OPEN ACCESS
Stochastics and Dynamics
(2023) 2350020 (31 pages)

DOI: 10.1142/S021949372350020X

World Scientific
www.worldscientific.com

Physical measures of asymptotically autonomous dynamical systems

Julian Newman* and Peter Ashwin†

Department of Mathematics and Statistics, University of Exeter
Harrison Building, North Park Road, Exeter, EX4 4QF, UK
**j.m.i.newman@exeter.ac.uk*
†p.ashwin@exeter.ac.uk

Received 6 May 2022
Revised 1 November 2022
Accepted 2 November 2022
Published 10 February 2023

Although chaotic attractors for autonomous dynamical systems show sensitive dependence on initial conditions, they also typically support a physical or natural measure that characterizes the statistical behavior of almost all initial conditions near the attractor with respect to a background measure such as Lebesgue. In this paper, we identify conditions under which a nonautonomous system that limits as $t \to -\infty$ to an autonomous system with a physical measure is guaranteed to possess a "nonautonomous physical measure" that limits to the physical measure of the autonomous system.

Keywords: Nonautonomous dynamical systems; asymptotically autonomous dynamical systems; physical measures; pullback attraction.

AMS Subject Classification: 37C60, 37B55

1. Introduction

For an autonomous dynamical system, an invariant measure is "physical" or "natural" (or, assuming some extra properties, "SRB") [13] if it describes the long-term statistics (i.e. statistics after transients have decayed) of a typical trajectory. In practical terms, one can think of this as the observed statistics in the present of a typical trajectory that started an arbitrarily long time ago in the past. To apply this concept to systems where there is time-varying forcing, we need an analogous notion defined for *nonautonomous* dynamical systems, where the measure is not fixed but evolves in time under the action of the nonautonomous system [9]. The significance of such nonautonomous physical measures for application to interpreting climate statistics has been highlighted by several authors [3, 6, 7, 10].

*Corresponding author.

In the autonomous setting, a physical measure on a local attractor A is defined as an invariant measure μ supported on A such that the trajectory $x(t)$ of Lebesgue-almost every initial condition in the basin of attraction of A is *μ-generic*; this means that the empirical measure $\frac{1}{T}\int_0^T \delta_{x(t)}\,dt$ converges weakly to μ as $T \to \infty$. A related concept to physical measures is what we will call an *attracting measure* [3]: an attracting measure on a local attractor A is an invariant measure μ such that, if an initial condition $x(0)$ is selected at random from the basin of attraction of A with a probability distribution ν_0 that is Lebesgue-absolutely continuous, then the law ν_t of the position $x(t)$ at time t converges weakly to μ as $t \to \infty$. The main result of the classical paper [5] (Theorems 5.1 and 5.3) is that under mild conditions, an Axiom A attractor supports a mixing invariant measure that is both physical and attracting. There are also autonomous systems with a non-Axiom A chaotic attractor supporting an invariant measure that is physical and/or attracting [2, 12].

In the paper [3], the authors proposed a definition for physical measures and attracting measures supported on local *pullback* attractors of nonautonomous dynamical systems that limit to an autonomous system as $t \to -\infty$, and showed that this can be used to define a rigorous notion of "tipping probability" in the case that the nonautonomous system also limits to an autonomous system in the limit as $t \to +\infty$.

This paper gives sufficient conditions under which one can show that such nonautonomous physical and attracting measures exist. This paper extends existing work [1, 4] on asymptotically autonomous systems; in particular, [4] considers nonautonomous dynamical systems that limit as $t \to -\infty$ to an autonomous system with a stable fixed point, and addresses the question of whether the nonautonomous system possesses a singleton pullback-attractor that limits to the fixed point. In this paper, we now extend these results to a more general case where the nonautonomous physical measure limits to an autonomous physical measure that will typically be chaotic. For example, in climate applications one would like to consider models that have internal chaotic variability associated with turbulent transport and weather variability, rather than purely static states. One of the main technical tools in our proofs is application of various versions of Gronwall's Lemma [8, Sec. 1.1].

The structure of the paper is as follows. In Sec. 2.1, we give basic definitions in the abstract setting of continuous autonomous and nonautonomous dynamical systems on a compact metric space. In Sec. 2.2, we give general results regarding existence of pullback-attracting orbits and pullback-attracting measure-valued orbits for nonautonomous systems in this abstract setting. Then, in Sec. 2.3, we obtain results (Theorem 2.4) regarding existence of nonautonomous physical and attracting measures in the setting of asymptotically autonomous differential equations on a compact subset of $\mathbb{R}^N$, through application of the abstract results of Sec. 2.2. All the results stated in Sec. 2 are proved in Sec. 3. Some final remarks regarding further questions to address are then given in Sec. 4. In Appendix A, we present an alternative set of sufficient conditions for the conclusions of Theorems 2.1 and 2.2.

