



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Ann. I. H. Poincaré – PR 40 (2004) 759–770



ANNALES
DE L'INSTITUT
HENRI
POINCARÉ
PROBABILITÉS
ET STATISTIQUES

www.elsevier.com/locate/anihpb

On stationary distributions for the KPP equation with branching noise

Paul Horridge, Roger Tribe*

Mathematics Institute, University of Warwick, Coventry, UK

Received 31 October 2003; accepted 8 January 2004

Available online 12 September 2004

Abstract

This paper applies the method of Harris's convergence theorem for additive particle systems to a stochastic PDE that arises as the limit of long range contact processes. This is used to study the uniqueness of a translation invariant stationary distribution and its domain of attraction.

© 2004 Elsevier SAS. All rights reserved.

Résumé

On applique la méthode conduisant au théorème de convergence de Harris pour les systèmes de particules additifs à une EDP stochastique qui apparaît comme limite d'un processus de contact à longue portée. On étudie ainsi l'unicité de la mesure stationnaire invariante par translation et son domaine d'attraction.

© 2004 Elsevier SAS. All rights reserved.

1. Introduction and statement of results

We consider non-negative solutions of the one-dimensional stochastic PDE

$$\partial_t u = \partial_{xx} u + \theta u - u^2 + \sqrt{2u} dW. \quad (1)$$

Here W is a space–time white noise on $[0, \infty) \times \mathbb{R}$. By linear scaling we have reduced the equation to a single parameter θ . The aim is to show that, when θ is sufficiently large, there is exactly one non-zero, translation invariant, stationary distribution and to give a sufficient condition to be in its domain of attraction.

* Corresponding author.

E-mail address: tribe@maths.warwick.ac.uk (R. Tribe).

1.1. Background

The deterministic part of (1) is the KPP equation (also known as the Fisher or Kolmogorov equation) used in modelling travelling wave phenomena. The form of the noise term arises in high density limits of branching particle systems, as in Dawson–Watanabe processes. Solutions to (1) itself arise as limits of scaled long range contact processes (see Mueller and Tribe [7,8]). The solutions share many of the properties of the contact process (or its discrete time analogue oriented percolation), in particular the following phase transition. We write $u(t) = 0$ when $u(t, x) = 0$ for all $x \in \mathbb{R}$. In [7] it is shown that there exists a critical value of the parameter $\theta_c > 0$ such that, for any solution u started from a compactly supported non-zero initial condition, if $\theta < \theta_c$ then $u(t) = 0$ for large t with probability one, and if $\theta > \theta_c$ there is a strictly positive probability that $u(t)$ is non-zero for all t .

As for the contact process or oriented percolation, solutions to (1) possess a self duality relation. In [5], Harris uses a duality relation for a class of particle systems on \mathbb{Z}^d , which include long range contact processes, to study their ergodic behaviour. Our interest was to try to implement this use of duality in the stochastic PDE setting. The approach of Harris, sketched in Section 1.3, does carry over and the main interest of this paper is the methods used in this PDE setting for the various technical estimates needed. The method of proof relies totally on the duality relation, but the result acts as a guide for intuition about related equations for which no duality is available.

1.2. Self duality

A suitable state space for solutions is the set $\mathcal{C}_{\text{exp}}^+$ of non-negative continuous functions with slower than exponential growth, defined, as a subset of the space of non-negative continuous function \mathcal{C}^+ , by

$$\mathcal{C}_{\text{exp}}^+ = \{f \in \mathcal{C}^+ : \|f\|_\lambda < \infty \text{ for all } \lambda > 0\} \quad \text{where } \|f\|_\lambda = \sup_{x \in \mathbb{R}} e^{-\lambda|x|} |f(x)|.$$

Equip $\mathcal{C}_{\text{exp}}^+$ with the topology generated by the norms $\|f\|_\lambda$ for $\lambda > 0$. It is standard to show the existence of solutions, with continuous paths in $\mathcal{C}_{\text{exp}}^+$, started from any random initial condition in $\mathcal{C}_{\text{exp}}^+$ (see the techniques of Shiga [10]). Throughout this paper the results refer to such solutions. Uniqueness in law holds and the laws form a Feller Markov family (see Theorem 2.2 of Tribe [11]). We shall write P_f and E_f when we wish to indicate that we are referring to a solution with an initial condition f .

The self duality relation for (1) is as follows. Let u and v be independent solutions to (1) with initial conditions $u(0) \in \mathcal{C}_{\text{exp}}^+$ and $v(0) \in \mathcal{C}_c^+$, the space of non-negative continuous functions of compact support. Write (f, g) for the integral $\int_{\mathbb{R}} f(x)g(x) dx$, when this is well defined. Self duality is the identity, for $0 \leq s \leq t$,

$$E[e^{-(u(t), v(0))}] = E[e^{-(u(s), v(t-s))}] = E[e^{-(u(0), v(t))}]. \quad (2)$$

This identity fits into the framework of Markov process duality and can be established along the standard lines described in Ethier and Kurtz [4, Section 4.4]. The main steps are as follows: if we define

$$H(f, g) = e^{-(f, g)}((f, g^2) + (g^2, f) - (f, \partial_{xx}g) - \theta(f, g))$$

then by Ito's formula we have that $e^{-(u(t), g)} - \int_0^t H(u(s), g) ds$ is a local martingale. Assuming sufficient integrability we deduce that

$$\frac{d}{dt} E[e^{-(u(t), g)}] = E[H(u(t), g)].$$

Similarly $\frac{d}{dt} E[e^{-(v(t), f)}] = E[H(v(t), f)]$. Using integration by parts we have $H(f, g) = H(g, f)$, given enough smoothness and when one of the functions has compact support. The duality identity then follows formally, since we should have

$$\frac{d}{ds} E[e^{-(u(s), v(t-s))}] = 0.$$

To carry this argument through one needs to smooth the solutions to allow the integration by parts and also check that there is sufficient integrability. See Horridge [6, Proposition 3.4.1] for the details (or see similar smoothing arguments used to establish duality relations in, for example, Athreya and Tribe [1]).

