An Optimisation-Based Approach to FKPP-Type Equations

David Philip Driver

Churchill College
University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy.

Date of Submission: September 2017

Abstract

Title: An Optimisation-Based Approach to FKPP-Type Equations
Author: David Philip Driver

In this thesis, we study a class of reaction-diffusion equations of the form $\frac{\partial u}{\partial t} = \mathcal{L}u + \phi u - \frac{1}{k}u^{k+1}$ where $\mathcal{L}$ is the stochastic generator of a Markov process, ϕ is a function of the space variables and $k \in \mathbb{R}\backslash\{0\}$. An important example, in the case when $k > 0$, is equations of the FKPP-type. We also give an example from the theory of utility maximisation problems when such equations arise and in this case $k < 0$. We introduce a new representation, for the solution of the equation, as the optimal value of an optimal control problem. We also give a second representation which can be seen as a dual problem to the first optimisation problem. We note that this is a new type of dual problem and we compare it to the standard Lagrangian dual formulation.

By choosing controls in the optimisation problems we obtain upper and lower bounds on the solution to the PDE. We use these bounds to study the speed of the wave front of the PDE in the case when $\mathcal{L}$ is the generator of a suitable Lévy process.

Declaration

I hereby declare that, except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared below, in the Acknowledgements, or specified in the text.

Chapter 1 contains a literature review and is based on the work of those cited there. Chapters 2 and 3 and Appendix A is based on work done in collaboration with Dr Michael Tehranchi; see [DT18a] and a future paper [DT18b].

Acknowledgements

First and foremost, many thanks to my supervisor, Dr Michael Tehranchi, for suggesting such an interesting topic to work on and for offering his time to supervise, motivate, and guide me, over the course of my PhD.

Many thanks to my PhD examiners, Nathanael Berestycki and Julien Berestycki, for reading through my thesis and providing useful feedback.

Thank you to the EPSRC and the CCA for funding my research.

I am grateful to the teachers and supervisors who helped to cultivate my enjoyment of learning. My particular thanks go to my previous project supervisors, Dr Florian Theil and Dr Xue-Mei Li, for their advice concerning academia, and to my teachers: Mr McCluskey and Mr Mitchell, who helped me to appreciate the joys of mathematics, Dr Lewis, for the engaging conversations, and Mr Holmes, for encouraging me to aspire to get the best possible education.

I am thankful to Churchill College, the MCR, and the CCA, for giving me a chance to meet so many interesting people and I thank those people for helping to divert me from work and making my time in Cambridge so enjoyable. I thank Alex and Kc for always being around to hang out with and for forcing me to be sociable. I also thank my office-mate Dominic for being there to have an eclectic range of discussions with throughout the long days and my CCA colleagues for the interesting lunch-time conversations and coffee breaks.

I wish to thank Zhen for the many lovely times we spent together and for her support over the course of my PhD. I would also like to thank Yike for always being able to make me laugh and lift my spirits.

Finally, I would like to thank all of my family for always showing confidence in me. I give special thanks to my sister, Kerrie, for keeping me company over the years and playing all of those games with me; my mum, Valerie, for always supporting me; and my dad, Richard, for asking me intriguing maths questions since I was a child. Moreover, thank you to both of my parents, to whom I will be forever indebted, for bringing me into the world and always being there to offer love and invaluable guidance throughout my life.

David Philip Driver

For my loving family

Contents

Chapter 1.

Introduction

1.1. FKPP-Type Equations

In this work, we will study the behaviour of solutions to reaction-diffusion equations of the form

$$\begin{aligned} \frac{\partial u}{\partial t} &= \mathcal{L}u + \phi(x)u - \tfrac{1}{k}u^{k+1} \\ u(0,x) &= u_0(x) \text{ for all } x \in E, \end{aligned} \tag{1.1.1}$$

where the measurable functions $\phi : E \to \mathbb{R}$ and $u_0 : E \to [0,\infty)$ are given, the exponent $k \in \mathbb{R}\backslash\{0\}$ is constant and the operator $\mathcal{L}$ is the generator of a Markov process, X, valued in the state space E. The notion of a solution that we use will be defined rigorously in Chapter 2.

We will call equations of this type, FKPP-type equations. This particular formulation is slightly unusual but is justified since equation (1.1.1) is a generalisation of the equation studied, in 1937, by Fisher, Kolmogorov, Petrovskii, and Piskunov, from whom the equation gets its name. In fact, many equations that are commonly described as being of FKPP- (or simply KPP-) type, are of the form in equation (1.1.1). We note here that the case when $k < 0$ is not usually described as an FKPP-type equation (since the nonlinearity does not satisfy the KPP conditions) but this case fits naturally into the optimisation-based framework that we will introduce below and so in this work we will refer to the equation as an FKPP type equation.

The prototypical FKPP-type equation was first introduced by Fisher, in 1937, to model a biological problem and studied rigorously, in the same year, by Kolmogorov, Petrovskii and Piskunov (KPP). Equations of this form arise in many situations and have been well studied using a diverse set of techniques. These equations have been studied because of their interest to PDE theorists and, since the 1960s, their links to probability. There has been much interplay between these two branches of mathematics and also some effort to reproduce results using only ‘PDE-based techniques’ or only ‘probabilistic techniques’. Over time there have been many generalisations to the original FKPP equation and there is still active research into the behaviour of the solutions to such equations.

We will give a brief history of the study of FKPP-type equations of the form of equation (1.1.1) and some of the applications to modelling physical phenomena.

1.1.1. Applications of FKPP-Type Equations

Fisher's seminal paper [Fis37] was published in 1937. In his paper, Fisher modelled the spread of an advantageous gene throughout a linear habitat. He modelled the frequency of occurrence of the advantageous gene by the equation,

$$\begin{aligned} \frac{\partial u}{\partial t} &= c_1 \frac{\partial^2 u}{\partial x^2} + c_2 u(1-u) \\ u(0,x) &= u_0(x) \end{aligned}$$

where $u : [0,\infty) \times \mathbb{R} \to [0,1]$. $u(t,x)$ represents the frequency of the advantageous gene at time t and position x; $c_1 > 0$ is the coefficient of diffusion and $c_2 > 0$ is the intensity of selection in favour of the advantageous mutant gene. Here the frequency is normalised with respect to a saturation level of 1; that is to say, the u takes values in $[0,1]$.

We would expect that when an advantageous mutation occurs, the change in frequency over time will depend on both the frequency of the advantageous gene and the frequency of the other allelomorph. Here the 'reaction' term is of the form of a simple population growth model, for example the logistic equation. In the FKPP model, there is an extra diffusion term, which describes the way the advantageous gene should spread throughout the population in a wave-like manner over time.

By changing variables $x \mapsto \sqrt{\frac{c_2}{c_1}} x$ and $t \mapsto c_2 t$, we can assume without loss of generality that $c_1 = \frac{1}{2}$ and $c_2 = 1$. We will refer to the equation

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{1}{2} \frac{\partial^2 u}{\partial x^2} + u(1-u) \\ u(0,x) &= u_0(x) \end{aligned} \tag{1.1.2}$$

as the canonical FKPP equation.

It is natural to consider the simplest initial condition where one side of $x = 0$, there is a habitat consisting solely of a population with an advantageous gene adaptation and the other side is the other allelomorph. Therefore the equation is often studied in the case when $u_0(x) = \mathbb{1}_{(-\infty,0)}$.

A survey of the links to biology can be found in the book of Murray [Mur07].

An important summary of various applications of FKPP type equations can be found in the work of Champneys, Harris, Toland, Warren and Williams [CHT+95].

Another interesting application is from physics and the theory of spin glasses. In 1988, Derrida and Spohn [DS88] showed a connection between travelling wave solutions of equation (1.1.2) and disordered trees. Their ideas are extended in more recent papers such as those by Brunet and Derrida [BD09] which also include links between the particles in a branching Brownian motion (cf. Section 1.1.3).

In 2009, Munier and Peschanski [MP03] and later Munier [Mun09] consider the equation (1.1.2) and show how after transformations it describes high energy scattering from the point of view of quantum chromodynamics.

In 2003, del-Castillo-Negrete, Carreras and Lynch [CNCL03] proposed using (1.1.1) with the generator $\mathcal{L}$ being the fractional Laplacian operator to model situations when Gaussian diffusion is unrealistic. Such a situation arises in the study of plasma physics where there can be 'anomalously large particle displacements' which can be described by probability distributions with heavier tails than a Gaussian distribution.

In Chapter 3, we see a new application of equation (1.1.1) in the setting of a Merton type utility maximisation problem.

1.1.2. PDE Approach

In their 1937 paper [KPP37], Kolmogorov, Petrovskii and Piskunov studied the travelling wave behaviour of the equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + f(u) \tag{1.1.3}$$

$$u(0,x) = u_0(x) \tag{1.1.4}$$

under the following assumptions:

$$f \in C^1([0,1]) \tag{1.1.5a}$$

$$f(0) = f(1) = 0$$

$$f(v) > 0, \text{ for all } 0 < v < 1$$

$$f'(0) = \alpha \in (0,\infty) \text{ and } f'(v) < \alpha \text{ for } 0 < v \leq 1. \tag{1.1.5b}$$

Clearly, if $f(u) = u(1-u)$, as in Fisher [Fis37], then f satisfies these conditions.

These conditions on the nonlinearity f are commonly called the KPP conditions. There is some variation in what is meant by the KPP conditions but for the sake of the discussion in this Chapter, we will assume that the preceding conditions hold but some of the references included may use weaker conditions but for ease of exposition we won't mention the precise conditions required in these cases.

Given the biological description, it is natural to consider travelling wave solutions to equation (1.1.3); that is, solutions to the ODE

$$\frac{1}{2}w'' + cw' + f(w) = 0. \tag{1.1.6}$$

If w solves (1.1.6), then, u defined by $u(t,x) = w(x-ct)$ solves the PDE (1.1.3).

In order to obtain uniqueness, one can restrict attention to monotone waves with fixed translation. For example, we only consider $w \in \mathcal{W}$, where

$$\mathcal{W} = \left\{ w \in C^2(\mathbb{R},[0,1]) \mid w' < 0, w(-\infty) = 1, w(0) = \tfrac{1}{2} \text{ and } w(\infty) = 0 \right\}.$$

As suggested by Fisher [Fis37] and proved in [KPP37], there exist travelling waves in $\mathcal{W}$ of speed c if and only if $c \geq \sqrt{2f'(0)}$. We say $c = \sqrt{2f'(0)}$ is the critical or minimal speed. We denote a solution to equation (1.1.6) by w_c for $c > \sqrt{2f'(0)}$ and by w if $c = \sqrt{2f'(0)}$.

The travelling wave of critical speed has the following relation to u: suppose, for example that $u_0 = \mathbb{1}_{(-\infty,0)}$ and set m to be the increasing function that defines the median value of u; that is, $u(t, m(t)) = \frac{1}{2}$. m has the interpretation of the position of the wave front at time t. Then, the following holds:

$$u(t, x + m(t)) \to w(x) \text{ as } t \to \infty.$$

where, w solves equation (1.1.6) with $c = \sqrt{2f'(0)}$.

The precise form of $m(t)$ is of interest, and determining the exact behaviour is still an open problem, even for the canonical FKPP equation. KPP [KPP37], showed the following low order approximation:

$$\frac{m(t)}{t} \to \sqrt{2f'(0)} \text{ as } t \to \infty.$$

Aronson and Weinberger [AW75, AW78] extended these results to more general nonlinearities. They extended the biological application to a population living in $\mathbb{R}^d$ and allowed for intermediate genotypes. For dominant and recessive alleles, A and a, respectively the possible genotypes in this model are aa, aA and AA. In this case the equation modelling the spread of the advantageous gene would be

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{1}{2}\Delta u + f(u) \\ u(0, x) &= u_0(x) \end{aligned} \tag{1.1.7}$$

where

$$f(u) = u(1-u)[\tau_1 - \tau_2 - (\tau_1 - 2\tau_2 + \tau_3)u]$$

where τ_i for $i = 1, 2, 3$, represent the different death rates corresponding to the genotypes, AA, Aa and aa, respectively. Without loss of generality, we assume that $\tau_3 < \tau_1$. Here we are interested in the so called 'Heterozygote Intermediate' case. That is to say, we have $\tau_1 > \tau_2 \geq \tau_3$. In this case, f is of the KPP type.

In their 1978 paper, Aronson and Weinberger [AW78] looked for plane wave solutions, i.e. solutions to equation (1.1.7) of the form $u(t, x) = w(x \cdot \eta - ct)$ for $w \in \mathcal{W}$, where $\eta \in \mathbb{R}^d$ is an arbitrary unit vector and

$$\frac{1}{2}w'' + cw' + f(w) = 0.$$

as in the $d = 1$ case. Again, there are such travelling waves for all speeds $c > c^* = \sqrt{2f'(0)}$.

It was shown that, as long as u_0 is, for example, compactly supported, then

$$\begin{aligned} &\lim_{t\to\infty} u(t, x) = 1, \text{uniformly for} |x| \leq ct, \text{ and } c \in [0, c^*) \\ &\lim_{t\to\infty} u(t, x) = 0, \text{ uniformly for } |x| \geq ct, \text{ and } c \in (c^*, \infty) \end{aligned} \tag{1.1.8}$$

Aronson and Weinberger also treat the 'Heterozygote Inferior' and 'Heterozygote Superior' cases in their 1978 paper.

Uchiyama [Uch78] showed using techniques from the theory of PDEs that

$$m(t) = \sqrt{2}t - \frac{3}{2^{3/2}} \log t + \mathcal{O}(1)$$

for equation (1.1.3) with some conditions on f and with $u_0(x)$ equal to zero for large x. The form of $m(t)$ was also given independently by Bramson [Bra78] using probabilistic methods and this is referred to as the Bramson logarithmic correction term.

Larson [Lar78] considered equation (1.1.3) and compares the super-and-sub-solutions of Montroll [Mon67] and Rosen [Ros74] and considers asymptotic behaviour in the case of exponentially decaying initial conditions in an extension of the results of McKean [McK75].

Lau [Lau85] gave a simplified proof of the form of the Bramson correction term in the case of a more general forcing term f using maximum principle methods.

More recently, Hamel and Roques [HR10] studied equation (1.1.3) under the additional assumption that $f(s) \geq f'(0)s - Ms^{1+\delta}$ for $s \in [0, s_0]$, for some $\delta > 0$, $s_0 \in (0,1)$ and $M \geq 0$. In their paper they look at the previously unstudied slowly decaying initial conditions. In particular they show that under (not necessarily decreasing) initial conditions, such that for any $\varepsilon > 0$, $u_0(x) \geq e^{-\varepsilon x}$ for large x. This case is of interest because, if

$$m^\lambda(t) = \{x \in \mathbb{R} : u(t,x) = \lambda\}, \tag{1.1.9}$$

then for any fixed $\lambda \in (0,1)$,

$$\lim_{t\to\infty} \frac{1}{t} \min m^\lambda(t) = \infty$$

which is in contrast to the linear speed of the wave front the case when the initial condition has exponential decay. Hamel and Roques study the precise asymptotic behaviour of the level sets m^λ in this case.

Hamel, Nolen, Roquejoffre and Ryzhik [HNRR13] gave a short proof of the form of the Bramson correction term for rather general conditions. Nolen, Roquejoffre and Ryzhik [NRR17] showed an improved result in a more specialised setting before refining the result in their next paper [NRR16]. In their paper [NRR16], they studied the speed of the wave front for the solution to (1.1.2), given that, for example, u_0 decreasing with $u_0(x) \leq \exp(-\beta x)$ for $\beta > \sqrt{2}$ and x sufficiently large (cf. condition (1.1.12) below). They improved a conjecture of Ebert and van Saarloos [ES98] and van Saarloos [Saa03], and showed that

$$m(t) = \sqrt{2}t - \frac{3}{2^{3/2}} \log t + c_0 - \frac{3\sqrt{\pi}}{\sqrt{2t}} + o\left(\frac{1}{t^{1-\varepsilon}}\right)$$

for all $\varepsilon > 0$, where c_0 depends on the initial condition, u_0. Ebert and van Saarloos suggested this result with $\varepsilon = 1/2$. Interestingly, the coefficient of the $t^{-1/2}$ term does not depend on u_0. Under similar conditions, Berestycki and Brunet [BB16] also studied equation (1.1.2). In this case the process $m^\lambda(t)$ such that $u(t, m^\lambda(t)) = \lambda$ is single valued. Define x_λ to be such that $w(x_\lambda) = \lambda$ where w solves the travelling wave ODE. Berestycki and Brunet conjecture a further refinement on the particular form of $m^\lambda(t)$ and suggest that

$$m^\lambda(t) = \sqrt{2}t - \frac{3}{2^{3/2}} \log t + C + x_\lambda - \frac{3\sqrt{\pi}}{\sqrt{2t}} + K\frac{\log t}{t} + \mathcal{O}\left(\frac{1}{t}\right)$$

where C and K are constants that do not depend on λ.

1.1.3. Links to Probability

The FKPP equation has often been studied in a probabilistic setting by using an observation, such as the following, linking branching processes and PDEs. Skorokhod [Sko64] was perhaps the first to notice such a connection in the context of a general branching diffusion process and gave equations for the transition probabilities and conditions on when solutions exist. Later, Watanabe [Wat68] studied branching processes in relation to PDEs in a very general setting.

Let

$$G(s) = \sum_{i=0}^{\infty} p_i s^i$$

be a probability generating function of a positive integer valued random variable. For suitably regular v_0 and X, the function

$$v(t, x) = \mathbb{E}^x \left[\prod_{i \in I_t} v_0(X_t^i) \right] \tag{1.1.10}$$

can been shown to solve the equation

$$\begin{aligned} \frac{\partial v}{\partial t} &= \mathcal{L}v + G(v) - v \\ v(0, x) &= v_0(x) \text{ for all } x \in E \end{aligned}$$

where $\{X_t^i, i \in I_t, t \geq 0\}$ is a branching process constructed from the Markov process X with generator $\mathcal{L}$ as follows: initially there is one particle located at x, and moves according to the law of X. At an exponentially distributed time $T \sim \exp(1)$, this particle disappears and is replaced with N independent and identical copies located at same position, where $\mathbb{P}(N = n) = p_n$. This procedure then repeats. We are using the notation I_t for the index set of the particles alive at time t.

The function defined by $u = 1 - v$ then satisfies then satisfies a PDE with nonlinearity $f(u) = 1 - u - G(1 - u)$ which we can see if of the KPP type (for G continuously differentiable).

The first person to use a branching process representation to explicitly study an FKPP-type equation was McKean. In 1975, McKean [McK75, McK76] studied the simplest branching diffusion process which corresponds to the canonical FKPP equation. Independently of Skorokhod (cf. equation (1.1.10)), McKean wrote the solution to the equation

$$\begin{aligned} \frac{\partial v}{\partial t} &= \frac{1}{2} \frac{\partial^2 v}{\partial x^2} + v(v - 1) \\ v_0(x) &= \mathbb{1}_{(0,\infty)} \end{aligned}$$

as

$$v(t, x) = \mathbb{P}\left(\max_{i \in I_t} X_t^i \leq x \right)$$

where X is a standard Brownian motion, and $N \equiv 2$. In other words, the process is a dyadic branching Brownian motion (BBM). We see that $v(t, \cdot)$ is the distribution function of the maximal particle at time t.

After substituting $v \mapsto 1 - u$, we see that this is precisely of the form in equation (1.1.2) and rewriting we have

$$u(t, x) = \mathbb{P}\left(\max_{i \in I_t} X_t^i \geq x \right) \tag{1.1.11}$$

McKean simplified the proof of KPP [KPP37] that $u(t, x + m(t))$ converges to the travelling wave solution of critical speed for more general initial conditions.

McKean also studied more general initial conditions: for u_0 taking values in $[0, 1]$, and $b \in (0, \sqrt{2}]$, $\lim_{x\to\infty} u_0(x)e^{bx} = a$, for some $a \geq 0$, then,

$$\lim_{t\to\infty} u(t, x + ct) = w_c(x)$$

where w_c solves the travelling wave ODE (1.1.6) with $c = \frac{1}{b} + \frac{b}{2}$.

If we, consider the case when $u_0 = \mathbb{1}_{(-\infty,0)}$ and write $R_t = \max_{i\in I_t} X_t^i$, we see that in this probabilistic framework, the convergence $u(t, x+m(t)) \to w(x)$ can be interpreted as a statement about convergence in distribution of $R_t - m(t)$. Also, by knowing the speed of m, we have a law of large numbers for a BBM in the sense that

$$\frac{R_t}{t} \xrightarrow{P} \sqrt{2}, \text{ as } t \to \infty.$$

McKean also obtained an upper bound on m:

$$m(t) \leq \sqrt{2}t - 2^{-3/2} \log t.$$

Bramson [Bra78, Bra83] (see also the review article [Bra86]) extended these results to a BBM with an average of two branches and finite second moment:

$$\sum_{i=1}^{\infty} ip_i = 2$$
$$\sum_{i=1}^{\infty} i^2 p_i < \infty$$

and showed that McKean's bound on m was not sharp. In fact, it was shown that

$$m(t) = \sqrt{2}t - 3 \cdot 2^{-3/2} \log t + \mathcal{O}(1).$$

if and only if

$$\int_0^\infty ye^{\sqrt{2}y} u_0(y) < \infty \tag{1.1.12}$$

Bramson also extended the results to more general nonlinearities by a comparison principle argument. Finally, Bramson improved upon the results of McKean [McK75] and Uchiyama [Uch78] and gave necessary and sufficient conditions for an initial condition to give rise to travelling waves of a certain speed. It was shown that

$$u(t, x + m(t)) \to w(x)$$

uniformly in x as $t \to \infty$, if and only if the Bramson conditions on u_0 hold: for some $h, \eta, M, N > 0$

$$\limsup_{z\to\infty} \frac{1}{z} \log \int_z^{z(1+h)} u_0(y)\mathrm{d}y \leq -\sqrt{2}$$

and

$$\int_z^{z+N} u_0(y)\mathrm{d}y > \eta \tag{1.1.13}$$

for all $z \le -M$. Similarly, for $\tilde{m}(t) = ct + o(t)$,

$$u(t, x + \tilde{m}(t)) \to w_c(x)$$

uniformly in x as $t \to \infty$ if and only if condition (1.1.13) holds and

$$\limsup_{z\to\infty} \frac{1}{z} \log \int_z^{z(1+h)} u_0(y)\mathrm{d}y = -b.$$

A multidimensional analogue of Bramson's results for spherically symmetric initial conditions was found by Gärtner [Gär82].

In 1988, Chauvin and Rouault [CR88] gave another probabilistic interpretation. One can rewrite the representation (1.1.11) as $u(t,x) = \mathbb{P}(Z_t((x,\infty)) > 0)$ where $Z_t(A)$ counts of the number of particles of, for example, a standard branching Brownian motion, in the set A at time t for suitable $A \subset \mathbb{R}$.

Chauvin and Rouault studied the large deviations of the right-most particle process R_t; that is $\mathbb{P}(R_t \ge ct)$ for $c > \sqrt{2}$. To obtain their results, they exploit the relationship between this and a sub-critical Galton-Watson process $(\zeta_n)_{n\ge 0}$; recall, that in this case, if $\mathbb{E}[\zeta_1 \log \zeta_1] < \infty$, as $n \to \infty$ then $\mathbb{P}(\zeta_n > 0) \sim C\mathbb{E}[\zeta_n]$ for $C > 0$. Analogously,

$$\mathbb{P}(Z_t((ct + x, \infty) > 0) \sim C\mathbb{E}[Z_t((ct + x, \infty))] \text{ as } t \to \infty,$$

for some constant $C > 0$.

In terms of the PDE, this means that for $c > \sqrt{2}$,

$$u(t, ct + x) \sim CU(t, ct + x) \text{ as } t \to \infty,$$

for a constant $C > 0$, where U is the solution of the linearised PDE

$$\frac{\partial U}{\partial t} = \frac{1}{2}\frac{\partial^2 U}{\partial x^2} + U.$$

Note that Chauvin and Rouault state their results for more general branching Brownian processes.

Harris [Har98] showed existence and uniqueness (up to translation) of monotone travelling wave solutions $w \in \mathcal{W}$ for the canonical FKPP equation (1.1.2), using *only* probabilistic techniques. Harris constructed the travelling wave by extending the martingale approach of Neveu [Nev88], Biggins[Big92] and Champneys et al. [CHT$^+$95].

The travelling wave of speed $c > \sqrt{2}$ is given in terms of the 'additive' martingale

$$Z(t) = \sum_{i \in I_t} e^{-\lambda(X_t^i + \lambda t)}.$$

where $\lambda = c - \sqrt{c^2 - 2}$. In particular, the travelling wave is given by

$$1 - w(x) = \mathbb{E}\left[\exp\left(-e^{-\lambda x} Z(\infty)\right)\right].$$

The travelling wave of critical speed is given in terms of the 'derivative' martingale

$$Z'(t) = \sum_{i \in I_t} (X_t^i + \sqrt{2}t) e^{-\sqrt{2}(X_t^i + \sqrt{2}t)}$$

and the travelling wave is given by

$$1 - w(x) = \mathbb{E}\left[\exp\left(-e^{-\sqrt{2}x} Z'(\infty)\right)\right].$$

Kyprianou [Kyp04] gave an alternative to Harris' methods also using purely probabilistic approach to Harris' methods. The 'spine decomposition' approach of Kyprianou is shown for more general branching processes of a similar form to Bramson.

1.1.4. More General Generators

So far we have only considered equation (1.1.1) in the case when the generator $\mathcal{L}$ is a Laplacian and the corresponding Markov process X is a Brownian motion. More recently, there has been interest in studying equation (1.1.1) and the corresponding branching processes for more general Markov processes.

Non-local models of equations for the spread of a population or epidemic have a long history. Schumacher [Sch80] showed a travelling wave solutions to a class of integro-differential equations and Carr and Chmaj [CC04] showed uniqueness of travelling waves for the equation

$$\frac{\partial u}{\partial t} = H * u - u + f(u) \tag{1.1.14}$$

where $H * v(x) = \int_{\mathbb{R}} H(x-y)v(y)\mathrm{d}y$ for a function $v : \mathbb{R} \to \mathbb{R}$ where f is a KPP nonlinearity and $H : \mathbb{R} \to [0, \infty)$ is compactly supported, $H(x) = H(-x)$ and $\int_{\mathbb{R}} H(y)\mathrm{d}y = 1$.

The generator defined by $\mathcal{L}u = H*u$ corresponds to a compound Poisson process (c.f. Example (4.4.3)).

In 2013, Cabré and Roquejoffre [CR13] studied equations such as

$$\begin{aligned}\frac{\partial u}{\partial t} &= -(-\Delta)^{\alpha/2} u + f(u)\\ u(0,x) &= u_0(x)\end{aligned}$$

on $(0, \infty) \times \mathbb{R}^d$ where $-(-\Delta)^{\alpha/2}$ denotes the fractional Laplacian for $\alpha \in (0, 2)$, f is a KPP nonlinearity and $0 \leq u_0 \leq 1$. They showed that if $u_0(x) = \mathcal{O}(|x|^{-d-\alpha})$ for large $|x|$, then

$$\begin{aligned}&\lim_{t\to\infty} u(t,x) = 1, \text{uniformly for} |x| \leq e^{ct}, \text{ and } c \in [0, c^*)\\ &\lim_{t\to\infty} u(t,x) = 0, \text{ uniformly for } |x| \geq e^{ct}, \text{ and } c \in (c^*, \infty)\end{aligned}$$

where $c^* = \frac{f'(0)}{d+\alpha}$ and thus proved an analogous result to Aronson and Weinberger cf. (1.1.8) and the later results for exponentially decaying initial conditions. Notice, however, that in this case the wave front moves with exponential speed rather than with linear speed.

In the case when $d = 1$ and $u_0(x) \leq x^{-\alpha}$ for large x and $u_0 \not\equiv 0$ is decreasing on $\mathbb{R}$, the wave front moves with speed $\frac{f'(0)}{\alpha} > c^*$. This faster spread for non-compact initial data is in contrast to the case when the diffusion is Gaussian (cf. Aronson and Weinberger [AW78]).

Cabré and Roquejoffre also consider level sets analogous to m^λ in equation (1.1.9) in the special case of $f(u) = u(1-u)$ and $d = 1$.

In 1999, Kyprianou [Kyp99] studied branching Lévy processes in the sense of Section 1.1.3 but with the more general assumption that, upon splitting, the child particles are scattered according to a random process. This built upon the work of Neveu [Nev88] and Chauvin [Cha91] in the setting of branching Brownian motion and Biggins [Big95, Big97] general discrete branching processes. In 2016, Groisman and Jonckheere [GJ16], use Biggins' [BLSW91] and Kyprianou's [Kyp99] results to generalise the work of Harris [Har98] to show the existence of travelling waves for the equation

$$\begin{aligned}\frac{\partial u}{\partial t} &= \mathcal{L}u + ru(1-u)\\ u(0,x) &= u_0(x)\end{aligned}$$

with speeds c such that $\Lambda^*(c) \geq r$ where Λ^* is the Legendre transform of the Laplace exponent of $\mathcal{L}$ under some conditions including that the underlying Lévy process has non-zero diffusion part and X_1 has exponentially decaying tails.

1.2. Summary of New Results

1.2.1. The Value of k

In order to treat such a general equation as equation (1.1.1) in a unified manner, we try to make as few assumptions as possible. We will see that to establish the optimisation results, there are three cases to consider, when $k > 0$, $-1 < k < 0$ and $k < -1$, which must be treated separately.

Note that by setting $\phi = 1/k$ constant in equation (1.1.1), the non-linearity becomes $f(u) = \frac{1}{k}u(1-u^k)$.

1. When $k > 0$, this is of the KPP form. Note f is only Lipschitz on $[0, 1]$, when $k > 0$.
2. When $k \in (-1, 0)$, f is not of the KPP form. However, f satisfies the KPP conditions with (1.1.5a) replaced by $f \in C^1((0, 1)) \cap C^0([0, 1])$ and condition (1.1.5b) modified with $\alpha = \infty$.
3. When $k < -1$, the form of the nonlinearity is very different from the nonlinearities satisfying the KPP conditions.

Furthermore, in the case where $-1 < k \leq 1$, the FKPP equation is directly amenable to the branching representation discussed above in Section 1.1.3. Here, the probability generating function G is of the form

$$\begin{aligned}G(v) &= v + \tfrac{1}{k}[(1-v)^{1+k} + v - 1]\\ &= \frac{(1+k)}{2!}v^2 + \frac{(1+k)(1-k)}{3!}v^3 + \frac{(1+k)(1-k)(2-k)}{4!}v^4 + \dots\end{aligned}$$

This is a probability generating function since $G(1) = 1$.

We will avoid the degenerate cases $k = -1$ and $k = 0$. In this case we see that equation (1.1.1) is a linear equation.

1.2.2. The Results

Now, we briefly summarise the new results that are presented in this work. Suppose that $u : [0, \infty) \times E \to [0, \infty)$ solves equation (1.1.1) in a sense to be made precise in Chapter 2. Then, we show that u can be computed in two ways: as the value function of a certain maximisation problem and as the value function of a related minimisation problem. These representations make up what we refer to as the *optimisation-based representations* of equation (1.1.1).

One potential application of this approach is in providing a systematic method of obtaining upper and lower bounds to solutions of FKPP-type equations. Indeed, we can simply choose an arbitrary control to insert into the objective function of the corresponding optimisation problem and obtain a bound. As demonstrated in Section 1.1, there are many examples of practical interest and theoretical interest. Not all cases of interest are covered by this framework but an extension to nonlinearities coming from the theory of branching processes is also considered in Appendix A.

As we saw, the classical and perhaps most important example is when the state space $E = \mathbb{R}^d$ and $\mathcal{L} = \frac{1}{2}\Delta$ is the Laplacian, in which case the Markov process is Brownian motion. However, the Markov process plays no important part in the derivation of our representations. In fact, we will see that our optimisation-based representations hold with no assumption on $\mathcal{L}$ other than that equation (1.1.1) has a solution in a rather weak sense which is made precise in Definition 2.1.1.

