

# *Astérisque*

S. C. HARRIS

D. WILLIAMS

**Large deviations and martingales for a typed  
branching diffusion, 1**

*Astérisque*, tome 236 (1996), p. 133-154

[http://www.numdam.org/item?id=AST\\_1996\\_\\_236\\_\\_133\\_0](http://www.numdam.org/item?id=AST_1996__236__133_0)

© Société mathématique de France, 1996, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (<http://smf4.emath.fr/Publications/Asterisque/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

# Large deviations and martingales for a typed branching diffusion, 1

S. C. Harris, D. Williams

**Abstract.** — We study a certain family of typed branching diffusions where the type of each particle moves as an Ornstein-Uhlenbeck process and binary branching occurs at a rate quadratic in the particle's type. We calculate the 'left-most' particle speed for the branching process explicitly, aided by close connections with harmonic oscillator theory. The behaviour of the system changes markedly below a certain critical temperature parameter.

In the high-temperature regime, the study of various 'additive' martingales and their use in a change of measure method provides the proof of the almost sure speed of spread of the particle system.

Also, we briefly mention how to use the martingale results of the branching diffusion model in representations of travelling-wave solutions for the associated reaction-diffusion equation.

## 1. Introduction

Our aim is to produce a series of papers on a certain family of typed branching diffusions each with rich structure. The present paper introduces the simplest (binary-branching) model and (except for a 'sneak preview' of the critical-temperature phase in the Section 9) studies this model only in the high-temperature phase in which there is a high degree of ergodicity. Here, many standard methods are applicable, though we have been able to carry them through only because the model's close relation to the harmonic oscillator allows explicit calculations; the first calculations also have a long history in probability going back to Cameron and Martin – see Sections 5.13-5.15 of Itô and McKean (1965). Some of the calculations necessary for our approach are rather complicated; and these are only sketched here – see Harris (1995) and Harris and Williams (1995) for more details. We deal with the substantive points of rigour, but skip some details of rigour to keep the text to an appropriate length.

We begin by recalling how certain 'linear' expectations for the branching process may be calculated by considering a one-particle system, and we derive certain martingale properties. We then study in some detail the large-deviation heuristics

for the problem, emphasizing the rôle of Legendre transformations. Next, we use a method of Neveu to establish uniform-integrability properties of certain martingales; this requires calculation of an expectation which reflects the *non*-linearity of the system, and we are saved only because Meyer's *opérateur carré du champ* behaves well. By exploiting a change-of-measure technique ('exponential tilting' in the exotic/quixotic terminology of statisticians), we *prove* the results suggested by large-deviation theory. The martingale methods have the significant bonus that they imply existence of monotone travelling waves associated with the model. In the present context, it may not be easy to establish the existence of these waves by analysis. Neveu's method of proving *lack* of uniform integrability of certain martingales will be important for proving uniqueness in some cases, non-existence in others, for monotone travelling waves. This idea is developed in full for a simpler problem in Champneys, Harris, Toland, Warren and Williams (1995); in the present context, it requires difficult *a priori* estimates. We also refer the reader to the Champneys *et al* paper for a list of references to which the present paper is equally indebted.

Further study of the high-temperature regime is made in Harris (1995), and will be continued in other joint papers. The changes of measure have some bizarre features which we wish to discuss further, bringing in important ideas from Chauvin and Rouault (1988, 1990). The long-term behaviour of the 'Gibbs-Boltzmann' measure  $J_\lambda(t)$  which assigns mass  $J_\lambda(t, k)$  as at (6.1) to the point  $(X_k(t), Y_k(t))$  is the most fascinating aspect of the high-temperature phase. Note that the fundamental martingale  $Z_\lambda^-(t)$  gives the 'partition function'. The study of the long-term behaviour of  $J_\lambda$  is closely related to that of the 'excited-state' martingales for our system.

A major challenge for the binary-branching model is the low-temperature regime ( $\theta < 8r$ ) in which all of the methods used here fail: the expected number of particles in a region blows up, though the number of particles remains almost surely finite. Other models present other challenges.

## 2. The Branching Model

We consider a typed branching diffusion where, for time  $t \geq 0$ ,

- $N(t)$  is the number of particles alive,
- $X_k(t)$  in  $\mathbb{R}$  is the spatial position of the  $k^{\text{th}}$ -born particle,
- $Y_k(t)$  in  $\mathbb{R}$  is the 'type' of the  $k^{\text{th}}$ -born particle,

$(N(t); X_1(t), \dots, X_{N(t)}; Y_1(t), \dots, Y_{N(t)})$  is the current state of the particle system.

The *type* moves on the real line as an Ornstein-Uhlenbeck process associated with the differential operator (generator)

$$\mathcal{Q}_\theta := \frac{\theta}{2} \left( \frac{\partial^2}{\partial y^2} - y \frac{\partial}{\partial y} \right)$$

where  $\theta$  is a positive real parameter considered as the *temperature* of the system. The *spatial* motion of a particle of type  $y$  is a driftless Brownian motion with variance

$$A(y) := ay^2, \quad \text{where } a \geq 0.$$

The *breeding* of a type  $y$  particle occurs at a rate

$$R(y) := ry^2 + \rho, \quad \text{where } r, \rho \geq 0,$$

and we have one child born at these times (binary splitting). A child inherits its parent's current type and (spatial) position then moves off *independently* of all others. Particles live forever (once born!).

Let  $\mathbb{P}^{x,y}$  and  $\mathbb{E}^{x,y}$  represent probability and expectation when the process starts from  $(N; \mathbf{X}, \mathbf{Y}) = (1; x; y)$ .

For starting point  $(N; \mathbf{X}; \mathbf{Y}) = (1; 0; 0)$ , we have

$$\mathbb{P}^{0,0}(N(t) = 1 \mid \sigma(Y_1(s) : s \leq t)) = \exp\left(-\int_0^t R(Y_1(s)) ds\right),$$

and on the set  $\{N(t) \geq k\}$  we have

$$(2.1) \quad \mathbb{P}^{0,0}\left(X_k(t) \in F \mid \sigma(N(s), \mathbf{Y}(s) : s \leq t)\right) \\ = \int_F \left\{ 2\pi \int_0^t A(Y_k(s)) ds \right\}^{-\frac{1}{2}} \exp\left(-\frac{x^2}{2 \int_0^t A(Y_k(s)) ds}\right) dx,$$

where  $Y_k(s)$  is the type of the unique 'ancestor' alive at  $s$  of the  $k$ -th particle alive at time  $t$ .

We are going to consider  $r, \rho, a$  as fixed, and look at the effects of changing the temperature  $\theta$ . One of our main concerns is: what is 'the velocity of the leftmost particle'; to be precise, what is the value of

$$\text{Vel} := \lim_{t \rightarrow \infty} L(t)/t$$

(we prove that the almost sure limit does exist), where

$$L(t) := \inf_{1 \leq k \leq N(t)} X_k(t)?$$

The temperature controls the balance of competition between the ergodic mixing of the Ornstein-Uhlenbeck process (which increases with  $\theta$ ) and the large breeding rate and large diffusion coefficient for the  $X$ -motion away from the type-origin. This is reflected in the answer

$$\text{Vel} = -\tilde{c}(\theta),$$

where

$$(2.2) \quad \tilde{c}(\theta)^2 = \begin{cases} 2a \left( r + \rho + \frac{2(2r + \rho)^2}{\theta - 8r} \right) & \text{for } \theta > 8r, \\ +\infty & \text{for } \theta \leq 8r. \end{cases}$$

When  $\theta$  is very large, the system may be approximated by a ‘mean field’ model in which  $A(Y)$  is replaced by its mean  $a$  and  $R(Y)$  by its mean  $r + \rho$  under the (standard normal) invariant law of the type process.

In all but the last section of this paper,

*we assume that  $\theta > 8r$ .*

The challenging low-temperature cases and many other things are left to other occasions.

