Large fluctuations for a nonlinear heat equation with noise

1982 J. Phys. A: Math. Gen. 15 3025

(http://iopscience.iop.org/0305-4470/15/10/011)

J. Phys. A: Math. Gen. 15 (1982) 3025–3055. Printed in Great Britain

Large fluctuations for a nonlinear heat equation with noise†

William G Faris‡ and Giovanni Jona-Lasinio§
Institut des Hautes Etudes Scientifiques, 91440 Bures-sur-Yvette, France

Received 22 February 1982

Abstract. We study a nonlinear heat equation in a finite interval of space subject to a white noise forcing term. The equation without the forcing term exhibits several equilibrium configurations, two of which are stable. The solution of the complete forced equation is a stochastic process in space and time that has a unique stochastic equilibrium. We study this process in the limit of small noise, and obtain lower and upper bounds for the probability of large fluctuations. We then apply these estimates to calculate the transition probability between the stable configurations (tunnelling). This model problem can be interpreted as a rigorous version of some recent attempts to describe Euclidean quantum systems in terms of stochastic equilibrium states of a nonlinear stochastic differential equation in infinite dimensions. However, its significance goes beyond this situation and our methods may be applicable to models in other areas of natural science.

1. Introduction

Stochastic partial differential equations constitute a relatively new subject. There have been a number of interesting papers in the last few years, but the field is still in its infancy if compared with the highly developed theory of stochastic ordinary differential equations. We therefore consider it reasonable to spend a few words illustrating some basic motivations arising in different areas of science before introducing the specific problem of our study.

An example of primary importance in physics is provided by hydrodynamics. As is well known, the behaviour of an incompressible viscous fluid is usually described in terms of the Navier–Stokes equation. However, this equation is known to be approximate in more than one aspect. It takes into account only approximately the microscopic nature of a classical field. In addition quantum effects and other sources of fluctuations are completely ignored. It is therefore of interest to know which properties described by the Navier–Stokes equations survive perturbations, in particular small stochastic perturbations which imitate some of the neglected effects. This latter problem is also of special importance in connection with modern theories of turbulence, where one would like to determine physically interesting measures invariant under the flow generated by the Navier–Stokes equations and stable under small perturbations.

† This material is based on work supported in part by the National Science Foundation under Grant No MCS-8002945.
‡ Permanent address: Department of Mathematics, University of Arizona, Tucson, Arizona 85721 USA.
§ Permanent address: Istituto di Fisica, Università di Roma, Roma, Italia.

A vast class of evolution equations, which according to the usual terminology of partial differential equations may be classified as semilinear parabolic equations, arises in the phenomenological approach to such different phenomena as the diffusion of a fluid in a porous medium, transport in a semiconductor, coupled chemical reactions with possibility of spatial diffusion, and population genetics (Henry 1981). In all these cases, due to the phenomenological approximate character of the equations, it is again of interest to test how the description changes under the effect of stochastic perturbation.

More recently, mathematical developments in statistical mechanics, quantum mechanics and quantum field theory seem to lead in various ways to the study of stochastic differential equations and more generally of a calculus in infinitely many dimensions. In this connection, let us mention two specific ways in which infinite-dimensional stochastic differential equations arise. If one tries to extend to field theory the so-called method of stochastic quantisation, one is confronted with the problem of defining diffusions in Hilbert spaces. This aspect is reflected for example in the approach to field theory developed by Albeverio and Hoegh-Krohn (1977). Also some recent work on the quantisation of Yang–Mills fields (Asorey and Mitter 1981) involves stochastic differential equations in Hilbert space. There is however a different path to infinite-dimensional diffusions which in recent years has attracted the attention of people working in both statistical mechanics (Holley and Stroock 1976a, b) and quantum field theory (Parisi and Wu 1981). This is based on the following remark. Suppose we have to study an equilibrium situation described by a formal density proportional to

$$\exp[-(2/\varepsilon^2)S(u)], \tag{1.1}$$

where $S(u)$ is a functional of the $u(x)$ which are field variables, functions or distributions defined on n-dimensional space. We can imagine that the equilibrium is the result over long times of the, so far formal, diffusion process

$$\partial u(x,t)/\partial t = -\delta S(u)/\delta u(x,t) + \varepsilon\alpha(x,t), \tag{1.2}$$

where $\alpha(x,t)$ is a white noise with $(n+1)$-dimensional parameter. Of course, the problem of giving a precise meaning to equation (1.2) is a non-trivial one, and a general theory of infinite-dimensional diffusion has so far been developed only under certain restrictive hypotheses (Faris 1979, 1980, Marcus 1974, 1978, 1979, Belopolskaia and Daletzki 1978, Kozlov 1978, Doss and Royer 1978, Royer 1979). Even within these limits, however, it is possible to gain insight into some interesting examples.

Broadly speaking, the present work fits into the latter scheme. However, the model we shall discuss has an independent interest, and it proves to be an excellent laboratory for the development of new techniques. The mathematical goal of our paper is in fact the generalisation to an infinite-dimensional situation of the well known theory of small random perturbations developed by Ventsel' and Freidlin (1970, 1979). This theory has already been useful in treating tunnelling phenomena from the standpoint of stochastic mechanics (Jona-Lasinio *et al* 1981).

The model we shall discuss is directly inspired by the quantum mechanical double well anharmonic oscillator. The field variables u are functions defined on a one-dimensional interval [0, L]. The functional $S(u)$ appearing in (1.2), and which from now on will be referred to as the equilibrium action functional, is

$$S(u) = \int_0^L \left[\tfrac{1}{2}(\partial u/\partial x)^2 + V(u)\right] \mathrm{d}x, \tag{1.3}$$

with

$$V(u) = (\lambda/4)u^4 - (\mu/2)u^2, \tag{1.4}$$

where $\lambda > 0$ and $\mu > 0$ are fixed parameters. The reason for this choice of V is that it is the simplest even polynomial that is not quadratic. The signs are chosen so as to ensure two distinct minima. Equation (1.2) becomes

$$\partial u/\partial t = \partial^2 u/\partial x^2 - V'(u) + \varepsilon\alpha, \tag{1.5}$$

where α is a white noise in two dimensions. This means that α is a random Gaussian distribution with zero mean and covariance

$$E(\alpha(x, t)\alpha(x', t')) = \delta(x - x')\delta(t - t'), \tag{1.6}$$

where E denotes the expectation. In this paper [0, L] will be a fixed interval and we shall assume that our functions $u(x, t)$ satisfy Dirichlet boundary conditions

$$u(0, t) = u(L, t) = 0. \tag{1.7}$$

That is, if $\partial u/\partial x$ is not in L^2 or if u does not satisfy the Dirichlet boundary conditions, then $S(u) = +\infty$. The reason for the restriction to a finite interval is that a theory of large fluctuations for equation (1.5) can be developed along similar lines as in the finite-dimensional case. The situation where the interval $[0, L]$ is replaced by the whole real line requires additional consideration. Of course it is only in the latter case that the full quantum mechanical situation is recovered in equilibrium. However, one may also think of the equation (1.5) as describing the motion of an elastic string in a high-viscosity noisy environment. This interpretation is actually a useful reference for intuition.

The picture that emerges from our analysis is the following. For sufficiently large values of $\mu^{1/2}L$, the deterministic string defined by (1.5) for $\varepsilon = 0$ will have two stable equilibrium positions $\pm u_1$ corresponding to the two minima of the potential $V(u)$ at $u = \pm(\mu/\lambda)^{1/2}$. These are depicted in figure 1. They are the absolute minima of the equilibrium action $S(u)$. There are also a certain number (depending on the value of $\mu^{1/2}L$) of instanton-like or multi-instanton-like unstable equilibria. Figure 2 shows one of these, corresponding to one instanton. All of these solutions are critical points of the equilibrium action $S(u)$. The instanton or multi-instanton solutions are saddle points.

When the noise $\varepsilon\alpha$ is introduced the string will most likely perform small fluctuations near the stable configurations, but from time to time a particularly lucky

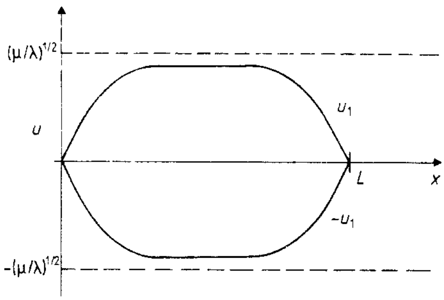

Figure 1. The stable configurations.

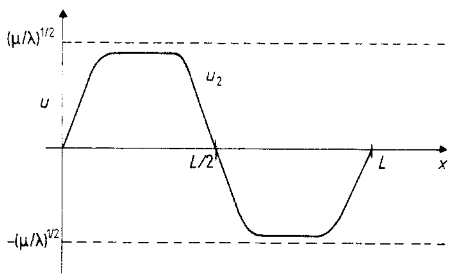

Figure 2. An instanton-like solution.

fluctuation will take it from one stable position to the other. We may call this event a tunnelling. The probability of tunnelling in a fixed large time interval will be very small, and its rough magnitude will be independent of the length of the interval. We shall show that this magnitude is of the order of

$$\exp\{-(2/\varepsilon^2)[S(u_2)-S(u_1)]\}, \tag{1.8}$$

where u_2 is the instanton-like solution (figure 2), and u_1 is a stable equilibrium solution. We must emphasise that this picture is strictly connected with the circumstance that the space interval $[0, L]$ is finite.

We now give an outline of our strategy. The standard way to approach an equation like (1.5) is to convert it into an integral equation using the solution of the linear part. In our case this is just the heat equation. Let $g(x, y, t)$ be the kernel of the integral operator that solves the initial value problem for the heat equation with Dirichlet boundary conditions (1.7) on $[0, L]$. Let u_0 be the initial value and let gu_0 be the solution. Thus

$$gu_0(x, t) = \exp(t\partial^2/\partial x^2)u_0(x, t) = \int_0^L g(x, y, t)u_0(y)\,\mathrm{d}y. \tag{1.9}$$

The operator G that gives the solution of the inhomogeneous heat equation with zero initial condition is then

$$\begin{aligned} Gf(x, t) &= (\partial/\partial t - \partial^2/\partial x^2)^{-1}f(x, t) \\ &= \int_0^T\int_0^L G(x, t, y, s)f(y, s)\mathrm{d}y\,\mathrm{d}s \\ &= \int_0^T\int_0^L g(x, y, t-s)\theta(t-s)f(y, s)\,\mathrm{d}y\,\mathrm{d}s, \end{aligned} \tag{1.10}$$

where θ is the indicator function of $[0, \infty)$. The method is to compare the process u of interest with the process w satisfying

$$(\partial/\partial t - \partial^2/\partial x^2)w = \alpha \tag{1.11}$$

and zero initial conditions. This linear equation has the solution

$$w = G\alpha \tag{1.12}$$

and so w is a zero-mean Gaussian process. The integral equation satisfied by u is then

$$u = -GV'(u) + \varepsilon w + gu_0.$$

The w that occurs in this equation is still random, but now it has a continuous covariance

$$E(w(x, t)w(x', t')) = \int_0^T\int_0^L G(x, t, y, s)G(x', t', y, s)\mathrm{d}y\,\mathrm{d}s = GG^*(x, t, x', t'), \tag{1.14}$$

where G^* is the adjoint operator of G.

The existence and uniqueness of global solutions of (1.13), continuous with probability one, follows from an adaptation of the work of Marcus (1978). (Similar results are contained in work of Kozlov (1978).) The ergodicity of the process and the existence of a unique equilibrium measure (1.1) also follow.

