



## Metastability of the Nonlinear Wave Equation: Insights from Transition State Theory

Katherine A. Newhall<sup>1</sup> · Eric Vanden-Eijnden<sup>2</sup>

Received: 3 August 2016 / Accepted: 22 December 2016 / Published online: 3 January 2017  
© Springer Science+Business Media New York 2017

**Abstract** This paper is concerned with the longtime dynamics of the nonlinear wave equation in one-space dimension,

$$u_{tt} - \kappa^2 u_{xx} + V'(u) = 0 \quad x \in [0, 1]$$

where  $\kappa > 0$  is a parameter and  $V(u)$  is a potential bounded from below and growing at least like  $u^2$  as  $|u| \rightarrow \infty$ . Infinite energy solutions of this equation preserve a natural Gibbsian invariant measure, and when the potential is double-welled, for example when  $V(u) = \frac{1}{4}(1 - u^2)^2$ , there is a regime such that two small disjoint sets in the system's phase-space concentrate most of the mass of this measure. This suggests that the solutions to the nonlinear wave equation can be metastable over these sets, in the sense that they spend long periods of time in these sets and only rarely transition between them. Here, we quantify this phenomenon by calculating exactly via transition state theory (TST) the mean frequency at which the solutions of the nonlinear wave equation with initial conditions drawn from its invariant measure cross a dividing surface lying in between the metastable sets. We also investigate numerically how the mean TST frequency compares to the rate at which a typical solution crosses this dividing surface. These numerical results suggest that the dynamics of the nonlinear wave equation is ergodic and rapidly mixing with respect to the Gibbs invariant measure when the parameter  $\kappa$  is small enough. In this case, successive transitions between the two regions are roughly uncorrelated and their dynamics can be coarse-grained to

---

Communicated by Charles R. Doering.

---

✉ Katherine A. Newhall  
knewhall@unc.edu

<sup>1</sup> Department of Mathematics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

<sup>2</sup> Courant Institute of Mathematical Sciences, New York University, New York, NY, USA

jumps in a two-state Markov chain whose rate can be deduced from the mean TST frequency. This is a regime in which the dynamics of the nonlinear wave equation displays a metastable behavior that is not fundamentally different from that observed in its stochastic counterpart in which random noise and damping terms are added to the equation. For larger  $\kappa$ , however, the dynamics either stops being ergodic, or its mixing time becomes larger than the inverse of the TST frequency, indicating that successive transitions between the metastable sets are correlated and the coarse-graining to a Markov chain fails.

**Keywords** Transition state theory · Metastability · Stochastic partial differential equation · Effective dynamics · Nonlinear wave equation

**Mathematics Subject Classification** 37A60 · 60H15 · 82B05

## 1 Introduction

Metastability is a feature commonly displayed by dynamical systems (Bovier 2015). One of its simplest manifestations is in the context of finite-dimensional systems whose evolution is governed by the so-called overdamped Langevin equation, that is, steepest descent over a given energy subject to thermal fluctuations modeled by an additive white-noise driving term. In such systems, if the energy landscape possesses several local minima separated by barriers much higher than the available thermal energy, with high probability the solution will spend a long time in a basin associated with the energy minimum before making a transition to another such basin. The timescale of these transitions follows the Arrhenius law, i.e., it is exponentially large in the energy barrier the system must surmount to escape the basin, measured in units of temperature. This prediction can be confirmed rigorously using Freidlin–Wentzell large deviation theory (Freidlin and Wentzell 1994), which also predicts the mechanism of the transitions. More refined estimates for the transition time that include its prefactor can be obtained using, e.g., the potential theoretic approach to metastability introduced by Bovier and collaborators (Bovier 2015; Bovier et al. 2001, 2002, 2004), and extended by many others (Barret 2015; Huisinga et al. 2004; Vanden-Eijnden and Westdickenberg 2008).

These results can be extended to infinite-dimensional dynamical systems, e.g., the stochastic Allen–Cahn equation (Berglund and Gentz 2013; Faris and Jona-Lasinio 1982; Vanden-Eijnden and Westdickenberg 2008). They can also be extended to other types of dynamics, e.g., if inertial effects are included. In this case, the system's evolution can be modeled by the (inertial) Langevin equation, that is, Hamilton's equations in which a linear damping term and a white-noise forcing term are added in the equation for the momenta. More complicated is the case of Hamiltonian systems in which there is no damping nor thermal noise, since the dynamics of such systems is conservative and deterministic. Still, metastability can arise in this context as well. This is most easily pictured in finite-dimensional systems, assuming that the constant energy surface the system is evolving upon consists of two large regions connected by a narrow bottleneck: This arises, e.g., if the total energy of the system is only slightly

higher than the potential energy barrier between two regions in position space. If the dynamics is ergodic, the system can and will eventually cross this bottleneck, but in order to do so it must give up most of its kinetic energy to overcome the potential energy barrier, which is highly unlikely to happen. This picture is intuitively appealing, though it may be hard to justify rigorously since ergodicity is typically very hard to prove. On top of this, it also becomes harder to use and justify as the dimensionality of the system under consideration increases. Indeed, energy typically is an extensive property that grows with the system’s size, meaning that large systems always have much more energy than what is needed to cross a given potential energy barrier. In particular, the picture above seems to become moot for infinite-dimensional Hamiltonian systems, since their total energy can be infinite. The purpose of the present paper is to investigate metastability in the context of one such system.

Specifically, we will study the nonlinear one-dimensional wave equation for  $u(x, t)$

$$u_{tt} - \kappa^2 u_{xx} + V'(u) = 0 \tag{1}$$

for  $x \in [0, 1]$ , with either Dirichlet or Neumann boundary conditions at  $x = 0$  and  $x = 1$ , and  $V(u) > Cu^2$  for some  $C > 0$  as  $|u| \rightarrow \infty$ . In (1),  $\kappa$  measures the typical spatial scale over which  $u(x, t)$  varies: This parameter can be absorbed in the domain size after rescaling via  $\xi = x/\kappa \in [0, 1/\kappa]$ . Defining  $p(x, t) = u_t(x, t)$ , we can write (1) as

$$u_t = p, \quad p_t = \kappa^2 u_{xx} - V'(u) \tag{2}$$

which are Hamilton’s equations associated with the infinite-dimensional Hamiltonian

$$H(p, u) = T(p) + U(u), \tag{3}$$

where

$$T(p) = \int_0^1 \frac{1}{2} |p(x)|^2 dx, \quad U(u) = \int_0^1 \frac{1}{2} \kappa^2 |u_x(x)|^2 + V(u(x)) dx. \tag{4}$$

We will consider (2) in a parameter regime in which the energy (3) is infinite (in which case we will have to consider weak solutions of (2)—more on this below). In this regime, Eq. (2) possesses a single natural invariant measure which coincides with the canonical (Gibbs) measure of the associated Langevin system, i.e., the stochastic partial differential equation (SPDE) formally given by

$$u_t = p, \quad p_t = \kappa^2 u_{xx} - V'(u) - \gamma p + \sqrt{2\gamma\beta^{-1}} \eta(x, t), \tag{5}$$

in which  $\eta(x, t)$  is a space-time white noise, i.e., the Gaussian process with mean zero and covariance  $\mathbb{E}(\eta(x, t)\eta(y, s)) = \delta(x - y)\delta(t - s)$ ,  $\gamma > 0$  is the friction coefficient, and  $\beta > 0$  is the inverse temperature. The Gibbs measure for (5) can be written

formally and in an unnormalized way using the notations of McKean and Vaninsky (1999) as

$$d\mu = \exp(-\beta H(p, u)) d^\infty p d^\infty u. \quad (6)$$

The precise meaning of this measure will be explained in Sect. 3. The fact that (6) is the only natural invariant measure for (2) is intriguing: Indeed, in finite dimensions, Hamiltonian systems possess two different classes of invariant measures, the microcanonical and the canonical ones, and if their dynamics are ergodic, it can only be so with respect to the former since their evolution is energy-conserving (Sinai 1976). What therefore happens to the microcanonical measures in the context of the wave equation? The equivalence of invariant measure for the energy-conserving wave equation in (2) and its Langevin counterpart in (5) also raises another question: Do other features of their dynamics coincide, in particular those related to metastability?

We will address both these questions here. Regarding the first, we will consider finite-dimensional truncations of the wave equation based on the Hamiltonian defined as

$$H_N(\mathbf{p}, \mathbf{u}) = \frac{1}{N} \sum_{j=1}^N \left( \frac{1}{2} p_j^2 + V(u_j) \right) + N \sum_{j=0}^N \frac{1}{2} \kappa^2 (u_{j+1} - u_j)^2, \quad (7)$$

where  $u_j = u(x_j, t)$  and  $p_j = p(x_j, t)$  for  $x_j = j/(N + 1)$ , we denote  $\mathbf{u} = (u_1, \dots, u_N)$  and  $\mathbf{p} = (p_1, \dots, p_N)$ , and  $u_0$  and  $u_{N+1}$  are fixed by some boundary conditions. These finite-dimensional truncations do indeed possess both the microcanonical and the canonical measures [see (21) and (23) below] as invariant measures. However, we will show in Proposition 1 below that *both the microcanonical measure on the energy shell  $E_N = N/\beta$  and the canonical measure at inverse temperature  $\beta$  associated with the truncated Hamiltonian in (7) converge to the same canonical measure in (6) as the size  $N$  of the truncation tends to infinity*. This result will be proven in the context of quadratic potentials  $V(u)$  but should hold also for more general potentials. Note that the scaling  $E_N = N/\beta$  implies that the energy of the discretized solutions converge to infinity as  $N \rightarrow \infty$ .

The inverse temperature  $\beta$  in (6) controls the concentration of the invariant measure around the minimizers of the Hamiltonian (3). This concentration of measure suggests that the system's dynamics can display metastability if  $V(u)$  is a double-well potential, like  $V(u) = \frac{1}{4}(1 - u^2)^2$ , since in that case the Hamiltonian (3) has two minimizers: These are  $(p, u) = (0, u^-)$  and  $(p, u) = (0, u^+)$  where  $u^\pm$  are the solutions to

$$\kappa^2 u_{xx} = V'(u) \quad (8)$$

that minimize the potential energy  $U(u)$  defined in (4). For  $V(u) = \frac{1}{4}(1 - u^2)^2$  and Neumann boundary condition, these are simply  $u^\pm = \pm 1$ , and they can be similarly obtained for Dirichlet boundary conditions (see Fig. 2 below). We will show in Proposition 2 that the measure in (6) concentrates around these minimizers  $(0, u^\pm)$  as  $\beta \rightarrow \infty$ , therefore suggesting that the solutions to (2) spend long periods of time around one before making a transition to the vicinity of the other.

As a means to quantify this behavior and answer the second question raised above, we will compute the mean frequency, averaged over initial conditions drawn from

the invariant measure (6), at which the solutions to (2) make transitions between two regions that each contain  $(0, u^-)$  and  $(0, u^+)$  and partition its phase space. Specifically, we will take as the boundary between these regions the affine set of codimension 1 defined as

$$\mathcal{S} = \left\{ (p, u) : \int_0^1 (u - u^s) \phi^s dx = 0 \right\} \tag{9}$$

where  $u^s(x)$  is the saddle configuration between  $u^-$  and  $u^+$ , i.e., the solution to (8) with Morse index 1 that minimizes  $U(u)$ , and  $\phi^s(x)$  is the unique eigenfunction with negative eigenvalue of the Hessian operator at  $u^s$ , that is,

$$\mathcal{A}^s = -\kappa^2 \frac{d^2}{dx^2} + V''(u^s(x)). \tag{10}$$

Denoting by  $N_T(p_0, u_0)$  the total number of times the solution to (2) for the initial condition  $(p(0), u(0)) = (p_0, u_0)$  crosses the dividing surface  $\mathcal{S}$  during the time interval  $[0, T]$ ,  $T > 0$ , we will compute

$$\nu_{\mathcal{S}} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int N_T(p_0, u_0) d\mu(p_0, u_0) \equiv \lim_{T \rightarrow \infty} \frac{\mathbb{E}_{\mu} N_T}{2T}. \tag{11}$$

Transition state theory (TST) (Eyring 1935; Horiuti 1938) permits the exact calculation of this quantity for arbitrary surfaces, and the results depend in general on which surface one picks. For the purpose of investigating metastability, we are particularly interested in (11) evaluated for the surface in (9) as trajectories are likely to have minimal recrossing of this surface during a trip between two disjoint sets containing the minimizers  $(0, u_{\pm})$ . Here, we will generalize the derivation in Tal and Vanden-Eijnden (2006) to the wave equation with random initial data to obtain an exact expression for  $\nu_{\mathcal{S}}$  that we will then estimate asymptotically in the limit  $\beta \rightarrow \infty$ . This will allow us to obtain asymptotic expressions for the average residency times  $\tau_-$  and  $\tau_+$  on either side of  $\mathcal{S}$ , defined as

$$\tau_{\pm} = \frac{\mu(\mathcal{B}_{\pm})}{\nu_{\mathcal{S}}} \tag{12}$$

where

$$\mathcal{B}_- = \left\{ (p, u) : \int_0^1 (u - u^s) \phi^s dx < 0 \right\}, \quad \mathcal{B}_+ = \left\{ (p, u) : \int_0^1 (u - u^s) \phi^s dx > 0 \right\}. \tag{13}$$

The asymptotic expressions for  $\tau_{\pm}$  we will arrive at are

$$\tau_{\pm} \sim 2\pi \Lambda_{\pm} e^{\beta \Delta E_{\pm}} \quad \text{as } \beta \rightarrow \infty. \tag{14}$$

Here,  $\Delta E_{\pm} = U(u^s) - U(u^{\pm})$  denote the energy barriers between the minima and the saddle point and

$$\Lambda_{\pm} = \frac{1}{\sqrt{\lambda_1^{\pm}}} \prod_{j=2}^{\infty} \sqrt{\frac{\lambda_j^s}{\lambda_j^{\pm}}} \tag{15}$$

where  $\lambda_j^s$  are the eigenvalues of the operator (10) and  $\lambda_j^\pm$  those of

$$\mathcal{A}^\pm = -\kappa^2 \frac{d^2}{dx^2} + V''(u^\pm(x)). \quad (16)$$

In (15), the eigenvalues are enumerated such that  $\lambda_j < \lambda_k$  whenever  $j < k$ . Note that the negative eigenvalue  $\lambda_1^s < 0$  corresponding to the unstable direction of the saddle solution (i.e., associated with the eigenvector  $\phi^s$  defined above) is not included in the infinite product.

