N_s^{\uparrow,T})$ denotes those particles with descendants alive at time T and $(x_i^{\downarrow}; i = 1, \ldots, N_s^{\downarrow,T})$ denotes those whose descendants die out by time T.

We can now state the main theorem of this section.

Theorem 11.2 (T-Skeletal Decomposition) *Suppose that $\mu = \sum_{i=1}^n \delta_{x_i}$, for $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in E$. Then $(X, \mathbb{P}_\mu^{\updownarrow,T})$ is equal in law to*

$$\sum_{i=1}^n \left(\Theta_i^T X_t^{\updownarrow,T}(i) + (1-\Theta_i^T)X_t^{\downarrow,T}(i)\right), \qquad t \geq 0, \tag{11.61}$$

where, for each $i = 1, \ldots, n$, Θ_i^T is an independent Bernoulli random variable with probability of success given by

$$p_T(x_i) := 1 - w_T(x_i) \tag{11.62}$$

and the processes $X^{\downarrow,T}(i)$ and $X^{\updownarrow,T}(i)$ are independent copies of $(X, \mathbb{P}_{\delta_{x_i}}^{\downarrow,T})$ and $(X, \mathbb{P}_{\delta_{x_i}}^{\updownarrow,T}(\cdot|c_\emptyset^T(0) = \uparrow))$, respectively.

Proof As previously, we may think of $((x_i(t), c_i^T(t)), i \leq N_t)$, $0 \leq t \leq T$, as a two-type branching process under $\mathbb{P}^{\updownarrow,T}$. The change of measure (11.32) gives us that, for $\mu = \sum_{i=1}^n \delta_{x_i}$ with $n \geq 1$ and $x_i \in E$, $i = 1, \ldots, n$,

$$\begin{aligned}&\mathbb{E}_\mu^{\updownarrow,T}\left[\Pi_t^T[f,g]\right]\\ &= \sum_{I\subseteq\{1,\ldots,n\}}\prod_{i\in I} p_{T-t}(x_i)\mathtt{u}_t^{\updownarrow,T-t}[f,g](x_i)\prod_{i\in\{1,\ldots,n\}\setminus I} w_{T-t}(x_i)u_t^{\downarrow,T-t}[g](x_i).\end{aligned} \tag{11.63}$$

What this shows is that the change of measure (11.32) (which we recall is, in fact, unity) is equivalent to a Doob h-transform on a two-type branching particle system (i.e., types $\{\uparrow, \downarrow\}$) where we consider the system after disregarding the marks. The

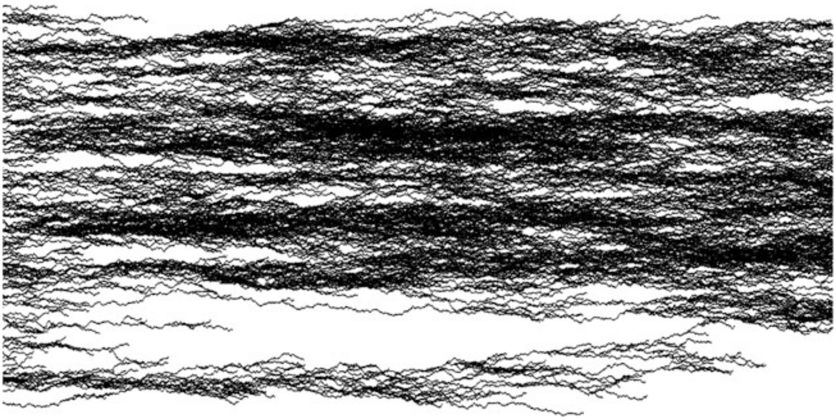

Fig. 11.1 A simulation of a branching Brownian motion in one dimension (vertical axis) with 150 initial individuals, with time running from left to right. Individuals reproduce locally and have either 0 or 2 offspring. The darker regions in the figure are indicative of a concentration of subcritical dressing, that is, the $\downarrow$-marked subtrees, immigrating off the genealogies that survive to time T (the width of the diagram), that is, the $\uparrow$-marked individuals

effect of this Doob h-transform on type-$\downarrow$ particles is that they generate $\downarrow$-trees, as described in Lemma 11.2, whereas type-$\uparrow$ particles generate a dressed $\uparrow$-tree as described in Lemma 11.3. □

As alluded to in the introduction to this section, the skeletal decomposition, Theorem 11.2, paints a picture in which a BMP can be identified as a series of thinner trees of $\uparrow$-marked individuals that survive to the prescribed time horizon T, and all other mass takes the form of $\downarrow$-subtrees that die out by time T, either rooted at time $t = 0$ or rooted on the "$\uparrow$-trees" as a dressing. The $\uparrow$-trees are generally much thinner than $\downarrow$-trees. The latter consume the bulk of the mass in the particle system. Figure 11.1 demonstrates precisely this point via a simulation of a branching Brownian motion in one dimension with time running from left to right.[1] The $\uparrow$-trees are not specifically visible in this figure, where all particles are coloured black and, in any case, are largely drowned out by the $\downarrow$-marked individuals. However, one notes that the darker streaks where there is concentration of mass appear to roughly outline the shape of a tree-like structure. Indeed, it is the dressing of short-lived $\downarrow$-marked subtrees that huddle around the long and thin (in this picture) $\uparrow$-trees, thickening them up, manifesting in the aforementioned "darker streaks".

Remark 11.1 It is an easy consequence of Theorem 11.2 and the inherent underlying Markovian structure that, for $\mu \in \mathcal{M}_c(E)$, and $t \in [0, T]$, the law of $(X_t, \mathbb{P}_\mu^{\updownarrow,T})$ conditional on $\mathfrak{F}_t = \sigma(X_s, s \leq t)$ is equal to that of a Binomial point process with

[1] This image was produced using software written by Matt Roberts. The authors are most grateful for his assistance.

intensity $p(\cdot)X_t(\cdot)$ under $\mathbb{P}_\mu$. The latter, written $\mathrm{BinPP}(pX_t)$, is an atomic random measure given by

$$\mathrm{BinPP}(p_{T-t}X_t) = \sum_{i=1}^{N_t} \mathtt{B}_i \delta_{x_i(t)},$$

where (we recall) that $X_t = \sum_{i=1}^{N_t} \delta_{x_i(t)}$ and $\mathtt{B}_i$ is a Bernoulli random variable with probability $p_{T-t}(x_i(t))$, $i = 1, \cdots, N_t$.

Remark 11.2 As alluded to at the start of this section, the reader was encouraged to question whether it is necessary in our analysis to assume that $T < \infty$. Careful inspection of the proofs reveals that the only thing that is really needed of the time horizon are the assumptions (G9) and (G10). If we replace the aforesaid assumptions simply by the exact same statements with $T = \infty$, then all of the proofs proceed verbatim. In that case, the assumptions become that

$$\inf_{x\in E} w(x) > 0 \text{ and } w(x) < 1 \text{ for } x \in E, \tag{11.64}$$

where $w(x) = \mathbb{P}_{\delta_x}(\zeta < \infty)$ is the probability of extinction. In the next chapter, we will develop more results that firm up the notion that, under the assumption (G2), the sign of λ_* dictates whether the process becomes extinct or not.

11.5 Spine Emerging from the Skeleton at Criticality

What is the relationship between the spine and the T-skeletal decomposition? In the setting of a critical BMP, the two can be coupled by considering what happens when we condition a BMP to survive as $T \to \infty$. In terms of the skeleton, this corresponds to conditioning the number of $\uparrow$-marked initial ancestors to be at least one as $T \to \infty$.

Let us thus put ourselves into the critical setting, and suppose that the assumptions of Theorem 10.1 are henceforth at our disposal. In that case, we also then have at our disposal that, uniformly for $x \in E$,

$$\lim_{T\to\infty} T p_T(x) = \frac{\varphi(x)}{\Sigma}, \tag{11.65}$$

where we recall that φ is the right eigenfunction in (G2) and Σ is the constant given by (10.5).

Now consider the setup in Theorem 11.2, namely that we have a BMP $(X, \mathbb{P}_\mu)$, where $\mu \in \mathscr{M}_c(E)$ taking the form $\mu = \sum_{i=1}^n \delta_{x_i}$, with $n \in \mathbb{N}$ and $x_i \in E$. We start by looking at the number of individuals in the skeletal decomposition at time $t = 0$ that are marked by $\uparrow$. To this end, let us further assume the following is in force:

(G11) $\lim_{T\to\infty}\inf_{x\in E} w_T(x) = 1$.

Note that (G11) implies (G8). By appealing to Theorem 11.2, we have for $k = 1, \cdots, n$,

$$\lim_{T\to\infty} \mathbb{P}_\mu^{\updownarrow,T}(N^\uparrow = k | N^\uparrow \geq 1)$$
$$= \lim_{T\to\infty} \frac{\sum_{I\subset\{1,\cdots,n\},|I|=k}\prod_{j\in I} p_T(x_j)\prod_{\ell\in\{1,\cdots,n\}\setminus I} w_T(x_\ell)}{1-\prod_{i=1}^n w_T(x_i)}, \tag{11.66}$$

where $N^\uparrow$ is the number of individuals at time $t = 0$ marked by $\uparrow$ and the sum is taken over all subsets of size k of $\{1, \cdots, n\}$. In words, (11.66) gives the asymptotic probability as $T \to \infty$ that, conditional on X surviving to time T, there are k individuals at time $t = 0$ that carry the mark $\uparrow$.

We note that if $k = 1$ in (11.66), then we claim that

$$\lim_{T\to\infty} \mathbb{P}_\mu^{\updownarrow,T}(N^\uparrow = 1 | N^\uparrow \geq 1) = \lim_{T\to\infty} \frac{\sum_{j=1}^n p_T(x_j)\prod_{\ell\neq j} w_T(x_\ell)}{1-\prod_{i=1}^n w_T(x_i)} = 1. \tag{11.67}$$

To see why, consider the mapping $(x_1, \cdots, x_n) \mapsto \prod_{i=1}^n x_i$ around $(0, \cdots, 0)$ on $[0, 1]^n$. The multi-dimensional Taylor's formula tells us that

$$\prod_{i=1}^n x_i = 1 + \sum_{j=1}^n (x_j - 1) + O(\sum_{i<j} x_i x_j). \tag{11.68}$$

Using (11.68) in (11.67), we have

$$\lim_{T\to\infty} \mathbb{P}_\mu^{\updownarrow,T}(N^\uparrow = 1 | N^\uparrow \geq 1) = \lim_{T\to\infty} \frac{\sum_{j=1}^n p_T(x_j)\prod_{\ell\neq j} w_T(x_\ell)}{\sum_{j=1}^n p_T(x_j) + O(\sum_{j<i} p_T(x_j)p_T(x_i))}. \tag{11.69}$$

Dividing and multiplying the denominator and numerator by T in (11.69), by appealing to (11.65) and noting that $Tp_T(x_i)p_T(x_j) \to 0$ as $T \to \infty$, (11.67) follows.

The proof of (11.67) in fact shows something more particular. Indeed, now it is clear that the number of $\uparrow$-marked initial ancestors concentrates on one individual conditionally on survival as $T \to \infty$, if we label the index of that one individual by i^*, then we see that

$$\lim_{T\to\infty} \mathbb{P}_\mu^{\updownarrow,T}(i^* = k | N^\uparrow \geq 1) = \lim_{T\to\infty} \frac{p_T(x_k)\prod_{\ell\neq k} w_T(x_\ell)}{\sum_{j=1}^n p_T(x_j) + O(\sum_{j<i} p_T(x_j)p_T(x_i))}$$

$$= \frac{\varphi(x_k)}{\sum_{i=1}^n \varphi(x_i)},$$

which again follows by dividing and multiplying by T before taking limits in the first equality and using (11.65) to obtain the right-hand side.

The limit (11.67) tells us that, for BMPs that become extinct under the assumptions we have made, conditional on survival, as $T \to \infty$, the skeleton concentrates on a single individual dressed tree rooted at $t = 0$, which carries all the $\uparrow$-marked individuals in its tree of descendants.

Now consider the change of measure (11.39) that defined $\mathbf{P}^{\downarrow,T}$. Thanks to (G11), we see that $\mathbf{P}^{\downarrow,T}$ converges weakly to $\mathbf{P}$. Similarly, considering (11.49) and the change of measure that defines $\mathbf{P}^{\uparrow,T}$, we have, for $x \in E$,

$$\frac{\mathtt{G}[1-p_T](x)}{p_T(x)} = \frac{\mathtt{A}[p_T](x) - \frac{1}{2}\gamma(x)\mathscr{V}[p_T](x) + \frac{1}{2}\gamma(x)\mathscr{V}[p_T](x)}{p_T(x)} - \frac{\gamma(x)\mathscr{E}_x\left[\sum_{i=1}^N p_T(x_i) - p_T(x)\right]}{p_T(x)}, \qquad (11.70)$$

where we have used the notation from (10.14) and Lemma 10.2. Dividing and multiplying by T again and then taking $T \to \infty$, we can use the estimates in Lemma 10.2 and the fact that

$$\lim_{T\to\infty} \frac{p_{T-t}(y)}{p_T(x)} = \frac{\varphi(y)}{\varphi(x)},$$

uniformly for $x, y \in E$ to deduce that, due to Lemma 10.5, we have

$$\lim_{T\to\infty} \frac{\mathtt{G}[1-p_T](x)}{p_T(x)} = -\gamma(x)\frac{\mathtt{m}[\varphi](x) - \varphi(x)}{\varphi(x)}, \qquad (11.71)$$

uniformly for $x \in E$. Note that the product of two or more probabilities p_{T-t} will tend to zero, e.g., in the term $\mathscr{V}[p_T](x)$, even with the help of a multiplying factor T, which explains why the first factor on the right-hand side of (11.70) goes to zero. Hence, recalling the definition (11.4), we see that $\mathbf{P}_x^{\downarrow,T}$ converges weakly to the law of the process with associated semigroup $\mathtt{P}^{\varphi}$.

