Article ID jdeq.1998.3594, available online at http://www.idealibrary.com on IDEAL

42

Ergodicity of Dissipative Differential Equations Subject to Random Impulses

J. M. Sanz-Serna*

Departamento Matemática Aplicada y Computación, Facultad de Ciencias, Universidad de Valladolid, Spain

and

A. M. Stuart[†, ‡]

Scientific Computing and Computational Mathematics Program, Division of Mechanics and Computation, Durand 252, Stanford University, Stanford California 94305-4040

Received December 1, 1997; revised June 12, 1998

Differential equations subject to random impulses are studied. Randomness is introduced both through the time between impulses, which is distributed exponentially, and through the sign of the impulses, which are fixed in amplitude and orientation. Such models are particular instances of piecewise deterministic Markov processes and they arise naturally in the study of a number of physical phenomena, particularly impacting systems. The underlying deterministic semigroup is assumed to be dissipative and a general theorem which establishes the existence of invariant measures for the randomly forced problem is proved. Further structure is then added to the deterministic semigroup, which enables the proof of ergodic theorems. Characteristic functions are used for the case when the deterministic component forms a damped linear problem and irreducibility measures are employed for the study of a randomly forced damped double-well nonlinear oscillator with a gradient structure.

Key Words: random impulses; stochastic differential equations; dissipative systems; gradient systems; nonlinear oscillators; ergodic theorem.

1. INTRODUCTION

In this paper we study the ergodic behaviour of ordinary differential equations (ODEs) subject to random impulses—specifically impulses of a fixed amplitude and orientation which occur after random, exponentially distributed intervals of time, and have a random sign.

* Supported by project PB95-705.

† Supported by the Office of Naval Research under Grant N00014-92-J-1876. Most of this research was performed while this author was Iberdrola Visiting Professor at Valladolid.

‡ Current address: Mathematics Institute, University of Warwick, Coventry CV4 7AL, England.

0022-0396/99 $30.00

Given a vector $w \in \mathbb{R}^m$ (the impulse), an initial state u_0, and a vector field $f \in C^\infty(\mathbb{R}^m, \mathbb{R}^m)$, we consider the equation

$$\frac{du}{dt} = f(u) + \sum_{n=1}^{\infty} \theta_n w \delta(t - \tau_n), \qquad u(0^+) = u_0 + \theta_0 w. \tag{1}$$

Here $\delta(\cdot)$ denotes a unit point mass at the origin; the $\{\theta_n\}_{n=0}^{\infty}$ are independent, identically distributed (IID) random variables with $\mathbb{P}\{\theta_0 = \pm 1\} = \frac{1}{2}$; and the waiting times $t_n = \tau_{n+1} - \tau_n$, $n \in \mathbb{Z}^+$, $\tau_0 = 0$, are IID random variables exponentially distributed with parameter λ. Moreover we assume that the sequences $\{\theta_n\}_{n=0}^{\infty}$ and $\{t_n\}_{n=0}^{\infty}$ are independent. Here $\mathbb{N} = \{1, 2, 3, ...\}$ and $\mathbb{Z}^+ = \{0, 1, 2, ...\}$.

We assume that the vector field is such that, for any initial data, the deterministic problem underlying (1) has a unique solution for all $t \geqslant 0$. We denote by $S \in C^\infty(\mathbb{R}^m \times \mathbb{R}^+, \mathbb{R}^m)$ the semigroup generated by the problem

$$\frac{du}{dt} = f(u), \qquad u(0) = U; \tag{2}$$

thus the solution of (2) is $u(t) = S(U, t)$ for $t \geqslant 0$. Occasionally we will use $S(U, \cdot)$ with negative argument, noting that this is well-defined on a U dependent interval of $\mathbb{R}^-$. Using this semigroup we give Eq. (1) a precise interpretation. Its solution is

$$u(t) = S(u(\tau_n^-) + \theta_n w, t - \tau_n), \qquad t \in (\tau_n, \tau_{n+1}), \quad n \in \mathbb{N}.$$

This follows by noting that, for $t \neq \tau_n$, Eq. (1) is a standard ODE, whilst integrating over $(\tau_n - \varepsilon, \tau_n + \varepsilon)$ and letting $\varepsilon \to 0$ yields $u(\tau_n^+) = u(\tau_n^-) + \theta_n w$. Note that u is not defined at the impulse times τ_n. One could define $u(0) = u_0$, $u(\tau_n) = u(\tau_n^-)$, $n \in \mathbb{N}^+$ for example, a choice that makes u left-continuous with right-hand limits, but this definition is not used later.

If we set $u_n := u(\tau_n^-)$, $n \in \mathbb{N}^+$, then we obtain

$$u_{n+1} = S(u_n + \theta_n w, t_n), \qquad n \in \mathbb{N}. \tag{3}$$

Since the pairs (θ_n, t_n) are IID, Eq. (3) generates a discrete-time, homogeneous Markov chain with uncountable state-space $\mathbb{R}^m$.

Equations such as (1) arise naturally in a variety of different contexts. In [2] a collection of *deterministic* problems arising in engineering and involving differential equations subject to impulses are studied; Eq. (1) provides a natural first step into the study of *randomly* occuring impulses. Motivated by applications in [14], the article [7] contains calculations of first-passage times for (1) with $f \equiv 0$ and the θ_j exponentially distributed. The paper [17] analyzes a linear problem of the form (1) but with θ_j depending upon u (multiplicative noise). Both [7] and [17] rely heavily on explicit calculation of probability densities. An equation similar to (1) with multiplicative noise is studied numerically in [10]; the numerical method revolves around direct solution of the integral equation governing probability density evolution. A problem closely related to ours is to consider a vector field with random parameter which flips between two possible values at random times—see [13]. Also (1) provides a particular instance of a piecewise deterministic Markov process as in the important work of Davis [3, 4]; however, we find it expedient to work directly with the Markov chain (3) rather than the Markov process generated by (1).

We were originally motivated by random generalizations of the impulse oscillators described in [2] and for such problems (2) represents a damped oscillator; therefore our analysis will concern problems where a dissipative mechanism is present. Our main contribution consists of some nontrivial ergodicity proofs; ergodicity is a consequence of the balance between deterministic dissipation and mean input of energy through noise.

However, ergodicity is difficult to prove in general for problems like (3) essentially because a non-trivial interaction between the direction of random impulses, and the underlying deterministic flow, is necessary to establish irreducibility. Natural generalizations of our problem, for which it would be easier to prove ergodicity, would allow random rotations of the impulse vector, at least within some subspace of $\mathbb{R}^m$, and impulses which have continuously distributed amplitude. For our problem, in which orientation and amplitude of the impulse vector are fixed, we have been able to prove ergodicity only for two examples. In the first, $m=1$ and hence the orientation of the impulse vector is immaterial. In the second, the dimension is $m=2$ and the rotational character of the deterministic flow is combined with fixed orientation impulses to deduce irreducibility; once this has been done, a Lyapunov–Foster drift condition essentially gives ergodicity.

In Section 2 we prove the existence of invariant measures for (3) under two distinct structural assumptions on f, both inducing some form of dissipation in the deterministic problem (2). In Section 3 we study ergodicity for the linear scalar problem $f(u) = -\gamma u$; characteristic functions are employed. Similar issues are addressed for a damped double-well nonlinear oscillator in Section 4, employing the abstract theory of Markov chains in uncountable state spaces from [11]. We note that the requirement that the impulses have constant direction appears naturally in this nonlinear oscillator example and makes ergodicity more difficult to prove.