2. Nonautonomous Physical Measures and Their Existence

2.1. *Notions of attractivity of points and measures in a general topological setting*

Given a compact metric space (X, d), we write M_X for the set of Borel probability measures on X. We write $\pi_1, \pi_2 : X \times X \to X$ for the coordinate projections $\pi_i(x_1, x_2) = x_i$. Given $\mu_1, \mu_2 \in M_X$, let

$$\mathcal{J}(\mu_1, \mu_2) = \{\mathbb{P} \in M_{X\times X} : \pi_{i*}\mathbb{P} = \mu_i \text{ for } i = 1, 2\}.$$

We equip M_X with the 1-*Wasserstein metric* d_W given by

$$d_W(\mu_1, \mu_2) = \min_{\mathbb{P}\in\mathcal{J}(\mu_1,\mu_2)} \mathbb{E}_{\mathbb{P}}[d(\pi_1, \pi_2)].$$

It is well known that the topology induced by d_W is precisely the topology of weak convergence.

For any measurable function $f : X \to X$, we write $f_* : M_X \to M_X$ for the map sending a measure μ to its pushforward $f_*\mu$. Note that for measurable functions $f, \tilde{f} : X \to X$, $(f \circ \tilde{f})_* = f_* \circ \tilde{f}_*$. Note also that if f is continuous on X then f_* is continuous on M_X; and the dominated convergence theorem gives that if (f_n) is a sequence of measurable functions converging pointwise to f, then f_{n*} converges pointwise to f_*.

2.1.1. *Autonomous dynamical systems*

We first introduce autonomous dynamical systems and related definitions.

Definition 2.1. An *autonomous dynamical system on* X is a $[0, \infty)$-indexed family $(\Phi^t)_{0\leq t<\infty}$ of continuous functions $\Phi^t : X \to X$ such that: $\Phi^{s+t} = \Phi^t \circ \Phi^s$ for all $s, t \geq 0$ and $\Phi^0 = \mathrm{id}_X$, and the map $t \mapsto \Phi^t(x)$ from $[0, \infty)$ to X is continuous for all $x \in X$. A *forward-orbit* of the autonomous dynamical system $(\Phi^t)_{t\geq 0}$ is a $[0, \infty)$-indexed path in X of the form $(\Phi^t(x))_{t\geq 0}$ for some $x \in X$. A *backward-orbit* of $(\Phi^t)_{t\geq 0}$ is a $(-\infty, 0]$-indexed path $(x_t)_{t\leq 0}$ in X such that $\Phi^t(x_{s-t}) = x_s$ for all $s \leq 0$ and $t \geq 0$.

Note that if $(\Phi^t)_{t\geq 0}$ is an autonomous dynamical system on X, then $(\Phi^t_*)_{t\geq 0}$ is an autonomous dynamical system on M_X.

Definition 2.2. Given $x \in X$, $O \subset X$, and an autonomous dynamical system (Φ^t) on X, we say that *x is eventually in O under* (Φ^t) if there exists $T \geq 0$ such that for all $t \geq T$, $\Phi^t(x) \in O$.

Definition 2.3. Given $x, p \in X$ and an autonomous dynamical system (Φ^t) on X, we say that *x is attracted to p under* (Φ^t) if $\Phi^t(x) \to p$ as $t \to \infty$.

(This implies in particular that p is a fixed point, i.e. $\Phi^t(p) = p$ for all $t \geq 0$.)

Definition 2.4. Given $\lambda \in M_X$, $\mathcal{O} \subset M_X$, and an autonomous dynamical system (Φ^t) on X, we say that λ *is Cesàro-eventually in* $\mathcal{O}$ *under* (Φ^t_*) if there exists $T_1 \geq 0$ such that for all $T > T_1$ and $s \geq 0$, $\frac{1}{T}\int_0^T \Phi_*^{s+t}\lambda\, dt \in \mathcal{O}$.

Definition 2.5. Given $\lambda, \mu \in M_X$ and an autonomous dynamical system (Φ^t) on X, we say that λ *is Cesàro-attracted to* μ *under* (Φ^t_*) if $\frac{1}{T}\int_0^T \Phi^t_*\lambda\, dt \to \mu$ as $T \to \infty$.

(This implies in particular that μ is an invariant measure, i.e. $\Phi^t_*\mu = \mu$ for all $t \geq 0$.)

2.1.2. *Nonautonomous dynamical systems*

We now introduce nonautonomous dynamical systems and analogous related definitions.

Definition 2.6. A *nonautonomous dynamical system on* X is a family $(\Phi_{s,t})_{-\infty<s\leq t\leq 0}$ of continuous functions $\Phi_{s,t} : X \to X$ such that: $\Phi_{s,u} = \Phi_{t,u} \circ \Phi_{s,t}$ for all $s \leq t \leq u \leq 0$ and $\Phi_{t,t} = \mathrm{id}_X$ for all $t \leq 0$, and the map $t \mapsto \Phi_{s,t}(x)$ from $[s, 0]$ to X is continuous for all $s < 0$ and $x \in X$. An *orbit of* $(\Phi_{s,t})_{s\leq t\leq 0}$ is a $(-\infty, 0]$-indexed path $(x_t)_{t\leq 0}$ in X such that $\Phi_{s,t}(x_s) = x_t$ for all $s \leq t \leq 0$.