1.3. The main result

The idea in Harris’s argument is to let $t \rightarrow \infty$ in the duality relation (2). The branching noise is well known (see Section 2) to maintain the compact support of a solution v started at $g \in \mathcal{C}_c^+$. Define the death time $\tau = \inf\{t: v(t) = 0\}$. Then from the self duality relation (2) we have

$$E[e^{-(u(t),g)}] = E[e^{-(u(0),v(t))}] \geq P_g[\tau \leq t]$$

since $(u(0), v(t)) = 0$ on $\{\tau \leq t\}$. Letting $t \rightarrow \infty$ we deduce

$$\liminf_{t \rightarrow \infty} E[e^{-(u(t),g)}] \geq P_g[\tau < \infty]. \tag{3}$$

When $\theta < \theta_c$ we know that $P_g[\tau < \infty] = 1$ and thus any solution suffers local extinction, that is converges in probability to zero on any compact region. In particular there can be no non-zero stationary distributions.

When $\theta > \theta_c$ the right-hand side of (3) is non-zero. The main step in Harris’s argument is to find a class of initial conditions $u(0)$ for which the complementary inequality to (3) holds, that is when

$$\lim_{t \rightarrow \infty} E[e^{-(u(t),g)}] = P_g[\tau < \infty]. \tag{4}$$

The method used to show that this identity holds is to show that on the set $\{\tau = \infty\}$ the solution $v(t)$ grows, in that $(v(t), 1) \rightarrow \infty$, and to use this to force $(u(0), v(t)) \rightarrow \infty$. The convergence of the Laplace functional, for all compactly supported test functions, is equivalent to convergence in distribution for random Radon measures on \mathbb{R} with the topology of vague convergence (which is also equivalent to convergence in distribution of $((u(t), g_1), \dots, (u(t), g_n))$ for any $g_i \in \mathcal{C}_c^+$ and $n \in \mathbb{N}$.) Thus when (4) holds, it implies that $u(t)$ converges in distribution and that the limit does not depend on $u(0)$. The class of $u(0)$ for which this will hold will contain all translation invariant $u(0)$ satisfying $P[u(0) = 0] = 0$. This will then establish the main result of the paper, which we state here. We use $(T_t: t \geq 0)$ to denote the heat semigroup.

Theorem 1. *Suppose that $\theta > \theta_c$. The Laplace functional*

$$\int e^{-(f,g)} \mu(df) = P_g[\tau < \infty], \quad \text{for } g \in \mathcal{C}_c^+, \tag{5}$$

characterizes a translation invariant measure μ on $\mathcal{C}_{\text{exp}}^+$ that is a stationary distribution for (1). It is the unique translation invariant stationary distribution satisfying $\mu\{f \neq 0\} = 1$.

If a measure ν on $\mathcal{C}_{\text{exp}}^+$ satisfies, for some $t_0 > 0$ and $K < \infty$,

$$\lim_{\delta \rightarrow 0} \nu\{f: T_{t_0}(f \wedge K)(x) < \delta\} = 0, \quad \text{uniformly over } x \in \mathbb{R} \tag{6}$$

then ν is in the domain of attraction of μ ; in other words if u is a solution to (1) with initial law ν then (the law of) $u(t)$ converges in distribution to μ .

Remarks. (1) As a consequence of (6), a deterministic initial condition $f \in \mathcal{C}_{\text{exp}}^+$ is in the domain of μ whenever $T_{t_0}(f \wedge K)$ is bounded away from zero, for some $t_0 > 0$ and $K < \infty$. The cut-off $f \wedge K$ is natural since the non-linearity $-u^2$ brings down arbitrarily large initial conditions in finite time.

(2) The hypothesis (6) is actually independent of the choice of $t_0 > 0$ and $K \in (0, \infty)$. Indeed the independence of K follows from the inequalities

$$\frac{K}{L} T_t(f \wedge L) \leq T_t(f \wedge K) \leq T_t(f \wedge L) \quad \text{whenever } 0 < K < L.$$

To see that it is independent of t we fix $0 < s < t$. Then $T_t(f) \geq (s/t)^{1/2} T_s(f)$ shows that if the hypothesis holds for s it will also hold for t . Conversely suppose it holds for t . Then, given $\varepsilon > 0$, there exists $\delta > 0$ and sets $\Omega_x \subseteq C_{\text{exp}}^+$ with measure $\nu(\Omega_x) \geq 1 - \varepsilon$ so that $T_t(f \wedge K)(x) \geq \delta$ for $f \in \Omega_x$. Choosing $L = L(K, \delta)$ large we have, writing $\chi_{[a,b]}$ for the indicator of an interval,

$$T_t((f \wedge K)\chi_{[x-L, x+L]})(x) \geq \frac{\delta}{2} \quad \text{for } f \in \Omega_x.$$

Comparing the Brownian densities at s and t over the interval $[-L, L]$ we find, for $f \in \Omega_x$,

$$\begin{aligned} T_s(f \wedge K)(x) &\geq T_s((f \wedge K)\chi_{[x-L, x+L]})(x) \geq \sqrt{\frac{t}{s}} \exp\left(\frac{L^2}{4t} - \frac{L^2}{4s}\right) T_t((f \wedge K)\chi_{[x-L, x+L]})(x) \\ &\geq \sqrt{\frac{t}{s}} \exp\left(\frac{L^2}{4t} - \frac{L^2}{4s}\right) \frac{\delta}{2} \end{aligned}$$

which shows the hypothesis will hold for s as well.

