Open system control of dynamical transitions under the generalized Kruskal-Neishtadt-Henrard theorem

Diego M. Fieguth and James R. Anglin
State Research Center OPTIMAS and Fachbereich Physik, Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau, D-67663 Kaiserslautern, Germany

(Received 6 October 2022; revised 15 February 2023; accepted 14 March 2023; published 20 March 2023)

Useful dynamical processes often begin through barrier-crossing dynamical transitions; engineering system dynamics in order to make such transitions reliable is therefore an important task for biological or artificial microscopic machinery. Here, we first show by example that adding even a small amount of back-reaction to a control parameter, so that it responds to the system's evolution, can significantly increase the fraction of trajectories that cross a separatrix. We then explain how a post-adiabatic theorem due to Neishtadt can quantitatively describe this kind of enhancement without having to solve the equations of motion, allowing systematic understanding and design of a class of self-controlling dynamical systems.

DOI: 10.1103/PhysRevE.107.034209

I. INTRODUCTION

From satellite capture in astrophysics to chemomechanical reactions, dynamical transitions induced by slow parameter change are important phenomena in many physical settings. Some kind of effective potential barrier gradually rises or falls until, at some critical point, a qualitative change in system behavior occurs, analogous to the change that a physical pendulum makes when it crosses the phase space separatrix between rotation and libration.

Dynamical transitions of this kind are typically described within adiabatic theory, assuming that the Hamiltonian which governs the system has some slow, fixed time dependence, through an "external" parameter which slowly changes in a predetermined way [1–17]. The adiabatic approximation breaks down near a separatrix, and so post-adiabatic methods are needed to describe the dynamical transition—and even to determine whether it occurs. Separatrix-crossing dynamical transitions often turn out to be quasi-probabilistic, in that they occur only for a subset of initial conditions, and this subset may depend sensitively on exactly how the Hamiltonian depends on time. By applying Liouville's theorem as well as adiabatic analysis, the Kruskal-Neishtadt-Henrard theorem (KNH) derives the post-adiabatic probability of a dynamical transition into a given region of phase space in terms of the rate at which the area of that region is changing (due to the slow change of the Hamiltonian). Neishtadt's generalization of the KNH theorem to include dissipation (hereafter GKNH [19,20]) further shows that even a small amount of dissipation can significantly shift the probability of a separatrix-crossing dynamical transition.

It has recently [18] been pointed out that nondissipative KNH theory can be the basis of a class of "blind" control strategies: by engineering a slowly time-dependent Hamiltonian to increase the growth rate of a target region, the set of initial conditions which lead to transitions into the target region can be enlarged, without having to monitor or control fast degrees of freedom. Our first purpose in this paper is therefore to point out the implications of the GKNH theorem for open-system control over dynamical transitions. Adding a small amount of dissipation to an adiabatic control process can enlarge the range of initial conditions for which a desired transition occurs, potentially converting a rare fluke into a reliable consequence. This is an application of dissipation to control which has nothing to do with the usual idea of cooling to keep the system near its ground state.

Our second purpose in this paper is, conversely, to let the concept of open-system control shed some light on the GKNH theorem, by providing a physical picture of how dissipation can provide efficient control over a sensitive process. Dissipation can be represented within a Hamiltonian framework by including parameters in the system Hamiltonian which, instead of having a fixed time dependence, obey their own non-trivial equations of motion that include coupling to the system degrees of freedom. This is, after all, what dissipation always really is, microscopically—the additional dynamical parameters which provide viscous drag on an object moving through air, for example, are simply the air molecules that can collide with the object. By taking this kind of perspective on dissipation, we will show that dissipation can improve control because it is equivalent to a more sophisticated control strategy that includes tailored feedback, so that the time-dependent control parameter adapts, at least slightly, to the system's time evolution.

We begin our study in Sec. II by demonstrating both of our points with an example. We compare two model control problems that share the same goal of initiating a certain dynamical transition by slowly changing a control parameter. The actual time-dependences of the parameter remain very close to each other in the two models, but in one case dissipative back-reaction induces small correlations between system and control parameter, and these small but systematic differences are enough to induce dynamical transitions for a much larger set of initial conditions than in the case

that has only predetermined time dependence in the control parameter.

We then show in Sec. III that this advantage of open-system control extends beyond our illustrative example, by using Neishtadt's GKNH theorem [19,20] to compute the enhancement of transition probabilities in generally similar scenarios. We conclude in Sec. IV with a brief discussion.

II. ENHANCING TRANSITION PROBABILITY WITH OPEN-SYSTEM CONTROL: AN EXAMPLE

A. Dragging as a control task

Our example to show the general effect is a particle in one dimension, with phase space coordinates q, p as usual, subject to a potential which depends on a time-dependent parameter $\lambda(t)$. In particular the potential is sinusoidal in q, with a λ-dependent amplitude that slowly ramps on and off. At the same time as the potential ramps on and off in strength, it also moves back and forth in space, in dependence on λ. The goal of this twofold λ dependence, as a control task, is to drag the particle over a distance, by capturing it into a bound orbit in one of the wells of the moving sinusoidal potential. Since the potential initially has vanishing strength, the particle is not bound initially; as the potential strength grows, a phase space region of bound orbits grows, but the particle is not initially in it. To succeed in the control task, the particle must cross an adiabatic separatrix into this bound-orbit region.

Because adiabaticity breaks down near a separatrix, the particle in the time-dependent potential can in fact cross the instantaneous separatrix and be dragged successfully. For some initial conditions it does this, while for others it does not. We compare the phase space areas of initial conditions that lead to capture (and thus to successful dragging) for two different forms of $\lambda(t)$ time dependence. The two forms of time dependence will differ only very slightly as functions of time, but they will differ qualitatively in precisely how their time dependences are determined. In one case, it is preordained, while in the other it depends weakly on the instantaneous momentum p of the particle. This weak dependence of the control parameter on the target particle's momentum has a significant effect on the control task, but it is by no means a daemon-like fantasy coupling that cannot be implemented realistically. On the contrary, we will show after defining the momentum dependence that it is nothing but a bit of dissipation.

B. Co-moving frame Hamiltonian

As a simple case we will take the potential's instantaneous velocity to be the parameter λ, on which the potential's instantaneous strength also depends. To simplify our description of the problem we will work in the co-moving frame of the moving potential, where our λ-dependent Hamiltonian reads

$$H(\lambda, q, p) = \frac{(p-\lambda)^2}{2} - \beta^2(\cos(q)+1) \tag{1}$$

with $\beta = \sqrt{g}\exp(-\frac{\alpha}{4}\lambda^2)$ for some constant $\alpha > 0$. In these co-moving frame coordinates, success in the control task of dragging means keeping q nearly constant, captured in a potential well, with only small oscillations of p around λ as λ slowly changes. Failure to capture the particle and drag it will appear instead as q changing with λ, as the particle fails to move with the potential, while p remains constant.