Note that the infinite product in (15) converges in one spatial dimension,  $D = 1$ , but it does not for  $D \geq 2$ , suggesting that the derivation of the TST time cannot be extended to the wave equation in multiple spatial dimensions. This problem, stemming from the scaling of the eigenvalues of the Laplacian, is consistent with the behavior of the SPDE in (5) which is well posed for  $D = 1$ , but not for  $D \geq 2$  (Hairer 2014; Ryser et al. 2012).

Interestingly, the prediction in (14) for the residency times of the solutions to (2) coincides with the zero-damping limit of the residency times of the solutions to the Langevin stochastic partial differential equation in (5), since the latter are given by

$$\tau_\pm^\gamma \sim \pi \frac{\gamma + \sqrt{\gamma^2 + 4|\lambda_1^s|}}{\sqrt{|\lambda_1^s|}} \Lambda_\pm e^{\beta \Delta E_\pm} \quad \text{as } \beta \rightarrow \infty \text{ with } \gamma > 0 \text{ fixed.} \quad (17)$$

An interesting additional question is whether these results also apply for almost all initial conditions drawn from the invariant measure. This trajectory-wise comparison is non-trivial as it requires an assumption of ergodicity. Below, we will investigate this assumption numerically. These computations suggest that the system's dynamics is indeed ergodic and rapidly mixing if the parameter  $\kappa$  is small enough: In this case, the residency times of every trajectory closely match that predicted by TST, and successive visits to  $\mathcal{B}_+$  and  $\mathcal{B}_-$  are roughly uncorrelated statistically. This means that the dynamics of (2) can be effectively coarse-grained to that of a two-state Markov jump process with rates that can be deduced from the mean TST frequency. It also means that in this regime of small  $\kappa$  and infinite energy solutions, the dynamics of (2) is not fundamentally different from that of its stochastic counterpart in (5), and in the latter, we expect that the limits  $\beta \rightarrow \infty$  and  $\gamma \rightarrow 0$  commute. If the parameter  $\kappa$  is too large, however, either ergodicity breaks down, or the mixing time increases and becomes much larger than the inverse of the mean TST frequency. As a result, successive visits in  $\mathcal{B}_+$  and  $\mathcal{B}_-$  become strongly correlated, and the coarse-graining of (2) to a two-state Markov jump process fails. In this regime, the dynamics of (2) is different from that of (5), and in this second equation, the limits as  $\beta \rightarrow \infty$  and  $\gamma \rightarrow 0$  no longer commute.

The remainder of this paper is organized as follows: In Sect. 2, we begin by setting up the wave equation along with its finite-dimensional truncation that we will study. In Sect. 3, we discuss the equivalence of the microcanonical and canonical distributions in infinite dimensions, with detailed calculations of the convergence of characteristic functions of the solutions in ‘‘Appendix 1’’ and of the concentration of measure in

“Appendix 2.” In Sect. 4, we then revisit TST and derive the mean frequency of transition between two specific regions in the phase space of solutions of the wave equation: These calculations are done in the main text for initial conditions drawn from the canonical distribution and in “Appendix 3” for initial conditions drawn from the microcanonical distribution. In Sect. 5, we confirm these results via numerical calculations and also investigate the ergodicity of the dynamics in different parameter regimes. Finally, in Sect. 6, we conclude the paper by discussing the implications of the Hamiltonian dynamics on the low-damping regime of the Langevin system.

## 2 Preliminaries: Discretized Equation, Weak Solutions, and Critical Points

To understand the behavior of the solution to (2), we will consider the finite discretization of the interval  $x \in [0, 1]$ , into  $N + 2$  points, defining  $x_j = j/(N + 1)$ ,  $u_j = u(x_j, t)$  and  $p_j = u_t(x_j, t)$ , and work with the discrete Hamiltonian in (7). Hamilton’s equations associated with (7) are

$$\dot{u}_j = p_j, \quad \dot{p}_j = \kappa^2 N^2(u_{j-1} - 2u_j + u_{j+1}) - V'(u_j) \tag{18}$$

for  $j = 1$  to  $N$ , where the dot indicates differentiation with respect to time. The Hamiltonian also has  $2N$  degrees of freedom as the endpoints  $u_0$  and  $u_{N+1}$  are fixed by the boundary conditions. For example,  $u_0 = u_{N+1} = 0$  for Dirichlet boundary conditions, or  $u_0 = u_1$  and  $u_{N+1} = u_N$  for Neumann boundary conditions.

The solutions to (18) converge to those of the wave equation in (2) as  $N \rightarrow \infty$ . However, since we will consider a parameter regime in which the energy of the initial conditions to (18) scale as  $N/\beta$ , the discretized solutions to (18) will converge to infinite energy solutions of (2), and we cannot expect these solutions to be differentiable in space and time. Therefore, we will have to consider weak solution of (2) (or rather (1)) that satisfy

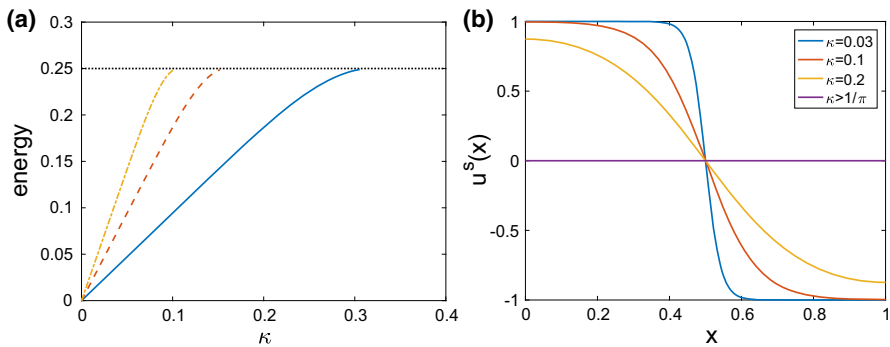
$$\int_0^\infty \int_0^1 (\phi_{tt} - \kappa^2 \phi_{xx}) u + V'(u) \phi \, dx dt = \int_0^1 \phi(x, 0) u_t(x, 0) - \phi_t(x, 0) u(x, 0) dx \tag{19}$$

for all smooth test functions  $\phi(x, t) \in C^\infty$  with compact support in time, the same boundary conditions as  $u(x, t)$ , and with  $\phi_t(x, 0) = 0$ .

As mentioned in the introduction, the critical points of the Hamiltonian in (3) will play an important role in what follows. Because of the separable structure of the Hamiltonian, these critical points are always of the form  $(p, u) = (0, u)$  with  $u$  a solution to (8). The specific shape of these solutions depends both on the form of  $V(u)$  and the choice of boundary conditions. To fix ideas, consider, e.g.,

$$V(u) = \frac{1}{4}(1 - u^2)^2 \tag{20}$$

and Neumann boundary conditions. In this case, the solutions of (8) that minimize  $U(u)$  in (4) (i.e., such that all the eigenvalues of the operator (16) are positive) are the constant functions  $u^+(x) = 1$  and  $u^-(x) = -1$ .



**Fig. 1** Critical points of  $U(u)$  for the potential given in (20) and Neumann boundary conditions. **a** Energies of the first three non-trivial saddle points that emerge as  $\kappa$  is decreased. The bifurcations happen at  $1/n\pi$  for  $n = 1, 2, 3, \dots$  corresponding to the  $u(x) = 0$  solution with energy  $1/4$  gaining negative eigenvalues. **b** The index 1 saddle point  $u^s(x)$  (one negative eigenvalue) for various values of  $\kappa$  found by numerically solving a discretized version of (8) with initial guess of the form  $\tanh(\cos(\pi x)/(\sqrt{2\pi\kappa}))$ . For  $\kappa < 1/\pi$ , these points correspond to points along the blue solid line in the left panel. Also for  $\kappa < 1/\pi$ , due to symmetry, there is a second index 1 saddle point  $u^s(-x)$  (Color figure online)

Besides the minimizers of  $U(u)$ , we will also need the saddle point with Morse index 1, i.e., the function  $u^s(x)$  that is a solution to (8) such that the operator (10) has only one negative eigenvalue. Unlike the minimizers,  $u^s(x)$  as well as its energy  $E^s$  depends on the value of  $\kappa$  (see Fig. 1). For instance,  $u(x) = 0$  is always a critical point of the Hamiltonian, but it is only the index 1 saddle point when  $\kappa > 1/\pi$ . This can be seen from the eigenvalues of the operator (10) evaluated at  $u^s = 0$ , namely  $\lambda = -1 + \kappa^2 n^2 \pi^2$  for  $n = 0, 1, 2, 3, \dots$ . The  $n = 0$  eigenvalue is always  $-1$  regardless of the value of  $\kappa$ . It is the only negative eigenvalue when  $\kappa > 1/\pi$ , and  $u(x) = 0$  is therefore the index 1 saddle point. The  $u(x) = 0$  solution becomes an index 2 saddle point when  $\kappa < 1/\pi$ , and an index three saddle when  $\kappa < 1/2\pi$ , and so forth as it gains more negative eigenvalues. These bifurcation points correspond to the appearance of lower energy, lower index, saddle points. The energy of these saddle points for the first three bifurcations is shown in Fig. 1a.

A similar analysis can be made if the potential  $V(u)$  is modified (e.g., if we make it asymmetric) or if the boundary conditions are changed [e.g., if we use Dirichlet boundary conditions,  $u(0) = u(1) = 0$ ].

### 3 Existence and Equivalence of Invariant Measures

The finite-dimensional Hamiltonian system associated with (7) has two types of invariant measures: the microcanonical measures each supported on one energy shell  $H_N(\mathbf{p}, \mathbf{u}) = E_N$  with  $E_N > 0$

$$dm_N = c_N^{-1} \delta(E_N - H_N(\mathbf{p}, \mathbf{u})) d\mathbf{p} d\mathbf{u} \equiv c_N^{-1} \frac{d\sigma(\mathbf{p}, \mathbf{u})}{|\nabla H_N(\mathbf{p}, \mathbf{u})|} \tag{21}$$



where  $d\sigma(\mathbf{p}, \mathbf{u})$  denotes the surface element on  $H_N(\mathbf{p}, \mathbf{u}) = E_N$  and

$$c_N = \int_{\mathbb{R}^{2N}} \delta(E_N - H_N(\mathbf{p}, \mathbf{u})) d\mathbf{p}d\mathbf{u} \equiv \int_{H_N(\mathbf{p}, \mathbf{u})=E_N} \frac{d\sigma(\mathbf{p}, \mathbf{u})}{|\nabla H_N(\mathbf{p}, \mathbf{u})|}; \tag{22}$$

and the canonical measures parametrized by the inverse temperature  $\beta > 0$

$$dM_N = C_N^{-1} e^{-\beta H_N(\mathbf{p}, \mathbf{u})} d\mathbf{p}d\mathbf{u} \tag{23}$$

where

$$C_N = \int_{\mathbb{R}^{2N}} e^{-\beta H_N(\mathbf{p}, \mathbf{u})} d\mathbf{p}d\mathbf{u}. \tag{24}$$

We claim that if we set  $E_N = N/\beta$  and let  $N \rightarrow \infty$  with  $\beta$  fixed, the microcanonical and canonical measures above converge to the same limit. Here, we will establish this result in the context of quadratic potential functions  $V(u)$ . Specifically, the following proposition is proven in ‘‘Appendix 1’’:

**Proposition 1** *Pick two smooth deterministic functions  $s(x)$  and  $t(x)$ , denote  $\mathbf{s} = (s_1, \dots, s_N)$  and  $\mathbf{t} = (t_1, \dots, t_N)$  by defining  $s_j = s(x_j)$  and  $t_j = t(x_j)$  for  $j = 1 \dots N$ , and consider the characteristic functions of the microcanonical and the canonical measures defined, respectively, as*

$$\begin{aligned} \phi_N^m(\mathbf{s}, \mathbf{t}) &= \int_{\mathbb{R}^{2N}} \exp\left(\frac{i}{N} \mathbf{s} \cdot \mathbf{p} + \frac{i}{N} \mathbf{t} \cdot \mathbf{u}\right) dm_N(\mathbf{p}, \mathbf{u}) \\ \phi_N^M(\mathbf{s}, \mathbf{t}) &= \int_{\mathbb{R}^{2N}} \exp\left(\frac{i}{N} \mathbf{s} \cdot \mathbf{p} + \frac{i}{N} \mathbf{t} \cdot \mathbf{u}\right) dM_N(\mathbf{p}, \mathbf{u}) \end{aligned}$$

Then if we set  $E_N = N/\beta$  and  $V(u) = \frac{1}{2}\alpha u^2$ ,  $\alpha > 0$ , both  $\phi_N^m$  and  $\phi_N^M$  have the same limit as  $N \rightarrow \infty$ :

$$\begin{aligned} \lim_{N \rightarrow \infty} \phi_N^m(\mathbf{s}, \mathbf{t}) &= \lim_{N \rightarrow \infty} \phi_N^M(\mathbf{s}, \mathbf{t}) \\ &= \phi(\mathbf{s}, \mathbf{t}) \equiv \exp\left(-\frac{1}{2\beta} \int_0^1 s^2(x) dx + \int_0^1 \int_0^1 t(x) C(x, y) t(y) dx dy\right), \end{aligned} \tag{25}$$

where the covariance function  $C(x, y)$  solves

$$\left(-\kappa^2 \frac{\partial^2}{\partial x^2} + \alpha\right) C(x, y) = \delta(x - y) \tag{26}$$

with the same boundary conditions as (2).

The functional  $\phi(\mathbf{s}, \mathbf{t})$  in (25) is the characteristic functional of a Gaussian measure, and it is consistent with the stochastic processes  $p(x)$  and  $u(x)$  being statistically independent and such that  $p(x)$  is a spatial white noise, scaled by the parameter  $\beta$ , that is a mean-zero Gaussian process with covariance function

$$\mathbb{E}[p(x)p(y)] = \beta^{-1} \delta(x - y); \tag{27}$$

and  $u(x)$  is a mean-zero Gaussian process with covariance function

$$\mathbb{E}[u(x)u(y)] = C(x, y) \quad (28)$$

where  $C(x, y)$  solves (26). Note that the Gaussian measure whose characteristic function is (25) can be written formally as (6) with  $V(u) = \frac{1}{2}\alpha u^2$ .