Another important situation in which equations of the form in equation (1.1.1) naturally arise, is in the theory of financial mathematics. We will see that this equation appears in the study of a certain optimal investment problem in the context of a model of a market with stochastic volatility and an investor with constant relative risk aversion. Indeed, equation (1.1.1) arises from applying the nonlinear transformation proposed by Zariphopoulou [Zar01] to the Hamilton-Jacobi-Bellman (HJB) equation associated with the optimisation problem. Details of this example will be provided in Chapter 3. In this application, the exponent k appearing equation (1.1.1) is defined by

$$k = -\frac{1}{R(1-\rho^2)+\rho^2}$$

where $R > 0$ is the investor's coefficient of relative risk aversion and ρ is the correlation between the infinitesimal increments of the price of the risky asset and its volatility. Furthermore, in this setting, the coefficient ϕ, in equation (1.1.1), is a linear combination of the interest rate, the Sharpe ratio of the stock and rate of subjective discounting. In particular, the exponent k is *negative* and the coefficient ϕ is generally a non-constant function of the state variable. These observations motivate the generality in which we study equation (1.1.1).

In the special case of equation (1.1.1) when $\mathcal{L}$ is the generator of a suitable Lévy process, u_0 an initial condition with fast enough decay, and $k > 0$, one expects that wave fronts will develop and by studying the upper and lower bounds on the solution one can find the speed of this front. By using the optimisation-based representation, we will derive such bounds and study the wave speed for a variety of choices of generator, $\mathcal{L}$ and initial condition u_0. Further details can be found in Chapter 4 and Appendix A.

1.2.3. The Outline

The rest of this work is organised as follows: in Chapter 2, we present the precise formulation of the mathematical framework in which we will work and the representations of the solution to the equation we will study. In Section 2.2, we present the first main result which is a representation of the solution to an equation of the form (1.1.1) in terms of an optimisation problem which we will call the primal problem. In Section 2.3, we introduce a second representation as another optimisation problem which we will see as a dual problem in a way to be made precise. We show the duality is a strong duality. After giving proofs of these representations in Section 2.4, we will present some immediate consequences of the representations and compare the dual problem to the Lagrangian dual problem in Sections 2.6 and 2.5, respectively. In Section 2.7, we will give sufficient conditions for existence of solutions to equation (2.1.1).

In Chapter 3, we set up a Merton-type utility maximisation problem in Section 3.1. Then, in Section 3.2, we show the relationship between the corresponding HJB equation and FKPP-type equations in the form of equation (1.1.1) and thus the representations of Chapter 2.

Finally, in Chapter 4, we use the representations of Chapter 2 to calculate the speed of the wave front for various examples of equation (1.1.1) when the nonlinearity is given by $f(u) = \frac{1}{k}u(1-u^k)$. In Section 4.1, we deduce some useful preliminary results from Chapter 2 and then in Section 4.2 we apply these results to prove Aronson-and-Weinberger-type results in the case when the Markov process is a standard Brownian motion. Then, we give analogous results for a class of Lévy processes in Section 4.3. We apply these results to some examples in Section 4.4. In Sections 4.4.1 and 4.4.2, we consider two specific examples of X that don't fit into the class of Lévy processes of Section 4.3 but where results can still be obtained by slight modifications in the arguments.

In Appendix A, we show how the primal representation given in Chapter 2 can be adapted to a class of equations with concave nonlinearity. This allows us to consider equations corresponding to a more general class of branching processes than those covered by Chapter 4. We briefly explain how the techniques of Chapter 4 can be adapted to this important case.

In Appendix B we give an alternative view of the primal optimisation problem of Section 2.2 and study the problem in a simplified setting using and see how the optimisation problem gives rise to equation (1.1.1) via the dynamic programming principle. We also present a direct proof of strong duality in the sense of Lagrangian duality. We highlight the technical challenges inherent in this stochastic control approach.

Chapter 2.

The Representations

2.1. The Set-Up

For each t we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $(\mathcal{F}_s)_{s\geq 0}$ and a Markov process $(X_s)_{0\leq s\leq t}$ taking values in a Borel space $(E, \mathcal{B}(E))$. We use the usual notation of $\mathbb{P}^x$ for $\mathbb{P}(\cdot|X_0 = x)$ and, analogously, we define $\mathbb{E}^x$ but we may drop the x from the notation when the meaning is clear from the context.

Usually one considers the case when $E = \mathbb{R}^d$ and takes X to be a Lévy process or, more generally, a Feller process, but, as we will see, the properties of X do not play a large role in the results of this section.

As noted in the introduction we will study the following equation:

$$\begin{aligned} \frac{\partial u}{\partial t} &= \mathcal{L}u + \phi(x)u - \tfrac{1}{k}u^{k+1}, \\ u(0,x) &= u_0(x), \text{ for all } x \in E. \end{aligned} \tag{2.1.1}$$

We suppose that $\phi : E \to \mathbb{R}$ and $u_0 : E \to [0, \infty)$ are measurable functions and are known. The exponent $k \in \mathbb{R}\backslash\{-1, 0\}$ is constant. We define the operator $\mathcal{L}$ to be the infinitesimal generator of the Markov process X.

We make the following definition of a solution.

Definition 2.1.1. *We say that a measurable function $u : [0,\infty) \times E \to [0,\infty)$ is a solution to equation* (2.1.1) *if for every (t, x) there exists a measurable E-valued process $(X_s)_{0\leq s\leq t}$ defined on some probability space with $X_0 = x$ almost surely and adapted to a filtration satisfying the usual conditions such that*

$$u(0, x) = u_0(x)$$

and such that the process $(M^*_s)_{0\leq s\leq t}$ defined by

$$M^*_s = u(t-s, X_s)e^{\int_0^s [\phi(X_r) - \frac{1}{k}u(t-r,X_r)^k]\mathrm{d}r} \tag{2.1.2}$$

is a càdlàg martingale.

Remark 1. *Intuitively, one may think of Definition 2.1.1 as replacing the assumption of smoothness of a classical solution with an assumption of integrability. Indeed, consider the following result.*

If u is a non-negative classical solution to equation (2.1.1) *with generator $\mathcal{L}$ corresponding to a Markov process X, then M^* is a local martingale.*

Indeed, by applying Itô's formula, we see that

$$\begin{aligned} \mathrm{d}u(t-s, X_s) = &= (-u_t + \mathcal{L}u)\mathrm{d}s + \mathrm{d}N_s \\ &= (-\phi u + \tfrac{1}{k}u^{k+1})\mathrm{d}s + \mathrm{d}N_s \end{aligned}$$

a local martingale N, since u is a classical solution to equation (2.1.1)*. Thus,*

$$\mathrm{d}\left(u(t-s, X_s)e^{\int_0^s [\phi(X_r) - \frac{1}{k}u(t-r,X_r)^k]\mathrm{d}r}\right) = e^{\int_0^s [\phi(X_r) - \frac{1}{k}u(t-r,X_r)^k]\mathrm{d}r}\mathrm{d}N_s.$$

Integrating both sides, we see that left hand side is a local martingale, since the integrand is continuous, as required.

We also see that if u is a classical solution and u and ϕ are bounded, as is often the case, then M^* is a true martingale. ◇

If u is a solution of equation (2.1.1) in the sense of Definition 2.1.1, then M^* is a martingale and so we have the following representation of u:

$$u(t,x) = \mathbb{E}\left[u_0(X_t)e^{\int_0^t [\phi(X_r)\mathrm{d}r - \frac{1}{k}u(t-r,X_r)^k]}\right]. \tag{2.1.3}$$

In other words, a generalised solution in the sense of Freidlin [Fre85, Chapter 5.1] is also a solution of the form of Definition 2.1.1.

In order to elucidate Definition 2.1.1 further, we look at some examples of equations of the form in equation (2.1.1).

Example 2.1.2. *For the following examples, we take $E = \mathbb{R}^d$.*

1. a) *We can consider X to be a Lévy process starting at $x \in \mathbb{R}^d$. It then follows that the operator $\mathcal{L}$ is of the integro-differential form*

$$\begin{aligned} \mathcal{L}f = &\sum_{1 \le i \le n} b_i \frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{1 \le i,j \le n} a_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j} \\ &+ \int_{\mathbb{R}^d \setminus \{0\}} \left[f(x+y) - f(x) - \mathbb{1}_{\{\|y\| \le 1\}} \sum_{1 \le i \le n} y_i \frac{\partial f}{\partial x_i}(x)\right] \nu(\mathrm{d}y) \end{aligned}$$

   for constant b_i, a_{ij} such the matrix $(a_{ij})_{ij}$ is non-negative definite and the measure ν such that $\int_{\mathbb{R}^d \setminus \{0\}} \|y\|^2 \wedge 1 \ \nu(\mathrm{d}y) < \infty$, corresponding to the Lévy process with characteristic triple (b, a, ν). See, for instance, [App04].

   b) *The simplest example is when X is an d-dimensional Brownian motion starting at $x \in \mathbb{R}^d$. Then, $\mathcal{L} = \frac{1}{2}\Delta$ is a multiple of the Laplacian. Then equation* (2.1.1) *is the standard FKPP equation considered in Aronson and Weinberger [AW78], for example.*

   c) *In Chapter 4 we will restrict our attention to the univariate case (d=1), where the Lévy measure ν has a finite exponential moment such that*

$$\int e^{\theta y}\mathbb{1}_{\{|y|>1\}}\nu(\mathrm{d}y) < \infty$$

   for all θ in a neighbourhood of 0.

d) *Another special case is when $b = 0$ and $a = 0$ and the Lévy measure ν has a density proportional to $\|y\|^{-(\alpha+d)}$ for $0 < \alpha < 2$. With this choice, the operator is the fractional Laplacian $\mathcal{L} = -(-\Delta)^{\alpha/2}$ corresponding to a rotationally symmetric α-stable process. This corresponds to the equation considered in the work of Cabré and Roquejoffre [CR13].*

2. a) *Another direction to generalise is to take the operator $\mathcal{L}$ to be an elliptic differential operator of the form*

$$\mathcal{L} = \sum_{1 \le i \le n} b_i(x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{1 \le i,j \le n} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j}$$

for suitably regular functions $b_i, a_{ij} : \mathbb{R}^d \to \mathbb{R}$ corresponding to the Markov process X solving the stochastic differential equation

$$\mathrm{d}X_s = b(X_s)\mathrm{d}s + \sigma(X_s)\mathrm{d}W_s$$

where $a = \sigma\sigma^\intercal$ and W is a Brownian motion. We consider the univariate case of such a diffusion in the financial application in Chapter 3.

b) *In Chapter 4 we look at the particular example when X is an Ornstein-Uhlenbeck process.*

As demonstrated in these examples, the operator $\mathcal{L}$ is not everywhere-defined. In particular, if we look at classical solutions to equation (2.1.1) we must assume a certain amount of regularity of u in order for $\mathcal{L}u$ to make sense.

In proving the following representation formulas, we will not need to use the special structure of $\mathcal{L}$, ϕ or u_0 other than the assumption that M^* is a càdlàg martingale. In this chapter, we will mostly keep these unspecified. The statement and proof of Theorems for existence and uniqueness of solutions are deferred to the end of this Chapter. For now, we will simply *assume* that equation (2.1.1) has a solution in the sense of Definition 2.1.1.

There are three natural cases to consider in terms of k. The non-linearity changes behaviour at $k = -1$ and $k = 0$ and so the results in Chapter 2 will usually be stated in each of these cases. Note also that when $k < 0$, the integral term involving $u(t-r, X_r)^k$ in the notion of solution blows up if u is allowed to vanish. This corresponds to the nonlinearity u^{1+k} failing to be Lipschitz at $u = 0$. Therefore, in the case that $k < 0$, we will assume that u is strictly positive everywhere, to avoid these problems. We include this assumption and one more simplifying assumptions throughout.

In summary, we always assume the following.

Assumption 2.1.3. *1. $u_0(x) \ge 0$ for all $x \in E$.*

2. For each fixed k, we have

$$\int_0^t |\phi(X_s)|\mathrm{d}s < \infty \text{ and } \int_0^t u(t-s, X_s)^k \mathrm{d}s < \infty \text{ almost surely.}$$

3. *A solution, u, to equation* (2.1.1) *exists in the sense of Definition 2.1.1.*

Sufficient conditions for the third assumption to hold are given in Section 2.7.

2.2. Primal Representation

We will now introduce the first representation of the solution u in terms of an optimisation problem. We refer to this as the primal representation; this refers to the fact that the optimisation problem has a natural interpretation where u is the value function of a certain optimisation problem and then equation (2.1.1) is the corresponding HJB equation. Moreover, in we will introduce another representation in Section 2.3 which can be seen as a dual problem.

We begin with an Feynman-Kac-type formula which will motivate the representations to follow. Recall that a Feynman-Kac result is one of the form

Theorem 2.2.1 (Feynman-Kac Theorem)**.** *Let u be a classical solution to*

$$\begin{aligned}\frac{\partial u}{\partial t}(t,x) &= \mathcal{L}u(t,x) + V(t,x)u(t,x) + f(t,x),\\ u(0,x) &= u_0(x).\end{aligned}$$

Given sufficiently regular $\mathcal{L}, V$ and f and u_0, (see, for example [IW89]),

$$u(t,x) = \mathbb{E}^x\left[\int_0^t e^{\int_0^s V(r,X_r)\mathrm{d}r} f(s,X_s)\mathrm{d}s + e^{\int_0^t V(r,X_r)\mathrm{d}r} u_0(X_t)\right].$$

Given sufficiently nice $\mathcal{L}, V$ and f and u_0, the result follows easily from Itô's formula. One can see that the result can be extended to the case when V and f depend on the solution u.

Now we can introduce the starting point for our discussion. Notice that the following Feynman-Kac type result holds for a solution to equation (2.1.1). Define a process $(Y^*_s)_{0\le s\le t}$ by

$$Y^*_s = e^{-\int_0^s u(t-r,X_r)^k \mathrm{d}r}. \tag{2.2.1}$$

Note that Y^* is positive, non-increasing, absolutely continuous and adapted.

There is an important link between the process Y^* and the martingale M^* from Definition 2.1.1. Two important properties that will simplify the following calculations are

$$e^{\int_0^t \phi(X_r)\mathrm{d}r}(Y^*_t)^{\frac{1}{k}} u_0(X_t) = M^*_t \tag{2.2.2}$$

and

$$e^{\int_0^s \phi(X_r)\mathrm{d}r}(-\dot{Y}^*_s)^{\frac{1}{k}} = M^*_s \text{ for } 0 \le s \le t \tag{2.2.3}$$

Here we are using the notation $\dot{h}$ to denote the weak derivative of an absolutely continuous function h. Notice that

$$\dot{Y}^*_s = -u(t-s,X_s)^k Y^*_s.$$

As the notation suggests, the processes Y^* and M^* will play important roles in the discussion that follows.

The following lemma is a motivation for the primal representation in Theorem 2.2.3.

Lemma 2.2.2. *Fix (t,x) and let the process X be as in Definition 2.1.1. Then,*

$$u(t,x) = \mathbb{E}\left[\int_0^t e^{\int_0^s \phi(X_r)\mathrm{d}r}|\dot{Y}_s^*|^{1+\frac{1}{k}}\mathrm{d}s + e^{\int_0^t \phi(X_r)\mathrm{d}r}|Y_t^*|^{1+\frac{1}{k}}u_0(X_t)\right] \tag{2.2.4}$$

Proof. Fix t and x. Since M^* is a martingale with $M_0^* = u(t,x)$, we have

$$\begin{aligned}
u(t,x) &= \mathbb{E}[M_t^*] \\
&= \mathbb{E}\left[M_t^*(1 - Y_t^* + Y_t^*)\right] \\
&= \mathbb{E}\left[-\int_0^t \dot{Y}_s^* M_t^* \mathrm{d}s + Y_t^* M_t^*\right] \\
&= \mathbb{E}\left[\int_0^t e^{\int_0^s \phi(X_r)\mathrm{d}r}|\dot{Y}_s^*|^{1+\frac{1}{k}}\mathrm{d}s + e^{\int_0^t \phi(X_r)\mathrm{d}r}|Y_t^*|^{1+\frac{1}{k}}u_0(X_t)\right].
\end{aligned}$$

To go from the second to the third line, we used the tower property of conditional expectation:

$$\begin{aligned}
\mathbb{E}\left[-\int_0^t \dot{Y}_s^* M_t^* \mathrm{d}s\right] &= \int_0^t \mathbb{E}\left[-\dot{Y}_s^* \mathbb{E}[M_t^*|\mathcal{F}_s]\right]\mathrm{d}s \\
&= \mathbb{E}\left[-\int_0^t \dot{Y}_s^* M_s^* \mathrm{d}s\right].
\end{aligned}$$

Since all of the terms are non-negative, there are no technical issues involving integrability. The final line follows by equations (2.2.2) and (2.2.3). □

Because of the nonlinearity, in equation (2.1.1), the Feynman-Kac formula is implicit and so we can not use it directly to describe the behaviour of u. However, we will now introduce the primal representation of u in which we write u in terms of an optimisation problem with the optimal control given by the process Y^*. Choosing other controls will give us explicit upper bounds on u.

Theorem 2.2.3. *Fix (t,x) and let X be as in Definition 2.1.1.*

- *If $k < -1$, then*

  $$u(t,x) = \max_Y \mathbb{E}\left[\int_0^t e^{\int_0^s \phi(X_r)\mathrm{d}r}|\dot{Y}_s|^{1+\frac{1}{k}}\mathrm{d}s + e^{\int_0^t \phi(X_r)dr}Y_t^{1+\frac{1}{k}}u_0(X_t)\right]$$

  where the maximum is over positive, decreasing, adapted and absolutely continuous processes, $(Y_s)_{0\le s\le t}$, with $Y_0 = 1$.

- *If $-1 < k < 0$, then*

  $$u(t,x) = \min_Y \mathbb{E}\left[\int_0^t e^{\int_0^s \phi(X_r)\mathrm{d}r}|\dot{Y}_s|^{1+\frac{1}{k}}\mathrm{d}s + e^{\int_0^t \phi(X_r)dr}Y_t^{1+\frac{1}{k}}u_0(X_t)\right]$$

  where the minimum is over positive, strictly decreasing, adapted and absolutely continuous processes, $(Y_s)_{0\le s\le t}$, with $Y_0 = 1$.

- *If $k > 0$, then*

  $$u(t,x) = \min_Y \mathbb{E}\left[\int_0^t e^{\int_0^s \phi(X_r)\mathrm{d}r}|\dot{Y}_s|^{1+\frac{1}{k}}\mathrm{d}s + e^{\int_0^t \phi(X_r)dr}|Y_t|^{1+\frac{1}{k}}u_0(X_t)\right]$$

  where the minimum is over adapted and absolutely continuous processes, $(Y_s)_{0\le s\le t}$, with $Y_0 = 1$.

In each case, the unique optimiser is given by $Y = Y^*$.

Consequently, we can write that

$$u(t,x) = \frac{1}{k+1} \min_Y (k+1)\mathbb{E}\left[\int_0^t e^{\int_0^s \phi(X_r)\mathrm{d}r}(-\dot{Y}_s)^{1+1/k}\,\mathrm{d}s + e^{\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t) Y_t^{1+1/k}\right]$$

where the minimum is taken over positive, decreasing, adapted and absolutely continuous $(Y_s)_{0\le s\le t}$ with $Y_0 = 1$.

We postpone the proof of Theorem 2.2.3 to Section 2.4 where we will exploit a certain duality to be defined in Section 2.3, below. However, another way to view Theorem 2.2.3 is that equation (2.1.1) is essentially the Hamilton–Jacobi–Bellman equation of the stochastic control problem. We will go into more details of this in Appendix B.

2.3. Martingale-Dual Representation

We saw in Section 2.2 that we can write the solution u of equation (2.1.1) in terms of a minimisation/maximisation problem. Therefore, if we try any admissible control, we automatically obtain an upper/lower bound. Naturally, we may then compute the standard Lagrangian dual formulation and try to obtain lower/upper bounds on the solution. This approach runs into some difficulties and requires stronger assumptions on u. In this section, we introduce a new type of dual representation and we will call this the martingale-dual representation. The martingale-dual representation improves upon the Lagrangian representation in the sense that it is more general and there is a correspondence between it and the Lagrangian dual: the martingale-dual problem gives tighter bounds on u for the same control when the control is admissible in both formulations. We discuss the Lagrangian approach in more detail in Section 2.5 and Appendix B.

Fix (t,x) and let the processes X and M^* be as in Definition 2.1.1.

Before giving the dual representation we write u in terms of the optimal control which will be the process M^*.

Lemma 2.3.1. *For fixed (t,x), the following identity holds*

$$u_0(X_t)^k = \left(e^{-\int_0^t \phi(X_r)\mathrm{d}r} M_t^*\right)^k + u_0(X_t)^k \int_0^t \left(e^{-\int_0^s \phi(X_r)\mathrm{d}r} M_s^*\right)^k \mathrm{d}s. \tag{2.3.1}$$

In particular, we can write u as

$$u(t,x) = \mathbb{E}\left[\frac{M_t^* u_0(X_t)}{\left[\left(e^{-\int_0^t \phi(X_r)\mathrm{d}r} M_t^*\right)^k + u_0(X_t)^k \int_0^t \left(e^{-\int_0^s \phi(X_r)\mathrm{d}r} M_s^*\right)^k \mathrm{d}s\right]^{\frac{1}{k}}}\right].$$

Proof. Since M^* is a martingale, we only need to show the identity (2.3.1) and this follows from the definition of M^*. Indeed, note that

$$\left(e^{-\int_0^s \phi(X_r)\mathrm{d}r} M_s^*\right)^k = u(t-s, X_s)^k e^{\int_0^s u(t-r,X_r)^k \mathrm{d}r}$$

Therefore, we have the following two relations between M^* and u

$$\left(e^{-\int_0^t \phi(X_r)\mathrm{d}r} M_t^*\right)^k = u_0(X_t)^k e^{\int_0^t u(t-r,X_r)^k \mathrm{d}r}$$
$$\int_0^t \left(e^{-\int_0^s \phi(X_r)\mathrm{d}r} M_s^*\right)^k \mathrm{d}s = 1 - e^{\int_0^t u(t-r,X_r)^k \mathrm{d}r}$$

□

The following theorem is what we will refer to as the martingale-dual representation.

Theorem 2.3.2. *Fix (t, x) and let u solve equation* (2.1.1).

- *If $k < -1$, then*

$$u(t,x) = \min_M \mathbb{E}\left[\frac{M_t u_0(X_t)}{[(M_t e^{-\int_0^t \phi(X_r)\mathrm{d}r})^k + u_0(X_t)^k \int_0^t (M_s e^{-\int_0^s \phi(X_r)\mathrm{d}r})^k]^{1/k}}\right]$$

  where the minimum is over positive martingales M.

- *If $k > -1$ and $k \neq 0$, then*

$$u(t,x) = \max_M \mathbb{E}\left[\frac{M_t u_0(X_t)}{[(M_t e^{-\int_0^t \phi(X_r)\mathrm{d}r})^k + u_0(X_t)^k \int_0^t (M_s e^{-\int_0^s \phi(X_r)\mathrm{d}r})^k]^{1/k}}\right]$$

  where the maximum is over non-negative martingales M.

Remark 2. *When $k < 0$ we have assumed the u_0 is strictly positive, but when $k > 0$ we allow for $u_0(x)$ to be zero for some $x \in E$. In particular, for $k > 0$ the martingale M_s may vanish for some $0 \leq s \leq t$ with positive probability. We thus make the convention that $0/0 = 0$ so that the expression on the right-hand side is always well-defined.* ◇

It is interesting to note that the optimisation problem in Theorem 2.3.2 - unlike the one in Theorem 2.2.3 - is *not* in a form amenable to dynamic programming and the HJB equation.

2.4. Proof of Representation Theorems

In light of Lemmas 2.2.2 and 2.3.1 we see that Theorems 2.2.3 and 2.3.2 will follow directly from the following lemma.

Lemma 2.4.1. *For any non-negative $\mathcal{F}_t$ measurable H and positive, continuous, adapted process $(h_s)_{0\leq s\leq t}$*

- *for $k < -1$ we have*

$$\mathbb{E}\left[\frac{HM_t}{(M_t^k + H^k \int_0^t (M_s/h_s)^k)^{\frac{1}{k}}}\right] \geq \mathbb{E}\left[\int_0^t h_s |\dot{Y}_s|^{1+\frac{1}{k}} \mathrm{d}s + HY_t^{1+\frac{1}{k}}\right]$$

  for all positive martingales M and positive, decreasing, adapted and absolutely continuous Y with $Y_0 = 1$.

- *for* $-1 < k < 0$ *we have*

$$\mathbb{E}\left[\frac{HM_t}{(M_t^k + H^k\int_0^t (M_s/h_s)^k)^{\frac{1}{k}}}\right] \leq \mathbb{E}\left[\int_0^t h_s|\dot{Y}_s|^{1+\frac{1}{k}}\mathrm{d}s + HY_t^{1+\frac{1}{k}}\right]$$

for all positive martingales M *and positive, strictly decreasing, adapted and absolutely continuous* Y *with* $Y_0 = 1$.

- *for* $k > 0$ *we have*

$$\mathbb{E}\left[\frac{HM_t}{(M_t^k + H^k\int_0^t (M_s/h_s)^k)^{\frac{1}{k}}}\right] \leq \mathbb{E}\left[\int_0^t h_s|\dot{Y}_s|^{1+\frac{1}{k}}\mathrm{d}s + H|Y_t|^{1+\frac{1}{k}}\right]$$

for all non-negative martingales M *and adapted and absolutely continuous* Y *with* $Y_0 = 1$.

In each case, there is equality if and only if there is a constant $c > 0$ *such that*

$$-\dot{Y}_s = c(M_s/h_s)^k. \tag{2.4.1}$$

and

$$M_t^k + H^k\int_0^t (M_s/h_s)^k = H^k/c \tag{2.4.2}$$

The proof of Lemma 2.4.1 and thus Theorems 2.2.3 and 2.3.2 is based on two further lemmas; the first lemma we use is the following.

Lemma 2.4.2. *Fix* $k \notin \{0, -1\}$ *and measurable functions* f, g *on the interval* $[0, t]$, *where* g *is positive and the functions* f, $|f|^{1+\frac{1}{k}}g$ *and* g^{-k} *are integrable.*

- *If* $k < -1$ *we have*

$$\int_0^t f(s)^{1+\frac{1}{k}}g(s)\mathrm{d}s \leq \frac{\left(\int_0^t f(s)\mathrm{d}s\right)^{1+\frac{1}{k}}}{\left(\int_0^t g(s)^{-k}\mathrm{d}s\right)^{\frac{1}{k}}}$$

for positive f.

- *If* $-1 < k < 0$ *we have*

$$\int_0^t f(s)^{1+\frac{1}{k}}g(s)\mathrm{d}s \geq \frac{\left(\int_0^t f(s)\mathrm{d}s\right)^{1+\frac{1}{k}}}{\left(\int_0^t g(s)^{-k}\mathrm{d}s\right)^{\frac{1}{k}}}$$

for strictly positive f.

- *If* $k > 0$ *we have*

$$\int_0^t |f(s)|^{1+\frac{1}{k}}g(s)\mathrm{d}s \geq \frac{\left|\int_0^t f(s)\mathrm{d}s\right|^{1+\frac{1}{k}}}{\left(\int_0^t g(s)^{-k}\mathrm{d}s\right)^{\frac{1}{k}}}$$

In all cases, there is equality if and only if there is a constant C *such that*

$$f(s) = C\ g(s)^{-k} \text{ for almost every } 0 \leq s \leq t.$$

Proof. The proof of this result follows from applying Hölder's inequality. We write p and q for the Hölder coefficients and so $1/p + 1/q = 1$ with $p, q > 1$.

Some care needs to be taken with each range of k. In particular, note that if $k < -1$, we have $0 < 1 + 1/k < 1$; if $-1 < k < 0$, then $1 + 1/k < 0$ and if $k > 0$, then $1 + 1/k > 1$.

If $k < -1$, we can use Hölder's inequality with $p = \frac{k}{k+1}$ and $q = -k$:

$$\int_0^t f(s)^{\frac{k+1}{k}} g(s) \mathrm{d}s \leq \left(\int_0^t f(s) \mathrm{d}s \right)^{1+\frac{1}{k}} \left(\int_0^t g(s)^{-k} \mathrm{d}s \right)^{-\frac{1}{k}}$$

If $-1 < k < 0$, we can use Hölder's inequality with $p = -1/k$ and $q = \frac{1}{k+1}$ and then,

$$\begin{aligned} \int_0^t g(s) \mathrm{d}s &= \int_0^t g(s)^{-k} f(s)^{-(1+k)} f(s)^{1+k} \mathrm{d}s \\ &\leq \left(\int_0^t g(s) f(s)^{\frac{k+1}{k}} \right)^{-k} \left(\int_0^t f(s) \mathrm{d}s \right)^{k+1} \end{aligned}$$

If $k > 0$, we can use Hölder's inequality with $p = 1 + 1/k$ and $q = k + 1$ and so,

$$\begin{aligned} \left| \int_0^t f(s) \mathrm{d}s \right| &\leq \int_0^t |f(s)| g(s)^{\frac{k}{k+1}} g(s)^{-\frac{k}{k+1}} \mathrm{d}s \\ &\leq \left(\int_0^t |f(s)|^{\frac{k+1}{k}} g(s) \mathrm{d}s \right)^{\frac{k}{k+1}} \left(\int_0^t g(s)^{-k} \mathrm{d}s \right)^{\frac{1}{k+1}} \end{aligned}$$

In each case we can easily check that $p > 1$ and so also $q > 1$.

For each case we verify that we have equality if and only if $f(s) = Cg(s)^{-k}$. □

The second lemma we need is the following.

Lemma 2.4.3.
- *For $k < -1$, $\zeta, \eta > 0$ and $0 < y < 1$ we have*

$$\frac{(1-y)^{1+\frac{1}{k}}}{\zeta^{\frac{1}{k}}} + \frac{y^{1+\frac{1}{k}}}{\eta^{\frac{1}{k}}} \leq \frac{1}{(\zeta+\eta)^{\frac{1}{k}}}$$

- *For $-1 < k < 0$, $\zeta, \eta > 0$ and $0 < y < 1$ we have*

$$\frac{(1-y)^{1+\frac{1}{k}}}{\zeta^{\frac{1}{k}}} + \frac{y^{1+\frac{1}{k}}}{\eta^{\frac{1}{k}}} \geq \frac{1}{(\zeta+\eta)^{\frac{1}{k}}}$$

- *For $k > 0$, $\zeta, \eta > 0$ and $y \in \mathbb{R}$ we have*

$$\frac{|1-y|^{1+\frac{1}{k}}}{\zeta^{\frac{1}{k}}} + \frac{|y|^{1+\frac{1}{k}}}{\eta^{\frac{1}{k}}} \geq \frac{1}{(\zeta+\eta)^{\frac{1}{k}}}.$$

In all cases there is equality if and only if $y = \frac{\eta}{\zeta+\eta}$.

Proof. Viewed as a function of y, the left-hand side is strictly concave for $k < -1$ since the sum of strictly concave functions is again strictly concave. Similarly, the left hand side is strictly convex for $k > -1$.

Differentiating, we see that there is always a stationary point for $y \in (0, 1)$ and in particular at the point $y^* = \frac{\eta}{\zeta+\eta}$. Since the left hand side is strictly concave/convex in the relevant ranges of y, the stationary point y^* maximises/minimises the left hand side with the value $(\zeta + \eta)^{-\frac{1}{k}}$. □

We are now ready for the proof of Lemma 2.4.1 and the representations of u.

Proof of Lemma 2.4.1. Let M be a non-negative martingale, and if $k < 0$ assume that M is positive. Also let Y be adapted and absolutely continuous, and if $k < 0$ assume that Y positive and decreasing (strictly so, in the case when $k \in (-1, 0)$). Let $Z = -\dot{Y}$.