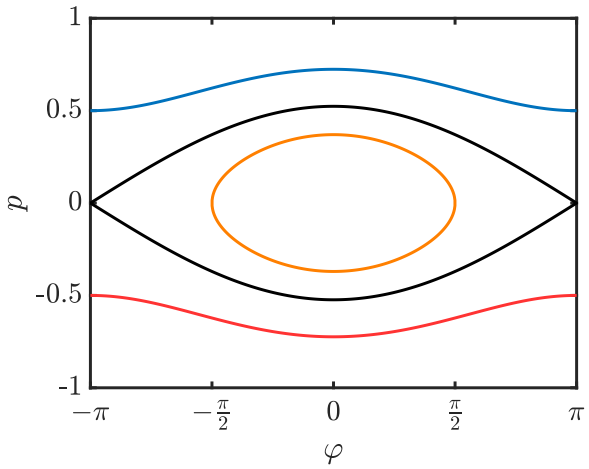

FIG. 1. Instantaneous energy contours in phase space for $\lambda = 0$, when the separatrix is at its largest size. Shown are orbits with energies $H(0, \pi/2, 0)$ (orange), $H(0, \pi, 0)$ (the separatrix, blue), and $H(0, \pi, 0.5)$ (red) for $g = (\pi/4)^2/9$, $\alpha = 1$.

Captured and non-captured orbits in phase space under this H for fixed λ are shown in Fig. 1. We see the separatrix (black) and an orbit inside it (orange), as well as orbits below (red) and above (blue) the separatrix. The eye-shaped separatrix becomes vanishingly small for large $|\lambda|$, as the Gaussian factor the potential strength becomes small, but in general there is a separatrix at any λ, separating phase space into three regions. Since the corners of the separatrix are at the unstable fixed points $p = \lambda$, $q = \pm\pi$, the separatrix is a contour with zero energy and a parametrization is given by

$$p_\pm(q, \lambda) = \lambda \pm \beta\sqrt{2 + 2\cos q}, \tag{2}$$

enclosing the area

$$A_{\text{sep}}(\lambda) = 16\beta. \tag{3}$$

Initially λ will be negative, and it will steadily increase, so that the separatrix slowly grows and then shrinks, while simultaneously moving upward in p. Since ordinary adiabatic theory applies until the system encounters the separatrix, ordinary adiabatic theory can yield an accurate estimate for the value of λ at which the system encounters the separatrix, for given initial conditions. This adiabatically estimated point is known as the "'pseudo-crossing", because the exact λ when the system meets the adiabatic separatrix depends on post-adiabatic corrections. As we show in Appendix A, the value of λ_{pc} is given, in cases where $\beta^2(\lambda(0))$ is negligible, is determined as a function of $p(0)$ by

$$p(0) = \lambda_{\text{pc}} + \frac{A_{\text{sep}}(\lambda_{\text{pc}})}{4\pi} = \lambda_{\text{pc}} + \frac{4\beta(\lambda_{\text{pc}})}{\pi}. \tag{4}$$

Alternatively, Eq. (4) can define the value of $p(0)$ for which the system will encounter the separatrix at a given value of $\lambda = \lambda_{\text{pc}}$. The corresponding time when the system meets the

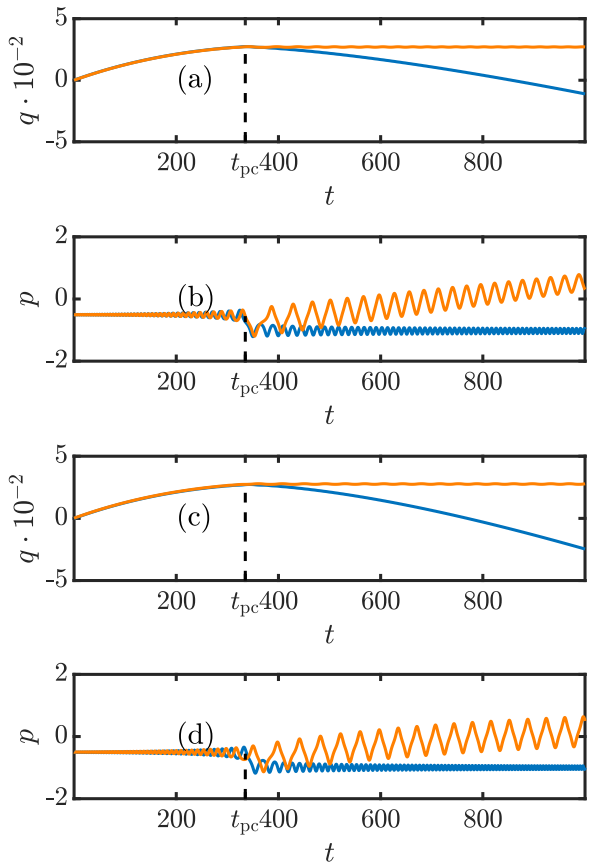

FIG. 2. Time evolution of q (a) and p (b) for trajectories of the responsive case transitioning to different phases of motion. The same qualitative behavior can be seen for the unresponsive case, q (c) and p (d). Slight differences between the orange curves in (b) and (d) can be seen around t_{pc}, and also at late times in the decreasing amplitude of the orange-curve oscillations in (b), with no corresponding amplitude decrease in (d). The constants are $g = (\pi/4)^2/9$, $\alpha = 1$, $\gamma = 1$, and $\varepsilon = 0.002$.

separatrix is then given, for any particular $\lambda(t)$ time dependence, by $\lambda_{pc} = \lambda(t_{pc})$.

Because adiabaticity breaks down whenever the system does encounter the separatrix, at that point the system may possibly cross the separatrix even though this is forbidden adiabatically. This potential crossing is the crux of our control task. The system begins outside the separatrix, above it. If the control task succeeds, the system will enter the separatrix as the separatrix migrates upward, and orbit within it for some significant time, moving upward with it in p. If the task fails, the particle is never captured into the separatrix, but is left behind in the region below the separatrix after it passes. Figure 2 shows examples of both kinds of trajectories, one in blue and one in orange, which differ so slightly in the initial values of q and p that the differences between orange and blue trajectories are nearly invisible until the decisive point is reached. This illustrates the sensitive nature of the dynamical transition that is the goal of this control task. Panels (a),(b) of Fig. 2 show evolution under one particular $\lambda(t)$, while panels (c),(d) are for a $\lambda(t)$ that differs in small but significant ways that we will discuss below.

The second pair of panels in Fig. 2 resembles the first pair quite closely; this shows that the control task can succeed for both forms of $\lambda(t)$, and it can also fail for both kinds of $\lambda(t)$, depending on the precise initial values of q and p. As we will see further below, the significant difference between the two forms of $\lambda(t)$ is in how common it is, among all possible initial conditions, to succeed in the control task. To obtain evolutions like the orange trajectories, instead of those like the blue ones, requires more finely tuned initial conditions for the $\lambda(t)$ of panels (c) and (d) than it does for the $\lambda(t)$ of panels (a) and (b). The difference between the two $\lambda(t)$ is qualitatively drastic even though it is quantitatively small.

C. Time dependence of λ

Whether the time-dependence of λ is fixed in advance, or is determined dynamically, it can always be expressed with a differential equation of some kind. For example, the equation of motion

$$\dot{\lambda} = -\varepsilon\gamma(\lambda - S) \tag{5}$$

for constant S, with the initial condition $\lambda(0) = R$, implies a $\lambda(t)$ that begins with the initial value R and slowly relaxes, on the time scale $(\varepsilon\gamma)^{-1}$, to the final value S. (We include the parameter $\varepsilon \ll 1$ to emphasize the comparative smallness of this relaxation rate, and because the time scale γ will appear without ε in our next model, below.) Since this $\lambda(t)$ is unaffected by the system evolution $q(t)$, $p(t)$, we will refer to it as "the unresponsive case" and denote the solution to Eq. (5) with $\lambda(0) = R$ as $\lambda(t) \to \lambda_U(t)$, namely,

$$\lambda_U(t) = (R - S)e^{-\varepsilon\gamma t} + S. \tag{6}$$

A control degree of freedom which does respond to the system as it evolves could instead of Eq. (5) obey an equation of the form $\dot{\lambda} = \varepsilon f(\lambda(t), q(t), p(t))$ for some function f. For simplicity we will consider f which depend only linearly on $q(t)$ and $p(t)$; this linear case should actually represent more general dependence, as long as the control strategy ends up working well enough that the actual $q(t)$, $p(t)$ do not depart far from the trajectory which the control is trying to enforce.