Proposition 1 is stated for a quadratic  $V(u)$ , but we believe that a similar statement holds for a general  $V(u)$  that is bounded below and such that  $V(u) \geq C|u|^2$  for some  $C > 0$  as  $|u| \rightarrow \infty$ . Indeed, it seems reasonable to think that the canonical measure is becoming concentrated near the energy surface  $E_N = N/\beta$  in the limit of  $N \rightarrow \infty$ , much in the same way as the volume of a ball becomes concentrated in a thin shell near its surface as the space dimension increases (Flaschka 2010). This observation plays a key role in the proof of Proposition 1, and it suggests that both (21) and (23) converge to (6) as  $N \rightarrow \infty$  even for a  $V(u)$  that is non-quadratic. In this case, the way to interpret (6) is as follows: The processes  $p(x)$  and  $u(x)$  are again statistically independent, with  $p(x)$  being a spatial white noise with the covariance in (27), but  $u(x)$  is no longer Gaussian. For example, if we consider Dirichlet boundary conditions, let us introduce the scaled Brownian bridge process  $B(x)$ , that is, the mean-zero Gaussian process with covariance function

$$\mathbb{E}[B(x)B(y)] = (\beta\kappa^2)^{-1}(\min(x, y) - xy). \quad (29)$$

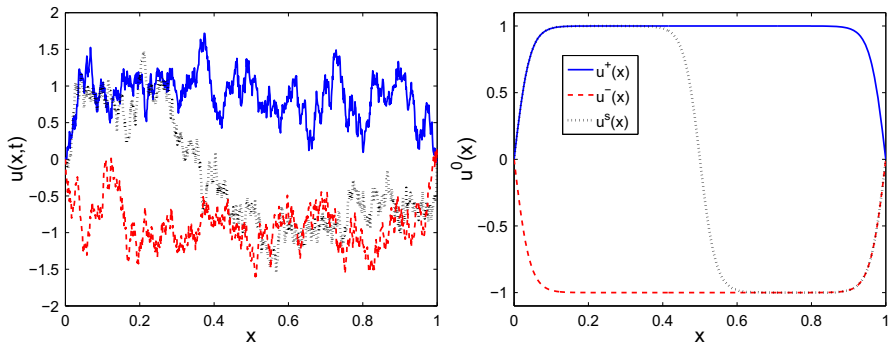
Then  $u(x)$  is the process whose Radon–Nikodym derivative (i.e., density) with respect to  $B(x)$  is given by

$$\exp\left(-\beta \int_0^1 V(B(x))dx\right). \quad (30)$$

The parameter  $\beta$  appears in this term as well, indicating that  $\beta$  and  $\kappa$  have different influences on the function  $u(x)$ . We discuss in Sect. 4.3 the effect of  $\beta$  on reducing the variance of the solution, equivalently the concentration of the measure, and in Sect. 5 the effect of  $\kappa$  on the covariance of  $u(x)$  leading to more uniform in space solutions. Shown in Fig. 2 are three snapshots of the solution to (2) at three different moments in time, when initial conditions are chosen approximately from the invariant measure (6) with the double-well potential (20) and Dirichlet boundary conditions.

McKean and Vaninsky (1999) (see also Friedlander 1985; Zhidkov 1994) verified that (6) is an invariant measure of (2) with periodic boundary conditions  $u(0, t) = u(1, t)$ . This invariance of the Gibbs measure is not unique to the wave equation and has been shown for other Hamiltonian PDEs such as the nonlinear Schrödinger equation (Bourgain 1994; Lebowitz et al. 1988; Tzvetkov 2006, 2008), Korteweg–de Vries equation (Oh et al. 2016), and others (Oh 2010; Deng 2015).

We have added to (6) the parameter  $\beta$  not present in McKean's original results. Consistent with the aforementioned scaling of the variance of the stochastic functions  $p(x)$  and  $u(x)$  by  $1/\beta$ , the parameter  $\beta$ , added to the measure (6), is playing the roll of an inverse temperature. Recall the measure (6) is also an invariant measure for the noisy-damped system (5) regardless of the value of  $\gamma > 0$ , the damping parameter. In this case, the temperature  $\beta^{-1}$  characterizes the size of the fluctuations



**Fig. 2** (Left) Shown are three sample solutions to (18), the discretized version of (2), with Dirichlet boundary conditions, at various instances in time, with potential function (20),  $N = 1024$  discretization points,  $\kappa = 0.03$ , and initial conditions with energy  $N/\beta$  and  $\beta\Delta E = 4.6$ . The blue solid line is a solution near the minimizer  $u^+(x)$ , the red dashed line is a solution near the minimizer  $u^-(x)$ , and the black dotted line is in the process of transitioning from  $u^+(x)$  to  $u^-(x)$ . (Right) These critical points along with one of the saddles,  $u^s(x)$ , separating them (Color figure online)

driving the system, and therefore the variance of the solutions. In terms of the system with Hamiltonian (3), the parameter  $\beta^{-1}$  acts like a temperature parameter and also characterizes the spatial variance of the weak solutions.

### 4 Expected Residency Times from Transition State Theory

As a means to determine whether other features of the dynamics of the wave equation (2) and its Langevin counterpart in (5) coincide, we will now compute the mean residency times,  $\tau_{\pm}$ , the solutions spend in one of two regions that each contain a minimizer of the Hamiltonian and partition phase space. To this end, we revisit TST, developed for finite-dimensional systems (Eyring 1935; Horiuti 1938; Wigner 1938), that computes the mean frequency  $\nu_S$  of crossing a codimension one surface partitioning the phase space into two regions; the mean residency times  $\tau_{\pm}$  are related to  $\nu_S$  by means of the definition in (12). Rather than the dynamics of a single trajectory, here we consider the average crossing frequency of a collection of trajectories with initial distribution given by the invariant measure (6). We start by considering the finite-dimensional system (18), with initial conditions selected from the invariant measure (23), and then justify the infinite-dimensional limit. The same calculation for the microcanonical invariant measure (21) is considered in “Appendix 3.”

#### 4.1 TST Revisited in an Averaged Setup

For a Hamiltonian system, we follow a similar derivation to those in Vanden-Eijnden and Tal (2005), Tal and Vanden-Eijnden (2006) but do not assume ergodicity at this point. We divide the finite-dimensional phase space,  $\mathbf{u}, \mathbf{p} \in \mathbb{R}^N$  into two regions,  $B_-$  containing  $(\mathbf{0}, \mathbf{u}^-)$  and  $B_+$  containing  $(\mathbf{0}, \mathbf{u}^+)$ , by means of a dividing surface  $S$  that separates the two: How to specify this surface will be discussed later in Sect. 4.2. For

a single trajectory, the average TST frequency of transitions from region  $B_-$  to  $B_+$  over a time  $T > 0$  is given by

$$v_S^T(\mathbf{u}(0), \mathbf{p}(0)) = \frac{1}{T} N_T(\mathbf{u}(0), \mathbf{p}(0)) \tag{31}$$

where  $N_T(\mathbf{u}(0), \mathbf{p}(0))$  counts the number of times the trajectory  $\mathbf{u}(t)$  solution of (18) for the initial condition  $(\mathbf{u}(0), \mathbf{p}(0))$  leaves the region  $B_-$  by crossing  $S$  within the time interval  $[0, T]$ . It is convenient to parametrize  $S$  by the zero level set of some function  $q(\mathbf{u})$ , specifically

$$S = \{(\mathbf{p}, \mathbf{u}) : q(\mathbf{u}) = 0\}. \tag{32}$$

We will explain the specific choice of  $q(\mathbf{u})$  leading to the surface in (9) in the next section. For convention, take  $B_+ = \{(\mathbf{p}, \mathbf{u}) : q(\mathbf{u}) > 0\}$  and  $B_- = \{(\mathbf{p}, \mathbf{u}) : q(\mathbf{u}) < 0\}$ . Transitions from  $B_-$  to  $B_+$  are then changes in the sign of  $q(\mathbf{u})$  from negative to positive. This allows the frequency to be written as

$$v_S^T(\mathbf{u}(0), \mathbf{p}(0)) = \frac{1}{T} \int_0^T \max\left(\frac{d}{dt} \Theta(q(\mathbf{u}(t))), 0\right) dt \tag{33}$$

where  $\Theta(z)$  is the Heaviside function,  $\Theta(z) = 0$  if  $z < 0$ ,  $\Theta(0) = \frac{1}{2}$ ,  $\Theta(z) = 1$  if  $z > 0$ . Using the property that  $\Theta'(z) = \delta(z)$ , where  $\delta(z)$  is the Dirac distribution, using the chain rule to get to  $\dot{\mathbf{u}}(t) \cdot \nabla q(\mathbf{u}(t)) \delta(q(\mathbf{u}(t)))$ , and setting  $\dot{\mathbf{u}}(t) = \mathbf{p}(t)$ , (recall the dot indicates differentiation with respect to time), we arrive at

$$v_S^T(\mathbf{u}(0), \mathbf{p}(0)) = \frac{1}{T} \int_0^T \max(\mathbf{p}(t) \cdot \nabla q(\mathbf{u}(t)), 0) \delta(q(\mathbf{u}(t))) dt. \tag{34}$$

We now want to compute the average of  $v_S^T$  over trajectories with initial conditions chosen with respect to an invariant measure, specifically the canonical distribution (23). By definition, if we denote by  $(\mathbf{p}(t), \mathbf{p}_0, \mathbf{u}_0), \mathbf{u}(t, \mathbf{p}_0, \mathbf{u}_0)$  the solution to (18) for the initial condition  $(\mathbf{p}(0), \mathbf{p}_0, \mathbf{u}_0), \mathbf{u}(0, \mathbf{p}_0, \mathbf{u}_0) = (\mathbf{p}_0, \mathbf{u}_0)$ , for any suitable test function  $f(\mathbf{p}_0, \mathbf{u}_0)$ , we have

$$\int f(\mathbf{p}(t, \mathbf{p}_0, \mathbf{u}_0), \mathbf{u}(t, \mathbf{p}_0, \mathbf{u}_0)) dM_N(\mathbf{p}_0, \mathbf{u}_0) = \int f(\mathbf{p}, \mathbf{u}) dM_N(\mathbf{p}, \mathbf{u}). \tag{35}$$

As a result, the ensemble average of the integrand in (34) is invariant in time, and after proper interpretation of integration with respect to the Dirac distribution, we arrive at the following expression for the mean frequency

$$v_S^c = C_N^{-1} \int_S \max(\mathbf{p} \cdot \hat{n}(\mathbf{u}), 0) e^{-\beta H_N(\mathbf{u}, \mathbf{p})} d\tilde{\sigma}(\mathbf{p}, \mathbf{u}) \tag{36}$$

where  $\hat{n}(\mathbf{u})$  is the unit normal vector to the surface  $S$  restricted to the space of  $\mathbf{u}$  and  $d\tilde{\sigma}(\mathbf{p}, \mathbf{u})$  is the surface element on  $S$ . Notice that (36) is identical to the expression we would have arrived at if we had assumed ergodicity with respect to the invariant

measure, and then taken the longtime limit of (34), replacing the longtime average with an average over the invariant measure.

### 4.2 Optimal TST

To understand metastability, we are interested in evaluating (36) for the dividing surface  $S$  that best separates the two metastable states containing  $u^-$  and  $u^+$ , minimizing the mean frequency of transition (Vanden-Eijnden and Tal 2005) (and thereby maximizing the transition time). While we may not be able to determine this surface exactly, a reasonable surface to use is the finite-dimensional approximation of the codimension one hyperplane given by (4). The hyperplane in (4) passes through the saddle point, perpendicular to the unstable direction of the gradient flow  $w_t(x, t) = \kappa^2 w_{xx}(x, t) - V'(w(x, t))$ . While this surface may not be the true surface that maximizes the mean transition time, for large  $\beta$  we expect this surface to have minimal recrossings, and therefore produce a close bound on the mean transition time.

The reasoning for this surface having minimal recrossings starts with the assumption that most transitions will occur near the saddle point, the lowest energy point on the dividing surface (Ariel and Vanden-Eijnden 2007). Near the saddle point, the linearized Hamiltonian system has one hyperbolic direction corresponding to the one unstable direction of the gradient flow; the remaining directions are oscillatory. Any surface not perpendicular to the unstable direction of the gradient flow (which bisects the stable and unstable hyperbolic directions) would also not be orthogonal to all the oscillatory directions of the Hamiltonian flow, allowing for many recrossings of the surface in at least one of the oscillatory directions.

The finite-dimensional approximation to (4) is the hyperplane

$$S = \{(\mathbf{p}, \mathbf{u}) : (\mathbf{u} - \mathbf{u}^s) \cdot \boldsymbol{\phi}^{(1)} = 0\} \tag{37}$$

where  $\mathbf{u}^s$  and  $\boldsymbol{\phi}^{(1)}$  are the discrete equivalents to  $u^s$  and  $\phi^s$ , that is,  $\mathbf{u}^s$  is the Morse index 1 critical point of

$$U_N(\mathbf{u}) = \frac{1}{N} \sum_{j=1}^N V(u_j) + N \sum_{j=0}^N \frac{1}{2} \kappa^2 (u_{j+1} - u_j)^2, \tag{38}$$

with minimum energy, and  $\boldsymbol{\phi}^{(1)}$  is the unique eigenvector with negative eigenvalue of the Hessian of (38) evaluated at  $\mathbf{u} = \mathbf{u}^s$ . In other words, these vectors solve

$$0 = \kappa^2 N^2 (u_{j-1}^s - 2u_j^s + u_{j+1}^s) - V'(u_j^s) \tag{39}$$

$$\lambda_k^s \phi_j^{(k)} = -\kappa^2 N^2 (\phi_{j-1}^{(k)} - 2\phi_j^{(k)} + \phi_{j+1}^{(k)}) + V''(u_j^s) \phi_j^{(k)}, \tag{40}$$

when  $k = 1$  and  $\lambda_1^s < 0$  for  $j = 1, \dots, N$  and appropriate boundary conditions fixing  $u_0^s, u_{N+1}^s, \phi_0^{(k)}$ , and  $\phi_{N+1}^{(k)}$ . Note the remaining eigenvectors for  $k = 2 \dots N$  solve (40) when  $\lambda_k^s > 0$ .

### 4.3 Quadratic Approximation as $\beta \rightarrow \infty$

In order to proceed with computing the TST frequency (36), we must perform an integration over the hyperplane given by (37), as well as another integration to evaluate the normalization constant. While we cannot evaluate the integrals for an arbitrary potential function, we can proceed if it is justified to approximate the potential  $V(u)$  in the Hamiltonian locally as a quadratic function. This approximation is justified if most of the measure we are integrating over is confined within a localized region, within which  $V(u)$  is approximately quadratic. We show this concentration for the quadratic potential and work under the assumption that the total measure outside this region for the true potential  $V(u)$  is exponentially small, and therefore even large errors in the approximation of  $V(u)$  do not contribute significantly to the integration for the TST frequency.