2. DISSIPATIVE IMPULSE SYSTEMS

In this section we introduce the two classes of vector fields f in (1) of interest to us. Both are *dissipative* (see [6]) in the sense that, in the absence of impulses, for the semigroup $S(\cdot, \cdot)$ generated by (2), there is a bounded positively invariant *absorbing set* B such that, for any $U \in \mathbb{R}^m$, $S(U, t) \in B$ for all $t \geqslant T$, $T = T(U)$. Clearly in a dissipative system all positive orbits are bounded; this implies in particular that the solution of (1) exists for all positive $t > 0$ provided that the waiting times t_n are such that $\sum t_n = \infty$, which happens with probability 1. Roughly speaking, the balance between this dissipativity property and the expected net input of energy through noise will lead to the existence of invariant probability measures. Situations where similar balances between dissipation and noise take place are of course very common in statistical mechanics.

Throughout, $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ denote the Euclidean norm and inner product.

2.1. *The First Class*

In the first class of problems, the semigroup is generated by (2) under the assumption that there exist $\alpha > 0$ and $\beta > 0$ such that, for each $u \in \mathbb{R}^m$,

$$\langle f(u), u \rangle \leqslant \alpha - \beta \|u\|^2. \tag{4}$$

Such a semigroup is dissipative with absorbing set given by the ball of centre 0 and radius $\alpha/\beta + \varepsilon$ with arbitrary $\varepsilon > 0$. Examples include the scalar linear problem

$$m = 1, \qquad f(u) = -\gamma u, \qquad \gamma > 0 \tag{5}$$

and the Lorenz equations ($m = 3$)

$$\begin{aligned} x_t &= -\sigma(x - y), \\ y_t &= rx - y - xz, \\ z_t &= -bz + xy. \end{aligned} \tag{6}$$

For Eq. (6) with $u = (x, y, z - r - \sigma)$ the governing equations then satisfy (4) (see [15]).

We now proceed to show that, for problems of this class, the expected value of $\|u_n\|^2$ is ultimately bounded independently of the initial ($n = 0$) distribution μ of the Markov chain. This is a stochastic analog of the dissipativity of the underlying deterministic systems.

Here and elsewhere, $\mathbb{E}^\mu$ denotes unconditional expectation when the initial state u_0 is distributed according to μ. Furthermore the notation $\mathbb{E}(\cdot \mid \mathscr{F}_n)$, $n \in \mathbb{N}$ is used to mean conditional expectation given the σ-algebra of the events that only depend on the first $n + 1$ coordinates $u_0, \ldots, u_n$ (note the superscript μ is not needed in view of the Markov property).

THEOREM 2.1. *Consider the semigroup $S(\cdot,\cdot)$ generated by* (1) *under* (4) *and assume that the Markov chain* (3) *has initial data distributed according to a probability measure* μ *with*

$$\int_{\mathbb{R}^m} \|u\|^2 \mu(du) < \infty. \tag{7}$$

Then,

$$\mathbb{E}(\|u_{n+1}\|^2 \mid \mathscr{F}_n) \leqslant \frac{\lambda}{\lambda+2\beta} \|u_n\|^2 + 2\frac{\alpha+\eta}{\lambda+2\beta}, \tag{8}$$

where

$$\eta := \tfrac{1}{2}\lambda \|w\|^2. \tag{9}$$

Furthermore, $\mathbb{E}^\mu(\|u_n\|^2) < \infty$ *for all* $n \in \mathbb{N}$ *and, for any* $C > (\alpha+\eta)/\beta$, *there is* $N = N(\mu)$ *such that, for* $n \geqslant N$,

$$\mathbb{E}^\mu(\|u_n\|^2) \leqslant C. \tag{10}$$

Proof. Integrating the inequality

$$\frac{1}{2}\frac{d}{dt}\|u\|^2 \leqslant \alpha - \beta\|u\|^2,$$

we obtain, for $t \in (\tau_n, \tau_{n+1})$,

$$\|u(t)\|^2 \leqslant \frac{\alpha}{\beta} + e^{-2\beta(t-\tau_n)}\left(\|u(\tau_n^+)\|^2 - \frac{\alpha}{\beta}\right).$$

Thus, recalling that $u_{n+1} = u(\tau_{n+1}^-)$ and $u(\tau_n^+) = u_n + \theta_n w$,

$$\|u_{n+1}\|^2 \leqslant \frac{\alpha}{\beta} + e^{-2\beta t_n}\left(\|u_n\|^2 + 2\theta_n \langle u_n, w\rangle + \|w\|^2 - \frac{\alpha}{\beta}\right).$$

Noting that t_n and θ_n are independent of $u_0, ..., u_n$ (and of each other) we get, after computing the expected value of $e^{-2\beta t_n}$,

$$\mathbb{E}(\|u_{n+1}\|^2 \mid \mathscr{F}_n) \leqslant \frac{\alpha}{\beta} + \frac{\lambda}{\lambda+2\beta}\left(\|u_n\|^2 + \|w\|^2 - \frac{\alpha}{\beta}\right),$$

which gives (8).

From the last bound,

$$\mathbb{E}^\mu(\|u_{n+1}\|^2) \leqslant 2\frac{\alpha+\eta}{\lambda+2\beta} + \frac{\lambda}{\lambda+2\beta}\mathbb{E}^\mu(\|u_n\|^2),$$

and $\mathbb{E}^{\mu}(\|u_n\|^2) < \infty$ for all $n \in \mathbb{N}$ provided that (7) holds. Also

$$\mathbb{E}^{\mu}(\|u_n\|^2) \leqslant \frac{\alpha+\eta}{\beta}\left(1-\left(\frac{\lambda}{\lambda+2\beta}\right)^n\right)+\left(\frac{\lambda}{\lambda+2\beta}\right)^n \mathbb{E}^{\mu}(\|u_0\|^2),$$

and the final result follows. ∎

Remark. In this result, and in Theorem 2.3, the ultimate bound for $\mathbb{E}^{\mu}(\|u_n\|)^2$ is an $\mathcal{O}(\eta)$ perturbation of the bound in the corresponding deterministic situation. This shows that, in our setup, $\eta \ll 1$ is an appropriate definition of small noise.

2.2. *The Second Class*

In the second class we assume that $m=2l$, $u=(q^T, p^T)^T$, and that the system is of the form

$$\frac{d}{dt}\begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} p \\ -\gamma p - \nabla V(q) \end{pmatrix} \tag{11}$$

with $\gamma > 0$. This, of course, represents a damped mechanical system with potential energy V; q contains particle coordinates and p the corresponding scaled momenta. We assume that $V \geqslant 0$, so that the undamped ($\gamma = 0$) system is oscillatory. We let $u=(q^T, p^T)^T$.

The system can be rewritten in second-order form as

$$\frac{d^2q}{dt^2} + \gamma\frac{dq}{dt} + \nabla V(q) = 0;$$

this is in gradient form, with

$$\frac{d}{dt}E = -\gamma\left(\frac{dq}{dt}\right)^2, \qquad E = E\left(q, \frac{dq}{dt}\right) := \frac{1}{2}\left(\frac{dq}{dt}\right)^2 + V(q). \tag{12}$$

Thus all bounded orbits have ω-limit sets contained in the set of equilibria [6].