Now consider the branching mechanisms $\mathtt{G}^{\downarrow,T}$ and $\mathtt{G}^{\updownarrow,T}$, defined in (11.40) and (11.60), respectively, as $T \to \infty$. In the case of $\mathtt{G}^{\downarrow,T}$, it is straightforward to see that

$$\lim_{T\to\infty} \mathtt{G}_s^{\downarrow,T}[f](x) = \mathtt{G}[f](x),$$

uniformly, for $x \in E$ and $f \in B_1^+(E)$. Moreover, for $\mathtt{G}^{\updownarrow,T}$, we note from the representation in (11.51) and (11.52) that, with the help of (11.71),

$$\lim_{T\to\infty} \gamma^{\updownarrow,T}(x) = \gamma(x)\frac{\mathtt{m}[\varphi](x)}{\varphi(x)}, \tag{11.72}$$

uniformly in $x \in E$. Moreover, again by dividing and multiplying by T, we get, uniformly for $x \in E$ and $f, g \in B_1^+(E)$,

$$\begin{aligned}
&\lim_{T\to\infty} \mathtt{G}_s^{\updownarrow,T}[f,g](x) \\
&:= \lim_{T\to\infty} \frac{\gamma(x)}{p_{T-s}(x)}\mathcal{E}_x\Bigg[\sum_{\substack{I\subseteq\{1,\dots,N\}\\|I|\geq 1}} \prod_{i\in I} p_{T-s}(x_i)f(x_i) \prod_{j\in\{1,\dots,N\}\setminus I} w_{T-s}(x_i)g(x_j)\Bigg] \\
&\qquad\qquad - \gamma(x)\frac{\mathtt{m}[\varphi](x)}{\varphi(x)}f(x) \\
&= \lim_{T\to\infty} \gamma(x)\mathcal{E}_x\Bigg[\sum_{i=1}^N \frac{Tp_{T-s}(x_i)}{Tp_{T-s}(x)} f(x_i)\prod_{j\neq i} g(x_j)\Bigg] \\
&\quad + \lim_{T\to\infty} \gamma(x)\mathcal{E}_x\Bigg[\sum_{\substack{I\subseteq\{1,\dots,N\}\\|I|\geq 2}} T\prod_{i\in I}\frac{p_{T-s}(x_i)f(x_i)}{Tp_{T-s}(x)} \prod_{j\in\{1,\dots,N\}\setminus I} g(x_j)\Bigg] \\
&- \gamma(x)\frac{\mathtt{m}[\varphi](x)}{\varphi(x)}f(x) \\
&= \gamma(x)\mathcal{E}_x\Bigg[\sum_{i=1}^N \frac{\varphi(x_i)}{\varphi(x)} f(x_i)\prod_{j\neq i} g(x_j)\Bigg] - \gamma(x)\frac{\mathtt{m}[\varphi](x)}{\varphi(x)}f(x).
\end{aligned}$$

Recalling the expressions (11.5) and (11.6), we conclude that, uniformly for $x \in E$ and $f, g \in B_1^+(E)$,

$$\lim_{T\to\infty} \mathtt{G}_s^{\updownarrow,T}[f,g](x) = \gamma^{\varphi}(x)\mathcal{E}_x\Bigg[\frac{\mathsf{Z}[\varphi]}{\mathtt{m}[\varphi](x)}\sum_{i=1}^N \frac{\varphi(x_i)}{\mathsf{Z}[\varphi]} f(x_i)\prod_{j\neq i} g(x_j)\Bigg] - \gamma^{\varphi}(x)f(x).$$

This indicates that (conditionally on survival as $T \to \infty$) an $\uparrow$-marked individual will branch at rate γ^{φ}, at which point it generates a copy of Z under $\mathcal{P}^{\varphi}$ (defined via the change of measure (11.6)), and one individual is chosen to continue as a $\uparrow$-marked process and the remaining individuals are assigned $\downarrow$-marks.

Putting all these pieces together, in the conditional limiting sense, as $T \to \infty$, we see that there is concentration in distribution on one initial ancestor with an $\uparrow$-mark such that this is the i-th initial ancestor with probability proportional to $\varphi(x_i)$, and all other initial ancestors take the mark $\downarrow$. Moreover, the rate at which more than one $\uparrow$-marked individual appears at any branching event drops to zero. At the same time, the law of the motion and the branching mechanism for $\downarrow$-marked particles settle down to those of $(X, \mathbb{P})$, and the law of the motion of the single genealogy of the $\uparrow$-marked particles and the branching mechanism for its dressing settle down to that of the spine.

On a final note, we mention that, in spite of the multiple asymptotic calculations above, the collapse of the skeleton to the spine on conditioning to survive to T and letting $T \to \infty$ is in fact easy to see as inevitable. Under the assumptions of Theorem 10.1 together with (G11), for any $f \in B^+(E)$ and $\mu \in \mathcal{M}_c(E)$, the Markov property and dominated convergence tell us that

$$\begin{aligned}\lim_{T\to\infty} \mathbb{E}^{\updownarrow,T}_{\mu}\left[\mathrm{e}^{-X_t[f]}\middle| N^{\uparrow} \geq 1\right] &= \lim_{T\to\infty} \mathbb{E}_{\mu}\left[\mathrm{e}^{-X_t[f]}\middle| T < \zeta\right] \\ &= \mathbb{E}_{\mu}\left[\frac{\mathbb{P}_{X_t}(T-t<\zeta)}{\mathbb{P}_{\mu}(T<\zeta)}\mathrm{e}^{-X_t[f]}\right] \\ &= \mathbb{E}_{\mu}\left[\frac{X_t[\varphi]}{\mu[\varphi]}\mathrm{e}^{-X_t[f]}\right].\end{aligned}$$

Hence, the asymptotic conditioning corresponds to the martingale change of measure that induces the spine decomposition.

11.6 Comments

The idea of a spine decomposition is a natural progression from the many-to-one formula, giving a path decomposition rather than a distributional identity. Early spine decompositions in the setting of BMPs are found in the work of Chauvin et al. [23], although they indicate the connection of their work to Palm measures for branching processes studied in numerous earlier works. A deeper understanding of spines for branching random walks and their relation to martingale changes of measure was uncovered by Lyons [97], leading to a sequence of subsequent articles for a variety of other spatial branching models [17, 87, 104]. In parallel, the notion of a spine decomposition emerged in the superprocess literature, see, e.g., Roelly-Coppoletta and Rouault [115] and Evans [59] as well as [56, 58, 114]. Spine decompositions have increasingly become the main tool to prove results concerning the growth and spread of spatial branching processes as well as fragmentation and growth fragmentation processes; cf. [1, 2, 11, 14, 15, 53, 58, 124, 125].

In this respect, Theorem 11.1 and Lemma 11.1 are but another addition to a very robust theory of spine decompositions, with the specialised feature being that of non-local reproduction. Their proofs can be seen as taking influence from many if not all of the aforementioned literature. The non-local nature of these results is largely inspired by the proofs from the NBP setting in Horton et al. [74].

Skeletal decompositions are also well explored in a variety of settings for branching processes as well as superprocesses. See, for example, [10, 51, 56, 60, 64, 68, 88]. The setup in Sect. 11.4 reproduces results given in Harris et al. [69] for non-local branching Markov processes, albeit for a finite time horizon. Addressing the additional complexities due to working over a finite time horizon is inspired by the finite time skeletal decomposition given in Etheridge and Williams [56]. The idea that a spine emerges from the skeletal decomposition in the critical setting by conditioning on survival to a time horizon T as $T \to \infty$ also comes from Etheridge and Williams [56]. The robustness of this concept has also been demonstrated in Harris et al. [68] for branching Brownian motion in a bounded domain and Fekete et al. [63] for continuous-state branching processes.

Chapter 12
Martingale Convergence and Laws of Large Numbers

As usual, we will work in the setting that our branching Markov process, $(X, \mathbb{P})$, belongs to the Asmussen–Hering class, that is to say, (G2) is satisfied. In Chap. 11, we introduced the process

$$W_t = \mathrm{e}^{-\lambda_* t} \frac{X_t[\varphi]}{\mu[\varphi]}, \qquad t \geq 0,$$

which is a martingale under $\mathbb{P}_\mu$, $\mu \in \mathcal{M}_c(E)$. As remarked earlier, this is a non-negative martingale and hence has an almost sure limit, $W_\infty := \lim_{t\to\infty} W_t$, thanks to the classical Martingale convergence theorem. In this chapter, we explore the important relationship between the limit of this martingale and the survival set of X. Moreover, in the setting of supercriticality, we revisit the relationship between the aforesaid martingale and, whenever it is well defined, the limit as $t \to \infty$ of the more general path functional

$$\mathrm{e}^{-\lambda_* t} X_t[f], \qquad t \geq 0,\ f \in B^+(E), \tag{12.1}$$

and its relationship to the martingale limit W_∞.

As we have seen in the setting of the NBP (cf. Theorem 6.3), the reason why the limit of this path functional is of such interest is that, under the assumption (G2), its limit should mimic, in the pathwise sense, the behaviour of the discounted expectation semigroup $(\psi_t, t \geq 0)$. Indeed, $\mathbb{E}_{\delta_x}[\mathrm{e}^{-\lambda_* t} X_t[f]] = \mathrm{e}^{-\lambda_* t}\psi_t[f](x)$, $x \in E, t \geq 0$.

As a first step towards this goal, we first address the relationship between the convergence of the martingale (12.1) to a non-trivial limit and its relationship to the survival event.

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8_12

12.1 Martingale Convergence and Survival

Recall from Chap. 11 that, under the assumptions (G1) and (G2), we can use $W = (W_t, t \geq 0)$ to define a change of measure

$$\left.\frac{\mathrm{d}\mathbb{P}^{\varphi}_{\mu}}{\mathrm{d}\mathbb{P}_{\mu}}\right|_{\mathfrak{F}_t} = W_t, \qquad t \geq 0, \tag{12.2}$$

and, moreover, that $(X, \mathbb{P}^{\varphi})$ enjoys a spine decomposition, which will be of use throughout this chapter. The following result provides a simple dichotomy between convergence to a trivial vs. non-trivial limit accordingly as $\lambda_* \leq 0$ vs. $\lambda_* > 0$.

Theorem 12.1 *Under the assumption (G2) and the existence of second moments as in (G2) for $k = 2$, we have the following cases for the martingale W:*

(i) If $\lambda_* > 0$, then W is $L^2(\mathbb{P})$ convergent (and hence has a non-trivial limit).
(ii) If $\lambda_* < 0$, then $W_\infty = 0$-almost surely.
(iii) If $\lambda_* = 0$, and additionally (G8) holds, then $W_\infty = 0$-almost surely.

Proof From (12.2), by considering the inverse of the change of measure, we see that $1/W$ is a non-negative $\mathbb{P}^{\varphi}$-supermartingale. Hence, since $1/W$ is non-negative, its limit automatically exists thanks to the Martingale convergence theorem. Standard measure theory for martingale change of measures also tells us that for a measurable set, A, on the ambient measurable space, we have, for each $\mu \in \mathcal{M}_c(E)$,

$$\mathbb{P}_\mu(A) = \mathbb{E}^{\varphi}_{\mu}[W_\infty \mathbf{1}_{(A\cap\{W_\infty<\infty\})}] + \mathbb{P}^{\varphi}_{\mu}(A \cap \{W_\infty = \infty\}).$$

The consequence of this decomposition is that the measures $\mathbb{P}_\mu$ and $\mathbb{P}^{\varphi}_{\mu}$ are orthogonal on the ambient measurable space if and only if

$$\mathbb{P}^{\varphi}_{\mu}(W_\infty = \infty) = 1, \tag{12.3}$$

in which case, the martingale limit satisfies $\mathbb{P}_\mu(W_\infty = 0) = 1$. Moreover, $\mathbb{P}_\mu$ is absolutely continuous with respect to $\mathbb{P}^{\varphi}_{\mu}$ if and only if

$$\mathbb{P}^{\varphi}_{\mu}(W_\infty < \infty) = 1, \tag{12.4}$$

in which case we have $\mathbb{E}_\mu[W_\infty] = 1$, and hence, the martingale experiences $L^1(\mathbb{P}_\mu)$ convergence. This forms the basis of the proof in all three cases.

(i) Let us first deal with the case that $\lambda_* > 0$. From Theorem 9.2, we have, using Fatou's Lemma and (12.2), that

$$\mathbb{E}^{\varphi}_{\mu}\left[\liminf_{t\to\infty} W_t\right] \leq \liminf_{t\to\infty} \mathbb{E}^{\varphi}_{\mu}[W_t] \leq \lim_{t\to\infty} \mathbb{E}_\mu[W_t^2] < \infty. \tag{12.5}$$

The expectation in (12.5) ensures that $\mathbb{P}_\mu(\liminf_{t\to\infty} W_t < \infty) = 1$, and hence, since the limit exists, $\lim_{t\to\infty} W_t$ is $\mathbb{P}^{\varphi}_\mu$-almost surely finite. Combined with (12.4), this will tell us that W is an $L^1(\mathbb{P})$-martingale.