If we focus first on the possibility of λ depending on q without p, we could allow an equation for $\lambda(t)$ of the form

$$\dot{\lambda} = \varepsilon f_0(t) + \varepsilon a q(t) \tag{7}$$

for constant a and some $f_0(t)$. We could then combine Eq. (7) with the canonical equations of motion for the H of Eq. (1) and differentiate $\dot{q}$ with respect to t to find

$$\ddot{q} = -\beta^2(t)\sin(q) - \varepsilon a q - \varepsilon f_0(t). \tag{8}$$

This means that the $\varepsilon a q(t)$ term in Eq. (7) is effectively just a change of the particle's potential energy from $-\beta^2\cos(q)$ to $-\beta^2\cos(q) + \varepsilon a q^2/2$. Motion in slowly time-dependent potentials of quite general form has been studied extensively, and we do not expect to learn anything new just by adding a parabolic perturbation to our cosine potential.

We will therefore ignore the possibility that $\dot{\lambda}$ might depend on $q(t)$, and focus on linear dependence on $p(t)$. For "the responsive case" of $\lambda(t)$ that we will compare to our unresponsive $\lambda_U(t)$ from Eq. (6), therefore, we will consider

$$\dot{\lambda}_R = \varepsilon(1 - \gamma(\lambda_R - p)). \tag{9}$$

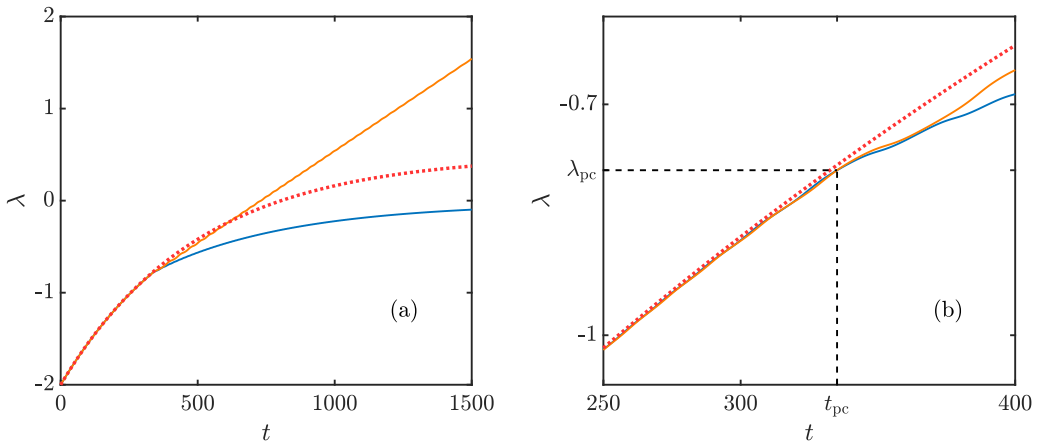

FIG. 3. Time evolution of the control parameter λ for the same initial conditions as in Fig. 2. The orange curve is the responsive $\lambda_R(t)$ for the q, p trajectory which successfully enters the separatrix, while the blue curve shows the responsive $\lambda_R(t)$ for the trajectory which fails to cross the separatrix. The red dotted curve is the unresponsive $\lambda_U(t)$, which is the same for both kinds of system trajectory. (b) depicts the same evolutions zoomed in around λ_{pc}. The three curves all remain close until after the system encounters the separatrix at t_{pc}, showing that the success or failure of the control task depends sensitively on the precise time dependence of the control parameter.

When we now differentiate the canonical equation for $\dot{q}$ we obtain

$$\ddot{q} = -\beta^2(t)\sin(q) - \varepsilon - \varepsilon\gamma\dot{q}. \tag{10}$$

The effect of even a small momentum dependence in $\dot{\lambda}_R$ is to add dissipation to the particle motion, which can substantially affect adiabatic control. Whereas the product $\varepsilon\gamma$ still defines the relaxation rate for λ_R as it did for λ_U, we can see from Eq. (10) that γ alone, without the small prefactor ε, represents the relative strength of dissipation, and of the responsiveness of λ to the system, compared to the predefined steady increase of λ, driven by the term in Eq. (10) without γ.

Conversely, Eq. (10) shows that adding simple damping to our controlled particle is equivalent to making the control parameter $\lambda_R(t)$ respond dynamically to the particle, through its momentum. Representing dissipation in this way, through a dynamical reaction of the control parameter to the controlled particle, allows us to make direct comparison between closed-system control, with $\lambda_U(t)$ and open-system control with the responsive $\lambda_R(t)$ evolving under Eq. (9), which is equivalent to including damping.

This particular comparison between $\lambda_U(t)$ and $\lambda_R(t)$ is informative because even though the qualitative difference between them is drastic, with $\lambda_U(t)$ being non-responsive and fixed while $\lambda_R(t)$ is responsive and dissipative, we will be able to keep the quantitative difference between the two time-dependent control parameters *small*. This lets us isolate the specific effect of openness in open-system control.

The particular form of $\lambda_U(t)$ can closely approximate $\lambda_R(t)$, in spite of their qualitative differences, because until the system approaches the separatrix, ordinary adiabatic theory will accurately describe $q(t)$ and $p(t)$, and we can therefore also use this adiabatic $p(t)$ to compute $\lambda_R(t)$ self-consistently, to high accuracy for small ε. As often with adiabatic evolution, the result is equivalent to averaging away small-amplitude, high-frequency components of $\lambda_R(t)$. This analysis is presented fully in Appendix A, but can be outlined here briefly. If the particle is not (yet) captured, then the only effect of the sinusoidal potential on $p(t)$ will be "speed bump" modulations which average away, leaving the constant average $\bar{p}(t) = p(0)$. Substituting $p(t) \to p(0)$ in Eq. (9) then leads to the self-consistent solution in which the approximation to $\lambda_R(t)$ is simply given by $\lambda_U(t)$ with

$$R = \lambda(0) \quad S = p(0) + \frac{1}{\gamma}. \tag{11}$$

It is therefore these particular values of R and S that we will use in our $\lambda_U(t)$, making it close to the responsive $\lambda_U(t)$ in our comparison.

Just how quantitatively close $\lambda_U(t)$ and $\lambda_R(t)$ are can be seen in Fig. 3. There the dotted red curve is the unresponsive $\lambda \to \lambda_U(t)$ with optimal parameters Eq. (11); since $\lambda_U(t)$ is unresponsive, this curve is the same regardless of what the system actually does. The orange and blue curves are the responsive $\lambda \to \lambda_R(t)$ for successful and unsuccessful $p(t)$ trajectories, respectively. In panel (b), we see that even when zoomed in the three control parameter time dependences remain virtually indistinguishable (on this scale) up until after the time t_{pc} when the system has reached the separatrix and either crossed it successfully or failed to do so.

Because the three $\lambda(t)$ evolutions agree well up to this point, we can use $\lambda_U(t_{pc}) = \lambda_{pc}$ to compute the pseudo-crossing time t_{pc} in all cases. As Fig. 2 shows, this pseudo-crossing t_{pc} actually is very close, for both responsive and unresponsive cases of λ, to the time at which the system either enters the separatrix, with $p(t)$ thereafter oscillating around λ as in the orange trajectory, or else bypassed by the moving separatrix, in the blue trajectory, with a one-time downward shift in p followed by "speed bump" oscillations around constant p.