We further assume that there are $\alpha > 0$, $\beta \in (0, 1)$, such that, for all $q \in \mathbb{R}^l$,

$$\frac{1}{2}\langle \nabla V(q), q\rangle \geqslant \beta V(q) + \frac{\gamma^2\beta(2-\beta)}{8(1-\beta)}\|q\|^2 - \alpha. \tag{13}$$

This assumption ensures dissipativity as we prove next. Note that dissipativity, combined with the gradient structure, implies that all positive orbits are bounded and have ω-limit sets contained in the (bounded) set of equilibria. In the proof of dissipativity and later, we need the function

$$G(u) = \frac{1}{2}\|p\|^2 + V(q) + \frac{\gamma}{2}\langle p, q\rangle + \frac{\gamma^2}{4}\|q\|^2; \tag{14}$$

this is a perturbation of the standard gradient Lyapunov function E (energy) appearing in (12).

LEMMA 2.2. *Assume that* (13) *holds. Then the following are true.*

(i) *The function* G *defined in* (14) *is positive definite; more precisely for all* $u=(q^T, p^T)^T \in \mathbb{R}^m$

$$G(u) \geqslant \frac{1}{8}\|p\|^2 + \frac{\gamma^2}{12}\|q\|^2. \tag{15}$$

(ii) *Along solutions of* (11),

$$\frac{1}{\gamma}\frac{dG}{dt} \leqslant \alpha - \beta G. \tag{16}$$

(iii) *The system* (11) *is dissipative with absorbing set*

$$B = \left\{u: G \leqslant \frac{\alpha}{\beta} + \varepsilon\right\}$$

for any $\varepsilon > 0$.

Proof. Since, for any $\delta > 0$,

$$\langle p, \gamma q\rangle \geqslant -\left(\frac{\delta}{2}\|p\|^2 + \frac{\gamma^2}{2\delta}\|q\|^2\right)$$

and $V(q)$ is positive, we have

$$G(u) \geqslant \frac{1}{2}\left(1-\frac{\delta}{2}\right)\|p\|^2 + \frac{\gamma^2}{4}\left(1-\frac{1}{\delta}\right)\|q\|^2.$$

Choosing $\delta = \frac{3}{2}$ leads to (15).

Next, for any $\delta > 0$,

$$\beta G \leqslant \frac{\beta}{2}\|p\|^2 + \beta V(q) + \frac{\gamma\beta}{2}\left(\frac{\delta}{2}\|p\|^2 + \frac{1}{2\delta}\|q\|^2\right) + \frac{\gamma^2\beta}{4}\|q\|^2,$$

which for $\delta = 2(1-\beta)/(\gamma\beta)$ implies, by (13),

$$\begin{aligned}\beta G &\leqslant \frac{1}{2}\|p\|^2 + \beta V(q) + \frac{\gamma^2\beta(2-\beta)}{8(1-\beta)}\|q\|^2\\ &\leqslant \frac{1}{2}\|p\|^2 + \frac{1}{2}\langle \nabla V(q), q\rangle + \alpha.\end{aligned}$$

The estimate (16) follows from here, because

$$\frac{dG}{dt} = \left\langle p, \frac{dp}{dt} \right\rangle + \left\langle \nabla V(q), \frac{dq}{dt} \right\rangle + \frac{\gamma}{2} \left\langle p, \frac{dq}{dt} \right\rangle + \frac{\gamma}{2} \left\langle q, \frac{dp}{dt} \right\rangle + \frac{\gamma^2}{2} \left\langle q, \frac{dq}{dt} \right\rangle$$

$$= -\frac{\gamma}{2} \|p\|^2 - \frac{\gamma}{2} \langle \nabla V(q), q \rangle.$$

Part (iii) is a straightforward consequence of (16). ∎

Examples of systems (11) satisfying (13) include any problem where $V(q)$ is polynomial with leading term

$$V(q) = \frac{a}{2} \|q\|^{2r} + \cdots, \qquad a > 0, \quad r \in \mathbb{N}^+.$$

For problems (11) we consider always impulses of the form

$$w = (0^T, v^T)^T \tag{17}$$

with $v \in \mathbb{R}^l$. Such impulses have no effect on the positions q and cause a jump in the momenta p, thus mimicking mechanical collisions. We will be particularly interested in the one-degree-of-freedom case

$$l = 1, \qquad V(q) = \tfrac{1}{4}(1 - q^2)^2. \tag{18}$$

This is a double well potential leading to an unstable equilibrium at $q = 0$ and stable equilibria at $q = \pm 1$.

The proof of the following result will be omitted; it is similar to that of Theorem 2.1, but this time the starting point is (16). The remark following Theorem 2.1 also applies to Theorem 2.3.

THEOREM 2.3. *Consider the semigroup $S(\cdot, \cdot)$ generated by* (11), *under* (13) *and* (17), *and assume that the Markov chain* (3) *has initial data distributed according to a probability measure μ with*

$$\int_{\mathbb{R}^m} G(u) \, \mu(du) < \infty.$$

Then, with η as in (9),

$$\mathbb{E}(G(u_{n+1}) \mid \mathcal{F}_n) \leqslant \frac{\lambda}{\lambda + \gamma\beta} G(u_n) + \frac{\gamma\alpha + 2\eta}{\lambda + \gamma\beta}. \tag{19}$$

Furthermore $\mathbb{E}^\mu(G(u_n)) < \infty$ *for all* $n \in \mathbb{N}$, *and, for any* $C > (\gamma\alpha + 2\eta)/(\gamma\beta)$, *there is* $N = N(\mu)$, *such that, for* $n \geq N$,

$$\mathbb{E}^\mu\left(\frac{1}{8}\|p_n\|^2 + \frac{\gamma^2}{12}\|q_n\|^2\right) \leq C. \tag{20}$$

2.3. *Existence of an Invariant Probability Measure*

Theorems 2.1 and 2.3 lead to the following result.

THEOREM 2.4. *Under the conditions of either Theorem* 2.1 *or* 2.3, *there is a probability measure* ν *on* $\mathbb{R}^m$ *invariant for* (3).

Proof. It is clear that (10) or (20) imply the tightness of the sequence $\{\mu_n\}$, where μ_n is the distribution law of u_n. By standard arguments (see Proposition 12.1.3 of [11] or Proposition 1.8(e) in [9]) the sequence

$$\nu_n = \frac{1}{n}\sum_{j=1}^{n} \mu_j$$

possesses a weakly convergent subsequence whose limit ν is invariant. ∎

We would really like to know that ν is unique, so that some form of ergodicity follows. This is hard to establish at the level of generality of this section and so in the next two sections we restrict attention to the two examples (5) (covered by Theorem 2.1) and (11), (18) (covered by Theorem 2.3).

3. THE SCALAR LINEAR PROBLEM

In this section we study ergodic properties of the Markov chain (3) where $S(\cdot,\cdot)$ is generated by the scalar linear problem (5). The simplicity of the problem allows use of characteristic functions (Fourier transforms) for the proof; such an approach is instructive as it elucidates some of the structure of the invariant probability measure in a way that the abstract theory of ergodicity for Markov chains, employed for the nonlinear oscillator problem in the next section, does not.

We begin with an auxiliary result, whose validity is not resticted to the linear problem (5).