Typically, one defines nonautonomous dynamical systems over two-sided time, i.e. with subscripts $-\infty < s \leq t < +\infty$; but in this paper it is sufficient just to consider the past, i.e. with the time subscripts only going up to some finite number, which without loss of generality we take to be 0. Accordingly, we do not define "forward orbits", but use the term "orbit" to refer to what is, in effect, a backward orbit.

Note that if $(\Phi_{s,t})_{s\leq t\leq 0}$ is a nonautonomous dynamical system on X then $(\Phi_{s,t*})_{s\leq t\leq 0}$ is a nonautonomous dynamical system on M_X.

In analogy to Definition 2.2, we have the following definition.

Definition 2.7. Given $x \in X$, $O \subset X$, and a nonautonomous dynamical system $(\Phi_{s,t})$ on X, we say that x *is pullback-eventually in* O *under* $(\Phi_{s,t})$ if there exist $T_1, T_2 \geq 0$ such that for all $t \leq -T_2$, for all $s \leq t - T_1$, $\Phi_{s,t}(x) \in O$.

In analogy to Definition 2.3, we can define for nonautonomous systems both a notion of attraction to a point (Definition 2.8) and a notion of attraction to an orbit (Definition 2.9). In our abstract results in Sec. 2.2, the concept in Definition 2.8 will appear in the conditions, and the concept in Definition 2.9 (which is fundamental to the definition of "attracting measures" in Sec. 2.3) will appear in the conclusions; but the attraction described in Definition 2.8 will itself be verified in the setting of asymptotically autonomous differential equations (Sec. 2.3) under the conditions of Theorem 2.4.

Definition 2.8. Given $x, p \in X$, a nonautonomous dynamical system $(\Phi_{s,t})$ on X, and a value $r \geq 0$, we say that *x is past-attracted to p under $(X, d, (\Phi_{s,t}))$ with nonautonomous error of decay rate r* if the following statement holds: there exists a pair of functions $R_1, R_2 : (0, \infty) \to [0, \infty)$ such that for all $\varepsilon_1, \varepsilon_2 > 0$, for all $t \leq -R_2(\varepsilon_2)$, for all $s \leq t - R_1(\varepsilon_1)$, we have

$$d(\Phi_{s,t}(x), p) < \varepsilon_1 + \varepsilon_2 e^{rt}. \tag{2.1}$$

Remark 2.1. If the nonautonomous dynamical system $(\Phi_{s,t})_{s \leq t \leq 0}$ reduces to an autonomous dynamical system $(\Phi^t)_{t \geq 0}$ by $\Phi_{s,t} = \Phi^{t-s}$, then "removing the nonautonomous error term" recovers the definition of attraction in Definition 2.3; to be precise, the statement that x is attracted to p under (Φ^t) is equivalent to the following: There exists a function $R_1 : (0, \infty) \to [0, \infty)$ such that for all $\varepsilon > 0$, for all $t \leq 0$ and $s \leq t - R_1(\varepsilon)$, we have

$$d(\Phi_{s,t}(x), p) < \varepsilon.$$

Definition 2.9. Given $x \in X$, a nonautonomous dynamical system $(\Phi_{s,t})$ on X, and an orbit (p_t) of $(\Phi_{s,t})$, we say that *x is pullback-attracted to (p_t) under $(\Phi_{s,t})$* if for each $t \leq 0$, $\Phi_{s,t}(x) \to p_t$ as $s \to -\infty$.

Remark 2.2. If the nonautonomous dynamical system $(\Phi_{s,t})_{s \leq t \leq 0}$ reduces to an autonomous dynamical system $(\Phi^t)_{t \geq 0}$ by $\Phi_{s,t} = \Phi^{t-s}$, then Definition 2.9 reduces to saying that there exists $p \in X$ such that $p_t = p$ for all $t \leq 0$ and x is attracted to p under (Φ^t).

In analogy to Definition 2.4, we have the following definition.

Definition 2.10. Given $\lambda \in M_X$, $\mathcal{O} \subset M_X$, and a nonautonomous dynamical system $(\Phi_{s,t})$ on X, we say that *λ is pullback-Cesàro-eventually in $\mathcal{O}$ under $(\Phi_{s,t*})$* if there exist $T_1, T_2 \geq 0$ such that for all $\sigma \leq t \leq -T_2$ and $T > T_1$, $\frac{1}{T} \int_{\sigma-T}^{\sigma} \Phi_{s,t*}\lambda \, ds \in \mathcal{O}$.