(3) An earlier, somewhat weaker, version of this result was contained in the Warwick PhD Thesis [6], to which we refer for various technical details.

2. Preliminary results

A comparison argument will allow us to obtain estimates for solutions to (1) by studying the equation

$$\partial_t \bar{u} = \partial_{xx} \bar{u} + \theta \bar{u} + \sqrt{2\bar{u}} dW. \quad (7)$$

Indeed, given $f \in C_{\text{exp}}^+$, we may construct solutions u and \bar{u} , both with initial condition f and with respect to the same white noise, so that they are coupled to satisfy $u(t, x) \leq \bar{u}(t, x)$ for all t, x , almost surely. See Mueller and Tribe [7, Lemma 2.1.4] for some details. As the noise coefficient is non-Lipschitz, and we do not know whether pathwise uniqueness holds, we cannot assert that any two solutions driven by the same noise will satisfy the coupling. But since uniqueness in law holds it will be sufficient for us to obtain one pair of coupled solutions. This and similar couplings were used in [7], so in this paper we omit the detailed proofs of these constructions and merely sketch the main ideas. First approximate by an equation with Lipschitz coefficients. For this approximation standard comparison theorems (see Donati-Martin and Pardoux [2, Theorem 2.1]) imply that the pair of solutions are coupled. Checking tightness, one extracts a convergent subsequence and the limit is the desired coupled pair of solutions to the original equation.

The solutions to (7) are the densities of Dawson–Watanabe processes with constant mass creation θ and many properties are known. We will mainly refer to the survey by Perkins [9] for the properties we need. In particular (see [11, Lemma 2.1]) solutions \bar{u} started from $f \in C_c^+$ retain compact support at all times $t \geq 0$. The comparison shows this also holds for solutions u to (1).

2.1. Death estimates

The process \bar{u} may die out in finite time (see [9, II.5.3]) and the exact probability of this is known: for any $f \in C_{\text{exp}}^+$ satisfying $(f, 1) < \infty$,

$$P_f[(\bar{u}(t), 1) = 0] = \exp\left(\frac{-\theta(f, 1)}{1 - e^{-\theta t}}\right). \quad (8)$$

More generally there is an exact formula for the probability that \bar{u} is zero on an interval $[a, b]$:

$$P_f[(\bar{u}(t), \chi_{[a,b]}) = 0] = \exp(-\theta(f, U_t^{a,b})), \quad (9)$$

where $U_t^{a,b} = \lim_{\lambda \uparrow \infty} U_t^{\lambda,a,b}$ and $U_t^{\lambda,a,b}$ is the unique non-negative solution to the PDE

$$\begin{cases} \partial_t u = \partial_{xx} u + \theta u - u^2 & \text{for } x \in \mathbb{R} \text{ and } t \geq 0, \\ u = \lambda \chi_{[a,b]} & \text{when } t = 0. \end{cases}$$

It is known (see [9, III.5.11]) that $U_t^{a,b}(x)$ is bounded and integrable. By the comparison $u \leq \bar{u}$, the formulae (8), (9) hold as lower bounds for the process u .

We also need a death estimate that exploits the non-linearity $-u^2$ in (1). We claim that, for any $t > 0$ there exists $p(\theta, t) > 0$ so that

$$P_f[u(t) = 0] \geq p(\theta, t) > 0 \quad \text{for all } f \in C_c^+ \text{ supported in } [0, 1]. \tag{10}$$

The point is that the lower bound on the death probability depends only on the size of the support of the initial condition f . This is possible because the $-u^2$ term in (1) can bring down infinite initial conditions in finite time. By scaling and translation invariance the specific choice of the interval $[0, 1]$ is irrelevant. To prove (10) we argue as follows:

$$\begin{aligned} P_f[u(2t) = 0] &= E_f[P_{u(t)}[u(t) = 0]] \quad \text{by the Markov property,} \\ &\geq E_f[P_{u(t)}[\bar{u}(t) = 0]] \quad \text{by the comparison } u \leq \bar{u}, \\ &= E_f\left[\exp\left(-\frac{\theta}{1 - e^{-\theta t}}(u(t), 1)\right)\right] \quad \text{by (8),} \\ &= E_{\frac{\theta}{1 - e^{-\theta t}}}\left[\exp(-u(t), f)\right] \quad \text{by self duality,} \\ &\geq P_{\frac{\theta}{1 - e^{-\theta t}}}\left[(u(t), \chi_{[0,1]}) = 0\right] \quad \text{since } f \text{ is supported in } [0, 1], \\ &\geq P_{\frac{\theta}{1 - e^{-\theta t}}}\left[(\bar{u}(t), \chi_{[0,1]}) = 0\right] \quad \text{by the comparison } u \leq \bar{u}, \\ &= \exp\left(-\frac{\theta}{1 - e^{-\theta t}}(U_t^{0,1}, 1)\right) := p(\theta, 2t) \quad \text{by (9).} \end{aligned}$$

Above we have written $\frac{\theta}{1 - e^{-\theta t}}$ for the constant function with this value.