For the canonical measure, (23), intuitively one would expect the measure to concentrate around the points  $(\mathbf{0}, \mathbf{u}^\pm)$  that minimize the Hamiltonian as  $\beta \rightarrow \infty$ . It is convenient to write an expansion for  $\mathbf{u}$  in the rotated and rescaled eigenvector basis as

$$\mathbf{u} = \mathbf{u}^\pm + \tilde{\mathbf{u}} = \mathbf{u}^\pm + \sum_{k=1}^N b_k^\pm \boldsymbol{\psi}^{(k)\pm}, \tag{41}$$

in which the eigenvalues  $\lambda_k^\pm$  and the corresponding eigenvectors  $\boldsymbol{\psi}^{(k)}$  for each  $k = 1 \dots N$  solve

$$\lambda_k^\pm \psi_j^{(k)\pm} = -\kappa^2 N^2 \left( \psi_{j-1}^{(k)\pm} - 2\psi_j^{(k)\pm} + \psi_{j+1}^{(k)\pm} \right) + V''(u_j^\pm) \psi_j^{(k)\pm} \tag{42}$$

for  $j = 1 \dots N$  and appropriate boundary conditions fixing  $\psi_0^{(k)\pm}$  and  $\psi_{N+1}^{(k)\pm}$ . The eigenvectors are normalized to satisfy  $\frac{1}{N} \sum_{j=1}^N (\psi_j^{(k)\pm})^2 = 1$ , the discrete version of  $L_2$  normalization. We use the same numbering convention as before,  $\lambda_j^\pm < \lambda_k^\pm$  if  $j < k$ , but this time all the  $\lambda_k^\pm > 0$  for  $k = 1 \dots N$ . The quadratic expansion of the potential energy (38) is

$$\begin{aligned} U_N(\mathbf{u}) &\sim U_N(\mathbf{u}^\pm) + \frac{1}{2N} \sum_{k=1}^N \lambda_k^\pm (b_k^\pm)^2 \boldsymbol{\psi}^{(k)\pm} \cdot \boldsymbol{\psi}^{(k)\pm} \quad \text{as } \mathbf{u} \rightarrow \mathbf{u}^\pm \\ &= E_N^\pm + \frac{1}{2} \sum_{k=1}^N \lambda_k^\pm (b_k^\pm)^2. \end{aligned} \tag{43}$$

Defining  $\tilde{p}_j = \dot{b}_j^\pm$ , the rotated momenta, (recall the dot indicates differentiation with respect to time), the quadratic approximation of the Hamiltonian is

$$H_N \sim E_N^\pm + \sum_{j=1}^N \left( \frac{1}{2N} \tilde{p}_j^2 + \frac{1}{2} \lambda_j^\pm (b_j^\pm)^2 \right) \quad \text{as } \mathbf{u} \rightarrow \mathbf{u}^\pm, \tag{44}$$

where we have defined  $E_N^\pm = U_N(\mathbf{u}^\pm)$ . We show in ‘‘Concentration of the Canonical Measure’’ section of Appendix 2 that as  $\beta \rightarrow \infty$  the canonical measure concentrates into a localized region around the point  $\mathbf{b}^\pm = 0$ , even as  $N \rightarrow \infty$  in the following proposition:

**Proposition 2** Consider the marginal canonical measure with the quadratic Hamiltonian approximation (44) and an  $N$ -dimensional box with side edge lengths  $2\delta_j$ ,

$$D_\delta^N = \{\mathbf{b}^\pm : -\delta_j \leq b_j^\pm \leq \delta_j \quad \forall j = 1 \dots N\},$$

where  $\delta_j$  are such that  $\sum_{j=1}^N \delta_j^2$  converges as  $N \rightarrow \infty$ . Assume that the eigenvalues  $\lambda_j^\pm$  solving (42) satisfy

$$\lambda_j^\pm \sim Cj^2 \quad \text{as } j \rightarrow \infty, \text{ for some } C > 0$$

Then, for every  $\hat{\delta} > 0$ , there exists  $\beta_1 > 0$  and a set of  $\{\delta_j\}_{j=1}^N$  such that for every  $\beta > \beta_1$  and every  $N > 0$

$$\frac{\int_{D_\delta^N} \exp\left(-\frac{\beta}{2} \sum_{j=1}^N \lambda_j^\pm (b_j^\pm)^2\right) d\mathbf{b}^\pm}{\int_{\mathbb{R}^N} \exp\left(-\frac{\beta}{2} \sum_{j=1}^N \lambda_j^\pm (b_j^\pm)^2\right) d\mathbf{b}^\pm} > 1 - \hat{\delta} \tag{45}$$

Together with the assumption that the total measure outside the localized region is exponentially small, the quadratic approximation (44) can be used to approximate the normalization constant in (36). We first separate the integration into the two regions,  $B_+ = \{(\mathbf{p}, \mathbf{u}) : (\mathbf{u} - \mathbf{u}^s) \cdot \boldsymbol{\phi}^{(1)} > 0\}$  and  $B_- = \{(\mathbf{p}, \mathbf{u}) : (\mathbf{u} - \mathbf{u}^s) \cdot \boldsymbol{\phi}^{(1)} < 0\}$  that partition space, separated by  $S$ ,

$$C_N = C_N^+ + C_N^- \tag{46}$$

where

$$C_N^\pm = \int_{B_\pm} e^{-\beta H_N(\mathbf{p}, \mathbf{u})} d\mathbf{p} d\mathbf{u}. \tag{47}$$

Now, each region contains only one minimizer of the potential energy. Within region  $B_+$ , we expect most of the measure is near  $\mathbf{u}^+$  while in  $B_-$ , its in the vicinity of  $\mathbf{u}^-$ . The result after completing the integration appears below in Proposition 3.

For the integral in (36), the measure is restricted to the surface  $S$  defined in (37). On this surface, it is the point  $(\mathbf{0}, \mathbf{u}^s)$  that minimizes the Hamiltonian. We expand about this point and rotate to an eigenvector basis

$$\mathbf{u} = \mathbf{u}^s + \tilde{\mathbf{u}} = \mathbf{u}^s + \sum_{k=1}^N a_k \boldsymbol{\phi}^{(k)} \tag{48}$$

with  $\lambda_k^s$  and  $\phi^{(k)}$  for each  $k = 1 \dots N$  solving the system in (40) with appropriate boundary conditions fixing  $\phi_0^{(k)}$ , and  $\phi_{N+1}^{(k)}$ , and normalized to satisfy  $\frac{1}{N} \sum_{j=1}^N (\phi_j^{(k)})^2 = 1$ . Our convention is to order the eigenvalues such that  $\lambda_1^s$  is the sole negative eigenvalue and  $\lambda_j^s < \lambda_k^s$  for  $j < k$ . The quadratic approximation of the potential energy (38) is

$$\begin{aligned} U_N(\mathbf{u}) &\sim U_N(\mathbf{u}^s) + \frac{1}{2N} \sum_{k=1}^N \lambda_k^s a_k^2 \phi^{(k)} \cdot \phi^{(k)} \quad \text{as } \mathbf{u} \rightarrow \mathbf{u}^s \\ &= E_N^s + \frac{1}{2} \sum_{k=1}^N \lambda_k^s a_k^2 \end{aligned} \quad (49)$$

due to the defined normalization of the eigenvectors, and where we have defined  $E_N^s = U_N(\mathbf{u}^s)$ . Defining  $\tilde{p}_j = \dot{a}_j$ , the rotated momenta, the approximation of the Hamiltonian in the vicinity of  $\mathbf{u}^s$  is

$$H_N \sim E_N^s + \sum_{j=1}^N \left( \frac{1}{2N} \tilde{p}_j^2 + \frac{1}{2} \lambda_j^s a_j^2 \right) \quad \text{as } \mathbf{u} \rightarrow \mathbf{u}^s. \quad (50)$$

In ‘‘Concentration of the Canonical Measure’’ section of Appendix 2, we state Proposition 4, an analogous Proposition to Proposition 2, for the concentration of the measure restricted to the surface  $S$ . This allows the use of (50) to complete the integration in (36).

The following proposition states the asymptotic in  $\beta$  values of the integrals needed to compute the TST frequency in (36). The justification of these formulas appears in ‘‘Evaluation of the Integrals’’ section of Appendix 2.

**Proposition 3** *For the canonical ensemble, the asymptotic in  $\beta \rightarrow \infty$ , valid for all  $N > 0$ , expansion of the integral in (36) is*

$$\int_S \max(\mathbf{p} \cdot \hat{\mathbf{n}}(\mathbf{u}), 0) e^{-\beta H_N(\mathbf{u}, \mathbf{p})} d\tilde{\sigma}(\mathbf{p}, \mathbf{u}) \sim \frac{2^{N-2} N^N (N-2)! S_{2N-3}}{\beta^N \prod_{j=2}^N \sqrt{\lambda_j^s}} e^{-\beta E_N^s} \quad (51)$$

and the expansion of the integrals in (47) is

$$C_N^\pm \sim \frac{2^{N-1} N^N (N-1)! S_{2N-1}}{\beta^N \prod_{j=1}^N \sqrt{\lambda_j^\pm}} e^{-\beta E_N^\pm}. \quad (52)$$

In the above, the eigenvalues  $\lambda_j^s$ ,  $j = 2 \dots N$  solve (40), the eigenvalues  $\lambda_j^\pm$  solve (42), and  $S_n = 2\pi^{(n+1)/2} / \Gamma((n+1)/2)$  is the surface area of the  $n$ -dimensional sphere (embedded in  $n+1$  dimensional space) of unit radius.



Using the above asymptotic in  $\beta$  expansions, the TST frequency from (36) is

$$v_S^c \sim \frac{1}{2\pi} \frac{\prod_{j=2}^N \frac{1}{\sqrt{\lambda_j^s}} e^{-\beta E_N^s}}{\prod_{j=1}^N \frac{1}{\sqrt{\lambda_j^+}} e^{-\beta E_N^+} + \prod_{j=1}^N \frac{1}{\sqrt{\lambda_j^-}} e^{-\beta E_N^-}}. \tag{53}$$

### 4.4 Mean Residency Times

The mean residency times on either side of  $S$ , given by (37), can be computed with the same integrals considered above. Namely,

$$\tau_{\pm}^c = \frac{C_N^{\pm}}{C_N v_S^c} \tag{54}$$

for the canonical ensemble. With the asymptotic expansions of the integrals in Proposition 3, we arrive at the asymptotic in  $\beta$  expressions for the mean residency time

$$\tau_{\pm}^c \sim 2\pi \frac{\prod_{j=2}^N \sqrt{\lambda_j^s}}{\prod_{j=1}^N \sqrt{\lambda_j^{\pm}}} e^{\beta(E_N^s - E_N^{\pm})}. \tag{55}$$

This expression is valid for all  $N$  allowing us to take the limit as  $N \rightarrow \infty$ , arriving at  $\tau_{\pm}$  in (14), provided the infinite product  $\Lambda_{\pm}$  given in (15) converges. It is the infinite-dimensional limit of the finite-dimensional eigenvalue products in (55). For example,  $\Lambda_{\pm}$  converges when considering eigenvalues related to the Laplacian in one spatial dimension, as we do here (however, it does not for two or more spatial dimensions).

## 5 Numerical Results

Using numerical simulations of the finite-dimensional truncation of (2), we investigate the implication of our results for the longtime dynamics of the nonlinear wave equation. We proceed in two steps. First, we validate our approximations by considering an ensemble of trajectories chosen approximately from the finite-dimensional truncation of the invariant measure (6), and estimating numerically the mean residency time averaged over these initial conditions. Second, we compute the residency time averaged over a single trajectory in order to investigate the extent to which the dynamics of (2) is ergodic and mixing with respect to the invariant measure (6). We find support for this statement for small values of the exchange coupling  $\kappa$ . For larger values of  $\kappa$ , however, we observe that either ergodicity fails, or the mixing time grows so much that we cannot observe convergence of the time averages toward their ensemble counterparts averaged over the invariant measure (6).

We run numerical simulations of the discretized system of  $2N$  first-order differential equations (18) with Neumann boundary conditions using the velocity-Verlet

integrator to conserve the energy on average. Initial conditions are chosen from an approximation of the invariant measure (6): the finite-dimensional canonical measure (23) with the quadratic approximation for the Hamiltonian (44). Specifically,  $\mathbf{p} \in \mathbb{R}^N$  is a discretization of white noise and  $\mathbf{u} \in \mathbb{R}^N$  is a Gaussian random variable with known mean and covariance. They are computed as

$$\mathbf{p} = \sqrt{\frac{N}{\beta}} \mathbf{h}, \quad \mathbf{u} = \mathbf{u}^+ + \frac{\alpha}{\sqrt{\beta}} \sum_{j=1}^N g_j \sqrt{\lambda_j^+} \boldsymbol{\psi}^{(k)+} \quad (56)$$

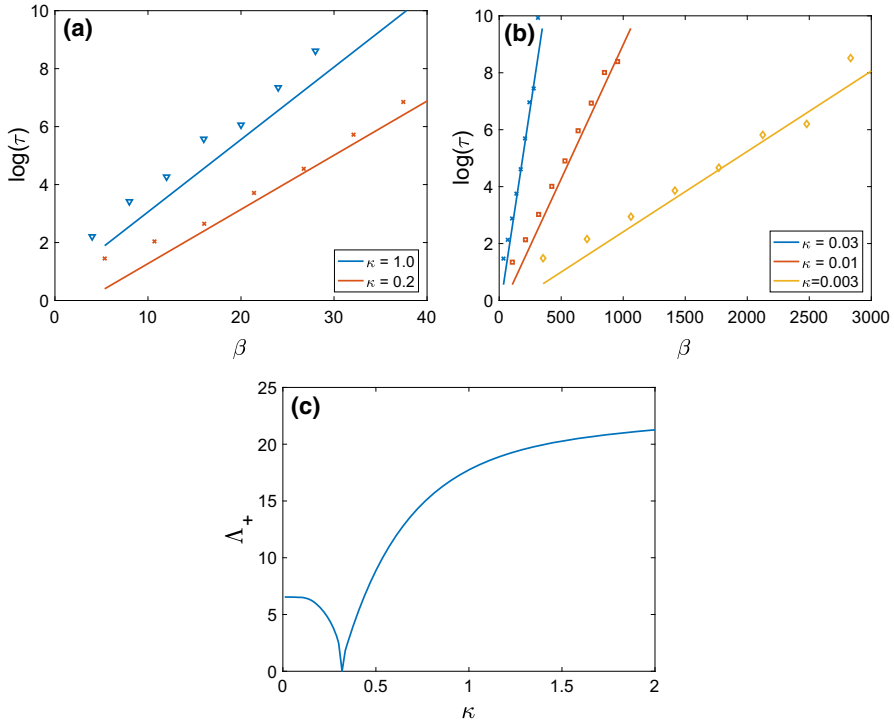
in which  $\mathbf{h} = (h_1, \dots, h_N)$  and  $\mathbf{g} = (g_1, \dots, g_N)$  are vectors of  $N$  independent mean-zero standard Gaussian random variables,  $\mathbf{u}^+ = 1$  for Neumann boundary conditions, the eigenvalues  $\lambda_+$  and eigenvectors  $\boldsymbol{\psi}^{(k)+}$  solve (42), normalized such that  $\boldsymbol{\psi}^{(k)+} \cdot \boldsymbol{\psi}^{(k)+} = N$ , and  $\alpha$  is chosen to ensure the energy of the initial condition is  $N/\beta$  exactly.