To keep the direction of the inequality the same in the following calculation, we multiply by $k+1$. Applying tower property of conditional expectation, we have

$$\begin{aligned}
(k+1)\mathbb{E}&\left[\int_0^t h_s |Z_s|^{1+\frac{1}{k}} \mathrm{d}s + H|Y_t|^{1+\frac{1}{k}}\right] \\
&\geq (k+1)\mathbb{E}\left[\int_0^t h_s |Z_s|^{1+\frac{1}{k}} \frac{M_t}{M_s} \mathbb{1}_{\{M_s>0\}} \mathrm{d}s + H|Y_t|^{1+\frac{1}{k}}\right] \\
&\geq (k+1)\mathbb{E}\left[\frac{|1-Y_t|^{1+\frac{1}{k}}}{\left(\int_0^t \left(\frac{M_s}{h_s M_t}\right)^k \mathrm{d}s\right)^{\frac{1}{k}}} + H|Y_t|^{1+\frac{1}{k}}\right] \\
&\geq (k+1)\mathbb{E}\left[\frac{H}{\left(1 + H^k \int_0^t \left(\frac{M_s}{h_s M_t}\right)^k \mathrm{d}s\right)^{\frac{1}{k}}}\right]
\end{aligned}$$

using Lemma 2.4.2 to pass from the first to second line with

$$f(s) = Z_s \text{ and } g(s) = \frac{h_s M_t}{M_s}$$

and using Lemma 2.4.3 to pass from the second to third line with

$$\eta = H^{-k} \text{ and } \zeta = \int_0^t \left(\frac{M_s}{h_s M_t}\right)^k \mathrm{d}s.$$

For the case of equality, note that by Lemma 2.4.2 there is equality only if there exists a random variable C such that

$$Z_s = C\left(\frac{M_s}{h_s M_t}\right)^k.$$

Since,

$$\frac{C}{M_t^k} = -\frac{\dot{Y}_s h_s^k}{M_s^k},$$

and the right hand side is adapted to $\mathcal{F}_s$, we must have that there is a constant c such that $C = cM_t^k$. In other words,

$$1 - Y_s = c\int_0^s \left(\frac{M_r}{h_r}\right)^k \mathrm{d}r$$

Now, by Lemma 2.4.3 there is equality only if

$$Y_t = \frac{M_t^k}{M_t^k + H^k \int_0^t (M_s/h_s)^k \mathrm{d}s}.$$

By combining these two conditions, we have the result. ∎

Remark 3. *Note that if we set*

$$h_s = e^{\int_0^s \phi(X_r)\mathrm{d}r}$$

and

$$H = e^{\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t)$$

we arrive at the formulation in Theorems 2.2.3 and 2.3.2. In this case, the equality conditions (2.4.2) *and* (2.4.1) *with* $c = 1$, *are satisfied at* Y^* *and* M^* *(by Lemmas 2.2.2 and 2.3.1). In particular,*

$$\mathbb{E}\left[\frac{HM_t^*}{((M_t^*)^k + H^k \int_0^t (M_s^*/h_s)^k)^{\frac{1}{k}}}\right] = u(t,x) = \mathbb{E}\left[\int_0^t h_s|\dot{Y}_s^*|^{1+\frac{1}{k}}\mathrm{d}s + H|Y_t^*|^{1+\frac{1}{k}}\right].$$

◇

Remark 4. *Note that we have used no specific assumptions on the Markov process* X *for in this proof. Indeed, the only role that* X *plays in the above calculation is through the formulae* $h_s = e^{\int_0^s \phi(X_r)\mathrm{d}r}$ *and* $H = e^{\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t)$. *We saw, however, that the above proof works for any positive process* $(h_s)_{0\le s\le t}$ *and* $H > 0$. *In particular, the Markov property is not needed at all.* ◇

2.5. Lagrangian Duality

Note that the primal representation given in Section 2.2 expresses the quantity $u(t,x)$ as the value of a convex minimisation problem. In particular, it is natural to apply standard Lagrangian methods to find the dual problem. We will see that the dual variable is again a martingale M. However, it turns out that our new dual problem given in Section 2.3, which exploits the special structure of the problem, is stronger in a sense to made precise.

We do the formal calculation to identify the standard Lagrangian dual problem for the primal problem of minimising

$$\mathbb{E}\left[\int_0^t h_s|Z_s|^{1+\frac{1}{k}}\mathrm{d}s + HY_t^{1+\frac{1}{k}}\right]$$

over positive, adapted, absolutely continuous processes Y and negative adapted, measurable process Z such that $Z = -\dot{Y}$ and $Y_0 = 1$, and where $(h_s)_{0\le s\le t}$ is a given positive adapted measurable process and H is a given positive random variable.

Consider the Lagrangian

$$L(Y,Z;\lambda,M) = \mathbb{E}\left[\int_0^t h_s|Z_s|^{1+\frac{1}{k}}\mathrm{d}s + HY_t^{1+\frac{1}{k}} - (1+\tfrac{1}{k})\int_0^t M_s(\dot{Y}_s + Z_s)\mathrm{d}s\right] + \lambda(1-Y_0)$$

where the Lagrange multiplier M is a martingale and the multiplier λ is a real number. We rewrite the Lagrangian as

$$\begin{aligned} L(Y,Z;\lambda,M) =& \mathbb{E}\left[\int_0^t [h_s|Z_s|^{1+\frac{1}{k}}\mathrm{d}s - (1+\tfrac{1}{k})Z_sM_s]\mathrm{d}s + HY_t^{1+\frac{1}{k}} - (1+\tfrac{1}{k})M_tY_t\right] \\ &+ \lambda + [(1+\tfrac{1}{k})M_0 - \lambda]Y_0 + (1+\tfrac{1}{k})\mathbb{E}\left[\int_0^t Y_s\mathrm{d}M_s\right] \end{aligned}$$

We henceforth assume that the expected value of the stochastic integral on the second line vanishes. Assuming M is positive and $\lambda = (1 + \frac{1}{k})M_0$ we see that the minimum of $L(Y, Z; \lambda, M)$ occurs when $Y_t = (M_t/H)^k$ and $Z_s = (M_s/h_s)^k$.

Therefore, the standard Lagrangian dual problem is to maximise

$$(1 + \tfrac{1}{k})M_0 - \frac{1}{k}\mathbb{E}\left[\int_0^t h_s^{-k} M_s^{k+1} \mathrm{d}s + H^{-k} M_t^{k+1}\right]$$

over positive martingales M. By pulling out the initial condition M_0 and maximising over this, we see that, heuristically, the dual problem is to consider

$$\max_M \mathbb{E}\left[\int_0^t h_s^{-k} M_s^{k+1} \mathrm{d}s + H^{-k} M_t^{k+1}\right]^{-1/k} \tag{2.5.1}$$

over positive martingales M with $M_0 = 1$.

In the case when h and H correspond to the representations in Theorem 2.2.3 and given enough regularity on u, we see that this is the correct formulation of the Lagrangian dual approach and the cases when $k < 0$ are analogous.

We note that the above standard Lagrangian dual problem has at least one advantage over the dual problem presented in Section 2.3 in the case of interest where $h_s = e^{\int_0^s \phi(X_r)\mathrm{d}r}$ and $H = e^{\int_0^t \phi(X_r)\mathrm{d}r} u(X_t)$ and X is a Markov process. This advantage is that objective is in the standard form for stochastic control and hence can be studied via the dynamic programming principle and the HJB equation.

However, there is an advantage to the new dual problem of Section 2.3. Notice that

$$\mathbb{E}\left[\frac{HM_t}{\left(M_t^k + H^k \int_0^t (M_s^k/h_s^k)\mathrm{d}s\right)^{1/k}}\right] = \mathbb{E}\left[M_t^{1+1/k} \frac{H}{\left(M_t^{k+1} + H^k M_t \int_0^t (M_s^k/h_s^k)\mathrm{d}s\right)^{1/k}}\right] \tag{2.5.2}$$

$$\geq \mathbb{E}\left[H^{-k}\left(H^k \int_0^t (M_t M_s^k/h_s^k)\mathrm{d}s + M_t^{k+1}\right)\right]^{-1/k} \tag{2.5.3}$$

$$= \mathbb{E}\left[\int_0^t h_s^{-k} M_s^{k+1} \mathrm{d}s + H^{-k} M_t^{k+1}\right]^{-1/k} \tag{2.5.4}$$

where we used Hölder's inequality as in Lemma 2.4.2 (adapted for expectations) and the assumption that the martingale M starts at $M_0 = 1$ to go from the first to the second line. To go from the first to the second line we use the tower property of condition expectation. If $k < -1$, the inequality is reversed.

We have equality when

$$M_t = cH^{-k} M_t \left(M_t^k + H^k \int_0^t (M_s^k/h_s^k)\mathrm{d}s\right),$$

for some constant c. We can see that M^* satisfies this as long as it the right hand side of (2.5.2) is well defined.

Recognising the expression on the right-hand side as the objective function of the martingale dual problem introduced in Section 2.3, we have shown that the new duality is stronger than the standard Lagrangian duality in the following sense: for the same control M the new objective function is closer to the common optimal value than the standard objective function. Indeed, note that if $H = 0$ with positive probability, the standard objective function is finite only for those martingales such that $\{H = 0\} \subseteq \{M_t = 0\}$. We can see that the Lagrangian dual is also more restrictive in terms of necessary conditions for it to hold.

More details on this approach and a direct proof in a simplified setting can be found in Appendix B.

2.6. Simple Consequences

2.6.1. Corollaries of the Representations

Here, we list some simple consequences of Theorems 2.2.3 and 2.3.2. This result corresponds to restricting the set of admissible processes in the representations of Theorems 2.2.3 and 2.3.2 to deterministic controls. The bounds will be improved upon in Chapter 4.

Corollary 2.6.1. *Fix (t, x) and let X be as in Definition 2.1.1. Then,*

- *for $k < -1$, we have*

$$u(t,x) \geq \left(\mathbb{E}\left[e^{\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t) \right]^{-k} + \int_0^t \mathbb{E}\left[e^{\int_0^s \phi(X_r)\mathrm{d}r} \right]^{-k} \mathrm{d}s \right)^{-1/k}.$$

- *For $k > -1$, $k \neq 0$ we have*

$$u(t,x) \leq \left(\mathbb{E}\left[e^{\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t) \right]^{-k} + \int_0^t \mathbb{E}\left[e^{\int_0^s \phi(X_r)\mathrm{d}r} \right]^{-k} \mathrm{d}s \right)^{-1/k}.$$

On the other hand,

- *For $k < -1$ we have*

$$u(t,x) \leq \mathbb{E}\left[\left(e^{-k\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t)^{-k} + \int_0^t e^{-k\int_0^s \phi(X_r)\mathrm{d}r} \mathrm{d}s \right)^{-1/k} \right].$$

- *For $k > -1$, $k \neq 0$ we have*

$$u(t,x) \geq \mathbb{E}\left[\left(e^{-k\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t)^{-k} + \int_0^t e^{-k\int_0^s \phi(X_r)\mathrm{d}r} \mathrm{d}s \right)^{-1/k} \right].$$

Proof. Let

$$-\dot{Y}_s = \frac{\mathbb{E}\left[e^{\int_0^s \phi(X_r)\mathrm{d}r} \right]^{-k}}{\mathbb{E}\left[e^{\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t) \right]^{-k} + \int_0^t \mathbb{E}[e^{\int_0^r \phi(X_\nu)\mathrm{d}\nu}]^{-k}\mathrm{d}r}.$$

in Theorem 2.2.3. This gives,

$$Y_s = \frac{\mathbb{E}\left[e^{\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t) \right]^{-k} + \int_s^t \mathbb{E}\left[e^{\int_0^r \phi(X_\nu)\mathrm{d}\nu} \right]^{-k} \mathrm{d}r}{\mathbb{E}\left[e^{\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t) \right]^{-k} + \int_0^t \mathbb{E}\left[e^{\int_0^r \phi(X_\nu)\mathrm{d}\nu} \right]^{-k} \mathrm{d}r}.$$

Note that Y is positive, decreasing with $Y_0 = 1$.

For the second part, we plug in the control $M = 1$ in Theorem 2.3.2. $\square$

Note that there is an interesting relationship between these upper and lower bounds. For example, in the case where $k > 0$, the bounds are of the form

$$\mathbb{E}\left[\frac{H}{(1 + H^k \int_0^t h_s^{-k} \mathrm{d}s)^{1/k}}\right] \leq u(t,x) \leq \frac{\mathbb{E}[H]}{(1 + \mathbb{E}[H]^k \int_0^t \mathbb{E}[h_s]^{-k} \mathrm{d}s)^{1/k}}.$$

We will observe this kind of behaviour again in Chapter 4.

As noted above, the choice of Y in the proof of Corollary 2.6.1 is deterministic. In fact, it achieves the optimum value of the right-hand side over all deterministic controls.

Remark 5. *Fix a constant $H > 0$ and a positive, deterministic function h on $[0,t]$. Consider the problem of minimising*

$$\int_0^t h(s)|\dot{y}(s)|^{1+\frac{1}{k}} \mathrm{d}s + H|y(t)|^{1+\frac{1}{k}}$$

over absolutely continuous functions y with $y(0) = 1$. We can appeal to the above duality proof. Since nothing is random, the only martingales are constant. Since scaling is irrelevant we may let $M_s = 1$ for all $0 \leq s \leq t$. The dual optimality condition

$$H^k/c = M_t^k + H^k \int_0^t [M_s/h(s)]^k \mathrm{d}s$$

yields

$$c = \frac{1}{H^{-k} + \int_0^t h(s)^{-k} \mathrm{d}s}.$$

The primal optimality condition is then

$$-\dot{y}(s) = c[M_s/h(s)]^k = \frac{h(s)^{-k}}{H^{-k} + \int_0^t h(r)^{-k} \mathrm{d}r}.$$

By choosing

$$H = \mathbb{E}\left[e^{\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t)\right] \text{ and } h(s) = \mathbb{E}\left[e^{\int_0^s \phi(X_r)\mathrm{d}r}\right]$$

we see that the control given the proof of Corollary 2.6.1 is optimal over deterministic controls as claimed. ◇

Remark 6. *Recall that we have the trivial bound $u(t,x) \leq e^t \mathbb{E}^x[u_0(X_t)]$ from the Feynman-Kac representation. The upper bound in Corollary 2.6.1 gives a slight improvement on this.* ◇

2.6.2. Choosing a control, M

We can use the representation of Theorem 2.3.2 directly to obtain a lower bound on u. One simple example is given below as a demonstration but the best bounds presented in this work will follow from iterating the bounds in Section 2.6.1 and are presented in Chapter 4.

Proposition 2.6.2. *Let u be a solution to the equation*

$$\frac{\partial u}{\partial t} = \mathcal{L}u + u - u^2$$
$$u_0(x) = \mathbb{1}_{(-\infty,0)}(x)$$

Let M be a continuous non-negative martingale with $M_t = 0 \iff u_0(X_t) = 0$. Then,

$$u(t,x) \geq \frac{M_0^2}{M_0^2 + \mathbb{E}\left[\int_0^t e^{-s} \mathrm{d}\langle M \rangle_s\right]}. \tag{2.6.1}$$

Proof. By Theorem 2.3.2, we have

$$u(t,x) \geq \mathbb{E}\left[\frac{M_t u_0(X_t)}{Z_t u_0(X_t) + (1-u_0(X_t))e^{-t} M_t}\right]$$

where

$$Z_s = e^{-s} M_s + \int_0^s e^{-r} M_r \mathrm{d}r.$$

Since, $u_0(x) = \mathbb{1}_{(-\infty,0)}(x)$, and $M_t \mathbb{1}_{(-\infty,0)}(X_t) = M_t$, by assumption, we see that

$$u(t,x) \geq \mathbb{E}\left[\frac{M_t}{Z_t}\right]$$

with equality for M^*. By the Cauchy-Schwarz inequality:

$$\begin{aligned} u(t,x) &\geq \frac{\mathbb{E}\left[M_t\right]^2}{\mathbb{E}[M_t Z_t]} \\ &= \frac{M_0^2}{M_0^2 + \mathbb{E}\left[\int_0^t e^{-s} \mathrm{d}\langle M \rangle_s\right]} \end{aligned}$$

since $Z_0 = M_0$ and

$$\mathrm{d}Z_s = e^{-s} \mathrm{d}M_s.$$

□

In order to obtain an explicit lower bound, we can choose a particular martingale, M. We want to choose a martingale to approximate

$$M_s^* = u(t-s, X_s) e^{\int_0^s [1-u(t-r,X_r)]\mathrm{d}r}.$$

Let's consider the case when X is a Brownian motion.

In order to approximate M^*, we will need to approximate u. It is well known that the position of the front is $\sqrt{2}t + o(t)$ (and we show this in Chapter 4). This suggests that we could use an approximation of u such as $u(t,x) \approx u_0(x - \sqrt{2}t)$. For simplicity, we work with the martingale given by

$$M_s = \mathbb{E}\left[\mathbb{1}(x + W_t < 0) | \mathcal{F}_s\right].$$

We can rewrite this as

$$M_s = \int_{\frac{x+W_s}{\sqrt{t-s}}}^{\infty} \frac{e^{-y^2/2}}{\sqrt{2\pi}} \mathrm{d}y$$

and so

$$\mathrm{d}M_s = -\frac{e^{-\frac{(x+W_s)^2}{2(t-s)}}}{\sqrt{2\pi(t-s)}} \mathrm{d}W_s$$

and

$$\mathrm{d}\langle M \rangle_s = \frac{e^{-\frac{(x+W_s)^2}{(t-s)}}}{2\pi(t-s)} \mathrm{d}s.$$

Now we see that

$$\frac{1}{u(t,x)} \leq 1 + \frac{1}{M_0^2}\int_0^t \frac{e^{-s-\frac{x^2}{s+t}}}{2\pi\sqrt{t^2-s^2}}\mathrm{d}s \tag{2.6.2}$$

For $x \leq c\sqrt{t}$ for a constant c, we have that

$$\frac{1}{M_0^2}\int_0^t \frac{e^{-s-\frac{x^2}{s+t}}}{2\pi\sqrt{t^2-s^2}}\mathrm{d}s \to 0$$

as $t \to \infty$. This follows from estimating M_0 with the standard Gaussian tail bound, $x^2/(s+t)$ by $x^2/2t$, and noting that

$$\int_0^t \frac{e^{-s}}{\sqrt{t^2-s^2}}\mathrm{d}s \to 0.$$

This gives us that $u(t,x) \to 1$ for $x \leq c\sqrt{t}$ for any c. This is a slight improvement on the result that follows from the trivial bound

$$u(t,x) \geq \Phi\left(-\frac{x}{\sqrt{t}}\right)$$

obtained by setting $M = 1$ in the representation of Theorem 2.3.2, where Φ is the CDF of the standard normal distribution. In that case we see that $\lim_{t\to\infty} u(t,x) \geq 1/2$ when $x = o(\sqrt{t})$.

Neither of these bounds are very good. However, in Chapter 4, we iterate the trivial bound here to obtain the speed of the wave front up to the linear term. We can not directly iterate the bound in (2.6.2) since it only holds when u_0 is an indicator function but the bound could be improved with a better choice of control, M. One possible choice of control would be

$$M_s = \mathbb{E}\left[\mathbb{1}(x+W_t<0)e^{\int_0^t \mathbb{1}(x+W_r-\sqrt{2}(t-r)>0)}\middle|\mathcal{F}_s\right]$$

but this requires more work.

Remark 7. *By using the Lagrangian framework, one can obtain some information about a suitable set of admissible controls to use when bounding a solution u using Proposition 2.6.2. We will see that it is essential to have a condition such as $\{M_s = 0\} = \{X_s > 0\}$ in this way. If $u_0 = \mathbb{1}_{(-\infty,0)}$, then, for $t > 0$*

$$\frac{1}{u(t,x)} \leq \min_{M\in\mathcal{M}_1} \mathbb{E}\left[\int_0^t e^{-s}M_s^2\mathrm{d}s + e^{-t}M_t^2\right] \tag{2.6.3}$$

where $\mathcal{M}_1$ is the set containing all non-negative martingales M, with $M_0 = 1$, such that $\langle M\rangle$ is differentiable and $M_s = 0 \iff u_0(X_s) = 0 \iff X_s > 0$. We can see this as follows: for any $M \in \mathcal{M}_1$,

$$\mathbb{E}[M_s^2] = M_0^2 + \mathbb{E}\left[\int_0^s \frac{\mathrm{d}\langle M\rangle_s}{\mathrm{d}s}\right] \tag{2.6.4}$$

and so

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{E}[M_s^2] = \mathbb{E}\left[\frac{\mathrm{d}\langle M\rangle_s}{\mathrm{d}s}\right].$$

Expanding

$$\int_0^t e^{-s}\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{E}[M_s^2]\mathrm{d}s$$

using integration by parts, we see that

$$M_0^2 + \mathbb{E}\left[\int_0^t e^{-s}\mathrm{d}\langle M\rangle_s\right] = \mathbb{E}\left[e^{-t}M_t^2 + \int_0^t e^{-s}M_s^2\mathrm{d}s\right]$$

and thus, by the bound (2.6.1)*,*

$$u(t,x) \geq \frac{M_0^2}{\mathbb{E}\left[e^{-t}M_t^2 + \int_0^t e^{-s}M_s^2\mathrm{d}s\right]}$$

for all $M \in \mathcal{M}_1$ *and the bound* (2.6.3) *follows.*

Compare this to the Lagrangian representation, (2.5.1)*. We see that*

$$\frac{1}{v(t,x)} \equiv 1 = \min_{M\in\mathcal{M}_2} \mathbb{E}\left[\int_0^t e^{-s}M_s^2\mathrm{d}s + e^{-t}M_t^2\right]$$

where $\mathcal{M}_2$ *is defined in the same way as* $\mathcal{M}_1$ *without the condition that* $M_s = 0 \iff X_s > 0$*. We have, above, that* $v \equiv 1$*, since* 1 *is the unique solution to the equation*

$$\frac{\partial v}{\partial t} = \mathcal{L}v + v - v^2$$

with $v_0 \equiv 1$*.*

Take X *to be a standard Brownian Motion. Since, when* $u_0 = \mathbb{1}_{(-\infty,0)}$*, we have* $u(t,x) \leq 1$ *for* $t > 0$ *and so we see that the additional condition* $\{M_s = 0\} = \{X_s > 0\}$ *in the bound* (2.6.3) *is an important and natural restriction and when bounding* u *using this method one should choose controls,* M*, with this feature.* $\diamond$

2.7. Existence of Solutions

So far we have been working under the assumption that solutions to equation (2.1.1) exist. Here, we give some sufficient conditions for existence and uniqueness. Since such questions are well studied for mild solutions, we will first give results for existence and uniqueness of mild solutions and show that, in the cases below, this is an equivalent notion to solutions defined in the sense of Definition 2.1.1.

Throughout this section we will use the following notation. Let $f : [0,\infty) \times E \times \mathbb{R} \to \mathbb{R}$ and $u_0 : E \to \mathbb{R}$ be measurable. We let X be a time-homogeneous Markov process with generator $\mathcal{L}$.

Note that sometimes f will not be time-dependent and we will consider the special case $f : E \times \mathbb{R} \to \mathbb{R}$.

Definition 2.7.1. *We say that a measurable function,* $u : [0,\infty) \times E \to \mathbb{R}$*, is a mild solution to the equation*

$$\begin{aligned} u_t &= \mathcal{L}u + f(t,x,u) \\ u(0,\cdot) &= u_0 \end{aligned} \tag{2.7.1}$$

if for all $(t,x) \in [0,\infty) \times E$*,*

$$u(t,x) = \mathbb{E}^x\left[u_0(X_s) + \int_0^t f(t-r, X_r, u(t-r,X_r))\mathrm{d}r\right].$$

The following result is based on a standard Picard iteration proof. For further details, one can see, for example, Cabré and Roquejoffre [CR13].

Theorem 2.7.2. *Suppose that $f(t,x,\cdot)$ is Lipschitz uniformly in t and x and that $u_0 : E \to \mathbb{R}$ be bounded. Fix some $T > 0$. Then, the equation*

$$\begin{aligned} u_t &= \mathcal{L}u + f(t,x,u) \\ u(0,\cdot) &= u_0 \end{aligned} \tag{2.7.2}$$

has a unique mild solution $u : [0,\infty) \times E \to \mathbb{R}$ and u is bounded on $[0,T] \times E$.

Furthermore, let $u_0^i : E \to \mathbb{R}$, $i = 1,2$ be bounded with

$$u_0^1(x) \le u_0^2(x)$$

for all $x \in E$ and let u^i, $i = 1,2$ be the corresponding solutions of equation (2.7.2) corresponding to initial conditions u_0^i. Then,

$$u^1(t,x) \le u^2(t,x),$$

for all $(t,x) \in [0,\infty) \times E$.

Note that for KPP-type nonlinearities, this result does not apply directly. However, because of the comparison principle, one can extend the result to these nonlinearities by showing that the solution takes values in a compact set for all time and then assuming, without loss of generality, that f is globally Lipschitz outside of this set. This is the idea for the next theorem:

Theorem 2.7.3. *Suppose that $f : E \times \mathbb{R} \to \mathbb{R}$ is Lipschitz on $[0,c]$ for $c > 0$ with $f(x,0) = 0$ and that $f(x,c) \le 0$ and that $0 \le u_0(x) \le c$ for all $x \in E$.*

Then, the equation

$$\begin{aligned} u_t &= \mathcal{L}u + f(x,u) \\ u(0,\cdot) &= u_0 \end{aligned} \tag{2.7.3}$$

has a unique mild solution u with $0 \le u(t,x) \le c$ for all $(t,x) \in [0,\infty) \times E$.

Proof. For all $x \in E$, we define a new function $\tilde{f}(x,\cdot)$ that is equal to $f(x,\cdot)$ on the interval $[0,c]$ and is *globally* Lipschitz. By Theorem 2.7.2, there is a unique mild solution $\tilde{u}$ to

$$\begin{aligned} \tilde{u}_t &= \mathcal{L}\tilde{u} + \tilde{f}(x,\tilde{u}) \\ \tilde{u}(0,\cdot) &= u_0. \end{aligned} \tag{2.7.4}$$

Let $u_0^1 \equiv 0$ and $u_0^2 \equiv c$ and let u^1 and u^2 be the corresponding solution to (2.7.4) with u_0^1 and u_0^2 as initial conditions.

Since $\tilde{f}(\cdot,0) = 0$, we have $u^1(t,x) = 0$ for all t and x. Similarly, we see that u^2 is equal to the solution to the ODE

$$\begin{aligned} \frac{\mathrm{d}v}{\mathrm{d}t} &= \tilde{f}(x,v) \\ v(0) &= c. \end{aligned}$$

Since $\tilde{f}(\cdot, c) \leq 0$, it follows that $u^2(t,x) \leq c$.

Since $0 = u_0^1 \leq u_0 \leq u_0^2 = c$, it follows, by Theorem 2.7.2, that $0 \leq \tilde{u}(t,x) \leq c$ for all t and x.

Finally, since f and $\tilde{f}$ agree on $[0, c]$, we can set $u = \tilde{u}$ and then u is the required mild solution to equation (2.7.3) and the result follows. □

2.7.1. Solutions when $k > 0$

In this section, we consider the non-linearity $f(x,u) = u(\phi(x) - \frac{1}{k}u^k)$ with $k < 0$.

The following result is a corollary of Theorem 2.7.3:

Theorem 2.7.4. *Let $k > 0$ and suppose that $\phi : E \to \mathbb{R}$ with $\phi(x) \leq c^k/k$ for $c > 0$. Suppose $0 \leq u_0(x) \leq c$ for all $x \in E$. Then, the equation*

$$\begin{aligned} u_t &= \mathcal{L}u + u(\phi - \tfrac{1}{k}u^k) \\ u(0, \cdot) &= u_0 \end{aligned} \tag{2.7.5}$$

has a unique mild solution u with $0 \leq u(t,x) \leq c$ for all $(t,x) \in [0,\infty) \times E$.

So far, we have been dealing with mild solutions. Now, we reconcile this with the generalised solution of Definition 2.1.1. Recall, we say that a measurable function $u : [0,\infty) \times E \to [0,\infty)$ is a solution to equation (2.7.5), if, M^* defined by

$$M_s^* = u(t-s, X_s)e^{\int_0^s [\phi(X_r) - \frac{1}{k}u(t-r,X_r)^k]\mathrm{d}r}$$

is a martingale.

Theorem 2.7.5. *Let u be as in Theorem 2.7.4. Then there exists a unique solution to equation* (2.7.5) *in the sense of Definition 2.1.1.*

Proof. By Theorem 2.7.4, u satisfies

$$u(t,x) = \mathbb{E}^x \left[u_0(X_t) + \int_0^t f(X_r, u(t-r, X_r))\mathrm{d}r \right]$$

where $f(x,u) = u(\phi(x) - \frac{1}{k}u^k)$ for $u \geq 0$.

From this we see that

$$Z_s \equiv u(t-s, X_s) + \int_0^s f(X_r, u(t-r, X_r))\mathrm{d}\nu$$

defines a martingale. Indeed, for $0 \leq s \leq t$,

$$Z_s - \mathbb{E}^{X_0=x}[Z_t|\mathcal{F}_s^X] = u(t-s, X_s^x) - \mathbb{E}^{X_0=x}\left[u_0(X_t) + \int_s^t f(X_r, u(t-r, X_r))\mathrm{d}r \middle| X_s \right]$$

where we used the Markov property of X. By the definition of a mild solution we have

$$\begin{aligned} Z_s - \mathbb{E}^x[Z_t|\mathcal{F}_s] &= u(t-s, X_s^x) - \mathbb{E}^y \left[u_0(X_{t-s}) + \int_0^{t-s} f(X_r, u(t-s-r, X_r))\mathrm{d}r \right] \Bigg|_{y=X_s^x} \\ &= 0. \end{aligned}$$

Since Z is a martingale, the process M defined by

$$M_s = u(t,x) + \int_0^s e^{\int_0^r h(X_\nu, u(t-r,X_\nu))\mathrm{d}\nu} \mathrm{d}Z_r$$

is a martingale, since the integrand is bounded and continuous. Here we define $h : E \times \mathbb{R} \to \mathbb{R}$ by $h(x,u) = \frac{1}{k}(\phi(x) - u^k)$. Using integration by parts, we have

$$M_s = u(t-s, X_s) e^{\int_0^s h(X_r, u(t-r,X_r))\mathrm{d}r}.$$

In other words, u is a solution in the sense of Definition 2.1.1. □

2.7.2. Solutions when $k < 0$

Here we consider the non-linearity $f(x,u) = u(\phi(x) - \frac{1}{k}u^k)$ with $k < 0$. In this case $f(x,\cdot)$ is not Lipschitz at $u = 0$. Therefore, we do not expect a unique mild solution for each initial condition.

If the initial condition is bounded away from zero, we can, without loss of generality, modify the nonlinearity around zero to be Lipschitz and then the above discussion and Theorem 2.7.5, in particular, can be directly applied to this case:

Theorem 2.7.6. *Let $k < 0$. Suppose that ϕ is bounded below. Then there exists $\bar{\varepsilon}$ such that for any $0 < \varepsilon \le \bar{\varepsilon}$, if $u_0(x) \ge \varepsilon$ for all $x \in E$, then the equation*

$$\begin{aligned} u_t &= \mathcal{L}u + u(\phi - \tfrac{1}{k}u^k) \\ u(0,\cdot) &= u_0 \end{aligned} \tag{2.7.6}$$

has a unique solution u with $u(t,x) \ge \varepsilon$ for all $(t,x) \in [0,\infty) \times E$ in the mild sense and in the sense of Definition 2.1.1.

The proof mirrors the proofs of Theorems 2.7.3 – 2.7.5.

Proof. Since ϕ is bounded below, we can find $\bar{\varepsilon} > 0$ such that for all $0 < \varepsilon \le \bar{\varepsilon}$ we have $\phi(x) \ge \frac{1}{k}\bar{\varepsilon}^k \ge \frac{1}{k}\varepsilon^k$. Now, we fix one such ε.

Let $f(x,u) = u(\phi(x) - \frac{1}{k}u^k)$ and we can define, for all $x \in E$, $\tilde{f}(x,\cdot)$ so that $\tilde{f}(x,u) = f(x,u)$ for $u \ge \varepsilon$ and $\tilde{f}(x,\cdot)$ is globally Lipschitz. For the modified PDE there is a unique mild solution by Theorem 2.7.2.

Let u^1 be the solution corresponding to initial condition $u_0^1 \equiv \varepsilon$. Then, $u^1(t,x) \ge \varepsilon$, since $\tilde{f}(x,\varepsilon) \ge 0$ by the fact that $\phi \ge \frac{1}{k}\varepsilon^k$. Therefore, $u(t,x) \ge \varepsilon$ by the comparison principle Theorem 2.7.2. Since f and $\tilde{f}$ agree for $u \ge \varepsilon$, the result holds when u is a mild solution.