Remark 2.3. Suppose $\mathcal{O}$ takes the form $\{\lambda \in M_X : \lambda(O) = 1\}$ for some Borel set $O \subset X$. If λ is pullback-eventually in $\mathcal{O}$ under $(\Phi_{s,t*})$ then λ is pullback-Cesàro-eventually in $\mathcal{O}$ under $(\Phi_{s,t*})$ (with the same T_1, T_2 from Definition 2.7 working in Definition 2.10).

In analogy to Definition 2.5, we can define for nonautonomous systems both a notion of Cesàro-attraction to a measure (Definition 2.11) and a notion of Cesàro-attraction to a measure-valued orbit (Definition 2.12).

Definition 2.11. Given $\lambda, \mu \in M_X$, a nonautonomous dynamical system $(\Phi_{s,t})$ on X, and a value $r \geq 0$, we say that *λ is past-Cesàro-attracted to μ under*

$(M_X, d_W, (\Phi_{s,t*}))$ *with nonautonomous error of decay rate* r if the following statement holds: there exists a pair of functions $R_1, R_2 : (0,\infty) \to [0,\infty)$ such that for all $\varepsilon_1, \varepsilon_2 > 0$, for all $t \le -R_2(\varepsilon_2)$ and $T > R_1(\varepsilon_1)$, we have

$$d_W\left(\frac{1}{T}\int_{t-T}^{t} \Phi_{s,t*}\lambda\, ds, \mu\right) < \varepsilon_1 + \varepsilon_2 e^{rt}. \tag{2.2}$$

Remark 2.4. If the nonautonomous dynamical system $(\Phi_{s,t})_{s\le t\le 0}$ reduces to an autonomous dynamical system $(\Phi^t)_{t\ge 0}$ by $\Phi_{s,t} = \Phi^{t-s}$, then "removing the nonautonomous error term" recovers the definition of Cesàro attraction in Definition 2.5; to be precise, the statement that λ is Cesàro-attracted to μ under (Φ^t_*) is equivalent to the following: There exists a function $R_1 : (0,\infty) \to [0,\infty)$ such that for all $\varepsilon > 0$, for all $t \le 0$ and $T > R_1(\varepsilon)$, we have

$$d_W\left(\frac{1}{T}\int_{t-T}^{t} \Phi_{s,t*}\lambda\, ds, \mu\right) < \varepsilon.$$

Definition 2.12. Given $\lambda \in M_X$, a nonautonomous dynamical system $(\Phi_{s,t})$ on X, and an orbit (μ_t) of $(\Phi_{s,t*})$, we say that λ *is pullback-Cesàro-attracted to* (μ_t) *under* $(\Phi_{s,t*})$ if for each $t \le 0$, $\frac{1}{T}\int_{t-T}^{t} \Phi_{s,t*}\lambda\, ds \to \mu_t$ as $T \to \infty$.

Remark 2.5. If the nonautonomous dynamical system $(\Phi_{s,t})_{s\le t\le 0}$ reduces to an autonomous dynamical system $(\Phi^t)_{t\ge 0}$ by $\Phi_{s,t} = \Phi^{t-s}$, then Definition 2.12 reduces to saying that there exists $\mu \in M_X$ such that $\mu_t = \mu$ for all $t \le 0$ and λ is Cesàro-attracted to μ under (Φ^t_*).

2.2. *Conditions for existence of pullback-attracting orbits in X and M_X*

Definition 2.13. Given $r > 0$, $p \in X$, and a nonautonomous dynamical system $(\Phi_{s,t})$ on X, a *nonautonomousness-controller* (*NAC*) *of* $(X, d, (\Phi_{s,t}))$ *near* p *of rate* r is a triple (O, T^*, q) consisting of a closed set $O \subset X$, a value $T^* \le 0$, and a continuous function $q : (-\infty, T^*] \to O$ such that the following statements hold:

(i) for all $t < T^*$, $\frac{1}{h} d(\Phi_{t,t+h}(q(t)), q(t)) \to 0$ as $h \to 0+$;
(ii) $\int_{-\infty}^{T^*} d(q(t), p) e^{r|t|}\, dt < \infty$;
(iii) $\Phi_{s,t}(p) \in O$ for all $s \le t \le T^*$.

Remark 2.6. Condition (i) says, heuristically, that $q(t)$ is an "equilibrium of the instantaneous dynamics at time t". In the case that X is a compact subset of $\mathbb{R}^N$ and $(\Phi_{s,t})$ is the solution flow of a smooth nonautonomous differential equation $\dot{x}(t) = f_t(x(t))$, condition (i) is equivalent to saying that $f_t(q(t)) = 0$ for all $t \le T^*$. Condition (ii) describes a kind of "sufficiently fast convergence" of q to p as time tends to $-\infty$; so p functions as a kind of "approximate equilibrium for large negative times". (Of course, condition (ii) does not specify a strict convergence $q(t) \to p$ as $t \to -\infty$, but this will follow when one assumes the extra condition of being "monotone-like" as in Definition 2.14.) It will be important that we do not require the set O to be a neighborhood of p, but we nonetheless impose condition (iii).