The theory of large fluctuations is then constructed as follows. Ventsel' and Freidlin (1979) have developed a theory of large fluctuations for Gaussian processes with values in Hilbert spaces. This has subsequently been cast in a more general form suitable for Gaussian processes with values in a Banach space in work of Azencott (1980). These results apply to the Gaussian process w. Define a functional I_0 of the sample paths by

$$I_0(w) = \tfrac{1}{2} \int_0^T \int_0^L [(\partial/\partial t - \partial^2/\partial x^2) w]^2 \, dx \, dt. \tag{1.15}$$

If A is a Borel set in the space of continuous functions w satisfying the Dirichlet boundary conditions and vanishing at $t = 0$, we define

$$I_0(A) = \inf_{w \in A} I_0(w). \tag{1.16}$$

It is plausible that this functional should play a role in the large fluctuation theory of the Gaussian process w, since it involves the inverse $(\partial/\partial t - \partial^2/\partial x^2)^*(\partial/\partial t - \partial^2/\partial x^2)$ of the covariance operator GG^*.

There is also a large fluctuation theory for the full nonlinear process u. This involves a functional I that will be called the action functional of the process (not to be confused with the equilibrium action functional S). Let

$$I(u) = \tfrac{1}{2} \int_0^T \int_0^L [\partial u/\partial t - \partial^2 u/\partial x^2 + V'(u)]^2 \, dx \, dt. \tag{1.17}$$

If A is a Borel set in the space of continuous functions u satisfying Dirichlet boundary conditions and an initial condition $u(x, 0) = u_0(x)$, then we may also define

$$I(A) = \inf_{u \in A} I(u). \tag{1.18}$$

The Ventsel'–Freidlin estimates for the process come in two parts, a lower bound for the probability of an open set, and an upper bound for the probability of a closed set. This pattern is equivalent to that in the large fluctuation results of Varadhan (1966). The first estimate states that if A is open, then

$$-I(A) \leq \liminf_{\varepsilon \to 0} \varepsilon^2 \log P^\varepsilon(u \in A). \tag{1.19}$$

The second states that if $\bar{A}$ is closed, then

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log P^\varepsilon(u \in \bar{A}) \leq -I(\bar{A}). \tag{1.20}$$

In these formulae P^ε is the probability when u is determined with noise parameter ε.

The proof of these estimates is in two steps. First the corresponding result is proved for the Gaussian process εw and the functional I_0, as mentioned above. Then one uses the observation of Ventsel' and Freidlin (1979) that such asymptotic results are preserved under continuous mappings. We establish that the transformation mapping the trajectory εw into the solution u of (1.13) is continuous in the sup norm. The estimates (1.19) and (1.20) for the non-Gaussian process follow immediately. These general large fluctuation results are summarised in § 6.

When $\bar{A}$ is a closed set, then there is a minimising trajectory u with $I(u) = I(\bar{A})$, but it may be very difficult to compute explicitly. However, in order to compute the probability of tunnelling we need only a special form of the event A; it should be a transition from the initial u_0 to some open set Y in some fixed time T. Thus we take

$$A = \{u : u(\cdot, 0) = u_0, u(\cdot, T) \in Y\}, \tag{1.21}$$

where the Dirichlet boundary conditions are understood. Then $\bar{A}$ is a transition from u_0 to the closed set $\bar{Y}$ at time T, and we may use A and $\bar{A}$ in the two Ventsel'–Freidlin estimates.

For the transition to be a tunnelling from $-u_1$ to u_1, we need to take the initial u_0 to lie in a small neighbourhood N of $-u_1$ and the final $u(\cdot, T)$ to lie in a small neighbourhood Y of u_1. Notice that these are uniform neighbourhoods and so there is nothing to ensure that the equilibrium action $S(u)$ will be finite for u in these neighbourhoods. It is thus remarkable that there are bounds in terms of the change

$$\Delta S = S(u_2) - S(u_1) \tag{1.22}$$

in the equilibrium action between the stable u_1 and the one-instanton state u_2.

The estimates needed to apply the lower and upper bounds on the probability are upper and lower bounds on the functional. The upper bound on the functional states that there is a neighbourhood N of $-u_1$ so that for every $\zeta > 0$, there is a $T < \infty$ so that

$$I(A) \leqslant 2\Delta S + \zeta. \tag{1.23}$$

This bound is proved in § 9 by computing $I(u)$ with a suitable trial function u in A. The lower bound states that there is a neighbourhood Y of u_1 so that for every $\zeta > 0$ and every compact set K of initial conditions, there is a neighbourhood N of $-u_1$ such that if u_0 is in $K \cap N$, then

$$2\Delta S - \zeta \leqslant I(\bar{A}). \tag{1.24}$$

This bound is more difficult, since in order to get a lower bound on $I(\bar{A})$, we need to get a lower bound on $I(u)$ for all u in $\bar{A}$, that is, for all conceivable tunnelling trajectories. This requires a topological argument that involves the critical point structure of the equilibrium action S. This is a point where the infinite-dimensional theory exhibits new difficulties, since S is not continuous on the space of continuous functions with the uniform norm. These difficulties are resolved in § 10.

We may summarise our results in a theorem. Let $A(u_0, Y; T)$ be the transition from u_0 to Y in time T given in (1.21). Let $P^\varepsilon(A(u_0, Y; T))$ be the probability of this event when the noise parameter is $\varepsilon \neq 0$. We wish to estimate the probability of a tunnelling from u_0 near $-u_1$ to a neighbourhood Y of u_1. Here $\pm u_1$ are the two ground states. The theorem gives bounds on this probability in terms of $\Delta S = S(\pm u_2) - S(\pm u_1)$, the difference of the equilibrium action between the one-instanton states $\pm u_2$ and the ground states $\pm u_1$.

Theorem 1.1. There is a neighbourhood Y of u_1 such that for every $\zeta > 0$ and every compact set K of continuous functions, there is a neighbourhood N of $-u_1$ and a time T, such that for all u_0 in $K \cap N$ and all ε sufficiently small

$$\exp[-(1/\varepsilon^2)(2\Delta S + \zeta)] \leqslant P^\varepsilon[A(u_0, Y; T)] \leqslant \exp[-(1/\varepsilon^2)(2\Delta S - \zeta)]. \tag{1.25}$$

This result suggests the picture that in the limit of small noise, tunnelling takes place through an instanton configuration.

2. The Gaussian process

In this section we derive some properties of the Gaussian process w defined by the equations (1.6) for the white noise α and the equations (1.11) and (1.12) expressing w in terms of α. Since the white noise is a random distribution, the equation (1.6) really means that

$$E(\alpha(f)\alpha(g)) = \langle f, g\rangle, \tag{2.1}$$

where $\langle f, g\rangle = \int_0^L f(x, t)g(x, t)\,\mathrm{d}x\,\mathrm{d}t$ is the L^2 inner product of two test functions f and g. Similarly, the equation (1.12) means that

$$\langle w, f\rangle = \langle \alpha, G^*f\rangle \tag{2.2}$$

for every test function f. Here

$$G = (\partial/\partial t - \partial^2/\partial x^2)^{-1}, \tag{2.3}$$

where the operator $\partial^2/\partial x^2$ is defined with Dirichlet boundary conditions at $x = 0$ and $x = L$, and the $\partial/\partial t$ is defined with the boundary condition that functions in its domain vanish when $t = 0$. The operator $\partial^2/\partial x^2$ is self-adjoint, but the $\partial/\partial t$ is not. In fact $(\partial/\partial t)^*$ is $-\partial/\partial t$ with the boundary condition that functions in its domain vanish at the other end point $t = T$. The operator G is given explicitly by (1.10) as an integral operator. We shall see below that $\mathrm{Tr}(GG^*) < \infty$, so G is actually Hilbert–Schmidt.

It follows from (2.1) and (2.2) that

$$E(w(f)w(g)) = \langle G^*f, G^*g\rangle = \langle f, GG^*g\rangle. \tag{2.4}$$

Thus the covariance of w is the operator

$$\Gamma = GG^*, \tag{2.5}$$

where G^* is the adjoint of G. In other words, we have

$$E(w(x, t)w(y, s)) = \Gamma(x, t, y, s), \tag{2.6}$$

where

$$\Gamma(x, t, y, s) = \int_0^T\int_0^L G(x, t, x', t')G(y, s, x', t')\,\mathrm{d}x'\,\mathrm{d}t'. \tag{2.7}$$

If we write this as an operator valued function of t and s, we obtain

$$\begin{aligned}\Gamma(t, s) &= \int_0^T \exp[(t-t')\partial^2/\partial x^2]\theta(t-t')\exp[(s-t')\partial^2/\partial x^2]\theta(s-t')\,\mathrm{d}t' \\ &= \int_0^T \exp[(t+s-2t')\partial^2/\partial x^2]\theta(t-t')\theta(s-t')\,\mathrm{d}t' \\ &= \int_0^{\min(s,t)} \exp[(t+s-2t')\partial^2/\partial x^2]\,\mathrm{d}t' \\ &= -\tfrac{1}{2}(\partial^2/\partial x^2)^{-1}\{\exp(|t-s|\partial^2/\partial x^2) - \exp[(t+s)\partial^2/\partial x^2]\}\end{aligned} \tag{2.8}$$

At this point we may introduce the eigenfunction expansion of $-\partial^2/\partial x^2$, which is determined by

$$(-\partial^2/\partial x^2)\phi_n = \mu_n\phi_n, \tag{2.9}$$

where $\mu_n = n^2\pi^2/L^2$ and $\phi_n(x) = (2/L)^{1/2}\sin(n\pi x/L)$. It follows that

$$\Gamma(x, t, y, s) = \tfrac{1}{2}\sum_n (1/\mu_n)[\exp(-|t-s|\mu_n) - \exp(-(t+s)\mu_n)]\phi_n(x)\Phi_n(y). \tag{2.10}$$

The series obviously converges absolutely, so Γ is a continuous function. It also follows that the operator Γ is trace class.

Lemma 2.1. The covariance Γ is Hölder continuous with exponent $\frac{1}{2}$.

Proof. It is sufficient to check Hölder continuity in each variable separately. We begin with the t variable. It is clearly sufficient to prove that $\Sigma_n (1/\mu_n)\exp(-\tau\mu_n)\phi_n(x)\phi_n(y)$ is Hölder continuous in τ for $\tau \geq 0$. But

$$\begin{aligned}\sum_n (1/\mu_n)\{\exp[-(\tau+h)\mu_n] - \exp(-\tau\mu_n)\}\phi_n(x)\phi_n(y)\\ &= -\sum_n \exp(-\tau\mu_n)\int_0^h \exp(-\sigma\mu_n)\,\mathrm{d}\sigma\,\phi_n(x)\phi_n(y)\\ &= -\int_0^h \sum_n \exp[-(\tau+\sigma)\mu_n]\phi_n(x)\phi_n(y)\,\mathrm{d}\sigma.\end{aligned} \tag{2.11}$$

The magnitude of this is bounded by

$$\begin{aligned}\int_0^h\int_0^\infty \exp[-(\tau+\sigma)\zeta^2\pi^2/L^2]\,\mathrm{d}\zeta\,\mathrm{d}\sigma &= C\int_0^h (\tau+\sigma)^{-1/2}\,\mathrm{d}\sigma\\ &= 2C[(t+h)^{1/2} - \tau^{1/2}] \leq 2Ch^{1/2}.\end{aligned} \tag{2.12}$$

This is clearly Hölder continuous with exponent $\frac{1}{2}$, even at $\tau = 0$.

The remaining point is Hölder continuity in x. This follows from the Hölder continuity of the eigenfunction

$$|\phi_n(x+h) - \phi_n(x)| \leq C(nh)^\alpha \tag{2.13}$$

for $\alpha \leq 1$. We use this to estimate the difference in the covariance between $x+h$ and x. We obtain a series

$$\sum_n (1/n^2)n^\alpha h^\alpha \tag{2.14}$$

which is convergent for $\alpha < 1$. This shows that the exponent may be taken to be any $\alpha < 1$, in particular, $\alpha = \frac{1}{2}$.

The continuity of the covariance Γ implies the mean continuity of the process. In fact

$$E([w(x,t) - w(y,s)]^2) = [\Gamma(x,t,x,t) - \Gamma(x,t,y,s)] - [\Gamma(x,t,y,s) - \Gamma(y,s,y,s)]. \tag{2.15}$$

It is a more subtle fact that for a Gaussian process, Hölder continuity implies pointwise continuity of the process. In fact, more is true, Hölder continuity of the covariance with exponent α implies pointwise Hölder continuity of the process with exponent α' for every $\alpha' < \alpha/2$. These are standard facts and may be found for instance in Fernique (1975) or in Colella and Lanford (1973). Thus we have the following result for our Gaussian process.