LEMMA 3.1. *Assume that for the Markov chain* (3) *the probability distributions of* u_n *have densities* ϕ_n, $n \in \mathbb{Z}^+$. *Then*

$$\lambda\phi_{n+1}(u) + \nabla\cdot(f(u)\,\phi_{n+1}(u)) = \frac{\lambda}{2}\left[\phi_n(u+w) + \phi_n(u-w)\right]. \tag{21}$$

Proof. Let $\psi_n(u, t)$ denote the probability density for the deterministic equation (2) corresponding to the initial density

$$\psi_n(u, 0) = \tfrac{1}{2}(\phi_n(u+w) + \phi_n(u-w)). \tag{22}$$

Then, using the expression for the density of the exponential distribution,

$$\phi_{n+1}(u) = \lambda \int_0^\infty e^{-\lambda t} \psi_n(u, t)\, dt,$$

so that $\phi_{n+1}(u)$ is λ times the Laplace transform of $\psi_n(u, t)$. Now $\psi_n(u, t)$ satisfies the Liouville equation

$$\frac{\partial \psi_n}{\partial t} + \nabla \cdot (f\psi_n) = 0$$

and has initial condition (22). Taking λ times the Laplace transform of the Liouville equation leads to the result. ▮

Remark. With η given by (9), Eq. (21) may be rewritten in the following revealing manner:

$$\begin{aligned} &\lambda(\phi_{n+1}(u) - \phi_n(u)) + \nabla \cdot (f(u)\, \phi_{n+1}(u)) \\ &\quad = \frac{\eta}{\|w\|^2} (\phi_n(u+w) - 2\phi_n(u) + \phi_n(u-w)). \end{aligned}$$

This form makes it is clear that, if $\lambda \to \infty$, $\|w\| \to 0$ (the impulses become smaller but more frequent) while keeping η and the direction of w constant, then the equation formally approaches a Fokker–Planck equation, as is to be expected; see, for example, [5, 8, 18].

We are now ready to discuss the linear problem (5). We assume that μ is chosen with

$$\int_{\mathbb{R}} u^2 \mu(du) < \infty,$$

so that, by Theorem 2.1, the variable u_n^2 has finite expectation for all $n \in \mathbb{N}$. This ensures that the characteristic functions

$$\Phi_n(k) = \mathbb{E}^\mu(\exp(iku_n)), \qquad n \in \mathbb{N},$$

are twice differentiable.

LEMMA 3.2. *Consider* (3) *under* (5). *The characteristic function* Φ_{n+1}, $n \in \mathbb{N}$, *can be obtained from* Φ_n *by means of the singular ordinary differential equation*

$$k \frac{d}{dk} \Phi_{n+1}(k) + \delta\Phi_{n+1}(k) = \delta \cos(kw)\, \Phi_n(k), \qquad \delta := \frac{\lambda}{\gamma}. \tag{23}$$

Proof. It is clearly enough to consider the case where the distributions μ_n of u_n possess densities ϕ_n. Then the result follows by taking Fourier transforms in the equation (21), that now reads

$$\lambda\phi_{n+1}(u) + \frac{d}{du}\left[-\gamma u\phi_{n+1}(u)\right] = \frac{\lambda}{2}\left[\phi_n(u+w) + \phi_n(u-w)\right]. \quad \blacksquare$$

We shall show that, for $n \to \infty$, the Φ_n converge pointwise to the smooth function

$$\Phi(k) = \exp\left(\delta \int_0^k \frac{\cos(mw) - 1}{m}\, dm\right) \tag{24}$$

solution of the initial value problem

$$k \frac{d}{dk} \Phi(k) + \delta\Phi(k) = \delta \cos(kw)\, \Phi(k), \qquad \Phi(0) = 1. \tag{25}$$

Note that, in (25), the initial condition is a normalization that Φ needs to satisfy to be the characteristic function of a probability measure, while the differential equation is obtained by imposing $\Phi_{n+1} = \Phi_n$ in (23).

LEMMA 3.3. *With the notation above, for each* $k \in \mathbb{R}$, $\Phi_n(k) \to \Phi(k)$, *as* $n \to \infty$. *Thus* Φ *is the characteristic function of an invariant probability measure* ν *on* $\mathbb{R}^m$ *and the probability distributions* μ_n *of* u_n *converge weakly to* ν *as* $n \to \infty$.

Proof. The functions $\Delta_n = \Phi_n - \Phi$ obey the equation

$$k \frac{d}{dk} \Delta_{n+1}(k) + \delta\, \Delta_{n+1}(k) = \delta \cos(kw)\, \Delta_n(k).$$

Integration, considering the cases $k > 0$ and $k < 0$ separately, yields

$$\Delta_{n+1}(k) = \frac{1}{|k|^\delta} \int_{\min\{0,k\}}^{\max\{0,k\}} \delta\, |m|^{\delta-1} \cos(wm)\, \Delta_n(m)\, dm. \tag{26}$$

Now, $\Delta_0(0) = 0$ and Δ_0 is twice differentiable and bounded, and therefore there exists a constant C, such that $|\Delta_0(k)| \leqslant C\,|k|$ for each real k. Induction in n using (26) shows that, for each k,

$$|\Delta_n(k)| \leqslant \left(\frac{\delta}{1+\delta}\right)^n C\,|k|,$$

and the convergence of Φ_n to Φ follows. The remaining statements follow from this convergence—see [1]. ▮

As a direct consequence (see [1, Theorem 25.8]) of Lemma 3.3 we have the following ergodic theorem.

THEOREM 3.4. *Consider the Markov chain* (3) *with semigroup* $S(\cdot,\cdot)$ *generated by* (5) *and with initial data distributed according to a probability measure* μ *with*

$$\int_{\mathbb{R}} u^2 \mu(du) < \infty.$$

Then for all continuous bounded real-valued functions g,

$$\lim_{n\to\infty} \mathbb{E}^{\mu} g(u_n) = \lim_{n\to\infty} \int_{\mathbb{R}} g(u)\,\mu_n(du) = \int_{\mathbb{R}} g(u)\,\nu(du),$$

where ν *is the invariant probability measure with characteristic function* (24) *and* μ_n *is the distribution of* u_n.

We conclude this section with some remarks. The integrand in (24) is an analytic function of m; hence all moments of ν exist. On the other hand, as $|k| \to \infty$, the integral in (24) diverges logarithmically, implying that $|\Phi(k)|$ only decays as a negative power of $|k|$ and that ν cannot be very smooth. In fact, from (21) the equation for the density ϕ of ν is

$$\frac{d\phi(u)}{du} = (\delta - 1)\frac{\phi(u)}{u} - \left(\frac{\delta}{2}\right)\frac{\phi(u+w)+\phi(u-w)}{u},$$

with $\delta = \lambda/\gamma$ (note that δ is nondimensional, as both $1/\lambda$ and $1/\gamma$ possess the dimension of t). The behaviour of ϕ near the origin is easily investigated by standard asymptotic analysis. A lack of smoothness is present, but the regularity improves as δ increases, i.e. as the ratio of noise to dissipation increases (in the absence of noise the invariant probability measure is a point mass at the origin and ϕ does not exist). If $\delta \in (0, 1)$ then, for $|u| \ll 1$,

$$\phi(u) \approx C\,|u|^{\delta-1}, \qquad \delta - 1 \in (-1, 0).$$

If $\delta = 1$ then, for $|u| \ll 1$,

$$\phi(u) \approx -\tfrac{1}{2}\,[\phi(w) + \phi(-w)]\ln(|u|).$$

If $\delta \in (1, 2)$ then, for $|u| \ll 1$,

$$\phi(u) \approx \frac{\delta}{2(\delta-1)} [\phi(w) + \phi(-w)] + C |u|^{\delta-1}, \qquad \delta - 1 \in (0, 1).$$

The ergodicity result of Theorem 3.4 can be extended to sample path averages by employing the techniques of the next section. We do not give the details.