Definition 2.14. We say that a NAC (O, T^*, q) of $(X, d, (\Phi_{s,t}))$ near p of rate r is a *monotone-like NAC* (*MLNAC*) if there exists $C \geq 1$ such that for all $s < T^*$, there exists $\delta(s) > 0$ such that for all $t \in (s, s + \delta(s)]$,

$$d(q(t), q(s)) \leq C(d(q(t), p) - d(q(s), p)). \tag{2.3}$$

Remark 2.7. Definition 2.14 together with the continuity of q implies that the map $t \mapsto d(q(t), p)$ is monotone increasing. If we suppose that O is identified as a subset of $\mathbb{R}$ with $d(x, y) = |x - y|$, then Definition 2.14 is equivalent to saying that q is monotone, in which case Eq. (2.3) holds as an equality with $C = 1$.

Definition 2.15. Given $r > 0$ and a nonautonomous dynamical system $(\Phi_{s,t})$ on X, a *growth-controller* (*GC*) *of* $(X, d, (\Phi_{s,t}))$ *of rate* r is a pair (O, T^*) consisting of a set $O \subset X$ and a value $T^* \leq 0$ such that for all $t < T^*$ and $x, y \in O$, $\liminf_{h \to 0+} \frac{1}{h}[d(\Phi_{t,t+h}(x), \Phi_{t,t+h}(y)) - d(x, y)] \leq r d(x, y)$.

Remark 2.8. Heuristically, Definition 2.15 says that in O, the speed of separation of two trajectories is controlled by their distance from each other. In the case that X is a compact subset of $\mathbb{R}^N$ and $(\Phi_{s,t})$ is the solution flow of a smooth nonautonomous differential equation $\dot{x}(t) = f_t(x(t))$, it is equivalent to the "one-sided Lipschitz condition" that $(x - y) \cdot (f_t(x) - f_t(y)) \leq r|x - y|^2$ for all $t \leq T^*$ and $x, y \in O$. The one-sided Lipschitz condition will be discussed further in Sec. 2.3 (especially Remark 2.10).

Note that

- the "rate" for a MLNAC is really a *lower bound*: if (O, T^*, q) is a MLNAC near p of some rate $r_1 > 0$, then it is also a MLNAC near p of any rate $r \in (0, r_1]$;
- the "rate" for a GC is really an *upper bound*: if (O, T^*) is a GC of some rate $r_2 > 0$, then it is also a GC of any rate $r \in [r_2, \infty)$.

So, we now consider what happens if, for a given triple (O, T^*, q), the bounded-above interval of nonautonomousness-controller rates and the bounded-below interval of growth-controller rates overlap.

Theorem 2.1. *Suppose we have* $p \in X$ *and a nonautonomous dynamical system* $(\Phi_{s,t})$ *on* X. *Suppose we have* $r > 0$ *and a triple* (O, T^*, q) *such that* (O, T^*, q) *is a MLNAC of* $(X, d, (\Phi_{s,t}))$ *near* p *of rate* r *and* (O, T^*) *is a GC of* $(X, d, (\Phi_{s,t}))$ *of rate* r. *Then there exists an orbit* (p_t) *of* $(\Phi_{s,t})$ *such that* p *is pullback-attracted to* (p_t) *under* $(\Phi_{s,t})$. *Furthermore,* $d(p_t, p)$ *is* $o(e^{rt})$ *as* $t \to -\infty$.

Remark 2.9. The condition that the NAC (O, T^*, q) is monotone-like can also be replaced with a condition that it controls the instantaneous speed of trajectories in proportion to their distance from $q(t)$ with constant of proportion r; details of this are given in Appendix A. However, as explained in Appendix A, this gives a generally less applicable result.

Theorem 2.2. *Assume the setting of Theorem* 2.1*. Suppose we have $x \in X$ that is pullback-eventually in O under $(\Phi_{s,t})$ and is past-attracted to p under $(X, d, (\Phi_{s,t}))$ with nonautonomous error of decay rate r. Then x is pullback-attracted to (p_t) under $(\Phi_{s,t})$.*

Now, Theorems 2.1 and 2.2 have the following immediate consequences, by virtue of their application to the system $(\Phi_{s,t*})$ on the space of measures M_X.

Corollary 2.1. *Suppose we have $\mu \in M_X$ and a nonautonomous dynamical system $(\Phi_{s,t})$ on X. Suppose we have $r > 0$ and a triple $(\mathcal{O}, T^*, \nu)$ such that $(\mathcal{O}, T^*, \nu)$ is a MLNAC of $(M_X, d_W, (\Phi_{s,t*}))$ near μ of rate r and $(\mathcal{O}, T^*)$ is a GC of $(M_X, d_W, (\Phi_{s,t*}))$ of rate r. Then there exists an orbit (μ_t) of $(\Phi_{s,t*})$ such that μ is pullback-attracted to (μ_t) under $(\Phi_{s,t*})$. Furthermore, $d_W(\mu_t, \mu)$ is $o(e^{rt})$ as $t \to -\infty$.*

Corollary 2.2. *Assume the setting of Corollary* 2.1*. Suppose we have $\lambda \in M_X$ that is pullback-eventually in $\mathcal{O}$ under $(\Phi_{s,t*})$ and is past-attracted to μ under $(M_X, d_W, (\Phi_{s,t*}))$ with nonautonomous error of decay rate r. Then λ is pullback-attracted to (μ_t) under $(\Phi_{s,t*})$.*

Having considered attraction in Corollary 2.2, we now consider Cesàro-attraction.