The fact that u is also a solution in the sense of Definition 2.1.1, follows from the same proof as in Theorem 2.7.5 (since $h(x,\cdot)$ is bounded on $[\varepsilon,\infty)$). □

There are not unique mild solutions when f is not Lipschitz. However, the optimisation problems we considered still make sense and would be interesting to know when solutions in the sense of Definition 2.1.1 exist.

When $k \in (-1, 0)$ we set

$$U_t(f)(x) = \min_Y \mathbb{E}^x \left[\int_0^t e^{\int_0^s \phi(X_r)\mathrm{d}r} |\dot{Y}_s|^{1+1/k} + e^{\int_0^t \phi(X_r)\mathrm{d}r} Y_t^{1+1/k} f(X_t) \right]$$

where the minimum is taken over positive, strictly decreasing, adapted and absolutely continuous $(Y_s)_{0\leq s\leq t}$ with $Y_0 = 1$. Similarly, for $k < -1$, we define U in terms of the corresponding maximisation problem.

By Lemma 2.4.1/Corollary 2.6.1, we see that for $k \in (-1, 0)$:

$$U_t(u_0)(x) \geq \mathbb{E}\left[\left(u_0(X_t)^{-k} e^{-k\int_0^t \phi(X_r)\mathrm{d}r} + \int_0^t e^{-k\int_0^s \phi(X_r)\mathrm{d}r} \right)^{-1/k} \right], \tag{2.7.7}$$

and for $k < -1$:

$$U_t(u_0)(x) \geq \left(\mathbb{E}\left[e^{\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t) \right]^{-k} + \int_0^t \mathbb{E}\left[e^{\int_0^s \phi(X_r)\mathrm{d}r} \right]^{-k} \mathrm{d}s \right)^{-1/k} \tag{2.7.8}$$

Note that this does not require U to be a solution in any sense.

Suppose that $0 < \varepsilon < \overline{\varepsilon}$ as in Theorem 2.7.6. From the bounds above, we see that for any such ε, there exists $\delta > 0$ such that $U_t(u_0)(x) \geq \varepsilon$ for any $t > \delta$ and that $\varepsilon \to 0$ if and only if $\delta \to 0$. Therefore, by Theorem 2.7.6, for $t > \delta$, $u(t, x) = U_t(u_0)(x)$ defines a mild solution and a solution in the sense of Definition 2.1.1 to

$$\begin{aligned} u_t &= \mathcal{L}u + f(x, u) \text{ for } t > \delta \\ u(\delta, x) &= U_\delta(u_0)(x) \end{aligned}$$

for any $\delta > 0$.

We would like to take the limit $\delta \searrow 0$.

Set $h(x, u) = \phi - \frac{1}{k}u^k$. Then, for all $t > \delta > 0$,

$$u(t + \delta, x) = \mathbb{E}\left[u(\delta, X_t) e^{\int_0^t h(X_s, u(t+\delta-s, X_s))\mathrm{d}s} \right]$$

and so

$$u(t, x) = \mathbb{E}\left[u(\delta, X_{t-\delta}) e^{\int_0^{t-\delta} h(X_s, u(t-s, X_s))\mathrm{d}s} \right].$$

Let

$$\tilde{u}(t, x) \equiv \mathbb{E}\left[u_0(X_t) e^{\int_0^t h(X_s, u(t-s, X_s))\mathrm{d}s} \right]. \tag{2.7.9}$$

To show existence of a solution in the sense of Definition 2.1.1, one must show, by taking the limit $\delta \searrow 0$, that $u = \tilde{u}$. Then, by the Markov property of X, it would follow that M^* is a martingale and that u is a solution to

$$\begin{aligned} u_t &= \mathcal{L}u + f(x, u) \text{ for } t > 0 \\ u(0, x) &= u_0(x). \end{aligned}$$

If we can show that

$$u(\delta, X_{t-\delta}) e^{\int_0^{t-\delta} h(X_s, u(t-s, X_s))\mathrm{d}s} \xrightarrow{P} u_0(X_t) e^{\int_0^t h(X_s, u(t-s, X_s))\mathrm{d}s} \tag{2.7.10}$$

and that

$$\left(u(s, X_{t-s})e^{\int_0^{t-s} h(X_r, u(t-r, X_r))\mathrm{d}r}\right)_{0\le s\le t} \tag{2.7.11}$$

is uniformly integrable, then we will be done.

In general, showing uniform integrability would require some subtlety to use the trade-off between $u(s, X_{t-s})e^{\int_0^{t-s} h(X_r,u(t-r,X_r))\mathrm{d}r}$ but here we give a simple sufficient condition:

Theorem 2.7.7. *Let X be a Lévy process. Suppose that ϕ is bounded and that u_0 is continuous, bounded, and satisfies*

$$\mathbb{E}\left[e^{Cu_0(X_t)^k}\right] < \infty \tag{2.7.12}$$

for all t, where $C(t, X_t) = \frac{t}{-k}e^{k\int_0^t \phi(X_r)\mathrm{d}r}$. Then, there exists a solution to

$$\begin{aligned} u_t &= \mathcal{L}u + u(\phi - \tfrac{1}{k}u^k) \\ u(0, \cdot) &= u_0 \end{aligned} \tag{2.7.13}$$

in the sense of Definition 2.1.1 and it is given by $u(t, x) = U_t(u_0)(x)$.

In the case when X is a Brownian motion we have the following corollary.

Corollary 2.7.8. *Let X be standard one-dimensional Brownian motion. Suppose that ϕ is bounded and that u_0 is bounded, continuous and satisfies*

$$u_0(x) \ge c|x|^{-\alpha}$$

for $|x|$ sufficiently large and $\alpha < \frac{2}{-k}$. Then, there exists a solution to

$$\begin{aligned} u_t &= \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + u(\phi - \tfrac{1}{k}u^k) \\ u(0, \cdot) &= u_0 \end{aligned} \tag{2.7.14}$$

in the sense of Definition 2.1.1 and it is given by $u(t, x) = U_t(u_0)(x)$.

Proof of Corollary 2.7.8. By Theorem 2.7.7, the result will follow if

$$\mathbb{E}\left[\exp\left(c_2(t)|X_t|^{(-k)\alpha}\right)\mathbb{1}_{|X_t|\ge c_1}\right] < \infty$$

where $c_2 = \frac{1}{-k}cte^{bt}$ and b is a constant such that $\phi \ge b/k$, and $c_1 > 0$ is some constant. Since X is a Brownian motion, the expectation is finite whenever $(-k)\alpha < 2$. $\square$

Proof of Theorem 2.7.7. We need to show the convergence and uniform integrability properties in (2.7.10) and (2.7.11), respectively.

Since u_0 is bounded above, it follows, from Corollary 2.6.1, that $u(s, x)$ is bounded above for all $s \le t$ and $x \in E$. Therefore, from the discussion above if we bound

$$\left(e^{\int_0^{t-s} h(X_r, u(t-r, X_r))\mathrm{d}r}\right)_{s\le t}$$

by a function in $L^1(\mathbb{P})$, we will have the required uniform integrability of (2.7.11).

From (2.7.7) and (2.7.8), we have the following trivial bound:

$$u(t,x) \geq \mathbb{E}\left[e^{\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t)\right]$$

and therefore

$$u(t-s,X_s) \geq e^{-\int_0^s \phi(X_r)\mathrm{d}r}\mathbb{E}\left[e^{\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t)\big|\mathcal{F}_s\right]$$

and $N_s := \mathbb{E}\left[e^{\int_0^t \phi(X_r)\mathrm{d}r} u_0(X_t)\big|\mathcal{F}_s\right]$ defines a martingale.

Write $\phi \leq c$. Then, for all $s \in [0,t]$,

$$\begin{aligned}
\mathbb{E}\left[e^{\int_0^{t-s} h(X_r,u(t-r,X_r))\mathrm{d}r}\right] &\leq e^{ct}\mathbb{E}\left[\exp\left(\tfrac{1}{-k}\int_0^{t-s} N_r^k \mathrm{d}r\right)\right] \\
&\leq e^{ct}\mathbb{E}\left[\sup_{r\leq t}\exp\left(\tfrac{t}{-k}N_r^k\right)\right] \\
&\leq e^{ct+1}\mathbb{E}\left[\exp\left(\tfrac{t}{-k}N_t^k\right)\right] \\
&\leq e^{ct+1}\mathbb{E}\left[\exp\left(\tfrac{t}{-k}e^{k\int_0^t \phi(X_r)\mathrm{d}r}u_0(X_t)^k\right)\right]
\end{aligned}$$

In the first line we used the fact that $\phi \leq c$ and the definition of N. To go from the second to third line, we use Doob's submartingale inequality with the fact that $\exp(CN^k)$ and CN^k are sub-martingales for any constant $C > 0$ independent of s. In particular we used the fact that if $(M_s)_{s\leq t}$ is a submartingale, then for $p > 1$

$$\mathbb{E}\left[\sup_{s\leq t} e^{M_s}\right] \leq \mathbb{E}\left[\sup_{s\leq t}(e^{\frac{M_s}{p}})^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[\sup_{s\leq t} e^{M_t}\right]$$

and $\inf_{p>1}\left(\frac{p}{p-1}\right)^p = e$. Therefore, by condition (2.7.12), we have the required uniform integrability.

Next, we show that the convergence in (2.7.10). We will use the product rule for convergence in probability and show that

$$u(\delta, X_{t-\delta}) \xrightarrow{P} u_0(X_t) \tag{2.7.15}$$

and that

$$e^{\int_0^{t-\delta} h(X_s,u(t-s,X_s))\mathrm{d}s} \xrightarrow{P} e^{\int_0^t h(X_s,u(t-s,X_s))\mathrm{d}s}. \tag{2.7.16}$$

The convergence in (2.7.15) follows from Corollary 2.6.1. Indeed, take $k \in (-1,0)$ (the case $k < -1$ is similar). Then,

$$\begin{aligned}
&\mathbb{E}\left[\left(e^{-k\int_0^\delta \phi(X_{t-\delta+r})\mathrm{d}r}u_0(X_t)^{-k} + \int_0^\delta e^{-k\int_0^s \phi(X_{t-\delta+r})\mathrm{d}r}\mathrm{d}s\right)^{-1/k}\Bigg|X_{t-\delta}\right] \\
&\leq u(\delta, X_{t-\delta}) \leq \\
&\left(\mathbb{E}\left[e^{\int_0^\delta \phi(X_{t-\delta+r})\mathrm{d}r}u_0(X_t)\big|X_{t-\delta}\right]^{-k} + \int_0^\delta \mathbb{E}\left[e^{\int_0^s \phi(X_{t-\delta+r})\mathrm{d}r}\big|X_{t-\delta}\right]^{-k}\mathrm{d}s\right)^{-1/k}
\end{aligned}$$

and the upper and lower bounds converge to $u_0(X_t)$ in probability, since u_0 is continuous and bounded and X is continuous in probability and so the limit (2.7.15) holds.

From above, we see that

$$\int_0^t h(X_s, u(t-s,X_s))\mathrm{d}s < \infty$$

almost surely and so the expression

$$\left| e^{\int_0^{t-\delta} h(X_s,u(t-s,X_s))\mathrm{d}s} - e^{\int_0^{t} h(X_s,u(t-s,X_s))\mathrm{d}s} \right|$$

is defined and converges to 0, proving limit (2.7.16). Therefore, we have the required convergence in probability and it follows that $u(t,x)$ is given by equation (2.7.9), as required. □

Chapter 3.

An Application To A Merton-Type Utility Maximisation Problem

3.1. Merton-Type Utility Maximisation Problems

In this chapter we consider an HJB equation coming from a utility maximisation problem. This equation is usually not considered as an FKPP-type equation since the non-linearity is not of the KPP form but we see that in this case the equation is in the class of equations of the form (2.1.1).

The mathematics of utility maximisation problems has been intensively studied since Merton wrote his seminal papers [Mer69, Mer71]. We will use a model based on Merton's work as well as ideas from the work of Zariphopoulou [Zar01].

A summary of the financial model that we consider is the following. We will study a finite horizon model. In the model, our investor wants to maximise their utility over a fixed time interval and the utility function will depend on the amount of money they consume and also the wealth left over at the end of the time interval. The investor can only increase their wealth by investing in either a riskless account or a risky stock and can only consume money coming from their wealth. The investor can choose how to invest and also how much they consume at all times. In this model there are no transaction costs. Now we will set up the problem rigorously below.

Consider the following utility maximisation problem. Firstly, we suppose that we have a simple market model. In particular, suppose that we have a money market account with price process $(B_s)_{s\geq 0}$ and a stock with price process $(S_s)_{s\geq 0}$. We assume that the dynamics evolve according to

$$\begin{aligned} \mathrm{d}B_s &= r_s B_s \mathrm{d}s \\ \mathrm{d}S_s &= S_s(\mu_s \mathrm{d}s + \sigma_s \mathrm{d}W_s) \end{aligned} \tag{3.1.1}$$

where r, μ and σ are predictable and locally bounded and W is a standard Brownian motion. The process r can be interpreted as the interest rate which evolves over time. μ represents the mean return of the stock and σ is the volatility process.

Now, we introduce an investor into the market. We write Y_s for the total wealth at time s. To avoid doubling strategies, we will make the assumption that the investor can't go into debt;

that is,

$$Y_s > 0 \text{ for all } s \geq 0 \text{ almost surely.}$$

We write π_s^0 for the number of shares invested in the money market account and π_s for the number of shares invested in the stock at time s. We allow the processes π^0 and π to take any positive real values. Therefore, by definition,

$$Y_s = \pi_s^0 B_s + \pi_s S_s.$$

We allow the investor to consume at a rate $c_s \geq 0$ at time s and assume the self financing condition. In this case, this means that the wealth of the investor evolves according to

$$\mathrm{d}Y_s = \pi_s^0 \mathrm{d}B_s + \pi_s \mathrm{d}S_s - c_s \mathrm{d}s.$$

Combining these we see that

$$\begin{aligned} \mathrm{d}Y_s &= (r_s Y_s + \pi_s(\mu_s - r_s)S_s - c_s)\mathrm{d}s + \pi_s \sigma_s S_s \mathrm{d}W_s \\ &= Y_s[(r_s + \theta_s \lambda_s - \eta_s)\mathrm{d}s + \theta_s \mathrm{d}W_s] \end{aligned} \tag{3.1.2}$$

where we define θ and η to be the normalised controls given by

$$\begin{aligned} \theta_s Y_s &= \pi_s \sigma_s S_s, \\ \eta_s Y_s &= c_s, \end{aligned} \tag{3.1.3}$$

and λ to be the Sharpe ratio defined by

$$\lambda_s = \frac{\mu_s - r_s}{\sigma_s}.$$

Thus, given initial wealth Y_0, the process $Y^{\theta,\eta}$ is well defined if

$$\int_0^t \left(|r_s + \theta_s \lambda_s - \eta_s| + \theta_s^2\right) \mathrm{d}s < \infty \text{ almost surely.} \tag{3.1.4}$$

The original controls, can be recovered via

$$\begin{aligned} \pi_s &= \frac{\theta_s Y_s^{\theta,\eta}}{\sigma_s S_s} \\ \pi_s^0 &= \frac{Y_s^{\theta,\eta}}{B_s}\left(1 - \frac{\theta_s}{\sigma_s}\right) \\ c_s &= \eta_s Y_s^{\theta,\eta}. \end{aligned} \tag{3.1.5}$$

Now, we can state the utility maximisation problem. For a fixed time horizon $t > 0$, we aim to maximise

$$\mathbb{E}\left[\int_0^t e^{-\int_0^s \gamma_\nu \mathrm{d}\nu} U(c_s)\mathrm{d}s + e^{-\int_0^t \gamma_\nu \mathrm{d}\nu} G_t U(Y_t)\right]$$

where the process $(\gamma_s)_{0 \leq s \leq t}$ is the investor's subjective stochastic rate of discounting, U is an increasing concave function representing the investor's utility function, and where $G_t > 0$ is an additional random factor.

This type of problem was first considered by Merton in his seminal paper [Mer69]. In order to make the problem tractable, however, we make some simplifications based on the work of Zariphopoulou [Zar01].

We introduce an underlying economic process $(X_s)_{0\leq s\leq t}$ which evolves according to

$$\mathrm{d}X_s = \beta(X_s)\mathrm{d}s + \alpha(X_s)\mathrm{d}\tilde{W}_s$$

where α and β are regular enough (for instance, Lipschitz continuous) for there to exist a unique strong solution X starting from a given $X_0 = x$ and where $\tilde{W}$ is a standard Brownian Motion. We will work with the filtration $\mathcal{F}$ generated by the pair $(W, \tilde{W})$.

We suppose that W and $\tilde{W}$ have a constant covariation ρ and so

$$\langle W, \tilde{W}\rangle_s = \rho s.$$

In the market, we assume that the spot interest rate and the Sharpe ratio depend on the economic factor X. That is

$$\begin{aligned} r_s &= r(X_s), \\ \lambda_s &= \lambda(X_s), \text{ for all } s \in [0,t] \end{aligned}$$

where $r, \lambda : \mathbb{R} \to \mathbb{R}$ with some slight abuse of notation.

Apart from the assumption that the correlation ρ is constant and that the economic factor X is one-dimensional, the construction given is fairly general.

Now, for the preferences of the investor, we also assume that the subjective rate of discounting and G depend on the economic factor X:

$$\begin{aligned} \gamma_s &= \gamma(X_s) \text{ for all } s \in [0,t], \\ G_t &= G(X_t), \end{aligned}$$

where $\gamma, G : \mathbb{R} \to \mathbb{R}$, again with some slight abuse of notation.

Finally, we assume that the utility function U corresponds to constant relative risk aversion (CRRA) and is of the form

$$U(y) = \frac{y^{1-R}}{1-R}, \text{ for } R > 0.$$

3.2. Solving the Maximisation Problem

Here, we show that it is possible to write this maximisation problem over two variables in terms of the solution to an FKPP-type equation of the form

$$\frac{\partial u}{\partial t} = \mathcal{L}u + \phi u - \tfrac{1}{k}u^{1+k}.$$

We will consider the following special case: let

$$\delta = \frac{R}{R + (1-R)\rho^2}.$$

and set

$$\begin{aligned} \mathcal{L}u &= \frac{\alpha^2}{2}\frac{\partial^2 u}{\partial x^2} + \left[\beta + \frac{1-R}{R}\alpha\rho\lambda\right]\frac{\partial u}{\partial x} \\ \phi &= \frac{1-R}{\delta}\left[r + \frac{\lambda^2}{2R} - \frac{\gamma}{1-R}\right] \\ k &= -\frac{\delta}{R}. \end{aligned} \tag{3.2.1}$$

In this case, the operator $\mathcal{L}$ is the stochastic generator of the process, Z, solving the SDE

$$\mathrm{d}Z_s = \alpha(Z_s)\mathrm{d}W_s + \left(\beta(Z_s) + \frac{1-R}{R}\rho\alpha(Z_s)\lambda(Z_s)\rho\right)\mathrm{d}s. \tag{3.2.2}$$

Note that the particular value of R used in practice is debated and depends on the individual investor. In the important paper by Mehra and Prescott [MP85] on the Equity Premium Puzzle, R is set to be between 0 and 10. Note that, if $R \in (0,1)$, then $k \in (-\infty,-1)$ and if $R \in (1,\infty)$, then $k \in (-1,0)$.

Before we give the representation, we introduce the following technical definition.

Definition 3.2.1. *We say that a pair of controls (θ,η) has Property (D) if the process defined by*

$$e^{-\int_0^s \gamma(X_r)\mathrm{d}r}u(t-s,X_s)^\delta U(Y_s^{\theta,\eta})$$

is in class D.

Assumption 3.2.2. *We adopt the set-up from Section 3.1 and define δ as above. Suppose that the following holds:*

1. *$G^{1/\delta} \in C^2(\mathbb{R})$ and is bounded and uniformly continuous and there exists $\varepsilon > 0$ such that $G(x) > \varepsilon$ for all $x \in \mathbb{R}$.*

2. *α and β are globally Lipschitz and bounded, and r, λ, and γ are continuously differentiable and bounded.*

Theorem 3.2.3. *Suppose that Assumption 3.2.2 holds. Then, there exists a classical solution $u : [0,\infty)\times\mathbb{R} \to (0,\infty)$ to*

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\alpha^2}{2}\frac{\partial^2 u}{\partial x^2} + \left[\beta + \frac{1-R}{R}\alpha\rho\lambda\right]\frac{\partial u}{\partial x} + \frac{1-R}{\delta}\left[r + \frac{\lambda^2}{2R} - \frac{\gamma}{1-R}\right]u + \frac{R}{\delta}u^{1-\delta/R} \\ u(0,x) &= G(x)^{1/\delta}. \end{aligned} \tag{3.2.3}$$

Moreover, we have

$$U(Y_0)u(t,x)^\delta = \max_{\theta,\eta}\mathbb{E}\left[\int_0^t e^{-\int_0^s \gamma(X_r)\mathrm{d}r}U(\eta_s Y_s^{\theta,\eta})\mathrm{d}s + e^{-\int_0^t \gamma(X_r)\mathrm{d}r}G(X_t)U(Y_t^{\theta,\eta})\right]$$

with the maximum taken over all admissible strategies such that Property (D) holds. Suppose also that Property (D) holds for (θ^*,η^*).

The optimal θ and η are given by

$$\begin{aligned} \theta_s^* &= \frac{1}{R}\left(\lambda(X_s) + \rho\delta\alpha(X_s)\frac{\partial}{\partial x}\log u(t-s,X_s)\right), \\ \eta_s^* &= u(t-s,X_s)^{-\delta/R}, \end{aligned} \tag{3.2.4}$$

Proof of Theorem 3.2.3. Since $G > \varepsilon$ for some $\varepsilon > 0$ and ϕ (defined in equation (3.2.1)) is bounded below, Theorem 2.7.6 applies and guarantees the existence of a mild solution with $u \geq \tilde{\varepsilon}$ for some $\tilde{\varepsilon} > 0$. On the range $[\tilde{\varepsilon}, \infty)$, the nonlinearity f, defined by $f(u) = \phi u - \frac{1}{k}u^{1+k}$, is continuously differentiable and outside of this range it can be modified, without loss of generality, to be smooth, globally Lipschitz and bounded. Therefore, we can consider the PDE with this modified nonlinearity which we denote by $\tilde{f}$.

Let $C_{u,b}(\mathbb{R})$ be the Banach space of bounded and uniformly continuous functions mapping $\mathbb{R} \to \mathbb{R}$ equipped with the uniform norm. If $u \in C_{u,b}(\mathbb{R})$, then $\tilde{f}(u) \in C_{u,b}(\mathbb{R})$. By Assumption 3.2.2, it follows that $\mathcal{L}$ generates a strongly continuous semigroup in $C_{u,b}(\mathbb{R})$ and $G^{1/\delta}$ is in the domain of $\mathcal{L}$. Therefore, the mild solution u is also a classical solution – see, for example, Pazy [Paz92, Theorem 6.1.5] and a straightforward adaptation of the proof in [SP12, Theorem 19.9].

Now we show why the representation holds. We use Itô's Lemma:

$$\begin{aligned}
(\star) &\equiv \mathrm{d}\left(\int_0^s e^{-\int_0^r \gamma(X_\nu)\mathrm{d}\nu} U(\eta_r Y_r)\mathrm{d}r + e^{-\int_0^s \gamma(X_r)\mathrm{d}r} u(t-s, X_s)^\delta U(Y_s)\right) \\
&= e^{-\int_0^s \gamma(X_r)\mathrm{d}r}\left[U(\eta_s Y_s) - \gamma(X_s)u(t-s,X_s)^\delta U(Y_s) + \mathrm{d}(u(t-s,X_s)^\delta U(Y_s))\right].
\end{aligned}$$

Note that

$$\begin{aligned}
\mathrm{d}U(Y_s) &= Y_s^{1-R}\left[\left(r + \left(\theta\lambda - \tfrac{R}{2}\theta^2\right) - \eta\right) + \theta \mathrm{d}W_s\right] \\
\mathrm{d}u(t-s,X_s)^\delta &= \delta u^{\delta-1}\left[\left(-u_t + \tfrac{\alpha^2}{2}u_{xx} + \beta u_x + \alpha^2 \tfrac{\delta-1}{2}\tfrac{u_x^2}{u}\right)\mathrm{d}s + \alpha u_x \mathrm{d}\tilde{W}_s\right] \\
\mathrm{d}\langle U(Y_\cdot), u(t-\cdot, X_\cdot)^\delta\rangle_s &= \delta u^{\delta-1} Y_s^{1-R}\alpha\rho\theta u_x
\end{aligned}$$

Then, for all θ, η, we have

$$\begin{aligned}
(\star) = e^{-\int_0^s \gamma(X_r)\mathrm{d}r} Y_s^{1-R} u^{\delta-1} &\left[\tfrac{1}{u^{\delta-1}}\left(\tfrac{\eta^{1-R}}{1-R} - \eta u^\delta\right) + \left((\lambda u + \delta\rho\alpha u_x)\theta - \tfrac{Ru}{2}\theta^2\right)\right. \\
&\left. + \left(r - \tfrac{\gamma}{1-R}\right)u + \tfrac{\delta}{1-R}\left(-u_t + \tfrac{\alpha^2}{2}u_{xx} + \beta u_x + \tfrac{\delta-1}{2}\alpha^2\tfrac{u_x^2}{u}\right)\right] \\
&+ \mathrm{d}M_s
\end{aligned}$$

where M is a local martingale.

We see that for $R > 0$,

$$\frac{1}{u^{\delta-1}}\left(\frac{\eta^{1-R}}{1-R} - \eta u^\delta\right) \leq \frac{R}{1-R}u^{1-\delta/R}$$

with equality if and only if $\eta = \eta^*$ and

$$\left((\lambda u + \delta\rho\alpha u_x)\theta - \frac{Ru}{2}\theta^2\right) \leq \frac{(\lambda u + \delta\rho\alpha u_x)^2}{2Ru}$$

with equality if and only if $\theta = \theta^*$.

Combining this with the fact that u solves the PDE (3.2.3) we see that

$$\begin{aligned}
(\star) &\leq \left(\frac{\delta^2\rho^2}{R} + \frac{\delta(\delta-1)}{1-R}\right) e^{-\int_0^s \gamma(X_r)\mathrm{d}r} Y_s^{1-R} u^{\delta-2}\frac{\alpha^2}{2}u_x^2 \mathrm{d}s + \mathrm{d}M_s \\
&= \mathrm{d}M_s
\end{aligned}$$

by the definition of δ; this is the reason for choosing δ as it is defined.

Let $(\tau_n)_{n\geq 1}$ be an increasing sequence of stopping times that reduce M and such that $\tau_n \nearrow t$ as $n \to \infty$. Then, integrating over $[0, \tau_n]$ and taking expectations gives

$$\mathbb{E}\left[\int_0^{\tau_n} e^{-\int_0^r \gamma(X_\nu)\mathrm{d}\nu} U(\eta_r Y_r)\mathrm{d}r + e^{-\int_0^{\tau_n} \gamma(X_\nu)\mathrm{d}\nu} u(t-\tau_n, X_{\tau_n})^\delta U(Y_{\tau_n})\right] \leq u(t,x)^\delta U(Y_0)$$

with equality for θ^* and η^* defined above.

We can interchange the limit and the expectation for the first term by the Monotone Convergence Theorem and the second term by since

$$e^{-\int_0^s \gamma(X_r)\mathrm{d}r} u(t-s, X_s)^\delta U(Y_s^{\theta,\eta})$$

is in class D, by assumption. □

Remark 8. *In the case that $R \in (0,1)$, we can maximise over a larger set of controls; we do not need to specify that all controls (θ, η) satisfy Property (D). Indeed, we only require this for the optimal controls; in the proof above, we can use Fatou's Lemma to see*

$$\mathbb{E}\left[\int_0^{t} e^{-\int_0^r \gamma(X_\nu)\mathrm{d}\nu} U(\eta_r Y_r)\mathrm{d}r + e^{-\int_0^{t} \gamma(X_\nu)\mathrm{d}\nu} G(X_t)^\delta U(Y_t)\right] \leq u(t,x)^\delta U(Y_0)$$

since all terms are positive. We then only use the class D assumption for interchanging limits for the optimal controls. ◇

3.2.1. Motivation for Theorem 3.2.3

The FKPP type equation (3.2.3), occurs naturally in this setting; in fact, it arises from the HJB equation corresponding to the maximisation problem. We outline the details here. The following type of argument is due to Zariphopoulou [Zar01].

Define the value function, v, by

$$v(s,x,y) = \sup \mathbb{E}\left[\int_s^{t} e^{-\int_s^r \gamma(X_\nu)\mathrm{d}\nu} U(\eta_r Y_r)\mathrm{d}r + e^{-\int_s^{t} \gamma(X_r)\mathrm{d}r} G(X_t) U(Y_t)\middle| X_s = x, Y_s = y\right]$$

where the supremum is taken over all admissible processes $(\theta_r, \eta_r)_{s\leq r\leq t}$.

The corresponding HJB problem is:

$$\begin{aligned} 0 &= \frac{\partial v}{\partial s} + \frac{\alpha^2}{2}\frac{\partial^2 v}{\partial x^2} + \beta \frac{\partial v}{\partial x} - \gamma v \\ &\quad + \sup_{(\theta,\eta)}\left[U(\eta y) + y(r + \theta\lambda - \eta)\frac{\partial v}{\partial y} + \frac{y^2\theta^2}{2}\frac{\partial^2 v}{\partial y^2} + y\rho\theta\alpha \frac{\partial^2 v}{\partial y \partial x}\right] \qquad (3.2.5) \\ v(t,x,y) &= G(x)U(y) \end{aligned}$$

Considering the scaling properties of the utility function, we make the ansatz $v(s,x,y) = U(y)w(t-s,x)$. Then, finding v simplifies to solving the equation

$$\begin{aligned} \frac{\partial w}{\partial s} &= \frac{\alpha^2}{2}\frac{\partial^2 w}{\partial x^2} + \left[\beta + \frac{(1-R)}{R}\lambda\rho\alpha\right]\frac{\partial w}{\partial x} + (1-R)\left[r + \frac{\lambda^2}{2R} - \frac{\gamma}{1-R}\right] w \\ &\quad + R w^{1-1/R} + \frac{1-R}{2R}\rho^2\alpha^2 \frac{\left(\frac{\partial w}{\partial x}\right)^2}{w} \qquad (3.2.6) \\ w(0,x) &= G(x). \end{aligned}$$

This was obtained by simply substituting $v(s,x,y) = U(y)w(t-s,x)$ and noting that due to the special form of U, the finding the supremum simplifies to maximising two simpler expressions:

$$\max_{\eta>0}\left[\frac{\eta^{1-R}}{1-R} - \eta w\right]$$

and

$$\max_{\theta>0}\left[\left(\lambda w + \rho\alpha\frac{\partial w}{\partial x}\right)\theta - \frac{Rw}{2}\theta^2\right]$$

This gives optimal controls

$$\begin{aligned}\theta_s^* &= \frac{1}{R}\left(\lambda(X_s) + \rho\alpha(X_s)\frac{\partial}{\partial x}\log w(t-s,X_s)\right),\\ \eta_s^* &= w(t-s,X_s)^{-1/R},\end{aligned} \tag{3.2.7}$$

Finally, in order to remove the nonlinearity in the derivative term, we appeal to the transformation popularised by Zariphopoulou [Zar01], $w(s,x) = u(s,x)^\delta$, where

$$\delta = \frac{R}{R+(1-R)\rho^2}.$$

Then, solving the HJB equation reduces to solving

$$\begin{aligned}\frac{\partial u}{\partial t} &= \frac{\alpha^2}{2}\frac{\partial^2 u}{\partial x^2} + \left[\beta + \frac{1-R}{R}\alpha\rho\lambda\right]\frac{\partial u}{\partial x} + \frac{1-R}{\delta}\left[r + \frac{\lambda^2}{2R} - \frac{\gamma}{1-R}\right]u + \frac{R}{\delta}u^{1-\delta/R}\\ u(0,x) &= G(x)^{1/\delta}\end{aligned} \tag{3.2.8}$$

In terms of u, the optimal controls become

$$\begin{aligned}\theta_s^* &= \frac{1}{R}\left(\lambda(X_s) + \rho\delta\alpha(X_s)\frac{\partial}{\partial x}\log u(t-s,X_s)\right),\\ \eta_s^* &= u(t-s,X_s)^{-\delta/R}.\end{aligned} \tag{3.2.9}$$

Chapter 4.