So from Fig. 3 we have seen that $\lambda_R(t)$ and $\lambda_U(t)$ remain very close to each other up to and around t_{pc}. And we have seen that both responsive [Figs. 2(a) and 2(b)] and unresponsive control parameters [Figs. 2(c) and 2(d)] do allow the dynamical transition into the separatrix, with success in the dragging control task, while also allowing failure in the control task, depending on precise initial conditions $q(0)$, $p(0)$.

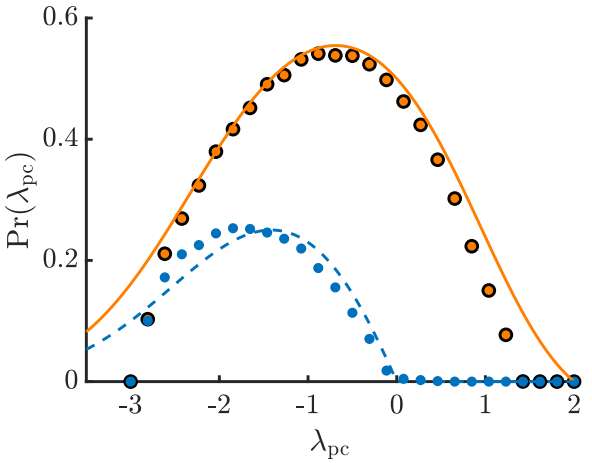

FIG. 4. Probabilities of success in the dragging control task, as predicted post-adiabatically by the KNH (blue dashed) and GKNH (orange) formulas, with the parameters from Fig. 2 for different λ_{pc}. To calculate each probabilities we evolved $N = 10^4$ trajectories numerically for different random initial conditions and determined the probability by counting the trajectories entering the separatrix, N_{sep}, and then taking the fraction N_{sep}/N. Blue dots without borders mark the probabilities for the unresponsive system, while orange dots with black borders represent the responsive system. What appear to be blue dots with black borders, towards the sides of the plot, are in fact borderless blue dots overlying bordered orange dots, because the probabilities in the two cases are indistinguishable.

How big a difference, then, do the small differences between $\lambda_R(t)$ and $\lambda_U(t)$ possibly make, in this control task? They can make quite a large difference: the relative proportions of captured trajectories for responsive and unresponsive cases are drastically different, as can be seen in Fig. 4.

D. Systematic effect of responsiveness

Figure 4 represents evolution of an ensemble of 10^4 initial conditions $q(0)$ randomly distributed in $[-\pi, \pi]$, for each of a range values of $p(0)$ corresponding to 27 different values of λ_{pc}. Trajectories with separatrix crossing are identified for evolution with the unresponsive parameter $\lambda \to \lambda_U(t)$, and the probability $\Pr(\cdot)$ of control task success is then plotted, as blue dots, versus λ_{pc}. The whole procedure is then repeated with the responsive parameter $\lambda \to \lambda_R(t)$, and the probabilities shown as orange dots. The results clearly show that the probability of control task success, *i.e.* the fraction of initial conditions leading to orbits with capture into the separatrix, is substantially higher in the responsive case. The probability can even be significant in the responsive case for $\lambda_{pc} > 0$, where it is zero for the unresponsive case.

The reason for the great difference in control effectiveness between the responsive and unresponsive cases can then be seen in Fig. 5. Because the slight differences between $\lambda_R(t)$ and $\lambda_U(t)$ are mainly on the short time scale that is normally ignored in adiabatic approximations, in panel (b) of Fig. 5 the differences are enhanced by plotting the time derivatives $\dot{\lambda}_R$ and $\dot{\lambda}_U$. We notice that the small difference between $\lambda_R(t)$ and $\lambda_U(t)$ happens to be that the responsive $\lambda_R(t)$ has small high-frequency oscillations which closely follow the oscillations of $p(t)$—while the unresponsive $\lambda_U(t)$ does not track these small oscillations of $p(t)$. The difference between responsive $\lambda_R(t)$ and unresponsive $\lambda_U(t)$ is thus small, but it is by no means simply a random small difference. Instead there is a small *correlation* between $\lambda_R(t)$ and $p(t)$, which is missing from $\lambda_U(t)$. This small *cooperative* interaction between $p(t)$ and the responsive $\lambda_R(t)$ can have decisive effects on the unstable evolution around the separatrix. Nonetheless, as we have seen above, this almost intelligent-seeming cooperative interaction is equivalent to nothing but a little bit of dissipation added to our adiabatically controlled dynamical system.

Since the separatrix includes an unstable fixed point, it should not really be surprising that small differences can affect separatrix crossing dramatically. What is less obvious, in this representation of dissipation through a responsive control parameter, is that the effect of parameter responsiveness on the control task is indeed systematically cooperative: the systematic dissipative drift of the system into the lower-energy phase

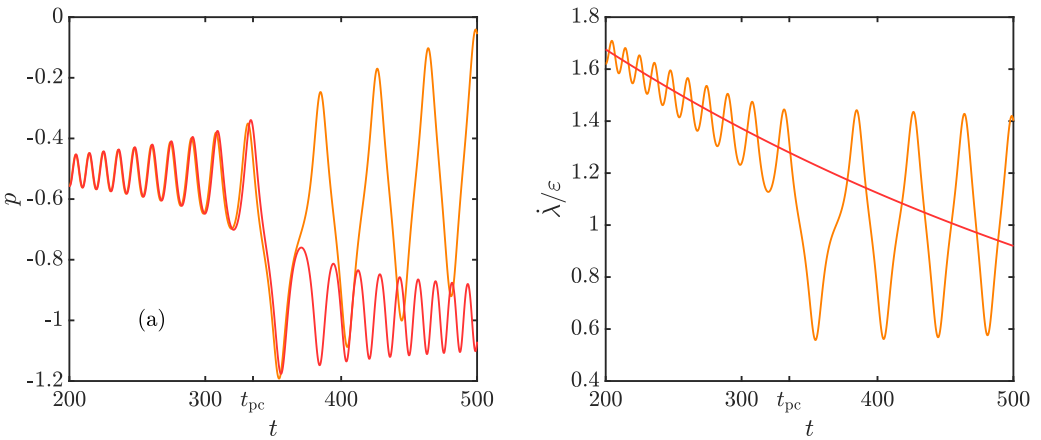

FIG. 5. Evolution for p and $\dot{\lambda}$ in both the responsive (orange) and unresponsive (red) case for the same initial conditions with the same parameters as before. Note how the $p(t)$ oscillations are very similar for the two cases up until t_{pc}, but the oscillations in the responsive orange $\dot{\lambda}_R(t)$ closely follow the oscillations in $p(t)$ for $t < t_{pc}$, while the unresponsive red curve for $\dot{\lambda}(t) \to \dot{\lambda}_U(t)$ does not. This difference in the responsive $\dot{\lambda}_R(t)$ behavior is small, but systematically cooperative in triggering capture.

space region inside the separatrix pushes the system towards success. For any single trajectory the success or failure of the separatrix-crossing control task still depends sensitively on precise values of $q(0)$ and $p(0)$, as well as on α, ε, and $\lambda(0)$. Over a broad enough ensemble of initial conditions, however, the dynamical responsiveness of $\lambda_R(t)$ to the system, in comparison to the unresponsiveness of the predetermined $\lambda_U(t)$, contributes a substantial bias in favor of successful capture into the separatrix, as clearly shown in Fig. 4.