Theorem 2.3. *Assume the setting of Corollary* 2.1*. Suppose we have $\lambda \in M_X$ that is pullback-Cesàro-eventually in $\mathcal{O}$ under $(\Phi_{s,t*})$ and is past-Cesàro-attracted to μ under $(M_X, d_W, (\Phi_{s,t*}))$ with nonautonomous error of decay rate r. Then λ is pullback-Cesàro-attracted to (μ_t) under $(\Phi_{s,t*})$.*

2.3. *Application to asymptotically autonomous differential equations*

We now apply the results of Sec. 2.2 to the case of a nonautonomous dynamical system arising from a real-time parameter-drift $\gamma = \Gamma(t)$ through a parameterized family $(f_\gamma)_{\gamma \in \mathcal{I}}$ of vector fields f_γ with a continuous branch of invariant measures $(\tilde{\nu}(\gamma))_{\gamma \in \mathcal{I}}$.

The special case that $(\tilde{\nu}(\gamma))_{\gamma \in \mathcal{I}}$ corresponds to a branch of hyperbolic stable fixed points (with $\tilde{\nu}(\gamma)$ simply being the Dirac mass at the stable fixed point of f_γ) has been addressed in detail in [4]. Hence, in this paper, we are mainly concerned about when one is outside that special case.

Write

$$S_N := \{v \in \mathbb{R}^N : |v| = 1\},$$

and for each $A \in \mathbb{R}^{N\times N}$, let

$$\mathfrak{L}(A) = \max_{v\in S_N} v^{\mathrm{T}} A v.$$

The value of $\mathfrak{L}(A)$ can be computed as the maximum of the eigenvalues of the symmetric matrix $\frac{1}{2}(A + A^{\mathrm{T}})$. Given a compact metric space X, let K_X be the set of non-empty closed subsets of X, which we equip with the Hausdorff metric d_H.

2.3.1. *Conditions on the parameter-dependent vector field*

Suppose that X is a compact subset of $\mathbb{R}^N$, with d being the Euclidean distance $d(x, y) = |x - y|$. Let $\mathcal{F}$ be a set of C^1 vector fields on $\mathbb{R}^N$ with the property that for any $T > 0$ and any continuously (with respect to the C^1 topology) $[0, T]$-parameterized family $(\tilde{f}_t)_{0\leq t\leq T}$ of members $\tilde{f}_t$ of $\mathcal{F}$, any solution $x : [0, T] \to \mathbb{R}^N$ of the differential equation

$$\dot{x}(t) = \tilde{f}_t(x(t)) \tag{2.4}$$

with $x(0) \in X$ has $\{x(t)\}_{0\leq t\leq T} \subset X$.

Now, suppose we have an interval $\mathcal{I} \subset \mathbb{R}$ that includes its lower end-point $\gamma_0 \in \mathcal{I} \cap \partial\mathcal{I}$, and suppose we have a continuously (with respect to the C^1 topology) $\mathcal{I}$-parameterized family $(f_\gamma)_{\gamma\in\mathcal{I}}$ of vector fields $f_\gamma \in \mathcal{F}$. Suppose we have a continuous function $\tilde{\nu} : \mathcal{I} \to M_X$ such that

(a) for each $\gamma \in \mathcal{I}$, $\tilde{\nu}(\gamma)$ is an invariant measure of the autonomous differential equation

$$\dot{x}(t) = f_\gamma(x(t)), \quad t \geq 0; \tag{2.5}$$

(b) for each $\gamma \in \mathcal{I}$, $\tilde{Q}(\gamma) := \operatorname{supp} \tilde{\nu}(\gamma)$ is contained in the interior of X relative to $\mathbb{R}^N$;
(c) the map $\gamma \mapsto \tilde{Q}(\gamma)$ is continuous on $\mathcal{I}$;
(d) writing

$$\mu := \tilde{\nu}(\gamma_0)$$
$$P := \tilde{Q}(\gamma_0) = \operatorname{supp} \mu,$$

there exists $C \geq 1$ such that for every $\gamma \in \mathcal{I}$ there exists $\tilde{\delta}(\gamma) > 0$ such that for all $\gamma' \in (\gamma, \gamma + \tilde{\delta}(\gamma)]$,

$$d_W(\tilde{\nu}(\gamma'), \tilde{\nu}(\gamma)) \leq C(d_W(\tilde{\nu}(\gamma'), \mu) - d_W(\tilde{\nu}(\gamma), \mu)) \quad \text{and}$$
$$d_H(\tilde{Q}(\gamma'), \tilde{Q}(\gamma)) \leq C(d_H(\tilde{Q}(\gamma'), P) - d_H(\tilde{Q}(\gamma), P)).$$

(In (d), it is sufficient just to consider $\gamma \in \mathcal{I}\backslash\{\gamma_0\}$, since the inequalities with $\gamma = \gamma_0$ trivially hold with $C = 1$.)