4. A DAMPED NONLINEAR OSCILLATOR

In this section we invoke the theory of Markov chains in uncountable state spaces developed in [11] (see also [12, 16]) to study ergodicity of the Markov chain (3) generated by Eqs. (11), (18).

4.1. *The Deterministic Problem*

We first need to prove some results concerning the behaviour of the oscillator in the absence of impulses. The corresponding phase portrait is presented in Fig. 1.

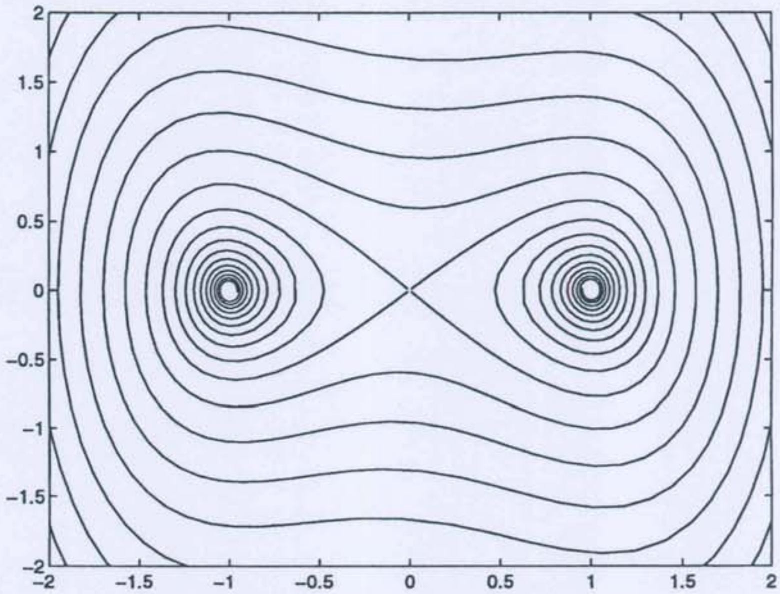

FIG. 1. The phase portrait of the deterministic double-well problem. Horizontal axis is q, vertical axis is q'.

Let q_0 be a fixed value of q and consider the half-line

$$\mathscr{L}^+ := \{(q, p)^T : q = q_0, p > 0\}.$$

If $u \in \mathscr{L}^+$ and u is not on the unstable manifold $\mathscr{A}$ of the origin, then the backwards trajectory $S(u, -t)$, $t > 0$ that starts at u meets $\mathscr{L}^+$ again. We may therefore define a Poincaré return map Ξ: $\mathscr{L}^+ \cap \mathscr{A}^c \to \mathscr{L}^+ \cap \mathscr{A}^c$ such that for each $u \in \mathscr{L}^+ \cap \mathscr{A}^c$, $\Xi(u) = S(u, -\bar{t})$ with $\bar{t} = \bar{t}(u) = \min\{t > 0 : S(u, -t) \in \mathscr{L}^+\}$. By identifying each point $u = (q_0, p)^T$ on $\mathscr{L}^+$ with the corresponding value of p, we may, if convenient, see Ξ as a real-valued function of the real variable p. It may be shown that Ξ is smooth in view of the transversality of $\mathscr{L}^+$ to the flow. The next result gives the behaviour of Ξ for large p.

LEMMA 4.1. *There exists a constant* $\beta > 0$ *such that the Poincaré mapping* Ξ *defined above satisfies*

$$\Xi(p) = p + \beta\sqrt{p} + r(p), \tag{27}$$

with

$$r(p) = O(1), \qquad \frac{dr}{dp} = O\left(\frac{1}{p\sqrt{p}}\right), \qquad p \to \infty.$$

Proof. We solve (11), (18) with initial condition $q(0) = q_0$, $p(0) = 1/\varepsilon^2$, ε small and non-zero. By integrating (12), we have

$$\Xi(p(0))^2 - p(0)^2 = 2\gamma \int_{-\bar{t}}^{0} p(\tau)^2 \, d\tau, \tag{28}$$

where $\bar{t} = \bar{t}(p(0)) > 0$ denotes the return time from $p(0)$ as defined above.

After introducing the new time $s = t/\varepsilon$ and the new function $Q(s) = \varepsilon q(\varepsilon s)$, the problem (11), (18) becomes

$$\frac{d^2Q}{ds^2} + \varepsilon\gamma \frac{dQ}{ds} - \varepsilon^2 Q + Q^3 = 0, \qquad Q(0) = \varepsilon q_0, \qquad \frac{dQ}{ds}(0) = 1 \tag{29}$$

and (28) can be written as

$$\Xi(p(0))^2 - p(0)^2 = \frac{R(\varepsilon)}{\varepsilon^3} \tag{30}$$

with

$$R(\varepsilon) := 2\gamma \int_{-\bar{s}(\varepsilon)}^{0} \left| \frac{dQ}{ds}(\sigma, \varepsilon) \right|^2 d\sigma.$$

Here $\bar{s}(\varepsilon) := \bar{t}(p(0))/\varepsilon = \bar{t}(\varepsilon^{-2})/\varepsilon$ coincides with the return time for the backward time Poincaré section for (29) with the half-line $Q = \varepsilon q_0$, $dQ/ds > 0$.

In the limit $\varepsilon \to 0$, we obtain from (29) the reduced problem

$$\frac{d^2 Q_0}{ds^2} + Q_0^3 = 0, \qquad Q_0(0) = 0, \qquad \frac{dQ_0}{ds}(0) = 1; \tag{31}$$

this, since the level curves of the resulting Hamiltonian are closed curves, describes an undamped oscillation. Clearly $\bar{s}(\varepsilon)$ is a smooth function of ε up to $\varepsilon = 0$, with $\bar{s}(0) > 0$ given by the period of the solution of (31). Furthermore $R(\varepsilon)$ is a smooth function of ε which, at $\varepsilon = 0$, has the value

$$R(0) = 2\gamma \int_{-\bar{s}(0)}^{0} \left| \frac{dQ_0}{ds}(\sigma) \right|^2 d\sigma > 0.$$

Upon rearranging (30), we obtain

$$\Xi\left(\frac{1}{\varepsilon^2}\right) = \frac{1}{\varepsilon^2} \sqrt{1 + \varepsilon R(\varepsilon)}. \tag{32}$$

Expansion of the square root leads to

$$\Xi\left(\frac{1}{\varepsilon^2}\right) = \frac{1}{\varepsilon^2} + \frac{1}{2\varepsilon} R(0) + O(1), \qquad \varepsilon \to 0,$$

which proves the estimate for r. The estimate for dr/dp follows similarly after differentiation in (32). ∎

From Lemma 4.1 (or by inspection of the phase portrait), it is clear that, for any $p > 0$ for which $\Xi(p)$ is defined, the sequence of iterates $\Xi(p)$, $\Xi^2(p)$, ... grows to ∞. Actually the growth with n of $\Xi^n(p)$ is $O(n^2)$ as we show next.

LEMMA 4.2. *For any p such that $(q_0, p)^T \in \mathscr{L}^+ \cap \mathscr{A}^c$ there exist positive constants C_1 and C_2 such that, for all $n \in \mathbb{N}$,*

$$\Xi^n(p) \leqslant C_1 n^2 + C_2.$$

Proof. By induction in n using (27). ∎

The next result will be crucial in later developments.