2.2. Moments

We quote from Tribe [11] two simple moment estimates for the solutions u to (1) and \bar{u} to (7). Lemma 3.2 from [11] implies, for any $p > 0$ and $T, K < \infty$,

$$E_f[|u(t, x)|^p] \leq E_f[|\bar{u}(t, x)|^p] \leq C(p, \theta, T, K) \quad \text{for } x \in \mathbb{R}, t \leq T \text{ and } f \leq K, \tag{11}$$

and the argument from Lemma 3.4 in [11] implies, for any $p \geq 2$ and $T, K < \infty$,

$$E_f[|u(t, y) - u(t, x)|^p] \leq C(p, \theta, T, K)|y - x|^{(p/2)-1} \quad \text{for } t \leq T, |y - x| \leq 1 \text{ and } f \leq K.$$

By Levy’s chaining argument that obtains a modulus of continuity from increment moments, one can deduce that for a fixed $x \in \mathbb{R}$ there exists a random $\Delta(x, t) > 0$ so that

$$|u(t, x) - u(t, y)| \leq |x - y|^{1/3} \quad \text{for all } y \in [x - \Delta, x + \Delta], \text{ almost surely,} \tag{12}$$

and that $P_f[\Delta(x, t) < \delta] \rightarrow 0$ as $\delta \rightarrow 0$, uniformly over $x \in \mathbb{R}, t \leq T$ and $f \leq K$. (We do not need the best modulus of continuity here, which would be $1/2 - \varepsilon$.)

We also need simple estimates on the first moment. Write $T_t^\theta f = e^{\theta t} T_t f$ for the heat semigroup with exponential growth. The Green’s function representation for the solution to (1) implies

$$E_f[u(t, x)] = T_t^\theta f(x) - \int_0^t E_f[T_{t-s}^\theta(u_s^2)(x)] ds. \tag{13}$$

Discarding the second term we have the upper bound $E_f[u(t, x)] \leq T_t^\theta f(x)$. To get a lower bound on the first moment from (13) we need an upper bound on $E_f[u^2(t, x)]$. We use the comparison $u \leq \bar{u}$ and second moments of the Dawson–Watanabe process \bar{u} which are known explicitly; indeed when $g \geq 0$ we have (see [9, II.5.2.(b)]):

$$E_f[(\bar{u}(t), g)^2] = (T_t^\theta f, g)^2 + \int_0^t (T_{t-s}^\theta f, (T_s^\theta g)^2) ds. \tag{14}$$

Suppose that $f \leq K$ and that $(g, 1) = 1$, so that $(T_s^\theta f, g) \leq K e^{\theta s}$ and $T_s^\theta g(x) \leq e^{\theta s} (2\pi s)^{-1/2}$. Using these estimates in (14) we obtain

$$E_f[(\bar{u}(t), g)^2] \leq (T_t^\theta f, g) \left(K e^{\theta t} + \int_0^t e^{\theta s} (2\pi s)^{-1/2} ds \right) \leq (T_t^\theta f, g) e^{\theta t} (K + t^{1/2}).$$

In particular we see that $E_f[u^2(t, x)] \leq E_f[\bar{u}^2(t, x)] \leq e^{\theta t} (K + t^{1/2}) T_t^\theta f(x)$. Substituting this into (13) we obtain

$$(1 - (Kt + t^{3/2})e^{\theta t}) T_t^\theta f(x) \leq E_f[u(t, x)] \leq T_t^\theta f(x). \tag{15}$$

In particular, when t is small, the first moment is well approximated by $e^{\theta t} T_t f(x)$.

2.3. Mixing properties

We want to approximate solutions to (1) by solutions to the same equation over large finite intervals with Dirichlet boundary conditions. Fix $x_0, L > 0$ and the deterministic initial condition f and consider solutions u^{L, x_0} to

$$\begin{cases} \partial_t u = \partial_{xx} u + \theta u - u^2 + \sqrt{2u} dW & \text{for } |x - x_0| \leq L \text{ and } t \geq 0, \\ u = 0 & \text{for } x = x_0 \pm L \text{ and } t \geq 0, \\ u = f & \text{for } |x - x_0| \leq L \text{ and } t = 0. \end{cases} \tag{16}$$

We can construct a solutions u^{L, x_0} coupled to a solution u of (1) so that $u^{L, x_0}(t, x) \leq u(t, x)$ for all $t \geq 0$ and $x \in [x_0 - L, x_0 + L]$, almost surely. The noise used for u^{L, x_0} is the restriction of the noise used for u to the interval $[x_0 - L, x_0 + L]$. In Section 3.3 we need such a construction for countably many solutions u^{L, x_i} for $i = 1, 2, \dots$ all simultaneously coupled below a solution u to (1). This is also possible by the same techniques (see [7, Lemma 2.1.5]). Two solutions u^{L, x_i} and u^{L, x_j} will be independent provided that $|x_i - x_j| \geq 2L$. This is intuitive since they are driven by independent parts of the noise, and follows rigorously since it holds at each stage of the approximation scheme used to construct the couplings. This independence can then be used to establish a mixing property since, when L is large, the solution $u^{L, x_0}(t, x_0)$ should be close to the solution $u(t, x_0)$ of the original equation (1) on the whole line. The following lemma establishes this closeness.