E. Generality

This effect is not just a lucky special feature of this particular model system; neither is it an irreducibly complex manifestation of unstable dynamics which can only be represented accurately by solving equations of motion exactly for all initial conditions. The powerful and general GKNH theory provides the solid orange and dashed blue curves in Fig. 4, to which the ensemble numerical dots conform closely (except at the edges of the curves, where the adiabatic assumptions of the theory break down for our illustrative parameters). In the next section we review this theory and show how it implies a generic advantage, from even weak dissipation, for open-system control of dynamical transitions.

III. THE GENERALIZED KNH THEOREM

Here, we review the Kruskal-Neishtadt-Henrard formula, showing how it applies to our unresponsive case from Sec. II. We then briefly explain the generalized GKNH version of the result, and show how it predicts the advantages that we have found for the dissipative responsive case of open-system control.

A. The KNH theorem

Our unresponsive case is simply a Hamiltonian with a fixed (and slow) time dependence; the probabilities that result from breakdown of adiabaticity around an unstable fixed point are given by the Kruskal-Neishtadt-Henrard theorem [18]. The KNH theorem is not trivial to prove rigorously [20], but in essence it follows from Liouville's theorem. If phase space regions are changing their area, the Liouvillian incompressibility of phase space flow along system orbits implies that trajectories must move between regions. For the three areas $A_\downarrow, A_{\rm sep}, A_\uparrow$ in our pendulum-like system, this leads directly to the simple formula for the probability $\Pr(\cdot)$ of control task success (fraction of initial orbits which cross the separatrix)

$$\Pr(t_{\rm pc}) = \frac{\dot{A}_{\rm sep}(t_{\rm pc})}{\dot{A}_{\rm sep}(t_{\rm pc}) + \dot{A}_\downarrow(t_{\rm pc})}, \tag{12}$$

where the areas $A_{\downarrow,\uparrow}$ are made finite with a constant cutoff far away from the separatrix. In cases where $\Pr(t_{\rm pc})$ as given by Eq. (12) is greater than one or negative, the theorem specifies that Pr is to be taken instead to be one or zero, respectively.

We can use Eq. (12) to determine the capture probability for our unresponsive system. With the separatrix area $A_{\rm sep} = 16\beta$ as given by Eq. (3), our time-dependent potential amplitude $\beta = \sqrt{g}e^{-\frac{\alpha}{4}\lambda^2(t)}$ implies

$$\dot{A}_{\rm sep} = -8\alpha\dot{\lambda}\lambda\beta. \tag{13}$$

The area below the separatrix is $A_\downarrow = 2\pi(\lambda - C_\downarrow) - A_{\rm sep}/2$, where $C_\downarrow$ is a constant cutoff far below the separatrix. This yields

$$\dot{A}_\downarrow = 2\pi\dot{\lambda} - 4\alpha\dot{\lambda}\lambda\beta. \tag{14}$$

The probability of entering the separatrix in the unresponsive case is then given by the formula

$$\Pr(t_{\rm pc}) = \left.\frac{-4\alpha\lambda\beta}{\pi - 2\alpha\lambda\beta}\right|_{\lambda=\lambda_{\rm pc}}. \tag{15}$$

It is noteworthy that this expression is independent of $\dot{\lambda}$, being obtained in the adiabatic limit; we could have chosen $\lambda_U(t)$ arbitrarily (as long as it is sufficiently slow) and obtained the same capture probability. This formula corresponds to the dashed blue line in Fig. 4, which agrees well with the blue dots from the numerical calculation. It also explains why there is no capture for $\lambda_{\rm pc} > 0$, where the separatrix area is constantly shrinking and no states can enter the separatrix due to Liouville's theorem. The only way to exceed this KNH capture efficiency is to evade Liouville's theorem, by having an open system with some non-Hamiltonian term in its evolution.

B. The generalized KNH theorem

The KNH theorem has been generalized to include some forms of non-Hamiltonian evolution. In this section we will adapt the formulas from [19,20] for the probability of entering a double-well-like system with Hamiltonian $H(\lambda, q, p)$, tailoring them to our example, and then give a heuristic derivation of the formula. In [20] it is explicitly anticipated that such adaption to different phase space geometries is possible, even to those on a cylinder as in our case.

First we note that our example fits the general form of equation of motion

$$\dot{\lambda} = \varepsilon f_\lambda(\lambda, q, p,), \tag{16}$$

$$\dot{q} = \frac{\partial H}{\partial p} + \varepsilon f_q(\lambda, q, p), \tag{17}$$

$$\dot{p} = -\frac{\partial H}{\partial q} + \varepsilon f_p(\lambda, q, p), \tag{18}$$

for which the theorem holds. We further have a phase space structure like that of a physical pendulum: an eye-like inner region bounded by a separatrix consisting of upper and lower arcs. The separatrix is a contour of constant instantaneous energy; for convenience (and without loss of generality) we set it to have zero energy at all times, so that the energy in its interior is negative. Any trajectory which crosses the separatrix will first approach it and orbit past it closely for some time; we will compute and compare the loss of energy h during orbits just above and just below the separatrix, using the upper-arc "frown" and lower-arc "smile" symbols as subscripts to denote orbits above and below the separatrix, respectively.

In particular we consider a set of orbits which are initially above the separatrix with energies in the range $(0, \delta h_\frown]$; see Fig. 6. This set (except an exponentially small subset near the fixed point [2]) will evolve closely to the separatrix and will pass along it much faster than the separatrix will deform. The upper bound $\delta h_\frown$ is the energy loss of the trajectories,

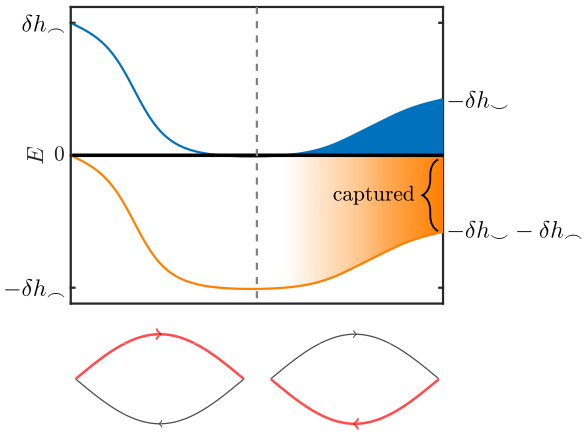

FIG. 6. Change of the energy interval $[0, \delta h_\frown]$ (upper bound blue, lower bound orange) over time, during the evolution along upper and lower arc of the separatrix. Along the upper arc the energy $\delta h_\frown$ gets lost, yielding interval $[-\delta h_\frown, 0]$. Along the lower arc, the energy $-\delta h_\smile$ is gained; orbits which thereby rise above zero energy again are not returning to their initial phase space region above the separatrix, however, but are rather proceeding below the lower arc of the separatrix, where energy is also positive. These blue-shaded orbits will therefore *not* cycle around the separatrix again and be captured; instead they have successfully crossed the whole negative energy region, climbing back up to positive energy, and will never be captured. The captured fraction is what remains below the separatrix energy, colored in orange.

calculated as integral over the upper arc of the instantaneous separatrix contour. This energy loss over the upper separatrix arc will bring the upper energy of our orbits' range down to the separatrix energy of zero. This means that we are dealing with the largest set of orbits which begins above the upper branch of the separatrix and is brought entirely under it, by the parameter change and non-Hamiltonian energy loss, as it follows close to the arc. The energy range $\delta h_\frown$ can be computed in the adiabatic limit by integrating along the upper arc:

$$\delta h_\frown = -\oint \mathrm{d}\tau \frac{\mathrm{d}H}{\mathrm{d}\tau} = -\varepsilon \oint \mathrm{d}\tau \left(\frac{\partial H}{\partial \lambda} f_\lambda + \frac{\partial H}{\partial q} f_q + \frac{\partial H}{\partial p} f_p \right), \tag{19}$$

where τ refers to time evolution under the instantaneous Hamiltonian at t_{pc} with "frozen" $\lambda = \lambda_{\mathrm{pc}}$.