### 3. Calculations using the One-Particle System

Let  $(\xi, \eta)$  be a process behaving like a single particle’s space and type motions in the branching model described above. Thus,  $\xi$  is a Brownian motion controlled by an Ornstein-Uhlenbeck process  $\eta$ , and  $(\xi, \eta)$  has formal generator  $\mathcal{H}$ , where

$$(\mathcal{H}F)(x, y) = \frac{1}{2}A(y)\frac{\partial^2 F}{\partial x^2} + (\mathcal{Q}_\theta F)(x, y) = \frac{1}{2}A(y)\frac{\partial^2 F}{\partial x^2} + \frac{\theta}{2}\left(\frac{\partial^2 F}{\partial y^2} - y\frac{\partial F}{\partial y}\right).$$

Of course,  $\eta$  is an autonomous Markov process with generator  $\mathcal{Q}_\theta$  and with (standard normal) invariant density

$$\phi(y) := (2\pi)^{-\frac{1}{2}} \exp(-\frac{1}{2}y^2).$$

For functions  $h_1, h_2$  on  $\mathbb{R}$ , we define the  $L^2(\phi)$  inner product:

$$\langle h_1, h_2 \rangle_\phi := \int_{\mathbb{R}} h_1(y)h_2(y)\phi(y)dy.$$

The following principle is used repeatedly.

(3.1) LEMMA: ‘From One to Many’. *For any non-negative Borel function  $f$  on  $\mathbb{R} \times \mathbb{R}$ , we have*

$$\mathbb{E}^{x,y} \left( \sum_{k=1}^{N(t)} f(X_k(t), Y_k(t)) \right) = \mathbb{E}^{x,y} \left( \exp \left( \int_0^t R(\eta_s) ds \right) f(\xi_t, \eta_t) \right).$$

This principle is often combined with a change-of-measure formula for Ornstein-Uhlenbeck processes. We use  $\text{OU}(\theta, \mu)$  to represent an Ornstein-Uhlenbeck process with variance  $\theta$  and drift parameter  $\mu$ , thus with generator  $\frac{\theta}{2}\frac{\partial^2}{\partial y^2} - \mu y\frac{\partial}{\partial y}$ .

(3.2) LEMMA: ‘Change of Measure between OU processes’. Let  $\eta$  be an  $\text{OU}(\theta, \mu)$  under  $\mathbb{P}_\mu$ . For  $\delta \in \mathbb{R}$ , we can define a new probability measure  $\mathbb{P}_\delta$ , equivalent to  $\mathbb{P}_\mu$  on every  $\mathcal{F}_t$ , via the Radon-Nikodým derivative

$$\frac{d\mathbb{P}_\delta}{d\mathbb{P}_\mu} = \exp\left(\frac{(\mu - \delta)}{2\theta}(\eta_t^2 - \theta t) + \frac{(\mu^2 - \delta^2)}{2\theta} \int_0^t \eta_s^2 ds - \frac{(\mu - \delta)}{2\theta} \eta_0^2\right) \quad \text{on } \mathcal{F}_t.$$

Under  $\mathbb{P}_\delta$ ,  $\eta$  is an  $\text{OU}(\theta, \delta)$  process.

Let us remind ourselves of how these results may be combined. Define

$$(3.3) \quad \lambda_{\min} := -\sqrt{\frac{\theta - 8r}{4a}}.$$

Let  $\alpha, \lambda, \mu \in \mathbb{R}$ , with the following convention which we always use for  $\lambda$ :

$$(3.4) \quad \lambda_{\min} < \lambda < 0.$$

Suppose for example that we wish to calculate (for a positive Borel function  $h$ )

$$(3.5) \quad \text{Required} := \mathbb{E}^{0,y} \sum_{k=1}^{N(t)} \exp\{\alpha Y_k(t)^2 + \lambda X_k(t)\} h(Y_k(t)).$$

We find that

$$\begin{aligned} \text{Required} &= \mathbb{E}_{\frac{1}{2}\theta}^y \exp\left\{\alpha \eta_t^2 + \frac{1}{2}\lambda^2 \int_0^t a \eta_s^2 ds + \int_0^t (r \eta_s^2 + \rho) ds\right\} h(\eta_t) \\ &= \mathbb{E}_\mu^y \exp\{\text{Quad}\} h(\eta_t), \end{aligned}$$

where

$$(3.6) \quad \text{Quad} := \alpha \eta_t^2 - \psi^-(\eta_t^2 - \theta t) + \left(\frac{1}{2}a\lambda^2 + \frac{\mu^2 - \frac{1}{2}\theta^2}{2\theta} + r\right) \int_0^t \eta_s^2 ds + \rho t + \psi^- y^2,$$

where

$$\psi^- := (\frac{1}{2}\theta - \mu)/2\theta.$$

We now choose  $\mu$  so that the coefficient of the integral on the right-hand side of (3.6) is zero:

$$(3.7) \quad \mu = \mu_\lambda := \frac{1}{2}\sqrt{\theta^2 - \theta(8r + 4a\lambda^2)},$$

and we write  $\psi_\lambda^-$  for  $\psi^-$  with this  $\mu$ : indeed, we shall write

$$\psi_\lambda^\pm := \frac{\theta \pm \sqrt{\theta^2 - \theta(8r + 4a\lambda^2)}}{4\theta},$$

Then

$$(3.8) \quad \text{Required} = \exp [(\rho + \theta\psi_\lambda^-)t + \psi_\lambda^- y^2] \mathbb{E}_{\mu_\lambda}^y \exp \{(\alpha - \psi_\lambda^-)\eta_t^2\} h(\eta_t).$$

However, it is well known that, for the  $\mathbb{P}_\mu^y$  law,  $\eta_t$  has the normal distribution

$$(3.9) \quad N \left( e^{-\mu t} y, \frac{\theta(1 - e^{-2\mu t})}{2\mu} \right),$$

so that the expression 'Required' is easily calculated explicitly: if  $h \equiv 1$ , it will be finite for all finite  $t$  if and only if  $\alpha \leq \psi_\lambda^+$ .

In particular, on taking  $\alpha = \psi_\lambda^-$  and  $h \equiv 1$ , and making obvious use of the branching property, we obtain the following lemma.

(3.10) LEMMA. For

$$\lambda_{\min} < \lambda < 0,$$

the expression

$$(3.11) \quad Z_\lambda^-(t) := \sum_{k=1}^{N(t)} \exp \{ \psi_\lambda^- Y_k(t)^2 + \lambda [X_k(t) + c_\lambda^- t] \},$$

where

$$(3.12) \quad \psi_\lambda^- := \frac{\theta - \sqrt{\theta^2 - \theta(8r + 4a\lambda^2)}}{4\theta},$$

$$(3.13) \quad c_\lambda^- := -(\rho + \theta\psi_\lambda^-) / \lambda,$$

defines a martingale  $Z_\lambda^-$  (under each  $\mathbb{P}^{x,y}$  measure).

Since  $\lambda < 0$  and the non-negative martingale  $Z_\lambda^-$  converges, we must have

$$\liminf [X_k(t) + c_\lambda^- t] > -\infty, \quad \text{a.s.},$$

whence

$$\liminf t^{-1} L(t) \geq -c_\lambda^-, \quad \text{a.s.},$$

and we have the lower bound:

$$(3.14) \quad \liminf t^{-1} L(t) \geq -\tilde{c}(\theta) := -\inf \{ c_\lambda^- : \lambda_{\min} < \lambda < 0 \}.$$

This formula for  $\tilde{c}(\theta)$  agrees with that at (2.2).