Lemma 2. *Let u and u^{L, x_0} be the coupled solutions to (1) and (16) constructed as above with the same white noise and the same initial condition f . Then there exist $c_1 = c_1(\theta) > 0$ so that*

$$E_f[|u(t, x_0) - u^{L, x_0}(t, x_0)|] \leq C(\theta, T, K) \exp(-c_1 L) \quad \text{for } x_0 \in \mathbb{R}, t \leq T \text{ and } f \leq K.$$

Proof. Without loss of generality we may take $x_0 = 0$. Let $G_t(x, y)$ (respectively $G_t^{(L)}(x, y)$) be the Green’s function for the equation $\partial_t u = \Delta u + \theta u$ on the whole space (respectively on $[-L, L]$ with Dirichlet boundary conditions). Subtracting the two Green’s function representations for u and $u^{L, 0}$ we obtain, for $x \in (-L, L)$ and $t \leq T$,

$$\begin{aligned}
 0 &\leq u(t, x) - u^{L,0}(t, x) \\
 &= \int (G_t(x, y) - G_t^{(L)}(x, y))f(y) dy + \int_0^t \int (G_{t-s}^{(L)}(x, y)(u^{L,0}(s, y))^2 - G_{t-s}(x, y)u^2(s, y)) dy ds \\
 &\quad + \int_0^t \int (G_{t-s}(x, y)\sqrt{2u(s, y)} - G_{t-s}^{(L)}(x, y)\sqrt{2u^{L,0}(s, y)})W(dy, ds) \\
 &\leq \int (G_t(x, y) - G_t^{(L)}(x, y))f(y) dy + \int_0^t \int (G_{t-s}^{(L)}(x, y) - G_{t-s}(x, y))u^2(s, y) dy ds \\
 &\quad + \int_0^t \int (G_{t-s}(x, y)\sqrt{2u(s, y)} - G_{t-s}^{(L)}(x, y)\sqrt{2u^{L,0}(s, y)})W(dy, ds).
 \end{aligned}$$

Taking expectations, and using the moment bounds (11), we obtain

$$\begin{aligned}
 E[|u(t, x) - u^{L,0}(t, x)|] \\
 \leq \int (G_t(x, y) - G_t^{(L)}(x, y))f(y) dy + C(\theta, T, K) \int_0^t \int (G_{t-s}(x, y) - G_{t-s}^{(L)}(x, y)) dy ds.
 \end{aligned}$$

Now we choose $x = 0$. The difference in the Green’s functions $\int G_t(0, y) - G_t^{(L)}(0, y) dy$ can be bounded, for example, by the Brownian probability $e^{\theta t} P_0[B_s \text{ exits } [-L, L] \text{ before time } 2t]$. Straightforward estimates now give the required result. \square

3. The main argument

3.1. Consequences of the uniform non-degeneracy hypothesis (6)

The following lemma converts the hypothesis (6) into the form we shall need.

Lemma 3. *Suppose v satisfies the hypothesis (6). Then there exists $t_1 = t_1(\theta) > 0$ such that*

$$\lim_{\rho \downarrow 0} v\{f: P_{f \wedge 1}[u(t_1, x) > \rho] < \rho\} = 0 \quad \text{uniformly in } x \in \mathbb{R}. \tag{17}$$

Proof. By the second remark after the statement of Theorem 1, we may assume in the hypothesis (6) that $K = 1$ and also choose the value of t_0 . We choose to take $t_0 = t_1(\theta)$ where $(t_1 + t_1^{3/2})e^{\theta t_1} = 1/2$. Then, by the moment estimate from (15), we have

$$\frac{1}{2}e^{\theta t_1} T_{t_1}(f \wedge 1)(x) \leq E_{f \wedge 1}[u(t_1, x)] \leq e^{\theta t_1} T_{t_1}(f \wedge 1)(x).$$

We now use a simple estimate derived from the first and second moments. For a non-negative variable Z we have

$$E[Z] \leq \rho + E[Z\chi_{\{Z>\rho\}}] \leq \rho + (E[Z^2])^{1/2} (P[Z > \rho])^{1/2}.$$

Rearranging and taking $\rho = E[Z]/2$ gives

$$P\left[Z > \frac{1}{2}E[Z]\right] \geq \frac{(E[Z])^2}{4E[Z^2]}.$$

Applying this to $Z = u(t_1, x)$, and using the first moment bounds above and the second moment bounds from (11), we obtain for $\rho \leq \frac{1}{4}e^{\theta t_1} T_{t_1}(f \wedge 1)(x) \leq \frac{1}{2} E_{f \wedge 1}[u(t_1, x)]$

$$P_{f \wedge 1}[u(t_1, x) > \rho] \geq P_{f \wedge 1}\left[u(t_1, x) > \frac{1}{2} E_{f \wedge 1}[u(t_1, x)]\right] \geq c_2 (T_{t_1}(f \wedge 1))^2(x)$$

for some $c_2 = c_2(\theta, t_1) > 0$. Hence

$$v\{f: P_{f \wedge 1}[u(t_1, x) > \rho] < \rho\} \leq v\{f: c_2 (T_{t_1}(f \wedge 1))^2(x) < \rho\} + v\left\{f: \frac{1}{4} e^{\theta t_1} T_{t_1}(f \wedge 1)(x) \leq \rho\right\}$$

which converges to zero, uniformly in x , as $\rho \rightarrow 0$ by the hypothesis (6). \square

3.2. Growth conditioned on non-extinction

We need to show that, conditioned on non-extinction, a solution started from finite initial mass has growing total mass. The difficulty is to find an argument that works for all $\theta > \theta_c$, since the critical value θ_c is defined in a non-explicit way. Although one expects linear growth of the total mass it will be enough for us to show the following weak estimate.