For any non-empty bounded $O \subset \mathbb{R}^N$, let

$$L(O) := \inf\{r \in \mathbb{R} : \forall\, x, y \in O,\ (x-y)\cdot(f_{\gamma_0}(x) - f_{\gamma_0}(y)) \leq r|x-y|^2\}.$$

Remark 2.10. For any distinct $x, y \in \mathbb{R}^N$, letting $v := \frac{1}{|x-y|}(x-y) \in S_N$, the mean value theorem applied to the map $t \mapsto v \cdot f_{\gamma_0}(y+tv)$ on the interval $[0, |x-y|]$ gives that

$$(x-y)\cdot(f_{\gamma_0}(x) - f_{\gamma_0}(y)) = (v^{\mathrm{T}}(Jf_{\gamma_0})(\xi)v)|x-y|^2$$

for some ξ in the line-segment joining x and y. Hence, if O is an open convex set then

$$L(O) = \max_{x \in \bar{O}} \mathfrak{L}((Jf_{\gamma_0})(x)).$$

More generally, assuming that O contains at least two points, it is possible to cover $\bar{O}$ by a finite collection of open convex sets $O_1, \ldots, O_n$, and then $L(O)$ is bounded above by the finite value

$$\max\left(\max_{1\leq i\leq n} L(O_i), \min\left\{\frac{(x-y)\cdot(f_{\gamma_0}(x) - f_{\gamma_0}(y))}{|x-y|^2} : (x,y) \in (\bar{O}\times\bar{O})\backslash\bigcup_{i=1}^{n}(O_i \times O_i)\right\}\right).$$

Obviously, $L(O)$ is bounded below by any value of $\frac{(x-y)\cdot(f_{\gamma_0}(x)-f_{\gamma_0}(y))}{|x-y|^2}$ with distinct $x, y \in O$. Hence, if O has at least two points then $L(O)$ is finite, and it is then easy to see that the infimum in the definition of $L(O)$ is, in fact, a minimum. If, on the other hand, O is a singleton, then $L(O) = -\infty$.

Now let

$$r_0 := \max\left(L(P), \max_{x\in P} \mathfrak{L}((Jf_{\gamma_0})(x))\right). \tag{2.6}$$

Proposition 2.1. *If $r_0 < 0$ then P is a singleton $P = \{p\}$ for some hyperbolic stable fixed point p of f_{γ_0}; and hence, for all γ in a neighborhood of γ_0, $\tilde{Q}(\gamma)$ is a singleton $\tilde{Q}(\gamma) = \{\tilde{q}(\gamma)\}$ for some hyperbolic stable fixed point $\tilde{q}(\gamma)$ of f_γ.*

Hence, we will focus on the situation that $r_0 \geq 0$.

2.3.2. *The parameter-drift and the resulting nonautonomous dynamical system*

Given an increasing continuous function $\Gamma : [-\infty, 0] \to \mathcal{I}$ with $\Gamma(-\infty) = \gamma_0$, we will consider the "asymptotically autonomous" differential equation

$$\dot{x}(t) = f_{\Gamma(t)}(x(t)), \quad t \in (-\infty, 0]. \tag{2.7}$$

We define the "past-limit system" as Eq. (2.5) with $\gamma = \gamma_0$, i.e.

$$\dot{x}(t) = f_{\gamma_0}(x(t)), \quad t \geq 0. \tag{2.8}$$

Let $(\Phi_{s,t})_{s \leq t \leq 0}$ be the nonautonomous dynamical system on X generated by Eq. (2.7), and let $(\Psi^t)_{t \geq 0}$ be the autonomous dynamical system on X generated by Eq. (2.8). As detailed shortly (in Sec. 2.3.3), we will be concerned with the relationship between properties of the autonomous system $(\Psi^t)_{t \geq 0}$ and analogous properties of the nonautonomous system $(\Phi_{s,t})_{s \leq t \leq 0}$. For convenience, write

$$\nu(t) := \tilde{\nu}(\Gamma(t)),$$
$$Q(t) := \tilde{Q}(\Gamma(t)).$$

2.3.3. *Existence of nonautonomous attracting and physical measures*

We now give definitions of "natural measure" concepts, first in the autonomous setting (applied to the past-limit system (2.8)) and then in our nonautonomous setting.