It is easy to check that the function  $c^-$  is convex on  $(\lambda_{\min}, 0)$ , and achieves its minimum at a unique point  $\tilde{\lambda}(\theta)$ . We shall prove later that the martingale  $Z_{\tilde{\lambda}}^-$  is uniformly integrable for  $\lambda \in (\tilde{\lambda}(\theta), 0)$ , and use this in obtaining the upper bound

$$(3.15) \quad \limsup t^{-1} L(t) \leq -\tilde{c}(\theta), \quad \text{a.s.}$$

#### 4. Large-Deviation Heuristics.

(4.1) *Feynman-Kac heuristics.* We wish to find the rate of growth in numbers of particles out along rays in space-time, that is, we wish to calculate

$$(4.2) \quad \Delta(\gamma) := \lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E} \left( \sum_{k=1}^{N(t)} \mathbb{I} \{X_k(t) \leq -\gamma t\} \right) \quad (\gamma \geq 0).$$

Large-deviation theory makes us conjecture the existence of the limit  $\Delta(\gamma)$  and also that

$$(4.3) \quad \begin{aligned} \Delta(\gamma) &= \inf_{\lambda < 0} \lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E} \sum_{k=1}^{N(t)} \exp(\lambda \gamma t) \exp\{\lambda X_k(t)\} \\ &= \inf_{\lambda < 0} \lim_{t \rightarrow \infty} t^{-1} \log \left\{ \exp(\lambda \gamma t) \mathbb{E} \exp \left( \int_0^t \{ \frac{1}{2} \lambda^2 A + R \} (\eta(s)) ds \right) \right\} \\ &= \inf_{\lambda < 0} \{ E(\lambda) + \lambda \gamma \}, \end{aligned}$$

where, by the Feynman-Kac formula,  $E(\lambda)$  is the rightmost eigenvalue of the self-adjoint operator on  $L^2(\phi)$  defined by

$$\mathcal{L}_\lambda := \mathcal{Q}_\theta + \frac{1}{2} \lambda^2 A(y) + R(y).$$

We find that for

$$\lambda_{\min} := -\sqrt{\frac{\theta - 8r}{4a}} < \lambda < 0,$$

we have

$$(4.4) \quad E(\lambda) = \rho + (\theta - \sqrt{\theta(\theta - 8r - 4a\lambda^2)})/4 = \rho + \theta \psi_\lambda^- = -\lambda c_\lambda^-.$$

The expression  $E(\lambda) + \lambda \gamma$  is minimized when  $\lambda = \lambda_\gamma$ , where

$$(4.5) \quad E'(\lambda_\gamma) = -\gamma, \quad \text{whence } \lambda_\gamma = -\sqrt{\frac{\gamma^2(\theta - 8r)}{a(4\gamma^2 + a\theta)}};$$

and hence

$$(4.6) \quad \Delta(\gamma) = \rho + (\theta - \sqrt{a^{-1}(\theta - 8r)(4\gamma^2 + \theta a)})/4.$$

It is now tempting to guess that  $\tilde{c}(\theta)$  is given by

$$(4.7) \quad \begin{aligned} \tilde{c}(\theta) &= \sup\{\gamma : \Delta(\gamma) > 0\} = \inf\{-E(\lambda)/\lambda : \lambda_{\min} < \lambda < 0\} \\ &= \inf\{c_\lambda^- : \lambda_{\min} < \lambda < 0\}, \end{aligned}$$

in agreement with what we had previously. This guess that, *in this particular situation*, ‘expectation’ and ‘particle’ wave-speeds agree, is *proved* rigorously by martingale techniques in Section 6.

Since  $\mathcal{L}_\lambda$  is self-adjoint relative to  $\langle \cdot, \cdot \rangle_\phi$ , we have

$$(4.8) \quad E(\lambda) = \sup_g \{\langle g, \mathcal{L}_\lambda g \rangle_\phi : \langle g, g \rangle_\phi = 1\} = \sup_g \{U(\lambda; g) : \langle g, g \rangle_\phi = 1\},$$

where

$$U(\lambda; g) := \langle Rg, g \rangle_\phi - \frac{1}{2}\theta \langle g', g' \rangle_\phi + \frac{1}{2}\lambda^2 \langle Ag, g \rangle_\phi.$$

Moreover, the supremum at (4.8) will be obtained at the eigenfunction corresponding to  $E(\lambda)$ :

$$(4.9) \quad g_\lambda^*(y) = (2\mu_\lambda/\theta)^{\frac{1}{2}} \exp\{\psi_\lambda^- y^2\}.$$

Our arguments have therefore suggested the formula

$$(4.10) \quad \Delta(\gamma) = \inf_{\lambda < 0} \sup_g L_\gamma(\lambda; g),$$

where

$$(4.11) \quad L_\gamma(\lambda; g) := \langle Rg, g \rangle_\phi - \frac{1}{2}\theta \langle g', g' \rangle_\phi + \frac{1}{2}\lambda^2 \langle Ag, g \rangle_\phi + \lambda\gamma.$$

(4.12) *Discussion.* It is helpful to note that Mercer’s Theorem applied to the Feynman-Kac semigroup appearing in (4.3) gives (as  $t \rightarrow \infty$ )

$$(4.13) \quad \mathbb{E}^{0,y} \sum_{k=1}^{N(t)} e^{\lambda X_k(t)} I(Y_k(t) \in dz) \sim \exp[tE(\lambda)] g_\lambda^*(y) g_\lambda^*(z) \phi(z) dz,$$

equivalently,

$$(4.14) \quad \mathbb{E}^{0,y} \sum_{k=1}^{N(t)} e^{\lambda X_k(t)} h(Y_k(t)) \sim \exp[tE(\lambda)] g_\lambda^*(y) \langle g_\lambda^*, h \rangle_\phi.$$

Note how this tallies with the martingale property of  $Z_\lambda^-$ .

However, (3.8) and (3.9) with  $\alpha = 0$  give the *exact* form of (4.13), and imply in particular that

$$(4.15) \quad \mathbb{E}^{0,0} \sum_{k=1}^{N(t)} \exp[\lambda X_k(t)] = \{1 + \theta\psi_\lambda^-(1 - e^{-2\mu_\lambda t})/\mu_\lambda\}^{-\frac{1}{2}} \exp\{(\rho + \theta\psi_\lambda^-)t\}.$$

The existence of  $\Delta(\gamma)$  at (4.2) and proof that  $\Delta(\gamma)$  is given by (4.3) and (4.10) may be obtained from this by saddle-point techniques.

(4.16) *Heuristics based on the rate-functional for occupation densities.* We now link the above heuristics to the dual approach which more properly belongs to large-deviation theory. One of the great Donsker-Varadhan theorems (see, for example, Deuschel and Stroock (1989)) makes precise the idea that the probability that a process  $\eta$  with generator  $\mathcal{Q}_\theta$  will have occupation density  $tg^2\phi$  by time  $t$  (where, of course,  $\langle g, g \rangle_\phi = 1$ ) is roughly

$$\exp(-tI(g)), \quad \text{where } I(g) := \frac{1}{2}\theta\langle g', g' \rangle_\phi.$$

(For ergodic self-adjoint processes, the rate functional for occupation densities relative to the invariant measure agrees with the Dirichlet norm.) Thus, we guess the expected number of particles with type histories having densities  $tg^2$  with respect to  $\phi$  by time  $t$  to be roughly

$$\exp\left(t\left\{\langle Rg, g \rangle_\phi - \theta I(g)\right\}\right),$$

and, by (2.1), the expected number of *these* particles with  $X$ -values at time  $t$  near to  $-\gamma t$  should be about

$$\exp\left(t\left\{\langle Rg, g \rangle_\phi - \theta I(g) - \frac{1}{2}\gamma^2/\langle Ag, g \rangle_\phi\right\}\right).$$

By Laplace-Varadhan asymptotics, the expected *total* number of particles with  $X$ -values near  $-\gamma t$  is roughly

$$\exp\left(t \sup_g \left\{ \langle Rg, g \rangle_\phi - \theta I(g) - \frac{1}{2}\gamma^2/\langle Ag, g \rangle_\phi : \langle g, g \rangle_\phi = 1 \right\}\right)$$

However, from (4.11), we see that

$$(4.17) \quad \inf_{\lambda < 0} L_\gamma(\lambda, g) = \langle Rg, g \rangle_\phi - \theta I(g) - \frac{1}{2}\gamma^2/\langle Ag, g \rangle_\phi,$$

so that we are in the usual situation where the key thing to linking the dual pictures is to be able to interchange the sup and inf.