For the potential function (20), the system is symmetric. We therefore determine the mean transition time by taking the total simulation time and divide it by the number of transitions across the diving hyperplane  $S$  described in Sect. 4.2. Specifically, we compute the eigenvector  $\boldsymbol{\psi}^{(1)}$  solving (40), the normal vector to the surface  $S$ . At each time step, we compute the distance to this dividing surface. A change in sign of the distance indicates a crossing of the surface. Notice that for  $\kappa < 1/\pi$ , there are two saddle points; yet, they both lie in the just described plane.

In Fig. 3a, we show excellent agreement between the mean transition time found from the numerical simulations at finite  $N = 128$  averaged over an ensemble of 1000 trajectories, and the infinite-dimensional theory, (14). We choose two values of  $\kappa$ , one smaller than the bifurcation point  $\kappa = 1/\pi$  in which  $u^s$  is non-uniform and one larger than this point in which case  $u^s = 0$ . This agreement supports the validity of the quadratic approximation used in obtaining (14).

Next, we investigate the ability of the TST time to describe the mean transition time of a single trajectory. As we pointed out above in the derivation of  $\tau_{\pm}$ , we would arrive at the same expression had we assumed the flow was ergodic with respect to the invariant measure. Recall in finite dimensions, no Hamiltonian systems are ergodic with respect to the invariant measure given by the canonical ensemble while many are conjectured to be ergodic with respect to the invariant measure given by the microcanonical ensemble (Sinai 1976). In infinite dimensions, McKean and Vaninsky conjectured that the flow of (2) is ergodic (a.k.a. metrically transitive) with respect to the measure (6) except for completely integrable systems with an infinite number of conserved quantities like  $V'(u) = u, \sin u$  or  $\sinh u$  (McKean and Vaninsky 1999). Although the measure (6) looks like the canonical distribution, we argued its equivalence to the microcanonical measure, and therefore do not find an assumption of ergodicity surprising.

Indeed, at small values of  $\kappa$ , we find numerical evidence for ergodicity, and agreement between the mean transition time found from a longtime average over a single trajectory and the time (14) as shown in Fig. 3b. In this semilog plot, the slope of the lines is  $\Delta E_+$  (for Neumann boundary conditions, this is solely  $E^s$ , the value of which is show in Fig. 1 as a function of  $\kappa$ ), while the y-intercept is given by the prefactor,

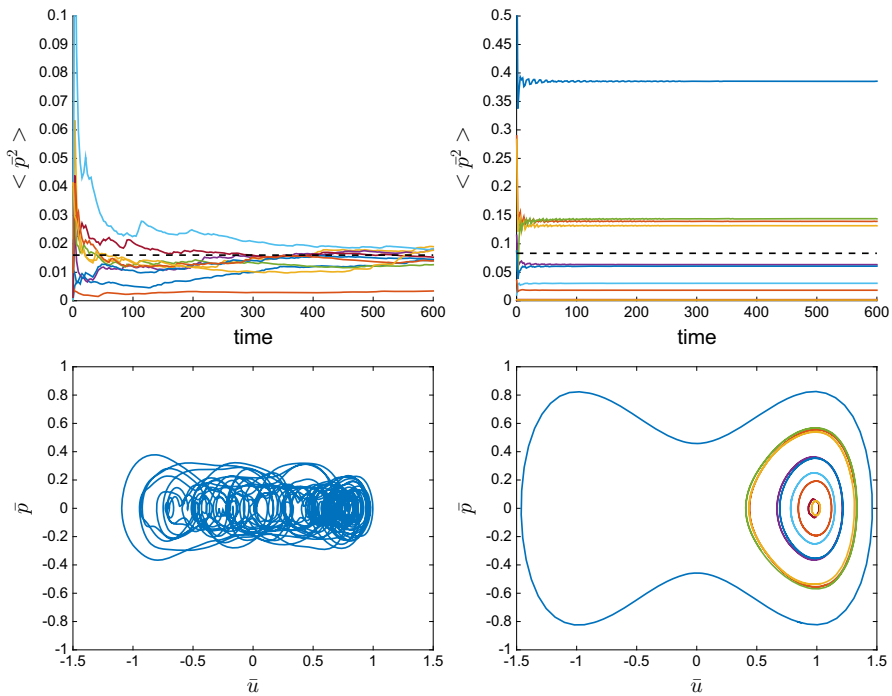


**Fig. 3** **a** Log of the mean transition time as a function of  $\beta$  (symbols) averaged over an ensemble of 1000 (10,000 for larger  $\beta$ ) trajectories chosen according to (56), evolved by numerically simulating (18), with potential function (20) and  $N = 128$  and (lines) the theoretical value of  $\tau_+$  given by (14). **b** Log of the mean transition time (symbols) computed from a longtime average of a single trajectory chosen according to (56) and (lines) the theoretical value of  $\tau_+$  given by (14). **c** The prefactor  $\Lambda_+$  in (15) computed numerically with  $N = 1024$  points as a function of  $\kappa$ . As the bifurcation point  $\kappa = 1/\pi$  is approached, the prefactor goes to zero because one of the eigenvalues of the saddle point is about to cross zero.

$2\pi \Lambda_+$ . The eigenvalue product  $\Lambda_+$  is shown in Fig. 3c as a function of  $\kappa$ , computed with a truncation of  $N = 1024$ .

However, at larger values of  $\kappa$ , we do not see support for ergodicity over the finite simulation time which is orders of magnitude greater than the expected mean transition time. This point is emphasized in the top row of plots in Fig. 4, showing that longtime averages seem to be independent of initial conditions for small  $\kappa$  (top left) yet remain strongly dependent on initial conditions for times much longer than an expected mean transition time for large  $\kappa$  (top right).

This loss of ergodicity can be explained by a fundamental change in the dynamics as  $\kappa$  increases. Not only does increasing  $\kappa$  reduce the size of the fluctuations in the solution  $u(x, t)$ , but it also makes this solution more uniform in  $x$ . This is evident from the energy penalty proportional to  $\kappa^2$  on the  $x$ -derivatives in the Hamiltonian (3), which also lead to the loss of the non-trivial saddle point for  $\kappa > 1/\pi$ . The solutions to (2) at each instance in time become approximated by small fluctuations in space about the quantity  $\bar{u}(t) = \int_0^1 u(x, t) dx$ .



**Fig. 4** Results of numerical simulations of the finite-dimensional system (18) with  $N = 1024$ . Simulations starting from 9 different initial conditions, selected according to (56), with  $\beta \Delta E = 3$ , (left)  $\kappa = 0.05$  (right)  $\kappa = 1$ , were run until  $t = 600$  (much larger than the mean transition time). (Top) To emphasize ergodicity in the case of small  $\kappa$ , the time-averaged value of  $\bar{p}^2 = \sum_{j=1}^N p_j^2/N$  is plotted, and its value of  $1/\beta$  obtained from averaging over the invariant measure (6) is shown as a dotted line. (Bottom) Dynamics of the same numerical simulations plotted in terms of the reduced variables  $\bar{u} = \sum_{j=1}^N u_j/N$  and  $\bar{p} = \sum_{j=1}^N p_j/N$ . (bottom left) Single trajectory, (bottom right) all 9 trajectories

These effects can be quantified by starting from a large  $\beta$  expansion for  $u(x, t)$ , the infinite-dimensional version of what was presented in Sect. 4.3. Taking

$$u(x, t) = u^+(x) + \frac{1}{\sqrt{\beta}} \tilde{u}(x, t) \tag{57}$$

and expanding the Hamiltonian to quadratic order in  $\beta^{-1/2}$ , the resulting canonical distribution describes  $\tilde{u}$  as a Gaussian process with covariance function

$$C(x, y) = \begin{cases} \frac{(e^{x/\kappa} + e^{-x/\kappa})(e^{(1-y)/\kappa} + e^{-(1-y)/\kappa})}{2\kappa(e^{1/\kappa} - e^{-1/\kappa})} & x \leq y \\ \frac{(e^{y/\kappa} + e^{-y/\kappa})(e^{(1-x)/\kappa} + e^{-(1-x)/\kappa})}{2\kappa(e^{1/\kappa} - e^{-1/\kappa})} & x \geq y \end{cases}, \tag{58}$$

derived in “Example Covariance Functions” section of Appendix 1. The uniformity of  $\tilde{u}(x)$  in  $x$  is apparent from the  $\kappa^{-1}$  exponential dependence indicating the covariance between  $\tilde{u}(x)$  at two different points decays slower as  $\kappa$  increases, resulting in a more uniform description. More precisely, we can write  $u(x)$  selected from the invariant measure at each moment of time approximately as a sum of independent Gaussian random variables,  $g_n$ ,

$$u(x) \approx \pm 1 + \frac{1}{\sqrt{\beta}} \left[ g_0 + \sum_{n=1}^{\infty} \frac{\cos(n\pi x)}{\sqrt{1 + \kappa^2 n^2 \pi^2}} g_n \right] \tag{59}$$

in which  $\cos(n\pi x)$  and  $1/(1 + \kappa^2 n^2 \pi^2)$  for  $n = 0, 1, 2 \dots$  are the eigenfunctions and eigenvalues of the covariance  $C(x, y)$  in (58). For small  $\kappa$ , a number of the  $n \geq 1$  terms are significant, while for large  $\kappa$  we see that only the constant zeroth order term is significant.

Another way to interpret this large  $\kappa$  behavior is that only the first constant mode of the solution contributes significantly to the dynamics, namely

$$\bar{u}(t) = \int_0^1 u(x, t) dx \quad \text{and} \quad \bar{p}(t) = \int_0^1 p(x, t) dx.$$

The dynamics of these reduced variables are shown in the bottom row of Fig. 4. For small  $\kappa$  in the bottom left plot, a solution trajectory explores the phase space, while for large  $\kappa$  in the bottom right plot, each trajectory with a different randomly chosen initial condition appears to nearly conserve the reduced Hamiltonian,

$$\bar{H} = \frac{1}{2} \bar{p}^2 + V(\bar{u}),$$

violating the assumption of ergodicity with respect to the canonical measure with the original Hamiltonian. The conservation of the reduced Hamiltonian in the limit of large  $\kappa$  can be understood by computing the time evolution of the quantity. Namely,

$$\begin{aligned} \frac{d\bar{H}}{dt} &= \bar{p} \int_0^1 \frac{dp}{dt} dx + V'(\bar{u}) \int_0^1 \frac{du}{dt} dx \\ &= \bar{p} \int_0^1 (\kappa^2 u_{xx} - V'(u)) dx + V'(\bar{u}) \bar{p} = O(\kappa^{-1}) \end{aligned}$$

in which the last equality holds since  $\int_0^1 \kappa^2 u_{xx} dx = 0$  by the Neumann boundary conditions, and  $V'(u) = V'(\bar{u}) + O(\kappa^{-1})$  when  $u(x)$  is nearly uniform by the expansion in (59).

### 6 Conclusions

We have investigated the dynamics of the nonlinear wave equation (1), with particular interest in metastability and mean transition times. We discussed how the finite-

dimensional microcanonical distribution, with energy scaling extrinsically as  $N/\beta$ , is equivalent to the canonical distribution as  $N \rightarrow \infty$ . By revisiting TST for an ensemble of trajectories chosen with the canonical invariant measure of the finite-dimensional Hamiltonian system, we derived the mean frequency at which the solutions cross a dividing surface separating the two minimizers of the Hamiltonian. Extension to the PDE case required we justify the use of a quadratic approximation of the Hamiltonian in infinite dimensions by showing concentration of the canonical invariant measure for arbitrarily large  $N$ . Numerical simulations of the finite-dimensional Hamiltonian system supported that the TST time applied for almost all initial conditions only if  $\kappa$  was small.

At large  $\kappa$ , even though the system could still be ergodic with respect to the canonical invariant measure, the results of numerical simulations in Sect. 5 suggest that the mixing time to observe convergence of longtime averages becomes much longer than the residency times in the metastable sets. This increase in the mixing time would cause correlations between successive transition. At small values of  $\kappa$ , the mixing time appears to be much shorter than the computed mean residency time with TST, suggesting that the metastable dynamics could be approximately coarse-grained onto that of a two-state Markov jump process for the two localized regions around the minimizers. Note also that our observations (data not shown) suggest a slow increasing of the mixing time, and no abrupt change in behavior from large to small system size.

The simulated dynamics of the deterministic nonlinear wave equation also provides interesting insight into the low-damping,  $\gamma \rightarrow 0$ , regime of the Langevin equation (5). The mean transition time computed with large deviation theory (LDT) in (17) is equivalent to the TST time (14) in the limit as  $\gamma \rightarrow 0$ . The implied equivalence of trajectory-wise dynamics is supported only by the numerical simulations of the deterministic system in the small  $\kappa$  regime, in which the mean transition time computed over a single trajectory is well described by the TST time. In this regime, both the deterministic equation (2) and its stochastic counterpart (5) have equivalent transition times.

This is in contrast to the large  $\kappa$  regime in which the numerical simulations of the deterministic equation effectively behave like those of a one-dimensional Hamiltonian system, approximately conserving a reduced energy. This illuminates what we might expect from the low-damping regime of the Langevin equation (5). It should behave like a one-dimensional noisy-damped system, with values of reduced energy that are often less than the value of the saddle point energy and slowly diffusing relative to the timescale of a typical Hamiltonian orbit. We would therefore expect the method of averaging (Freidlin and Wentzell 1994), applied to the reduced variables rather than the full system, to characterize the motion of the reduced energy. Computing the transition times for the reduced energy as we did in Newhall and Vanden-Eijnden (2013) would produce a prefactor, inversely proportional to the damping, on the exponential transition time that is much larger than the lower bound computed here with TST. We hope to pursue this further in future work.