Wave Fronts in FKPP-Type Equations

As we noted in Chapter 1, there has been much work on studying travelling wave solutions to FKPP-type equations and also in studying the propagation and speed of the wave front for a given initial condition. In this Chapter, we see how we can use the representations of Chapter 2 to study the speed of the wave front for several examples of FKPP type equations. In particular, we use the optimisation-based representations to obtain bounds on the solution to the FKPP-type equations, and by comparing the speed of the fronts of the upper and lower bounds, we can find bounds the speed of the wave front.

As we noted before, the role of the Markov process X does not play a role in obtaining the representations. For this reason, it is to be expected that the methods we use below for the case when X is a Brownian motion and u solves the classical FKPP equation may be generalised. This turns out to be the case; we will introduce the results separately for the case of the classical FKPP equation and then generalise suitably for other types of equation.

First of all, we need to provide upper and lower bounds on the solution to equation (2.1.1) under some mild conditions.

4.1. Bounds on u

We introduce bounds on the solution to

$$\begin{aligned} u_t &= \mathcal{L}u + \tfrac{1}{k}u(1-u^k), \text{ in } (0,\infty)\times\mathbb{R}^d \\ u(0,x) &= u_0(x), \text{ for } x \in \mathbb{R}^d, \end{aligned} \tag{4.1.1}$$

and write $(X_t)_{t\geq 0}$ for the Lévy process with $X_0 = 0$ with stochastic generator given by $\mathcal{L}$. Here we deal with the case when $k > 0$.

We will then be able to use the bounds below to describe the position of the front of the solution given an initial condition u_0 satisfying a simple, exponential decay condition.

For simplicity later on, we introduce the following notation.

Definition 4.1.1. *Define the function D by*

$$D_t(f)(x) = \frac{f(x)e^{t/k}}{\left(f(x)^k(e^t-1)+1\right)^{1/k}} \tag{4.1.2}$$

for any function $t \geq 0$, $x \in \mathbb{R}^d$, *and* $f : \mathbb{R}^d \to \mathbb{R}$. *Define the function* P *by*

$$P_t(f)(x) = \mathbb{E}[f(x + X_t)] \tag{4.1.3}$$

for any function $t \geq 0$, $x \in \mathbb{R}^d$, *and* $f : \mathbb{R}^d \to \mathbb{R}$.

We have the following bounds on u.

Lemma 4.1.2. *Let* u *be a solution to equation* (4.1.1) *and* D *and* P *be defined as above. Then,*

$$(P_t \circ D_t)(u_0)(x) \leq u(t, x) \leq (D_t \circ P_t)(u_0)(x).$$

This follows from rewriting Corollary 2.6.1.

Remark 9. *Let* P *and* D *defined as above. By definition, for sufficiently regular* v_0, v *defined by* $v(t, x) = P_t(v_0)(x)$ *solves*

$$\frac{\partial v}{\partial t} = \mathcal{L}v$$
$$v(0, x) = v_0(x)$$

and w *defined by* $w(t, x) = D_t(w_0)(x)$ *solves*

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{1}{k}w(1 - w^k)$$
$$w(0, x) = w_0(x)$$

since D *corresponds to the deterministic optimisation problem.*

In other words, P *and* D *are semigroups and the bounds of Lemma 4.1.2 correspond to splitting equation* (4.1.1) *into the linear and nonlinear parts. In the spirit of the Trotter-Kato Theorem [Kat78], we will iterate these bounds over small time steps and obtain convergent bounds as the time-step size converges to zero. In fact, this is the content of Section 4.1.1 and for* t *fixed, we will have uniform convergence.* $\diamond$

4.1.1. Iterative Bounds

We now generalise the above bounds by approximating over small time steps.

In order to generalise Lemma 4.1.2 we will need the following simple observation.

Lemma 4.1.3. *Let* D *and* P *be defined as above. Fix* t. *Then,* D_s *and* P_s *are increasing; that is, if* $f_1, f_2 : \mathbb{R}^d \to \mathbb{R}$ *with*

$$f_1(x) \leq f_2(x) \text{ for all } x,$$

then

$$D_t(f_1)(x) \leq D_t(f_2)(x)$$

and

$$P_t(f_1)(x) \leq P_t(f_2)(x).$$

Moreover, if $l(x) = x$, *then* $D_t(l) : [0, 1] \to [0, 1]$ *with* $D_t(l)(0) = 0$ *and* $D_t(l)(1) = 1$. *Also, with some abuse of notation,* $P_t(0) = 0$ *and* $P_t(1) = 1$. *Finally,* $D_0(f) = f$ *and* $P_0(f) = f$ *for all* $x \in \mathbb{R}^d$ *and* $f : \mathbb{R}^d \to \mathbb{R}$.

The proof is clear.

We introduce the following notation.

$$\mathbb{Y} = \{Y : Y : [0,t] \to \mathbb{R}, \text{is absolutely continuous, decreasing and } Y_0 = 1\}. \tag{4.1.4}$$

We will obtain the result by using Theorem 2.2.3 and minimising over a subset and superset of set

$$\mathcal{Y} = \{Y \in \mathbb{Y} : Y \text{ is adapted to the filtration } \mathcal{F}\}$$

We denote, by $(t_k)_{\{0\leq k\leq n\}}$, where $t_0 = 0$ and $t_n = t$, a partition of the interval $[0,t]$. Then, we set

$$\overline{\mathcal{Y}}_n = \left\{ Y \in \mathbb{Y} : Y_s = \sum_{k=0}^{n-1} \mathbb{1}_{(t_k,t_{k+1}]}(s)\xi_s^k, \xi_s^k \in \mathcal{F}_{t_k} \text{ for } s \in [t_k, t_{k+1}) \right\}. \tag{4.1.5}$$

This is a generalisation of the set of deterministic functions.

We also define

$$\underline{\mathcal{Y}}_n = \left\{ Y \in \mathbb{Y} : Y_s = \sum_{k=1}^{n} \mathbb{1}_{(t_{k-1},t_k]}(s)\xi_s^k, \xi_s^k \in \mathcal{F}_{t_k} \text{ for } s \in (t_{k-1}, t_k] \right\}, \tag{4.1.6}$$

to be piecewise functions that can see ahead to the next point in time in the partition.

By definition, we have

$$\overline{\mathcal{Y}}_n \subset \mathcal{Y} \subset \underline{\mathcal{Y}}_n.$$

As we saw in Lemma 4.1.2,

$$(P_t \circ D_t)(u_0)(x) \leq u(t,x) \leq (D_t \circ P_t)(u_0)(x).$$

If we iterate this approximation backwards over each time point in the partition we will obtain a better bounds on u. This suggests the following result

Theorem 4.1.4.

$$\begin{aligned} &(P_{t_1-t_0} \circ D_{t_1-t_0}) \circ \cdots \circ (P_{t_n-t_{n-1}} \circ D_{t_n-t_{n-1}})(u_0)(x) \\ &\leq u(t,x) \\ &\leq (D_{t_1-t_0} \circ P_{t_1-t_0}) \circ \cdots \circ (D_{t_n-t_{n-1}} \circ P_{t_n-t_{n-1}})(u_0)(x) \end{aligned} \tag{4.1.7}$$

Mostly, we will use the partition with $t_k = kt/n$ and then this result simply becomes

$$(P_{\frac{t}{n}} \circ D_{\frac{t}{n}})^{(n)}(u_0)(x) \leq u(t,x) \leq (D_{\frac{t}{n}} \circ P_{\frac{t}{n}})^{(n)}(u_0)(x). \tag{4.1.8}$$

Proof of Theorem 4.1.4. Using Theorem 2.2.3 we see that

$$u(t,x) = \min_{\mathcal{Y}} \mathbb{E}\left[\int_0^t e^{s/k}|\dot{Y}_s|^{1+1/k}\mathrm{d}s + e^{t/k}u_0(x+X_t)|Y_t|^{1+1/k}\right].$$

If we replace $\mathcal{Y}$ by $\overline{\mathcal{Y}}_n$, we will obtain an upper bound on $u(t,x)$ since $\overline{\mathcal{Y}}_n \subset \mathcal{Y}$. Similarly, if we replace $\mathcal{Y}$ by $\underline{\mathcal{Y}}_n$, we obtain a lower bound on $u(t,x)$.

1. **Upper Bound**

   We will write $\overline{\mathcal{Y}}_{i,j}$ for processes in $\mathcal{Y}$ defined analogously for a partition on the interval $[t_i, t_j]$.

   Then we write

$$V^j(x) = \min_{\overline{\mathcal{Y}}_{j,n}} \mathbb{E}\left[\int_0^{t-t_j} e^{r/k}|\dot{Y}_r|^{1+1/k}\mathrm{d}r + e^{(t-t_j)/k}u_0(x+X_{t-t_j})|Y_{t-t_j}|^{1+1/k}\right]$$

   and so $V^n(x) = u_0(x)$

$$\begin{aligned}
V^0(x) &= u(t,x) \\
&\le \min_{\overline{\mathcal{Y}}_n} \mathbb{E}\left[\int_0^t e^{s/k}|\dot{Y}_s|^{1+1/k}\mathrm{d}s + e^{t/k}u_0(x+X_t)|Y_t|^{1+1/k}\right] \\
&= \min_{\overline{\mathcal{Y}}_n} \mathbb{E}\left[\int_0^{t_1} e^{s/k}|\dot{Y}_s|^{1+1/k}\mathrm{d}s\right. \\
&\qquad\qquad \left. + \mathbb{E}\left[\int_{t_1}^t e^{s/k}|\dot{Y}_s|^{1+1/k}\mathrm{d}s + e^{t/k}u_0(x+X_t)|Y_t|^{1+1/k}\middle|\mathcal{F}_{t_1}\right]\right] \\
&= \min_{\overline{\mathcal{Y}}_{0,1}} \mathbb{E}\left[\int_0^{t_1} e^{s/k}|\dot{Y}_s|^{1+1/k}\mathrm{d}s + e^{t_1/k}|Y_{t_1}|^{1+1/k}V^1(x+X_{t_1})\right] \\
&= \min_{\overline{\mathcal{Y}}_{0,1}} \left(\int_0^{t_1} e^{s/k}|\dot{Y}_s|^{1+1/k}\mathrm{d}s + e^{t_1/k}|Y_{t_1}|^{1+1/k}\mathbb{E}\left[V^1(x+X_{t_1})\right]\right) \\
&= (D_{t_1}\circ P_{t_1})(V^1)(x)
\end{aligned}$$

   where the last line follows from Lemma 4.1.2 since the upper bound corresponds to the optimal value when minimising over optimal controls (see Corollary 2.6.1).

   We can pull out the Y_{t_1} factor since for $(Z_s)_{t_1\le s\le t}$, with $Z_s = \frac{Y_s}{Y_{t_1}}$ we have

$$Z_s = 1 + \int_{t_1}^t \dot{Z}_s$$

   and $Z \in \overline{\mathcal{Y}}_{1,n}$.

   Using Lemma 4.1.3, we can continue this process up to $V^n = u_0$, by induction and we are done.

2. **Lower Bound**

   The lower bound follows by a similar argument to the above: define

$$V^j(x) = \min_{\underline{\mathcal{Y}}_{j,n}} \mathbb{E}\left[\int_0^{t-t_j} e^{r/k}|\dot{Y}_r|^{1+1/k}\mathrm{d}r + e^{(t-t_j)/k}u_0(x+X_{t-t_j})|Y_{t-t_j}|^{1+1/k}\right]$$

   and so $V^n(x) = u_0(x)$

$$\begin{aligned}
V^0(x) &= u(t,x) \\
&\ge \min_{\underline{\mathcal{Y}}_n} \mathbb{E}\left[\int_0^t e^{s/k}|\dot{Y}_s|^{1+1/k}\mathrm{d}s + e^{t/k}u_0(x+X_t)|Y_t|^{1+1/k}\right] \\
&= \min_{\underline{\mathcal{Y}}_{0,1}} \mathbb{E}\left[\int_0^{t_1} e^{s/k}|\dot{Y}_s|^{1+1/k}\mathrm{d}s + e^{t_1/k}|Y_{t_1}|^{1+1/k}V^1(x+X_{t_1})\right] \\
&\ge (P_{t_1}\circ D_{t_1})(V^1)(x)
\end{aligned}$$

   by Lemma 4.1.2. □

Remark 10. *We see that the proof of 4.1.4 above requires that ϕ is constant so that we can use the Bellman method above.* $\diamond$

We will often use the following bound on D throughout this section.

Lemma 4.1.5. *For any $a \in (0,1)$, $f : \mathbb{R}^d \to [0,1]$ and $s > 0$, and we have*

$$D_s(f)(x) \geq a\mathbb{1}\left\{f(x) \geq abe^{-s/k}\right\}$$

where,

$$b = \frac{1}{(1-a^k)^{1/k}}.$$

Proof. We rearrange the equation $D_s(f)(x^*) = a$ for $f(x^*)$ to find that

$$f(x^*) = \frac{a}{(a^k + (1-a^k)e^s)^{1/k}}.$$

Since, D is increasing in the sense of Lemma 4.1.3, we find that

$$\begin{aligned} D_s(f)(x) &\geq a\mathbb{1}\{f(x) \geq f(x^*)\} \\ &\geq a\mathbb{1}\left\{f(x) \geq abe^{-s/k}\right\}. \end{aligned}$$ $\square$

Theorem 4.1.6. *For t fixed, the bounds in Theorem 4.1.4, with partition given as in equation* (4.1.8)*, converge to $u(t,x)$ uniformly in x.*

Proof.

$$(P_{\frac{t}{n}} \circ D_{\frac{t}{n}})^{(n)}(u_0)(x) \leq u(t,x) \leq (D_{\frac{t}{n}} \circ P_{\frac{t}{n}})^{(n)}(u_0)(x)$$

We consider the difference,

$$\begin{aligned} (D_{\frac{t}{n}} \circ P_{\frac{t}{n}})^{(n)}(u_0)(x) - (P_{\frac{t}{n}} \circ D_{\frac{t}{n}})^{(n)}(u_0)(x) &= (D_{\frac{t}{n}} - \mathrm{Id_x}) \circ P_{\frac{t}{n}} \circ (D_{\frac{t}{n}} \circ P_{\frac{t}{n}})^{(n-1)}(u_0)(x) \\ &\quad + (P_{\frac{t}{n}} \circ D_{\frac{t}{n}})^{(n-1)} \circ P_{\frac{t}{n}} \circ (\mathrm{Id_x} - D_{\frac{t}{n}})(u_0)(x) \\ &\equiv (A) + (B). \end{aligned}$$

By differentiating and finding the maximum, we see that

$$\begin{aligned} |D_{\frac{t}{n}}(f) - f| &\leq \frac{e^{\frac{t}{2n}} - 1}{e^{\frac{t}{2n}} + 1} \\ &= \tanh\left(\tfrac{t}{4n}\right) \\ &= \mathcal{O}(n^{-1}) \end{aligned}$$

for $f : \mathbb{R}^d \to [0,1]$ and so

$$|D_{\frac{t}{n}}(f) - f| \to 0, \text{ as } n \to 0.$$

Since $P_{\frac{t}{n}} \circ (D_{\frac{t}{n}} \circ P_{\frac{t}{n}})^{(n-1)}(u_0)(x)$ is always bounded above and below by 1 and 0, respectively, we have that $(A) \to 0$.

For the term (B), note that

$$|(B)| \leq D_{\frac{t}{n}}^{(n-1)}\left(\tanh\left(\tfrac{t}{4n}\right)\right)$$

Since, $D_\delta(f)(x) \leq f(x)e^\delta$, we have

$$
\begin{aligned}
|(B)| &\leq e^t \tanh\left(\tfrac{t}{4n}\right) \\
&\to 0
\end{aligned}
$$

as $n \to \infty$, for fixed t and so we are done. □

4.2. Speed of the Wave Front for the Canonical FKPP Equation

First of all, we consider the canonical FKPP equation in one dimension. Suppose that $u : [0, \infty) \times \mathbb{R} \to [0, 1]$ solves the equation

$$
\begin{aligned}
\frac{\partial u}{\partial t} &= \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \frac{1}{k}u(1-u^k) \\
u(0,x) &= u_0(x)
\end{aligned}
\tag{4.2.1}
$$

in the sense of Definition 2.1.1 for $k > 0$.

Here we show how bounds above can be used to estimate the speed of the wave front in this case given suitable initial conditions.

We will write $X_s = x + W_s$, where W is a standard Brownian motion. This corresponds to $\mathcal{L} = \frac{1}{2}\frac{\partial^2}{\partial x^2}$.

Firstly, we consider the case when $u_0 = \mathbb{1}_{(-\infty,0]}$.

Theorem 4.2.1. *Suppose that u is the solution to equation* (4.2.1) *with $u_0 = \mathbb{1}_{(-\infty,0)}$. Then, for all fixed $(t,x) \in [0,\infty) \times \mathbb{R}$, $n \geq 1$ and $a \in [0,1)$ we have*

$$
a\Phi\left(-x\sqrt{\frac{n}{t}} - (n-1)\Phi^{-1}\left(be^{-\frac{t}{nk}}\right)\right) \leq u(t,x) \leq e^{t/k}\Phi\left(-\frac{x}{\sqrt{t}}\right) \tag{4.2.2}
$$

where $b = (1-a^k)^{-1/k}$.

Proof. The upper bound follows from Theorem 4.1.4 with $n = 1$ (i.e. Lemma 4.1.2):

$$
\begin{aligned}
(D_t \circ P_t)(\mathbb{1}_{(-\infty,0)})(x) &= \frac{\mathbb{E}[\mathbb{1}_{(-\infty,0)}(x+W_t)]e^{t/k}}{\left(\mathbb{E}[\mathbb{1}_{(-\infty,0)}(x+W_t)]^k(e^t-1)+1\right)^{1/k}} \\
&\leq e^{t/k}\mathbb{P}(W_t \leq -x) \\
&= e^{t/k}\Phi\left(-\frac{x}{\sqrt{t}}\right).
\end{aligned}
$$

For the lower bound, we use Theorem 4.1.4 again with fixed n. Using Lemma 4.1.3, we see that

$$
\begin{aligned}
(P_{\frac{t}{n}} \circ D_{\frac{t}{n}})^{(1)}(\mathbb{1}_{(-\infty,0)})(x) &= \mathbb{P}(x + W_{t/n} \leq 0) \\
&= \Phi\left(-x\sqrt{\frac{n}{t}}\right).
\end{aligned}
$$

We can apply induction to the inequality

$$
(P_{\frac{t}{n}} \circ D_{\frac{t}{n}})^{(m)}(\mathbb{1}_{(-\infty,0)})(x) \geq a\Phi\left(-x\sqrt{\frac{n}{t}} - (m-1)\Phi^{-1}\left(be^{-t/nk}\right)\right)
$$

where $m \in \mathbb{Z}$ with $1 \le m \le n$. Since $a \in [0,1)$, this clearly holds for $m = 1$.

For the inductive step we use Lemma 4.1.5. In particular for $f : [0,1] \to [0,1]$, strictly decreasing, we have the bound

$$\begin{aligned}(P_{t/n} \circ D_{t/n})(af)(x) &\ge a\mathbb{P}\left(f(x + W_{t/n}) \ge be^{-t/nk}\right) \\ &= a\mathbb{P}\left(\sqrt{\tfrac{t}{n}}W_1 \le -x + f^{-1}(be^{-t/nk})\right)\end{aligned}$$

where $b = (1 - a^k)^{-1/k}$. □

Remark 11. *We can modify the above proof to obtain a lower bound when the initial condition is shifted and scaled. Suppose that $u_0(x) \ge c_1 \mathbb{1}_{(-\infty,-c_2)}$ where $c_1 \in (0,1]$ and $c_2 \in \mathbb{R}$. Suppose also that instead of evenly spaced time steps, we have an increasing sequence of fixed times $t_0 = 0 < t_1 < .... < t_j < ...t = t_n$ and if we set $s_j = t_j - t_{j-1}$, then we find that there exists T depending on n, k, a and c_1 such that for $t \ge T$ we have*

$$u(t,x) \ge a\Phi\left(\frac{1}{\sqrt{s_1}}\left(-x + c_2 - \sum_{j=2}^{n} \sqrt{s_j}\Phi^{-1}\left(be^{-\frac{s_{j-1}}{k}}\right)\right)\right). \tag{4.2.3}$$

◇

4.2.1. Wave Speed

A corollary of Theorem 4.2.1 is the classical result [KPP37] giving the speed of the wave front.

Theorem 4.2.2. *Let u be as in Theorem 4.2.1 and define $m(t)$ to be the median value. That is to say, $u(t, m(t)) = \frac{1}{2}$. Then,*

$$m(t) = \sqrt{\frac{2}{k}}t + o(t)$$

To prove Theorem 4.2.2, we use the following elementary lemma, the proof of which is given below.

Lemma 4.2.3. *The following holds:*

$$\Phi^{-1}(\varepsilon) \sim -\sqrt{2}\sqrt{-\log(\varepsilon)}$$

as $\varepsilon \searrow 0$.

Given this lemma, we can prove Theorem 4.2.2.

Proof of Theorem 4.2.2. Fix $t > 0$. Since u is decreasing, if x is such that $u(t,x) \le 1/2$, then $m(t) \le x$. Setting the right hand side of equation (4.2.2) equal to a half and rearrange for x. Then, we see that if $u(t,x) \le 1/2$ when

$$x = -\sqrt{t}\Phi^{-1}\left(\frac{1}{2e^{t/k}}\right).$$

So,

$$\frac{m(t)}{t} \le \sqrt{\frac{2}{k}}\sqrt{1 + \tfrac{k}{t}\log 2}(1 + o(1))$$

by Lemma 4.2.3 and so

$$\limsup_{t\to\infty}\frac{m(t)}{t}=\sqrt{\frac{2}{k}}$$

For the lower bound, we choose $a\in\left(\frac{1}{2},1\right)$. Similarly to above, if x is such that $u(t,x)\geq 1/2$, then $m(t)\geq x$. Setting the left hand side of equation (4.2.2) equal to a half and rearranging for x, we find that if

$$x\sqrt{\frac{n}{t}}=-\Phi^{-1}\left(\frac{1}{2a}\right)-(n-1)\Phi^{-1}\left(be^{-t/nk}\right),$$

then $u(t,x)\geq m(t)$. Thus,

$$\frac{m(t)}{t}\geq\sqrt{\frac{2}{k}}\frac{n-1}{n}(1+o(1))$$

by Lemma 4.2.3. Therefore,

$$\liminf_{t\to\infty}\frac{m(t)}{t}=\sqrt{\frac{2}{k}}\left(1-\frac{1}{n}\right)$$

for all n.

We have shown that,

$$\lim_{t\to\infty}\frac{m(t)}{t}=\sqrt{\frac{2}{k}}$$

as required. □

All that remains is the proof of Lemma 4.2.3.

Proof. [of Lemma 4.2.3] The result follows by the following elementary tail bounds on Φ:

$$1-\frac{1}{y}e^{-\frac{y^2}{2}}\leq\Phi(y)\leq 1-\frac{y^2-1}{y^3}e^{-\frac{y^2}{2}}$$

for $y>1$.

We set $z=\Phi(y)$ and so $y=\Phi^{-1}(z)$. If we take the logarithm of both sides of the inequality, we have

$$\log(y)+\frac{y^2}{2}\leq-\log(1-z)\leq\frac{y^2}{2}-\log(y^2-1)+\log(y^3)$$

For y large enough, we have for any η,

$$\frac{(y-\eta)^2}{2}\leq-\log(1-z)\leq\frac{(y+\eta)^2}{2}$$

and since $y=\Phi^{-1}(z)$, we have

$$-\eta+\sqrt{-2\log(1-z)}\leq\Phi^{-1}(z)\leq\sqrt{-2\log(1-z)}+\eta.$$

Since $y\to\pm\infty\iff z\to\frac{1}{2}\pm\frac{1}{2}$, we can rewrite this in the form that we wish by setting $z\mapsto 1-z$ and $y\mapsto -y$ and use $\Phi^{-1}(1-z)=-\Phi^{-1}(z)$ to obtain the form given in the Lemma. □

Remark 12. *We can see from the proof of Corollary 4.2.2, above, that we could take $m(t)$ to be defined so that $u(t,x)=\varepsilon$ where $\varepsilon\in(0,1)$ and the value of $m(t)$ is still $\sqrt{\frac{2}{k}}t+o(t)$. The estimate is not fine enough for this to affect it.*

Also, in general $m(t)$ can be multi-valued (in which case $m(t)=\{x:u(t,x)=1/2\}$) but this is not the case if u_0 is decreasing. ◇

4.2.2. General Initial Conditions

We can easily extend the above results to the case when u_0 is bounded below at $-\infty$ by a shifted and scaled indicator function. As we know from McKean [McK75], the wave front should still move at speed $\sqrt{\frac{2}{k}}$ for quickly exponentially decaying initial conditions.

Theorem 4.2.4. *Suppose that u solves equation (4.2.1) with initial condition, u_0 such that $\liminf_{x\to-\infty} u_0(x) > 0$ and $u_0(x) \leq 1 \wedge Ce^{-\beta x}$*

$$u(t, rt) \to 0 \text{ if } r > \tilde{r}$$

where

$$\tilde{r} = \begin{cases} \frac{\beta}{2} + \frac{1}{k\beta} & \text{if } \beta < \sqrt{\frac{2}{k}} \\ \sqrt{\frac{2}{k}} & \text{if } \beta \geq \sqrt{\frac{2}{k}} \end{cases}$$

and

$$u(t, rt) \to 1 \text{ if } r < \sqrt{\frac{2}{k}}.$$

Remark 13. *It is not clear how to recover the fact that $u(t, rt) \to 1$ when $\beta < \sqrt{\frac{2}{k}}$ (see McKean [McK75]) from Theorem 4.2.1.* ◇

Proof. Using Gaussian tail bounds we have

$$\begin{aligned} 0 \leq u(t, rt) &\leq e^{t/k}\mathbb{E}[u_0(X_t)] \\ &\leq e^{t/k}\Phi\left(-r\sqrt{t} + \frac{\log C}{\beta\sqrt{t}}\right) + Ce^{(\frac{1}{k}-\beta r+\frac{\beta^2}{2})t}\Phi\left((r-\beta)\sqrt{t} - \frac{\log C}{b\sqrt{t}}\right) \\ &\leq \begin{cases} e^{(\frac{1}{k}-\frac{r^2}{2})t}\cdot\mathcal{O}\left(\frac{1}{\sqrt{t}}\right) & \text{if } \beta > r \\ e^{(\frac{1}{k}-\beta r+\frac{\beta^2}{2})t}\cdot\mathcal{O}(1) & \text{if } \beta \leq r \end{cases} \\ &\to 0 \text{ as } t \to 0. \end{aligned}$$

since $\frac{1}{k} - \frac{r^2}{2} \leq \frac{1}{k} - \beta r + \frac{\beta^2}{2}$ with equality when $r = \beta$.

If $\beta > \sqrt{\frac{2}{k}}$, then since u is decreasing in its second component, we can assume without loss of generality that $\sqrt{\frac{2}{k}} < r < \beta$ and so in this case $u(t, rt) \to 0$ as $t \to \infty$.

If $\beta \leq \sqrt{\frac{2}{k}}$, then $r > \frac{\beta}{2} + \frac{1}{k\beta} \geq \beta$ and so $\frac{1}{k} - \beta r + \frac{\beta^2}{2} < 0$ and $u(t, rt) \to 0$.

The upper bound easily follows from Theorem 4.2.2 and the bound (4.2.3) since we can write $u_0 \geq c_1\mathbb{1}_{(-\infty,-c_2)}(x)$ for some constants $c_1 > 0, c_2 \in \mathbb{R}$. □

We present the following trivial corollary to refer back to after proving Theorem 4.3.10 below.

Corollary 4.2.5. *Consider the equation,*

$$v_t = \frac{\sigma^2}{2}v_{xx} + bv_x + \tfrac{1}{k}v(1 - v^k) \tag{4.2.4}$$

with $k > 0$ and $\sigma > 0$, and initial condition v_0 such that $\liminf_{x\to-\infty} v_0(x) > 0$ and $v_0(x) \leq 1 \wedge Ce^{-\beta x}$ where $\beta \geq \sqrt{\frac{2}{k}}$. Then, the position of the wave front is given by

$$m(t) = \left(\sqrt{\tfrac{2\sigma^2}{k}} - b\right)t + o(t).$$

Proof. Let u be the solution to equation (4.2.4), with initial condition u_0 as in the Theorem, in the case when $\sigma = 1$ and $b = 0$. Then, v defined by $v(t,x) = u(t, \sigma^{-1}(x + bt))$ solves equation (4.2.4) with initial condition $v_0 = u_0(\sigma^{-1}(x+bt))$. We see that $v(t, m(t)) = u(t, \sigma^{-1}(m(t)+bt)) = \frac{1}{2}$ and then we apply Theorem 4.3.10. $\square$

4.2.3. Initial Conditions with Compact Support

In this section we see how the above results can be extended to the case when u_0 has compact support. We present the proof separately since the proof does not extend directly to the case of more general $\mathcal{L}$. Note, however, that above results follow from the following theorem:

Theorem 4.2.6. *Let u solve equation* (4.2.1) *with u_0 such that $0 < u_0(x) \leq 1$ on some interval. Then,*

$$u(t, rt) \to 1 \text{ for } |r| < \sqrt{\tfrac{2}{k}}$$
$$u(t, rt) \to 0 \text{ for } |r| > \sqrt{\tfrac{2}{k}}.$$

Before we prove this result, we introduce some notation and a lemma that we will use during the proof.

Definition 4.2.7. *For, $c, \rho, n, t > 0$, we define $\Psi_c : \mathbb{R}_+ \to [0,1]$ by*

$$\Psi_c(\rho) = \mathbb{P}\left(W_1 \in \left[(-c+\rho)\sqrt{\tfrac{n}{t}}, (c+\rho)\sqrt{\tfrac{n}{t}}\right]\right). \tag{4.2.5}$$

Note that $\Psi_c : \mathbb{R}_+ \to [0,1]$ is strictly decreasing and hence invertible. However, since Ψ_c does not have an explicit inverse we bound it below by a function $f_c : \mathbb{R}_+ \to [0,1]$ that has an explicit inverse in the appropriate range that we will consider.

Lemma 4.2.8. *Let $l > 0$ be a constant and let $j = 0, 1, \cdots, n$. Set $c_{j+1} = l + \sqrt{\frac{2}{k}}\frac{jt}{n}(1+\alpha_t)$ where $\alpha_t \to 0$ as $t \to \infty$.*

For $\Psi_{c_{j+1}}$ defined as in equation (4.2.5)*, there exists a function $f_{c_{j+1}}$ such that for $\rho > 0$,*

$$\Psi_{c_{j+1}}(\rho) \geq f_{c_{j+1}}(\rho),$$

$f^{-1}_{c_{j+1}}(be^{-t/nk})$ exists, and

$$f^{-1}_{c_{j+1}}(be^{-t/nk}) = c_{j+1} + \frac{t}{n} \cdot \sqrt{\frac{2}{k}}(1 + o(1))$$

for t sufficiently large.