(4.18) *Duality.* Let us expand on the duality. See Harris (1995) for details. The optimizing  $\bar{g}_\gamma$  which gives the supremum of the right-hand side of (4.17) is

$$(4.19) \quad \bar{g}_\gamma := \left( \frac{a(\theta - 8r)}{4\gamma^2 + \theta a} \right)^{\frac{1}{2}} \exp \left\{ \left( 1 - \sqrt{\frac{a(\theta - 8r)}{4\gamma^2 + \theta a}} \right) \frac{y^2}{4} \right\}.$$

Everything checks. With  $\lambda_\gamma$  as at (4.5), we have

$$(4.20) \quad g_{\lambda_\gamma}^* = \bar{g}_\gamma,$$

and indeed it is true that

$$\sup_g \inf_{\lambda < 0} L_\gamma(\lambda, g) = \inf_{\lambda < 0} \sup_g L_\gamma(\lambda, g) = L_\gamma(\lambda_\gamma, \bar{g}_\gamma).$$

*Particles with  $X$ -values close to  $-\gamma t$  at time  $t$  are likely to have type histories by time  $t$  with occupation densities close to  $t(\bar{g}_\gamma)^2 \phi$ .*

The following Legendre-conjugate expressions hold (we already saw the first at (4.3)):

$$(4.21) \quad \Delta(\gamma) = \inf_{\lambda < 0} \{E(\lambda) + \lambda\gamma\}, \quad E(\lambda) = \sup_{\gamma > 0} \{\Delta(\gamma) - \gamma\lambda\},$$

If, for  $\lambda_{\min} < \lambda < 0$ , we write  $\gamma_\lambda$  for the  $\gamma$  value which achieves the supremum on the right-hand side of (4.21), then the functions

$$\begin{aligned} \lambda &\mapsto \gamma_\lambda \text{ from } (-\lambda_{\min}, 0) \text{ to } (0, \infty), \text{ and} \\ \gamma &\mapsto \lambda_\gamma \text{ from } (0, \infty) \text{ to } (-\lambda_{\min}, 0) \end{aligned}$$

are inverses of each other.

(4.22) *Remarks.* When we move on to our rigorous treatment, we find the familiar story: getting a lower bound for  $\text{Vel}$  is quite easy, while obtaining the best upper bound is much more tricky. The fact (now known to us) that the expected number of particles near  $-\gamma t$  for  $0 \leq \gamma < \tilde{c}(\theta)$  is large for large  $t$  does not guarantee that we shall continue finding particles in that region. ‘Expectation wavefronts’ and particle wavefronts can differ. The martingale techniques in the next section are safe but not always applicable.

(4.23) *A formula concerning  $Y$ .* From (4.15) with  $\lambda = 0$ ,

$$\mathbb{E}^{0,0} N(t) \sim \text{const} \exp \{(\rho + \theta\psi_0^-)t\},$$

as one would expect from the formula (4.4) with  $\lambda = 0$ . We can use (3.8) to calculate the expected number of  $Y$ -values in any subregion of  $\mathbb{R}$  at time  $t$ . Indeed, writing the answer in a form symmetrical in  $\psi^+$  and  $\psi^-$ , we have

$$(4.24) \quad \frac{\mathbb{E}^{0,0} \# \{k \leq N(t) : Y_k(t) \in dy\}}{dy} = \frac{\exp \left( (\rho + \frac{1}{2}\theta)t - \frac{(\psi_0^+ e^{\mu_0 t} - \psi_0^- e^{-\mu_0 t})}{2 \sinh \mu_0 t} y^2 \right)}{\sqrt{2\pi (\theta/\mu_0) \sinh \mu_0 t}}.$$

For large  $t$ , it is as if, given  $N(t)$ , each  $Y$ -value is normally distributed with mean 0 and variance  $(2\psi_0^+)^{-1}$ , that is, each with density a constant multiple of  $(g_0^*)\phi$ , in agreement with (4.13).

(4.25) *Heuristics on the long-term behaviour of  $Y$ .* (These remain heuristics here: rigorous treatment is a matter for a different paper.) If we believe that expectation and particle wavefronts agree here, we will guess that (provided  $\theta > 8r$ ) we have, almost surely,

$$(4.26)(?) \quad \lim_{t \rightarrow \infty} t^{-1} \sup_{k \leq N(t)} Y_k(t)^2 = \frac{\rho + \theta\psi_0^-}{\psi_0^+} = \frac{4\rho + \theta - \sqrt{\theta(\theta - 8r)}}{1 + \sqrt{1 - 8r/\theta}}.$$

We might very well believe this to hold when  $\theta = 8r$ , too.

### 5. Uniform integrability of $Z_\lambda^-$ for $\lambda \in (\tilde{\lambda}(\theta), 0)$

Recall that for  $\lambda \in (\lambda_{\min}, 0)$ , where

$$\lambda_{\min} := -\sqrt{\frac{\theta - 8r}{4a}},$$

we define

$$\psi_\lambda^\pm := \frac{\theta \pm \sqrt{\theta^2 - \theta(8r + 4a\lambda^2)}}{4\theta}, \quad c_\lambda^- := -(\rho + \theta\psi_\lambda^-) / \lambda,$$

and that

$$Z_\lambda^-(t) := \sum_{k=1}^{N(t)} \exp \left\{ \psi_\lambda^- Y_k(t)^2 + \lambda [X_k(t) + c_\lambda^- t] \right\}.$$

Recall too that  $\tilde{\lambda}(\theta)$  is the unique point in  $(\lambda_{\min}, 0)$  at which the convex function  $c^-$  achieves its minimum  $\tilde{c}(\theta)$ .

We want to show that, for  $\lambda \in (\tilde{\lambda}(\theta), 0)$ , the martingale  $Z_\lambda^-$  is  $\mathcal{L}^p$  bounded for some  $p > 1$ . To do so, we need the following important estimate the proof of which is given in Section 7.

(5.1) LEMMA: An  $\mathcal{L}^p$  bound for the  $Z_\lambda^-$  martingales. *Given  $\lambda \in (\lambda_{\min}, 0)$  and  $\epsilon > 0$ , there exist  $T > 0$  and  $\tilde{K}$  in  $[0, \infty)$  such that for  $p \in [1, 2]$ ,*

$$\mathbb{E}^{0,y} (|Z_\lambda^-(t)|^p) \leq \tilde{K} e^{p(\psi_\lambda^- + \epsilon)y^2} \quad (y \in \mathbb{R}, t \in [0, T]).$$

Given Lemma 5.1, the following lemma from Neveu (1987) works very effectively.

(5.2) LEMMA (Neveu), *Let  $p \in (1, 2]$ . For any finite sequence of positive independent random variables  $W_1, \dots, W_n$  in  $\mathcal{L}^p$  and any sequence of positive real numbers  $c_1, \dots, c_n$ , we have*

$$\beta_p \left( \sum_{k=1}^n c_k W_k \right) \leq \sum_{k=1}^n c_k^p \beta_p(W_k),$$

where  $\beta_p(W) := \mathbb{E}(W^p) - \mathbb{E}(W)^p$  for  $W \in \mathcal{L}^p$ .

(5.3) THEOREM. *Let  $\lambda \in (\tilde{\lambda}(\theta), 0)$ . Then, for some  $p > 1$ ,  $Z_\lambda^-$  is (for every  $\mathbb{P}^{x,y}$ ) bounded in  $\mathcal{L}^p$ . Thus, under each  $\mathbb{P}^{x,y}$ ,*

$$Z_\lambda^-(\infty) \text{ exists almost surely and in } \mathcal{L}^1.$$

Moreover,  $\mathbb{P}^{x,y}(Z_\lambda^-(\infty) = 0) = 0$  for all pairs  $(x, y)$ .