Despite now knowing that W is an $L^1(\mathbb{P})$-martingale, we still need to prove the stronger result that it is an $L^2(\mathbb{P})$-martingale. This is nonetheless a straightforward conclusion from (12.5) and Doob's martingale inequality. Indeed, let us write $\widetilde{W}_t = W_t - 1$, $t \geq 0$, so that $\widetilde{W}$ is a zero-mean martingale with limit $\widetilde{W}_\infty = W_\infty - 1$. As in (12.5), Theorem 9.2 gives us that $\lim_{t\to\infty} \mathbb{E}_\mu[\widetilde{W}_t^2] < \infty$. By Doob's martingale inequality, it follows that $\mathbb{E}_\mu[\sup_{s\geq 0} W_s^2] < \infty$; hence, dominated convergence implies that $\lim_{t\to\infty} \mathbb{E}_\mu[\widetilde{W}_t^2] = \mathbb{E}_\mu[\widetilde{W}_\infty^2] < \infty$. Jensen's inequality tells us that $(\widetilde{W}_t^2, t \geq 0)$ is a submartingale, so that, in fact, $\mathbb{E}_\mu[\widetilde{W}_t^2] \uparrow \mathbb{E}_\mu[\widetilde{W}_\infty^2]$. We can now write

$$\lim_{t\to\infty} \mathbb{E}_\mu[(W_t - W_\infty)^2] = \lim_{t\to\infty} \mathbb{E}_\mu[(\widetilde{W}_t - \widetilde{W}_\infty)^2] = \mathbb{E}_\mu[\widetilde{W}_\infty^2] - \lim_{t\to\infty} \mathbb{E}_\mu[\widetilde{W}_t^2] = 0,$$

and the required $L^2(\mathbb{P})$ convergence holds.

(ii) Next, for the case $\lambda_* < 0$, it is easy to see that

$$W_t \geq \mathrm{e}^{-\lambda_* t}\varphi(\xi^{\varphi}_t), \qquad t \geq 0,$$

where ξ^φ is the spine given in the decomposition of Sect. 11.1. The boundedness of φ, the stationary behaviour of the spine $\mathbb{P}^\varphi$ (cf. Lemma 11.1), and the strict negativity of λ_* ensure that $\mathbb{P}^{\varphi}_\mu(\limsup_{t\to\infty} W_t = \infty) = 1$ for all $\mu \in \mathscr{M}_c(E)$, and hence, from (12.3), $\mathbb{P}_\mu(W_\infty = 0) = 1$.

(iii) Finally, for the case $\lambda_* = 0$, we recall from Lemma 10.1 that, under our assumptions, $\mathbb{P}_\mu(\zeta < \infty) = 1$, where $\zeta = \inf\{t > 0 : X_t[1] = 0\}$ and $\mu \in \mathscr{M}_c(E)$. On account of the trivial fact that $\{\zeta < \infty\} \subseteq \{W_\infty = 0\}$, the desired conclusion follows. □

It is perhaps worth noting that the proof of (iii) could equally serve as the proof of (ii) albeit that we have not yet established the seemingly obvious fact that if extinction is almost sure in the critical setting (with the assumptions (G2), (G8)), then it should also be almost sure in the subcritical setting too. In fact, the next section will tidy up these concurrences.

Theorem 12.2 *Under the assumptions (G2), (G6), and (G8), we also have that the events $\{W_\infty = 0\}$ and $\{\zeta < \infty\}$-almost surely agree under $\mathbb{P}$, where $\zeta = \inf\{t > 0 : X_t[1] = 0\}$ is the time of extinction of the BMP. In particular, there is almost sure extinction if and only if $\lambda_* \leq 0$.*

Before proving Theorem 12.2, let us first extract a result that lies implicitly in the proof of Lemma 10.1, specifically from the derivation of (10.9).

Lemma 12.1 *Assume (G2) and (G8), then*

$$\lim_{t\to\infty} X_t[\varphi] = \infty \quad \text{on} \quad \{\zeta = \infty\}. \tag{12.6}$$

Proof The proof of the above fact can be found in Lemma 10.1 in the critical case; however, the proof does not really require $\lambda_* = 0$. Indeed, the only adjustment needed is the estimate (10.10), which has an obvious replacement. The rest of the proof goes through verbatim. We leave this as an exercise for the reader to check. □

Proof *(of Theorem 12.2)* Recall the trivial observation that $\{\zeta < \infty\} \subseteq \{W_\infty = 0\}$, and hence,

$$\mathbb{P}_\mu(\zeta < \infty) \leq \mathbb{P}_\mu(W_\infty = 0), \tag{12.7}$$

for all $\mu \in \mathcal{M}_c(E)$. It thus suffices to show that (12.7) holds with the inequality reversed.

Taking note of Lemma 12.1, when $\lambda_* \leq 0$, since $W_t = \mathrm{e}^{-\lambda_* t} X_t[\varphi] \geq X_t[\varphi]$, the conclusion of Theorem 12.1 immediately implies that $\{W_\infty = 0\} \subseteq \{\zeta < \infty\}$. This gives us the inequality the other way around to (12.7).

Next we consider the setting that $\lambda_* > 0$. Due to our assumptions and the boundedness of φ, cf. (G2), and boundedness of the number of offspring (G6), we have uniformly, for all times t such that there is a discontinuity in W, that $|W_t - W_{t-}|$ is uniformly bounded by some constant $M > 0$, $\mathbb{P}_{\delta_x}$-almost surely, $x \in E$. Defining the stopping time $T_1 = \inf\{t \geq 0, W_t \geq 1\}$, using the fact that W is a non-negative, $L^2(\mathbb{P})$-martingale, and hence uniformly integrable, by appealing to Doob's optional stopping theorem, we deduce that, for all $x \in E$,

$$1 = \mathbb{E}_{\delta_x}[W_{T_1}] = \mathbb{E}_{\delta_x}[W_{T_1}\mathbf{1}_{\{W_\infty > 0\}}] \leq \frac{1}{\varphi(x)}\mathbb{E}_{\delta_x}((1+M)\mathbf{1}_{\{W_\infty > 0\}}).$$

It follows that, for all $x \in E$, $\mathbb{P}_{\delta_x}(W_\infty > 0) \geq \varphi(x)/(1+M)$. Hence, that there exists $c_1 > 0$ such that

$$\mathbb{P}_{\delta_x}(W_\infty = 0) \leq 1 - c_1\varphi(x), \qquad x \in E.$$

Now, using the branching property, we obtain for all $\mu = \sum_{i=1}^n \delta_{x_i} \in \mathcal{M}_c(E)$,

$$\log \mathbb{P}_\mu(W_\infty = 0) \leq \sum_{i=1}^n \log(1 - c_1\varphi(x_i)) \leq -c_1 \sum_{i=1}^n \varphi(x_i) = -c_1\mu[\varphi].$$

From (12.6), we have, on $\{\zeta = \infty\}$,

$$\limsup_{n\to\infty} \log \mathbb{P}_{X_n}(W_\infty = 0) \leq -c_1 \lim_{n\to\infty} X_n[\varphi] = -\infty.$$

With the upper bound of any probability being unity, we can thus write

$$\limsup_{n\to\infty} \mathbb{P}_{X_n}(W_\infty = 0) \leq \mathbf{1}_{\{\zeta < \infty\}}.$$

The branching Markov property now entails

$$\mathbb{P}_\mu(W_\infty = 0) = \mathbb{E}_\mu\left[\mathbb{P}_{X_n}(W_\infty = 0)\right], \quad \forall n \in \mathbb{N}, \mu \in \mathcal{M}_c(E).$$

The reverse Fatou's lemma now gives us

$$\begin{aligned}
\mathbb{P}_\mu(W_\infty = 0) &= \limsup_{n\to\infty} \mathbb{E}_\mu\left[\mathbb{P}_{X_n}(W_\infty = 0)\right] \\
&\le \mathbb{E}_\mu\left[\limsup_{n\to\infty} \mathbb{P}_{X_n}(W_\infty = 0)\right] \\
&\le \mathbb{E}_\mu\left[\mathbf{1}_{\{\zeta < \infty\}}\right] \\
&= \mathbb{P}_\mu(\zeta < \infty).
\end{aligned}$$

Together with (12.7), this completes the proof of the theorem. □

12.2 Strong Law of Large Numbers

The following fundamental result provides a stochastic analogue to (G2). It also shows that just as the leading order behaviour of the linear semigroup is described by the eigentriplet $(\lambda_*, \varphi, \tilde{\varphi})$, the leading order stochastic behaviour of linear functionals is analogously described by the same eigentriple, albeit via the martingale W.

Theorem 12.3 *Suppose $\lambda_* > 0$, (G2) holds, as well as second moments in the form of (G2) for $k = 2$. Suppose, moreover that, for all open Ω compactly embedded subsets of E,*

$$\liminf_{t\to 0} \mathtt{P}_t^{\Omega}[1](x) \ge \mathbf{1}_\Omega(x), \tag{12.8}$$

where $(\mathtt{P}_t^{\Omega}, t \ge 0)$ is the movement semigroup killed on exiting Ω, for $(X, \mathbb{P})$. Then, for any $g \in B^+(E)$ such that, $g/\varphi \in B^+(E)$ and initial configuration $\mu \in \mathcal{M}_c(E)$,

$$\mathrm{e}^{-\lambda_* t}\frac{X_t[g]}{\mu[\varphi]} \to \tilde{\varphi}[g]W_\infty, \tag{12.9}$$

$\mathbb{P}_\mu$-almost surely.

We think of this result as a strong law of large numbers because, combining it with the conclusion of Theorem 12.1, it infers that, almost surely, on $\{\zeta < \infty\}$,

$$\lim_{t\to\infty} \frac{\sum_{i=1}^{N_t} g(x_i(t))\varphi(x_i(t))}{\sum_{i=1}^{N_t} \varphi(x_i(t))} = \lim_{t\to\infty} \frac{X_t[\varphi g]}{X_t[\varphi]} = \tilde{\varphi}[\varphi g].$$

With weaker assumptions than Theorem 12.3 slightly, we can also recover the same strong law of large numbers, albeit in the sense of L^2 convergence.

Corollary 12.1 *Suppose $\lambda_* > 0$, (G2) holds, as well as second moments in the form of (G2) for $k = 2$. For any $g \in B^+(E)$ and $\mu \in \mathscr{M}_c(E)$, the limit (12.29) holds in the $L^2(\mathbb{P}_\mu)$ sense.*

Let us consider a number of examples for which this and other conditions can be verified for Theorem 12.3.

Branching Feller Processes The assumption (12.8) is somewhat unusual but is automatically satisfied for any process with paths that are right continuous. Indeed, for such a setting, one appeals directly to Fatou's lemma to conclude that

$$\liminf_{t\to 0} \mathtt{P}_t^{\Omega}[1](x) = \liminf_{t\to 0} \mathbf{E}_x[\mathbf{1}_\Omega(\xi_t)] \geq \mathbf{E}_x[\liminf_{t\to 0} \mathbf{1}_\Omega(\xi_t)] = \mathbf{1}_\Omega(x),$$

for $x \in E$.

Recall from the end of Sect. 2.3, the definition of a Feller process. Feller processes can be constructed with right-continuous paths, and hence, (12.8) is also automatically satisfied for branching Feller processes. This would include, for example, the setting of branching Brownian motion in a compact domain. Importantly, the Feller property is not a necessary requirement as the next example shows.

Neutron Branching Process The setting of the NBP in Theorem 6.3 is automatically covered by Theorem 12.3. Indeed, if we go back to the setting of the NBP on $D \times V$, by thinking of scatter events as a special case of branching (in which precisely one offspring is produced), the process $(\xi, \mathbf{P})$ is nothing more than a linear drift. More precisely, under $\mathbf{P}_{(r,\upsilon)}$, $\xi_t = r + \upsilon t$ for $t < \kappa^D_{r,\upsilon}$, after which time it is sent to a cemetery state. Similarly, it is sent to a cemetery state if neutron absorption occurs in D (i.e., a fission event with no offspring). For Ω open and compactly embedded in $D \times V$, it is trivial to see that (12.8) holds. It is notable that $(\xi, \mathbf{P})$ is not a Feller process because its semigroup can easily fail to be continuous in the variable υ. To see why, consider Fig. 12.1, and the behaviour of $\mathtt{P}_t[g](r,\upsilon) = \mathbf{E}_{(r,\upsilon)}[g(r + \upsilon t, \upsilon)\mathbf{1}_{(t<\kappa^D_{r,\upsilon})}]$ as the velocity moves from υ' to υ''.