Proof. For t sufficiently large, there exists $\varepsilon \in (0,1)$ with

$$\begin{aligned}
\Psi_c(\rho) &= \frac{1}{\sqrt{2\pi}} \int_{(\rho-c)\sqrt{\frac{n}{t}}}^{(c+\rho)\sqrt{\frac{n}{t}}} e^{-y^2/2} \mathrm{d}y \\
&\geq \frac{1}{\sqrt{2\pi}} \mathbb{1}_{\{\rho > c\}} \left[\left(\frac{1}{(\rho-c)\sqrt{\frac{n}{t}}} - \frac{1}{\left((\rho-c)\sqrt{\frac{n}{t}}\right)^3}\right) e^{-\frac{(\rho-c)^2 n}{2t}} - \frac{1}{(\rho+c)\sqrt{\frac{n}{t}}} e^{-\frac{(\rho+c)^2 n}{2t}}\right] \\
&\geq \frac{e^{-\frac{(\rho-c)^2 n}{2t}}}{\sqrt{2\pi}} \mathbb{1}_{\left\{(\rho-c)\sqrt{\frac{n}{t}} > \sqrt{\frac{1}{1-\varepsilon}}\right\}} \left[\frac{\varepsilon}{(\rho-c)\sqrt{\frac{n}{t}}} - \frac{e^{\frac{-2\rho c n}{t}}}{(\rho+c)\sqrt{\frac{n}{t}}}\right]
\end{aligned}$$

using the standard Gaussian tail bounds.

The behaviour of this bound changes slightly depending on whether $c = c_j$ is constant (the case when $j = 0$) or linear in t (the case when $j > 0$). Therefore, we split the proof into these two cases.

Case 1. $j = 0$

Firstly, consider the case when $c = c_1 = l$. Here,

$$\begin{aligned}
\Psi_l(\rho) &\geq \frac{e^{-\frac{(\rho-l)^2 n}{2t}}}{\sqrt{2\pi}} \mathbb{1}_{\left\{(\rho-l)\sqrt{\frac{n}{t}} > \sqrt{\frac{1}{1-\varepsilon}}\right\}} \frac{1}{(\rho-l)\sqrt{\frac{n}{t}}} \left[\varepsilon - e^{\frac{-2\rho l n}{t}}\right] \\
&\geq e^{-\frac{(\rho-l)^2 n}{2t}} \cdot \beta t^{-1} \mathbb{1}_A \\
&\equiv f_l(\rho)
\end{aligned}$$

where

$$A = \left\{(\rho - l)\sqrt{\tfrac{n}{t}} > \sqrt{\tfrac{1}{1-\varepsilon}}\right\} \cap \left\{(\rho - l)\sqrt{\tfrac{n}{t}} \leq \gamma t\right\} \cap \left\{\rho l > -\tfrac{1}{2}\tfrac{t}{n}\log(\varepsilon\delta)\right\}$$

and $\beta = \frac{(1-\delta)\varepsilon}{\gamma\sqrt{2\pi}}$ for some constant $\delta \in (0, 1)$ to be chosen later and $\gamma > 0$.

Let y be such that $f_l(\rho) = y$ and $y > 0$. Then, $\rho = f_l^{-1}(y)$ is well defined. Then,

$$f_l^{-1}(y) = l + \sqrt{\frac{2t}{n}\left(-\log y - \log\frac{t}{\beta}\right)}.$$

If $y = be^{-t/nk}$, then

$$\rho = l + \frac{t}{n} \cdot \sqrt{\frac{2}{k}\left(1 - \frac{kn}{t}\log\frac{t}{b\beta}\right)}$$

since $\rho \in A$ if t is large enough and we choose ε and δ such that $e^{-2l\sqrt{\frac{2}{k}}} < \varepsilon\delta < 1$.

Case 2. $j \geq 1$.

For $j \geq 0$, $c_{j+1} = \mathcal{O}(t)$ and is close in size to ρ and so we adapt the above calculation. We use the following:

$$\begin{aligned}
\Psi_c(\rho) &\geq \frac{e^{-\frac{(\rho-c)^2 n}{2t}}}{\sqrt{2\pi}} \mathbb{1}_{\left\{(\rho-c)\sqrt{\frac{n}{t}} > \sqrt{\frac{1}{1-\varepsilon}}\right\}} \frac{1}{(\rho-c)\sqrt{\frac{n}{t}}} \left(\varepsilon - \frac{\rho-c}{\rho+c}\right) \\
&\geq e^{-\frac{(\rho-c)^2 n}{2t}} \cdot \beta t^{-1} \mathbb{1}_B \\
&\equiv f_c(\rho)
\end{aligned}$$

where

$$B = \left\{(\rho - c)\sqrt{\tfrac{n}{t}} > \sqrt{\tfrac{1}{1-\varepsilon}}\right\} \cap \left\{(\rho - c)\sqrt{\tfrac{n}{t}} \leq \gamma t\right\} \cap \left\{\rho \leq \tfrac{1+\varepsilon\delta}{1-\varepsilon\delta}c\right\}$$

and $\beta = \frac{(1-\delta)\varepsilon}{\gamma\sqrt{2\pi}}$ for some constant $\delta \in (0, 1)$ to be chosen later and $\gamma > 0$.

As with above, let y be such that $f_{c_{j+1}}(\rho) = y$ and $y > 0$. Then, $\rho = f_{c_{j+1}}^{-1}(y)$ is well defined. If $y = be^{-t/nk}$, then

$$\rho = c_{j+1} + \frac{t}{n} \cdot \sqrt{\frac{2}{k}}(1 + o(1)).$$

We check that $\rho \in B$ for $j \geq 1$. We have $\rho \leq c_{j+1}(1 + 1/j)(1 + o(1))$ and so we pick ε and δ so that $\frac{1}{1+2j} < \varepsilon\delta < 1$ for $j \geq 1$. □

Proof of Theorem 4.2.6. For simplicity, let $u_0 = \mathbb{1}_{[-l,l]}$; otherwise bound u_0 above and below by a shifted and scaled indicator function. Let $V_{t/n} = P_{t/n} \circ D_{t/n}$. Then,

$$\begin{aligned} V_{t/n}(u_0)(x) &\geq \mathbb{P}\left(W_1 \in \left[(-l-x)\sqrt{\tfrac{n}{t}}, (l-x)\sqrt{\tfrac{n}{t}}\right]\right) \\ &\equiv \Psi_{c_1}(|x|) \end{aligned}$$

where we set $c_1 = l$ and Ψ defined in equation (4.2.5). c_j for $j > 0$ is defined below and we see that it is of the form specified in Lemma 4.2.8 and so the notation is not ambiguous.

We define the sequence $(c_j)_{j\geq 0}$ by induction with $c_1 = l$: abusing notation we have,

$$\begin{aligned} (V_{t/n})(af_{c_j}(|x|)) &\geq a\mathbb{P}\left(|x + W_{t/n}| \leq f_{c_j}^{-1}(be^{-t/nk})\right) \\ &\geq af_{c_{j+1}}(|x|) \end{aligned}$$

where we define

$$c_{j+1} = f_{c_j}^{-1}\left(be^{-t/nk}\right).$$

By Lemma 4.2.8, we bound $\Psi_{c_{j+1}}(\rho) \geq f_{c_{j+1}}(\rho)$ for all $j = 0, 1, \cdots, n$ and appropriate ρ. Therefore, by Theorem 4.1.4, we have

$$u(t, x) \geq af_{c_n}(|x|).$$

The Wave Position

Set $|x| = rt$ where $0 < r < \sqrt{\frac{2}{k}}$ and set $0 < \tilde{\varepsilon} < \frac{1}{2}\left(\sqrt{\frac{2}{k}} - r\right)$. Choose $n > \frac{2}{\sqrt{\frac{2}{k}-r}}$. Then, $|x| - c_n < -\tilde{\varepsilon}t$ for t sufficiently large and so

$$\begin{aligned} u(t, rt) &\geq af_{c_n}(rt) \\ &\geq a\mathbb{P}(W_{t/n} \in [-\tilde{\varepsilon}t, rt]) \end{aligned}$$

for all $a \in (0, 1)$ and so

$$u(t, rt) \to 1 \text{ for } |r| < \sqrt{\tfrac{2}{k}}.$$

The fact that

$$u(t, rt) \to 0 \text{ for } |r| > \sqrt{\tfrac{2}{k}}$$

follows from previous calculations: considering and $\mathbb{1}_{[-l,l]} \leq \mathbb{1}_{[-l,\infty)}$ and $\mathbb{1}_{[-l,l]} \leq \mathbb{1}_{(\infty,l]}$ we see that

$$\begin{aligned} u(t, rt) &\leq e^{t/k}\Phi\left(\frac{l}{\sqrt{t}} - |r|\sqrt{t}\right) \\ &\to 0 \text{ if } |r| > \sqrt{\frac{2}{k}}. \end{aligned}$$

□

4.3. Speed of the wave front

Using similar techniques to Section 4.2 we can find the speed of the wave front in more general situations; this is the content of the main result of this section – Theorem 4.3.2 (and its corollary 4.3.10).

Let $k > 0$ and let u solve

$$\begin{aligned} u_t &= \mathcal{L}u + \tfrac{1}{k}u(1-u^k), \text{ in } (0,T)\times\mathbb{R} \\ u_0(x) &= \mathbb{1}_{(-\infty,0)}(x), \text{ for } x \in \mathbb{R}. \end{aligned} \tag{4.3.1}$$

in the sense of Definition 2.1.1, where $(X_t)_{t\geq 0}$ is a Lévy process with stochastic generator given by $\mathcal{L}$.

If the moment generating function of the process $-X$ exists and is given by Λ i.e.

$$\mathbb{E}[e^{-\theta X_s}] = e^{s\Lambda(\theta)},$$

then we can write Λ in the Lévy-Khintchine form as

$$\Lambda(\theta) = -b\theta + \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R}\setminus\{0\}} (e^{-\theta y} - 1 + \theta y \mathbb{1}_{|y|>1})\nu(\mathrm{d}y)$$

for $b, \sigma \in \mathbb{R}$ and a Borel measure ν on $\mathbb{R}\setminus\{0\}$ such that $\int_{\mathbb{R}\setminus\{0\}}(y^2 \wedge 1)\mathrm{d}\nu(y) < \infty$.

We will make the following assumption throughout Section 4.3.

Assumption 4.3.1.

Suppose that the moment generating function, Λ, of $-X$ exists for θ in a neighbourhood of 0; that is,

$$\mathbb{E}[e^{-\theta X_s}] = e^{s\Lambda(\theta)}.$$

Equivalently, we assume that

$$\int e^{-\theta y}\mathbb{1}_{\{|y|>1\}}\nu(\mathrm{d}y) < \infty$$

for all θ in a neighbourhood of 0.

Let q be defined by

$$q = \inf_{\theta > 0} \frac{1/k + \Lambda(\theta)}{\theta}. \tag{4.3.2}$$

We will show that q is the speed of the wave front for suitable choices of $\mathcal{L}$.

Remark 14. *For $x \in \mathbb{R}$, we define Λ^* to be the Legendre-Fenchel dual of Λ:*

$$\Lambda^*(x) = \sup_{\theta \geq 0}(\theta x - \Lambda(\theta)).$$

If the supremum is attained at $\theta = \theta^*$, then q is simply the unique number such that

$$\Lambda^*(q) = \frac{1}{k}.$$

◇

We will prove the following analogue to Theorem 4.2.2 in Section 4.3.2.

Theorem 4.3.2. *Suppose that Assumption 4.3.1 holds and that u solves equation* (4.3.1). *Let q be defined as in equation* (4.3.2). *Then,*

$$u(t, rt) \to 0, \text{ for } r > q,$$
$$u(t, rt) \to 1, \text{ for } r < q.$$

A similar result was shown independently by [GJ16] based on the ideas of [Kyp99] using branching processes.

Remark 15. *We can also see Theorem 4.3.2, in terms of the position of the wave front*

$$m(t) = \sup\left\{x : u(t, x) \geq \tfrac{1}{2}\right\}.$$

Then,

$$m(t) = qt + o(t), \text{ as } t \to \infty. \qquad \diamond$$

4.3.1. Preliminaries

Lemma 4.3.3 (Properties of Λ). *Let Λ and Λ^* be defined as above. Then, the following statements hold.*

1. *Λ and Λ^* are both convex.*
2. *$\Lambda^*(-\mathbb{E}[X_1]) = 0$.*
3. *Λ^* is strictly increasing on $[-\mathbb{E}[X_1], \infty)$.*

Proof. 1. These properties are simple consequences of the definition. Firstly, we show that Λ is convex. Let $\lambda \in (0, 1)$ and $\theta_1, \theta_2 \in \mathbb{R}$. Then, by Hölder's inequality,

$$\begin{aligned}\Lambda(\lambda\theta_1 + (1-\lambda)\theta_2) &= \log \mathbb{E}\left[(e^{-\theta_1 X_1})^{\lambda}(e^{-\theta_2 X_1})^{1-\lambda}\right] \\ &\leq \log \mathbb{E}\left[(e^{-\theta_1 X_1})\right]^{\lambda} + \log \mathbb{E}\left[(e^{-\theta_2 X_1})\right]^{1-\lambda} \\ &= \lambda\Lambda(\theta_1) + (1-\lambda)\Lambda(\theta_2)\end{aligned}$$

as required.

Next, we show that Λ^* is convex. Let $\lambda \in (0, 1)$ and $x_1, x_2 \in \mathbb{R}$.

$$\begin{aligned}\Lambda^*(\lambda x_1 + (1-\lambda)x_2) &= \sup_{\theta>0}\left[(\lambda x_1)\theta + (1-\lambda)x_2\theta - (\lambda + 1 - \lambda)\Lambda(\theta)\right] \\ &\leq \sup_{\theta>0}\left[\lambda(x_1\theta - \Lambda(\theta))\right] + \sup_{\theta>0}\left[(1-\lambda)(x_2\theta - \Lambda(\theta))\right] \\ &= \lambda\Lambda^*(x_1) + (1-\lambda)\Lambda^*(x_2)\end{aligned}$$

as required.

2. By Jensen's inequality, we see that $\Lambda(\theta) \geq -\theta\mathbb{E}[X_1]$. Thus, we see that $\Lambda^*(x) \leq \sup(\theta(x+\mathbb{E}[X_1]))$ and so $0 \leq \Lambda^*(-\mathbb{E}[X_1]) \leq 0$.

3. Pick $\theta_0 > 0$ so that Λ is defined for $|\theta| < \theta_0$. Take $y > \Lambda'(0) = -\mathbb{E}[X_1]$ and find $\varepsilon \in (0, \theta_0)$ such that $y > \Lambda'(\varepsilon)$. By convexity,

$$\begin{aligned}\Lambda(\varepsilon) - \Lambda(\theta) &\leq \Lambda'(\varepsilon)(\varepsilon - \theta) \\ &< y(\varepsilon - \theta)\end{aligned}$$

for $\theta < \varepsilon$ and so for any $z > y > -\mathbb{E}[X_1]$,

$$\begin{aligned}\Lambda^*(y) &= \sup_{\theta > \varepsilon}(y\theta - \Lambda(\theta)) \\ &\leq -\varepsilon(z-y) + \sup_{\theta>\varepsilon}(z\theta - \Lambda(\theta)) \\ &< \Lambda^*(z)\end{aligned}$$

□

In order to show Theorem 4.3.2, we will need some preliminary results. First of all we introduce some notation.

We write

$$F_s(y) = \mathbb{P}(X_s \leq y)$$

for the distribution function of X_s and we write

$$F_s^{-1}(y) \equiv \inf\{z : F_s(z) \geq y\}$$

for the generalised, left-continuous inverse. This is sometimes referred to as the quantile function. We use the convention that $\inf \varnothing = \infty$. In the case, that F_s is invertible, the generalised inverse coincides with the usual inverse.

We will use the following observation, below.

Lemma 4.3.4. *Let F_s and F_s^{-1} be defined as above. Let $z \in [0, 1]$. Then,*

$$F_s(y) \geq z \iff y \geq F_s^{-1}(z)$$

The proof is elementary and can be found in [EH13], for example.

We can now give the bounds on u that will allow us to prove the Theorem 4.3.2. The following result is analogous to Theorem 4.2.1.

Theorem 4.3.5. *Fix t,x and n and $a \in (1/2, 1)$ and set $b = (1 - a^k)^{-1/k}$. Let u solve equation (4.3.1). Then,*

$$aF_{\frac{t}{n}}\left(-x-(n-1)F_{\frac{t}{n}}^{-1}\left(be^{-t/nk}\right)\right) \leq u(t,x) \leq e^{t/k}F_t(-x) \tag{4.3.3}$$

Proof. The upper bound follows from Theorem 4.1.4 with $n = 1$:

$$\begin{aligned}(D_t \circ P_t)(\mathbb{1}_{(-\infty,0)})(x) &= \frac{\mathbb{E}^x[\mathbb{1}_{(-\infty,0)}(X_t)]e^{t/k}}{\left(\mathbb{E}^x[\mathbb{1}_{(-\infty,0)}(X_t)]^k(e^t-1)+1\right)^{1/k}} \\ &\le e^{t/k}\mathbb{P}^x(X_t \le 0) \\ &= e^{t/k}F_t(-x).\end{aligned}$$

For the lower bound, we use Theorem 4.1.4 again with fixed n. By definition, or Lemma 4.1.3, we see that

$$(P_{\frac{t}{n}} \circ D_{\frac{t}{n}})^{(1)}(\mathbb{1}_{(-\infty,0)})(x) = F_{t/n}(-x)\,.$$

We can apply induction to the inequality

$$(P_{\frac{t}{n}} \circ D_{\frac{t}{n}})^{(m)}(\mathbb{1}_{(-\infty,0)})(x) \ge aF_{t/n}\left(-x-(m-1)F_{t/n}^{-1}\left(be^{-t/nk}\right)\right)$$

where $m \in \mathbb{Z}$ with $1 \le m \le n$. Since $a \in (1/2, 1)$, this clearly holds for $m = 1$.

For the inductive step we use Lemma 4.1.5. In particular for $f : [0,1] \to [0,1]$, we have the bound

$$(P_{t/n} \circ D_{t/n})(f)(x) \ge a\mathbb{P}\left(f(x+W_{t/n}) \ge abe^{-t/nk}\right)$$

where $b = (1-a^k)^{-1/k}$.

$$\begin{aligned}(P_{\frac{t}{n}} \circ D_{\frac{t}{n}})^{(m+1)}(\mathbb{1}_{(-\infty,0)})(x) &\ge a\mathbb{P}\left(F_{t/n}\left(-X_{t/n}-x-(m-1)F_{t/n}^{-1}\left(be^{-t/nk}\right)\right) \ge be^{-t/nk}\right) \\ &= a\mathbb{P}\left(-X_{t/n} \ge x+mF_{t/n}^{-1}\left(be^{-t/nk}\right)\right)\end{aligned}$$

for $0 \le m \le n-1$. In the second line we used Lemma 4.3.4 if $be^{-t/nk} \le 1$ and if $be^{-t/nk} > 1$ the result clearly holds. The lower bound in (4.3.3) follows. □

Remark 16. *By slightly changing the proof of Theorem 4.3.5 we can obtain the following more general bound. Suppose that $u_0(x) \ge c_1\mathbb{1}_{(-\infty,-c_2)}$ where $c_1 \in (0,1]$ and $c_2 \in \mathbb{R}$. For an increasing sequence of times $t_0 = 0 < t_1 < .... < t_j < ...t = t_n$, we set $s_j = t_j - t_{j-1}$. Then, the following bound holds*

$$u(t,x) \ge aF_{s_1}\left(-x+c_2-\sum_{j=2}^{n}F_{s_j}^{-1}\left(be^{-s_{j-1}/k}\right)\right) \tag{4.3.4}$$

for all $x \in \mathbb{R}$ and $t \ge T$ where T depends on n, k, a and c_1.

Furthermore, if u_0 is left continuous and decreasing, we can define a generalised inverse by

$$u_0^{-1}(z) = \sup\{y : u(y) \le z\}$$

and then, by following the proof of Lemma 4.3.4 and Theorem 4.3.5, mutatis mutandis, we see that

$$u(t,x) \ge aF_{s_1}\left(-x+u_0^{-1}\left(be^{-s_n/k}\right)-\sum_{j=2}^{n}F_{s_j}^{-1}\left(be^{-s_{j-1}/k}\right)\right). \tag{4.3.5}$$

◇

In the case of a general Lévy process, we will need a substitute for the explicit asymptotics in Lemma 4.2.3. For this, we can use Crámer's Theorem from the theory of Large Deviations (see, for example, Grimmett and Stirzaker [GS01]).

Theorem 4.3.6 (Crámer's Theorem)**.** *Let Y be a Lévy process with a finite cumulant generating function given by Λ. Let Λ^* be the Legendre-Fenchel dual function. Then, for any $\alpha > \mathbb{E}[Y_1]$,*

$$\mathbb{P}(Y_t \geq \alpha t) \geq e^{-t\Lambda^*(\alpha) - C(t)}$$

where $C(t)/t \to 0$ as $t \to \infty$. If $\mathbb{P}(X_1 > \alpha) > 0$, then $\Lambda^*(\alpha) > 0$.

Lemma 4.3.7.

$$-\mathbb{E}[X_1] \leq q$$

The inequality is strict if X is not deterministic.

Proof. By Jensen's Inequality,

$$\begin{aligned} q &= \inf_{\theta > 0} \left[\frac{1}{k\theta} + \frac{1}{\theta} \log \mathbb{E}[e^{-\theta X_1}] \right] \\ &\geq \inf_{\theta > 0} \left[\frac{1}{k\theta} - \mathbb{E}[X_1] \right] \\ &= -\mathbb{E}[X_1] \end{aligned}$$

with equality if and only if X is deterministic, since log is strictly concave. □

We treat the degenerate case, when $X_s = bs$ separately here.

Lemma 4.3.8. *Suppose that X is a deterministic Lévy process, i.e. for all s, $X_s = bs$ for some constant b. Suppose that u solves equation* (4.3.1) *with $u_0 = \mathbb{1}_{(-\infty,0)}$. Then,*

$$\frac{m(t)}{t} = -b.$$

In fact, u is given explicitly by $u(t,x) = \mathbb{1}(x + bt < 0)$.

Proof. We just check the proposed solution is actually a solution:

$$\begin{aligned} M_s^* &= \mathbb{1}(x + bt < 0) e^{\frac{s}{k}(1 - \mathbb{1}(x+bt<0))} \\ &= \mathbb{1}(x + bt < 0) \end{aligned}$$

for all $s \geq 0$ and so is trivially a martingale. □

4.3.2. Proof of Theorem 4.3.2

Now we can prove Theorem 4.3.2.

Proof. By Lemma 4.3.8 and the remark above, we may assume that X is random with $\mathbb{E}[X_1] = 0$ and so Λ is strictly convex and the inequalities in Lemma 4.3.7 are strict.

By Theorem 4.2.1 and Markov's inequality, we have for $\theta > 0$,

$$\begin{aligned} u(t,x) &\leq e^{t/k}\mathbb{E}[\mathbb{1}(e^{-\theta X_t} \geq e^{\theta x})] \\ &\leq \exp\left(\frac{t}{k} + t\Lambda(\theta) - \theta x\right) \end{aligned}$$

We minimise over $\theta > 0$. We set $x = rt$ for $r > q$ where q is given by (4.3.2). Then $\Lambda^*(r) > \Lambda^*(q)$ by Lemma 4.3.3. Thus,

$$\limsup_{t\to\infty} u(t,rt) = 0,$$

and $u(t,x) \geq 0$, for all (t,x). Therefore,

$$u(t,rt) \to 0, \text{ as } t \to \infty.$$

Next, we show that

$$u(t,rt) \to 1, \text{ as } t \to \infty$$

for all $r < q$.

We use the bound

$$u(t,rt) \geq aF_{\frac{t}{n}}\left(-rt - (n-1)F^{-1}_{\frac{t}{n}}\left(be^{-t/nk}\right)\right)$$

We define a new constant $\overline{r} \in (r,q)$. By Lemma 4.3.4,

$$F^{-1}_{\frac{t}{n}}\left(be^{-t/nk}\right) \leq -\frac{\overline{r}t}{n} \iff be^{-t/nk} \leq F_{\frac{t}{n}}\left(\frac{-\overline{r}t}{n}\right)$$

Since $u(t,x)$ is increasing in x, we can assume without loss of generality that $q > \overline{r} > r > -\mathbb{E}[X_1]$ by Lemma 4.3.7. Thus, we can use Crámer's Theorem (Theorem 4.3.6), to see that there exists a $T_1 > 0$ such that for all $t > T_1$,

$$F_{\frac{t}{n}}\left(\frac{\overline{r}t}{n}\right) \geq be^{-t/nk}$$

Then, there exists $T_1 > 0$ such that for $t > T_1$,

$$\begin{aligned} \log F_{\frac{t}{n}}\left(-\frac{\overline{r}t}{n}\right) &= \log \mathbb{P}\left(-X_{t/n} \geq \overline{r}\tfrac{t}{n}\right) \\ &\geq -\frac{t}{n}\Lambda^*(\overline{r}) - o(t) \\ &\geq -\frac{t}{n}\Lambda^*(q) + \log b \\ &\geq -\frac{t}{nk} + \log b. \end{aligned}$$

To go from the second to third line we use the fact that Λ^* is strictly increasing by Lemma 4.3.3. The last line follows from the fact that $\Lambda^*(q) \leq \frac{1}{k}$.

Now, we fix

$$n \geq \frac{\overline{r} + 1 + \mathbb{E}[X_1]}{\overline{r} - r}.$$

By the weak law of large numbers, for any $\varepsilon > 0$, there exists $T_2 > 0$ such that for $t > T_1 \vee T_2$, we have

$$\begin{aligned} u(t, rt) &\geq aF_{\frac{t}{n}}\left((-rn + \overline{r}(n-1))\frac{t}{n}\right) \\ &\geq a\mathbb{P}\left(X_{t/n} \leq (1 + \mathbb{E}[X_1])\frac{t}{n}\right) \\ &\geq a(1-\varepsilon) \end{aligned}$$

In other words,

$$\liminf_{t\to\infty} u(t, rt) \geq a, \text{ for all } r > q.$$

Since $a \in (0,1)$ was arbitrary and $u(t,x) \leq 1$ for all (t,x), we have shown,

$$u(t, rt) \to 1, \text{ as } t \to \infty$$

for all $r < q$, as required. □

4.3.3. General Initial Conditions

We can generalise Theorem 4.3.2 for initial conditions with fast enough decay and are bounded below by a shift of the Heaviside function. These assumptions are made rigorous in Hypothesis 4.3.9 below.

Firstly, for $\theta > 0$, we will write

$$J(\theta) = \frac{1/k + \Lambda(\theta)}{\theta}. \tag{4.3.6}$$

Hypothesis 4.3.9. *1. Suppose that there exists θ^* with $q = J(\theta^*)$.*

2. Suppose that the initial condition u_0 taking values in $[0,1]$ satisfies the following assumptions:

$$\liminf_{x\to-\infty} u_0(x) > 0$$

and

$$u_0(x) \leq 1 \wedge Ce^{-\beta x}$$

for some $C > 0$ and $\beta > \theta^*$.

We have the following corollary to Theorem 4.3.2.

Theorem 4.3.10. *Suppose that u solves equation (4.3.1) with initial condition u_0 satisfying Hypothesis 4.3.9. Then,*

$$u(t, rt) \to 0 \text{ if } r > q$$
$$u(t, rt) \to 1 \text{ if } r < q$$

Remark 17. *Recall in Lemma 4.3.8, we have $\theta^* = \infty$ and so Theorem 4.3.10 does not apply in the case when $X_s = bs$ for a constant b. If we look for solutions of the form $u(t,x) = u_0(x + bt)$, we see that we must have $u_0(x)(1 - u_0(x)^k) = 0$ for all x and this can not be true if u_0 takes values outside of $\{0,1\}$.* ◇

Proof. By Lemma 4.3.8, we can assume that Λ^* is strictly convex. First of all, let's suppose that $r > q$

$$\begin{aligned} 0 \le u(t, rt) &\le e^{t/k}\mathbb{E}[u_0(X_t)] \\ &\le e^{t/k}\mathbb{P}\left(X_t \le -x + \frac{1}{\beta}\log C\right) + Ce^{t/k}\mathbb{E}\left[e^{-\beta(x+X_t)}\mathbb{1}\left(x + X_t > \frac{1}{\beta}\log C\right)\right] \\ &\equiv (A) + (B) \end{aligned}$$

where (A) and (B) refer to the first and second term respectively. By Markov's inequality,

$$(A) \le C^{\theta/\beta}e^{t\left(\frac{1}{k}-(\theta r - \Lambda(\theta)\right)}$$

for all $\theta > 0$ and so setting $\theta = \theta^*$,

$$(A) \le C^{\theta^*/\beta}e^{t\left(\frac{1}{k}-\Lambda^*(r)\right)} \to 0$$

as $\Lambda^*(q) = \frac{1}{k}$ and $r > q$. Similarly,

$$(B) \le C^{-\theta/\beta}e^{t\left(\frac{1}{k}-(\beta-\theta)r+\Lambda(\theta-\beta)\right)}$$

. Setting $\theta = \beta - \theta^* > 0$, we see that

$$(B) \le C^{\theta^*/\beta - 1}e^{t\left(\frac{1}{k}-\Lambda^*(r)\right)}.$$

Since $\frac{1}{k} - \Lambda^*(r) < \frac{1}{k} - \Lambda^*(q) = 0$, we see that $u(t, rt) \to 0$ when $r > q$.

The case when $r < q$ follows since there exist constants $c_1 > 0, c_2 \in \mathbb{R}$ such that $u_0(x) \ge c_1\mathbb{1}_{(-\infty,-c_2)}(x)$ and so the Theorem 4.3.2 can be adapted using the bound in (4.3.4). □

Finally, we see how Theorem 4.2.4 can be used to show that if u_0 decays more slowly than exponentially, the wave front moves faster than linearly.

Proposition 4.3.11. *Suppose that u solves equation* (4.3.1) *with initial condition u_0 such that for all $\varepsilon > 0$ and x sufficiently large, $u_0(x) \ge e^{-\varepsilon x}$. Then,*

$$\liminf_{t\to\infty} \frac{m(t)}{t} = \infty.$$

Proof. Let $\alpha = \mathbb{E}[X_1]$. By the bound (4.3.5), we see that for any $r \in \mathbb{R}$,

$$\begin{aligned} u(t, rt) &\ge aF_{t/n}\left(-rt + u_0^{-1}\left(be^{-t/nk}\right) - (n-1)F_{t/n}^{-1}\left(be^{-t/nk}\right)\right) \\ &\ge aF_{t/n}\left(\frac{t}{n}\left(-rn + \frac{1}{\varepsilon k} - \alpha n\right) - \frac{1}{\varepsilon}\log b\right) \end{aligned}$$

where we used the fact that

$$(n-1)F_{t/n}^{-1}(be^{-t/nk}) \le \alpha t \iff be^{-t/nk} \le F_{t/n}\left(\tfrac{\alpha t}{n-1}\right)$$

since

$$\begin{aligned} F_{t/n}\left(\tfrac{\alpha t}{n-1}\right) &= \mathbb{P}\left(\tfrac{n}{t}X_{t/n} \le \alpha\left(1 + \tfrac{1}{n-1}\right)\right) \\ &\ge be^{-t/nk} \end{aligned}$$

for n fixed and t sufficiently large, by the weak law of large numbers.

We also used that

$$u_0(x) \geq e^{-\varepsilon x} \iff u_0^{-1}(y) \geq -\tfrac{1}{\varepsilon}\log y$$

for $x \in (a, \infty)$ for some a and $y \in (0, u_0(a))$.

Therefore, for t sufficiently large,

$$u(t, rt) \geq a\mathbb{P}\left(\tfrac{n}{t}X_{t/n} \leq (1 + \mathbb{E}[X_1])\right)$$

for all $\varepsilon < [k(1 + (n+1)\mathbb{E}[X_1] + r)]^{-1}$. Thus, $\liminf_{t\to\infty} u(t, rt) \geq a(1-\delta)$, for all $a \in (0,1)$, $\delta > 0$ and so $\liminf_{t\to\infty} u(t, rt) = 1$ and the result follows. □

This type of initial condition was studied in detail in Hamel and Roques [HR10].

4.4. Examples

We give some examples of Theorem 4.3.10 for specific Lévy processes X.

Firstly, we consider the following well-known result.

Example 4.4.1. *Let us consider the case when $\mathcal{L} = \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2} + b\frac{\partial}{\partial x}$ which corresponds to a the Markov process X defined by $X_t = bt + \sigma W_t$. Then $\Lambda(\theta) = -b\theta + \frac{\sigma^2}{2}\theta^2$, and so $\theta^* = \sqrt{\frac{2}{k\sigma^2}}$. Thus, if Hypothesis 4.3.9 holds for some $\beta > \sqrt{\frac{2}{k\sigma^2}}$. Then*

$$\begin{aligned}\frac{m(t)}{t} &\to \inf_{\theta>0}\frac{\frac{1}{k} - b\theta + \frac{1}{2}\theta^2}{\theta}\\ &= \sqrt{\frac{2\sigma^2}{k}} - b\end{aligned}$$

As we saw in Corollary 4.2.5, the result also holds with $\beta = \sqrt{\frac{2\sigma^2}{k}}$.