LEMMA 4.3. *Let I be a segment of positive length contained in $\mathscr{L}^+ \cap \mathscr{A}^c$. The the successive images $I_n = \Xi^n(I)$, $n \in \mathbb{N}$, possess lengths that tend to ∞ as $n \to \infty$.*

Proof. We may write, with $|\cdot|$ denoting length,

$$\frac{|I_n|}{|I|} = \Pi_{i=0}^{n-1} \frac{|I_{i+1}|}{|I_i|},$$

so that our task is to show that the infinite product $\Pi_{i=0}^{\infty}(|I_{i+1}|/|I_i|)$ diverges to ∞. It suffices to show that, if $c_i := (|I_{i+1}|/|I_i|) - 1$, then the series $\sum_{i=0}^{\infty} c_i$ diverges. Now, using (27),

$$\begin{aligned} c_i &\geqslant \min\{\Xi'(p) : p \in I_i\} - 1 \\ &= \min\left\{\frac{\beta}{2\sqrt{p}} + r'(p) : p \in I_i\right\}. \end{aligned}$$

Since $\Xi^i(p_0) \to \infty$ as $i \to \infty$ we may neglect the contribution of $r'(p)$, $p \in I_i$. For $p \in I_i$ the minimum of $\beta/(2\sqrt{p})$ occurs at the upper end p_i of I_i. From the preceding lemma, $p_i = \Xi^i(p_0)$ is $O(i^2)$ and therefore $1/\sqrt{p_i}$ possesses an $O(1/i)$ lower bound and the divergence of $\sum_{i=0}^{\infty} c_i$ follows. ∎

4.2. *Irreducibility*

As in [11], we shall work hereafter with the transition probabilities p^n of the Markov chain (3) rather than with the probability distributions μ_n of u_n given an initial distribution μ for u_0.

The crucial step in establishing ergodicity is to exhibit an *irreducibility* measure ϕ for the chain, i.e., a nontrivial Borel measure on $\mathbb{R}^m$ such that, for any $u \in \mathbb{R}^m$ and any Borel set A with $\phi(A) > 0$, there exists an $n \in \mathbb{Z}^+$ for which $p^n(u, A) > 0$ (in other words every set of positive measure can be reached from every initial condition with positive probability after a finite number of steps). Irreducibility ensures that the chain is not a juxtaposition of two chains each acting on disjoint subsets of the state space.

If a given chain possesses an irreducibility measure ϕ, then it possesses many other irreducibility measures ϕ'. Among these measures there is ([11], Proposition 4.2.2) an essentially unique *maximal irreducibility measure* ψ, such that any measure ϕ' is an irreducibility measure if and only if ϕ' is absolutely continuous with respect to ψ. In practice it is convenient to show irreducibility by exhibiting an "easy" irreducibility measure ϕ, rather than finding ψ directly. This is the approach we take here.

We construct ϕ as follows. As in the preceding subsection, we consider the half-line $\mathscr{L}^+$ where $q = q_0 := 1/2$, $p > 0$ (the exact value of q_0 is immaterial as long as q_0 does not coincide with any of the equilibria 0, ± 1). On $\mathscr{L}^+$ we fix a small interval I; for definiteness let us take

$$I := \{u = (q_0, p) : q_0 = \tfrac{1}{2}, 1 \leqslant p \leqslant 1 + \delta\},$$

with $0 < \delta < 1$ (again, the range $1 \leqslant p \leqslant 1 + \delta$ could be replaced by any other range $p_0 \leqslant p \leqslant p_0 + \delta$, provided that $p_0 > 0$ is not too small). For each point $u \in I$ we carry out the following process:

(i) We move u upwards in the $(q, p)^T$ plane by giving it an impulse w, i.e., we transform u into $u + w$.

(ii) From $u + w$, we form the flow arc $A_u = \{S(u + w, s) : s \in [0, \Delta]\}$, where Δ is a small positive number.

(iii) We move A_u downwards by giving it a negative impulse. This results in the translated arc $A_u - w$.

(iv) From each point $\bar{u} \in A_u - w$ we form the arc $\{S(\bar{u}, t) : t \in [0, \Delta]\}$

For each $u \in I$, the collection of the arcs in (iv), i.e., the set

$$P_u = \bigcup_{\bar{u} \in A_u - w} \{S(\bar{u}, t) : t \in [0, \Delta]\},$$

coincides with the image set of the mapping

$$\Phi_u(s, t) = S(S(u + w, s) - w, t),$$

defined for $(s, t) \in [0, \Delta]^2$. From the geometric construction above (or by differentiation of Φ_u)

$$\frac{\partial \Phi_u}{\partial s}(0, 0) = f(u + w), \qquad \frac{\partial \Phi_u}{\partial t}(0, 0) = f(u);$$

here f is the vector field in our problem (1), (11), (18). The vectors $f(u + w)$ and $f(u)$ are not parallel because $\nabla V(q_0) \neq 0$ and hence, for Δ sufficiently small and each fixed $u \in I$, Φ_u is a diffeomorphism onto its image P_u. Thus P_u is a small curvilinear parallelogram with a vertex at u. Note that P_u depends continuously on u and, by reducing δ if necessary, we can ensure that the intersection

$$\bigcap_{u \in I} P_u \tag{33}$$

has a nonempty interior. We choose a closed ball B contained in the interior of the intersection (33) and then the measure ϕ is defined to be Lebesgue measure restricted to B.

THEOREM 4.4. *The measure ϕ constructed above is an irreducibility measure for the Markov chain* (3) *originating from the double-well oscillator* (11), (18).

Proof. Before we go into the proof, we determine $\bar{n} \in \mathbb{N}$ such that $J := \Xi^{\bar{n}}(I) \subset \mathscr{L}^+$ has length $> v$. This is possible by Lemma 4.3; note that I does not intersect the unstable manifold of the origin because (i) for the origin the energy E in (12) equals 1/4, (ii) points in I have $E > 1/2$, and (iii) E decreases along trajectories.

We fix an arbitrary initial state u^* and prove that it is possible to find an integer $N = N(u^*)$ such that $p^N(u^*, A) > 0$ for each set A with $\phi(A) > 0$. In order to determine N, we need some quantities depending on the chosen u^*:

(a) We first determine an integer $\ell \in \mathbb{N}$ such that the positive orbit $\{S(u^* + \ell w, t): t \geqslant 0\}$ starting at $u^* + \ell w$ intersects $\mathscr{L}^+$ at a point $u^{**} = S(u^* + \ell w, T_1)$ above J (i.e., at a point u^{**} for which the value of the momentum p^* is greater than or equal to the value of p for the points in J). The possibility of choosing such an ℓ follows from the geometry of the flow of the deterministic problem (11), (18).

(b) Next we determine $m \in \mathbb{N}$ such that $u^{**} - mw \in J$. This is possible because both u^{**} and J lie on $\mathscr{L}^+$, u^{**} is above J and J has length $> v$.

(c) Then we determine $T_2 \geqslant 0$ such that $u^{***} := S(u^{**} - mw, T_2) \in I$. This is possible because $u^{**} - mw \in J = \Xi^{\bar{n}}(I)$.