Corollary 2.2. The random functions w are Hölder continuous with exponent $\alpha' < \frac{1}{4}$, with probability one.

Another thing to note is that the variance $\Gamma(x, t, x, t)$ of $w(x, t)$ vanishes when $x = 0, L$ and when $t = 0$. This means that $w(x, t)$ is equal to its mean value when $x = 0, L$ and when $t = 0$, with probability one. This mean value is zero, so this proves the following result.

Proposition 2.3. The random functions w satisfy the boundary conditions $w(0, t) = 0$, $w(L, t) = 0$, and $w(x, 0) = 0$, with probability one.

The asymptotic behaviour of the process w may be read off the covariance (2.8). As t and s approach infinity, the result is a covariance in which only the first term appears. If we then set $s = t$ we see that the asymptotic space covariance is $-\frac{1}{2}(\partial^2/\partial x^2)^{-1}$. Thus the stationary distribution is Gaussian with mean zero and this covariance. It is straightforward to see that this covariance is Hölder continuous with exponent $\alpha < 1$. It follows as before that the sample path of a random function of x in the stationary distribution is Hölder continuous with exponent $\alpha' < \frac{1}{2}$, with probability one.

The inverse of the covariance plays an important role in large fluctuation theory. In the present case this is

$$\Gamma^{-1} = G^{-1*}G^{-1} = (\partial/\partial t - \partial^2/\partial x^2)^*(\partial/\partial t - \partial^2/\partial x^2). \tag{2.16}$$

This may also be written $\Gamma^{-1} = (-\partial^2/\partial t^2 + \partial^4/\partial x^4)$, but this form does not make the boundary conditions explicit. However, from the factored form (2.16) we see that any f in the domain of Γ^{-1} must satisfy $f = 0$ at $t = 0$ and at $x = 0$, $x = L$ and the additional boundary condition

$$(\partial/\partial t - \partial^2/\partial x^2)f = 0 \qquad \text{at } t = T \text{ and at } x = 0,\ x = L.$$

This additional boundary condition is what is called a free boundary condition. The reason for this terminology is that with this condition the kernel $\Gamma(x, t, y, s)$ for $s \leqslant T$, $t \leqslant T$ is independent of the value of T.

The inverse of the covariance $\Gamma^{-1} = G^{-1*}G^{-1}$ determines a quadratic form defined on vectors f with values $\|G^{-1}f\|^2$. It is actually this quadratic form that is most important for us. Notice that we have

$$\|G^{-1}f\|^2 = \|(\partial/\partial t - \partial^2/\partial x^2)f\|^2 = \|\partial f/\partial t\|^2 + \|\partial^2 f/\partial x^2\|^2 + \|\partial f(T)/\partial x\|^2 \tag{2.17}$$

where the last inner product is with respect to the space variables alone.

3. Ventsel'–Freidlin estimates for the Gaussian process

In this section we prove large fluctuation estimates for the Gaussian process w satisfying (1.11). Such estimates are known in a general Banach space setting (Azencott 1980), but we believe that it is helpful to give an elementary proof. Our method is to follow the arguments of Ventsel' and Freidlin (1979), with necessary modifications.

There are two estimates to be proved. The first is a lower bound for the probability that εw lies in an open set. The second is an upper bound for the probability that εw lies in a closed set. In this section we prove such estimates for special types of open and closed sets. In the following section we show that these special cases imply the general case.

We define the action functional

$$I_0(f)=\tfrac{1}{2}\|G^{-1}f\|^2=\tfrac{1}{2}\|(\partial/\partial t-\partial^2/\partial x^2)f\|^2, \tag{3.1}$$

where the norm is the L^2 norm on functions of space and time. The convention is that $I_0(f)=+\infty$ for all f that do not satisfy the boundary conditions at $x=0$, $x=L$, and at $t=0$. The open set will be a ball in the uniform norm $\|\cdot\|_\infty$. Both norms occur in the following proposition.

Proposition 3.1. Fix $\delta>0$. For every $\zeta>0$ and for every sufficiently small $\varepsilon>0$

$$P(\|\varepsilon w-f\|_\infty<\delta)\geqslant\exp\{-(1/\varepsilon^2)[I_0(f)+\zeta]\}. \tag{3.2}$$

Proof. Let $z=\varepsilon w-f$. If $I(f)=\tfrac{1}{2}\|G^{-1}f\|^2<\infty$, then the measure induced on function space by z is absolutely continuous with respect to that induced by εw. The relative density is given by

$$\begin{aligned}\mathrm{d}\mu_z/\mathrm{d}\mu_{\varepsilon w}&=\exp\{-[1/(2\varepsilon^2)][2\langle G^{-1}\varepsilon w,G^{-1}f\rangle+\langle G^{-1}f,G^{-1}f\rangle]\}\\&=\exp[-(1/\varepsilon)\langle\alpha,G^{-1}f\rangle]\exp[-(1/\varepsilon^2)I_0(f)].\end{aligned} \tag{3.3}$$

Thus

$$\begin{aligned}P(\|\varepsilon w-f\|_\infty<\delta)&=P(\|z\|_\infty<\delta)\\&=E\{\exp[-(1/\varepsilon)\langle\alpha,G^{-1}f\rangle]1_{\|\varepsilon w\|_\infty<\delta}\}\exp[-(1/\varepsilon^2)I_0(f)].\end{aligned} \tag{3.4}$$

The expectation may be written as a conditional expectation and estimated by Jensen's inequality:

$$\begin{aligned}&E\{\exp[-(1/\varepsilon)\langle\alpha,G^{-1}f\rangle]\,|\,\|\varepsilon w\|_\infty<\delta\}P(\|\varepsilon w\|_\infty<\delta)\\&\quad\geqslant\exp[-(1/\varepsilon)E(\langle\alpha,G^{-1}f\rangle\,|\,\|\varepsilon w\|_\infty<\delta)]P(\|\varepsilon w\|_\infty<\delta).\end{aligned} \tag{3.5}$$

However, the conditional expectation in the exponential is zero by symmetry. Thus we have

$$P(\|\varepsilon w-f\|_\infty<\delta)\geqslant P(\|\varepsilon w\|_\infty<\delta)\exp[-(1/\varepsilon^2)I_0(f)]. \tag{3.6}$$

Furthermore $P(\|\varepsilon w\|_\infty<\delta)\to1$ as $\varepsilon\to0$, so this is the required lower bound.

Proposition 3.2. Let $I_0^s=\{f:I_0(f)\leqslant s\}$. Fix $\delta>0$. For every $\zeta>0$ and every sufficiently small $\varepsilon>0$

$$P[\mathrm{dist}(\varepsilon w,I_0^s)\geqslant\delta]\leqslant\exp[-(1/\varepsilon^2)[s-\zeta]]. \tag{3.7}$$

Proof. We introduce an approximation by using eigenfunctions f_n of the covariance with

$$\Gamma f_n=\gamma_n f_n. \tag{3.8}$$

Set $w_N=\Sigma_{n=1}^N\langle w,f_n\rangle f_n$ and $\tilde{w}_N=w-w_N$. The coefficients $\langle w,f_n\rangle$ are independent Gaussian random variables with variance γ_n.

The event whose probability we want to compute occurs in conjunction either with $\|\varepsilon\tilde{w}_N\|_\infty\geqslant\delta$ or $\|\varepsilon\tilde{w}_N\|_\infty<\delta$. Thus

$$P[\mathrm{dist}(\varepsilon w,I^s)\geqslant\delta]\leqslant P(\|\varepsilon\tilde{w}_N\|_\infty\geqslant\delta)+P(\varepsilon w_N\notin I^s). \tag{3.9}$$

We estimate each of these terms. The variance of $\tilde{w}_N$ is

$$\tilde{\Gamma}_N(x, t, x, t) = \sum_{n=N+1}^{\infty} \gamma_n |f_n(x, t)|^2. \tag{3.10}$$

The $\tilde{\Gamma}_N(x, t, x, t)$ are continuous and converge monotonely to zero, so they converge uniformly on the compact set $[0, L] \times [0, T]$.

According to a theorem of Fernique (1975, theoreme 1.3.3) for Gaussian processes,

$$E\{\exp[\|\tilde{w}_N\|_\infty^2/(2a^2)]\} < \infty \tag{3.11}$$

for every a with $a^2 > \sup_{x,t} \tilde{\Gamma}_N(x, t, x, t)$. Thus we may estimate

$$P(\|\tilde{w}_N\|_\infty \geq \delta/\varepsilon) \leq \exp[-\delta^2/(2a^2\varepsilon^2)]E[\exp(\|\tilde{w}_N\|_\infty^2/2a^2)]. \tag{3.12}$$

Pick a so small that $\delta^2/(2a^2) > s$. Pick N so large that the expectation in (3.11) is finite. This gives an estimate

$$P(\|\varepsilon\tilde{w}_N\|_\infty \geq \delta) \leq C\exp(-s/\varepsilon^2). \tag{3.13}$$

Thus the first term gives no trouble.

The other term is the one that carries the essential information. We have

$$I_0(w_N) = \tfrac{1}{2}\|G^{-1}w_N\|^2 = \tfrac{1}{2}\langle w_N, \Gamma^{-1}w_N\rangle = \tfrac{1}{2}\sum_{n=1}^{N} \gamma_n^{-1} \langle w_n, f_n\rangle^2 = \chi_N^2/2$$

where the χ_N^2 random variable is the sum of squares of N normalised Gaussian random variables. Clearly

$$E[\exp(t\chi_N^2/2)] < \infty \tag{3.14}$$

for every $t < 1$. Hence

$$P(\varepsilon w_N \notin I_0^s) = P[I_0(\varepsilon w_N) > s] = P[I_0(w_N) > s/\varepsilon^2] = P(\chi_N^2 > 2s/\varepsilon^2)$$
$$\leq \exp(-ts/\varepsilon^2)E[\exp(t\chi_N^2/2)]. \tag{3.15}$$

We may take $t = 1 - \zeta/(2s)$. The multiplicative constants are independent of ε and may be taken care of by the other $\zeta/2$. This proves the proposition.

4. Properties of the action functional

In this section we prove some lower semicontinuity and compactness properties of the action functional I_0 for the Gaussian process w. These are used to give a somewhat stronger version of the Ventsel'–Freidlin estimates for this process. The Banach space of sample functions will be taken to be $C_{D0}([0, L] \times [0, T])$, the space of continuous functions on $[0, L] \times [0, T]$ satisfying the Dirichlet boundary conditions (1.7) and the zero initial condition at $t = 0$.

Proposition 4.1. The action functional $I_0(f) = \frac{1}{2}\|(\partial/\partial t - \partial^2/\partial x^2)f\|^2$ is lower semicontinuous on $C_{D0}([0, L] \times [0, T])$.

Proof. Let $f_n \to f$ in C_{D0}. We wish to show that $I_0(f) \leq \liminf_n I_0(f_n)$. We may as well assume that the right-hand side is finite, and by passing to a subsequence we may even assume that the $I_0(f) = \lim_n I_0(f_n)$.

For each f_n there is a corresponding h_n in L^2 with $f_n = Gh_n$, and the $I_0(f_n) = \frac{1}{2}\|h_n\|^2$ are bounded by a constant. Thus there is a subsequence h_m of the h_n that converges weakly to some h. It follows that $f_m = Gh_m$ converges weakly to Gh. This shows that $f = Gh$.

Since $2\langle h_m, h\rangle \leq \|h_m\|^2 + \|h\|^2$, we have $\|h\|^2 \leq \lim_m \|h_m\|^2$. This may be restated as $I_0(f) \leq \lim_n I_0(f_n)$, and this is the desired conclusion.