This means that all states above the separatrix with energies in the interval $(0, \delta h_\frown]$ will transition to $(-\delta h_\frown, 0]$, as sketched in the first half of Fig. 6. To determine to which region of phase space the state will transition, we follow its path close to the lower arc of the separatrix. There the energy changes by

$$\delta h_\smile = -\oint \mathrm{d}\tau \frac{\mathrm{d}H}{\mathrm{d}\tau} = -\varepsilon \oint \mathrm{d}\tau \left(\frac{\partial H}{\partial \lambda} f_\lambda + \frac{\partial H}{\partial q} f_q + \frac{\partial H}{\partial p} f_p \right). \tag{20}$$

This leads to a spread in energy of $(-\delta h_\frown - \delta h_\smile, -\delta h_\smile]$, as sketched in the second half of Fig. 6. It is essential to realize that the blue-shaded region in Fig. 6 has risen in energy above zero, so that it has escaped from inside the separatrix, but it has *not* returned to its starting region above the separatrix. Instead it has fallen below the lower arc of the separatrix, where the energy again rises. This blue region therefore represents a portion of the ensemble which has crossed through the negative-energy region without being trapped in it: it is the portion of the initial energy range just above the separatrix which does *not* get captured.

If for some values of λ_{pc} there is no blue region, because none of the initial energy range ever comes back to positive energies, then capture is certain; if there is no orange region, capture is impossible. Otherwise, the fraction of orbits near the separatrix which are captured inside it is given by the ratio of the final energy width of the orange-shaded region in Fig. 6 to the initial energy width $\delta h_\frown$,

$$\Pr(\lambda_{\mathrm{pc}}) = \left. \frac{\delta h_\frown + \delta h_\smile}{\delta h_\frown} \right|_{\lambda=\lambda_{\mathrm{pc}}}. \tag{21}$$

To see how these generalizations of the KNH probability formula apply to our responsive case, we rewrite the integrals (19) and (20) and decompose f_λ. We change the integration variable from τ to q and use $f_q = f_p = 0$, obtaining

$$\delta h_\frown = -\varepsilon \int_{-\pi}^{\pi} \mathrm{d}q \frac{\partial H}{\partial \lambda} \left(\frac{\partial H}{\partial p} \right)^{-1} f_\lambda(\lambda, q, p_+(q)) \tag{22}$$

with p_+ from Eq. (2). The two partial derivatives in the integral evaluate to

$$\frac{\partial H}{\partial \lambda} \left(\frac{\partial H}{\partial p} \right)^{-1} = -1 \mp 2\beta' \cos(q/2) \tag{23}$$

for the two different paths, where $\beta' = \partial_\lambda \beta$. We can then use a decomposition to separate the effects

$$f_\lambda(\lambda, q, p_\pm(q)) = \bar{f}(\lambda) + \Delta_\pm(\lambda, q). \tag{24}$$

The function $\bar{f}$ is simply a time dependence of the Hamiltonian, so in the adiabatic limit it has a constant value at $\lambda = \lambda_{\mathrm{pc}}$, and we can pull it in front of the integrals for both $\delta h_\frown$ and $\delta h_\smile$. In the numerator of the probability formula this then leaves us with

$$\begin{aligned} &\varepsilon \bar{f} \int_{-\pi}^{\pi} \mathrm{d}q \left[1 + 2\beta' \cos\left(\frac{q}{2}\right) - 1 + 2\beta' \cos\left(\frac{q}{2}\right) \right] \\ &= 16 \varepsilon \bar{f} \beta' = \varepsilon \bar{f} \frac{\partial}{\partial \lambda} A_{\mathrm{sep}}. \end{aligned} \tag{25}$$

If $\Delta = 0$ we have $\varepsilon \bar{f} \partial_\lambda = \dot{\lambda} \partial_\lambda = \partial_t$ and the rates of change of the areas from the KNH theorem are recovered.

To evaluate the term arising from $\Delta_\pm$ in the numerator, we choose the decomposition $\Delta_\pm(\lambda, q) = \gamma(p_\pm(q) - \lambda)$. The term arising from the back-reaction is then just

$$\varepsilon \int_{-\pi}^{\pi} \mathrm{d}q (\Delta_+ - \Delta_-) = 16 \varepsilon \gamma \beta = \varepsilon \gamma A_{\mathrm{sep}}, \tag{26}$$

where Δ_- is subtracted because the path of integration reverses and the second summand in from Eq. (23) vanishes because of symmetry. From Eq. (26) we immediately see the mechanism of our back-reaction by which the probability is increased: terms proportional to the separatrix area are added to the numerator, allowing it to be greater than zero even for a shrinking separatrix.

A more detailed calculation of the values of $\delta h_\frown$, $\delta h_\smile$ can be found in Appendix B. The final formula for the probability is

$$\Pr(\lambda_{\rm pc}) = \left.\frac{\gamma A_{\rm sep} + A'_{\rm sep}}{A'_\downarrow + A'_{\rm sep} + \gamma A_{\rm sep}/2 - 4\pi\gamma A_{\rm sep}A'_{\rm sep}/16^2}\right|_{\lambda=\lambda_{\rm pc}}, \tag{27}$$

where $A'_{\rm sep} = \partial_\lambda A_{\rm sep}$. Explicitly in terms of the system parameters for our model this yields

$$\Pr(\lambda_{\rm pc}) = \left.\frac{16\gamma\beta - 8\alpha\lambda\beta}{2\pi + 8\gamma\beta - 4\alpha\lambda\beta - 2\pi\gamma\alpha\lambda\beta^2}\right|_{\lambda=\lambda_{\rm pc}}, \tag{28}$$

which provides the theoretical curve for the responsive case (orange) in Fig. 4.

In Fig. 4 we also see some deviations of the analytical probabilities at the fringes of the distribution. These arise when the change of the separatrix area is slow compared to the speed of its instantaneous fixed point, so that the adiabatic limit no longer holds well.

C. Optimal capture

We have seen in the last section that what matters for capture is the behavior of the vector field $\mathbf{f} = (f_\lambda, f_q, f_p)^\top$ along the separatrix. Now we will assume that we have at least some control over $\mathbf{f}$, and ask whether and how a careful choice of $\mathbf{f}$ can make control success certain, i.e., raise Pr to one. Alternatively, it could be that capture inside the separatrix is a hazard that we would like to avoid; in such a case we would try to bring Pr to zero.

Let us consider that the separatrix at the pseudo-crossing parameter value $\lambda_{\rm pc}$ has been given. If we can choose $\mathbf{f}$ at every point along the separatrix, how should we chose it to enhance capture, or suppress it? A glance at Eq. (21) reveals that we want to minimize $\delta h_\smile$ and maximize $\delta h_\frown$—or else the opposite, if we desire to avoid having the system get captured.