*Proof (guided by Neveu).* Fix  $\lambda \in (\tilde{\lambda}(\theta), 0)$ . We know that  $\lambda \mapsto c_\lambda^-$  is strictly increasing on  $(\tilde{\lambda}(\theta), 0)$ . Hence, for some  $p > 1$  and  $\epsilon > 0$ , both of which we now fix, we shall have

$$(5.4) \quad p\lambda \left( c_{p\lambda}^- - c_\lambda^- \right) > 0. \quad \psi_\lambda^+ - p\psi_\lambda^- > p\epsilon.$$

Now

$$Z_\lambda^-(t+s) = \sum_{k=1}^{N(s)} e^{\lambda(X_k(s) + c_\lambda^- s)} W_t^{0, y_k}$$

where  $W_t^{0, y_k}$  behaves like the branching process started with one particle starting at  $(0, y_k)$  where  $y_k = Y_k(s)$ . Applying the conditional version of Neveu's Lemma 5.2 yields

$$\mathbb{E}^{x,y} \left( Z_\lambda^-(s+t)^p \mid \mathcal{F}_s \right) - Z_\lambda^-(s)^p \leq \sum_{k=1}^{N(s)} e^{p\lambda(X_k(s) + c_\lambda^- s)} \mathbb{E}^{0, y_k} Z_\lambda^-(t)^p.$$

By Lemma 5.1, there exist  $T > 0$  and  $\tilde{K} < \infty$  such that for  $y \in \mathbb{R}$  and  $t \in [0, T]$ ,

$$\mathbb{E} \left( Z_\lambda^-(s+t)^p \mid \mathcal{F}_s \right) - Z_\lambda^-(s)^p \leq \tilde{K} \sum_{k=1}^{N(s)} e^{p\lambda(X_k(s) + c_\lambda^- s) + p(\psi_\lambda^- + \epsilon)Y_k(s)^2}.$$

We now concentrate on the  $\mathbb{P}^{0,0}$  measure for simplicity. Using the methods in Section 3, we find that

$$\begin{aligned} & \mathbb{E}^{0,0} (Z_\lambda^-(s+t)^p) - \mathbb{E}^{0,0} (Z_\lambda^-(s)^p) \\ & \leq \tilde{K} \mathbb{E}^{0,0} \left( \sum_{k=1}^{N(s)} e^{p\lambda(X_k(s)+c_\lambda^-s)+p(\psi_\lambda^-+\epsilon)Y_k(s)^2} \right) \quad | \\ & \leq \tilde{K} e^{p\lambda c_\lambda^- s} \left\{ 1 - (1 - e^{-2\theta(\psi_{p\lambda}^+ - \psi_{p\lambda}^-)s}) \left( \frac{p\psi_\lambda^- + p\epsilon - \psi_{p\lambda}^-}{\psi_{p\lambda}^+ - \psi_{p\lambda}^-} \right) \right\}^{-\frac{1}{2}} e^{(\theta\psi_{p\lambda}^- + \rho)s}. \end{aligned}$$

Because of (5.4), this term decays exponentially as  $s \rightarrow \infty$ . Thus, we have found that there exist  $T > 0$ ,  $K < \infty$ ,  $\ell > 0$ , such that for  $t \in [0, T]$ ,

$$\mathbb{E}^{0,0} (Z_\lambda^-(s+t)^p) - \mathbb{E}^{0,0} (Z_\lambda^-(s)^p) \leq K e^{-\ell s} \quad (s > 0).$$

Finally, we have, for all  $t \geq 0$  and  $s \in [0, T]$ ,

$$\sum_{m=1}^{\infty} \mathbb{E}^{0,0} (Z_\lambda^-(ms+s)^p - Z_\lambda^-(ms)^p) \leq K \sum_{m=1}^{\infty} e^{-\ell ms} < \infty$$

and thus  $Z_\lambda^-$  is indeed bounded in  $\mathcal{L}^p$  under  $\mathbb{P}^{0,0}$ . Similar arguments apply to  $\mathbb{P}^{x,y}$ .

For the last part of the theorem, we use the following lemma.

(5.5) LEMMA. *Let  $\lambda \in (\lambda_{\min}, 0)$ . Then (as a function of  $(x, y)$ ),*

$$\mathbb{P}^{x,y} (Z_\lambda^-(\infty) = 0) \equiv 0 \text{ or } 1.$$

*Proof.* First note that

$$\mathbb{P}^{x,y} (Z_\lambda^-(\infty) = 0) = \mathbb{P}^{0,y} (e^{\lambda x} Z_\lambda(\infty) = 0) = \mathbb{P}^{0,y} (Z_\lambda^-(\infty) = 0) =: u(y).$$

Thus it is clear that the probability is independent of the spatial start position.

For all  $t \geq 0, y \in \mathbb{R}$ ,

$$\begin{aligned} u(y) &= \mathbb{E}^y (I \{ Z_\lambda^-(\infty) = 0 \}) = \mathbb{E}^y \left\{ \mathbb{E} (I \{ Z_\lambda^-(\infty) = 0 \} \mid \mathcal{F}_t) \right\} \\ &= \mathbb{E}^y \left( \prod_{k=1}^{N(t)} u(Y_k(t)) \right) \leq \mathbb{E}^y \left\{ u(Y_1(t)) \right\}. \end{aligned}$$

Hence  $u(\eta_t)$  is a bounded submartingale, whence it converges. Since  $\eta$  is recurrent,  $u(\cdot)$  must be constant on  $\mathbb{R}$ , whence (why?)  $u(\cdot) \equiv 0$  or  $1$ .  $\square$

**6. Proof that  $t^{-1}L(t) \rightarrow -\bar{c}(\theta)$ , a.s.**

Let us work with the  $\mathbb{P}^{0,0}$  measure. Let  $\lambda \in (\bar{\lambda}(\theta), 0)$ . Because  $Z_\lambda^-(\infty)$  exists in  $\mathcal{L}^1$  and is almost surely strictly positive, and since  $Z_\lambda^-(0) = 1$ , we can define a new probability measure  $\mathbb{Q}_\lambda = \mathbb{Q}_\lambda^{0,0}$  equivalent to  $\mathbb{P}$  on  $\mathcal{F}_\infty$  via

$$\frac{d\mathbb{Q}_\lambda}{d\mathbb{P}} = Z_\lambda^-(\infty) \quad \text{on } \mathcal{F}_\infty, \quad \text{whence} \quad \left. \frac{d\mathbb{Q}_\lambda}{d\mathbb{P}} \right|_{\mathcal{F}_t} = Z_\lambda^-(t).$$

It is now easy to show that  $(\partial/\partial\lambda)Z_\lambda^-(t)$  is also a  $\mathbb{P}$ -martingale, so that

$$M_\lambda(t) := Z_\lambda^-(t)^{-1} \frac{\partial}{\partial\lambda} Z_\lambda^-(t)$$

is a  $\mathbb{Q}_\lambda$ -martingale. For  $t \geq 0$ ,  $1 \leq j \leq N(t)$ , let

$$(6.1) \quad J_\lambda(t, j) := \frac{\exp(\psi_\lambda^- Y_j(y)^2 + \lambda(X_j(t) + c_\lambda^- t))}{\sum_{k=1}^{N(t)} \exp(\psi_\lambda^- Y_k(y)^2 + \lambda(X_k(t) + c_\lambda^- t))},$$

so that we have  $J_\lambda(\cdot, \cdot) \geq 0$  and  $\sum_{k=1}^{N(t)} J_\lambda(t, k) = 1$ . Then,

$$\begin{aligned} M_\lambda(t) &= \sum_{k=1}^{N(t)} J_\lambda(t, k) \{(\psi_\lambda^-)' Y_k(t)^2 + X_k(t) + (\lambda c_\lambda^-)' t\} \\ &= (\psi_\lambda^-)' \sum_{k=1}^{N(t)} J_\lambda(t, k) Y_k(t)^2 + \sum_{k=1}^{N(t)} J_\lambda(t, k) \{X_k(t) + (\lambda c_\lambda^-)' t\} \\ &\geq (\psi_\lambda^-)' \sum_{k=1}^{N(t)} J_\lambda(t, k) Y_k(t)^2 + L(t) + (\lambda c_\lambda^-)' t. \end{aligned}$$

Then, with  $V_\lambda(t) := \sum_{k=1}^{N(t)} J_\lambda(t, k) Y_k(t)^2$ ,

$$\frac{L(t)}{t} + (\lambda c_\lambda^-)' \leq \frac{M_\lambda(t)}{t} - (\psi_\lambda^-)' \frac{V_\lambda(t)}{t},$$

and the idea is to show that both terms on the right tend almost surely to zero:

$$(6.2) \quad t^{-1}V_\lambda(t) \rightarrow 0 \quad \text{a.s.}, \quad t^{-1}M_\lambda(t) \rightarrow 0 \quad \text{a.s.},$$

whence we shall have

$$\limsup \frac{L(t)}{t} \leq -(\lambda c_\lambda^-)', \quad \text{a.s.}$$

Since  $(\lambda c_\lambda^-)' \rightarrow \tilde{c}(\theta)$  as  $\lambda \downarrow \tilde{\lambda}(\theta)$ , we therefore have the desired upper bound

$$\limsup \frac{L(t)}{t} \leq -\tilde{c}(\theta), \quad \text{a.s.},$$

and the fact that

$$t^{-1}L(t) \rightarrow -\tilde{c}(\theta), \quad \text{a.s.}$$

will be proved. It therefore remains only to prove (6.2).