Multi-type Branching Process In Sect. 11.2, we considered the most basic of non-local branching Markov processes. That is, individuals have no associated motion and have type belonging to $\{1, \cdots, n\}$. Because there is no movement and the type space is so basic, the condition (12.8) is automatically satisfied. As before, we think of $(X_t, t \geq 0)$ as a vector $X_t = (X_t(1), \cdots, X_t(n))$, where $X_t(i)$ simply counts the number of individuals of type i alive at time $t \geq 0$. The conclusion of Theorem 12.3 can be re-worded as showing that the number of individuals of each type grows at the same rate. More precisely,

$$\lim_{t\to\infty} \mathrm{e}^{-\lambda_* t} X_t(i) = \tilde{\varphi}_i \varphi_j W_\infty, \qquad \mathbb{P}_{\delta_j} - \text{a.s.},$$

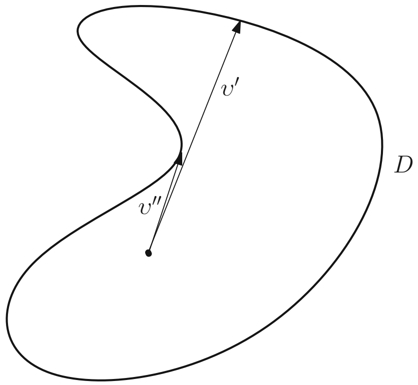

Fig. 12.1 For the given domain D, it is easy to see that $\kappa^D_{r,\upsilon}$ is not continuous in the variable $\upsilon \in V$, and hence, neither is $\mathtt{P}_t[g](r, \upsilon)$. This implies that $\mathtt{P}$ is not necessarily a Feller semigroup

where (G2) can be understood as the existence of right and left eigenvectors, say $\varphi = (\varphi_1, \cdots, \varphi_n)$ and $\tilde{\varphi} = (\tilde{\varphi}_1, \cdots, \tilde{\varphi}_n)$ with eigenvalue $\lambda_* \in \mathbb{R}$ for the mean semigroup matrix $\psi_t(i, j) = \mathbb{E}_i[X_t(j)]$, $i, j \in \{1, \cdots, n\}$.

12.3 Proof of the Strong Law of Large Numbers

We prove Theorem 12.3 by breaking it up into several parts. Starting with the following lemma, we first prove that Theorem 12.3 holds along lattice times. Just before we state the next lemma, let us quickly recall the filtration $\mathfrak{F}_t = \sigma(X_s : s \leq t)$, for $t \geq 0$.

Lemma 12.2 *Fix $\delta > 0$, and assume that (G2) is in force together with the assumption of second moments, (G2) for $k = 2$, and $\lambda_* > 0$. For $g \in B^+(E)$, define*

$$U_t = \mathrm{e}^{-\lambda_* t} X_t[g], \quad t \geq 0. \tag{12.10}$$

Then, for any non-decreasing sequence $(m_n)_{n \geq 0}$ with $m_0 > 0$ and $x \in E$,

$$\lim_{n \to \infty} |U_{(m_n+n)\delta} - \mathbb{E}[U_{(m_n+n)\delta} | \mathfrak{F}_{n\delta}]| = 0, \quad \mathbb{P}_{\delta_x}\text{-a.s.}. \tag{12.11}$$

Proof Thanks to the Borel–Cantelli lemma, it is sufficient to prove that for each $x \in E$ and all $\varepsilon > 0$,

$$\sum_{n \geq 1} \mathbb{P}_{\delta_x} \left(\left| U_{(m_n+n)\delta} - \mathbb{E}[U_{(m_n+n)\delta} | \mathfrak{F}_{n\delta}] \right| > \varepsilon \right) < \infty. \tag{12.12}$$

To this end, note that Markov's inequality gives

$$\mathbb{P}_{\delta_x}\left(\left|U_{(m_n+n)\delta} - \mathbb{E}[\,U_{(m_n+n)\delta}|\mathfrak{F}_{n\delta}]\right| > \varepsilon\,\right)$$
$$\leq \varepsilon^{-2}\mathbb{E}_{\delta_x}\left(\left|U_{(m_n+n)\delta} - \mathbb{E}[U_{(m_n+n)\delta}|\mathfrak{F}_{n\delta}]\right|^2\right). \tag{12.13}$$

Hence, let us consider the term in the conditional expectation on the right-hand side above. First note that

$$U_{(m_n+n)\delta} - \mathbb{E}[U_{(m_n+n)\delta}|\mathfrak{F}_{n\delta}] = \sum_{i=1}^{N_{n\delta}} \mathrm{e}^{-n\delta\lambda_*}(U^{(i)}_{m_n\delta} - \mathbb{E}[U^{(i)}_{m_n\delta}|\mathfrak{F}_{n\delta}]), \tag{12.14}$$

where, given $\mathfrak{F}_t$, the $U^{(i)}$ are independent and equal in distribution to U under $\mathbb{P}_{\delta_{x_i(t)}}$ and $(x_i(t) : i = 1, \cdots, N_t)$ describes the configuration of X at time $t \geq 0$. Note, in particular, conditional on $\mathfrak{F}_{n\delta}$, $Z_i = U^{(i)}_{m_n\delta} - \mathbb{E}(U^{(i)}_{m_n\delta}|\mathfrak{F}_{n\delta})$ are independent with zero mean. The formula for the variance of sums of zero-mean independent random variables together with the inequality $|a+b|^2 \leq 2(|a|^2 + |b|^2)$ now gives us

$$\begin{aligned}
&\mathbb{E}(|U_{(m_n+n)\delta} - \mathbb{E}[U_{(m_n+n)\delta}|\mathfrak{F}_{n\delta}]|^2|\mathfrak{F}_{n\delta}) \\
&= \sum_{i=1}^{N_{n\delta}} \mathrm{e}^{-2\lambda_* n\delta}\mathbb{E}\left[\left|U^{(i)}_{m_n\delta} - \mathbb{E}[U^{(i)}_{m_n\delta}|\mathfrak{F}_{n\delta}]\right|^2 \middle| \mathfrak{F}_{n\delta}\right] \\
&\leq \sum_{i=1}^{N_{n\delta}} \mathrm{e}^{-2\lambda_* n\delta}\mathbb{E}\left[2(|U^{(i)}_{m_n\delta}|^2 + |\mathbb{E}[U^{(i)}_{m_n\delta}|\mathfrak{F}_{n\delta}]|^2)|\mathfrak{F}_{n\delta}\right] \\
&\leq 4\sum_{i=1}^{N_{n\delta}} \mathrm{e}^{-2\lambda_* n\delta}\mathbb{E}\left[|U^{(i)}_{m_n\delta}|^2|\mathfrak{F}_{n\delta}\right],
\end{aligned} \tag{12.15}$$

where we have used Jensen's inequality in the final inequality. Hence, with $(x_i(n\delta) : i = 1, \ldots, N_{n\delta})$ describing the configurations of the particles at time $n\delta$ in X, we have

$$\begin{aligned}
&\sum_{n=1}^{\infty}\mathbb{E}\left[|U_{(m_n+n)\delta} - \mathbb{E}(U_{(m_n+n)\delta}|\mathfrak{F}_{n\delta})|^2\right] \\
&\quad \leq 4\sum_{n=1}^{\infty} \mathrm{e}^{-2\lambda_* n\delta}\mathbb{E}_{\delta_x}\left[\sum_{i=1}^{N_{n\delta}} \mathbb{E}_{\delta_{x_i(n\delta)}}\left[U^2_{m_n\delta}\right]\right] \\
&\quad \leq 4\|g\|^2\sum_{n=1}^{\infty} \mathrm{e}^{-2\lambda_* n\delta}\mathbb{E}_{\delta_x}\left[\sum_{i=1}^{N_{n\delta}} \varphi(x_i(n\delta))\frac{\mathrm{e}^{-2\lambda_* m_n\delta}\psi^{(2)}_{m_n\delta}[g](x_i(n\delta))}{\varphi(x_i(n\delta))}\right],
\end{aligned} \tag{12.16}$$

where we have used the notation from Theorem 9.2 in the final inequality. The latter also tells us that, by choosing m_n sufficiently large, we can uniformly bound the ratio on the right-hand side of (12.16) by a constant.

Recalling that φ is an eigenfunction for the linear semigroup of X, there exists a constant $K \in (0, \infty)$ such that

$$\sum_{n=1}^{\infty} \mathbb{E}\left[|U_{(m_n+n)\delta} - \mathbb{E}(U_{(m_n+n)\delta}|\mathfrak{F}_{n\delta})|^2\right] \le K\|g\|^2 \sum_{n\ge 1} \mathrm{e}^{-2\lambda_* n\delta} \mathbb{E}_{\delta_x}[X_{n\delta}[\varphi]]$$
$$= K\|g\|^2 \varphi(x) \sum_{n\ge 1} \mathrm{e}^{-\lambda_* n\delta} < \infty. \tag{12.17}$$

The result now follows by (12.12) and (12.13). □

Let us now return to the proof of our main strong law of large numbers result, albeit, first for convergence on lattice times. Notably, this allows for weaker assumptions than Theorem 12.3.

Theorem 12.4 *Suppose $\lambda_* > 0$, (G2) holds, as well as second moments, i.e., (G2) for $k = 2$. Then, for any $g \in B^+(E)$, $\delta > 0$, and any initial configuration $\mu \in \mathcal{M}_c(E)$,*

$$\mathrm{e}^{-\lambda_* n\delta} \frac{X_{n\delta}[g]}{\mu[\varphi]} \to \tilde{\varphi}[g] W_\infty, \tag{12.18}$$

$\mathbb{P}_\mu$-almost surely.

Proof It suffices to prove the result for $\mu = \delta_x$, $x \in E$, since then, for $\mu = \sum_{i=1}^n \delta_{x_i}$, we have under $\mathbb{P}_\mu$,

$$\lim_{t\to\infty} \mathrm{e}^{-\lambda_* t} X_t[g] = \sum_{i=1}^n \lim_{t\to\infty} \mathrm{e}^{-\lambda_* t} X_t^{(i)}[g] = \sum_{i=1}^n \varphi(x_i)\tilde{\varphi}[g] W_\infty^{(i)} = \tilde{\varphi}[g]\mu[\varphi] W_\infty,$$

where $X^{(i)}$ (resp., $W^{(i)}$) are independent copies of X (resp. W) under $\mathbb{P}_{\delta_{x_i}}$, $i = 1, \cdots, n$.

We have already noted that

$$\mathbb{E}_{\delta_x}\left[U_{t+s}\middle|\mathfrak{F}_t\right] = \sum_{i=1}^{N_t} \mathrm{e}^{-\lambda_* t} \bar{U}_s^{(i)},$$

where, given $\mathfrak{F}_t$, the $\bar{U}_s^{(i)}$ are independent and equal in law to $\mathbb{E}_{\delta_{x_i(t)}}[U_s]$, where $(x_i(t) : i = 1, \cdots, N_t)$ describes the configuration of X at time $t \ge 0$. Hence, using (8.14), we have

$$
\begin{aligned}
\mathbb{E}_{\delta_x}\left[U_{t+s}|\mathfrak{F}_t\right] &= \sum_{i=1}^{N_t} \mathrm{e}^{-\lambda_* t}\mathbb{E}_{\delta_{x_i(t)}}\left[\mathrm{e}^{-\lambda_* s}X_s[g]\right] \\
&= \varphi(x)\tilde{\varphi}[g]W_t \\
&\qquad + \sum_{i=1}^{N_t} \mathrm{e}^{-\lambda_* t}\left(\mathrm{e}^{-\lambda_* s}\frac{\psi_s[g](x_i(t))}{\varphi(x_i(t))} - \tilde{\varphi}[g]\right)\varphi(x_i(t)).
\end{aligned} \tag{12.19}
$$

Appealing (G2), we can pick s sufficiently large so that, for any given $\varepsilon > 0$,

$$
\|\mathrm{e}^{-\lambda_* s}\varphi^{-1}\psi_s[g] - \tilde{\varphi}[g]\| < \varepsilon. \tag{12.20}
$$

Combining this with (12.19), since we may take ε arbitrarily small, we conclude

$$
\lim_{t\to\infty}\left|\mathbb{E}_{\delta_x}[U_{t+s}|\mathfrak{F}_t] - W_\infty\tilde{\varphi}[g]\varphi(x)\right| = 0. \tag{12.21}
$$

The above combined with the conclusion of Lemma 12.2 gives the desired limit along lattice sequences.

□

By piecing together various calculations from the previous two results, we can also quickly deliver the proof of Corollary 12.1.

Proof (of Corollary 12.1) In the spirit of the calculations in (12.15), (12.19), and the already established L^2 convergence of W, for each $\mu \in \mathcal{M}_c(E)$, we can control the $L^2(\mathbb{P}_\mu)$ distance of $(U_{t+s} - \mathbb{E}_\mu[U_{t+s}|\mathfrak{F}_t])$, $(\mathbb{E}_\mu[U_{t+s}|\mathfrak{F}_t] - \mu[\varphi]\tilde{\varphi}[g]W_t)$, and $(W_t - W_\infty)$. Hence, by the triangle inequality, we have in the sense of $L^2(\mathbb{P}_\mu)$ convergence, $\lim_{t\to\infty} U_t = \mu[\varphi]\tilde{\varphi}[g]W_\infty$, as required. □

We now make the transition from lattice times to continuous times.