The next simplest example is with a Brownian motion with drift plus a jump process.

Example 4.4.2. *Set $X_t = x + bt + \sigma W_t + \sum_{i=1}^n P_t^i$, where P^i are Poisson processes with jumps of size y_i with intensity $\lambda_i > 0$. Then the Lévy measure is given by $\nu(A) = \sum \lambda_i\delta_{y_i}(A)$ for a measurable set $A \in \mathbb{R}$. So, the equation is*

$$u_t = \frac{\sigma^2}{2}u_{xx} + \beta u_x + \sum_{i=1}^n \lambda_i u(\cdot, \cdot + y_i) + \left(\frac{1}{k} - \sum_{i=1}^n \lambda_i\right)u - \frac{u^{1+k}}{k} \tag{4.4.1}$$

where u_0 satisfies Hypothesis 4.3.9 and $\beta = b - \sum_{i=1}^n \lambda_i \mathbb{1}_{|y_i|\leq 1}y_i$ and

$$\Lambda(\theta) = \left(\sum_{i=1}^n \lambda_i \mathbb{1}_{|y_i|\leq 1}y_i - b\right)\theta + \frac{\sigma^2}{2}\theta^2 + \sum_{i=1}^n \lambda_i\left(e^{-\theta y_i} - 1\right)$$

By Theorem 4.3.2, we know that Equation (4.4.1) has a wave front moving at speed $qt + o(t)$ where

$$q = \inf_{\theta>0}\frac{1/k + \Lambda(\theta)}{\theta}.$$

Here,

$$q = \inf_{\theta>0}\left[\sum_{i=1}^{n}\lambda_i \mathbb{1}_{|y_i|\le 1} y_i - b + \frac{\sigma^2}{2}\theta + \sum_{i=1}^{n}\lambda_i \frac{e^{-\theta y_i}-1}{\theta} + \frac{1}{k\theta}\right].$$

We see that J is differentiable and strictly convex with $J(0+) = J(\infty) = \infty$ and so there is a stationary point.

Even in very simple situations, the value of θ^* is not likely to be explicit. Consider the equation

$$u_t = \frac{1}{2}u_{xx} + u(\cdot,\cdot+1) - u^2. \tag{4.4.2}$$

Then θ^* is the root of $\frac{1}{2}\theta^2 e^{\theta} = 1+\theta$. Thus, $\theta^* \approx 1.163$ and $q \approx 1.581$.

Example 4.4.3. *Set $X_t = x + bt + \sigma W_t + \sum_{i=1}^{N(t)} Z^i$, where Z^i are IID random variables with law μ and $N(t)$ is a Poisson process with mean λt.*

Then the Lévy measure is given by $\nu(A) = \lambda\mu(A)$ for a measurable set $A \in \mathbb{R}$. So, the equation is

$$u_t = \frac{\sigma^2}{2}u_{xx} + bu_x + \lambda\int_{\mathbb{R}} \left(u(\cdot,\cdot+y) - u(\cdot,\cdot)\right)\mu(\mathrm{d}y) + \frac{1}{k}u(1-u^k) \tag{4.4.3}$$

with u_0 satisfying Hypothesis 4.3.9 and

$$\Lambda(\theta) = -b\theta + \frac{\sigma^2}{2}\theta^2 + \lambda\int_{\mathbb{R}}(e^{-\theta y}-1)\mu(\mathrm{d}y).$$

By Theorem 4.3.2, we know that equation (4.4.3) has a wave front moving at speed $qt + o(t)$ where

$$q = \inf_{\theta>0}\frac{1/k + \Lambda(\theta)}{\theta}.$$

Here,

$$q = \inf_{\theta>0}\left[\frac{1}{k\theta} + b + \frac{\sigma^2}{2}\theta + \lambda\int_{\mathbb{R}}\frac{e^{-\theta y}-1}{\theta}\mu(\mathrm{d}y)\right].$$

This is convex with $J(0+) = J(\infty) = \infty$ and so there is a minimum.

Let's consider the special case when $b = \theta = 0$, $\lambda = 1$ and $\mu(\mathrm{d}y) = H(y)\mathrm{d}y$ where $H : \mathbb{R} \to [0,\infty)$ is compactly supported, even and $\int_{\mathbb{R}} H(y)\mathrm{d}y = 1$. In this case, equation (4.4.3) becomes

$$u_t = H * u - u + \frac{1}{k}u(1-u^k) \tag{4.4.4}$$

where $H * v(x) = \int_{\mathbb{R}} H(x-y)v(y)\mathrm{d}y$ since

$$H * u(t,x) - u(t,x) = \int_{\mathbb{R}}(u(t,x+y) - u(t,x))H(y)\mathrm{d}y.$$

Equation (4.4.3) is of the form in equation (1.1.14). In this case

$$q = \min_{\theta>0}\frac{1}{\theta}\left(\int_{\mathbb{R}} H(y)e^{\theta y}\mathrm{d}y + \frac{1}{k} - 1\right)$$

for all u_0 satisfying Hypothesis 4.3.9.

Carr and Chmaj [CC04] showed this q to be the minimal speed of travelling waves i.e. decreasing solutions to

$$-cu' = H * u - u + \frac{1}{k}u(1-u^k)$$

with $w(-\infty) = 1$ and $w(\infty) = 0$.

Example 4.4.4. *Set $X_t = x + bt + P_t$, where P is a Poisson process with jumps of size y with intensity $\lambda > 0$ and where $b = \lambda \mathbb{1}_{|y|\le 1} y$. Then the Lévy measure is given by $\nu(A) = \lambda\delta_y(A)$ for a measurable set $A \in \mathbb{R}$. So, the equation is*

$$u_t = \lambda u(\cdot, \cdot + y) + \left(\tfrac{1}{k} - \lambda\right) u - \frac{u^{1+k}}{k} \tag{4.4.5}$$

where u_0 satisfies Hypothesis 4.3.9. Then,

$$\Lambda(\theta) = \lambda(e^{-\theta y} - 1)$$

and

$$q = \inf_{\theta>0} \left[\frac{\frac{1}{k} + \lambda(e^{-\theta y} - 1)}{\theta} \right].$$

Note that if $y < 0$, then J is convex with $J(0+) = J(\infty) = \infty$ and so there is a minimum and the speed of the front is $qt + o(t)$. If, on the other hand, $y \ge 0$ then $\inf_{\theta>0} J(\theta) = 0$ although this is not attained at a finite θ.

4.4.1. Fractional Laplacian

As we saw in the Section 1.1, one particular Markov process of interest is the α-stable Lévy process. There, the Lévy measure is given by

$$\nu(\mathrm{d}y) = \frac{c}{|y|^{1+\alpha}} \mathrm{d}y$$

and so

$$\Lambda(\theta) = \infty$$

for all $\theta > 0$. Therefore, Theorem 4.2.4 does not tell anything useful but the result is consistent with the result of Cabré and Roquejoffre [CR13] that $m(t)$ increases exponentially in t.

However, we can adapt the techniques in the proof of Theorem 4.2.1 to obtain the following.

Proposition 4.4.5. *Let u be a solution to*

$$\begin{aligned} u_t &= -\left(-\frac{\partial^2}{\partial x^2}\right)^{\alpha/2} u + \tfrac{1}{k} u(1 - u^k) \\ u_0 &= \mathbb{1}_{(-\infty,0)}(x). \end{aligned} \tag{4.4.6}$$

Then,

$$\lim_{t\to\infty} \frac{1}{t} \log m(t) = \frac{1}{k\alpha}.$$

Proof. Let

$$F_s(x) = \mathbb{P}(X_s \le x).$$

By the bound (4.3.4), we have that for t large enough,

$$u(t, x) \ge a F_{s_1} \left(-x - \sum_{j=2}^{n} F_{s_j}^{-1} \left(b e^{-s_{j-1}/k} \right) \right). \tag{4.4.7}$$

Write $s_j = \beta_j t$ for some constants $\beta_j > 0$, $\sum_{k=1}^{n} \beta_j = 1$. Setting the right hand side equal to 1/2 and rearranging we have

$$\begin{aligned} m(t) &\geq - F_{s_1}^{-1}\left(\tfrac{1}{2a}\right) - \sum_{j=2}^{n} F_{s_j}^{-1}\left(be^{-s_{j-1}/k}\right) \\ &= \exp\left(\max_{1\leq j\leq n-1} \beta_j t/k\alpha\right) \cdot \mathcal{O}(t) \end{aligned}$$

where we used

$$F_s^{-1}(x) \sim Ctx^{-\frac{1}{\alpha}} \text{ as } x \searrow 0$$

for some constant $C > 0$ (see, for example [ST94]).

Thus,

$$\liminf_{t\to\infty} \frac{1}{t} \log m(t) \geq \frac{\max_j \beta_j}{k\alpha}.$$

In particular, we can choose $\beta_1 = 1 - \varepsilon$ and $\beta_2 = \varepsilon$ for $\varepsilon \in (0, 1)$. Thus,

$$\liminf_{t\to\infty} \frac{1}{t} \log m(t) \geq \frac{1-\varepsilon}{k}$$

for any $\varepsilon \in (0, 1)$ and so

$$\liminf_{t\to\infty} \frac{1}{t} \log m(t) \geq \frac{1}{k}.$$

For the upper bound note that,

$$u(t, x) \leq e^{t/k} F_t(-x)$$

Setting the RHS equal to 1/2 we see that

$$m(t) \leq e^{t/k\alpha} \cdot \mathcal{O}(t)$$

and so,

$$\limsup_{t\to\infty} \frac{1}{t} \log m(t) \leq \frac{1}{k}$$

as required. □

Remark 18. *By using a similar idea to the proof of Theorem 4.2.4, the result of Proposition 4.4.5 can be easily adapted to any initial condition satisfying*

$$\liminf_{x\to-\infty} u_0(x) > 0 \tag{4.4.8}$$

and

$$u_0(x) \leq 1 \wedge Cx^{-\alpha} \tag{4.4.9}$$

for some $C > 0$ and all $x > 0$. Firstly, we can easily show the result for $u_0(x) = c_1 \mathbb{1}_{(-\infty, c_2)}$ for constants c_1, c_2. The first condition then implies that

$$\liminf_{t\to\infty} \frac{1}{t} \log m(t) \geq \frac{1}{k}.$$

For the upper bound, we can use that the transition kernel p of X satisfies

$$p(t, x) \leq \frac{Ct^{-\frac{1}{\alpha}}}{1 + t^{-\frac{1}{\alpha}}|x|^{1+\alpha}}.$$

for a constant $C > 0$. Thus, $\mathbb{E}^x[X_t^{-\alpha} \mathbb{1}_{(c,\infty)}(X_t)] \leq Cx^{-\alpha}$ for constants $c, C > 0$ and x large enough ($x > c$), as required. ◇

4.4.2. Ornstein-Uhlenbeck Processes

Suppose that $k > 0$, and the Markov process X is an Ornstein-Uhlenbeck process solving the equation

$$\begin{aligned} \mathrm{d}X_s &= \theta(\mu - X_s)\mathrm{d}s + \sigma \mathrm{d}W_s \\ X_0 &= x \end{aligned}$$

for constants $\mu \in \mathbb{R}$, and $\theta, \sigma > 0$. Then,

$$\mathcal{L}(f)(x) = \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial^2 x} + \theta(\mu - x)\frac{\partial f}{\partial x}.$$

Then, consider the equation

$$\begin{aligned} \frac{\partial u}{\partial t} &= \mathcal{L}u + \tfrac{1}{k}u(1-u^k) \\ u_0 &= \mathbb{1}_{(-\infty,0)} \end{aligned} \tag{4.4.10}$$

In this case, the analogue to Theorem 4.2.1 is the following.

Theorem 4.4.6. *Let u be a solution to equation* (4.4.10). *Then,*

$$a\Phi\left(-\frac{\alpha}{\beta}x + \frac{1-\alpha^{n-1}}{1-\alpha}\frac{\rho}{\alpha^{n-1}} - \frac{\gamma}{\beta\alpha^{n-1}}\right) \leq u(t,x) \leq e^{t/k}\Phi\left(-\frac{\alpha}{\beta}x - \frac{\gamma}{\beta}\right).$$

Moreover,

$$-\frac{\beta}{\alpha}\Phi^{-1}\left(\frac{1}{2a}\right) + \frac{\beta}{\alpha^n}\frac{1-\alpha^{n-1}}{1-\alpha}\rho - \frac{\gamma}{\alpha^n} \leq m(t) \leq -\frac{\beta}{\alpha}\Phi^{-1}\left(\frac{1}{2e^{t/k}}\right) - \frac{\gamma}{\alpha}$$

where $m(t) = \sup\{x : u(t,x) \geq \frac{1}{2}\}$.

Proof. Note that

$$X_s \sim \mathcal{N}\left(e^{-\theta s}x + \mu(1-e^{-\theta s}), \frac{\sigma^2}{2\theta}(1-e^{-2s\theta})\right).$$

We rewrite this as $X_s \sim \mathcal{N}(\alpha(s)x + \gamma(s), \beta(s)^2)$ where,

$$\begin{aligned} \alpha(s) &= e^{-\theta s}, \\ \gamma(s) &= \mu(1-e^{-\theta s}), \\ \beta(s) &= \frac{\sigma}{\sqrt{2\theta}}\sqrt{1-e^{-2\theta s}}. \end{aligned}$$

Following the proof of Theorem 4.2.1, and writing

$$V_{t/n} = P_{t/n} \circ D_{t/n}.$$

we see that

$$u(t,x) \geq V_{t/n}^{(n-1)}\Phi\left(-\frac{\alpha}{\beta}x - \frac{\gamma}{\beta}\right)$$

Write $\rho = -\Phi^{-1}(be^{-t/nk})$ and note that for n and b fixed and t large, we have $\rho > 0$.

We define two sequences, c_n and d_n, with $c_1 = \frac{\alpha}{\beta}$ and $d_1 = -\frac{\gamma}{\beta}$, by

$$\begin{aligned} V_{t/n}(a\Phi(-c_j \cdot + d_j)(x) &\geq a\mathbb{P}(\Phi(-c_j X + d_j) \geq be^{-t/nk}) \\ &= a\Phi\left(-c_{j+1}x + d_{j+1}\right) \end{aligned}$$

and so

$$\begin{aligned} c_{j+1} &= \frac{\alpha}{\beta} \\ d_{j+1} &= \frac{\rho + d_j}{c_j \beta} \end{aligned}$$

for all $j = 1, \cdots, n-1$.

Thus, we find

$$\begin{aligned} c_n &= \frac{\alpha}{\beta}, \\ d_n &= \frac{1-\alpha^{n-1}}{1-\alpha}\frac{\rho}{\alpha^{n-1}} - \frac{\gamma}{\beta\alpha^{n-1}}. \end{aligned}$$

To see the second part of the Theorem, since u is decreasing in x, we can simply set the upper and lower bounds equal to $1/2$ and rearrange for x. $\square$

Remark 19. *Compare this to Theorem 4.2.1. The proof above works with* $\theta = 0$ *(and then* $\beta(s) = s$*),* $\mu = 0$ *and* $\sigma = 1$*. Then, we see that* c_n *and* d_n *above simplify to*

$$\begin{aligned} c_n &= \sqrt{\frac{n}{t}}, \\ d_n &= (n-1)\rho, \end{aligned}$$

as we saw before. $\diamond$

Theorem 4.4.7. *Let* u *be the solution to equation* (4.4.10)*. Then,*

$$\lim_{t \to \infty} \frac{\log m(t)}{t} = \theta$$

Proof. From Theorem 4.4.6, we have

$$\begin{aligned} m(t) &\leq -\frac{\sigma}{\sqrt{2\theta}}\sqrt{1 - e^{-2\theta t}}e^{\theta t}\Phi^{-1}\left(\frac{1}{2e^{t/k}}\right) - \mu(1 - e^{-\theta t})e^{\theta t} \\ &\leq C_1\sqrt{t}e^{\theta t} \end{aligned}$$

for a constant $C_1(\theta, \mu, \sigma, a) > 0$ by Lemma 4.2.3.

Similarly, for t sufficiently large,

$$m(t) \geq C_2\sqrt{t}e^{\theta t/n}$$

for a constant $C_2(\theta, \mu, \sigma, a, n) > 0$. $\square$

4.5. Waves fronts when $k < 0$

Consider equation (4.3.1). So far we have only studied the case when $k > 0$. The reason for this is that the nonlinearity given by $f(u) = \frac{u}{k}(1 - u^k)$ is not Lipschitz at 0 and so may have multiple solutions. Indeed, take $k \in (-1, 0)$ and $u_0 = 0$. Then, we can define a mild solution by both $u(t, x) = 0$ and by $u(t, x) = D_t(0)(x) = (1 - e^{-t})^{-\frac{1}{k}}$.

Moreover, for any u_0 such that we have a solution in the sense of Definition 2.1.1, we have

$$u(t, x) \geq (1 - e^{-t})^{-\frac{1}{k}}$$

where u is defined by the relevant optimisation problem and so the solution arising from the optimisation problem does not give rise to wave fronts. Given that we should expect wave fronts to occur in the setting of branching processes, it follows that the optimisation method does not define the natural solution in this setting.

However, one could study these solutions using the operators P and D and in this case Corollary 2.6.1 give u that when $k \in (-1, 0)$,

$$P_t \circ D_t \circ u_0(x) \leq u(t, x) \leq D_t \circ P_t \circ u_0(x)$$

and when $k \in (-\infty, -1)$, we have

$$D_t \circ P_t \circ u_0(x) \leq u(t, x) \leq P_t \circ D_t \circ u_0(x)$$

where $(P_t \circ f)(x) = \mathbb{E}^x[f(X_t)]$ and $(D_t \circ f)(x) = (e^{-t} f(x)^{-k} + 1 - e^{-t})^{-\frac{1}{k}}$.

Appendices

Appendix A.

A Generalisation for Concave Nonlinearities

In this Appendix, we give another application of the ideas presented so far. In this work, we have been studying the general equation

$$\begin{aligned} \frac{\partial u}{\partial t} &= \mathcal{L}u + \phi(x)u - \tfrac{1}{k}u^{k+1} \\ u(0,x) &= u_0(x) \text{ for all } x \in E, \end{aligned} \tag{A.0.1}$$

and when studying wave fronts, we considered the special case when $k > 0$ and $\phi = 1/k$, $\mathcal{L}$ is the generator of some Lévy process and $E = \mathbb{R}$. In this case all of the representations of Chapter 2 hold. However, another possible generalisation of the FKPP equation was introduced in Section 1.1.3. Recall that we can define the probability generating function of a non-negative integer valued random variable N by

$$G(s) = \mathbb{E}[s^N].$$

Then, if $\mathbb{E}[N] < \infty$, the generalisation of the FKPP equation is given by

$$\begin{aligned} \frac{\partial u}{\partial t} &= \mathcal{L}u + f(u) \\ u(0,x) &= u_0(x) \text{ for all } x \in \mathbb{R}^d. \end{aligned}$$

Here, we define f by $f(u) = 1 - u - G(1-u)$. f is a KPP-type nonlinearity when $\mathbb{E}[N] < \infty$.

As in Chapter 4, we suppose that $\mathcal{L}$ is the generator of a Lévy process X with moment generating function, Λ, of $-X$ defined in a neighbourhood of $\theta = 0$; that is,

$$\mathbb{E}[e^{-\theta X_s}] = e^{s\Lambda(\theta)}.$$

for all θ in a neighbourhood of 0.

We see that in this case the results presented in this work so far do not apply directly. Indeed, the specific form of the nonlinearity that we have so far been studying is integral to the martingale representation when we used Hölder's inequality in Chapter 2 and the FKPP equation arising in the financial setting of Chapter 3 is not of this form. However, the specific form of the nonlinearity $f(u) = \frac{1}{k}u(1-u^k)$ was not as integral to the study of wave speed and the semigroup $(D_s)_{s\geq 0}$ was the salient object to study in relation to f. In fact, the wave speed is only given in terms of the important term $f'(0)$ (which is equal to $\frac{1}{k}$ in this case). Therefore, one might expect that it is possible to adapt the primal representation to this case

– and use analogous methods to those used in Chapter 4 – to obtain results for a different class of nonlinearities such as the one above arising from a general class of branching processes. This is the content of the rest of this Appendix.

A.1. Representation for Concave Nonlinearities

Firstly, we define the equation that we will study in this section. Here it is natural to define solutions in the mild sense.

Assumption A.1.1. *Suppose that the function $f : [0,\infty) \times E \times \mathbb{R} \to \mathbb{R}$ is measurable and $f(t,x,\cdot)$ is concave and differentiable for all (t,x). Suppose that $u : [0,\infty) \times E \to K$ is a mild solution to equation*

$$\begin{aligned} \frac{\partial u}{\partial t} &= \mathcal{L}u + f(t,x,u), \text{ in } (0,\infty) \times E \\ u(0,x) &= u_0(x), \text{ for all } x \in E. \end{aligned} \tag{A.1.1}$$

for some set $K \subset \mathbb{R}$ and that $\left|\frac{\partial f}{\partial u}\right| < C$ in $[0,\infty) \times E \times K$ for some constant $C > 0$.

We include the condition on the derivative of f here so that the optimal control in Theorem A.1.3 is in the set of admissible controls.

Results on existence and uniqueness of mild solutions were given in Theorem 2.7.2.

Definition A.1.2. *Let $g : [0,\infty) \times E \times \mathbb{R} \to \mathbb{R}$, and suppose that $g(t,x,\cdot)$ is concave for any $(t,x) \in [0,\infty) \times E$. We define a function $\hat{g}$ by*

$$\hat{g}(\cdot,\cdot,z) = \sup_{u \geq 0}(g(\cdot,\cdot,u) - uz).$$

Remark 20. *Usually, for a function $h : \mathbb{R} \to \mathbb{R}$, one defines the Legendre transform by $h^*(z) = \sup_{u\geq 0}(uz - h(u))$. In this case, $\hat{h}(z) = (-h)^*(-z)$.*

One could alternatively consider the PDE with convex nonlinearities and obtain analogous results, to the following, in terms of f^*. ◇

We have the following representation.

Theorem A.1.3. *Let u be a mild solution to equation* (A.1.1)*, as in Definition A.1.1 above. Then, the following holds:*

$$u(t,x) = \min_Z \mathbb{E}\left[\int_0^t e^{\int_0^s Z_r \mathrm{d}r} \hat{f}(t-s, X_s, Z_s)\mathrm{d}s + e^{\int_0^t Z_r \mathrm{d}r} u_0(X_t)\right] \tag{A.1.2}$$

where the minimum is taken over all bounded processes Z which are adapted to the filtration generated by X. The minimum is attained at $Z = Z^*$ where

$$Z^* = \frac{\partial f}{\partial u}(t-s, X_s, u(t-s, X_s))$$

Proof. Set

$$M_s = u(t-s, X_s) + \int_0^s f(t-r, X_r, u(t-r, X_r))\mathrm{d}r.$$

Since, u is a mild solution, it follows from the Markov property of X, that M is a martingale with respect to the filtration generated by X.

Define,

$$\xi_s = \int_0^s e^{\int_0^r Z_\nu \mathrm{d}\nu} \hat{f}(t-r, X_r, Z_r)\mathrm{d}r + e^{\int_0^t Z_r \mathrm{d}r} u(t-s, X_s).$$

Then,

$$\begin{aligned}\xi_t = &\int_0^t e^{\int_0^s Z_r \mathrm{d}r}[u_s Z_s + \hat{f}(t-s, X_s, Z_s) - f(t-s, X_s, u_s)]\mathrm{d}s \\ &+ M_t + \int_0^t (M_t - M_s) Z_s e^{\int_0^s Z_r \mathrm{d}r}\end{aligned} \tag{A.1.3}$$

where $u_s = u(t-s, X_s)$. This follows from Fubini's Theorem and integration by parts.

Since f is concave, by the definition of $\hat{f}$, it follows that

$$uz + \hat{f}(\tau, x, z) - f(\tau, x, u) \geq 0$$

with equality for $z = \frac{\partial f}{\partial u}(\tau, x, u)$. Since M is a martingale, it follows that

$$\mathbb{E}[\xi_t] \geq \xi_0 = u(t, x).$$

with equality for $Z = Z^*$ □

To see where equation (A.1.3) comes from, note that if u and X are sufficiently regular, then equation (A.1.3) would follow directly from Itô's formula.

Remark 21. *When $k > 0$, we can see that Theorem A.1.3 is a generalisation of representation* (2.2.3)*. Indeed, if $f(u) = \frac{1}{k}(u - u^{k+1})$, then*

$$\hat{f}(z) = \begin{cases} \left(\frac{1-zk}{k+1}\right)^{1+1/k} & \text{if } z \leq 1/k \\ 0 & \text{if } z > 1/k \end{cases}$$

In other words,

$$u(t, x) = \min_Z \mathbb{E}\left[\int_0^t e^{\int_0^s Z_r \mathrm{d}r} \left(\frac{1-Z_s k}{k+1}\right)^{1+1/k} \mathrm{d}s + e^{\int_0^t Z_r \mathrm{d}r} u_0(X_t)\right] \tag{A.1.4}$$

where, without loss of generality, the minimum is taken over the set of all bounded above by $1/k$ which are adapted to the filtration generated by X.

Setting

$$Y_s = \exp\left(-\frac{1}{k+1}\int_0^s (1 - kZ_r)\mathrm{d}r\right)$$

we recover

$$u(t, x) = \min_Y \mathbb{E}\left[\int_0^t e^{s/k}\left(-\dot{Y}_s\right)^{1+1/k} \mathrm{d}s + e^{t/k} Y_t^{1+1/k} u_0(X_t)\right]. \tag{A.1.5}$$

where the minimum is over the set of positive, absolutely continuous, adapted and decreasing processes Y. Note that the set of Y that we are taking the minimum over is smaller than in Theorem 2.2.3, but this does not matter for the subsequent results in this work. ◇

A.2. Application to Branching Lévy Processes

Now we restrict attention to the one-dimensional case when $d = 1$ and f does not depend on t or x.

Analogously to Theorem 4.1.4, one can find iterative bounds using the nonlinear semigroup corresponding to f and the linear semigroup corresponding to $\mathcal{L}$. In particular, one can easily show that analogous bounds hold.

Definition A.2.1. *For any $t \geq 0$, $x \in \mathbb{R}$, and $g : \mathbb{R} \to \mathbb{R}$, Define the function P by*

$$P_t(g)(x) = \mathbb{E}[g(x + X_t)] \tag{A.2.1}$$

and the function D to be such that $D_\cdot(g)(x)$ is the solution to the ODE:

$$\frac{\mathrm{d}\psi}{\mathrm{d}t} = f(\psi) \tag{A.2.2}$$

$$\psi(0) = g(x). \tag{A.2.3}$$

The following result is analogous to Theorem 4.1.4:

Theorem A.2.2. *Let u be as in Theorem A.1.3 and D and P as in Definition A.2.1. Then,*

$$(P_{t/n} \circ D_{t/n})^{(n)}(u_0)(x) \leq u(t, x) \leq (D_{t/n} \circ P_{t/n})^{(n)}(u_0)(x)$$

for $t \geq 0$, $x \in \mathbb{R}$.

Proof. As before, we consider the optimisation problem restricted to deterministic processes: by Theorem A.1.3, we have

$$\begin{aligned} u(t, x) &\leq \min_{Z \text{ deterministic}} \int_0^t e^{\int_0^s Z_r \mathrm{d}r} \hat{f}(Z_s)\mathrm{d}s + e^{\int_0^t Z_r \mathrm{d}r} \mathbb{E}\left[u_0(X_t)\right] \\ &= D_t(\mathbb{E}\left[u_0(X_t + \cdot)\right])(x) \end{aligned}$$

Here, we used Theorem A.1.3 with $\mathcal{L} = 0$ to show

$$D_t(g)(x) = \min_{Z \text{ deterministic}} \int_0^t e^{\int_0^s Z_r \mathrm{d}r} \hat{f}(Z_s)\mathrm{d}s + e^{\int_0^t Z_r \mathrm{d}r} g(x).$$

Similarly, by considering X pathwise,

$$D_t(u_0)(X_t) \leq \int_0^t e^{\int_0^s Z_r \mathrm{d}r} \hat{f}(Z_s)\mathrm{d}s + e^{\int_0^t Z_r \mathrm{d}r} u_0(X_t)$$

almost surely for all adapted processes Z (and even anticipative Z). Therefore,

$$u(t, x) \geq \mathbb{E}[D_t(u_0)(X_t + x)]$$

If we write $u(t, x) = U_t(u_0)(x)$, then $P_t \circ D_t \leq U_t \leq D_t \circ P_t$ and since each of the operators P_t, D_t and U_t are increasing, for any t, we can iterate the above expression to complete the proof. $\square$

It is now straightforward to generalise Theorem 4.3.2 to the case of a general concave nonlinearity that is of the form $f(u) = 1 - u - G(1-u)$ described above.

In particular, setting $u_0 = \mathbb{1}_{(\infty,0)}$, the following result can also be interpreted in terms of the distribution function for the maximum particle of the corresponding branching Lévy process and thus about the speed of maximum particle. Note that results of this type can also be found in the work of Biggins [BLSW91], Kyprianou [Kyp99] and, recently, Groisman and Jonckheere [GJ16].

Theorem A.2.3. *Let $f : \mathbb{R} \to \mathbb{R}$ be concave, and differentiable with uniformly bounded derivative. Suppose that u is a mild solution to*

$$\frac{\partial u}{\partial t} = \mathcal{L}u + f(u)$$
$$u_0(x) = \mathbb{1}_{(-\infty,0)}.$$

Suppose that Λ is finite in a neighbourhood of 0. Write $f'(0) = \gamma > 0$, $f(0) = f(\beta) = 0$.

Then, for

$$q = \inf_{\theta>0} \frac{\gamma + \Lambda(\theta)}{\theta}, \tag{A.2.4}$$

we have

$$u(t, rt) \to \begin{cases} 0, & \text{if } r > q \\ \beta, & \text{if } r < q. \end{cases} \tag{A.2.5}$$

The relevant existence and uniqueness theorem for this equation is Theorem 2.7.3.

Note that since the probability of the branching process going extinct is given by the smallest non-negative root of $G(s) = s$ (see, for example, Athreya and Ney [AN72]), we see that β corresponds to the probability of survival over all time.

The proof follows by using the same techniques as Section 4.3. In particular, for $x \in (0, \beta)$, t sufficiently large and $c < \gamma$, we have

$$xe^{ct} \leq D_t(Id)(x) \leq xe^{\gamma t}$$

where Id is the identity map. Therefore, the techniques used in Section 4.3 apply almost verbatim. One can also see the paper by Driver and Tehranchi [DT18a] for more details.

Appendix B.

A Dynamic Programming Approach

In Sections 2.2 and 2.3, we saw that a solution, u, to equation (2.1.1), can be represented in the primal form of Theorem 2.2.3 and in the martingale-dual form in Theorem 2.3.2. We also saw in Section 2.5 that, formally, there is a natural Lagrangian dual formulation but we did not make this rigorous.

In this appendix we give more intuition behind Theorem 2.2.3 in case when $k > 0$. In particular, we study the result in the context of the well studied dynamic programming approach. We will then show directly (that is, without Lagrangian techniques) that the formula in (2.5.1) is correct under the right conditions. The cases when $k < 0$ can be treated analogously given appropriate assumptions. As noted before, the Lagrangian dual is less general than the martingale dual in Section 2.3. Indeed, it is necessary that the initial condition u_0 is strictly positive and for the proof below we will require that the generator $\mathcal{L}$ is a strictly elliptic operator of the form

$$\mathcal{L} = \sum_{1\le i\le n} b_i \frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{1\le i,j\le n} a_{ij}\frac{\partial^2}{\partial x_i \partial x_j} \tag{B.0.1}$$

This corresponds to the process X satisfying the SDE

$$\mathrm{d}X_s = b(X_s)\mathrm{d}s + \sigma(X_s)\mathrm{d}W_s$$

with $X_0 = x \in \mathbb{R}^d$ and where $a = \sigma\sigma^\intercal$. Here $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}^d$ and W is a standard m-dimensional Brownian motion.