Proposition 4.2. For every $c < \infty$, the set $I_0^c = \{f: I_0(f) \leq c\}$ is compact in $C_{D0}([0, L] \times [0, T])$.

Proof. We write $f = Gh$ and $I_0(f) = \frac{1}{2}\|h\|^2$. We must show that when the $\|h\|^2$ are bounded, the corresponding f run over a bounded equicontinuous set. Then proposition 4.1 combined with Ascoli's theorem will show that I_0^c is compact.

First note that

$$f(x, t) = \int\int G(x, t, y, s)h(y, s)\,\mathrm{d}y\,\mathrm{d}s, \tag{4.1}$$

and so

$$|f(x, t)| \leq \|G(x, t)\|^2\|h\|^2 = \Gamma(x, t, x, t)\|h\|^2. \tag{4.2}$$

The first L^2 norm in (4.2) is with respect to the y,s variables. This inequality establishes the boundedness.

The equicontinuity is a similar argument. We have by essentially the same calculation

$$\begin{aligned} &|f(x+h, t+k) - f(x, t)| \\ &\quad \leq \{[\Gamma(x+h, t+k, x+h, t+k) - \Gamma(x+h, t+k, x, t)] \\ &\qquad - [\Gamma(x+h, t+k, x, t) - \Gamma(x, t, x, t)]\}\|h\|^2. \end{aligned} \tag{4.3}$$

But we know that the covariance Γ is uniformly Hölder continuous, by lemma 2.1, so the equicontinuity follows.

We conclude this section with a restatement of the main results for the Gaussian process in a somewhat more elegant form. Define the set function I_0 in terms of the corresponding point function by

$$I_0(A) = \inf_{f \in A} I_0(f). \tag{4.4}$$

We consider a set $A \subset C_{D0}([0, L] \times [0, T])$.

Theorem 4.3. (i) If A is open, then

$$-I_0(A) \leq \liminf_{\varepsilon \to 0} \varepsilon^2 P(\varepsilon w \in A). \tag{4.5}$$

(ii) If A is closed, then

$$\limsup_{\varepsilon \to 0} \varepsilon^2 P(\varepsilon w \in A) \leq -I_0(A). \tag{4.6}$$

Proof. (i) Let f be an arbitrary element of the open set A. Then there exists a $\delta > 0$ such that $\{g \in C_{D0}: \|g-f\|_\infty < \delta\} \subset A$. But $w \in C_{D0}$ with probability one, by proposition 2.3. Thus

$$P(\|w-f\|_\infty < \delta) \leqslant P(w \in A). \tag{4.7}$$

It follows from proposition 3.1 that

$$\liminf_{\varepsilon \to 0} \varepsilon^2 P(\varepsilon w \in A) \leqslant -I_0(f). \tag{4.8}$$

Since f is arbitrary, this proves the first part of the theorem.

(ii) Let $\eta > 0$. Define $s = I_0(A) - \eta$. The set I_0^s is compact, by proposition 4.2. The closed set A does not intersect the compact set I_0^s, so $\delta = \operatorname{dist}(A, I_0^s) > 0$. It follows that

$$\begin{aligned} P(\varepsilon w \in A) \leqslant P[\operatorname{dist}(\varepsilon w, I_0^s) \geqslant \delta] &\leqslant \exp[-(1/\varepsilon^2)[s-\eta]] \\ &\leqslant \exp\{-(1/\varepsilon^2)[I_0(A) - 2\eta]\}. \end{aligned} \tag{4.9}$$

Since η is arbitrary, this proves the second part.

5. Solutions of the nonlinear equation

In this section we treat the nonlinear equation in its integrated form (1.13). We temporarily forget the probabilistic aspect of the problem and treat the solution of the nonlinear problem as a perturbation of the solution of the linear problem. The main result will be that the solution u of the nonlinear equation (1.13) is a continuous function of the solution w of the linear problem and of the initial condition u_0.

Define

$$z = \varepsilon w + g u_0 \qquad \text{and} \qquad q = u - z. \tag{5.1}$$

The equation for the function u may also be stated in terms of q:

$$u = -GV'(u) + z \qquad \text{or} \qquad q = -GV'(q+z). \tag{5.2}$$

It is obvious that once we have information about q we will also have corresponding information about u.

We shall assume in the following that w and u_0 are given continuous functions. We would like to show that q and hence u are also continuous. We can do this locally in time without any special problem. We work in the Banach space $C_D([0, L] \times [0, T])$ of functions on space–time satisfying the Dirichlet boundary conditions at $x = 0$ and $x = L$. The initial condition u_0 will be taken in the space $C_D([0, L])$ of functions of space satisfying the Dirichlet boundary conditions.

Proposition 5.1. There is a time $T > 0$ depending on $\|u_0\|_\infty$ and $\|w\|_\infty$ such that the equation (5.2) has a unique solution in $C_D([0, L] \times [0, T])$.

Proof. Let B be the ball of radius a in this Banach space. Define a function F on B by

$$F(q) = -GV'(q+z). \tag{5.3}$$

Since $g: C_D([0, L]) \to C_D([0, L] \times [0, T])$ has norm one, $q + z = q + \varepsilon w + g u_0$ is always in the ball of radius $a+\varepsilon\|w\|_\infty+\|u_0\|_\infty$. Furthermore the polynomial $V'(u)=\lambda u^3-\mu u$ is bounded on bounded sets. Finally, the operator G: $C_{\mathrm{D}}([0,L]\times[0,T])\to C_{\mathrm{D}}([0,T])$ has norm bounded by T. It follows that for T sufficiently small, F maps B into B. Furthermore the polynomial V' is Lipshitz on bounded sets, so F is actually a contraction of B for sufficiently small T. It follows that F has a unique fixed point, and this is the required solution.

Corollary 5.2. Let $[0, T^*)$ be the maximum interval of time for which the equation (5.2) has a continuous solution. Then either $T^*=\infty$ or $\|u(t)\|_\infty\to\infty$ as $t\to T^*$.

Proof. If $T^*<\infty$ and $\|u(t)\|_\infty$ remained bounded, then we could use the argument of proposition 5.1 to continue the solution beyond T^* (Segal 1963).

We now begin a series of results that will lead to the continuous dependence of u on w and u_0. The crucial step is an *a priori* estimate (proposition 5.4) that shows that q is L^4 bounded. Marcus (1978) has given such an estimate for the case of a homogeneous polynomial. The following is an extension of his estimates.

Lemma 5.3. If q is a solution of (5.2), then

$$\langle V'(u), q\rangle\leq 0, \tag{5.4}$$

where the inner product is that of $L^2([0,L]\times[0,T])$.

Proof. The function q formally satisfies the differential equation obtained by subtracting (1.11) from (1.5). However, since w is not smooth it is not immediately clear that the equation is satisfied in a rigorous sense. Thus we use an approximation to get a rigorous equation and then show that the bound is maintained as the approximating q_n approach the exact q.

Let P_n be the projection onto the span of the first n eigenfunctions ϕ_j of $\partial^2/\partial x^2$. Set $q_n(t)=P_nq(t)$. Then $q_n(t)$ is automatically in the domain of $\partial^2/\partial x^2$ for each t. Furthermore, it satisfies the equation

$$(\partial/\partial t-\partial^2/\partial x^2)q_n(t)=-P_nV'[u(t)]. \tag{5.5}$$

Take the $L^2([0,L])$ inner product with $q_n(t)$. This gives

$$\tfrac{1}{2}\partial\|q_n(t)\|^2/\partial t+\|\partial q_n(t)/\partial x\|^2=-\langle V'(u(t)), q_n(t)\rangle. \tag{5.6}$$

The term $\|\partial q_n(t)/\partial x\|^2$ is positive and finite. If we omit it the equality becomes an inequality. Integrate this inequality from 0 to T. We obtain

$$\tfrac{1}{2}\|q_n(T)\|^2\leq-\int_0^T\langle V'(y(t)), q_n(t)\rangle\,\mathrm{d}t. \tag{5.7}$$

Let $n\to\infty$. This gives

$$\tfrac{1}{2}\|q(T)\|^2\leq-\int_0^T\langle V'(u(t)), q(t)\rangle\,\mathrm{d}t=-\langle V'(u), q\rangle. \tag{5.8}$$

This is the desired conclusion.

Lemma 5.4. Let q be a solution of (5.2). Then either

$$\|q\|_4\leq 10\|z\|_4 \qquad \text{or} \qquad \|q\|_4\leq(LT)^{1/4}(10\mu/\lambda)^{1/2}.$$

Proof. Assume that the conclusion is false, that is, that $\|z\|_4 < \frac{1}{10}\|q\|_4$ and $\mu(LT)^{1/2} < \frac{1}{10}\lambda\|q\|_4^2$. Then

$$\begin{aligned}\langle V'(q+z), q\rangle &\geqslant \lambda(\|q\|_4^4 - 3\|z\|_4\|q\|_4^3 - 3\|z\|_4^3\|q\|_4) - \mu(\|q\|_2^2 + \|z\|_2\|q\|_2)\\ &\geqslant \lambda\tfrac{6}{10}\|q\|_4^4 - \mu(LT)^{1/2}2\|q\|_4^2 \geqslant \lambda\tfrac{4}{10}\|q\|_4^4 > 0.\end{aligned} \tag{5.9}$$

This contradicts lemma 5.3.

The next task is to improve this bound to a uniform bound (lemma 5.8). We shall need some properties of the kernel $g(x, y, t)$ defined in (1.9), which gives the solution of the heat equation on the interval $[0, L]$. Let $p(x, t)$ be the fundamental solution of the heat equation on the whole real line.

Lemma 5.5. The inequality $0 \leqslant g(x, y, t) \leqslant p(x-y, t)$ is satisfied for all x and y in $[0, L]$ and all $t > 0$.

Proof. Let $f \geqslant 0$ and

$$gf(x, t) = \int_0^L g(x, y, t)\, f(y)\, dy \tag{5.10}$$

and

$$pf(x, t) = \int_{-\infty}^{\infty} p(x-y, t) f(y) dy. \tag{5.11}$$

Clearly $gf(x, t) - pf(x, t) \leqslant 0$ when $x = 0$ or $x = L$, since then the first term vanishes. Furthermore, $gf(x, t) = 0$ for $t = 0$. But this difference satisfies the heat equation in the interior of the strip $0 < x < L$, $0 < t$, so by the parabolic maximum principle (Protter and Weinberger 1967)

$$gf(x, t) - pf(x, t) \leqslant 0 \tag{5.12}$$

throughout the strip. Now let f approach a point measure. This gives the inequality for the kernels.

Lemma 5.6. Let $p(x, t) = (4\pi t)^{-1/2}\exp[-x^2/(4t)]$ be the fundamental solution of the heat equation. Then p is in $L^r(R \times [0, T])$ for every $r < 3$.

Proof. The L^r norm is a multiple of $\int_0^T t^{1/2 - r/2}\, dt$. This is finite when $r < 3$.

Proposition 5.7. Let G be the integral operator defined in (1.10). Let $1 \leqslant r < 3$, $1 \leqslant q \leqslant \infty$, and $1 \leqslant p \leqslant \infty$ with $1/p = 1/r + 1/q - 1$. Then G is bounded from $L^q([0, L] \times [0, T])$ to $L^p([0, L] \times [0, T])$.

Proof. Let $u = Gf$. Then

$$\begin{aligned}|u(x, t)| &\leqslant \int_{t-T}^{t}\int_0^L g(x, y, t-s)|f(y, s)|\, dy\, ds \\ &\leqslant \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_{[0,T]}(t-s) p(x-y, t-s)|f(y, s)|\, dy\, ds\end{aligned} \tag{5.13}$$

where $1_{[0,T]}$ is the indicator function of $[0, T]$. This follows from lemma 5.5 and the fact that the kernels vanish when the time argument is negative. The above inequality may be rewritten

$$|u| \leqslant (1_{[0,T]} \cdot p) * |f| \tag{5.14}$$

where the $*$ denotes convolution. The result then follows from Young's inequality. (This argument follows one in Stroock and Varadhan (1979).)