Proof (of Theorem 12.3 Along the Full Sequence) Suppose that Ω is an open set that is compactly embedded in E and fix $0 < \delta \ll 1$. If $t \in [n\delta, (n+1)\delta)$, then since $1 \geq \mathbf{1}_\Omega$,

$$
\begin{aligned}
\mathrm{e}^{-\lambda_* t}X_t[\mathbf{1}_\Omega] &= \mathrm{e}^{-\lambda_*(t-n\delta)}\sum_{i=1}^{N_{n\delta}}\mathrm{e}^{-\lambda_* n\delta}\sum_{j=1}^{N^{(i)}_{t-n\delta}}\mathbf{1}_\Omega(x^{(i)}_j(t-n\delta)) \\
&\geq \mathrm{e}^{-\lambda_*\delta}\sum_{i=1}^{N_{n\delta}}\mathrm{e}^{-\lambda_* n\delta}\mathbf{1}_\Omega(x_i(n\delta))X^{(i)}_{t-n\delta}[\mathbf{1}_\Omega],
\end{aligned} \tag{12.22}
$$

where $(x^{(i)}_j(s), j = 1, \cdots, N^{(i)}_s)$ are the collection of offspring at time $s \geq 0$ in the subtree $(X^{(i)}, \mathbb{P}_{\delta_{x_i(n\delta)}})$ rooted at the i-th individual alive at time $n\delta$, for each $i = 1, \cdots, N_{n\delta}$. Now note that, given $\mathfrak{F}_{n\delta}$, for all $s \in [0, \delta]$,

$$X_s^{(i)}[\mathbf{1}_\Omega] \geq \mathbf{1}_{\{X_s^{(i)}[\mathbf{1}_\Omega]\geq 1\}} \geq \mathbf{1}_{\{X_u^{(i)}[\mathbf{1}_\Omega]\geq 1,\text{ for all } u\in[0,\delta]\}}.$$

Hence, from the inequality (12.22), we can derive the more convenient inequality

$$\mathrm{e}^{-\lambda_* t} X_t[\mathbf{1}_\Omega] \geq \mathrm{e}^{-\lambda_*\delta} \sum_{i=1}^{N_{n\delta}} \mathrm{e}^{-\lambda_* n\delta} \mathbf{1}_\Omega(x_i(n\delta)) \Xi^\delta(x_i(n\delta)), \tag{12.23}$$

where

$$\Xi^\delta(x) := \mathbf{1}_{\{X_t[\mathbf{1}_\Omega]\geq 1 \text{ for all } t\in[0,\delta]\}},$$

for $X_0 = \delta_x$, $x \in E$.

Next note that if we trace the path of $(X, \mathbb{P}_{\delta_x})$ over the small time horizon $[0, \delta]$, then with high probability as $\delta \downarrow 0$, no branching event has occurred and the requirement that $X_t[\mathbf{1}_\Omega] \geq 1$ for $t \in [0, \delta]$ is close to the requirement that the initial ancestor remains in Ω prior to time δ. Indeed, suppose T_1 is the first branching time of $(X, \mathbb{P}_{\delta_x})$ and let $\eta^\delta(x) = \mathbb{E}_{\delta_x}[\Xi^\delta(x)] \leq 1$, then we have

$$\begin{aligned}
\eta^\delta(x) &= \mathbb{P}_{\delta_x}(X_t[\mathbf{1}_\Omega] \geq 1 \text{ for all } t \in [0, \delta]) \\
&\geq \mathbb{P}_{\delta_x}(X_t[\mathbf{1}_\Omega] \geq 1 \text{ for all } t \in [0, \delta],\ \delta < T_1) \\
&= \mathbf{E}_x\left[\mathbf{1}_{(\xi_t\in\Omega,\text{ for all } t\in[0,\delta])} \mathrm{e}^{-\int_0^\delta \gamma(\xi_s)\mathrm{d}s}\right] \\
&= \mathtt{P}_\delta^\Omega[1](x) - \mathbf{E}_x\left[\mathbf{1}_{(\xi_t\in\Omega,\text{ for all } t\in[0,\delta])}\left(1 - \mathrm{e}^{-\int_0^\delta \gamma(\xi_s)\mathrm{d}s}\right)\right],
\end{aligned} \tag{12.24}$$

where we recall that $(\xi, \mathbf{P}_x)$ and γ denote the law of particles' motion and the branching rate in $(X, \mathbb{P})$, respectively, and $(\mathtt{P}_t^\Omega, t \geq 0)$ is the semigroup of $(\xi, \mathtt{P})$, killed on exiting Ω. Thanks to the assumption $\gamma \in B^+(E)$ (cf. Chap. 8), it is easy to show that the second term on the right-hand side of (12.24) tends to zero as $\delta \to 0$, since

$$\begin{aligned}
&\mathbf{E}_x\left[\mathbf{1}_{(\xi_t\in\Omega,\text{ for all } t\in[0,\delta])}\left(1 - \mathrm{e}^{-\int_0^\delta \gamma(\xi_s)\mathrm{d}s}\right)\right] \\
&\quad\leq \mathbf{E}_x\left[\mathbf{1}_{(\xi_t\in\Omega,\text{ for all } t\in[0,\delta])} \int_0^\delta \gamma(\xi_s)\mathrm{d}s\right] \\
&\quad\leq \delta\|\gamma\|.
\end{aligned}$$

Thanks to the assumption (12.8), we have from (12.24),

$$\liminf_{\delta\downarrow 0} \eta^\delta(x) \geq \mathbf{1}_\Omega(x), \qquad x \in E. \tag{12.25}$$

If we denote the summation on the right-hand side of (12.23) by $\tilde{U}_{n\delta}(x)$, then we can apply similar arguments to those given in the proof of Lemma 12.2, see in particular (12.16), to show that, for some constant $C \in (0, \infty)$,

$$\sum_{n=1}^{\infty} \mathbb{E}_{\delta_x}\left[|\tilde{U}_{n\delta} - \mathbb{E}[\tilde{U}_{n\delta}|\mathfrak{F}_{n\delta}]|^2\right]$$
$$\leq C\sum_{n=1}^{\infty} \mathrm{e}^{-2\lambda_* n\delta}\mathbb{E}_{\delta_x}\left[\sum_{i=1}^{N_{n\delta}} \mathbf{1}_\Omega(x_i(n\delta))^2 \mathbb{E}_{\delta_{x_i(n\delta)}}[(\Xi^\delta)^2]\right]$$
$$\leq C\sum_{n=1}^{\infty} \mathrm{e}^{-2\lambda_* n\delta}\mathbb{E}_{\delta_x}\left[X_{n\delta}[\mathbf{1}_\Omega]\right]$$
$$= C\sum_{n=1}^{\infty} \mathrm{e}^{-2\lambda_* n\delta}\psi_{n\delta}[\mathbf{1}_\Omega](x). \tag{12.26}$$

The assumption (G2) now tells us that the sum on the right-hand side of (12.26) is finite.

Noting from (12.23) that

$$\mathbb{E}[\tilde{U}_{n\delta}|\mathfrak{F}_{n\delta}] = \mathrm{e}^{-\lambda_* n\delta} X_{n\delta}[\mathbf{1}_\Omega \eta^\delta],$$

the consequence of (12.26), when taken in the light of the Borel–Cantelli lemma, the lower bound (12.23), and Theorem 12.4, means that, $\mathbb{P}_{\delta_x}$-almost surely,

$$\liminf_{t\to\infty} \mathrm{e}^{-\lambda_* t} X_t[\mathbf{1}_\Omega] \geq \mathrm{e}^{-\lambda_* \delta}\tilde{\varphi}[\mathbf{1}_\Omega \eta^\delta] W_\infty \varphi(x).$$

Letting $\delta \downarrow 0$ with the help of (12.25) and Fatou's lemma in the above inequality yields

$$\liminf_{t\to\infty} \mathrm{e}^{-\lambda_* t} X_t[\mathbf{1}_\Omega] \geq \tilde{\varphi}[\mathbf{1}_\Omega] W_\infty \varphi(x), \tag{12.27}$$

$\mathbb{P}_{\delta_x}$-almost surely.

From (12.27), we can easily replace $\mathbf{1}_\Omega$ by simple measurable functions in $B^+(E)$, from where, we can upgrade by lower approximation to measurable $g \in B^+(E)$. Indeed, suppose that $(g_n, n \geq 1)$ is a sequence of simple functions in $B^+(E)$ such that $g_n \uparrow g \in B^+(E)$. Then, since $g \geq g_n$,

$$\liminf_{t\to\infty} \mathrm{e}^{-\lambda_* t} X_t[g] \geq \liminf_{t\to\infty} \mathrm{e}^{-\lambda_* t} X_t[g_n] \geq \tilde{\varphi}[g_n] W_\infty \varphi(x),$$

and hence, taking limits as $n \to \infty$ on the right-hand side with the help of monotone convergence, we have

$$\liminf_{t\to\infty} \mathrm{e}^{-\lambda_* t} X_t[g] \geq \tilde{\varphi}[g] W_\infty \varphi(x). \tag{12.28}$$

To complete the proof of Theorem 12.3, it now suffices to show that, $\mathbb{P}_{\delta_x}$-almost surely, $\limsup_{t\to\infty} \mathrm{e}^{-\lambda_* t} X_t[g] \leq \tilde{\varphi}[g] W_\infty \varphi(x)$, for $g \in B^+(E)$ and $g/\varphi \in B^+(E)$. To this end, note that, for $0 \leq g \leq c\varphi$, for some constant $c > 0$ (which, without loss of generality, we may take equal to 1),

$$\begin{aligned}
\limsup_{t\to\infty} \mathrm{e}^{-\lambda_* t} X_t[g] &= \limsup_{t\to\infty} \left(\varphi(x) W_t - \mathrm{e}^{-\lambda_* t} X_t[\varphi - g]\right) \\
&= \varphi(x) W_\infty - \liminf_{t\to\infty} \mathrm{e}^{-\lambda_* t} X_t[\varphi - g] \\
&\leq \varphi(x) W_\infty - \tilde{\varphi}[\varphi - g] \varphi(x) W_\infty \\
&= \tilde{\varphi}[g] W_\infty \varphi(x)
\end{aligned}$$

as required, where we have used the normalisation $\tilde{\varphi}[\varphi] = 1$. □

12.4 Discrete-Time Strong Law of Large Numbers

For this section, we recall the notation from Sect. 8.7. Just as in the continuous-time setting, it is easy to show that $\mathscr{W} = (\mathscr{W}_n, n \geq 0)$, where

$$\mathscr{W}_n = \rho_*^{-n} \frac{\mathscr{X}_n[\omega]}{\mu[\omega]}, \qquad n \geq 0,$$

is a $\mathbb{P}_\mu$-martingale, for $\mu \in \mathscr{M}_c(E)$. Moreover, being a non-negative martingale, it has an almost sure limit, denoted by $\mathscr{W}_\infty$. In a similar spirit to the calculation (12.5), using the (second) moment convergence results in Theorem 9.8, it is straightforward to show that $\mathscr{W}$ is an L^2-convergent martingale.

The calculations in the previous section that lead to Lemma 12.2 and Theorem 12.4 can be easily replicated in discrete time. The details are left to the reader, but there is little to change. Since the time index is now discrete rather than continuous, we automatically get for free the following result.

Theorem 12.5 *Suppose that $\rho_* > 1$ and (G4) hold, together with second moments of the offspring distribution, i.e., (G5), for $k = 2$. Then, for any $g \in B^+(E)$ and any initial configuration $\mu \in \mathscr{M}_c(E)$,*

$$\rho_*^n \frac{\mathscr{X}_n[g]}{\mu[\omega]} \to \tilde{\varphi}[g] \mathscr{W}_\infty, \tag{12.29}$$

$\mathbb{P}_\mu$-almost surely.

12.5 Comments

The proof of Theorem 12.1 is a variant of a standard one appealing to a standard measure theoretical dichotomy (cf. p. 242 of [47]), which has been used to analyse the convergence of many analogous martingales in the setting of different spatial branching processes. We mention [12, 97, 124], and [58] to name but a few of the contexts with similar results. The proof of Theorem 12.3 principally uses techniques that have been used a number of times in the literature, developed by [4, 24, 54] among others.

Glossary

Abbreviations

ACP	Abstract Cauchy problem	10
BMP	Branching Markov process	149
NRW	Neutron random walk	49
NBP	Neutron branching process	52
NGP	Neutron generation process	129
NTE	Neutron transport equation	5, 43
PBE	Pàl–Bell equation	93

Main Assumptions

(A1), (A2)	Two sufficient conditions for existence of a Perron–Frobenius decomposition for the semigroup of a general Markov process	32
(A1*), (A2*)	Two sufficient conditions for existence of a Perron–Frobenius decomposition in discrete time	136
(B1), (B2)	Alternative sufficient conditions to (A1), (A2), for the setting of a neutron random walk, which provides the existence of a Perron–Frobenius decomposition of its semigroup	73
(H1)	$\sigma_{\mathtt{s}}$, $\sigma_{\mathtt{f}}$, $\pi_{\mathtt{s}}$, and $\pi_{\mathtt{f}}$ are uniformly bounded away from infinity	6
(H2)	$\sigma_{\mathtt{s}}\pi_{\mathtt{s}} + \sigma_{\mathtt{f}}\pi_{\mathtt{f}} > 0$ on $D \times V \times V$	54
(H2*)	$\inf_{r\in D,\upsilon,\upsilon'\in V} \alpha(r,\upsilon)\pi(r,\upsilon,\upsilon') > 0$	63
(H3)	There is an open ball B compactly embedded in D such that $\sigma_{\mathtt{f}}\pi_{\mathtt{f}} > 0$ on $B \times V \times V$	54
(H3*)	There exists an open ball B, compactly embedded in D, such that $\inf_{r\in B,\upsilon,\upsilon'\in V} \sigma_{\mathtt{f}}(r,\upsilon)\pi_{\mathtt{f}}(r,\upsilon,\upsilon') > 0$	109
(H3**)	$\inf_{r\in D,\upsilon,\upsilon'\in V} \sigma_{\mathtt{f}}(r,\upsilon)\pi_{\mathtt{f}}(r,\upsilon,\upsilon') > 0.$	134