*Proof that  $t^{-1}V_\lambda(t) \rightarrow 0$  a.s..* We know from Section 3 that

$$\mathbb{Q}_\lambda^{0,0}(V_\lambda(t)) = \mathbb{E}_{\mu_\lambda}^0(\eta_t^2), \quad \mu_\lambda := \frac{1}{2}\sqrt{\theta(\theta - 8r - 4a\lambda^2)},$$

where  $\eta$  is an  $\text{OU}(\theta, \mu_\lambda)$  process started at 0 under  $\mathbb{P}_{\mu_\lambda}^0$ . Thus, since an  $\text{OU}(\theta, \mu_\lambda)$  process is ergodic,  $\mathbb{Q}_\lambda^{0,0}(V_\lambda(t))$  tends to a limit as  $t \rightarrow \infty$ , and is therefore bounded in  $t$ . For  $n \in \mathbb{N}$ , Jensen's inequality tells us that

$$\mathbb{Q}_\lambda^{0,0}(V_\lambda(t)^n) \leq \mathbb{Q}_\lambda^{0,0}\left(\sum_{k=1}^{N(t)} J_\lambda(t, k) Y_k(t)^{2n}\right) = \mathbb{E}_{\mu_\lambda}^0(\eta_t^{2n}),$$

and this expression is again bounded in  $t$ .

We now need the fact that

$$(6.3) \quad e^{2\mu_\lambda t} \left[ V_\lambda(t) - \frac{\theta}{2\mu_\lambda} \right] \text{ is a } \mathbb{Q}_\lambda^{0,0} \text{ martingale.}$$

This follows by combining the methods of Section 3 with the fact that if  $\eta$  is an  $\text{OU}(\theta, \mu)$  process, then

$$e^{2\mu t} \left( \eta_t^2 - \frac{\theta}{2\mu} \right) \text{ is a martingale.}$$

In harmonic-oscillator language, this martingale is a 'second excited state' martingale; the previous ones have been 'ground state'.

It is now trivial to prove that, for any  $\delta > 0$ ,

$$(6.4) \quad t^{-\delta} \left[ V_\lambda(t) - \frac{\theta}{2\mu_\lambda} \right] \rightarrow 0, \quad \text{a.s..}$$

For choose an integer  $n$  such that  $2n\delta > 1$ , and use Doob's submartingale inequality to see that for any integer  $T > 1$ ,

$$\begin{aligned} & \mathbb{P} \left[ \sup_{t \in [T-1, T]} \left| V_\lambda(t) - \frac{\theta}{2\mu_\lambda} \right| \geq \epsilon T^\delta \right] \\ & \leq \mathbb{P} \left[ \sup_{t \in [T-1, T]} \left( V_\lambda(t) - \frac{\theta}{2\mu_\lambda} \right)^{2n} e^{4nt} \geq \epsilon^{2n} T^{2n\delta} e^{4\mu_\lambda n(T-1)} \right] \\ & \leq \frac{1}{\epsilon^{2n} T^{2n\delta} e^{4\mu_\lambda n(T-1)}} e^{4\mu_\lambda nT} \mathbb{E} \left\{ \left( V_\lambda(T) - \frac{\theta}{2\mu_\lambda} \right)^{2n} \right\} \\ & \leq \frac{e^{4\mu_\lambda n}}{\epsilon^{2n}} U_n \frac{1}{T^{2n\delta}} \quad (U_n \text{ a constant}), \end{aligned}$$

and since  $\sum 1/T^{2n\delta}$  converges, the result (6.4) follows from the Borel-Cantelli Lemma.

*Proof that  $t^{-1}M_\lambda(t) \rightarrow 0$  a.s..* Jensen's inequality gives

$$M_\lambda(t)^2 \leq \sum_{k=1}^{N(t)} J_\lambda(t, k) \{ (\psi_\lambda^-)' Y_k(t)^2 + X_k(t) + (\lambda c_\lambda^-)' \}^2.$$

We can show that  $(Z_\lambda^-)^{-1}(\partial^2/\partial\lambda^2)Z_\lambda^-$  is a  $\mathbb{Q}_\lambda$ -martingale, where

$$\begin{aligned} & Z_\lambda^-(t)^{-1} \frac{\partial^2}{\partial\lambda^2} Z_\lambda^-(t) \\ & = \sum_{k=1}^{N(t)} J_\lambda(t, k) \left\{ \{ (\psi_\lambda^-)' Y_k(t)^2 + X_k(t) + (\lambda c_\lambda^-)' \}^2 + (\psi_\lambda^-)'' Y_k(t)^2 + (\lambda c_\lambda^-)'' t \right\}. \end{aligned}$$

Then

$$\begin{aligned} \mathbb{Q}_\lambda (M_\lambda(t)^2) & \leq \mathbb{Q}_\lambda \left( Z_\lambda^-(t)^{-1} \frac{\partial^2}{\partial\lambda^2} Z_\lambda^-(t) - \sum_{k=1}^{N(t)} J_\lambda(t, k) \{ (\psi_\lambda^-)'' Y_k(t)^2 + (\lambda c_\lambda^-)'' t \} \right) \\ & = \mathbb{Q}_\lambda \left( Z_\lambda^-(0)^{-1} \frac{\partial^2}{\partial\lambda^2} Z_\lambda^-(0) \right) - (\psi_\lambda^-)'' \mathbb{Q}_\lambda (V_\lambda(t)) - (\lambda c_\lambda^-)'' t \\ & \leq K_1(\lambda) + K_2(\lambda)t, \end{aligned}$$

for some finite  $K_i$ 's. Doob's submartingale inequality now yields

$$\begin{aligned} \mathbb{Q}_\lambda (\sup \{s^{-1} |M_\lambda(s)| : 2^{n-1} \leq s \leq 2^n\} \geq \epsilon) & \leq \mathbb{Q}_\lambda \left( \sup_{0 \leq s \leq 2^n} |M_\lambda(s)| \geq \epsilon 2^{n-1} \right) \\ & \leq (\epsilon 2^{n-1})^{-2} \{ K_1(\lambda) + 2^n K_2(\lambda) \}, \end{aligned}$$

and the Borel-Cantelli Lemma completes the proof.  $\square$

### 7. Proof of Lemma 5.1

(At last, an expectation calculation which is not of the elementary ‘linear’ form.)

Our branching process has statespace

$$\mathcal{I} := \bigcup_{n \geq 1} (\{n\} \times \mathbb{R}^n \times \mathbb{R}^n).$$

Its formal generator  $\mathcal{G}$  is given by

$$(7.1) \quad \mathcal{G} = \mathcal{G}_A + \mathcal{G}_\theta + \mathcal{G}_R,$$

where for  $n \geq 1, x, y \in \mathbb{R}^n$ , we have

$$(7.2) \quad \begin{aligned} (\mathcal{G}_A F)(n; \mathbf{x}; \mathbf{y}) &= \sum_{k=1}^n \frac{1}{2} A(y_k) \frac{\partial^2 F}{\partial x_k^2}, \\ (\mathcal{G}_\theta F)(n; \mathbf{x}; \mathbf{y}) &= \sum_{k=1}^n \frac{\theta}{2} \left\{ \frac{\partial^2 F}{\partial y_k^2} - y_k \frac{\partial F}{\partial y_k} \right\}, \\ (\mathcal{G}_R F)(n; \mathbf{x}; \mathbf{y}) &= \sum_{k=1}^n R(y_k) \left\{ F(n+1; (\mathbf{x}, x_k); (\mathbf{y}, y_k)) - F(n; \mathbf{x}; \mathbf{y}) \right\}, \end{aligned}$$

where  $(\mathbf{x}, x_k) := (x_1, \dots, x_n, x_k) \in \mathbb{R}^{n+1}$ .