In either case we can rewrite Eq. (19) as

$$\delta h_\frown = -\varepsilon \oint \mathrm{d}\tau \mathbf{n}\cdot\mathbf{f}, \tag{29}$$

where $\mathbf{n}(q(\tau), p(\tau)) = (\partial_\lambda H, \partial_q H, \partial_p H)^\top$ is the outward-pointing normal vector of the surface that the separatrix sweeps out as λ is varied, and the dot is the usual scalar product. An analogous formula holds for $\delta h_\smile$. Hence the goal of increasing Pr means increasing $\mathbf{n}\cdot\mathbf{f}$ over the upper branch of the separatrix, and decreasing it over the lower branch. If instead the goal were to decrease Pr, because capture inside the separatrix was a danger rather than a desired outcome, then we would want to decrease $\mathbf{n}\cdot\mathbf{f}$ over the upper branch of the separatrix and decrease it over the lower.

Our ability to control $\mathbf{n}\cdot\mathbf{f}$ at any point in phase space may in practice be limited. In the example from the previous subsection, only f_λ was non-zero; even if our control over $\mathbf{f}$ were limited by this constraint, we could in principle still make f_λ positive on the upper arc and negative on the lower arc, so that $\delta h_\smile > 0$ while $\delta h_\frown < 0$, which then immediately makes the right-hand side of Eq. (21) greater than one, meaning $\Pr = 1$. Or else we could make f_λ sufficiently negative on the lower arc, and sufficiently positive on the upper arc, to make $\delta h_\smile < 0$ while $\delta h_\frown + \delta h_\smile > 0$, so that the right-hand side of Eq. (21) becomes negative, meaning $\Pr = 0$.

A responsive parameter that is equivalent to dissipation can thus in principle achieve perfect control, either to guarantee that the system will be captured inside the separatrix, or to guarantee that it will not be. Whether those ideals can be achieved in practice depends on how well $\mathbf{f}$ can in fact be engineered.

IV. DISCUSSION

As our example in Sec. II illustrates, and our derivation in Sec. III from the GKNH shows in general, the efficiency of an adiabatic sweep type of control procedure can be altered by giving the control parameter some dynamical responsiveness, so that it suffers back-reaction from the target system instead of exactly following a pre-programmed sweep protocol. In our example in Sec. II we showed dramatic increase, in the responsive case, in the probability of capturing the system into a target region of phase space. Our derivation in Sec. III indicates, however, that capture probabilities can in general be either raised or lowered, depending on exactly how the control parameter responds to the system.

In some cases where control parameter back-reaction is lowering the capture probability, this may be an unwanted effect which needs to be eliminated; our message in such cases is that even small back-reaction from the system onto the control parameter could be the cause of otherwise mysteriously low capture efficiency. In other cases, however, capture might not be a goal but an obstacle; the control task might be to avoid capture of the system into an undesired state. Here, our message is that an appropriate form of parameter back-reaction can indeed reduce the chances of undesired capture. Conversely, desired capture can have its probability enhanced, potential up to unity, by appropriately engineered responsive sweeping.

In either case we emphasize that the effect of parameter responsiveness on capture probabilities can be substantial even when the parameter back-reaction itself is a small effect, because capture by separatrix crossing is a sensitive process. The subtlety of this effect has indeed sometimes been overlooked, for example in celestial mechanics analysis of satellite capture [2,21]. Problems of this nature have been studied in celestial mechanics for decades, and yet the generalized Kruskal-Neishtadt-Henrard theorem has only quite recently been given a rigorous proof [20]. The main purpose of our paper has been to emphasize the implications of these results for more general systems than celestial ones, including laboratory set-ups in atomic and optical physics, where various forms of open-system dynamics can indeed be engineered. In such contexts, the problem is not that of accurately approximating the particular kind of parametric time dependence that appears in one natural system, but of choosing the kind of parametric time dependence which offers efficient control over a dynamical transition.

Even microscopic systems may require efficient control mechanisms, whether artificially designed or naturally evolved. Adiabatic control procedures can be robust and efficient; a major advantage is that they can achieve blind control, in the sense that adequate success probabilities are

achieved without precise monitoring or control of fast degrees of freedom [18]. Adding a bit of responsiveness to a control parameter can substantially improve this kind of blind control, inasmuch as the parameter itself may be less blind to the system's fine-grained behavior than the experimenter is.

Taken to the extreme this idea becomes obvious. A Maxwell's demon is a hypothetical controller which affects the system in drastically different ways depending on the state of the system; at least if one is careless enough in formulating the demon, it can seem even to break the Second Law of Thermodynamics. Intelligent manipulation of a system, to the point where one carries it by hand to its target state, can obviously succeed at control tasks quite well.

One implication of the examples that we have seen here, however, is that even very simple and limited control responsiveness, far short of intelligence, can guide system evolution efficiently. Leakage of system information into the environment is often seen as an enemy which must be fought by isolating the system more perfectly; we argue here to the contrary, that open system control, in which the system is allowed to affect its environment, can offer dramatic benefits. Dissipation is usually considered an unwanted energy loss, but cooling an engine is crucial to its performance—and this principle can extend even to microscopic active particles [22].

The particular form of parameter responsiveness which we analyzed in Sec. II, even though it implements a feedback-driven correlation between system and control parameter, is still simply a representation of viscous damping. The advantage that we have shown for a dynamically responsive control parameter is thus a generic advantage of open-system control, using dissipation to enhance the efficiency of an adiabatic control procedure. This represents an application of dissipation to control beyond the common one of strong cooling to keep a system in a desired ground state. Open-system control allows even quite weak dissipation to significantly enhance the efficiency of control over dynamical transitions.

ACKNOWLEDGMENTS

The authors acknowledge support from State Research Center OPTIMAS and the Deutsche Forschungsgemeinschaft (DFG) through SFB/TR185 (OSCAR), Project No. 277625399.

APPENDIX A: CALCULATING t_{pc} FROM THE AVERAGING PRINCIPLE

The fact that λ_U and λ_R have the same t_{pc} is not a coincidence: the two versions of $\lambda(t)$ are very similar up to t_{pc}, because λ_U is equivalent to approximating λ_R via the so-called averaging principle [20,23].

To show this, we note the time derivative of the averaged momentum in the responsive case, using a bar to denote a variable which is time-averaged over one period of the fast evolution:

$$\dot{\bar{p}} = \frac{-1}{T}\int_0^d \mathrm{d}t\, g \exp\left(-\frac{\alpha}{2}\lambda_R^2\right)\sin(q) \approx 0, \tag{A1}$$

where T is the instantaneous period (of an open 2π-periodic orbit, since the initial ensemble is above the separatrix). Since $\bar{p}$ is thus time-independent while the system is far enough outside the separatrix, we conclude that $\bar{p}(t) = p(0)$ up until t_{pc}.

(For orbits outside the separatrix in which $q(t)$ runs periodically through 2π at a nearly constant speed, $\bar{p}$ is related simply to the usual adiabatically conserved action—the phase space area enclosed by an orbit under the instantaneous Hamiltonian. It is important to recognize that p is the *canonical* momentum, and not the usual linear momentum proportional to the velocity, which is rather $\dot{q} = p - \lambda$. Even when $\bar{p}$ is adiabatically invariant, well outside the separatrix, $\dot{q}$ is still slowly damped, due to the slow time dependence of λ.)

The simplified averaging principle with $\dot{q} \doteq$ constant that yields Eq. (A1) breaks down when the trajectory is close to the instantaneous fixed point; averaging properly over q for non-constant $\dot{q}$ yields a long expression with elliptic integrals for $\dot{\bar{p}}$. This exactly averaged quantity does vanish far away from the separatrix, however, confirming that Eq. (A1) is a good approximation up until t_{pc}. This then enables us to find a simple form for the averaged $\bar{\lambda}_R$.