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8

(H4)	Fission offspring are bounded in number by the constant $n_{\mathtt{max}} > 1$	54
(H5):	There exists a constant $C > 0$ such that for all $g \in L^+_\infty(D \times V)$, $\langle \tilde{\varphi}, \sigma_{\mathtt{f}} \mathbb{V}[g] \rangle \geq C \langle \tilde{\varphi}, \hat{g}^2 \rangle$, where $\hat{g} : D \to [0, \infty) : r \mapsto \int_V g(r, \upsilon')\mathrm{d}\upsilon'$	104
(G1)	$\sup_{x \in E} \mathscr{E}_x[N] < \infty$	156
(G2)	There exists an eigenvalue $\lambda_* \in \mathbb{R}$ and a corresponding right eigenfunction $\varphi \in B^+(E)$ and left eigenmeasure $\tilde{\varphi} \in \mathscr{M}_f(E)$ such that, for $f \in B^+(E)$ and $\mu \in \mathscr{M}_f(E)$, $\langle \psi_t[\varphi], \mu \rangle = \mathrm{e}^{\lambda_* t} \langle \varphi, \mu \rangle$ and $\langle \psi_t[f], \tilde{\varphi} \rangle = \mathrm{e}^{\lambda_* t} \langle f, \tilde{\varphi} \rangle$. Further let us define $\Delta_t = \sup_{x \in E, f \in B^+_1(E)} \lvert \varphi(x)^{-1} \mathrm{e}^{-\lambda_* t} \psi_t[f](x) - \langle \tilde{\varphi}, f \rangle \rvert$, $t \geq 0$. We suppose that $\sup_{t \geq 0} \Delta_t < \infty$ and $\lim_{t \to \infty} \Delta_t = 0$.	157
(G3)	$\sup_{x \in E} \mathscr{E}_x[\mathsf{Z}[1]^k] < \infty$	170
(G4)	There exists an eigenvalue $\rho \in \mathbb{R}$, and a corresponding right eigenfunction $0 < \omega \in B^+(E)$ and left eigenmeasure $\tilde{\omega} \in \mathscr{M}_f(E)$ such that, for $f \in B^+(E)$ and $\mu \in \mathscr{M}_f(E)$, $\mu[\Phi_n[\omega]] = \rho^n \mu[\omega]$ and $\tilde{\omega}[\Phi_n[f]] = \rho^n \tilde{\omega}[f]$. Further let us define $\Delta_n = \sup_{x \in E, g \in B_1(E)} \lvert \omega(x)^{-1} \rho^{-n} \Phi_n[f](x) - \tilde{\omega}[f] \rvert$, for $n \geq 0$. We suppose that $\sup_{n \geq 0} \Delta_n < \infty$ and $\lim_{n \to \infty} \Delta_n = 0$.	187
(G5)	$\sup_{x \in E} \mathrm{E}_x(\mathscr{Z}[1]^k) < \infty$.	188
(G6)	Offspring numbers are bounded by $n_{\mathtt{max}}$	196
(G7)	Irreducible branching condition, $\langle \gamma \mathscr{V}[g], \tilde{\varphi} \rangle \geq C \langle g, \tilde{\varphi} \rangle^2$	196
(G8)	Uniform extinction condition, $\sup_{x \in E} \mathbf{P}_x(t < \zeta) < 1$ for all sufficiently large $t > 0$.	196
(G9)	Extinction by time T is uniformly bounded away from zero, $\inf_{x \in E} \mathbb{P}_{\delta_x}(\zeta < T) > 0$, for $0 \leq T < \infty$.	223
(G10)	Extinction by time T is not a certainty, $\mathbb{P}_{\delta_x}(\zeta < T) < 1$ for $x \in E$ and $0 \leq T < \infty$.	224
(G11)	$\lim_{T \to \infty} \inf_{x \in E} w^T(x) = 1$.	238

Mathematical Notation

$\dagger$	Cemetery state	20
$\lVert\cdot\rVert$	The supremum norm on $B^+(E)$	32
$\langle \cdot, \cdot \rangle$	Inner product on $L^2(E)$	32
$\mu[f]$	Integral of f with respect to the measure μ	31
$\alpha(r, \upsilon)$	A mixed rate of scatter and fission	68
$\beta(r, \upsilon)$	The function $\sigma_{\mathtt{f}}(r, \upsilon) \left(\int_V \pi_{\mathtt{f}}(r, \upsilon, \upsilon')\mathrm{d}\upsilon' - 1 \right)$	67
γ	The branching rate for a BMP	149
γ^φ	Branching rate along the spine for a BMP	210
$\gamma^\updownarrow$	Branching rate of the BMP dressed $\uparrow$-marked tree	230

ζ	Lifetime of stochastic process	20, 103
η	Left eigenmeasure	32, 70
θ	A mixed scatter and fission density	132
$\kappa_{r,\upsilon}^{D}$	The time a linear drift from r with velocity υ leaves the domain D	47
λ_c	Leading eigenvalue in Perron–Frobenius decomposition of Markov semigroup	30
λ_*	Leading eigenvalue for the NBP or BMP	70, 157
ξ	A Markov process (which serves as the motion process for each particle in a BMP)	156
π	A mixed scatter and fission density	68
$\pi_{\mathtt{f}}$	Average neutron mass creation at fission	6
$\pi_{\mathtt{s}}$	Scatter distribution	6
ρ_*	Leading eigenvalue in Perron–Frobenius decomposition of discrete-time Markov semigroup	187
$\sigma_{\mathtt{f}}$	Rate at which fission occurs	6
$\sigma_{\mathtt{f}}^{\downarrow}$	Rate at which fission occurs in $\downarrow$ subtree of skeletal decomposition	120
$\sigma_{\mathtt{f}}^{\updownarrow}$	Rate at which fission occurs in $\updownarrow$ subtree of skeletal decomposition	121
$\sigma_{\mathtt{s}}$	Rate at which neutron scatter occurs	6
Σ	The constant equal to $\langle \sigma_{\mathtt{f}} \mathscr{V}[\varphi], \tilde{\varphi} \rangle$ for NBP and $\tilde{\varphi}\big[\gamma \mathscr{V}[\varphi]\big]$ of BMP	104, 196
φ	Right eigenfunction in the Perron–Frobenius decomposition for NBP and BMP	32, 70, 136, 157
$\tilde{\varphi}$	Left eigenmeasure (or density thereof) in the Perron–Frobenius decomposition for NBP and BMP	32, 70, 136, 157
Φ_n	First moment semigroup of generation NBP or discrete-time BMP	129, 165
ψ_t	First moment semigroup of NBP or BMP	57, 155
$\psi_t^{(k)}$	The k-th moment functional of the NBP or BMP	100, 169
Ψ_t	Solution to the NTE in $L^2(D \times V)$	5, 43
$\hat{\Psi}_t$	Solution to the dual NTE in the $L_2(D \times V)$ space	45
ω	Right eigenfunction in the Perron–Frobenius decomposition for generational NBP or discrete-time BMP	134, 187
$\tilde{\omega}$	Left eigenmeasure in the Perron–Frobenius decomposition for generational NBP or discrete-time BMP	134, 187
$a(t)$	The survival probability integrated against $\tilde{\varphi}$	195
$\mathtt{A}$	Difference of non-linear and linear branching operators	160
$B(E)$	Bounded measurable functions on E taking value zero on $\dagger$	21, 47
$B^+(E)$	Non-negative bounded measurable functions on E taking value zero on $\dagger$	21, 47
$B_1^+(E)$	Functions in $B^+(E)$ bounded by unity	93, 150
$\mathtt{B}(x)$	The function $\gamma(x)(\mathtt{m}[1](x) - 1)$	156
$\mathtt{Dom}(\cdot)$	The domain of the operator in brackets on $L_2(D \times V)$	11

E	Underlying state space of Markov process	19
$\mathtt{F}$	Infinitesimal generator of fission events for the backward NTE or linear branching operator for BMP	45, 155
$\mathscr{F}$	Forward infinitesimal generator of fission events	10
$\hat{\mathscr{F}}$	The $L^2(D \times V)$ dual of the infinitesimal generator of fission events	127
$\mathfrak{F}_t$	Filtration of NBP or BMP	55, 151
$\mathfrak{G}_t$	Filtration of the general Markov process ξ	19
$\mathscr{G}$	The sum of the transport, scatter, and fission forward operators	10
$\hat{\mathscr{G}}$	The $L^2(D \times V)$ dual operator of $\mathscr{G}$	45
$\mathtt{G}$	Branching mechanism for NBP or BMP	93, 150
$\mathtt{G}^{\uparrow}$, $\mathtt{G}^{\uparrow,T}$	Branching mechanism for $\uparrow$-marked individuals in the skeletal (resp., T-skeletal) decomposition for NBP or BMP	122, 230
$\mathtt{G}^{\downarrow}$, $\mathtt{G}^{\downarrow,T}$	Branching mechanism for $\downarrow$-marked individuals in the skeletal (resp., T-skeletal) decomposition for NBP or BMP	119, 229
$\mathtt{G}^{\updownarrow}$, $\mathtt{G}^{\updownarrow,T}$	Joint branching mechanism for $\uparrow$- and $\downarrow$-marked individuals in the skeletal (resp., T-skeletal) decomposition for NBP or BMP	121, 230
$I_t^{(k)}$	The k-th moment functional of running occupation measure for NBP or BMP	102, 169
$k_{\mathtt{eff}}$	Generational-time eigenvalue for neutron transport operators	128
$\mathfrak{K}_n$	Filtration of NGP or discrete-time BMP	129, 165
$L^2(E)$	The space of square integrable functions on E	10
$\mathscr{M}_f(E)$	Space of finite measures on E	32
$\mathscr{M}_c(E)$	Space of finite counting measures on E	53, 151
$m(r, \upsilon)$	Mean number of neutrons produced at a fission event from an incoming neutron with configuration (r, υ)	132
$\mathtt{m}[\cdot]$	Mean operator for offspring	155
p, p_T	Survival probability (resp., to time T) for NBP or BMP	116, 223
$\mathbb{P}$	Law of NBP or BMP	54
$\mathbb{P}^{\varphi}$	Change of measure with respect to W	209
$\tilde{\mathbb{P}}^{\varphi}$	Law of the dressed spine	112
$\mathbb{P}^{\downarrow}$, $\mathbb{P}^{\downarrow,T}$	Law of the NBP or BMP conditioned to become extinct (resp., by time T)	119, 225
$\hat{\mathbf{P}}^{\varphi}$	Marginal semigroup of the spine for BMP	211
$\mathbf{P}$	Law of Markov process for NRW and BMP	20, 49, 156
$\mathbf{P}^{\alpha\pi}$	Law of Markov process for NRW and BMP in the case that the scatter rate α and scatter kernel π are specified	49
$\mathbf{P}^{\sigma\pi,\dagger}$	Law of Markov process for NRW and BMP in the case that the scatter rate α and scatter kernel π are specified, with additional killing	135
$\hat{\mathbf{P}}$	Probabilities of underlying Markov process with additional jumps	156
$\mathbf{P}^{\dagger}$	Law of (R, Υ) with additional killing at rate $\bar{\beta} - \beta$	69

$\mathbf{P}^{\downarrow}$, $\mathbf{P}^{\downarrow,T}$	Law of the Markov process describing movement of $\downarrow$-marked particles in the NBP or BMP skeletal (resp., T-skeletal) decomposition	119, 226
$\mathbf{P}^{\uparrow}$, $\mathbf{P}^{\uparrow,T}$	Law of the Markov process describing movement of $\uparrow$-marked particles in the NBP or BMP skeletal (resp., T-skeletal) decomposition	122, 230
$\mathbb{P}^{\updownarrow}$, $\mathbb{P}^{\updownarrow,T}$	Law of $\uparrow$-marked dressed tree for NBP or BMP	118, 224
$\mathtt{P}^{\dagger}$	Semigroup associated to (R, Υ) under $\mathbf{P}^{\dagger}$	59
$\mathtt{P}$	Expectation semigroup $\sigma_{\mathtt{S}}\pi_{\mathtt{S}}$-NRW killed when it exits D or expectation semigroup of underlying Markov process for BMP	64, 149
$\mathtt{P}^{\varphi}$	Motion biased semigroup of the spine for a BMP	210
$\mathtt{P}^{\Omega}$	Movement semigroup of $\mathtt{P}$ killed on exiting Ω	247
$\mathtt{P}^{c}$	Semigroup of Markov process after change of measure with ground state eigenfunction	36
P	Branching point process probabilities for discrete-time model	164
$\mathscr{P}$	Fission point process probabilities or branching point process probabilities	53, 149
$\mathscr{P}^{\varphi}$	Fission point process probabilities or branching point process probabilities along the spine	112, 210
$\mathscr{P}^{\downarrow}$, $\mathscr{P}^{\downarrow,T}$	Fission point process probabilities or branching point process probabilities for $\downarrow$-marked trees in the skeletal (resp., T-skeletal) decomposition	120, 229
$\mathscr{P}^{\updownarrow}$, $\mathscr{P}^{\updownarrow,T}$	Joint fission point process probabilities or branching point process probabilities for $\downarrow$-marked and $\uparrow$-marked trees in the skeletal (resp., T-skeletal) decomposition	121, 234
$\mathtt{Q}$	Expectation semigroup of Markov process	21
$\mathtt{Q}^{\gamma}$	Expectation semigroup of Markov process with potential γ	21
$\mathscr{Q}$	Source term for the forward NTE	6
$\mathtt{S}$	Infinitesimal generator of scattering event for backward NTE	45
$\mathscr{S}$	Forward scatter operator	10
$\hat{\mathscr{S}}$	The $L^2(D \times V)$ dual of the forward scatter operator	127
$\mathtt{T}$	Infinitesimal generator of transport for backward NTE	45
$\mathscr{T}$	Forward transport operator	10
$\hat{\mathscr{T}}$	The $L^2(D \times V)$ dual of the forward transport operator	127
$\mathtt{u}_t[f, g]$	Solution to the re-oriented non-linear evolution equation for BMP and its running occupation	160
$\mathtt{u}_t(x)$	The probability of survival for both BMP and NBP	195
$\mathtt{u}_t^{\updownarrow,T}$	Non-linear semigroup of the dressed $\uparrow$-marked T-skeleton for the BMP	231
$\mathtt{u}_t^{\downarrow,T}$	Non-linear semigroup of the BMP conditioned to become extinct by time T	226
$\mathtt{v}_t[f]$	Solution to the non-linear semigroup (Pàl–Bell) equation for NBP or BMP	92, 152
$\mathtt{v}_t[f, g]$	Solution to the non-linear semigroup equation for BMP and its running occupation	161