Lemma 4. *Let v be a solution of (1) with initial condition $g \in C_c^+$. For each $V > 0$,*

$$\lim_{t \rightarrow \infty} P_g[0 < (v(t), 1) \leq V] = 0. \tag{18}$$

Moreover there exists $c_3 = c_3(\theta) < \infty$ so that for any $\delta \in (0, 1]$ and $V > 0$

$$P_g\left[0 < \sum_{k \in \delta\mathbb{Z}} \min\left\{\int_k^{k+\delta} v(t, x) dx, c_3\right\} \leq V\right] \rightarrow 0 \text{ as } t \rightarrow \infty. \tag{19}$$

Indeed we may take $c_3(\theta) = -\frac{1-e^{-\theta}}{\theta} \log p(\theta, 1)$, where $p(\theta, t)$ is from (10).

Remark. Using (18) in the self duality relation one can quickly show that the distribution of the solution $u(t)$, started from any initial condition f satisfying $f(x) \geq \delta > 0$ for all $x \in \mathbb{R}$, will converge to the stationary distribution given by (5). The problem is that typically random initial conditions of interest (in particular stationary distributions) will have regions where they are zero and will not be bounded below. Thus we will need the slightly more detailed information on the growth given in the second part of the lemma.

Proof. We start by following the argument in the proof of Theorem 3.3 in Durrett [3]. Recall that $\tau = \inf\{t: v(t) = 0\}$ is the death time. Note that for each $t \geq 0$ and $V > 0$,

$$\begin{aligned} P_g[\tau \in (t, t + 1]] &= P_g[v(t) \neq 0 \text{ and } v(t + 1) = 0] \geq P_g[0 < (v(t), 1) \leq V \text{ and } v(t + 1) = 0] \\ &= P_g[v(t + 1) = 0 \mid 0 < (v(t), 1) \leq V] P_g[0 < (v(t), 1) \leq V]. \end{aligned}$$

Using the Markov property and the death time estimate (8), we have

$$P_g[v(t + 1) = 0 \mid 0 < (v(t), 1) \leq V] \geq \exp\left(\frac{-\theta V}{1 - e^{-\theta}}\right). \tag{20}$$

Rearranging we find

$$P_g[0 < (v(t), 1) \leq V] \leq \exp\left(\frac{\theta V}{1 - e^{-\theta}}\right) P_g[\tau \in (t, t + 1]].$$

Letting $t \rightarrow \infty$ shows the first part of the lemma. For the second part we use another coupling construction used before in Lemma 2.1.7 of [7], which we briefly describe. Dawson–Watanabe processes are additive, in that the sum of two independent solutions $\bar{v}_1 + \bar{v}_2$, started from initial conditions g_1 and g_2 respectively, is a new solution started at $g_1 + g_2$. The solutions to (1) are super-additive in that the sum of two independent solutions is stochastically larger than a solution. This is intuitive from the derivation of solutions to (1) as limits of particle systems. The non-linearity $-v^2$ comes from a death term due to two particles colliding. Summing two independent solutions of say red and blue particles, allows deaths due to collisions between particles of the same colour, but ignores the deaths due to collisions between one particle of each colour, and is therefore healthier than the true solution where all collisions lead to deaths. Mathematically we can construct a coupling as follows. Suppose $g \in \mathcal{C}_c^+$ can be written as $g = \sum_k g_k$ for $g_k \in \mathcal{C}_c^+$. Then we can find independent solutions v_k to (1), started at g_k , and a further solution u started at g so that $v(t, x) \leq \sum_k v_k(t, x)$ for all t, x , almost surely.

We use this coupling with $g_k = g\chi_{[k, k+\delta]}$ where $\delta \in (0, 1]$ and $k \in \delta\mathbb{Z}$. Using the estimates on the death probabilities in (8), (10), we have

$$P_g[v(1) = 0] \geq \prod_k P_{g_k}[v(1) = 0] \geq \prod_k \max\{e^{-\frac{\theta}{1-e^{-\theta}}(g_k, 1)}, p(\theta, 1)\} = \exp\left(-\frac{\theta}{1-e^{-\theta}} \sum_k (g_k, 1) \wedge c_3\right).$$

The second half of the lemma now follows by repeating the argument for the first part, but conditioning on the event in (19) and using the above death estimate in place of (20). \square

3.3. Key estimate

The tricks used below to obtain the main estimate follow, in the main, those from the account of Harris’s theorem given in Chapter 3 of Durrett [3].

Take $u(t)$ a solution to (1) with an initial distribution ν satisfying the hypothesis (6). The main estimate shows that $E[\exp(-u(t_1), g)]$ will be small for certain good test functions g . We fix $\varepsilon > 0$ and choose a number of parameter values ρ, V, δ and L . They may make little sense until the main argument starts but we choose them now so as to make clear their dependence only on ε and θ . Moreover for notational simplicity we will not indicate their dependence on θ .

Firstly, using Lemma 3, we choose $\rho = \rho(\varepsilon) \in (0, 1]$ so that

$$\nu\{f: P_{f \wedge 1}[u(t_1, x) > \rho] < \rho\} \leq \varepsilon \quad \text{for all } x \in \mathbb{R}, \tag{21}$$

where $t_1(\theta)$ is chosen as in Lemma 3. Secondly, we choose $V = V(\varepsilon) \geq c_4 \vee 1$ so that

$$\exp(-c_4 \rho^2 V) \leq \varepsilon, \tag{22}$$

where $c_4 = \frac{1-e^{-c_3}}{c_3}$ and $c_3 < \infty$ is the value occurring in Lemma 4. Thirdly, using the moment bound (11) and the modulus of continuity (12), we have

$$E_f\left[\left|\inf_{y \in [x, x+\delta]} u(t_1, y) - u(t_1, x)\right|\right] \leq E_f[u(t_1, x)\chi_{\{\Delta(t_1, x) \leq \delta\}}] + \delta^{1/3} \leq C(P_f[\Delta \leq \delta])^{1/2} + \delta^{1/3}.$$