To average λ_R, we focus on open orbits before capture. We integrate Eq. (9) over one period T, which yields

$$\dot{\bar{\lambda}}_R = \frac{\varepsilon}{T}\int_0^d \mathrm{d}t\; 1 + \gamma(p - \lambda_R) = \varepsilon(1 + \gamma(\bar{p} - \bar{\lambda}_R)). \tag{A2}$$

Since $\bar{p} = p(0) = \text{const}$ we can integrate this, and find straightforwardly that it coincides with the solution in the unresponsive case (6)

$$\bar{\lambda}_R = \lambda_U = \left(\lambda(0) - \frac{1 + \gamma p(0)}{\gamma}\right)\mathrm{e}^{-\varepsilon\gamma t} + \frac{1 + \gamma p(0)}{\gamma}. \tag{A3}$$

We can now safely use $\lambda_U(t)$ to compute t_{pc} for both cases of λ. Since t_{pc} is simply the time at which the system's adiabatic orbit meets the rising separatrix, t_{pc} is determined by the condition that the adiabatically conserved area under the system orbit $2\pi\bar{p} = 2\pi p(0)$ is equal to the area under the upper arc of the separatrix

$$2\pi p(0) = A_{\frown}(\lambda) = 2\pi\lambda + A_{sep}(\lambda)/2. \tag{A4}$$

This fixes the pseudo-crossing time t_{pc} by fixing the pseudo-crossing parameter $\lambda_{pc} = \lambda_U(t_{pc})$. Alternatively, we can choose the initial condition

$$p(0) = \lambda_{pc} + 4\beta(\lambda_{pc})/\pi \tag{A5}$$

to obtain a trajectory with a given pseudo-crossing parameter value λ_{pc}.

APPENDIX B: CALCULATING THE RATE FROM $A_\downarrow$ FOR THE RESPONSIVE SYSTEM

Here, we calculate expressions for $\delta h_{\frown}$ and $\delta h_{\smile}$ from Eqs. (19) and (20). In our case with vanishing f_q, f_p,

$$\delta h_{\frown} = -\varepsilon\oint \mathrm{d}t\,\frac{\partial H}{\partial\lambda}f_\lambda = -\varepsilon\int_{-\pi}^{\pi}\mathrm{d}q\,\frac{\partial H}{\partial\lambda}f_\lambda/\dot{q}. \tag{B1}$$

The integrand can be assembled explicitly using

$$\dot{q} = p - \lambda, \quad p(q) = \lambda + \beta\sqrt{2(1+\cos q)}, \tag{B2}$$

$$\begin{aligned} \frac{H}{\lambda} &= \lambda - p - 2\beta\beta'(1+\cos q) \\ &= \lambda - p + \alpha\lambda\beta^2(1+\cos q), \\ f_\lambda &= 1 + \gamma(p-\lambda). \end{aligned} \tag{B3}$$

Straightforward integration then reveals

$$\delta h_\frown/\varepsilon = 2\pi + 8\gamma\beta - 4\alpha\lambda\beta - 2\pi\gamma\alpha\lambda\beta^2. \tag{B4}$$

By changing the sign in front of β in Eq. (B2) and then otherwise performing the same integral over q in the negative direction, we similarly obtain

$$\delta h_\smile/\varepsilon = -2\pi + 8\gamma\beta - 4\alpha\lambda\beta + 2\pi\gamma\alpha\lambda\beta^2. \tag{B5}$$

[1] D. Dobrott and J. M. Greene, Probability of trapping-state transitions in a toroidal device, Phys. Fluids **14**, 1525 (1971).

[2] J. Henrard, Capture into resonance: An extension of the use of adiabatic invariants, Celestial Mech. **27**, 3 (1982).

[3] A. Neishtadt, Passage through a separatrix in a resonance problem with a slowly-varying parameter, J. Appl. Math. Mech. **39**, 594 (1975).

[4] B. V. Chirikov, Particle confinement and adiabatic invariance, Proc. R. Soc. Lond. **413**, 145 (1987).

[5] A. V. Timofeev, On the constancy of an adiabatic invariant when the nature of the motion changes, Zh. Eksp. Teor. Fiz **75**, 1303 (1978) [Sov. Phys. JETP **48**, 656 (1978)].

[6] J. R. Cary, D. F. Escande, and J. L. Tennyson, Adiabatic-invariant change due to separatrix crossing, Phys. Rev. A **34**, 4256 (1986).

[7] J. H. Hannay, Accuracy loss of action invariance in adiabatic change of a one-freedom hamiltonian, J. Phys. A: Math. Gen. **19**, L1067 (1986).

[8] J. R. Cary and R. T. Skodje, Reaction Probability for Sequential Separatrix Crossings, Phys. Rev. Lett. **61**, 1795 (1988).

[9] Y. Elskens and D. F. Escande, Slowly pulsating separatrices sweep homoclinic tangles where islands must be small: An extension of classical adiabatic theory, Nonlinearity **4**, 615 (1991).

[10] A. A. Chernikov and G. Schmidt, Adiabatic chaos in Josephson-junction arrays, Phys. Rev. E **50**, 3436 (1994).

[11] S. Chow and T. Young, A geometric proof of separatrix crossing results, Nonlinear Anal.: Theory, Methods Appl. **56**, 1047 (2004).

[12] Z. Lu, C. Jarzynski, and E. Ott, Apparent topologically forbidden interchange of energy surfaces under slow variation of a hamiltonian, Phys. Rev. E **91**, 052913 (2015).

[13] D. Bénisti and L. Gremillet, Global change in action due to trapping: How to derive it whatever the rate of variation of the dynamics, Phys. Rev. E **91**, 042915 (2015).

[14] D. Bénisti, Nonlocal adiabatic theory. I. The action distribution function, Phys. Plasmas **24**, 092120 (2017).

[15] A. A. Chernikov, G. Schmid, and A. I. Neishtadt, Unlimited Particle Acceleration by Waves in a Magnetic Field, Phys. Rev. Lett. **68**, 1507 (1992).

[16] S. M. Carioli, A. A. Chernikov, and A. I. Neishtadt, Stochastic and coherent acceleration of relativistic particles in a regular potential wavepacket perpendicular to a uniform magnetic field - I, Phys. Scr. **40**, 707 (1989).

[17] R. Cappi and M. Giovannozzi, Novel Method for Multiturn Extraction: Trapping Charged Particles in Islands of Phase Space, Phys. Rev. Lett. **88**, 104801 (2002).

[18] T. Eichmann, E. P. Thesing, and J. R. Anglin, Engineering separatrix volume as a control technique for dynamical transitions, Phys. Rev. E **98**, 052216 (2018).

[19] A. I. Neishtadt, Probability phenomena due to separatrix crossing, Chaos **1**, 42 (1991).

[20] A. Neishtadt, Averaging method for systems with separatrix crossing, Nonlinearity **30**, 2871 (2017).

[21] P. Goldreich and S. Peale, Spin-orbit coupling in the solar system, Astron. J. **71**, 425 (1966).

[22] D. M. Fieguth, T. Schlachter, D. S. Brady, and J. R. Anglin, Hamiltonian active particles in an environment, Phys. Rev. E **106**, 044201 (2022).

[23] M. Freidlin and A. Wentzell, *Random Perturbations of Dynamical Systems* (Springer Science & Business Media, Verlag, 1998), Vol. 260.