(7.3) PROPOSITION: Local-martingale condition. *If  $F : [0, \infty) \times \mathcal{I} \rightarrow \mathbb{R}$  and*

$$\left\{ \left( \frac{\partial}{\partial t} + \mathcal{G} \right) F \right\} (t; n; \mathbf{x}; \mathbf{y}) = 0 \quad \text{for } t \geq 0, n \geq 1, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,$$

*then  $F(t; N(t); \mathbf{X}(t); \mathbf{Y}(t))$  is a local martingale.*

We know that

$$(7.4) \quad h_\lambda(t; n; \mathbf{x}; \mathbf{y}) := \sum_{k=1}^n e^{\psi_\lambda^- y_k^2 + \lambda(x_k + c_\lambda^- t)}$$

leads to the martingale  $Z_\lambda^-(t) = h_\lambda(t; N(t); \mathbf{X}(t); \mathbf{Y}(t))$ . Now,  $Z_\lambda^-$  jumps when a new particle is born; but any jump of  $Z_\lambda^-$  is of magnitude no greater than the current value of  $Z_\lambda^-$ . If, therefore, we introduce the stopping time

$$S_n := \inf\{t : Z_\lambda^-(t) \geq n\},$$

then  $Z_\lambda^-$  stopped at  $S_n$  never exceeds  $2n$ . Hence,  $Z_\lambda^-$  is locally in  $\mathcal{L}^2$  (relative to any  $\mathbb{P}^{x, y}$ ). We may now conclude that

$$Z_\lambda^-(t)^2 - A(t) \text{ is a local martingale}$$

where

$$A(t) := \int_0^t \left\{ \left( \mathcal{G} + \frac{\partial}{\partial t} \right) \left( (h_\lambda)^2 \right) \right\} (s; N(s); \mathbf{X}(s); \mathbf{Y}(s)) ds.$$

We find that – and this nice fact is crucial –

$$(7.5) \quad \frac{dA(t)}{dt} = \sum_{k=1}^{N(t)} \left\{ (a\lambda^2 + 4\theta(\psi_\lambda^-)^2 + r)Y_k(t)^2 + \rho \right\} e^{2\psi_\lambda^- Y_k(t)^2 + 2\lambda(X_k(t) + c_\lambda^- t)}.$$

The methods of Section 3 give tight bounds on  $\mathbb{E}^{0,y} A(t)$  and in particular show that it is finite for small  $t$ . We can use Fatou's Lemma to deduce from the fact that  $(Z_\lambda^-)^2 - A$  is a local martingale that

$$(7.6) \quad \mathbb{E}^{0,y} [Z_\lambda^-(t)^2] \leq \mathbb{E}^{0,y} A(t) + e^{2\psi_\lambda^- y^2}.$$

In this way, we can prove the following result.

(7.7) LEMMA: An  $\mathcal{L}^2$  bound for  $Z_\lambda^-$ . *Given  $\lambda < 0$  and  $\epsilon > 0$ , there exist  $T > 0$  and  $K < \infty$  such that*

$$\mathbb{E}^{0,y} (Z_\lambda^-(t)^2) \leq K e^{2(\psi_\lambda^- + \epsilon)y^2} \quad (y \in \mathbb{R}, t \in [0, T]).$$

Granted this result, we can use the monotonicity of  $\mathcal{L}^p$ -norms, namely

$$\mathbb{E}|X| \leq (\mathbb{E}[|X|^p])^{\frac{1}{p}} \leq (\mathbb{E}[|X|^2])^{\frac{1}{2}} \quad \text{for } p \in [1, 2]$$

to tell us that for  $p \in [1, 2]$ ,

$$\mathbb{E}^{0,y} (Z_\lambda^-(t)^p) \leq K^{\frac{p}{2}} e^{p(\psi_\lambda^- + \epsilon)y^2} \quad (y \in \mathbb{R}, t \in [0, T]),$$

and Lemma 5.1 follows.

*Remark.* In the case when  $\mathbb{E}^{0,y} A(t)$  is always finite, then, because  $Z_\lambda^-(t)$  is in  $\mathcal{L}^2$  for  $\mathbb{E}^{0,y}$ , the existence and uniqueness parts of the Meyer decomposition theorem show that  $(Z_\lambda^-)^2 - A$  is a true martingale. Thus, equality holds in (7.6).

## 8. The travelling-wave and reaction-diffusion equations

We just remark for now that, as McKean (1975) has taught us, for  $\tilde{\lambda}(\theta) < \lambda < 0$ ,

$$(8.1) \quad w(x, y) := \mathbb{E}^{x,y} \left( e^{-Z_\lambda^-(\infty)} \right)$$

is a solution, monotone in  $x$  and tending to 0 [respectively, 1] as  $x \rightarrow -\infty$  [ $x \rightarrow \infty$ ] for each  $y$ , of the travelling-wave equation

$$(8.2) \quad \frac{1}{2} A(y) \frac{\partial^2 w}{\partial x^2} + c_\lambda^- \frac{\partial w}{\partial x} + R(y)w(w-1) + \frac{\theta}{2} \left( \frac{\partial^2 w}{\partial y^2} - y \frac{\partial w}{\partial y} \right) = 0.$$

Then  $u(t, x, y) := w(x - c_\lambda^- t, y)$  solves the reaction-diffusion equation

$$(8.3) \quad \frac{\partial u}{\partial t} = \frac{1}{2} A(y) \frac{\partial^2 u}{\partial x^2} + R(y)u(u-1) + \frac{\theta}{2} \left( \frac{\partial^2 u}{\partial y^2} - y \frac{\partial u}{\partial y} \right)$$

associated with our problem. Further discussion of these equations is deferred to sequels to this paper.

### 9. The first expectation calculation for the critical case $\theta = 8r$

(9.1) THEOREM. *If  $\theta = 8r$ , then*

$$\lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E}^{0,0} \# \left\{ k \leq N(t) : X_k(t) > c(\theta a)^{\frac{1}{2}} t^2 \right\} > 0 \quad \text{if } c < \pi^{-1}(\rho + \frac{1}{4}\theta), \\ < 0 \quad \text{if } c > \pi^{-1}(\rho + \frac{1}{4}\theta).$$

*Thus, the expectation wavefront (in the positive direction) is essentially the parabolic curve  $x = (\theta a)^{\frac{1}{2}} \pi^{-1}(\rho + \frac{1}{4}\theta)t^2$ .*

*Proof.* From (4.15), we find that

$$(9.2) \quad \int e^{i\lambda x} \mathbb{E}^{0,0} \# \left\{ k \leq N(t) : \frac{X_k(t)}{t\sqrt{\theta a}} \in dx \right\} = \mathbb{E}^{0,0} \sum_{k=1}^{N(t)} \exp \left\{ \frac{i\lambda X_k(t)}{t\sqrt{\theta a}} \right\} \\ = \frac{e^{(\rho + \frac{1}{4}\theta)t}}{(1 + \frac{1}{4}\theta t)^{\frac{1}{2}}} \mathbb{E} e^{i\lambda S(t)},$$

where  $S(t)$  is a random variable with characteristic function

$$(9.3) \quad \mathbb{E} e^{i\lambda S(t)} = \left\{ \frac{1 + \frac{1}{4}\theta t}{\cosh \lambda + \frac{1}{4}\theta t \lambda^{-1} \sinh \lambda} \right\}^{\frac{1}{2}}.$$

Some of the following analysis could be derived probabilistically along the route suggested by Donati and Yor (1991) and Chan, Dean, Jansons and Rogers (1994). However, complex analysis deals very effectively with what we need.