Lemma 5.8. Let $z = \varepsilon w + g u_0$ lie in a bounded set in $C_D([0, L] \times [0, T])$. Then the corresponding u satisfying (1.13) are bounded in $C_D([0, L] \times [0, T])$.

Proof. Since z is bounded in $C_D([0, L] \times [0, T])$, it is also bounded in L^4. By lemma 5.4, $q = u - z$ is bounded in L^4. But then it follows that u is bounded in L^4. Thus $V'(u)$ is bounded in $L^{4/3}$.

Now apply proposition 5.7 with $q = \frac{4}{3}$, $r = \frac{12}{5}$ and $p = \frac{1}{6}$. This shows that u is bounded in L^6. It follows that $V'(u)$ is bounded in L^2.

Now apply proposition 5.7 again with $q = 2$, $r = 2$, and $p = \infty$. It follows that u is bounded in L^∞.

Theorem 5.9. The solution of (1.13) exists globally in time.

Proof. This follows from lemma 5.8 and corollary 5.2.

Theorem 5.10. The solution u of (3.3) depends continuously (in the uniform norm) on $z = \varepsilon w + u_0$.

Proof. Consider a pair z_1, z_2 and the corresponding pair u_1, u_2. Let $\bar{u} = u_1 - u_2$ and $\bar{z} = z_1 - z_2$. Then the differences satisfy

$$\bar{u} = -G[V'(u_1) - V'(u_2)] + \bar{z}. \tag{5.15}$$

On the other hand, we have

$$V'(u_1) - V'(u_2) = \bar{u}h,$$

where h is a polynomial in u_1 and u_2. The point is that by lemma 5.8 we have an *a priori* uniform bound K on h. Thus for each t we may estimate

$$\|\bar{u}(t)\|_\infty \leqslant K \int_0^t \|\bar{u}(s)\|_\infty \, \mathrm{d}s + \|\bar{z}\|_\infty. \tag{5.16}$$

The first two norms in this formula are uniform on $[0, L]$. It follows from Gronwall's inequality that we have an estimate

$$\|\bar{u}\|_\infty \leqslant \exp(KT) \|\bar{z}\|_\infty. \tag{5.17}$$

Here the norms are uniform on $[0, L] \times [0, T]$. This proves the theorem.

Remark. If z is in the space $C_D([0, L] \times [0, T])$ of continuous functions satisfying the Dirichlet boundary conditions at 0 and L, then so is u. Thus we will usually think of $u = \Psi(z)$ where Ψ is a continuous map of this space into itself. Note that the inverse map $\Psi^{-1}(u) = z$ is also well defined and continuous.

If we fix u_0 we may also think of u as a continuous function of εw in $C_{D0}([0, L]\times[0, T])$, where the 0 reminds us of the zero initial conditions. In that case $u = \Phi_{u_0}(\varepsilon w) = \Psi(\varepsilon w + gu_0)$. In that case the inverse map is given by $\varepsilon w = \Phi_{u_0}^{-1}(u) = \Psi^{-1}(u) - gu_0$ for u in $C_{Du_0}([0, L]\times[0, T])$.

6. Ventsel'–Freidlin estimates for the nonlinear process

We may now define the process of interest by taking w to be the Gaussian random function considered in § 2. We know that w is in the space $C_{D0}([0, L]\times[0, T])$ of functions satisfying Dirichlet boundary conditions in space and zero initial condition in time, with probability one. Consider an initial function u_0 in $C_D([0, L])$. Set $z = \varepsilon w + gu_0$. This random function is in $C_D([0, L]\times[0, T])$ and assumes the initial condition u_0 with probability one. The process of interest is $u = \Psi(z)$, the solution of the nonlinear equation (5.2). This u is in $C_{Du_0}([0, L]\times[0, T])$, since it assumes the initial condition u_0 with probability one. We may also regard u as a function

$$u = \Phi_{u_0}(\varepsilon w) = \psi(\varepsilon w + gu_0) \tag{6.1}$$

of εw.

In order to study this process we shall need certain functionals of the sample paths. We define a functional I by

$$I(u) = I_0[\Phi_{u_0}^{-1}(u)]. \tag{6.2}$$

In other words, $I(u) = I_0(\varepsilon w)$, where u and εw are related by the integral equation (5.1), (5.2). The functional I is thus (at least formally) given by

$$\begin{aligned} I(u) &= \tfrac{1}{2}\|(\partial/\partial t - \partial^2/\partial x^2)u + V'(u)\|_2^2, && u(0) = u_0, \\ &= \infty, && u(0) \neq u_o. \end{aligned} \tag{6.3}$$

We also define

$$I(A) = \inf_{u\in A} I(u). \tag{6.4}$$

Theorem 6.1. Let u be the random solution of (1.13). Then

(i) for every open set A in $C_{Du_0}([0, L]\times[0, T])$

$$-I(A) \leq \liminf_{\varepsilon\to 0} \varepsilon^2 \log P^\varepsilon(u \in A); \tag{6.5}$$

(ii) for every closed set A in $C_{Du_0}([0, L]\times[0, T])$

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log P^\varepsilon(u \in A) \leq -I(A). \tag{6.6}$$

Proof. The map Φ_{u_0} maps $C_{D0}([0, L]\times[0, T])$ onto $C_{Du_0}([0, L]\times[0, T])$. Furthermore $u = \Phi_{u_0}(\varepsilon w)$. Hence

$$P^\varepsilon(u \in A) = P[\varepsilon w \in \Phi_{u_0}^{-1}(A)]. \tag{6.7}$$

But if A is open or closed, then so is $\Phi_{u_0}^{-1}(A)$. Hence by theorem 4.3 we may compute the probabilities in terms of

$$I_0[\Phi_{u_0}^{-1}(A)] = I(A) \tag{6.8}$$

This completes the proof.

The preceding analysis reduces the problem to finding the infimum of a certain functional I. We now turn to some general properties of this functional. We show that a minimum is actually attained on non-empty closed sets.

Proposition 6.2. The functional I is lower semi-continuous on $C_{\mathrm{D}u_0}([0, L]\times[0, T])$.

Proof. This is because $I(f)=I_0[\Phi_{u_0}^{-1}(f)]$. Since $\Phi_{u_0}^{-1}$ is continuous and I_0 is lower semi-continuous (by proposition 4.1), it follows that I is lower semi-continuous.

Proposition 6.3. The set $I^s=\{f: I(f)\leqslant s\}$ is compact in $C_{\mathrm{D}u_0}([0, L]\times[0, T])$.

Proof. The set I^s is the image under Φ_{u_0} of the set I_0^s. But I_0^s is compact, by proposition 4.2. Furthermore Φ_{u_0} is continuous, by theorem 5.10. The image of a compact set under a continuous map is compact.

Theorem 6.4. If A is a closed set in $C_{\mathrm{D}u_0}([0, L]\times[0, T])$ such that $I(A)<\infty$, then there is an f in A with $I(f)=I(A)$.

Proof. Let $s=I(A)+1$. Then $A\cap I^s$ is a non-empty compact set, by proposition 6.3. Since I is lower semi-continuous by proposition 6.2, it follows that its minimum is attained on this set.

Corollary 6.5. Let f_c be the solution of (1.13) with $\varepsilon=0$. Let A be a closed set in $C_{\mathrm{D}u_0}([0, L]\times[0, L])$ that does not contain f_c. Then $I(A)>0$.

Proof. If $I(A)=0$, then by theorem 6.4 there is an f in A with $I(f)=0$. But then $I_0(\Phi_{u_0}^{-1}f)=0$, so $\Phi_{u_0}^{-1}f=0$. This shows that f is the unique solution of (1.13), so $f=f_c$. Thus f_c is in A, which is a contradiction.

Corollary 6.6. Under the same hypotheses, there is a $k>0$ with

$$P(u\in A)\leqslant\exp(-k/\varepsilon^2). \tag{6.9}$$

Thus the probability of a deviation from the deterministic path is exponentially small in the small ε limit.

Proof. Take $0<k<I(A)$ and apply theorem 6.1. This says that $\varepsilon^2\log P^\varepsilon(u\in A)\leqslant -k$ for ε sufficiently small, and this is equivalent to (6.9).

In the following we shall need to use the representation (6.3) of the action functional. Define u to be regular if $\partial u/\partial t$ and $\partial^2 u/\partial x^2$ are in $C_{\mathrm{D}}([0, L]\times[0, T])$. It is clear that when u is regular the representation makes sense. We now wish to find conditions that ensure that u is regular and that arbitrary u may be approximated by regular u.

Proposition 6.7. Assume that w is regular and that $\partial^2 u_0/\partial x^2$ is in $C_{\mathrm{D}}([0, L])$. Then the corresponding solution u of (5.2) is also regular.

Proof. Since w and gu_0 are regular, so is $z=\varepsilon w+gu_0$. Since $q=u-z$, $q=0$ at $t=0$. The equation for q involves $V'(q+z)$ and this nonlinear operator satisfies the hypotheses of lemma 3.1 of Segal (1963). Thus q is regular. Since $u=q+z$, we conclude that u is regular.

Lemma 6.8. Assume that $(\partial/\partial t-\partial^2/\partial x^2)\ w=h$ and that $\partial^2 h/\partial x^2$ is in $C_D([0,L]\times[0,T])$. Then w is regular.

Proof. Since $w=Gh$, where G is bounded and commutes with $\partial^2/\partial x^2$, it follows that $\partial^2 w/\partial x^2$ is in $C_D([0,L]\times[0,T])$. But since w satisfies the parabolic equation, it follows that w is regular.

Theorem 6.9. Assume that $I(u)<\infty$ and that the initial condition u_0 is smooth. Then there is a sequence u_n of regular functions such that $u_n\to u$ uniformly and $I(u_n)\to I(u)$.

Proof. We have $I(u)=I_0(\varepsilon w)=\frac{1}{2}\|\varepsilon h\|^2$ for h in L^2. Let h_n be a sequence of smooth functions that approach h in L^2. By lemmas 6.8 and 6.7 the corresponding w_n and u_n are regular. It is also clear that $I(u_n)\to I(u)$. By proposition 5.7 $w_n\to w$ uniformly. By theorem 5.10 $u_n\to u$ uniformly, and this completes the proof.

7. The equilibrium action

In this section we study the equilibrium action defined by

$$S(u)=\int_0^L \tfrac{1}{2}(\partial u/\partial x)^2+V(u)\,\mathrm{d}x, \tag{7.1}$$

with Dirichlet boundary conditions on u at $x=0$ and at $x=L$. For fixed u, the differential of S at u may be represented by a vector in $L^2([0,L])$, namely

$$S'(u)=-\partial^2 u/\partial x^2+V'(u). \tag{7.2}$$

This is defined only when u satisfies the Dirichlet boundary conditions. For fixed u the second differential of S at u is a quadratic form, represented in $L^2([0,L])$ by the operator

$$S''(u)=-\partial^2/\partial x^2+V''(u) \tag{7.3}$$

with Dirichlet boundary conditions. In this section we study the critical points of S, that is, the points u where $S'(u)=0$, and relate them to the topology of S.

The equation $S'(u)=0$ is Newton's equation of motion for a particle with potential energy $-V(u)$. With this interpretation x is the time variable, and the Dirichlet boundary conditions say that the particle returns to the origin at exactly time L. In this problem the force is $V'(u)$. We may think of this as a spring with spring constant $-V'(u)/u=-2\partial V(u)/\partial u^2$. The particular nonlinear force that we have chosen is such that the spring constant is

$$-2\partial V(u)/\partial u^2=-V'(u)/u=(\mu-\lambda u^2). \tag{7.4}$$

From this expression we see that the spring constant gets smaller as the spring is extended. This is thus what is called a soft spring or sublinear problem. It is this feature that is the key to the analysis.