We end with a final remark that the derivation of the TST time cannot be extended to the wave equation in multiple ( $D \geq 2$ ) spatial dimensions. While the system can still be discretized to form a finite-dimensional system, the limit toward the continuous system breaks down. We cannot justify metastability through the concentration of

the invariant measure, and the prefactor containing the eigenvalues in (14) fails to converge. As mentioned in the introduction, these problems stem from the scaling of the eigenvalues of the Laplacian in multiple spatial dimensions, and are reminiscent of known problems with solutions to stochastic partial differential equations like (5) when  $D \geq 2$  (Hairer 2014; Ryser et al. 2012).

**Acknowledgements** We would like to thank Weinan E and Gregor Kovačič for useful discussions. This work was supported in part by the NSF Grant DMS-1522767 (E. V.-E.).

### Appendix 1: Infinite-Dimensional Microcanonical Distribution

In this appendix, we prove proposition 1. We show the limit of the microcanonical distribution as  $N \rightarrow \infty$ , with energy given by  $E_N = N/\beta$ , leads to the invariant measure written formally in (6) by calculating the characteristic function of both  $u(x, t)$  and  $p(x, t) = u_t(x, t)$ . We do this exactly for a quadratic Hamiltonian, with  $V(u(x)) = \frac{1}{2}f(x)u^2(x)$ , and explicitly solve for the covariance function of the continuous stochastic process  $u(x, t)$  in three special cases;  $u_t(x, t)$  is always white noise.

We begin with the discrete microcanonical distribution (21) with energy  $E_N$  for a finite-dimensional quadratic Hamiltonian,

$$d\mu_N = c_N^{-1} \delta \left( E_N - \frac{1}{2N} \mathbf{p}^T \mathbf{p} - \frac{1}{2} \mathbf{u}^T A_N \mathbf{u} \right) d\mathbf{p} d\mathbf{u} \tag{60}$$

where the  $N \times N$  matrix  $A_N$  is such that  $\mathbf{u}^T A_N \mathbf{u} = \frac{1}{N} \sum_{j=0}^N \kappa^2 N^2 (u_{j+1} - u_j)^2 + f(x_j)u_j^2$  with  $u_0$  and  $u_{N+1}$  set by the boundary conditions of (1). As  $N \rightarrow \infty$ , we have  $\mathbf{u}^T A_N \mathbf{u} \rightarrow \int_0^1 u(x) \left( -\kappa^2 \frac{d^2}{dx^2} + f(x) \right) u(x) dx$ . Note the matrix  $A_N$  is symmetric in the case of either Dirichlet or Neumann boundary conditions.

We consider the characteristic function for the  $\{u_j\}$  first, finding it converges to

$$\phi_u(\mathbf{t}) \rightarrow \exp \left( -\frac{1}{2\beta} \int_0^1 \int_0^1 t(x) C(x, y) t(y) dx dy \right), \tag{61}$$

a Gaussian process with covariance  $C(x, y)/\beta$ . The covariance function is obtained by solving

$$\left( \kappa^2 \frac{d^2}{dx^2} + f(x) \right) C(x, y) = \delta(x - y)$$

with appropriate boundary conditions. A few special cases are given in ‘‘Example Covariance Functions’’ section. Then we follow a similar procedure for the  $\{p_j\}$  in ‘‘The  $\mathbf{p}$ ’’ section, finding the characteristic function converges to

$$\phi_p(\mathbf{s}) \rightarrow \exp \left( -\frac{1}{2\beta} \int_0^1 s(x)^2 dx \right), \tag{62}$$

a Gaussian process with covariance  $\delta(x - y)/\beta$  (i.e., white noise). These are combined to arrive at proposition 1.

### The $u$

The characteristic function for the sets of points  $u$  selected from the microcanonical distribution (60) is defined as

$$\phi_u(t) = \frac{\int_{\mathbb{R}^{2N}} e^{iNt^T u} \delta\left(E_N - \frac{1}{2N} p^T p - \frac{1}{2} u^T A_N u\right) dp du}{\int_{\mathbb{R}^{2N}} \delta\left(E_N - \frac{1}{2N} p^T p - \frac{1}{2} u^T A_N u\right) dp du}. \quad (63)$$

The integral in the denominator of (63) is determined first by switching to the coordinates  $q = A_N^{1/2} u$  and rescaling  $p$  to be  $p/\sqrt{N}$ . This results in the integral

$$\frac{N^{N/2}}{\det A_N^{1/2}} \int_{\mathbb{R}^{2N}} \delta\left(E_N - \frac{1}{2} p^T p - \frac{1}{2} q^T q\right) dp dq$$

where  $\det A_N^{1/2} = \prod_{j=1}^N \sqrt{\lambda_j}$  and the  $\lambda_j$  are the eigenvalues of the matrix  $A_N$ . Then, switching to spherical coordinates,  $r^2 = \sum_{j=1}^N p_j^2$ , we have

$$\frac{N^{N/2}}{\det A_N^{1/2}} S_{N-1} \int_{\mathbb{R}^N} \int_0^\infty \delta\left(E_N - \frac{1}{2} q^T q - \frac{r^2}{2}\right) r^{N-1} dr dq$$

where  $S_{N-1} = N\pi^{N/2}/\Gamma(N/2+1)$  is the surface area of the sphere in  $N$ -dimensional space, and integrating over  $r$ , we obtain

$$\frac{N^{N/2}}{\det A_N^{1/2}} S_{N-1} \int_{\mathbb{R}^N} \left(E_N - \frac{1}{2} q^T q\right)_+^{N/2-1} dq,$$

where  $x_+ = \max(0, x)$ . Last, we switch to spherical coordinates for the  $q$ ,

$$\frac{N^{N/2}}{\det A_N^{1/2}} (S_{N-1})^2 \int_0^{\sqrt{2E_N}} \left(E_N - \frac{r^2}{2}\right)^{N/2-1} r^{N-1} dr$$

which is equivalent to

$$\frac{N^{N/2}}{\det A_N^{1/2}} (S_{N-1})^2 E_N^{N-1} \frac{\sqrt{\pi} \Gamma(N/2)}{2^{N/2} \Gamma(N/2 + 1/2)}. \quad (64)$$



For the numerator of (63), we again switch to the coordinates  $\mathbf{q} = A_N^{1/2} \mathbf{u}$  and rescale the  $\mathbf{p}$  to  $\mathbf{p}/\sqrt{N}$ , obtaining

$$\frac{N^{N/2}}{\det A_N^{1/2}} \int_{\mathbb{R}^{2N}} e^{i\mathbf{t}^T A_N^{-1/2} \mathbf{q}} \delta \left( E_N - \frac{1}{2} \mathbf{p}^T \mathbf{p} - \frac{1}{2} \mathbf{q}^T \mathbf{q} \right) d\mathbf{p} d\mathbf{q}.$$

Switching the  $\mathbf{p}$  to spherical coordinates and integrating yields

$$\frac{N^{N/2}}{\det A_N^{1/2}} S_{N-1} \int_{\mathbb{R}^N} e^{i\mathbf{t}^T A_N^{-1/2} \mathbf{q}} \left( E_N - \frac{1}{2} \mathbf{q}^T \mathbf{q} \right)_+^{N/2-1} d\mathbf{q}.$$

To determine the integral over  $\mathbf{q}$ , we write the integral in spherical coordinates over all  $\hat{\mathbf{n}} \in \Omega$ , the unit normal to the  $N$ -dimensional sphere, and all  $r$ ,

$$\frac{N^{N/2}}{\det A_N^{1/2}} S_{N-1} \int_0^{\sqrt{2E_N}} \int_{\Omega} e^{i\mathbf{t}^T A_N^{-1/2} \hat{\mathbf{n}} r} \left( E_N - \frac{r^2}{2} \right)^{N/2-1} r^{N-1} d\hat{\mathbf{n}} dr. \tag{65}$$

The appearance of the  $\hat{\mathbf{n}}$  in the exponent does not allow us to integrate exactly. Before proceeding, we return to our original expression (63) by taking the ratio of (65) to (64), resulting in

$$\phi_{\mathbf{u}}(\mathbf{t}) = \frac{2^{N/2} \Gamma(N/2 + 1/2)}{E_N^{N-1} \sqrt{\pi} \Gamma(N/2)} \frac{1}{S_{N-1}} \int_0^{\sqrt{2E_N}} \int_{\Omega} e^{i\mathbf{t}^T A_N^{1/2} \hat{\mathbf{n}} r} \left( E_N - \frac{r^2}{2} \right)^{N/2-1} r^{N-1} d\hat{\mathbf{n}} dr. \tag{66}$$

In (66), the integral over  $\hat{\mathbf{n}}$  divided by the surface area can be thought of as an expectation with respect to the random variable  $\hat{\mathbf{n}}$ , the unit vector chosen uniformly over the surface of the unit sphere in  $N$ -dimensional space,

$$\frac{1}{S_{N-1}} \int_{\Omega} e^{i\mathbf{t}^T A_N^{1/2} \hat{\mathbf{n}} r} d\hat{\mathbf{n}} = \mathbb{E} \left[ e^{i\mathbf{t}^T A_N^{-1/2} \hat{\mathbf{n}} r} \right].$$

Alternatively, we can express the random variable  $\hat{\mathbf{n}}$  as  $\mathbf{g}/|\mathbf{g}|$  where  $\mathbf{g}$  is an  $N$ -dimensional vector with components that are independent identically distributed mean-zero Gaussian random variables. Intuitively, for large  $N$ , this random variable can be approximated by  $\mathbf{g}/\sqrt{N}$  (noticed by Poincaré 1912, see also McKean 1973). More specifically, in the limit as  $N \rightarrow \infty$ , we have that the expectation over  $\hat{\mathbf{n}}$  scales as

$$\mathbb{E} \left[ \exp \left( \frac{i}{N} \mathbf{t}^T A_N^{-1/2} \hat{\mathbf{n}} r \right) \right] \sim \mathbb{E} \left[ \exp \left( \frac{i}{N^{3/2}} \mathbf{t}^T A_N^{-1/2} \mathbf{g} r \right) \right]$$

where we use  $\sim$  to indicate that the ratio of the two sides approaches one as  $N \rightarrow \infty$ . This expectation on the right-hand side is the characteristic function of a mean-zero multivariate Gaussian; it is known explicitly. Combining the above results, we have

$$\frac{1}{S_N} \int_{\Omega} e^{\frac{i}{N} \mathbf{t}^T A_N^{1/2} \hat{\mathbf{u}} r} d\hat{\mathbf{u}} \sim \exp\left(-\frac{1}{2} \frac{r^2}{N^3} \mathbf{t}^T A_N^{-1} \mathbf{t}\right) \tag{67}$$

in the limit of large  $N$ .

Then, inserting (67) into (66) for the integration over  $\hat{\mathbf{u}}$ , we are left with the following integral,

$$\int_0^{\sqrt{2E_N}} e^{-\frac{1}{2} \frac{r^2}{N^3} \mathbf{t}^T A_N^{-1} \mathbf{t}} \left(E_N - \frac{r^2}{2}\right)^{N/2-1} r^{N-1} dr. \tag{68}$$

We would like to use Laplace’s method,

$$\int_a^b g(x) e^{Nf(x)} dx \sim \sqrt{\frac{2\pi}{N|f''(x_0)|}} g(x_0) e^{Nf(x_0)}$$

if the integral in (68) has the proper scalings as  $N \rightarrow \infty$ . We notice the scaling of the exponent in (68) is  $r^2/N + O(1)$  as  $\mathbf{t}^T A_N^{-1} \mathbf{t}/N^2$  converges to a double integral,

$$\frac{1}{N^2} \mathbf{t}^T A_N^{-1} \mathbf{t} \rightarrow \int_0^1 \int_0^1 t(x) C(x, y) t(y) dx dy$$

where the function  $C(x, y)$  is obtained by inverting the operator  $A_{\infty}$ , defined such that  $\mathbf{u}^T A_N \mathbf{u} \rightarrow \int_0^1 u(x) A_{\infty}(x) u(x) dx$ . This exponential term in (68) remains smaller than  $e^N$  as  $N \rightarrow \infty$  provided the maximum value of  $r$  does not grow faster than  $O(\sqrt{N})$ .

The exponential in  $N$  component comes from writing

$$\left(E_N - \frac{r^2}{2}\right)^{N/2-1} r^{N-1} \sim \exp\left(\frac{N}{2} \log\left(E_N - \frac{r^2}{2}\right) + N \log r\right)$$

to leading order in  $N$ . We see that this exponent attains its maximum between  $r = 0$  and  $\sqrt{2E_N}$  when  $r = \sqrt{E_N}$ . This corresponds to half of the total energy,  $E_N$ , contained within the variables  $\mathbf{u}$ ; the equipartition of energy between the  $\mathbf{p}$  and the  $\mathbf{u}$ .

Now, we define the value of the energy,  $E_N = N/\beta$ , as this is where the canonical measure energy concentrates, as discussed in the main text. Even with this energy scaling with  $N$ , and therefore the maximum value of  $r = \sqrt{N/\beta}$ , the exponential term in (68) remains smaller than  $e^N$ . Applying Laplace’s method to the integral in (68), with energy  $E_N = N/\beta$ , we obtain

$$\begin{aligned} &\int_0^{\sqrt{2N/\beta}} e^{-\frac{1}{2} \frac{r^2}{N^3} \mathbf{t}^T A_N^{-1} \mathbf{t}} \left(\frac{N}{\beta} - \frac{r^2}{2}\right)^{N/2-1} r^{N-1} dr \\ &\sim \left(\frac{N}{\beta}\right)^{N-1} 2^{-N/2+1} \sqrt{\frac{2\pi}{N4}} \exp\left(-\frac{1}{2} \frac{1}{\beta N^2} \mathbf{t}^T A_N^{-1} \mathbf{t}\right) \end{aligned}$$

as  $N \rightarrow \infty$ . Combining this result together with (67) into (66) also with  $E_N = N/\beta$ , we find after simplifying that

$$\phi_u(\mathbf{t}) \sim \frac{\Gamma(N/2 + 1/2)}{\Gamma(N/2)} \sqrt{\frac{2}{N}} \exp\left(-\frac{1}{2\beta} \frac{1}{N^2} \mathbf{t}^T A_N^{-1} \mathbf{t}\right).$$

In the limit as  $N \rightarrow \infty$ , the prefactor in this expression goes to unity (the true value of the integral of  $(1 - r^2/2)^{N/2-1} r^{N-1}$  and its approximation are asymptotically equivalent), leaving

$$\phi_u(\mathbf{t}) \rightarrow \exp\left(-\frac{1}{2\beta} \int_0^1 \int_0^1 t(x)C(x, y)t(y)dx dy\right). \tag{69}$$

The stochastic process  $u(x)$  has the covariance function  $C(x, y)/\beta$ . The covariance function  $C(x, y)$  is found by solving

$$A_\infty C(x, y) = \delta(x - y) \tag{70}$$

with appropriate boundary conditions. We consider special cases of  $A_\infty$  in the next section before proceeding to consider the characteristic function of the  $\mathbf{p}$ .