The function  $f_t : \mathbb{C} \rightarrow \mathbb{C}$ , where

$$f_t(z) := \cosh z + \frac{1}{4}\theta t z^{-1} \sinh z$$

is entire and of exponential order 1. All of its roots are purely imaginary; for if  $u + iv$  is a root, we find that

$$\frac{\sin v}{\cos v} = \frac{u \cosh u + \frac{1}{4}\theta t \sinh u}{v \sinh u} = \frac{(-v \cosh u)}{u \sinh u + \frac{1}{4}\theta t \cosh u},$$

whence

$$(\frac{1}{2}\theta^2 t^2 + u^2) \sinh u \cosh u + \frac{1}{2}\theta t u (\sinh^2 u + \cosh^2 u) = -v^2 \sinh u \cosh u,$$

and if  $u \neq 0$ , the two sides of this equation have different signs. Thus the zeros of  $f_t$  are at points  $\pm i v_1(t), \pm i v_2(t), \dots$ , where  $v_1(t), v_2(t), \dots$  are the positive roots (labelled in increasing order) of the equation

$$\cos v + \frac{1}{2}\theta t \frac{\sin v}{v} = 0.$$

We see from a graph that  $v(n)/n \rightarrow \pi$ , so that  $\sum v_n(t)^{-2} < \infty$ , and  $f_t(\cdot)$  is of genus 1 in the language of Hadamard's factorization theorem – see Section 8.2 of Titchmarsh (1952). Using the fact that  $f_t(z)$  is even in  $z$ , we obtain from Hadamard's Theorem the formula

$$\mathbb{E} e^{i\lambda S(t)} = \prod_{n=1}^{\infty} \left\{ \frac{1}{1 + \lambda^2/v_n(t)^2} \right\}^{\frac{1}{2}}.$$

This means that we can regard  $S(t)$  as a sum

$$S(t) = \sum_{n=1}^{\infty} W_n(t)$$

of independent variables,  $W_n(t)$  having characteristic function  $[1 + \lambda^2/v_n(t)^2]^{-\frac{1}{2}}$ . With  $R_1(t)$  denoting  $S(t) - W_1(t)$ , we have

$$\mathbb{P}[S(t) > x] \geq \mathbb{P}[W_1(t) > x; R_1(t) \geq 0] = \frac{1}{2}\mathbb{P}[W_1(t) > x].$$

But  $W_1(t)$  has the same distribution as  $W/v_1(t)$ , where

$$\mathbb{E} e^{i\lambda W} = (1 + \lambda^2)^{-\frac{1}{2}},$$

and  $W$  has density  $\pi^{-1}K_0$  where  $K_0$  is the usual modified Bessel function. Now,  $v_1(t) < \pi$  for every  $t$ ; and so, for  $x > 0$ ,

$$\mathbb{P}[W_1(t) > x] = \mathbb{P}[W > v_1(t)x] \geq P[W > \pi x] = \frac{e^{-\pi x}}{\sqrt{2x}} (1 + O(1/x)).$$

Since

$$\mathbb{E}^{0,0} \# \left\{ k \leq N(t) : X_k(t) > c(\theta a)^{\frac{1}{2}} t^2 \right\} = \frac{e^{(\rho + \frac{1}{2}\theta)t}}{(1 + \frac{1}{2}\theta t)^{\frac{1}{2}}} \mathbb{P}[S(t) > ct],$$

the first part of the theorem follows.

Provided that  $|v| < v_1(t)$ , we have

$$\mathbb{E} e^{vS(t)} = m_t(v) := \left\{ \frac{1 + \frac{1}{2}\theta t}{\cos v + \frac{1}{2}\theta t v^{-1} \sin v} \right\}^{\frac{1}{2}}.$$

Now, let  $c$  be an arbitrary but fixed number with  $c > \pi^{-1}(\rho + \frac{1}{4}\theta)$ . Choose and fix  $v_0$  with  $0 < v_0 < \pi$  and such that  $cv_0 > (\rho + \frac{1}{4}\theta)$ . Noting that  $v_1(t) \uparrow \pi$  when  $t \uparrow \infty$ , we realize that we can find  $t_0$  with  $v_1(t_0) > v_0$  such that

$$\sup_{t \geq t_0} m_t(v_0) < \infty.$$

But, for  $c > 0$ ,

$$\mathbb{P}[S(t) > ct] \leq e^{-cv_0 t} \mathbb{E}[e^{v_0 S}; S > ct] \leq e^{-cv_0 t} m_t(v_0).$$

The second part of the theorem now follows. □

(9.4) *Remark.* It is obvious from (9.3) that the distribution of  $S(t)$  converges to the distribution with characteristic function  $(\lambda / \sinh \lambda)^{\frac{1}{2}}$  which has tail behaviour roughly like  $e^{-\pi|\lambda|}$ . Assuming that at time  $t$  we have  $\mathbb{E}N(t)$  particle positions each of which when divided by  $(\theta a)^{\frac{1}{2}}t$  has this limiting distribution, would lead to the correct answer for the approximate asymptotic wavefront in this ‘ $\theta = 8r$ ’ case. However, you should contrast the following observations. If  $\theta > 8r$ , then, for any subinterval  $\Gamma$  of  $\mathbb{R}$ , we have

$$(9.5) \quad \frac{\mathbb{E}^{0,0} \# \left\{ k \leq N(t) : t^{-\frac{1}{2}} X_k(t) \in \Gamma \right\}}{\mathbb{E}^{0,0} N(t)} \rightarrow L(\Gamma),$$

where  $L$  is the normal law of mean 0 and variance  $a\sqrt{\theta/(\theta - 8r)}$ ; but assuming that we have  $\mathbb{E}N(t)$  particle positions each of which when divided by  $\sqrt{t}$  has this limiting normal distribution, would lead to too low a wavespeed in our earlier work. In (9.5), we are here attaching too little weight to the particles far from the origin which play a more significant role. ‘Central Limit Theorems’ such as (9.5) focus too heavily on deviations of ‘average’ magnitude, not on the large deviations which concern us.

## 10. Acknowledgements

We thank John Toland, Jonathan Warren and the referees for their help.

We also thank SERC/EPSRC for SCH’s research studentship.

*An ‘MN60’ Thankyou from DW.* DW started his career on Markov chains, which he learnt from his supervisors Kendall and Reuter and from the writings of Lévy, Chung and Neveu. Neveu’s work was a great influence then as now. After being enthused about Brownian motion by the books by Lévy and by Itô and McKean, about martingales by Doob’s book, and about stochastic integrals by McKean’s, DW

knew that he ought to learn *la théorie générale*. Inevitably then, much the most of what DW now knows about probability, he learnt from the writings of Meyer (much of it Meyer's own original work). And while DW is not gifted enough to be able to use what he has learnt to great effect himself, he sees his main task as that of helping the next generations (SCH is in the second) to use the Strasbourg theory as the foundation for their own research.

### References

- Champneys, A, Harris, S. C., Toland, J. F., Warren, J. & Williams, D (1995) Analysis, algebra and probability for a coupled system of reaction-diffusion equations, *Phil. Trans. Roy. Soc. London (A)*, **350**, 69–112.
- Chan, T., Dean, D. S., Jansons, K. M. & Rogers, L. C. G. (1994) On polymer conformations in elongational flows. *Comm. Math. Phys.* **160**, 239–257.
- Chauvin, B. & Rouault, A. (1988) KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees. *Probab. Thy. Rel. Fields* **80**, 299–314.
- Chauvin, B. & Rouault, A. (1990) Supercritical branching Brownian motion and K-P-P equation in the critical speed-area. *Math. Nachr.* **149**, 41–59.
- Donati-Martin, C. & Yor, M. (1991) Fubini's theorem for double Wiener integrals and the variance of the Brownian path. *Ann. Inst. H. Poincaré* **27**, 2, 181–200.
- Deuschel, J.-D. & Stroock, D. W. (1989) *Large Deviations*. Academic Press, Boston.
- Harris, S.C. (1995) Large deviations and martingales for a typed diffusion, 2 (In preparation).
- Harris, S.C. & Williams, D. (1995) Large deviations and martingales for a typed diffusion, 3: at the critical temperature (In preparation).
- Itô, K. & McKean, H.P. (1965) *Diffusion Processes and their Sample Paths*. Springer, Berlin.
- McKean, H. P. (1975) Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov. *Comm. Pure Appl. Math.* **28**, 323–331.
- Neveu, J. (1987) Multiplicative martingales for spatial branching processes. *Seminar on Stochastic Processes* (ed. E. Çinlar, K. L Chung and R. K. Gettoor), Progress in Probability and Statistics **15**. pp. 223–241. Birkhäuser, Boston.
- Titchmarsh., E. C. (1952) *The Theory of Functions* (second edition). Oxford University Press.

S. C. Harris and D. Williams,  
School of Mathematical Sciences,  
University of Bath,  
Bath, BA2 7AY,  
UK.