The theory of such sublinear problems is by now rather standard. There is a nice treatment in an article by Hempel (1971). He works with the hypothesis that $-2V(u)$

is a concave function of u^2. This is the same as requiring that the spring constant $-2\partial V(u)/\partial u^2 = -V'(u)/u$ is a decreasing function of u^2, and this is certainly true in our case. There is also a discussion in an article by Chafee and Infante (1974). They work with a slightly different hypothesis, namely, that $[-2V(u)]^{1/2}$ is concave in u. In the following we will find it more convenient to use Hempel's framework. In both treatments oscillation properties of the solutions play an important role.

It should be mentioned that Hempel (1971) has also treated the corresponding problem in higher dimensions. For such problems oscillation properties are of little use, and so he instead uses topology in function space. In this way he is able to obtain existence but not uniqueness. There is also an article by Coffman (1975) which treats the relation between the topology and oscillation properties in the one-dimensional case. This article also discusses variational principles.

In the following we present the main results. Since the proofs are scattered throughout the literature, we also give a brief outline of the proofs. In the stability result we use a technique of Laetsch (1975).

Lemma 7.1. If u is a non-zero critical point of S, then $S(u)<0$.

Proof. Define the quadratic functional

$$Q_u(w) = \int_0^L \tfrac{1}{2}(\partial w/\partial x)^2 + (\partial V(u)/\partial u^2)w^2 \, dx. \tag{7.5}$$

Then

$$Q'_u(w) = -\partial^2 w/\partial x^2 + 2(\partial V(u)/\partial u^2)w. \tag{7.6}$$

Notice that

$$Q_u(w) = \tfrac{1}{2}\langle Q'_u(w), w\rangle. \tag{7.7}$$

Furthermore

$$S'(u) = Q'_u(u) \tag{7.8}$$

and

$$S(u) = Q_u(u) - (\lambda/4)\|u^2\|^2. \tag{7.9}$$

Hence if $S'(u)=0$ then $Q'_u(u)=0$, $Q_u(u)=0$, and $S(u) = -(\lambda/4)\|u^2\|^2 \leqslant 0$, with equality only when $u=0$.

Lemma 7.2. If $\mu^{1/2}L \leqslant \pi$, then the only critical point of S is zero.

Proof. Note that

$$S(w) = Q_0(w) + (\lambda/4)\|w^2\|^2. \tag{7.10}$$

But

$$Q'_0(w) = -\partial^2 w/\partial x^2 - \mu w \tag{7.11}$$

is an expression that we can analyse explicity. The eigenvalues of $-\partial^2/\partial x^2$ are $\mu_n = n^2\pi^2/L^2$. Hence when $\mu \leqslant \mu_1$, the quadratic form $Q_0(w)$ is positive. By (7.11) and (7.10) the functional $S(w)$ is strictly positive except when $w=0$. By lemma 7.1 there are no other critical points.

Lemma 7.3. If $\mu^{1/2}L > \pi$, then there is a non-zero critical point θ_1 which is a positive function.

Proof. When $\mu > \mu_1$ the quadratic form $Q_0(w)$ assumes negative values, in fact $Q_0(w) = (\mu_1 - \mu)\|w\|^2 < 0$ when w is an eigenfunction corresponding to the eigenvalue μ_1. It follows that

$$S(\varepsilon w) = \varepsilon^2[(\mu_1 - \mu)\|w\|^2 + (\lambda/4)\varepsilon^4\|w^2\|^2] < 0 \tag{7.12}$$

for ε sufficiently small.

However, S is weakly lower semi-continuous on an appropriate Sobolev space (Coffman 1975), and hence its minimum is attained on some u. By taking absolute values we see that the minimum is attained on a $u_1 \geqslant 0$. By (7.12) $S(u_1) < 0$ and so u_1 cannot be the zero solution.

Lemma 7.4. (Hempel 1971) Let u be a solution of the differential equation $\partial^2 u/\partial x^2 = V'(u)$ with the initial condition $u(0) = a$. Let w be another solution with $w(0) = a$. Assume that $\partial w/\partial x(0) \geqslant \partial u/\partial x(0)$. Then $w > u$ throughout any interval in which $u \geqslant 0$.

Proof. We compute

$$(w'u - wu') = V'(w)u - wV'(u). \tag{7.13}$$

Integrate. This gives

$$w'(x)u(x) - w(x)u'(x) \geqslant \int_0^x uw[V'(w)/w - V'(u)/u]\,\mathrm{d}x. \tag{7.14}$$

Assume that there was an x in the interval with $w(x) = u(x)$. Consider the first such x. Up to this x, $0 \leqslant u < w$ and so by convexity $V'(u)/u \leqslant V'(w)/w$. Hence the integrand in (7.14) is positive and so $w'(x) \geqslant u'(x)$. Since w and u are distinct solutions we even have $w'(x) > u'(x)$. But this would imply that $w < u$ just before arriving at x, which is a contradiction.

Lemma 7.5. If $\mu^{1/2}L > \pi$, then the non-zero positive critical point is unique.

Proof. Assume that there were two such critical points u and w. Without loss of generality we may assume that $w'(0) > u'(0) > 0$. By lemma 7.4 $w(L) > u(L) = 0$, so w could not possibly satisfy the boundary condition at L.

Theorem 7.6. If $N\pi < \mu^{1/2}L \leqslant (N+1)\pi$ then S has precisely $2N+1$ critical points $\pm u_1, \pm u_2, \ldots, \pm u_N$, 0. The function u_n has n half-periods.

Proof. In order to prove existence, we apply lemma 7.3 to an interval of length L/n with $1 \leqslant n \leqslant N$. Since $\pi < \mu^{1/2}L/n$ there is a non-zero positive solution. Piece this solution together with alternating plus and minus signs. The gives the non-zero solution u_n in the interval of length L.

For uniqueness, note that any non-zero critical point is a function that divides the interval into n half-periods of length L/n, for some n. If $n \geqslant N+1$, the $\mu^{1/2}L/n \leqslant \pi$, so there is no solution, by lemma 7.2. If $1 \leqslant n \leqslant N$ then the solution is of the form $\pm u$ where u is positive. By lemma 7.5 these are the only possibilities.

Theorem 7.7. Assume that $\pm u_1, \pm u_2, \ldots, \pm u_N$ and 0 are critical points such that u_n has n half-periods. Then $S(u_1) < S(u_2) < \ldots < S(u_n) < 0$.

Proof. (Coffman 1975) Since by convexity

$$V(w) - V(u) \geqslant \partial V(u^2)/\partial u^2\, (w^2 - u^2), \tag{7.15}$$

it follows that

$$S(w) - S(u) \geqslant Q_u(w) - Q_u(u). \tag{7.16}$$

Let u be a critical point of $S(u)$ with n half-periods. Then $S'(u) = 0$ implies that $Q'_u(u) = 0$ (by 7.8) and so $Q_u(u) = 0$ (by 7.7). By the oscillation properties of one-dimensional linear differential operators $Q_u(w) > 0$ for any w with $n+1$ half-periods. But then $S(w) > S(u)$, by (7.16). This proves the result except for the last inequality. But this follows from lemma 7.1.

Theorem 7.8. (Chafee and Infante 1974, Laetsch 1975) Let $N\pi < \mu^{1/2}L \leqslant (N+1)\pi$. Let $\pm u_1, \pm u_2, \ldots, \pm u_N$ and $u_{N+1} = 0$ be the $2N+1$ critical points of S. Then the second differential $S''(u_n)$, $1 \leqslant n \leqslant N+1$ has precisely $n-1$ strictly negative eigenvalues.

Proof. Let $u = u_n$ be the critical point with $n-1$ interior zeros. Then 0 is the nth eigenvalue and u_n is the nth eigenfunction of the operator

$$Q'_u = -\mathrm{d}^2/\mathrm{d}x^2 + V'(u)/u \tag{7.17}$$

with Dirichlet boundary conditions. Furthermore let z be the derivative $z = \mathrm{d}u/\mathrm{d}x$. Since u vanishes at the end points, so does $\mathrm{d}^2u/\mathrm{d}x^2 = \mathrm{d}z/\mathrm{d}x$. Since z has n interior zeros, 0 is the $(n+1)$th eigenvalue and z is the $(n+1)$th eigenfunction of

$$L_{\mathrm{N}} = -\mathrm{d}^2/\mathrm{d}x^2 + V''(u) \tag{7.18}$$

with Neumann boundary conditions. We may also restrict z to a smaller interval whose end points are the first and last zeros. On this smaller interval z has $n-2$ interior zeros. Thus on this smaller interval 0 is the $(n-1)$th eigenvalue and z the corresponding eigenfunction of

$$L_{\mathrm{D}} = -\mathrm{d}^2/\mathrm{d}x^2 + V''(u) \tag{7.19}$$

with Dirichlet boundary conditions.

The operator of interest for us is

$$S''(u) = -\partial^2/\partial x^2 + V''(u) \tag{7.20}$$

with Dirichlet boundary conditions on the entire interval. We have the quadratic form inequality

$$L_{\mathrm{N}} \leqslant S''(u) \leqslant L_{\mathrm{D}}. \tag{7.21}$$

By a well known variational principle (Faris 1975) this implies that the $(n+1)$th eigenvalue of $S''(u)$ is strictly positive and the $(n-1)$th eigenvalue of $S''(u)$ is strictly negative. It remains to locate the nth eigenvalue.

We use the fact that $V'(u)/u$ is an increasing function of u^2 to conclude that $V'(u)/u \leqslant V''(u)$. Thus

$$Q'_u \leqslant S''(u). \tag{7.22}$$

It follows that the nth eigenvalue of $S''(u)$ is strictly positive.

The only remaining case is when $u = 0$. In that case

$$S''(0) = -\partial^2/\partial x^2 - \mu \tag{7.23}$$

has precisely N strictly negative eigenvalues, namely the $n^2\pi^2/L^2 - \mu$ with $1 \leqslant n \leqslant N$.

8. The gradient flow

In this section we wish to consider the flow generated by the gradient S' of the equilibrium action S given by (1.3). The critical points of S are the stationary points of the flow. We wish to show that every initial condition leads to a critical point, and to describe the initial conditions that lead to the absolute minima.

The problem of showing that every initial condition leads to a critical point ultimately reduces to a compactness argument. This may be carried out in the framework of the Palais–Smale theory (Berger 1977) or using results of Hale on dynamical systems (Chafee and Infante 1974, Henry 1981). We briefly indicate how the Palais–Smale theory applies.

Let H be the space $L^2([0, L])$ and let H^1 be the Sobolev space which is the domain of $\partial/\partial x$ with Dirichlet boundary conditions. Let H^{-1} be the space of distributions dual to H^1, so that we have $H^1 \subset H \subset H^{-1}$.

Lemma 8.1. (Palais–Smale compactness condition) Assume that u_n is a sequence of elements of H^1 such that $S(u_n) \leqslant c$ for some constant c and such that $S'(u_n) \to 0$ in H^{-1}. Then there is a subsequence u_m with $u_m \to u$ in H^1, and $S'(u) = 0$.

Proof. We have

$$S'(u) = -\partial^2 u/\partial x^2 + V'(u). \tag{8.1}$$

The linear operator $-\partial^2/\partial x^2$ is bicontinuous from H^1 to H^{-1}, and the map $u \to V'(u)$ is even compact from H^1 to H^{-1}. Since $S(u_n) \leqslant c$, the u_n are bounded in H^1. Hence there is a weakly convergent subsequence that converges to some u in H^1. This in turn has a subsequence u_m such that $V'(u_m) \to V'(u)$ strongly in H^{-1}. But since $S'(u_m) \to 0$ in H^{-1}, $-\partial^2 u_m/\partial x^2 \to -\partial^2 u/\partial x^2$ in H^{-1}, and so $u_m \to u$ in H^1. Since $S'(u_m) \to S'(u)$ in H^{-1}, we must have $S'(u) = 0$.

It is not difficult to show that the gradient flow equation $\partial u/\partial t = -S'(u)$ has solutions that remain in H^1. We now show that every orbit leads to a critical point in H^1.

Proposition. Let $u(t)$ be an orbit of the gradient flow. Then there is a critical point u in H^1 such that $u(t) \to u$ in H^1 as $t \to \infty$.