Therefore we may choose $\delta = \delta(\varepsilon) \in (0, 1]$ so that

$$E_f\left[\left|\inf_{y \in [x, x+\delta]} u(t_1, y) - u(t_1, x)\right|\right] \leq \frac{\varepsilon}{2V} \quad \text{for } x \in \mathbb{R} \text{ and } f \leq 1. \tag{23}$$

Fourthly, using Lemma 2, we choose $L = L(\varepsilon) \geq 1$ so that

$$E_f\left[|u(t_1, x) - u^{L, x}(t_1, x)|\right] \leq \frac{\varepsilon}{2V} \quad \text{for } x \in \mathbb{R} \text{ and } f \leq 1, \tag{24}$$

where $u^{L, x}$ is the solution of the equation on the interval $[x - L, x + L]$ with Dirichlet boundary conditions as described before Lemma 2.

Finally we say that the test function g is ε -good if there exist intervals $[x_j, x_j + \delta]$, with $x_j \in \delta\mathbb{Z}$, that satisfy

- (i) $|x_{j+1} - x_j| \geq 2L$ for each j ,
- (ii) $\sum_j \lambda_j \geq V$, where λ_j is defined by $\lambda_j = \min\{\int_{x_j}^{x_j+\delta} g(x) dx, c_3\}$.

We now start the main estimate. Fix an ε -good g . By throwing out some intervals if necessary we may assume that $\lambda_j > 0$ and that $\sum_j \lambda_j \in [V, 2V]$. Using the fact that we can couple a solution started at f above a solution started at $f \wedge 1$, we have

$$\begin{aligned} E[\exp(-(u(t_1), g))] &\leq \int_{C_{\text{exp}}^+} v(df) E_f[\exp(-(u(t_1), g))] \leq \int_{C_{\text{exp}}^+} v(df) E_{f \wedge 1}[\exp(-(u(t_1), g))] \\ &\leq \int_{C_{\text{exp}}^+} v(df) E_{f \wedge 1} \left[\exp\left(-\sum_j \lambda_j \inf_{y \in [x_j, x_j+\delta]} u(t_1, y)\right) \right] \\ &\leq \int_{C_{\text{exp}}^+} v(df) E_{f \wedge 1} \left[\exp\left(-\sum_j \lambda_j u(t_1, x_j)\right) \right] + \text{Error}_1, \end{aligned}$$

where, using $|e^{-\sum a_j} - e^{-\sum b_j}| \leq \sum |a_j - b_j|$, we can bound $|\text{Error}_1| \leq \sum_j \lambda_j \delta \leq \varepsilon$ by the choice (23) of δ and the fact that $\sum \lambda_j \leq 2V$. We continue

$$\begin{aligned} \int_{C_{\text{exp}}^+} v(df) E_{f \wedge 1} \left[\exp\left(-\sum_j \lambda_j u(t_1, x_j)\right) \right] &\leq \int_{C_{\text{exp}}^+} v(df) E_{f \wedge 1} \left[\exp\left(-\sum_j \lambda_j u^{L, x_j}(t_1, x_j)\right) \right] \\ &= \int_{C_{\text{exp}}^+} v(df) \prod_j E_{f \wedge 1}[\exp(-\lambda_j u^{L, x_j}(t_1, x_j))] \\ &\leq \int_{C_{\text{exp}}^+} v(df) \prod_j E_{f \wedge 1}[\exp(-\lambda_j u(t_1, x_j))] + \text{Error}_2. \end{aligned}$$

The equality here comes from the fact that the solutions u^{L, x_j} are independent over j . Using the inequality $|\prod_i a_i - \prod_i b_i| \leq \sum_i |a_i - b_i|$ when $a_i, b_i \in [0, 1]$, we can estimate the error term by $|\text{Error}_2| \leq \varepsilon$ by the choice (24) of L . Finally we use Hölder’s inequality, with $\sum_j p_j^{-1} = 1$, and the choice (21) of ρ to estimate

$$\begin{aligned} \int_{C_{\text{exp}}^+} v(df) \prod_j E_{f \wedge 1}[\exp(-\lambda_j u(t_1, x_j))] &\leq \prod_j \left(\int_{C_{\text{exp}}^+} v(df) (E_{f \wedge 1}[\exp(-\lambda_j u(t_1, x_j))]^{p_j}) \right)^{1/p_j} \\ &\leq \prod_j (\varepsilon + (1 - \rho + \rho e^{-\lambda_j \rho})^{p_j})^{1/p_j} \\ &\leq \prod_j (\varepsilon + \exp(-c_4 \rho^2 \lambda_j p_j))^{1/p_j}. \end{aligned}$$