$\mathtt{V}_n[f]$	Non-linear semigroup of the discrete-time BMP	165
$\mathscr{V}[f, g]$	Two-point fission or branching operator	94
$\mathscr{V}[f]$	Fission or branching variance operator	94, 173196
w, w_T	Extinction probability (resp., by time T)	116, 223
W_t	The unit mean martingale $\langle\varphi, X_t\rangle/\langle\varphi, \mu\rangle$ for NBPs or equivalently, $X_t[\varphi]/\mu[\varphi]$ for BMPs, under $\mathbb{P}_\mu$	108, 111, 209
$\mathscr{W}_n$	The unit mean martingale $\mathscr{X}_n[\omega]/\mu[\omega]$ under $\mathbb{P}_\mu$	255
X_t	Neutron branching process or branching Markov process	52
$\mathscr{X}_n$	Branching Markov process in discrete time	52164
$X_t^{\uparrow,T}$	The T-skeleton of X	230
$X_t^{\updownarrow,T}$	The dressed T-skeleton of X	230
Z	Point process of fission velocities for NBP or branching offspring point process for BMP	53, 149, 151
$\mathscr{Z}$	Branching offspring point process for discrete-time processes	164

References

1. E. Aïdékon, J. Berestycki, É. Brunet, Z. Shi, Branching Brownian motion seen from its tip. Probab. Theory Related Fields **157**(1–2), 405–451 (2013)
2. E. Aidekon, Z. Shi, The Seneta-Heyde scaling for the branching random walk. Ann. Probab. **42**(3), 959–993 (2014)
3. S. Asmussen, H. Hering, Strong limit theorems for general supercritical branching processes with applications to branching diffusions. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **36**(3), 195–212 (1976)
4. S. Asmussen, H. Hering, Strong limit theorems for supercritical immigration-branching processes. Math. Scand. **39**(2), 327–342 (1977, 1976)
5. S. Asmussen, H. Hering, *Branching Processes*, vol. 3. Progress in Probability and Statistics (Birkhäuser Boston, Inc., Boston, 1983)
6. K.B. Athreya, P.E. Ney, *Branching Processes* (Dover Publications, Inc., Mineola, 2004). Reprint of the 1972 original [Springer, New York; MR0373040]
7. G.I. Bell, On the stochastic theory of neutron transport. Nucl. Sci. Eng. **21**, 390–401 (1965)
8. G.I. Bell, S. Glasstone, *Nuclear Reactor Theory* (Reinhold, New York, 1970)
9. O. Bénichou, M. Coppey, M. Moreau, P.H. Suet, R. Voituriez, Averaged residence times of stochastic. Europhys. Lett. **70**, 42–48 (2005)
10. J. Berestycki, A.E. Kyprianou, A. Murillo-Salas, The prolific backbone for supercritical superprocesses. Stochastic Process. Appl. **121**(6), 1315–1331 (2011)
11. J. Berestycki, É. Brunet, J.W. Harris, S.C. Harris, M.I. Roberts, Growth rates of the population in a branching Brownian motion with an inhomogeneous breeding potential. Stochastic Process. Appl. **125**(5), 2096–2145 (2015)
12. J. Bertoin, *Random Fragmentation and Coagulation Processes*, vol. 102. Cambridge Studies in Advanced Mathematics (Cambridge University Press, Cambridge, 2006)
13. J. Bertoin, Markovian growth-fragmentation processes. Bernoulli **23**(2), 1082–1101 (2017)
14. J. Bertoin, A. Rouault, Discretization methods for homogeneous fragmentations. J. London Math. Soc. (2) **72**(1), 91–109 (2005)
15. J. Bertoin, A.R. Watson, The strong Malthusian behavior of growth-fragmentation processes. Ann. H. Lebesgue **3**, 795–823 (2020)
16. J.D. Biggins, Martingale convergence in the branching random walk. J. Appl. Probab. **14**(1), 25–37 (1977)
17. J.D. Biggins, A.E. Kyprianou, Measure change in multitype branching. Adv. Appl. Probab. **36**(2), 544–581 (2004)
18. J. Bliedtner, W. Hansen, *Potential Theory*. Universitext (Springer, Berlin, 1986). An analytic and probabilistic approach to balayage

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8

19. R.J. Brissenden, A.R. Garlick, Biases in the estimation of Keff and its error by Monte Carlo methods. Ann. Nucl. Energy **13**(2), 63–83 (1986)
20. M.B. Chadwick, M. Herman, P. Obložinský, M.E. Dunn, Y. Danon, A.C. Kahler, D.L. Smith, B. Pritychenko, G. Arbanas, R. Arcilla, R. Brewer, ENDF/B-VII.1 nuclear data for science and technology: cross sections, covariances, fission product yields and decay data. Nuclear data sheets. EPJ Web Conf. **112**(12), 2887–2996 (2011)
21. N. Champagnat, D. Villemonais, Exponential convergence to quasi-stationary distribution and Q-process. Probab. Theory Related Fields **164**(1–2), 243–283 (2016)
22. N. Champagnat, D. Villemonais, Uniform convergence to the Q-process. Electron. Commun. Probab. **22**, 7 (2017)
23. B. Chauvin, A. Rouault, A. Wakolbinger, Growing conditioned trees. Stochastic Process. Appl. **39**(1), 117–130 (1991)
24. Z-Q. Chen, Y-X. Ren, T. Yang, Law of large numbers for branching symmetric Hunt processes with measure-valued branching rates. J. Theoret. Probab. **30**(3), 898–931 (2017)
25. K.L. Chung, J.B. Walsh, *Markov Processes, Brownian Motion, and Time Symmetry*, vol. 249. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 2nd edn. (Springer, New York, 2005)
26. P. Collet, S. Martínez, J. San Martín, *Quasi-Stationary Distributions*. Probability and Its Applications (New York) (Springer, Heidelberg, 2013). Markov chains, diffusions and dynamical systems
27. J.L. Conlin, A.C. Kahler, A.P. McCartney, D.A. Rehn, NJOY21: next generation nuclear data processing capabilities. EPJ Web Conf. **146**, 09040 (2017)
28. E.D. Courant, P.R. Wallace, Fluctuations of the number of neutrons in a pile. Phys. Rev. **72**, 1038–1048 (1947)
29. N. Chopin, O. Papaspiliopoulosg, *An Introduction to Sequential Monte Carlo*. Springer Series in Statistics, pp. xii+378 (Springer, Cham, 2020)
30. A.M.G. Cox, S.C. Harris, E.L. Horton, A.E. Kyprianou, Multi-species neutron transport equation. J. Stat. Phys. **176**(2), 425–455 (2019)
31. A.M.G. Cox, S.C. Harris, A.E. Kyprianou, M. Wang, Monte Carlo methods for the neutron transport equation. Preprint (2021)
32. A.M.G. Cox, E. Horton, A.E. Kyprianou, D. Villemonais, Stochastic methods for neutron transport equation III: generational many-to-one and $k_{\texttt{eff}}$. SIAM J. Appl. Math. **81**(3), 982–1001 (2021)
33. R. Dautray, J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology*, vol. 6 (Springer, Berlin, 1993). Evolution problems. II, With the collaboration of Claude Bardos, Michel Cessenat, Alain Kavenoky, Patrick Lascaux, Bertrand Mercier, Olivier Pironneau, Bruno Scheurer and Rémi Sentis, Translated from the French by Alan Craig
34. R. Dautray, M. Cessenat, G. Ledanois, P.-L. Lions, E. Pardoux, R. Sentis, *Méthodes probabilistes pour les équations de la physique*. Collection du Commissariat a l'énergie atomique (Eyrolles, Paris, 1989)
35. E.B. Davies, *Heat Kernels and Spectral Theory*, Number 92 (Cambridge University Press, Cambridge, 1989)
36. B. Davison, J.B. Sykes, *Neutron Transport Theory* (Clarendon Press, Oxford, 1957)
37. C. Dellacherie, P-A. Meyer, *Probabilités et potentiel*. Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. XV (Hermann, Paris, 1975). Chapitres I à IV, Édition entièrement refondue
38. C. Dellacherie, P-A. Meyer, *Probabilités et potentiel. Chapitres V à VIII*. Actualités Scientifiques et Industrielles [Current Scientific and Industrial Topics], No. 1385, Revised edn. (Hermann, Paris, 1980). Théorie des martingales. [Martingale theory]
39. C. Dellacherie, P-A. Meyer, *Probabilités et potentiel. Chapitres IX à XI*. Publications de l'Institut de Mathématiques de l'Université de Strasbourg [Publications of the Mathematical Institute of the University of Strasbourg], XVIII, Revised edn. (Hermann, Paris, 1983). Théorie discrète du potential. [Discrete potential theory], Actualités Scientifiques et Industrielles [Current Scientific and Industrial Topics], 1410

40. C. Dellacherie, P-A. Meyer, *Probabilités et potentiel. Chapitres XII–XVI*. Publications de l'Institut de Mathématiques de l'Université de Strasbourg [Publications of the Mathematical Institute of the University of Strasbourg], XIX, 2nd edn. (Hermann, Paris, 1987). Théorie du potentiel associée à une résolvante. Théorie des processus de Markov. [Potential theory associated with a resolvent. Theory of Markov processes], Actualités Scientifiques et Industrielles [Current Scientific and Industrial Topics], 1417
41. R.A. Doney, A limit theorem for a class of supercritical branching processes. J. Appl. Probab. **9**, 707–724 (1972)
42. J.L. Doob, *Classical Potential Theory and Its Probabilistic Counterpart*. Classics in Mathematics (Springer, Berlin, 2001). Reprint of the 1984 edition
43. E. Dumonteil, A. Mazzolo, Residence times of branching diffusion processes. Phys. Rev. E **94**, 012131 (2016)
44. E. Dumonteil, E. Horton, A.E. Kyprianou, A. Zoia, Limit theorems for the neutron transport equation (2023)
45. E. Dumonteil, T.R. Bahran, B. Cutler, T. Dechenaux, J. Grove, G. Hutchinson, A. McKenzie, W. McSpaden, M. Monange, N. Nelson, Thompson, A. Zoia, Patchy nuclear chain reactions. Commun. Phys. **4**, 151 (2021)
46. S.D. Durham, Limit theorems for a general critical branching process. J. Appl. Probab. **8**(1), 1–16 (1971)
47. R. Durrett, *Probability: Theory and Examples* (Duxbury, Belmont, 1996)
48. E.B. Dynkin, Branching particle systems and superprocesses. Ann. Probab. **19**(3), 1157–1194 (1991)
49. E.B. Dynkin, *Diffusions, Superdiffusions and Partial Differential Equations*, vol. 50. American Mathematical Society Colloquium Publications (American Mathematical Society, Providence, 2002)
50. E.B. Dynkin, *Theory of Markov Processes* (Dover Publications, Inc., Mineola, 2006). Translated from the Russian by D. E. Brown and edited by T. Köváry, Reprint of the 1961 English translation
51. M. Eckhoff, A.E. Kyprianou, M. Winkel, Spines, skeletons and the strong law of large numbers for superdiffusions. Ann. Probab. **43**(5), 2545–2610 (2015)
52. K-J. Engel, R. Nagel, *A Short Course on Operator Semigroups*. Universitext (Springer, New York, 2006)
53. J. Engländer, *Spatial Branching in Random Environments and with Interaction*, vol. 20. Advanced Series on Statistical Science & Applied Probability (World Scientific Publishing Co. Pte. Ltd., Hackensack, 2015)
54. J. Engländer, S.C. Harris, A.E. Kyprianou, Strong law of large numbers for branching diffusions. Ann. Inst. Henri Poincaré Probab. Stat. **46**(1), 279–298 (2010)
55. A.M. Etheridge, *An Introduction to Superprocesses*, vol. 20 .University Lecture Series (American Mathematical Society, Providence, 2000)
56. A.M. Etheridge, D.R.E. Williams, A decomposition of the $(1+\beta)$-superprocess conditioned on survival. Proc. Roy. Soc. Edinburgh Sect. A **133**(4), 829–847 (2003)
57. S.N. Ethier, T.G. Kurtz, *Markov Processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics (John Wiley & Sons, Inc., New York, 1986). Characterization and convergence
58. S.N. Ethier, T.G. Kurtz, *Markov Processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics (John Wiley & Sons, Inc., New York, 1986). Characterization and convergence
59. S.N. Evans, Two representations of a conditioned superprocess. Proc. Roy. Soc. Edinburgh Sect. A **123**(5), 959–971 (1993)
60. S.N. Evans, N. O'Connell, Weighted occupation time for branching particle systems and a representation for the supercritical superprocess. Canad. Math. Bull. **37**(2), 187–196 (1994)
61. C.J. Everett, S. Ulam, Multiplicative systems. I. Proc. Nat. Acad. Sci. USA **34**, 403–405 (1948)