Proof. We have

$$\mathrm{d}S[u(t)]/\mathrm{d}t = \langle S'[u(t)], \mathrm{d}u(t)/\mathrm{d}t\rangle = -\|S'[u(t)]\|^2. \tag{8.2}$$

This shows that $S[u(t)]$ is decreasing. Since it is also bounded below, $\mathrm{d}S[u(t)]/\mathrm{d}t \to 0$ as $t \to \infty$. It follows from (8.2) that $S'[u(t)] \to 0$ in H and hence in H^{-1}. Lemma 8.1 now implies that the set of H^1 limit points of $u(t)$ is non-empty and consists of critical points.

We now wish to show that this set contains only one critical point. We know that there are only finitely many critical points. If the orbit had more than one of them as a limit point, then the orbit would have to pass in a region bounded away from the critical points infinitely often. But then lemma 8.1 would produce yet another critical point in this region, which is a contradiction. This finishes the proof.

The rest of this section is devoted to a discussion of the initial conditions that lead to the absolute minima of S under the gradient flow. The point of interest is that we want to leave the Sobolev space and look at initial conditions that are only known to be continuous functions. On such initial conditions S will usually assume the value $+\infty$. This is a nuisance, but it is forced on us by the fact that the equilibrium probability of the Sobolev space is zero. We shall discuss this point further in § 10.

We look at the gradient system with an initial condition u_0 in the space $C_{\mathrm{D}}([0, L])$. We know that the solutions are bounded. We wish to show that for fixed t the map $u_0 \to u(t)$ is continuous from $C_{\mathrm{D}}([0, L])$ to the Sobolev space H^1. We shall in fact show more, namely, that it is continuous from the Hilbert space $H = L^2([0, L])$ to H^1.

Lemma 8.3. For $t > 0$ the solution $u(t)$ in H^1 is continuous in the initial condition u_0 in H.

Proof. First we show that the evolution is continuous from H to H. If $\bar{u}_0$ is the change in the initial condition and $\bar{u}(t)$ is the change in the solution at time t, then

$$\bar{u} = g\bar{u}_0 - G(\bar{u}h), \tag{8.3}$$

where $|h|$ is bounded by some constant c. Since

$$Gf(t) = \int_0^t \exp[(t-s)\partial^2/\partial x^2]f(s)\,\mathrm{d}s, \tag{8.4}$$

we have

$$\|\bar{u}(t)\| \leqslant \|\bar{u}_0\| + c\int_0^t \|\bar{u}(s)\|\,\mathrm{d}s. \tag{8.5}$$

It follows from Gronwall's inequality that

$$\|\bar{u}(t)\| \leqslant \|\bar{u}_0\| \exp(ct). \tag{8.6}$$

This concludes the proof.

Now we use this to show that the evolution is continuous from H to H^1. The H^1 norm is the norm $\|f\|_1^2 = \|\partial f/\partial x\|^2 + \|f\|^2$. Thus we apply $\partial/\partial x$ to equation (8.3) and use the operator estimate

$$\|\partial/\partial x \exp[(t-s)\partial^2/\partial x^2]\| \leqslant (t-s)^{-1/2}. \tag{8.7}$$

This estimate follows from the spectral theorem applied to the self-adjoint operator

$-\partial^2/\partial x^2 = (\partial/\partial x)^*(\partial/\partial x)$. It follows that

$$\|\partial/\partial x \bar{u}(t)\| \leq t^{-1/2}\|\bar{u}_0\| + c\int_0^t (t-s)^{-1/2}\|\bar{u}(s)\|\, \mathrm{d}s. \tag{8.8}$$

Insert (8.6) in the integral in (8.8). This gives the estimates for $\|\bar{u}\|_1$ in terms of $\|\bar{u}_0\|$, valid for fixed $t>0$. The proof is finished.

If u is a critical point of S, we define the basin of attraction of u to be the set of all initial conditions u_0 such that the orbit $u(t)$ starting at u_0 approaches u as $t \to \infty$.

Theorem 8.4. The basins of attraction of the absolute minima $\pm u_1$ are open sets in the space $C_D([0, L])$ of continuous functions.

Proof. Let N be a Sobolev neighbourhood of u_1 contained in the basin of attraction of u_1. This exists because u_1 is a stable minimum, by theorem 7.8. Let u_0 be an arbitrary element of the basin of attraction of u_1. The solution $u(t)$ starting at u_0 is eventually in N for sufficiently large t. By lemma 8.3, if v_0 is uniformly close to u_0, then the corresponding solution $v(t)$ is in N. Hence $v(t) \to u_1$ as $t \to \infty$. It follows that v_0 is also in the basin of attraction of u_1. This proves that the basin of attraction of u_1 is open in the uniform norm. The proof for $-u_1$ is the same.

9. The lower bound for the probability of tunnelling

In this section we obtain a lower bound for the probability of tunnelling. This follows from the first Ventsel'–Freidlin estimate and an upper bound on the space–time action $I(A)$. The upper bound may be obtained by inserting a trial function. The main contribution comes from the change ΔS of the space equilibrium action during a path up the gradient of S. Thus this uphill path is a mechanism for tunnelling. We shall see in the next section that there is no easier mechanism.

The technique is based on the obvious estimate

$$I(A) \leq I(u) \tag{9.1}$$

for u in A. We need only choose u carefully in order to get the required upper bound. Throughout this section we make the assumption that the parameter $\mu > \pi^2/L^2$, so that there is a non-trivial critical point u_1 with $u_1 > 0$ in the open interval $(0, L)$ and with $S(u_1) < 0$.

Let u_0 be in an open uniform neighbourhood N of $-u_1$ and let Y be an open uniform neighbourhood of u_1. We are interested in the transition event

$$A = \{u \in C_{Du_0}\colon u(0) = u_0 \text{ and } u(T) \in Y\}. \tag{9.2}$$

Theorem 9.1. Let N be in the basin of attraction of $-u_1$. Then for all $\delta > 0$, there is a $T < \infty$ such that

$$I(A) \leq 2\Delta S + \delta. \tag{9.3}$$

The proof consists of the construction of the appropriate trial function that starts at u_0 in N at time 0 and is in Y at time T. The time interval $[0, T]$ is divided into five parts: (i) descent to near $-u_1$; (ii) transition near $-u_1$; (iii) ascent to near u_2; (iv) transition near u_2; (v) descent to near u_1. The main contribution will come from part (iii), the ascent. We shall show that the other contributions are small.

We begin with a lemma that will be used to estimate the transitions in stages (ii) and (iv). The idea is to construct a trial function by linear interpolation.

Lemma 9.2. Let $u(t)=u(a)(1-t/\tau)+u(b)t/\tau$ be the linear path from $u(a)$ to $u(b)$ in time τ. Assume that $u(a)$ and $u(b)$ are uniformly bounded and have $S'[u(a)]<\infty$ and $S'[u(b)]<\infty$. Then there exists a constant $c<\infty$ and a τ with $0<\tau<\infty$ such that the corresponding u satisfies $I(u)\leqslant 2c\|u(a)-u(b)\|$.

Proof. We have

$$I(u)=\tfrac{1}{2}\int_0^\tau \|u(t)/\partial t+S'[u(t)]\|^2\,\mathrm{d}t\leqslant\int_0^\tau \{\|\partial u(t)/\partial t\|^2+\|S'[u(t)]\|^2\}\,\mathrm{d}t. \tag{9.4}$$

Since

$$\partial u(t)/\partial t=[u(b)-u(a)]/\tau, \tag{9.5}$$

the first contribution is $\|u(b)-u(a)\|^2/\tau$.

The other term may be estimated by using

$$S'[u(t)]=(1-t/\tau)S'[u(a)]+(t/\tau)S'[u(b)]+\rho, \tag{9.6}$$

where ρ is a bounded function. It follows that $\|S'[u(t)]\|\leqslant c$ for all t between 0 and τ. The total contribution to the action is thus

$$I(u)\leqslant\|u(b)-u(a)\|^2/\tau+c^2\tau. \tag{9.7}$$

Take $\tau=\|u(b)-u(a)\|/c$. This gives the result.

Proof of theorem 9.1.
We construct the five paths separately and then piece them together into one large path.

The first path is constructed by starting at u_0 and following the gradient flow until it is very close to $-u_1$. For large t the gradient $S'[u(t)]$ stays bounded, and the difference between $u(t)$ and $-u_1$ may be made arbitrarily small in Sobolev norm. The contribution of this path to the action is zero.

The fifth path is constructed by starting at u_0 near u_2 and following the gradient flow until it gets near to $+u_1$. Since u_2 is a saddle point, it is certainly possible to find starting points u_0 close to u_2 in the Sobolev norm on which S is strictly smaller than $S(u_2)$. Such a starting point can lead only to u_1 or $-u_1$. By symmetry there must be points that lead to u_1 and also points that lead to $-u_1$. Since the eigenvectors of $S''(u_2)$ are in H^2, we may even take this starting point u_0 close to u_2 in the H^2 norm. This gives a bound on $S'(u_0)$. The contribution of this path to the action is also zero.

The third path is constructed in the same way as the fifth path, except that is then reversed and goes up the gradient from near $-u_1$ to near u_2. Thus it makes a sizable contribution to the action. This contribution is bounded by

$$\begin{aligned}\tfrac{1}{2}\int_0^T &\|\mathrm{d}u(t)/\mathrm{d}t+S'[u(t)]\|^2\mathrm{d}t\\ &=2\int_0^T\|S'[u(t)]\|^2\,\mathrm{d}t=2\int_0^T\langle S'[u(t)],\mathrm{d}u(t)/\mathrm{d}t\rangle\,\mathrm{d}t\\ &=2\{S[u(T)]-S[u(0)]\}\leqslant 2\Delta S.\end{aligned}$$

The fourth path is constructed by linear interpolation, as in lemma 9.2. The initial data for the third and fifth paths may be taken near enough to u_2 so that the contribution of the fourth path is less than $\delta/2$. This may require rather long time intervals for the third and fifth paths.

The second path is also constructed by linear interpolation. The times needed for the first and third paths may be taken so long that the final points near $-u_1$ are very close. The contribution from this second path is then also less than $\delta/2$.

The total contribution to the action is thus bounded by $0+\delta/2+2\Delta S+\delta/2+0=2\Delta S+\delta$.

10. The upper bound for the probability of tunnelling

In this section we obtain the upper bound for the probability of tunnelling. This follows from the second Ventsel'–Friedlin estimate and a lower bound on the space–time action $I(\bar{A})$. This lower bound is more difficult since we need information on the value of the functional $I(u)$ for all possible trial functions u in $\bar{A}$. The conclusion is nevertheless that no trial function gives an easier mechanism for tunnelling than going up the gradient. The change ΔS in equilibrium action is indeed an obstacle to tunnelling.

Let

$$\bar{A}=\{u\in C_{\mathrm{D}u_0}u(\cdot,0)=u_0(\cdot)\text{ and }u(\cdot,T)\subset\bar{Y}\},\tag{10.1}$$

where u_0 is in a open uniform neighbourhood N of $-u_1$ and Y is an open uniform neighbourhood of u_1. We choose Y small enough so that u in Y implies that $u(L/2)>0$ is bounded away from zero.

Let K be a compact set in the space $C_{\mathrm{D}}([0,L])$. We will insist that our initial condition u_0 belong to the set K. In order to get an interesting result the equilibrium probability of K should be strictly positive. This is not difficult to arrange by taking K to be a closed and bounded set of functions satisfying a uniform Hölder condition.

Theorem 10.1. For all $\zeta>0$ there is a neighbourhood N of $-u_1$ such that if u_0 is in $N\cap K$, then

$$2\Delta S-\zeta\leqslant I(\bar{A}).\tag{10.2}$$

The proof will be given under more restrictive assumptions, and then these assumptions will be relaxed. The heart of the proof is the topological argument in the following lemma.

Lemma 10.2 Let $u_0=-u_1$ and assume that u is regular. Then

$$2\Delta S\leqslant I(u)\tag{10.3}$$

for all u in $\bar{A}$.