Once we have found the values of ℓ, T_1, m, T_2, we set $N := \ell + m + 2$ and introduce the space $\mathbb{R}_+^{\ell+m}$ of vectors $\varepsilon \in \mathbb{R}^{\ell+m}$, $\varepsilon = (\varepsilon_0, ..., \varepsilon_{\ell+m-1})^T$, $\varepsilon_j \geqslant 0$, $j = 0, ..., \ell + m - 1$. Let $\omega_j = (\theta_j, t_j)$, with θ_j and t_j the random sign and random waiting time in (1) and consider the events

$$\begin{aligned} \mathscr{G}_\varepsilon = \{ z \in D_E \colon & \omega_j = (1, \varepsilon_j)^T, j = 0, ..., \ell - 2, \\ & \omega_{\ell-1} = (1, T_1 + \varepsilon_{\ell-1})^T, \\ & \omega_{\ell+j} = (-1, \varepsilon_{\ell+j})^T, j = 0, ..., m-2, \\ & \omega_{\ell+m-1} = (-1, T_2 + \varepsilon_{\ell+m-1})^T \} \end{aligned}$$

(for $\ell = 1$ or $m = 1$, the sets $\{j = 0, ..., \ell - 2\}$ or $\{j = 0, ..., m-2\}$ are understood to be empty) then clearly

$$\begin{aligned} & p^T(u^*, A) \\ & \quad \geqslant \int_{\mathbb{R}_+^{\ell+m}} \Pi_\varepsilon^N(u^*, A) \frac{\lambda^{\ell+m}}{2^{\ell+m}} e^{-\lambda\varepsilon_0} \cdots e^{-\lambda(T_1 + \varepsilon_{\ell-1})} e^{-\lambda\varepsilon_l} \cdots e^{-\lambda(T_2 + \varepsilon_{\ell+m-1})} \, d\varepsilon, \end{aligned}$$

where $\Pi_\varepsilon^N(u^*, A)$ denotes the probability of the event $\{u_N \in A\}$, conditioned to $\mathscr{G}_\varepsilon$, when the chain starts from $u_0 = u^*$. If we prove that, for ε in a neighbourhood of 0 in $\mathbb{R}_+^{\ell+m}$ and $\phi(A) > 0$, we have $\Pi_\varepsilon^N(u^*, A) > 0$, then the proof will be complete.

In order to prove that, for $\|\varepsilon\|$ small and $\phi(A)>0$, $\Pi_\varepsilon^N(u^*, A)>0$, it is clearly enough to show, for $\|\varepsilon\|$ small and $\phi(A)>0$, the positivity of the smaller quantity $p_\varepsilon^N(u^*, A)$ defined as the probability, when $u_0=u^*$, of the event $\{u_N \in A\} \cap Z$ conditioned to $\mathscr{G}_\varepsilon$, with

$$Z := \{\omega_{\ell+m} \in (1, [0, \Delta])^T, \omega_{\ell+m+1} \in (-1, [0, \Delta])^T\}.$$

We first show that $p_0^N(u^*, A)>0$, if $\phi(A)>0$. When the event $\mathscr{G}_0$ takes place, the initial state u^* undergoes ℓ consecutive upwards impulses, $u^* \to u^* + \ell w$. These are followed by flowing for T_1 units of time to reach the point u^{**} defined above. Then m downward impulses occur, followed by flowing for T_2 units of time. This leads *surely* to the point $u^{***} \in I$ after $\ell+m$ transitions. Now the event Z has a positive probability independent of $\mathscr{G}_0$ (namely $(1/4)(1-\exp(-\lambda\Delta))^2$). When Z takes place, u^{***} undergoes an upwards impulse, followed by flowing for $s \in [0, \Delta]$ units of time, followed by a downwards impulse and flowing for $t \in [0, \Delta]$ units of time. Hence, if $\mathscr{G}_0$ and Z are true, the initial state u^* moves after N transitions of the chain into a point of the curvilinear parallelogram $P_{u^{***}}$ and every point in $P_{u^{***}}$ can be reached in this way. We therefore conclude that a constant $K=K(u^*)>0$ exists, such that for each Borel set $E \subset P_{u^{***}}$,

$$p_0^N(u^*, E) \geqslant K\phi_{\mathbb{R}^2 Leb}(E),$$

where $\phi_{\mathbb{R}^2 Leb}(E)$ is the standard Lebesgue measure in the plane. In particular $p_0^N(u^*, A)>0$ for each $A \subset B$ with $\phi_{\mathbb{R}^2 Leb}(A)>0$, i.e., for each A with $\phi(A)>0$.

By the continuity of the deterministic semigroup $S(u, t)$ with respect to both of its arguments, it is clear that, when the events Z and $\mathscr{G}_\varepsilon$, $\|\varepsilon\|$ small, occur, the point u^* moves in N transitions to a point in a slightly deformed version of $P_{u^{***}}$ that still contains the support B of ϕ. Then, for $\phi(A)>0$, $p_\varepsilon^N(u^*, A)>0$ for $\|\varepsilon\|$ small. In fact for ε_c sufficiently small, and possibly by redefinition of K,

$$p_\varepsilon^N(u^*, E) \geqslant K\phi_{\mathbb{R}^2 Leb}(E),$$

for all $\varepsilon \in [0, \varepsilon_c]$. ∎

Remark. It is perhaps useful to point out that, for this theorem to hold, it is not essential to assume that the waiting times t_n are exponentially distributed. The proof only uses the fact that the density of the waiting times, with respect to Lebesgue measure, exists and is positive.

Before closing this subsection, it is convenient to deal with some technical points. Given a Markov chain on the state space $\mathbb{R}^m$, a Borel set C is called *a small set* [11, Section 5.2], if there exists $m \in \mathbb{Z}^+$ and a non-trivial Borel measure ν such that, for all $u \in C$ and all Borel sets A,

$$p^m(u, A) \geqslant \nu(A).$$

A Borel set C is a *petite set* if, for all $u \in C$ and all Borel sets A,

$$\sum_{n=0}^{\infty} a(n)\, p^n(u, A) \geqslant \nu_a(A),$$

where $a = \{a(n)\}$ is a probability measure on $\mathbb{Z}^+$ and ν_a a nontrivial Borel measure.

The definitions of small sets and petite sets try to encapsulate the idea of sets that are probilistically not large. Note that small sets are always petite: just take a to be a point mass.

LEMMA 4.5. *For the Markov chain* (3) *for the double-well oscillator* (11), (18), *every compact set in* $\mathbb{R}^*$ *is petite.*

Proof. We reconsider the proof of the Theorem 4.4. By continuity arguments similar to those used at the end of the proof, given a point $u^* \in \mathbb{R}^m$, there exist positive constants K', r such that, for the integer $N = N(u^*)$ we determined there,

$$p^N(u, A) \geqslant K' \phi_{\mathbb{R}^2 Leb}(A) \tag{34}$$

for each Borel set $A \subset B$ and each u in the open ball $B(u^*, r)$ of (sufficiently small) radius r and centre u^*. By definition $B(u^*, r)$ is then a small set and hence petite. By choosing u^* in B we ensure that $B(u^*, r)$ has $\phi_{\mathbb{R}^2 Leb}(B(u^*, r)) > 0$ and, *a fortiori*, $\psi(B(u^*, r)) > 0$ (recall that ψ is the maximal irreducibility measure). The result is then a direct consequence of Proposition 6.2.8 in [11]. ∎

4.3. *Aperiodicity*

Given an irreducible Markov chain with maximal irreducibility measure ψ, the sets $\{D_i\}_{i=1}^d$ are said to form a *d-cycle* if (i) they are disjoint, (ii) for $x \in D_i$, $p^1(x, D_{i+1}) = 1$, $i = 1, ..., d \pmod d$, (iii) the complement $[\bigcup_{i=1}^d D_i]^c$ is ψ-null. The chain is *aperiodic* if the maximal such d is 1.