**Example Covariance Functions**

Analytic solutions to (70) can be obtained in a few special cases. Of particular interest is the case of the quadratic function  $V(u) = \frac{1}{2}u^2$ , which can be used to approximate the nonlinear potential function  $V(u) = \frac{1}{4}(1 - u^2)^2$  near its minimizers.

For the case of dynamics governed by (2) with  $V(u) = \frac{1}{2}u^2$  and Dirichlet boundary conditions,  $u(0, t) = u(1, t) = 0$ , inverting the operator  $A_\infty$  amounts to solving

$$\left(-\kappa^2 \frac{\partial^2}{\partial x^2} + 1\right) C(x, y) = \delta(x - y) \tag{71}$$

with boundary conditions  $C(0, y) = C(1, y) = 0$  and continuity conditions at  $x = y$ . The solution is

$$C(x, y) = \begin{cases} \frac{(e^{x/\kappa} - e^{-x/\kappa})(e^{(1-y)/\kappa} - e^{-(1-y)/\kappa})}{2\kappa(e^{1/\kappa} - e^{-1/\kappa})} & x \leq y \\ \frac{(e^{y/\kappa} - e^{-y/\kappa})(e^{(1-x)/\kappa} - e^{-(1-x)/\kappa})}{2\kappa(e^{1/\kappa} - e^{-1/\kappa})} & x \geq y \end{cases} \tag{72}$$

This covariance is nearly constant along a large portion of  $x = y$ , indicating the solutions look something like an Ornstein Uhlenbeck process, but conditioned to have  $u(0, t) = u(1, t) = 0$ .

When the boundary conditions are changed to Neumann boundary conditions, the solution to (71) is

$$C(x, y) = \begin{cases} \frac{(e^{x/\kappa} + e^{-x/\kappa})(e^{(1-y)/\kappa} + e^{-(1-y)/\kappa})}{2\kappa(e^{1/\kappa} - e^{-1/\kappa})} & x \leq y \\ \frac{(e^{y/\kappa} + e^{-y/\kappa})(e^{(1-x)/\kappa} + e^{-(1-x)/\kappa})}{2\kappa(e^{1/\kappa} - e^{-1/\kappa})} & x \geq y \end{cases} \tag{73}$$

To understand what these covariance functions describe, consider no potential function,  $V(u) = 0$ , and Dirichlet boundary conditions. The covariance is given by the solution to

$$-\kappa^2 \frac{\partial^2}{\partial x^2} C(x, y) = \delta(x - y)$$

with boundary conditions  $C(0, y) = C(1, y) = 0$ , which is

$$C(x, y) = \frac{\min(x, y) - xy}{\kappa^2}, \tag{74}$$

the covariance function of a Brownian Bridge on the unit interval.

**The  $p$**

We can follow the exact same procedure to compute the characteristic function for the  $p$ ,

$$\phi_p(s) = \frac{\int e^{i s^T p} \delta \left( E_N - \frac{1}{2N} p^T p - \frac{1}{2} u^T A_N u \right) dp du}{\int \delta \left( E_N - \frac{1}{2N} p^T p - \frac{1}{2} u^T A_N u \right) dp du}. \tag{75}$$

Much of the above derivation follows directly, and we briefly detail the steps. By rescaling  $p$  to be  $p/\sqrt{N}$  and  $q = A_N^{1/2} u$ , we have

$$\phi_p(s) = \frac{\int e^{i s^T p} \delta \left( E_N - \frac{1}{2} p^T p - \frac{1}{2} q^T q \right) dp dq}{\int \delta \left( E_N - \frac{1}{2} p^T p - \frac{1}{2} q^T q \right) dp dq}.$$

switching to spherical coordinates and replacing  $\hat{n}$  with the scaling  $g/\sqrt{N}$  within the expectation results in

$$\phi_p(s) = \frac{\int_0^{\sqrt{2E_N}} \mathbb{E} \left[ e^{\frac{i}{\sqrt{N}} s^T g r} \right] \left( E_N - \frac{r^2}{2} \right)_+^{N/2-1} r^{N-1} dr}{\int_0^{\sqrt{2E_N}} \left( E_N - \frac{r^2}{2} \right)_+^{N/2-1} r^{N-1} dr}.$$

Using that the expectation

$$\mathbb{E} \left[ \exp \left( \frac{i}{N} s^T g r \right) \right] = \exp \left( -\frac{r^2}{2N^2} s^T s \right)$$

and using Laplace’s method for the integral over  $r$  with the scaling of  $E_N = N/\beta$ , we arrive at

$$\phi_p(s) \rightarrow \exp \left( -\frac{1}{2\beta} \int_0^1 s(x)^2 dx \right) \tag{76}$$

in the limit as  $N \rightarrow \infty$ . (Note that  $\frac{1}{N} s^T s \rightarrow \int_0^1 |s(x)|^2 dx$ ). Thus we see that  $p(x)$  is white noise with covariance function  $\delta(x - y)/\beta$ .

### Appendix 2: Quadratic Approximation

In this appendix, we prove Propositions 2 and 3. We show in “Concentration of the Canonical Measure” section that an arbitrarily large amount of the invariant measure (6) can be contained within an arbitrarily small region around a minimizer of the Hamiltonian for large enough values of  $\beta$  by developing a uniform in  $N$  concentration of the discrete quadratic canonical measure. This allows for both the argument that the solutions to (2) display metastability and the argument that the potential function  $V(u)$  can be replaced with a quadratic approximation in order to complete the integration in determining the time between transitions of the solutions to (2). “Evaluation of the Integrals” section contains the details of computing the integrals to arrive at Proposition 3 and Conjecture 1, the counterpart to Proposition 3 for the microcanonical measure found in “Appendix 3.”

### Concentration of the Canonical Measure

In the following, we show the concentration of the marginal canonical measure for  $\mathbf{b}^\pm = (b_1^\pm, \dots, b_N^\pm)$  as stated in Proposition 2 in the main text, and also for the marginal canonical measure restricted to the surface  $S = \{(\tilde{\mathbf{p}}, \mathbf{a}) : a_1 = 0\}$  in the eigenvector basis, stated in Proposition 4 below.

We first point out that the constraint on  $\sum_{j=1}^N \delta_j^2$  is to ensure the corner of the box remains a finite distance away from the center of the box regardless of the value of  $N$ , in particular as  $N \rightarrow \infty$ . Therefore, the measure is being contained within a localized region. In particular, the sides of the box must decay faster than  $1/\sqrt{j}$ ,

$$\delta_j^2 \sim \frac{C}{j^{1+\epsilon}}, \tag{77}$$

for some  $\epsilon > 0$  and  $C > 0$ .

We drop the  $\pm$  and proceed to calculate

$$\frac{\int_{D_\delta^N} e^{-\frac{\beta}{2} \sum_{j=1}^N \lambda_j b_j^2} d\mathbf{b}}{\int_{\mathbb{R}^N} e^{-\frac{\beta}{2} \sum_{j=1}^N \lambda_j b_j^2} d\mathbf{b}} \tag{78}$$

by squaring everything so each integral over  $b_j$  becomes an integral over  $b_j$  and  $a_j$ :

$$\frac{\int_{-\delta_j}^{\delta_j} \int_{-\delta_j}^{\delta_j} e^{-\frac{\beta}{2} (\lambda_j b_j^2 + \lambda_j a_j^2)} db_j da_j}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{\beta}{2} (\lambda_j b_j^2 + \lambda_j a_j^2)} db_j da_j}$$

Then, rescaling by the eigenvalues and noting the integral over the 2D square with sides  $2\delta_j\sqrt{\lambda_j}$  is greater than the integral over the circle with radius  $\delta_j\sqrt{\lambda_j}$ , gives

$$\frac{\int_{-\delta_j\sqrt{\lambda_j}}^{\delta_j\sqrt{\lambda_j}} \int_{-\delta_j\sqrt{\lambda_j}}^{\delta_j\sqrt{\lambda_j}} e^{-\frac{\beta}{2} (b_j^2 + a_j^2)} db_j da_j}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{\beta}{2} (b_j^2 + a_j^2)} db_j da_j} > \frac{\int_0^{\delta_j\sqrt{\lambda_j}} e^{-\frac{\beta}{2} r^2} r dr}{\int_0^{\infty} e^{-\frac{\beta}{2} r^2} r dr} = \left(1 - e^{-\frac{\beta}{2} \delta_j^2 \lambda_j}\right)$$

for each  $j = 1 \dots N$ .

Combining all the integrals, which are identical, we are left with showing that there exists a  $\beta_1$  such that

$$\prod_{j=1}^N \left(1 - \exp\left(-\frac{\beta}{2} \delta_j^2 \lambda_j\right)\right) > (1 - \hat{\delta})^2 \tag{79}$$

can be satisfied for all  $N > 0$  and for all  $\beta > \beta_1$ .

The product in (79) converges if the sum

$$\sum_{j=1}^N \log\left(1 - \exp\left(-\frac{\beta}{2} \delta_j^2 \lambda_j\right)\right) \tag{80}$$

converges as  $N \rightarrow \infty$ . We can compare the decay of the terms in this sum to the decay of the terms in a known converging series,  $\sum_{j=1}^{\infty} j^{-p}$  when  $p > 1$ . We note the terms in (80) are negative, so we add a negative sign and compare. We are interested if

$$\lim_{j \rightarrow \infty} \frac{-\log\left(1 - e^{-\frac{\beta}{2} \delta_j^2 \lambda_j}\right)}{1/j^p} < \infty \tag{81}$$

Taking the derivative of top and bottom yields

$$\lim_{j \rightarrow \infty} \frac{j^{p+1} \frac{\beta}{2} \delta_j^2 \lambda_j e^{-\frac{\beta}{2} \delta_j^2 \lambda_j}}{p \left( 1 - e^{-\frac{\beta}{2} \delta_j^2 \lambda_j} \right)} = 0, \tag{82}$$

provided the exponent in the numerator decays to zero,

$$\delta_j^2 \lambda_j \rightarrow \infty \tag{83}$$

as  $j \rightarrow \infty$ .

We are left to consider whether there exists a box with sides  $2\delta_j$  that satisfies both (77) and (83) simultaneously. This depends on the values of the  $\lambda_j$ . For our application, they are the eigenvalues of the operator (16). They are dominated by the Laplacian term, and therefore for large  $j$  scale like

$$\lambda_j \sim j^2. \tag{84}$$

From (77), we have that  $\epsilon > 0$  and from combining (77) and (84) into (83), we have that  $1 - \epsilon > 0$ . A range of values of  $\epsilon$  are possible, for example,  $\epsilon = 1/2$ . Thus it is possible to contain the measure within a finite region around a minimizer of the Hamiltonian. With a choose of  $\epsilon$ , the value of  $\beta_1$  is determined as the one that satisfies (79) when  $N \rightarrow \infty$  (the product is decreasing as  $N$  increases).

The integration for the TST frequency also requires the following proposition concerning the concentration of the measure restricted to the surface  $S$ , given by  $a_1 = 0$  in the eigenvector coordinates:

**Proposition 4** *Consider the marginal canonical measure restricted to the surface  $S$  defined in (37) with the quadratic Hamiltonian approximation (50) and an  $N - 1$ -dimensional box with side edge lengths  $2\delta_j$ ,*

$$\bar{D}_\delta^{N-1} = \{ \mathbf{a} : a_1 = 0, -\delta_j \leq a_j \leq \delta_j \quad \forall j = 2 \dots N \},$$

satisfying  $\sum_{j=2}^N \delta_j^2$  remaining finite as  $N \rightarrow \infty$ . Assume that the eigenvalues  $\lambda_j^s$  solving (40) satisfy

$$\lambda_j^s \sim Cj^2 \quad \text{as } j \rightarrow \infty, \text{ for some } C > 0$$

Then, for every  $\hat{\delta} > 0$ , there exists  $\beta_1 > 0$  and a set of  $\{\delta_j\}_{j=2}^N$  such that for every  $\beta > \beta_1$  and every  $N > 0$

$$\frac{\int_{\bar{D}_\delta^{N-1}} \exp\left(-\frac{\beta}{2} \sum_{j=2}^N \lambda_j^s a_j^2\right) da_2 \dots da_N}{\int_{\mathbb{R}^{N-1}} \exp\left(-\frac{\beta}{2} \sum_{j=2}^N \lambda_j^s a_j^2\right) da_2 \dots da_N} > 1 - \hat{\delta} \tag{85}$$

The proof of Proposition 4 follows exactly, replacing all summations above from  $j = 2 \dots N$ . In particular, the expression in (80) becomes

$$\sum_{j=2}^N \log \left( 1 - \exp \left( -\frac{\beta}{2} \delta_j^2 \lambda_j^s \right) \right) \tag{86}$$

and since the eigenvalues  $\lambda_j^s$  have the same scaling as (84), we arrive at the same conclusion on the convergence of (86) and therefore the concentration of the measure.

**Evaluation of the Integrals**

Having justified the use of a quadratic approximation as  $\beta \rightarrow \infty$  for the canonical invariant measure, we compute the integration in (93) or (36) required to compute the TST frequencies. First, we want to rewrite the integrals in the eigenvector basis. The normal  $\hat{n}(\mathbf{u}) \in \mathbb{R}^N$  is given by  $(1, 0, \dots, 0)$  and therefore  $\mathbf{p} \cdot \hat{n}(\mathbf{u}) = \tilde{p}_1$ . Changing the integration variables to the eigenvector basis is more than just a rotation, as we defined the eigenvector basis to be normalized to  $N$  rather than one. This results in the integrals in (93) and (36) gaining a factor of  $N^{(N-1)/2}$ , a factor of  $\sqrt{N}$  for each of the  $j = 2 \dots N$  eigenvector coordinates. Together with the asymptotic in  $\beta$  expansion of the measures, we have from (93)

$$\begin{aligned} v_S^m &\sim c_N^{-1} N^{(N-1)/2} \int_{\mathbb{R}^{N-1}} \max(\tilde{p}_1, 0) \\ &\times \delta \left( E_N - E_N^s - \sum_{j=1}^N \frac{\tilde{p}_j^2}{2N} - \sum_{j=2}^N \frac{\lambda_j^s a_j^2}{2} \right) d\tilde{p}_1 \dots d\tilde{p}_N da_2 \dots da_N \tag{87} \\ &= c_N^{-1} 2^{N-2} N^N \prod_{j=2}^N \frac{1}{\sqrt{\lambda_j^s}} S_{2N-3} \frac{(E_N - E_N^s)^{N-1}}{N-1} \end{aligned}$$

and from (36)

$$\begin{aligned} v_S^c &\sim C_N^{-1} N^{(N-1)/2} \int_{\mathbb{R}^{N-1}} \max(\tilde{p}_1, 0) \\ &\times \exp \left( -\beta E_N^s - \beta \sum_{j=1}^N \frac{\tilde{p}_j^2}{2N} - \beta \sum_{j=2}^N \frac{\lambda_j^s}{2} a_j^2 \right) d\tilde{p}_1 \dots d\tilde{p}_N da_2 \dots da_N \tag{88} \\ &= C_N^{-1} \frac{N^N}{\beta^N} \prod_{j=2}^N \frac{1}{\sqrt{\lambda_j^s}} S_{2N-3} 2^{N-2} (N-2)! e^{-\beta E_N^s} \end{aligned}$$

where  $S_n = 2\pi^{(n+1)/2} / \Gamma((n+1)/2)$  is the surface area of the  $n$ -dimensional sphere (embedded in  $n+1$  dimensional space) of unit radius.