Proof. If u is regular, then for $0\leqslant T'\leqslant T$ we may write

$$\begin{aligned}I(u)&=\tfrac{1}{2}\int_0^T\|\partial u/\partial t+S'(u)\|^2\,\mathrm{d}t\geqslant\tfrac{1}{2}\int_0^{T'}\|\partial u/\partial t+S'(u)\|^2\,\mathrm{d}t\\&=\tfrac{1}{2}\int_0^{T'}\|\partial u/\partial t-S'(u)\|^2\,\mathrm{d}t+2\{S[u(T')]-S[u(0)]\}\\&\geqslant 2\{S[u(T')]-S[u(0)]\}.\end{aligned}\tag{10.4}$$

Since $u(L/2, 0) = u_0(L/2) < 0$ and $u(L/2, T) > 0$, we must have $u(L/2, T') = 0$ for some T'. However, for such a function $u(\cdot, T')$ we may apply the variational characterisation of the zero-node solutions (lemmas 7.3 and 7.5) to each interval $[0, L/2]$ and $[L/2, L]$ to conclude that $S[u(T')] \geq S(u_2)$. Thus

$$I(u) \geq 2[S(u_2) - S(u_0)] = 2\Delta S. \tag{10.5}$$

Lemma 10.3. The theorem is true when $u_0 = -u_1$.

Proof. Let u in $\bar{A}$ be such that $I(u) = I(\bar{A})$. By theorem 6.9 we may choose a sequence of regular u_n such that $u_n \to u$ uniformly and $I(u_n) \to I(u)$. Thus if n is sufficiently large, then $u_n(L/2, T) > 0$ and so the argument of the lemma applies to show that

$$2\Delta S \leq I(u_n). \tag{10.6}$$

But n may be taken so large that also

$$2\Delta S \leq I(u_n) \leq I(u) + \zeta = I(\bar{A}) + \zeta. \tag{10.7}$$

This is the estimate of the theorem.

In the following it will be convenient to introduce a somewhat different action functional. Define

$$\tilde{I}_0(z) = I_0[z - gz(0)]. \tag{10.8}$$

This is formally

$$\tilde{I}_0(z) = \tfrac{1}{2}\|(\partial/\partial t - \partial^2/\partial x^2)z\|^2 \tag{10.9}$$

with Dirichlet boundary conditions at $x = 0, L$, but no boundary conditions at $t = 0, T$. Similarly, define

$$\tilde{I}(u) = \tilde{I}_0[\Psi^{-1}(u)] = \tilde{I}_0(z), \tag{10.10}$$

where $u = \Psi(z)$ is the solution of the nonlinear equation $u = -GV'(u) + z$. This is formally

$$\tilde{I}(u) = \tfrac{1}{2}\|(\partial/\partial t - \partial^2/\partial x^2)u + V'(u)\|^2 \tag{10.11}$$

with Dirichlet boundary conditions at $x = 0, L$, but no boundary conditions at $t = 0, T$.

The relation between $\tilde{I}$ and the previous I is the obvious one

$$\tilde{I}(u) = I(u) \qquad \text{when } u(0) = u_0. \tag{10.12}$$

This is because $\tilde{I}(u) = \tilde{I}_0(z) = I_0[z - gz(0)] = I_0(z - gu_0) = I_0(\varepsilon w) = I(u)$, where $u = -GV'(u) + z = -GV'(u) + \varepsilon w + gu_0$.

The functional $\tilde{I}$ has the nice property of being related to I_0 in a way that is independent of the initial condition. However, the price one pays for this is a certain loss of compactness. Thus we need to assume compactness of the initial conditions.

Proposition 10.4. Let K be a compact set in $C_D([0, L])$. Then the set $\{u \in C_D([0, L] \times [0, T])$: $\tilde{I}(u) \leq s$ and $u(0) \in K\}$ is compact.

Proof. This set is the image under Ψ of the set $\{z \in C_D([0, L] \times [0, T]$:$z = \varepsilon w + gu_0$, $I_0(\varepsilon w) \leq s$ and $u_0 \in K\}$. This set is compact, by proposition 4.2. Since Ψ is continuous, the image is also compact.

It is also true that $\tilde{I}$ is lower semi-continuous, since it inherits this property from $\tilde{I}_0$ and eventually from I_0.

The following is the fundamental result that lets us prove the theorem for typical initial conditions. We indicate the dependence of the tunnelling event on the initial condition by writing it as $\bar{A}(u_0)$.

Theorem 10.5. Let K be a compact set in $D_D([0, L])$. Then $I[\bar{A}(u_0)]$ is lower semi-continuous in u_0, in the uniform norm, as u_0 ranges over the compact set K. Here I is defined by (6.2).

Proof. By the above discussion it is enough to show that $\tilde{I}[\bar{A}(u_0)]$ is lower semi-continuous in u_0. In other words, we must show that for every $\zeta > 0$ there is a neighbourhood N of u_0 such that if v is in $N \cap K$, then

$$\tilde{I}[\bar{A}(u_0)] - \zeta \leq \tilde{I}[\bar{A}(v)]. \tag{10.13}$$

Define

$$\bar{A}(N_\delta \cap K) = \{u : u_0 \in N_\delta \cap K \text{ and } u(T) \in \bar{Y}\} \tag{10.14}$$

where N_δ is the closed δ ball about u_0. We know by compactness and lower semi-continuity that there is a u_δ in $\bar{A}(N_\delta \cap K)$ such that

$$\tilde{I}(u_\delta) = \tilde{I}[\bar{A}(N_\delta \cap K)].$$

We also know by compactness that there is a subsequence $u_\delta \to u$ for some u. Clearly u is in $\bar{A}(u_0)$. By lower semi-continuity,

$$\tilde{I}(u) \leq \tilde{I}(u_\delta) + \zeta$$

for δ sufficiently small and in the subsequence. In particular

$$\tilde{I}[\bar{A}(u_0)] \leq \tilde{I}(u) \leq \tilde{I}(u_\delta) + \zeta = \tilde{I}[\bar{A}(N_\delta \cap K)] + \zeta \leq \tilde{I}[\bar{A}(v)] + \zeta \tag{10.15}$$

for all v in $N_\delta \cap K$.

Proof of Theorem 10.1. By lemma 10.3 the theorem is true if we start at $-u_1$. Thus for every $\zeta > 0$

$$I[\bar{A}(-u_1)] \geq 2\Delta S - \zeta/2. \tag{10.16}$$

However, by theorem 10.5, there is a neighbourhood N of $-u_1$ so that for v in $N \cap K$,

$$I[\bar{A}(v)] \geq I[\bar{A}(-u_1)] - \zeta/2. \tag{10.17}$$

Thus for such v

$$I[\bar{A}(v)] \geq 2\Delta S - \zeta. \tag{10.18}$$

Remark 1. One could also base the proof of the topological lemma 10.2 on the concept of basin of attraction. If the path from $-u_1$ to u_1 remained below $S(u_2)$, then every $u(t)$ on the path would be attracted either to $-u_1$ or u_1. This would divide the time interval $[0, T]$ into two disjoint open non-empty sets, which would contradict its connectedness. Related minimax arguments have been applied to a variety of situations by Ambrosetti and Rabinowitz (1973).

Remark 2. It is important in the above results that the initial u_0 need only be near the minimum of the action in the uniform norm. We can appreciate the reason for this by returning to the equilibrium measure given formally by (1.1). The measure may be written

$$\mathrm{d}\mu(u) = \exp\left(-(1/\varepsilon^2)\int_0^T V(u)\,\mathrm{d}x\right)\mathrm{d}\mu_0(u)$$

where $\mathrm{d}\mu_0(u)$ is the measure whose formal density is given by

$$\mathrm{d}\mu_0(u) = \exp\left(-(1/\varepsilon^2)\int_0^L (\partial u/\partial x)^2\,\mathrm{d}x\right)\mathrm{d}^\infty u.$$

However, this is simply a Gaussian process with covariance operator $(\varepsilon^2/2)\,(\partial^2/\partial x^2)^{-1}$. Thus it is well defined and its properties are easy to compute.

This shows that the equilibrium measure (the invariant measure of the process) is absolutely continuous with respect to this Gaussian measure. It is known that this Gaussian process (the Brownian bridge) has continuous, in fact Hölder continuous, sample paths, with probability one. Thus uniform neighbourhoods will have strictly positive probability, and so our initial starting configurations can be regarded as representative of reasonable equilibrium configurations.

Acknowledgments

We wish to thank Jean Ginibre for useful conversations on the subject of this paper, and Edouard Brézin for a stimulating seminar on related developments in quantum field theory. In addition, we thank Charles Stuart for comments on nonlinear analysis and René Carmona for remarks on Gaussian processes. One of the authors (GJ-L) thanks the Centre de Physique Théorique, Marseille-Luminy, for hospitality during part of this work.

References

Albeverio S and Hoegh-Krohn R 1977 *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **40** 1–57

Ambrosetti A and Rabinowitz P H 1973 *J. Funct. Anal.* **14** 349–81

Asorey M and Mitter P K 1981 *Commun. Math. Phys.* **80** 43–58

Azencott R 1980 in Azencott R, Guivarc'h Y and Gundy R F *Ecole d'Eté de probabilités de Saint-Flour VIII-1978, Lecture notes in Math.* #774 (Berlin: Springer)

Belopolskaia Ya I and Daletzki Yu L 1978 *Trudy Mosk. Mat. Obsch.* **37** 107–41

Berger M S 1977 *Nonlinearity and Functional Analysis* (New York: Academic)

Chafee N and Infante E 1974 *Applicable Analysis* **4** 17–37

Coffman C V 1975 *J. Differential Equations* **17** 46–60

Colella P and Lanford O E 1973 in Velo G and Wightman A *Constructive Quantum Field Theory, Lecture Notes in Physics* #25 (Berlin: Springer)

Doss H and Royer G 1978 *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **46** 107–24

Faris W G 1975 *Self-Adjoint Operators, Lecture Notes in Math.* #433 (Berlin: Springer)

——1979 *J. Funct. Anal.* **32** 342–52

——1980 *Trans. Am. Math. Soc.* **261** 579–87

Fernique X M 1975 in Hennequin P L *Ecole d'Eté de Probabilités de Saint-Flour IV-*1974, *Lecture Notes in Math.* #480 (Berlin: Springer)

Hempel J A 1971 *Indiana Univ. Math. J.* **20** 983–96

Henry D 1981 *Geometric Theory of Semilinear Parabolic Equations, Lecture Notes in Math.* #840 (Berlin: Springer)
Holley R A and Stroock D W 1976a *Ann. Probability* **4** 195–228
——1976b *Commun. Math. Phys.* **48** 249–65
Jona-Lasinio G, Martinelli F and Scoppola E 1981 *Commun. Math. Phys.* **80** 223–54
Kozlov C M 1978 *Trudy Sem. im. I. G. Petrovskogo* **4** 147–72
Laetsch T 1975 *Indiana Univ. Math. J.* **24** 651–8
Marcus R 1974 *Trans. Am. Math. Soc.* **198** 177–90
——1978 *J. Funct. Anal.* **29** 275–86
——1979 *Pacific J. Math.* **84** 143–53
Parisi G and Wu Yong-shi 1981 *Scientia Sinica* **24** 483–96
Protter M H and Weinberger H F 1967 *Maximum Principles in Differential Equations* (Englewood Cliffs NJ: Prentice-Hall)
Royer, G 1979 *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **46** 165–78
Segal I E 1963 *Ann. Math.* **78** 339–64
Stroock D W and Varadhan S R S 1979 *Multidimensional Diffusion Processes* (New York: Springer)
Varadhan S R S 1966 *Commun. Pure Appl. Math.* **19** 261–86
Ventsel' A D and Freidlin M I 1970 *Russ. Math. Surveys* **25** 1–55
——1979*Fluctuations in Dynamical Systems Under the Action of Random Perturbations* (Moskva: Nauka)