THEOREM 4.6. *The irreducible Markov chain* (3) *for the double-oscillator* (11), (18) *is aperiodic.*

Proof. We again reconsider the proof of Theorem 4.4. If in that proof we choose the interval J to have length $> 2v$ (rather than merely to have length $> v$), then, in point (b) of the proof, it is possible to have both

$u-mu^{**}$ and $u-(m+1)u^{**}$ in J. Then (34) holds for $N=\ell+m+2$ and $N=\ell+m+3$, two coprime integers, and aperiodicity follows from Theorem 5.4.4 in [11]. ∎

4.4. *Drift*

The last property of the chain to be discussed before proving ergodicity is the probabilistic drift towards a petite set. We work with the function

$$\bar{G}(u)=1+G(u), \tag{35}$$

with G as defined in (14). Note that $\bar{G}(u)\geqslant 1$ for all u by (15) and also that the sets $\{u:\bar{G}\leqslant n\}$, $n\in\mathbb{N}$ are compact and therefore (Lemma 4.5) petite (in the terminology of [11], $\bar{G}$ is unbounded off petite sets). We shall prove that, for our Markov chain and outside a petite set C, the value of $\bar{G}$ decays geometrically at each step (see Condition (V4) in [11, Chapter 15]).

LEMMA 4.7. *For the Markov chain* (3) *generated by the double-well oscillator* (11), (18), *there exists a petite set C and constants $a>0$ and $b\in\mathbb{R}$, such that, for the function $\bar{G}$ in* (35), *all $u\in\mathbb{R}^m$ and all $n\in\mathbb{Z}^+$,*

$$\mathbb{E}(\bar{G}(u_{n+1})-\bar{G}(u_n)\,|\,u_n=u)\leqslant -a\bar{G}(u)+bI\{u\in C\}. \tag{36}$$

Proof. By (19),

$$\mathbb{E}(G(u_{n+1}\,|\,u_n=u)\leqslant a_1G(u)+a_2, \qquad a_1\in(0,1).$$

From here

$$\mathbb{E}(\bar{G}(u_{n+1})\,|\,u_n=u)\leqslant a_1\bar{G}(u)+(1-a_1+a_2),$$

and we can apply Lemma 15.2.8 in [11] to obtain (36). ∎

4.5. *Ergodicity*

We are ready to present our main result.

THEOREM 4.8. *For the Markov chain* (3) *generated by the double-well oscillator* (11), (18), *there exists a unique invariant probability measure π, a constant $R<\infty$ and a uniform rate of decay $\rho<1$ such that, for all Borel-measurable functions f with $|f|\leqslant\bar{G}$ and all $u^*\in\mathbb{R}^m$,*

$$\left|\int_{\mathbb{R}^m}(p^n(u^*,du)-\pi(du))\,f(u)\right|\leqslant R\bar{G}(u^*)\,\rho^n, \qquad n\in\mathbb{Z}^+, \tag{37}$$

where $\bar{G}$ is the function in (35).

Proof. This result is a direct application of Theorem 16.0.1 in [11] because the chain is irreducible, aperiodic and possesses a geometric drift. ∎

Note that (37) implies in particular that

$$\lim_{n \to \infty} \int_{\mathbb{R}^m} p^n(u^*, du)\, f(u) = \int_{\mathbb{R}^m} \pi(du)\, f(u)$$

for every bounded measruable function f (even if f is not continuous). It is also possible to prove the following sample path ergodic theorem:

THEOREM 4.9. *Consider the Markov chain* (3) *generated by the double-well oscillator* (11), (18). *Then for any* $g \in \mathscr{L}^1(\pi)$,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} g(u_k) = \int_{\mathbb{R}^m} g(u)\, \pi(du), \qquad \mathbb{P}^{\delta_x}\text{-a.s.}, \quad \forall x \in \mathbb{R}^2.$$

Proof. The Markov chain is Harris recurrent by Theorem 9.1.8 in [11] because Theorem 2.3 shows that, if we define the set

$$C_\zeta = \{ y: G(y) \leqslant \zeta \},$$

then

$$\mathbb{E}(G(u_{i+1}) \mid u_i = x) \leqslant G(x), \qquad \forall x \in C_\zeta^c$$

for $\zeta \geqslant (\gamma\alpha + 2\eta)/(\gamma\beta)$. Theorem 17.1.7 of [11] proves the sample path result. ∎

ACKNOWLEDGMENTS

We thank M. San Miguel and C. Simó for giving us useful references. We are indebted to C. Matrán for many enlightening discussions.

REFERENCES

1. P. Billingsley, "Probability and Measure," 2nd ed., Wiley, New York, 1996.
2. S. R. Bishop (Ed.), Impact oscillators, *Philos. Trans. Roy. Soc.* **347** (1994), 345–448.
3. M. H. A. Davis, Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models, *J. Roy. Statist. Soc. Ser. B* (1984), 353–388.
4. M. H. A. Davis, "Markov Models and Optimization," Monographs on Statistics and Applied Probability, Vol. 49, Chapman and Hall, London, 1993.
5. C. W. Gardiner, "Handbook of Stochastic Methods," Springer-Verlag, New York, 1983.
6. J. K. Hale, "Asymptotic Behaviour of Dissipative Systems," AMS Mathematical Surveys and Mongraphs, Vol. 25, Am. Math. Soc., Providence, RI, 1988.

7. E. Hernández-García, L. Pesquera, M. A. Rodríguez, and M. San Miguel, First passage time statistics: Processes driven by Poisson noise, *Phys. Rev. A* **36** (1987), 5774–5781.
8. E. Hernández-García, L. Pesquera, M. A. Rodríguez, and M. San Miguel, Random walk in dynamically disordered chains: Poisson with noise disorder, *J. Statist. Phys.* **55** (1989), 1027–1052.
9. T. M. Liggett, "Interacting Particle Systems," Springer-Verlag, New York, 1985.
10. G. A. Mersov, S. S. Moiseev, and Y. M. Nezlina, A method of investigation of the evolution of a nonlinear system driven by pulsed noise, *Chaos* **3** (1993), 177–182.
11. S. P. Meyn and R. L. Tweedie, "Markov Chains and Stochastic Stability," 2nd ed., Springer-Verlag, London, 1993.
12. E. Nummelin, "General Irreducible Markov Chains and Non-negative Operators," Cambridge Univ. Press, Cambridge, UK, 1984.
13. J. Olarrea, J. M. R. Parrondo, and F. J. de la Rubia, Escape statistics for systems driven by dichotomous noise, I, General theory; II, The imperfect pitchfork bifurcation as a case study, *J. Statist. Phys.* **79** (1995), 669–699.
14. Z. Schuss, "Theory and Applications of Stochastic Differential Equations," Wiley, New York, 1980.
15. R. Temam, "Infinite Dimensional Dynamical Systems in Mechanics and Physics," Springer, New York, 1989.
16. R. L. Tweedie, Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space, *Stochastic Process. Appl.* **3** (1975), 385–403.
17. N. G. Van Kampen, Processes with delta-correlated cumulants, *Physica A* **102** (1980), 489–495.
18. C. Van Den Broeck, On the relation between white shot noise, Gaussian white noise and the dichotomic Markov process, *J. Statist. Phys.* **31** (1983), 467–483.

Printed by Catherine Press, Ltd., Tempelhof 41, B-8000 Brugge, Belgium