62. F. Foutel-Rodier, E. Schertzer, Convergence of genealogies through spinal decomposition with an application to population genetics. Probab. Theory Relat. Fields, pp. 1–55 (Springer, 2023)
63. D. Fekete, J. Fontbona, A.E. Kyprianou, Skeletal stochastic differential equations for continuous-state branching processes. J. Appl. Probab. **56**(4), 1122–1150 (2019)
64. D. Fekete, S. Palau, J.C. Pardo, J.L. Pérez, Backbone decomposition of multitype superprocesses. J. Theoret. Probab. **34**(3), 1149–1178 (2021)
65. J. Fleischman, Limiting distributions for branching random fields. Trans. Am. Math. Soc. **239**, 353–389 (1978)
66. I. Gonzalez, Moments of branching Markov processes and related problems. PhD thesis. University of Bath (2021)
67. I. Gonzalez, E. Horton, A.E. Kyprianou, Asymptotic moments of spatial branching processes. Probab. Theory Related Fields **184**, 805–858 (2022)
68. S.C. Harris, M. Hesse, A.E. Kyprianou, Branching Brownian motion in a strip: survival near criticality. Ann. Probab. **44**, 235–275 (2016)
69. S.C. Harris, E. Horton, A.E. Kyprianou, Stochastic methods for the neutron transport equation II: almost sure growth. Ann. Appl. Probab. **30**(6), 2815–2845 (2020)
70. T.E. Harris, *The Theory of Branching Processes*. Dover Phoenix Editions (Dover Publications, Inc., Mineola, 2002). Corrected reprint of the 1963 original [Springer, Berlin; MR0163361 (29 #664)]
71. S.C. Harris, M.I. Roberts, The unscaled paths of branching Brownian motion. Ann. Inst. Henri Poincaré Probab. Stat. **48**(2), 579–608 (2012)
72. S.C. Harris, E. Horton, A.E. Kyprianou, M. Wang, Yaglom limit for critical neutron transport. Preprint (2021)
73. E. Horton, Stochastic analysis of the neutron transport equation, PhD thesis. University of Bath (2020)
74. E. Horton, A.E. Kyprianou, D. Villemonais, Stochastic methods for the neutron transport equation I: linear semigroup asymptotics. Ann. Appl. Probab. **30**(6), 2573–2612 (2020)
75. N. Ikeda, M. Nagasawa, S. Watanabe, Branching Markov processes. I. J. Math. Kyoto Univ. **8**, 233–278 (1968)
76. N. Ikeda, M. Nagasawa, S. Watanabe, Branching Markov processes. II. J. Math. Kyoto Univ. **8**, 365–410 (1968)
77. N. Ikeda, M. Nagasawa, S. Watanabe, Branching Markov processes. III. J. Math. Kyoto Univ. **9**, 95–160 (1969)
78. I. Iscoe, On the supports of measure-valued critical branching Brownian motion. Ann. Probab. **16**(1), 200–221 (1988)
79. J. Jacod, A.N. Shiryaev, *Limit Theorems for Stochastic Processes*, vol. 288. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 2nd edn. (Springer, Berlin, 2003)
80. P. Jagers, *Branching Processes with Biological Applications*. Wiley Series in Probability and Mathematical Statistics—Applied Probability and Statistics (Wiley-Interscience [John Wiley & Sons], London, 1975)
81. K. Jörgens, An asymptotic expansion in the theory of neutron transport. Comm. Pure Appl. Math. **11**, 219–242 (1958)
82. J.F.C. Kingman, The first birth problem for an age-dependent branching process. Ann. Probab. **3**(5), 790–801 (1975)
83. A. Klenke, Multiple scale analysis of clusters in spatial branching models. Ann. Probab. **25**(4), 1670–1711 (1997)
84. A.N. Kolmogoroff, N.A. Dmitriev, Branching stochastic processes. C. R. (Doklady) Acad. Sci. URSS (N.S.) **56**, 5–8 (1947)
85. A.N. Kolmogorov, Zur lösung einer biologischen aufgabe. Comm. Math. Mech. Chebyshev Univ. Tomsk. **2**, 1–12 (1938)

86. A.E. Kyprianou, Martingale convergence and the stopped branching random walk. Probab. Theory Relat. Fields **116**(3), 405–419 (2000)
87. A.E. Kyprianou, Travelling wave solutions to the K-P-P equation: alternatives to Simon Harris' probabilistic analysis. Ann. Inst. H. Poincaré Probab. Statist. **40**(1), 53–72 (2004)
88. A.E. Kyprianou, S. Palau, Y-X. Ren, Almost sure growth of supercritical multi-type continuous-state branching process. ALEA Lat. Am. J. Probab. Math. Stat. **15**(1), 409–428 (2018)
89. B. Lapeyre, É. Pardoux, R. Sentis, *Introduction to Monte-Carlo Methods for Transport and Diffusion Equations*, vol. 6. Oxford Texts in Applied and Engineering Mathematics (Oxford University Press, Oxford, 2003). Translated from the 1998 French original by Alan Craig and Fionn Craig
90. J.-F. Le Gall, *Spatial Branching Processes, Random Snakes and Partial Differential Equations*. Lectures in Mathematics ETH Zürich (Birkhäuser Verlag, Basel, 1999)
91. J. Lehner, The spectrum of the neutron transport operator for the infinite slab. J. Math. Mech. **11**, 173–181 (1962)
92. J. Lehner, G.M. Wing, On the spectrum of an unsymmetric operator arising in the transport theory of neutrons. Comm. Pure Appl. Math. **8**, 217–234 (1955)
93. J. Lehner, G. M. Wing, Solution of the linearized Boltzmann transport equation for the slab geometry. Duke Math. J. **23**, 125–142 (1956)
94. J. Lewins, Linear stochastic neutron transport theory. Proc. Roy. Soc. London Ser. A **362**(1711), 537–558 (1978)
95. E.E. Lewis Jr., W.F. Miller, *Computational Methods of Neutron Transport* (Wiley & Sons, New York, 1984)
96. L. Lux, L. Koblinger, *Particle Transport Methods: Neutron and Photon Calculations* (CRC Press, Boca Raton, 1991)
97. R. Lyons, A simple path to Biggins' martingale convergence for branching random walk, in *Classical and Modern Branching Processes (Minneapolis, MN, 1994)*, vol. 84. IMA Volumes in Mathematics and its Applications, pp. 217–221 (Springer, New York, 1997)
98. M. Mokhtar-Kharroubi, *Mathematical Topics in Neutron Transport Theory*, vol. 46. Series on Advances in Mathematics for Applied Sciences (World Scientific Publishing Co., Inc., River Edge, 1997). New aspects, With a chapter by M. Choulli and P. Stefanov
99. T. Mori, Neutron transport process on bounded homogeneous domain. Proc. Japan Acad. **46**(9), 944–948 (1970)
100. T. Mori, S. Watanabe, T. Yamada, On neutron branching processes. Publ. Res. Inst. Math. Sci. **7**, 153–179 (1971/1972)
101. S.C. Moy, Ergodic properties of expectation matrices of a branching process with countably many types. J. Math. Mech. **16**, 1207–1225 (1967)
102. T.W. Mullikin, Neutron branching processes. J. Math. Anal. Appl. **3**, 507–525 (1961)
103. T.W. Mullikin, Branching processes in neutron transport theory, in *Probabilistic Methods in Applied Mathematics*, vol. 1, pp. 199–281 (Academic Press, New York, 1968)
104. P. Olofsson, The $x \log x$ condition for general branching processes. J. Appl. Probab. **35**(3), 537–544 (1998)
105. L.I. Pál, Statistical fluctuations of neutron multiplication, in *Proceedings of the 2nd International Conference Geneva*, vol. 18 (1958)
106. L.I. Pál, On the theory of stochastic processes in nuclear reactors, in *Supplemento al vol. VII, series X del Nuovo Cimento. 25* (1958)
107. I. Pázsit, L. Pál, *Neutron Fluctuations: A Treatise on the Physics of Branching Processes* (Elsevier, Amsterdam, 2008)
108. A. Pazy, P.H. Rabinowitz, A nonlinear integral equation with applications to neutron transport theory. Arch. Rational Mech. Anal. **32**, 226–246 (1969)
109. A. Pazy, P.H. Rabinowitz, A nonlinear integral equation with applications to neutron transport theory. Arch. Ration. Mech. Anal. **32**, 226–246 (1969)

110. A. Pazy, P.H. Rabinowitz, On a branching process in neutron transport theory. Arch. Ration. Mech. Anal. **51**, 153–164 (1973)
111. P. Del Moral, Feynman-Kac *Formulae: Genealogical and Interacting Particle Systems with Applications*. Probability and its Applications (New York), pp. xviii+555 (Springer, New York, 2004)
112. E. Powell, An invariance principle for branching diffusions in bounded domains. Probab. Theory Related Fields **173**(3–4), 999–1062 (2019)
113. Y-X. Ren, R. Song, Z. Sun, Spine decompositions and limit theorems for a class of critical superprocesses. Acta Appl. Math. **165**, 91–131 (2020)
114. Y-X. Ren, R. Song, T. Yang, Spine decomposition and $L\log L$ criterion for superprocesses with non-local branching mechanisms. ALEA Lat. Am. J. Probab. Math. Stat. **19**(1), 163–208 (2022)
115. S. Roelly-Coppoletta, A. Rouault, Processus de Dawson-Watanabe conditionné par le futur lointain. C. R. Acad. Sci. Paris Sér. I Math. **309**(14), 867–872 (1989)
116. L.C.G. Rogers, D. Williams, *Diffusions, Markov Processes, and Martingales*, vol. 1. Cambridge Mathematical Library (Cambridge University Press, Cambridge, 2000). Foundations, Reprint of the second (1994) edition
117. L.C.G. Rogers, D. Williams, *Diffusions, Markov Processes, and Martingales*, vol. 2. Cambridge Mathematical Library (Cambridge University Press, Cambridge, 2000). Itô calculus, Reprint of the second (1994) edition
118. Y. Rugama, H. Henriksson, NEA nuclear data services: EXFOR, JANIS and the JEFF project. EPJ Web Conf. **798**, 61–68 (2005)
119. E. Seneta, *Non-negative Matrices and Markov Chains*. Springer Series in Statistics (Springer, New York, 2006). Revised reprint of the second (1981) edition [Springer-Verlag, New York; MR0719544]
120. B.A. Sevast'yanov, Branching stochastic processes for particles diffusing in a bounded domain with absorbing boundaries. Teor. Veroyatnost. i Primenen. **3**, 121–136 (1958)
121. B.A. Sevast'yanov, Transient phenomena in branching stochastic processes. Teor. Veroyatnost. i Primenen **4**, 121–135 (1959)
122. B.A. Sevast'yanov, The extinction conditions for branching processes with diffusion. Teor. Verojatnost. i Primenen. **6**, 276–286 (1961)
123. Z. Shi, *Branching Random Walks*, vol. 2151. Lecture Notes in Mathematics (Springer, Cham, 2015). Lecture notes from the 42nd Probability Summer School held in Saint Flour, 2012, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School]
124. Z. Shi, *Branching Random Walks*, vol. 2151. Lecture Notes in Mathematics, Lecture notes from the 42nd Probability Summer School held in Saint Flour, 2012, École d'Été de Probabilités de Saint-Flour (Springer, Berlin, 2015)
125. Q. Shi, A.R. Watson, Probability tilting of compensated fragmentations. Electron. J. Probab. **24**, Paper No. 78, 39 (2019)
126. A.V. Skorohod, Branching diffusion processes. Teor. Verojatnost. i Primenen. **9**, 492–497 (1964)
127. W.M. Stacey, *Nuclear Reactor Physics*, 3rd edn. (Wiley-VCH, Weinheim, 2018)
128. C. Tucker, *How to Drive a Nuclear Reactor* (Springer, Berlin 2020)
129. S. Watanabe, On the branching process for Brownian particles with an absorbing boundary. J. Math. Kyoto Univ. **4**, 385–398 (1965)
130. A. Weinberg, E. Wigner, *The Physical Theory of Neutron Chain Reactors* (University of Chicago Press, Chicago, 1958)
131. M.M.R. Williams, *Random Processes in Nuclear Reactors* (Elsevier, Amsterdam, 1974)
132. A.M. Yaglom, Certain limit theorems of the theory of branching random processes. Doklady Akad. Nauk SSSR (N.S.) **56**, 795–798 (1947)
133. A. Zoia, E. Dumonteil, A. Mazzolo, Collision densities and mean residence times for d-dimensional exponential flights. Phys. Rev. E **83**, 041137 (2011)
134. A. Zoia, E. Dumonteil, A. Mazzolo, Collision-number statistics for transport processes. Phys. Rev. Lett. **106**, 220602 (2011)

135. A. Zoia, E. Dumonteil, A. Mazzolo, Collision statistics for random flights with anisotropic scattering and absorption. Phys. Rev. E **84**, 061130 (2011)
136. A. Zoia, E. Dumonteil, A. Mazzolo, Residence time and collision statistics for exponential flights: the rod problem revisited. Phys. Rev. E **84**, 021139 (2011)
137. A. Zoia, E. Dumonteil, A. Mazzolo, S. Mohamed, Branching exponential flights: travelled lengths and collision statistics. J. Phys. A: Math. Theor. **45**, 425002 (2012)

Index

E. Horton, A. E. Kyprianou, *Stochastic Neutron Transport*, Probability and Its Applications, https://doi.org/10.1007/978-3-031-39546-8