For the normalization constants, the concentration of the measure also allows us to extend the integration to all space with little error, resulting in the asymptotic in  $\beta$  expressions

$$c_N^\pm \sim N^{N/2} \int_{\mathbb{R}^{2N}} \delta \left( E_N - E_N^\pm - \frac{1}{2} \sum_{j=1}^N \left[ \frac{\tilde{p}_j^2}{N} + \lambda_j^\pm (b_j^\pm)^2 \right] \right) d\tilde{p}_1 \dots d\tilde{p}_N db_1^\pm \dots db_N^\pm \tag{89}$$

and

$$C_N^\pm \sim N^{N/2} \int_{\mathbb{R}^{2N}} \exp \left( -\beta E_N^\pm - \frac{\beta}{2} \sum_{j=1}^N \left[ \frac{\tilde{p}_j^2}{N} + \lambda_j^\pm (b_j^\pm)^2 \right] \right) d\tilde{p}_1 \dots d\tilde{p}_N db_1^\pm \dots db_N^\pm \tag{90}$$

Notice the factor of  $N^{N/2}$  due to the change of the integration to the eigenvector basis. In our case, due to symmetry,  $c_N^+ = c_N^-$  and  $C_N^+ = C_N^-$  ( $E_N^+ = E_N^-$  and  $\lambda_j^+ = \lambda_j^-$  for all  $j = 1 \dots N$ ), but they could be different if  $V(u)$  was not symmetric. Completing the integration, we have

$$c_N^\pm \sim 2^{N-1} N^N \prod_{j=1}^N \frac{1}{\sqrt{\lambda_j^\pm}} S_{2N-1} (E_N - E_N^\pm)^{N-1} \tag{91}$$

and

$$C_N^\pm \sim 2^{N-1} \frac{N^N}{\beta^N} S_{2N-1} (N-1)! \prod_{j=1}^N \frac{1}{\sqrt{\lambda_j^\pm}} e^{-\beta E_N^\pm}. \tag{92}$$

Thus arriving at the expressions in Proposition 3 and Conjecture 1.

### Appendix 3: Expected Residency Time Using the Microcanonical Measure

Here, we compute the average of  $v_S^T$  over trajectories with initial conditions chosen with respect to the microcanonical (21) invariant measure. As for the case of the canonical invariant measure, the use of a quadratic approximation of the Hamiltonian is justified by the concentration of the measure. This concentration is not as obvious for the microcanonical measure (21) and will depend upon the dimension being large enough. Note that the points that minimize the Hamiltonian are in general not on the surface  $H = E_N$ ; however, this surface does contain  $\mathbf{u}^\pm$  with  $\mathbf{p}$  nonzero satisfying  $H(\mathbf{p}, \mathbf{u}^\pm) = E_N$ . We can therefore think about the concentration of the marginal measure for  $\mathbf{u}$  near the lowest potential energy point for as was shown for the canonical distribution. We do not explicitly bound the integration of the microcanonical measure over a localized region; yet, we expect the quadratic approximation to be reasonable at large  $N$  due to the shown equivalence between the microcanonical and canonical

measures as  $N \rightarrow \infty$ . We present the results assuming this approximation is valid here.

We proceed with the calculation from Sect. 4, using the microcanonical (21) invariant measure. As a result of (35), the ensemble average of the integrand in (34) is invariant in time, and after proper interpretation of integration with respect to the Dirac distribution, we arrive at the analogous equation to (36) for the expected frequency in the case of the microcanonical distribution:

$$v_S^m = c_N^{-1} \int_{\{H_N(\mathbf{p}, \mathbf{u})=E_N\} \cap S} \max(\mathbf{p} \cdot \hat{n}(\mathbf{u}), 0) \frac{d\hat{\sigma}(\mathbf{p}, \mathbf{u})}{|\nabla H_N(\mathbf{p}, \mathbf{u})|} \tag{93}$$

where  $d\hat{\sigma}(\mathbf{p}, \mathbf{u})$  is the surface element on the intersection of  $\{H_N(\mathbf{p}, \mathbf{u}) = E_N\}$  and  $S$ , and  $\hat{n}(\mathbf{u}) \in \mathbb{R}^N$  is the unit normal to  $\{\mathbf{u} : q(\mathbf{u}) = 0\}$ .

For the normalization constant, we first separate the integration into the two regions,  $B_+ = \{(\mathbf{p}, \mathbf{u}) : (\mathbf{u} - \mathbf{u}^s) \cdot \boldsymbol{\phi}^{(1)} > 0\}$  and  $B_- = \{(\mathbf{p}, \mathbf{u}) : (\mathbf{u} - \mathbf{u}^s) \cdot \boldsymbol{\phi}^{(1)} < 0\}$  that partition space, separated by  $S$ ,

$$c_N = c_N^+ + c_N^-, \tag{94}$$

where

$$c_N^\pm = \int_{H_N(\mathbf{p}, \mathbf{u})=E_N \cap B_\pm} \frac{d\sigma(\mathbf{p}, \mathbf{u})}{|\nabla H_N(\mathbf{p}, \mathbf{u})|}. \tag{95}$$

The details of obtaining the following expressions appear in ‘‘Evaluation of the Integrals’’ section of Appendix 2:

**Conjecture 1** *For the microcanonical ensemble with energy  $E_N = N/\beta$ , the asymptotic in  $\beta \rightarrow \infty$ , valid for large  $N$ , expansion of the integral in (93) is*

$$\int_{\{H_N(\mathbf{p}, \mathbf{u})=E_N\} \cap S} \max(\mathbf{p} \cdot \hat{n}(\mathbf{u}), 0) \frac{d\hat{\sigma}(\mathbf{p}, \mathbf{u})}{|\nabla H_N(\mathbf{p}, \mathbf{u})|} \sim \frac{2^{N-2} N^N S_{2N-3}}{(N-1) \prod_{j=2}^N \sqrt{\lambda_j^s}} \left(\frac{N}{\beta} - E_N^s\right)^{N-1} \tag{96}$$

and the expansion of the integrals in (95) is

$$c_N^\pm \sim \frac{2^{N-1} N^N S_{2N-1}}{\prod_{j=1}^N \sqrt{\lambda_j^\pm}} \left(\frac{N}{\beta} - E_N^\pm\right)^{N-1}. \tag{97}$$

In the above, the eigenvalues  $\lambda_j^s, j = 2 \dots N$  solve (40), the eigenvalues  $\lambda_j^\pm$  solve (42), and  $S_n = 2\pi^{(n+1)/2} / \Gamma((n+1)/2)$  is the surface area of the  $n$ -dimensional sphere (embedded in  $n + 1$  dimensional space) of unit radius.

Using the above asymptotic in  $\beta$  expansions, the TST frequency from (93) for the microcanonical ensemble with energy  $N/\beta$  is

$$v_S^m \sim \frac{1}{2\pi} \frac{\prod_{j=2}^N \frac{1}{\sqrt{\lambda_j^s}} \left(1 - \frac{\beta E_N^s}{N}\right)^{N-1}}{\prod_{j=1}^N \frac{1}{\sqrt{\lambda_j^+}} \left(1 - \frac{\beta E_N^+}{N}\right)^{N-1} + \prod_{j=1}^N \frac{1}{\sqrt{\lambda_j^-}} \left(1 - \frac{\beta E_N^-}{N}\right)^{N-1}} \tag{98}$$

and the mean residency time given by

$$\tau_{\pm}^m = \frac{c_N^{\pm}}{c_N v_S^m} \tag{99}$$

is

$$\tau_{\pm}^m \sim 2\pi \frac{\prod_{j=2}^N \sqrt{\lambda_j^s} \left(1 - \frac{\beta E_N^{\pm}}{N}\right)^{N-1}}{\prod_{j=1}^N \sqrt{\lambda_j^{\pm}} \left(1 - \frac{\beta E_N^s}{N}\right)^{N-1}}. \tag{100}$$

Notice that taking the  $N \rightarrow \infty$  limit of (100) also converges to (14). This relies on the fact that the fraction  $(1 - \beta E_N^{\pm}/N)^{N-1}/(1 - \beta E_N^s/N)^{N-1}$  converges to  $e^{\beta \Delta E_{\pm}}$ , defining the energy barrier height  $\Delta E_{\pm} = U(u^s) - U(u^{\pm})$ .

### References

Ariel, G., Vanden-Eijnden, E.: Testing transition state theory on Kac–Zwanzig model. *J. Stat. Phys.* **126**(1), 43–73 (2007)

Barret, F.: Sharp asymptotics of metastable transition times for one dimensional SPDEs. *Ann. Inst. Henri Poincaré Probab. Stat.* **51**(1), 129–166 (2015)

Berglund, N., Gentz, B.: Sharp estimates for metastable lifetimes in parabolic SPDEs: Kramers’ law and beyond. *Electron. J. Probab.* **18**, 1–58 (2013)

Bourgain, J.: Periodic nonlinear Schrödinger equation and invariant measures. *Commun. Math. Phys.* **166**(1), 1–26 (1994)

Bovier, A.: *Metastability*. Springer, Berlin (2015)

Bovier, A., Eckhoff, M., Gaynard, V., Klein, M.: Metastability in stochastic dynamics of disordered mean-field models. *Probab. Theory Relat. Fields* **119**(1), 99–161 (2001)

Bovier, A., Eckhoff, M., Gaynard, V., Klein, M.: Metastability and low lying spectra in reversible Markov chains. *Commun. Math. Phys.* **228**(2), 219–255 (2002)

Bovier, A., Eckhoff, M., Gaynard, V., Klein, M.: Metastability in reversible diffusion processes I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc.* **6**, 399–424 (2004)

Deng, Y.: Invariance of the Gibbs measure for the Benjamin–Ono equation. *J. Eur. Math. Soc.* **17**(5), 1107–1198 (2015)

Eyring, H.: The activated complex in chemical reactions. *J. Chem. Phys.* **3**, 107–115 (1935)

Faris, W.G., Jona-Lasinio, G.: Large fluctuations for a nonlinear heat equation with noise. *J. Phys. A Math. Gen.* **15**(10), 3025–3055 (1982)

Flaschka, H.: *Some geometry in high-dimensional spaces*. Lecture Notes in Mathematics 527A (2010)

Freidlin, M.I., Wentzell, A.D.: *Random Perturbations of Hamiltonian Systems*, Volume 109 of *Memoirs of the American Mathematical Society*. American Mathematical Society, Providence (1994)

Friedlander, L.: An invariant measure for the equation  $u_{tt} - u_{xx} + u^3 = 0$ . *Commun. Math. Phys.* **98**, 1–16 (1985)

- Hairer, M.: A theory of regularity structures. *Invent. Math.* **198**(2), 269–504 (2014)
- Horiuti, J.: On the statistical mechanical treatment of the absolute rate of chemical reaction. *Bull. Chem. Soc. Jpn.* **13**(1), 210–216 (1938)
- Huisinga, W., Meyn, S., Schütte, C.: Phase transitions and metastability in Markovian and molecular systems. *Ann. Appl. Probab.* **14**(1), 419–458 (2004)
- Lebowitz, J.L., Rose, H.A., Speer, E.R.: Statistical-mechanics of the nonlinear Schrödinger-equation. *J. Stat. Phys.* **50**(3–4), 657–687 (1988)
- McKean, H.P.: Geometry of differential space. *Ann. Probab.* **1**(2), 197–206 (1973)
- McKean, H.P., Vaninsky, K.L.: Statistical mechanics of nonlinear wave equations. *J. Math. Sci.* **94**(4), 1630–1634 (1999)
- Newhall, K.A., Vanden-Eijnden, E.: Averaged equation for energy diffusion on a graph reveals bifurcation diagram and thermally assisted reversal times in spin-torque driven nanomagnets. *J. Appl. Phys.* **113**(18), 184105 (2013)
- Oh, T.: Invariance of the Gibbs measure for the Schrödinger–Benjamin–Ono system. *SIAM J. Math. Anal.* **41**(6), 2207–2225 (2010)
- Oh, T., Richards, G., Thomann, L.: On invariant Gibbs measures for the generalized KdV equations. *Dyn. Partial Differ. Equ.* **13**(2), 133–153 (2016)
- Poincaré, H.: *Calcul des Probabilités*. Gauthier-Villars, Paris (1912)
- Ryser, M.D., Nigam, N., Tupper, P.F.: On the well-posedness of the stochastic Allen–Cahn equation in two dimensions. *J. Comput. Phys.* **231**, 2537–2550 (2012)
- Sinai, Y.G.: *Introduction to Ergodic Theory*. Princeton University Press, Princeton (1976)
- Tal, F.A., Vanden-Eijnden, E.: Transition state theory and dynamical corrections in ergodic systems. *Nonlinearity* **19**, 501–509 (2006)
- Tzvetkov, N.: Invariant measures for the nonlinear Schrödinger equation on the disc. *Dyn. Partial Differ. Equ.* **3**(2), 111–160 (2006)
- Tzvetkov, N.: Mesures invariantes pour l'équation de Schrödinger non linéaire. *Ann. Inst. Fourier* **58**(7), 2543–2604 (2008)
- Vanden-Eijnden, E., Tal, F.A.: Transition state theory: variational formulation, dynamical corrections, and error estimates. *J. Chem. Phys.* **128**, 184103 (2005)
- Vanden-Eijnden, E., Westdickenberg, M.G.: Rare events in stochastic partial differential equations on large spatial domains. *J. Stat. Phys.* **131**, 1023–1038 (2008)
- Wigner, E.: The transition state method. *Trans. Faraday Soc.* **34**, 29–41 (1938)
- Zhidkov, P.E.: An invariant measure for a nonlinear wave equation. *Nonlinear Anal. Theory Methods Appl.* **22**(3), 319–325 (1994)