We will assume that u_0 takes values in $[0, 1]$. Then, for non trivial initial conditions, $0 < u(t, x) < 1$ for all $t > 0$, $x \in \mathbb{R}$ (which can be seen, for example, by a comparison principle argument).

We remark that Hypothesis 2.1.3 is automatically satisfied in this appendix.

B.1. Primal Formulation

For simplicity, we are assuming that $k > 0$ and $\mathcal{L}$ is of the form (B.0.1) and $\phi = 1/k$ but these assumptions are not necessary. We will suppose that u is a classical solution to

$$\frac{\partial u}{\partial t} = \mathcal{L}u + \frac{1}{k}(u - u^{k+1}) \tag{B.1.1}$$

with initial condition u_0.

It follows from Theorem 2.2.3, that we can write

$$u(t,x) = \min_{Y \in \mathcal{Y}^1_{0,t}} \mathbb{E}\left[\int_0^t e^{s/k}|\dot{Y}_s|^{1+\frac{1}{k}}\mathrm{d}s + e^{t/k}|Y_t|^{1+\frac{1}{k}}u_0(X_t)\right] \tag{B.1.2}$$

where we define $\mathcal{Y}^y_{s,t}$ to be the set of all adapted, absolutely continuous processes $(Y_r)_{s\le r\le t}$ with $Y_s = y$.

We saw that this representation can be thought of as a generalisation of the Feynman-Kac formula. However, it is, perhaps, more natural to think of equation (B.1.1) as the Hamilton-Jacobi-Bellman (HJB) equation (up to change of variables) corresponding to the optimisation problem (B.1.2). We will formally derive the HJB equation corresponding to (B.1.1) and then show that the representation (B.1.2) can be shown by the standard argument for verifying that the formal HJB equation is actually the correct equation.

B.1.1. The HJB Equation

Let X be the solution to the SDE corresponding to $\mathcal{L}$ with $X_0 = x$ and let Y be as above with $Y_0 = y$. Define the value function V by

$$V(s,x,y) = \inf_{Y \in \mathcal{Y}^y_{s,t}} \mathbb{E}\left[\int_s^t f(r, X_r, \dot{Y}_r)\mathrm{d}r + g(t, X_t, Y_t)\middle| X_s = x, Y_s = y\right]$$

for some suitably differentiable functions f and g.

By the dynamic programming principle we expect that

$$V(0,x,y) = \inf_{Y \in \mathcal{Y}^y_{0,s}} \mathbb{E}\left[\int_0^s f(r, X_r, \dot{Y}_r)\mathrm{d}r + V(s, X_s, Y_s)\right]$$

and so by the Martingale Principle of Optimal Control we look for

$$M_s := \int_0^s f(r, X_r, \dot{Y}_r)\mathrm{d}r + V(s, X_s, Y_s)$$

to be a submartingale for all Y and a true martingale at the optimal choice of Y. Applying Itô's formula to M, we minimising the drift term pointwise in y and formally obtain the HJB equation,

$$\inf_z \left[V_t + \mathcal{L}V + zV_y + f\right] = 0.$$

with $V(t,x,y) = g(t,x,y)$. Set

$$f(s,x,y) = e^{s/k}|y|^{1+\frac{1}{k}},$$
$$g(s,x,y) = e^{t/k}|y|^{1+\frac{1}{k}}u_0(x).$$

By scaling, we make the ansatz that $V(s,x,y) = e^{s/k}|y|^{1+1/k}u(t-s,x)$. We see that the HJB equation simplifies to

$$\inf_z \left[e^s|y|^{1+1/k}(-u_t + \mathcal{L}u + \frac{u}{k}) + (1+1/k)z\,\mathrm{sign}(y)|y|^{1/k}e^s u + e^s|z|^{1+1/k}\right] = 0.$$

or

$$\begin{aligned} u_t &= \mathcal{L}u + \frac{u}{k} + \inf_w \left[w^{1+1/k} - (1+1/k)wu \right] \\ &= \mathcal{L}u + \frac{1}{k}(u - u^{1+k}). \end{aligned}$$

as required. Choosing $w = u^k$ is optimal.

We see that the optimal Y satisfies

$$-\frac{\dot{Y}_s}{Y_s} = u(t-s, X_s)^k.$$

We write this optimal control as Y^* where

$$Y_s^* = e^{-\int_0^s u(t-r,X_r)^k \, \mathrm{d}r} \text{ for all } 0 \le s \le t.$$

Remark 22. *Here we see that the specific form of the nonlinearity is integral to this approach and any alternative is only likely to work if f and g are sufficiently tractable.* ◇

B.1.2. Verifying the HJB Equation

Theorem B.1.1. *Let u be as above. Then,*

$$u(t,x) = \min_{Y \in \mathcal{Y}^1_{0,t}} \mathbb{E}\left[\int_0^t e^{s/k} |\dot{Y}_s|^{1+1/k} \, \mathrm{d}s + e^{t/k} u_0(X_t) |Y_t|^{1+1/k} \right]$$

where the minimum is taken over all adapted, absolutely continuous, processes Y with $Y_0 = 1$.

Moreover, the minimum is attained at $Y = Y^*$.

Note that Theorem B.1.1 follows from Theorem 2.2.3 but here we will give an alternative proof.

Proof. Fix $(t,x) \in (0,\infty) \times \mathbb{R}^d$ and drop it from the notation below. By applying Itô's formula, we have

$$\begin{aligned} &\mathrm{d}\left(\int_0^s e^{r/k} |\dot{Y}_r|^{1+1/k} \mathrm{d}r + e^{s/k} |Y_s|^{1+1/k} u \right) \\ &\quad = e^{s/k} \left(|\dot{Y}_s|^{1+1/k} + \tfrac{k+1}{k} \mathrm{sign}(Y_s) |Y_s|^{1/k} |\dot{Y}_s| u + |Y_s|^{1+1/k} (-u_t + \mathcal{L}u + u) \right) \mathrm{d}s + \mathrm{d}M_s \\ &\quad = e^{s/k} \left(|\dot{Y}_s|^{1+1/k} + \tfrac{k+1}{k} \mathrm{sign}(Y_s) |Y_s|^{1/k} |\dot{Y}_s| u + \frac{1}{k} |Y_s|^{1+1/k} u^{k+1} \right) \mathrm{d}s + \mathrm{d}M_s \\ &\quad \ge \mathrm{d}M_s \end{aligned} \tag{B.1.3}$$

where the terms involving u are evaluated at $(t-s, X_s)$.

$$M_s = \int_0^s e^{r/k} |Y_r|^{1+1/k} (\nabla_x u)(t-r, X_r)^\intercal \sigma \mathrm{d}W_r$$

is a local martingale. In the last line we used the fact that for all $k > 0$, $u \ge 0$, $y, z \in \mathbb{R}$, $\varepsilon = \mathrm{sign}(y)$ we have

$$|z|^{1+1/k} + \tfrac{k+1}{k} \varepsilon |y|^{1/k} z u + \frac{|y|^{1+1/k} u^{k+1}}{k} \ge 0 \tag{B.1.4}$$

with equality if

$$z = -yu^k.$$

This follows since the expression on the left hand side of (B.1.4) is convex in z with its unique stationary point at $z = -yu^k$.

Now let $(\tau_n)_{n\geq 1}$ be an increasing sequence of stopping times such that $\tau_n \nearrow t$ as $n \to \infty$ and which reduces the local martingale M. By integrating equation (B.1.3), and taking expectations, we have

$$u(t,x) \leq \mathbb{E}\left[\int_0^{\tau_n} e^{r/k}|\dot{Y}_r|^{1+1/k}\mathrm{d}r + e^{\tau_n/k}u(t-\tau_n, X_{\tau_n})|Y_{\tau_n}|^{1+1/k}\right] \tag{B.1.5}$$

with equality when $\dot{Y}_s = -u(t-s, X_s)^k Y_s$.

We need to take limits inside the expectation. The first term clearly converges by the Monotone Convergence Theorem. By the boundedness of u, the process Y^* is automatically in class D. Therefore, we will be done if the process defined by

$$e^{s/k}u(t-s, X_s)Y_s^{1+1/k}$$

is also in class D for any Y. Since $e^{s/k}u(t-s, X_s) \leq e^{t/k}$ for any s, we only need to show that Y is uniformly dominated by an integrable random variable.

We can assume without loss of generality that

$$\mathbb{E}\left[\int_0^t e^{s/k}|\dot{Y}_s|^{1+1/k}\,\mathrm{d}s\right] < \infty$$

since, otherwise, there is nothing to prove.

Now note that by Jensen's inequality applied to the integral we have

$$\begin{aligned}\mathbb{E}\left[\sup_{0\leq s\leq t}|Y_s|^{1+1/k}\right] &\leq \mathbb{E}\left[\left(1+\int_0^t |\dot{Y}_r|\mathrm{d}r\right)^{1+1/k}\right]\\ &\leq 2^{1+1/k}\left(1+\mathbb{E}\left[\int_0^t |\dot{Y}_r|^{1+1/k}\mathrm{d}r\right]\right) < \infty.\end{aligned}$$

□

B.2. Lagrangian Dual

Now we consider the Lagrangian Dual formulation and make the discussion in Section 2.5 rigorous in the particular setting described above. Unlike in the primal case above, the proof below uses the fact that $\mathcal{L}$ is of the form (B.0.1) and so is less general. We also require u_0 to decay slowly enough. This will automatically imply that $u(t,x) > 0$ for all $t > 0$, $x \in \mathbb{R}$. Neither of these conditions are required in the case of the martingale dual of Section 2.3.

Define $\mathcal{N}_t$ to be the set of bounded positive $\mathcal{F}$-martingales. We define α to be such that for all s,

$$N_s = \mathcal{E}\left(\int_0^s \alpha_r^\intercal \mathrm{d}W_r\right).$$

We also define a process N^* by

$$N_s^* = \frac{u(t-s, X_s)}{u(t,x)} e^{\frac{1}{k}\int_0^s \left(1-u(t-r,X_r)^k\right)\mathrm{d}r}$$

for each $s \in [0,t]$. Recall that this is simply the martingale M^* scaled to start at 1.

Theorem B.2.1. *Suppose that*

$$\mathbb{E}\left[u_0(X_t)^{-k}\right] < \infty. \tag{B.2.1}$$

Then,

$$\frac{1}{u(t,x)^k} = \min_{N\in\mathcal{N}_t} \mathbb{E}\left[\int_0^t e^{-s} N_s^{k+1}\,\mathrm{d}s + \frac{e^{-s}N_t^{k+1}}{u_0(X_t)^k}\right].$$

Moreover, $N = N^*$ *attains the minimum.*

The proof is essentially the same as the verification proof above for the PDE with solution defined by $u(t-s,x)^{-k}$. In this spirit, we have the following lemma.

Lemma B.2.2. *Suppose that* u_0 *is such that* u *is a positive solution to equation* (B.1.1)*. Fix* $t > 0$ *and let*

$$v(s,x) = u(t-s,x)^{-k}.$$

Then, v *solves*

$$\begin{aligned} \frac{\partial v}{\partial s} + \mathcal{L}v &= \frac{(1+k)}{2k}\frac{\|\nabla v^\intercal \sigma\|^2}{v} + v - 1 \\ v(t,x) &= u_0(x)^{-k}. \end{aligned}$$

Proof. We simply calculate

$$\frac{1}{v}\partial_t v = \frac{k}{u}\partial_t u \tag{B.2.2}$$

$$\frac{1}{v}\partial_{x_i} v = -\frac{k}{u}\partial_{x_i} u \tag{B.2.3}$$

$$\frac{1}{v}\partial^2_{x_i x_j} v = -k\frac{1}{u}\partial^2_{x_i x_j} u + \frac{1+k}{k}\frac{1}{v^2}(\partial_{x_i} v)(\partial_{x_j} v) \tag{B.2.4}$$

Thus,

$$\begin{aligned} \frac{1}{v}(v_t + \mathcal{L}v) &= \frac{k}{u}(u_t - \mathcal{L}u) + \frac{1+k}{2kv^2}\|\nabla v^\intercal \sigma\|^2 \\ &= \frac{1+k}{2kv^2}\|\nabla v^\intercal \sigma\|^2 + \left(1 - \frac{1}{v}\right) \end{aligned}$$

as required. □

Proof of Theorem B.2.1. Fix $x \in \mathbb{R}^d$ and $N \in \mathcal{N}_t$. By Itô's formula we have

$$\mathrm{d}N_s^{k+1} = (k+1)N_s^{k+1}\alpha_s^\intercal \mathrm{d}W_s + \frac{k(k+1)}{2}N^{k+1}\alpha^\intercal\alpha\mathrm{d}s$$

and by Lemma B.2.2,

$$\begin{aligned} \mathrm{d}(e^{-s}v) &= e^{-s}\nabla_x v^\intercal \sigma(X_s)\mathrm{d}W_s + e^{-s}(v_t + \mathcal{L}v - v) \\ &= e^{-s}\nabla_x v^\intercal \sigma \mathrm{d}W_s + e^{-s}\left(\tfrac{(1+k)}{2k}\tfrac{\|\nabla v^\intercal \sigma\|^2}{v} - 1\right)\mathrm{d}s \end{aligned}$$

where the terms involving v are evaluated at (s, X_s).

Finally,

$$\mathrm{d}\langle N^{k+1}_{\cdot}, e^{-\cdot} v(\cdot, X_{\cdot})\rangle_s = (\nabla_x v)^\intercal \sigma(X_s)\alpha_s \mathrm{d}s$$

Putting this together, we have

$$\mathrm{d}\left(\int_0^s e^{-r} N_r^{k+1}\mathrm{d}r + e^{-s} N_s^{k+1} v(s, X_s)\right) = \frac{k+1}{k}\frac{e^{-s}N_s^{k+1}}{2v(s, X_s)}\left\|kv\alpha + \nabla v^\intercal \sigma\right\|^2 \mathrm{d}s + \mathrm{d}M_s$$

where M is a local martingale. The drift term is obviously non-negative. Now let $(\tau_n)_{n\geq 1}$ be an increasing sequence of stopping times such that $\tau_n \nearrow t$ as $n \to \infty$ and which reduces M. Then

$$\mathbb{E}\left[\left(\int_0^{\tau_n} e^{-s} N_r^{k+1}\,\mathrm{d}r + e^{-\tau_n} N_{\tau_n}^{k+1} v(\tau_n, X_{\tau_n})\right)\right] \geq v(0, x). \tag{B.2.5}$$

We now send $n \to \infty$.

The optimal α is defined by

$$\alpha_s = \frac{\nabla u(t-s, X_s)^\intercal \sigma(X_s)}{u(t-s, X_s)}.$$

It is simple to check that this corresponds to N^*. The fact that $N^* \in \mathcal{N}_t$ is clear.

We can use the Monotone Convergence Theorem for the first term in (B.2.5) and take limits for the second term if we show that for any $N \in \mathcal{N}_t$, the process defined by $e^{-s}N_s^{k+1}u(t-s, X_s)^{-k}$ is of class D. Note that this holds automatically for N^* as u is bounded.

We will show that under the assumption (B.2.1) on u_0,

$$\mathbb{E} \sup_{0\leq s\leq t} u(t-s, X_s)^{-k} < \infty \tag{B.2.6}$$

and this will complete the proof.

By the representation (2.1.3), we have

$$u(t, x) \geq \mathbb{E}[u_0(X_t)]$$

and so

$$u(t-s, X_s)^{-k} \leq M_s^{-k}$$

where M is the martingale defined by

$$M_s = \mathbb{E}[u_0(X_t)|\mathcal{F}_s].$$

Since M is a martingale, M^{-1} is a submartingale and hence by Doob's inequality

$$\begin{aligned}\mathbb{E} \sup_{0\leq s\leq t} u(t-s, X_s)^{-k} &\leq \mathbb{E}\left[\sup_s (M_s^{-k/p})^p\right] \\ &\leq \frac{p}{p-1}\mathbb{E}[u_0(X_t)^{-k}] \\ &< \infty\end{aligned}$$

for all $p > 1$, by the assumption on u_0. □

Remarks

1. We can consider a bigger set of admissible controls if we have more conditions on u_0. For example, we can stipulate that $\mathbb{E}[N_t^{(k+1)p}] < \infty$ if $\mathbb{E}[u_0(X_t)^{-qk}] < \infty$ for $\frac{1}{p} + \frac{1}{q} = 1$, $(p, q > 1)$, by Hölder's Inequality and Doob's L^p Inequality.

2. The assumption that $\mathbb{E}\left[u_0(X_t)^{-k}\right] < \infty$ is reasonable for many applications. Take, for example, the case when X is Brownian motion, and there are constants $c, C > 0$ such that for sufficiently large $\|x\|$ we have

$$u_0(x) \geq Ce^{-c\|x\|}.$$

Importantly, however, Theorem B.2.1 is not suited to initial conditions such as the Heaviside function.

3. Theorem B.2.1 can be adapted to the case when $k < 0$ however, one additional complication arises in showing that N^* is admissible and an assumption on u_0 is necessary.

B.3. Simple Consequences

Using the dynamic programming method, we can give more intuition behind the bounds that we used in Chapter 4 which were proven in Corollary 2.6.1. In particular we have the following proposition.

Proposition B.3.1. *Let u be the solution to equation* (B.1.1)*. Then,*

$$\mathbb{E}\left[\frac{e^{s/k}u_0(x+X_t)}{((e^s-1)u_0(x+X_t)^k+1)^{1/k}}\right] \leq u(t,x) \leq \frac{e^{s/k}\mathbb{E}[u_0(x+X_t)]}{((e^s-1)\mathbb{E}[u_0(x+X_t)]^k+1)^{1/k}}$$

We will see that these bounds is come from solving an ODE which corresponds to a trivial HJB equation when there is no random process.

Proof. 1. **Upper Bound on u**

Recall the notation,

$$\mathbb{Y} = \{Y : Y : [0,t] \to \mathbb{R}, \text{is absolutely continuous, decreasing and } Y_0 = 1\}. \tag{B.3.1}$$

We will obtain the result by using Theorem 2.2.3 and minimising over a subset and superset of set

$$\mathcal{Y} = \{Y \in \mathbb{Y} : Y \text{ is adapted to the filtration } \mathcal{F}\}$$

If h solves the ODE

$$\frac{\mathrm{d}h}{\mathrm{d}s} = \frac{1}{k}h(1-h^k) \tag{B.3.2}$$

then,

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\int_0^s e^{r/k}|\dot{Y}_r|^{1+1/k}\mathrm{d}r + e^{s/k}|Y_s|^{1+1/k}h(t-s)\right) \geq 0$$

with equality for $Y_s = \exp(-\int_0^s h(t-s)^k\mathrm{d}s)$. Thus,

$$\int_0^t e^{r/k}|\dot{Y}_r|^{1+1/k}\mathrm{d}r + e^{t/k}|Y_t|^{1+1/k}h(0) \geq h(t)$$

We can solve for h explicitly and then we see that

$$h(s) = \frac{e^{s/k}h(0)}{((e^s-1)h(0)^k+1)^{1/k}}$$

Let $\overline{\mathcal{Y}} = \{y \in \mathbb{Y} : Y \text{ is } \mathcal{F}_0\text{-measurable}\}$. This defines the set of admissible, deterministic processes.

$$u(t,x) \leq \min_{\overline{\mathcal{Y}}} \left(\int_0^t e^{s/k}|\dot{Y}_s|^{1+1/k}\mathrm{d}s + e^{t/k}|Y_t|^{1+1/k}\mathbb{E}[u_0(x+X_t)] \right)$$

and so setting $Y_s = \exp(-\int_0^s h(t-s)^k \mathrm{d}s)$ with $h(0) = \mathbb{E}[u_0(x+X_t)]$ we see that

$$u(t,x) \leq \frac{e^{s/k}\mathbb{E}[u_0(x+X_t)]}{((e^s-1)\mathbb{E}[u_0(x+X_t)]^k+1)^{1/k}}$$

2. **Lower Bound on u**

We can also use this for a lower bound. Set

$$\underline{\mathcal{Y}} = \{Y \in \mathbb{Y} : Y_s \in \mathcal{F}_t \text{ for all } s\} \tag{B.3.3}$$

Then since $\mathcal{Y} \subset \underline{\mathcal{Y}}$, we have

$$\begin{aligned} u(t,x) &\geq \min_{\underline{\mathcal{Y}}} \mathbb{E}\left[\int_0^t e^{s/k}|\dot{Y}_s|^{1+1/k}\mathrm{d}s + e^{t/k}|Y_t|^{1+1/k}u_0(x+X_t)\right] \\ &\geq \mathbb{E}\left[\frac{e^{s/k}u_0(x+X_t)}{((e^s-1)u_0(x+X_t)^k+1)^{1/k}}\right] \end{aligned} \tag{B.3.4}$$

The second inequality is an equality for $Y_s = \exp(-\int_0^s h(t-s)^k \mathrm{d}s)$ with $h(0) = u_0(x+X_t)$ with $Y \in \underline{\mathcal{Y}}$.

□

Bibliography

[AN72] Athreya, K.B. ; Ney, P.E.: *Branching processes*. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Springer-Verlag, 1972

[App04] Applebaum, D.: *Lévy processes and Stochastic Calculus*. Cambridge University Press, 2004

[AW75] Aronson, D. G. ; Weinberger, H. F.: Nonlinear diffusion in population genetics, combustion, and nerve pulse propagation. In: *Partial differential equations and related topics*. Springer Berlin Heidelberg, 1975 (Lecture Notes in Mathematics 446), S. 5–49

[AW78] Aronson, D. G. ; Weinberger, H. F.: Multidimensional nonlinear diffusion arising in population genetics. In: *Advances in Mathematics* 30 (1978), Nr. 1, S. 33–76

[BB16] Berestycki, J. ; Brunet, É.: A note of the convergence of the Fisher-KPP front centred around its α-level. In: *arXiv:1603.06005 [math]* (2016)

[BD09] Brunet, É. ; Derrida, B.: Statistics at the tip of a branching random walk and the delay of traveling waves. In: *EPL (Europhysics Letters)* 87 (2009), Nr. 6

[Big92] Biggins, J. D.: Uniform convergence of martingales in the branching random walk. In: *The Annals of Probability* 20 (1992), Nr. 1, S. 137–151

[Big95] Biggins, J. D.: The growth and spread of the general branching random walk. In: *The Annals of Applied Probability* 5 (1995), Nr. 4, S. 1008–1024

[Big97] Biggins, J. D.: How fast does a general branching random walk spread? In: *Classical and Modern Branching Processes*. Springer New York, 1997 (The IMA Volumes in Mathematics and its Applications 84), S. 19–39

[BLSW91] Biggins, J. D. ; Lubachevsky, B. D. ; Shwartz, A. ; Weiss, A.: A branching random walk with a barrier. In: *The Annals of Applied Probability* 1 (1991), Nr. 4, S. 573–581

[Bra78] Bramson, M. D.: Maximal displacement of branching Brownian motion. In: *Communications on Pure and Applied Mathematics* 31 (1978), Nr. 5, S. 531–581

[Bra83] Bramson, Maury: *Convergence of Solutions of the Kolmogorov Equation to Travelling Waves*. American Mathematical Soc., 1983

[Bra86] Bramson, M.: Location of the travelling wave for the Kolmogorov equation. In: *Probability Theory and Related Fields* 73 (1986), Nr. 4, S. 481–515

[CC04] Carr, J. ; Chmaj, A.: Uniqueness of travelling waves for nonlocal monostable equations. In: *Proceedings of the American Mathematical Society* 132 (2004), Nr. 8, S. 2433–2439

[Cha91] Chauvin, B.: Product martingales and stopping lines for branching Brownian motion. In: *Annals of Probability* 30 (1991), S. 1195–1205

[CHT+95] Champneys, A. ; Harris, S. ; Toland, J. ; Warren, J. ; Williams, D.: Algebra, Analysis and Probability for a coupled system of reaction-diffusion equations. In: *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 350 (1995), Nr. 1692, S. 69–112

[CNCL03] Castillo-Negrete, D. del ; Carreras, B. A. ; Lynch, V. E.: Front dynamics in reaction-diffusion systems with Lévy flights: a fractional diffusion approach. In: *Physical Review Letters* 91 (2003), Nr. 1

[CR88] Chauvin, B. ; Rouault, A.: KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees. In: *Probability Theory and Related Fields* 80 (1988), Nr. 2, S. 299–314

[CR13] Cabré, X. ; Roquejoffre, J.-M.: The influence of fractional diffusion in Fisher-KPP equations. In: *Communications in Mathematical Physics* 320 (2013), Nr. 3, S. 679–722

[DS88] Derrida, B. ; Spohn, H.: Polymers on disordered trees, spin glasses, and traveling waves. In: *Journal of Statistical Physics* 51 (1988), Nr. 5-6, S. 817–840

[DT18a] Driver, D.P. ; Tehranchi, M. R.: An optimisation-based representation for reaction-diffusion equations. In: *arXiv:1803.09688 [math.AP]* (2018)

[DT18b] Driver, D.P. ; Tehranchi, M. R.: An optimisation-based representation for reaction-diffusion equations II. In: *(preprint)* (2018)

[EH13] Embrechts, P. ; Hofert, M.: A note on generalized inverses. In: *Mathematical Methods of Operations Research* 77 (2013), Juni, Nr. 3, S. 423–432

[ES98] Ebert, U. ; Saarloos, W. van: Universal algebraic relaxation of fronts propagating into an unstable state and implications for moving boundary approximations. In: *Physical Review Letters* 80 (1998), S. 1650–1653

[Fis37] Fisher, R. A.: The wave of advance of advantageous genes. In: *Annals of Eugenics* 7 (1937), Nr. 4, S. 355–369

[Fre85] Freidlin, M. I.: *Functional integration and partial differential equations.* Princeton University Press, 1985

[Gär82] Gärtner, J.: Location of wave fronts for the multidimensional KPP equation and Brownian first exit densities. In: *Mathematische Nachrichten* 105 (1982), S. 317–351

[GJ16] Groisman, P. ; Jonckheere, M.: Front propagation and quasi-stationary distributions for one-dimensional Lévy processes. In: *arXiv:1609.09338 [math]* (2016)

[GS01] Grimmett, G. ; Stirzaker, D.: *Probability and Random Processes.* OUP Oxford, 2001

[Har98] Harris, S. C.: Travelling-waves for the FKPP equation via probabilistic arguments. In: *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* 129 (1998), Nr. 03, S. 503 – 517

[HNRR13] Hamel, F. ; Nolen, J. ; Roquejoffre, J.-M. ; Ryzhik, L.: A short proof of the logarithmic Bramson correction in Fisher-KPP equations. In: *Networks and Heterogeneous Media* 8 (2013), Nr. 1, S. 275–289

[HR10] Hamel, F. ; Roques, L.: Fast propagation for KPP equations with slowly decaying initial conditions. In: *Journal of Differential Equations* 249 (2010), Nr. 7, S. 1726–1745

[IW89] Ikeda, N. ; Watanabe, S.: *Stochastic differential equations and diffusion processes.* North-Holland, 1989

[Kat78] Kato, T.: Trotters product formula for an arbitrary pair of self-adjoint contraction semigroups. In: *Topics in functional analysis (essays dedicated to MG Kren on the occasion of his 70th birthday)* 3 (1978), S. 185–195

[KPP37] Kolmogorov, A. ; Piscounov, N. ; Petrovsky, I.: Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application a un problème biologique. In: *Moscow University Bulletin Of Mathematics* 1 (1937), S. 1–25

[Kyp99] Kyprianou, A. E.: A note on branching Lévy processes. In: *Stochastic Processes and their Applications* 82 (1999), Nr. 1, S. 1–14

[Kyp04] Kyprianou, A. E.: Travelling wave solutions to the K-P-P equation: alternatives to Simon Harris' probabilistic analysis. In: *Annales de l'Institut Henri Poincare (B) Probability and Statistics* 40 (2004), Nr. 1, S. 53–72

[Lar78] Larson, D. A.: Transient bounds and time-asymptotic behavior of solutions to nonlinear equations of Fisher type. In: *SIAM Journal on Applied Mathematics* 34 (1978), Nr. 1, S. 93–103

[Lau85] Lau, K.-S.: On the nonlinear diffusion equation of Kolmogorov, Petrovsky, and Piscounov. In: *Journal of Differential Equations* 59 (1985), S. 44–70

[McK75] McKean, H. P.: Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov. In: *Communications on Pure and Applied Mathematics* 28 (1975), Nr. 3, S. 323–331

[McK76] McKean, H. P.: A correction to: Application of Brownian motion to the equation of Kolmagorov-Petrovski-Piskonov. In: *Communications on Pure and Applied Mathematics* 29 (1976), Nr. 5, S. 553–554

[Mer69] Merton, R. C.: Lifetime portfolio selection under uncertainty: the continuous-time case. In: *The Review of Economics and Statistics* 51 (1969), S. 247–57

[Mer71] Merton, R. C.: Optimum consumption and portfolio rules in a continuous-time model. In: *Journal of Economic Theory* (1971), S. 373–413

[Mon67] Montroll, E. W.: On nonlinear processes involving population growth and diffusion. In: *Journal of Applied Probability* 4 (1967), Nr. 2, S. 281–290

[MP85] Mehra, R. ; Prescott, E.: The equity premium: a puzzle. In: *Journal of Monetary Economics* (1985), S. 145–161

[MP03] Munier, S. ; Peschanski, R.: Geometric scaling as traveling waves. In: *Physical Review Letters* 91 (2003), Nr. 23

[Mun09] Munier, S.: Quantum chromodynamics at high energy and statistical physics. In: *Physics Reports* 473 (2009), Nr. 1, S. 1–49

[Mur07] Murray, J. D.: *Mathematical Biology: I. An introduction.* Springer Science & Business Media, 2007

[Nev88] Neveu, J.: Multiplicative Martingales for Spatial Branching Processes. In: *Seminar on Stochastic Processes, 1987.* Birkhäuser Boston, 1988 (Progress in probability and statistics 15), S. 223–242

[NRR16] Nolen, J. ; Roquejoffre, J.-M. ; Ryzhik, L.: Refined long time asymptotics for Fisher-KPP fronts. In: *arXiv:1607.08802 [math]* (2016)

[NRR17] Nolen, J. ; Roquejoffre, J.-M. ; Ryzhik, L.: Convergence to a single wave in the Fisher-KPP equation. In: *Chinese Annals of Mathematics, Series B* 38 (2017), Nr. 2, S. 629–646

[Paz92] Pazy, A.: *Semigroups of linear operators and applications to partial differential equations.* Springer New York, 1992

[Ros74] Rosen, G.: Approximate solution to the generic initial value problem for nonlinear reaction-diffusion equations. In: *SIAM Journal on Applied Mathematics* 26 (1974), Nr. 2, S. 221–224

[Saa03] Saarloos, W. van: Front propagation into unstable states. In: *Physics Reports* 386 (2003), Nr. 2-6, S. 29–222

[Sch80] Schumacher, K.: Travelling-front solutions for integro-differential equations. I. In: *Journal für die reine und angewandte Mathematik* 316 (1980), S. 54–70

[Sko64] Skorokhod, A. V.: Branching diffusion processes *(Russian)*. In: *Theory of Probability & Its Applications* 9 (1964), Nr. 3, S. 445–449

[SP12] Schilling, R. L. ; Partzsch, L.: *Brownian motion: an introduction to stochastic processes*. De Gruyter, 2012

[ST94] Samorodnitsky, G. ; Taqqu, M. S.: *Stable Non-Gaussian random processes: stochastic models with infinite variance*. CRC Press, 1994

[Uch78] Uchiyama, K.: The behavior of solutions of some non-linear diffusion equations for large time. In: *Journal of Mathematics of Kyoto University* 18 (1978), Nr. 3, S. 453–508

[Wat68] Watanabe, S.: A limit theorem of branching processes and continuous state branching processes. In: *Journal of Mathematics of Kyoto University* 8 (1968), Nr. 1, S. 141–167

[Zar01] Zariphopoulou, T.: A solution approach to valuation with unhedgeable risks. In: *Finance and Stochastics* (2001), S. 61–82