The final inequality here follows from

$$1 - \rho + \rho e^{-\lambda_j \rho} \leq 1 - \rho^2 \frac{1 - e^{-c_3}}{c_3} \lambda_j = 1 - c_4 \rho^2 \lambda_j \leq \exp(-c_4 \rho^2 \lambda_j)$$

which holds since $1 - z \leq e^{-z} \leq 1 - (\frac{1-e^{-K}}{K})z$ when $z \in [0, K]$. Choosing $p_j = (\sum_i \lambda_i) / \lambda_j$ we obtain

$$\prod_j (\varepsilon + \exp(-c_4 \rho^2 p_j \lambda_j))^{1/p_j} = \varepsilon + \exp\left(-c_4 \rho^2 \sum_i \lambda_i\right) \leq \varepsilon + \exp(-c_4 \rho^2 V) \leq 2\varepsilon$$

by the choice (22) of V . Collecting these estimates we have, for ε -good g , that

$$E[\exp(-(u(t_1), g))] \leq 4\varepsilon. \tag{25}$$

3.4. Completion of the proof of Theorem 1

Take $u(t)$ a solution to (1) with an initial distribution ν satisfying the hypothesis (6). Let $v(t)$ be an independent solution started at $g \in \mathcal{C}_c^+$. Choose $t_1(\theta)$ as in Lemma 3. Applying the duality relation (2) using the time $s = t_1$ we have, for $t > t_1$,

$$\begin{aligned} E[e^{-(u(t), g)}] &= E[e^{-(u(t_1), v(t-t_1))}] = P_g[\tau \leq t - t_1] + E[e^{-(u(t_1), v(t-t_1))} \chi_{\{v(t-t_1) \neq 0\}}] \\ &= P_g[\tau \leq t - t_1] + E[e^{-(u(t_1), v(t-t_1))} \chi_{\{v(t-t_1) \text{ is } \varepsilon\text{-good}\}}] \\ &\quad + E[e^{-(u(t_1), v(t-t_1))} \chi_{\{0 \neq v(t-t_1) \text{ is not } \varepsilon\text{-good}\}}]. \end{aligned} \tag{26}$$

Using the main estimate (25) from the last subsection, the second term on the right-hand side is bounded by 4ε . By Lemma 4 we know that

$$P_g \left[0 < \sum_{k \in \delta\mathbb{Z}} \min \left\{ \int_k^{k+\delta} v(t, x) dx, c_3 \right\} \leq \left(\frac{2L}{\delta} + 1 \right) V \right] \rightarrow 0 \quad \text{as } t \rightarrow \infty. \tag{27}$$

This implies that $P_g[0 \neq v(t) \text{ is not } \varepsilon\text{-good}] \rightarrow 0$ as $t \rightarrow \infty$. Indeed consider an infinite lattice of intervals of size δ , where the intervals are separated by $2L$. There are $2L/\delta$ disjoint lattices of this type and (27) ensures that with high probability at least one of these lattices satisfy the conditions for $v(t)$ to be ε -good.

Letting $t \rightarrow \infty$ in (26) we conclude that

$$\lim_{t \rightarrow \infty} E[e^{-(u(t), g)}] = P_g[\tau < \infty].$$

This shows that the distribution of $u(t)$ converges to a limiting distribution μ characterized by (5). Standard arguments for Feller process imply that μ is a stationary distribution. It satisfies $\mu\{f \neq 0\} > 0$ since $\theta > \theta_c$ and is translation invariant since $P_g[\tau < \infty]$ is unchanged by shifts of g . We now know that any ν satisfying the hypothesis (6) is in the domain of attraction of μ .

Suppose now that ν is any translation invariant stationary measure satisfying $\nu\{f \neq 0\} = q > 0$. Then $\tilde{\nu}(A) = q^{-1} \nu(A \cap \{f \neq 0\})$ defines another translation invariant stationary distribution. Moreover $\tilde{\nu}\{f: T_{t_0}(f \wedge K)(x) < \delta\}$ is independent of x by translation invariance and converges to zero as $\delta \downarrow 0$ since $\tilde{\nu}\{f \neq 0\} = 1$. Thus $\tilde{\nu}$ satisfies the hypothesis (6) and hence must equal μ . In particular this shows that $\mu\{f \neq 0\} = 1$ and that the set of translation invariant stationary measures is precisely $\{\theta\mu + (1 - \theta)\delta_0: \theta \in [0, 1]\}$, where δ_0 is a point mass on the zero function.

References

[1] S. Athreya, R. Tribe, Uniqueness for a class of one-dimensional stochastic PDEs using moment duality, Ann. Probab. 28 (2002) 1711–1734.
 [2] C. Donati-Martin, E. Pardoux, White noise driven SPDEs with reflection, Prob. Theory Related Fields 95 (1993) 1–24.
 [3] R. Durrett, Ten lectures on particle systems, in: Lecture Notes in Mathematics, vol. 1608, Springer, Berlin, 1993, pp. 97–201.
 [4] S.N. Ethier, T.G. Kurtz, Markov Processes, Characterization and Convergence, Wiley, 1986.
 [5] T.E. Harris, On a class of set valued Markov processes, Ann. Probab. 4 (1976) 175–194.

- [6] P. Horridge, Some methods for proving uniqueness of stationary distributions for stochastic PDEs, Warwick University, PhD Thesis, 2001.
- [7] C. Mueller, R. Tribe, A phase transition for a stochastic PDE related to the contact process, *Probab. Theory Related Fields* 100 (1994) 131–156.
- [8] C. Mueller, R. Tribe, Stochastic p.d.e.'s arising from the long range contact and long range voter processes, *Probab. Theory Related Fields* 102 (1995) 519–545.
- [9] E. Perkins, Dawson–Watanabe superprocesses and measure valued diffusions, in: *Lecture Notes in Mathematics*, vol. 1781, Springer, Berlin, 2002, pp. 125–324.
- [10] T. Shiga, Two contrasting properties of solutions for one-dimensional stochastic differential equations, *Canadian J. Math.* 46 (1993) 415–437.
- [11] R. Tribe, A travelling wave solution to the Kolmogorov equation with noise, *Stochastics Stochastics Rep.* 56 (1996) 317–340.