Definition 2.16. We say that μ is

- an *attracting measure of the past-limit system* if there exists a neighborhood $U \subset X$ of P such that for every Lebesgue-absolutely continuous probability measure λ with $\lambda(U) = 1$, λ is attracted to μ under (Ψ^t_*);
- a *physical measure of the past-limit system* if there exists a neighborhood $U \subset X$ of P such that for Lebesgue-almost all $x \in U$, δ_x is Cesàro-attracted to μ under (Ψ^t_*);
- a *weakly physical measure* (or *Cesàro-attracting measure*) *of the past-limit system* if there exists a neighborhood $U \subset X$ of P such that for every Lebesgue-absolutely continuous probability measure λ with $\lambda(U) = 1$, λ is Cesàro-attracted to μ under (Ψ^t_*).

(Note that a weakly physical measure is both physical and attracting.) We now extend these notions to the nonautonomous setting.

Definition 2.17. An orbit (μ_t) of $(\Phi_{s,t*})$ is called

- an *attracting measure rooted at* μ if there exists a neighborhood $U \subset X$ of P such that for every Lebesgue-absolutely continuous probability measure λ with $\lambda(U) = 1$, λ is pullback-attracted to (μ_t) under $(\Phi_{s,t*})$;
- a *physical measure rooted at* μ if there exists a neighborhood $U \subset X$ of P such that for Lebesgue-almost all $x \in U$, δ_x is pullback-Cesàro-attracted to (μ_t) under $(\Phi_{s,t*})$.
- a *weakly physical measure* (or *Cesàro-attracting measure*) *rooted at* μ if there exists a neighborhood $U \subset X$ of P such that for every Lebesgue-absolutely

continuous probability measure λ with $\lambda(U) = 1$, λ is pullback-Cesàro-attracted to (μ_t) under $(\Phi_{s,t*})$.

Definition 2.18. We will say that P *is Lyapunov-stable under the past-limit system* if for every neighborhood V of P there exists a neighborhood $U \subset V$ of P such that for all $t \geq 0$, $\Psi^t(U) \subset V$.

Now, define the pseudometric d_{X,C^0} on the space of continuous vector fields on $\mathbb{R}^N$ by

$$d_{X,C^0}(f, g) = \max_{x \in X} |f(x) - g(x)|.$$

Theorem 2.4. *Assume that $r_0 \geq 0$, and fix an arbitrary $r > r_0$.*

(A) *Suppose $\int_{-\infty}^{0} d_W(\nu(t), \mu)e^{r|t|}\,dt < \infty$ and $\int_{-\infty}^{0} d_H(Q(t), P)e^{r|t|}\,dt < \infty$. Then there exists an orbit (μ_t) of $(\Phi_{s,t*})$ such that μ is pullback-attracted to (μ_t) under $(\Phi_{s,t*})$, and $d_W(\mu_t, \mu)$ is $o(e^{rt})$ as $t \to -\infty$.*

(B) *Suppose, moreover, that $\int_{-\infty}^{0} d_{X,C^0}(f_{\Gamma(t)}, f_{\gamma_0})e^{r|t|}\,dt < \infty$ and P is Lyapunov-stable under the past-limit system. If μ is an attracting measure (respectively, physical measure, weakly physical measure) of the past-limit system, then (μ_t) is an attracting measure (respectively, physical measure, weakly physical measure) rooted at μ.*

Strategy of the proof. For part (A), we show that there exists a closed neighborhood O of P and a value $T^* \leq 0$ such that, writing $\mathcal{O} := \{\lambda \in M_X : \lambda(O) = 1\}$, $(\mathcal{O}, T^*, \nu)$ is a MLNAC of $(M_X, d_W, (\Phi_{s,t*}))$ near μ of rate r and $(\mathcal{O}, T^*)$ is a GC of $(M_X, d_W, (\Phi_{s,t*}))$ of rate r; Corollary 2.1 then gives the desired result. For part (B), we show that there exists a neighborhood U of P such that for every $\lambda \in M_X$ with $\lambda(U) = 1$,

- λ is pullback-eventually in $\mathcal{O}$ under $(\Phi_{s,t*})$;
- if λ is attracted (respectively, Cesàro-attracted) to μ under (Ψ_*^t), then λ is past-attracted (respectively, past-Cesàro-attracted) to (μ_t) under $(M_X, d_W, (\Phi_{s,t*}))$ with nonautonomous error of decay rate r.

Corollary 2.2 then gives the desired result for attracting measures, and Theorem 2.3 (together with Remark 2.3) for physical and weakly physical measures.

3. Details of Proofs

For a point $x \in X$ (respectively, a set $A \subset X$) and a value $\delta > 0$, $B_\delta(x)$ (respectively, $B_\delta(A)$) denotes the open δ-neighborhood of x (respectively, of A) under the metric d.

3.1. *Proofs of the results for the general topological setting*

Lemma 3.1. *Suppose we have a nonautonomous dynamical system* $(\Phi_{s,t})$ *on* X, *a point* $p \in X$, *a value* $r > 0$, *and a GC* (O, T^*) *of* $(X, d, (\Phi_{s,t}))$ *of rate* r.

(A) *For all* $t \leq u \leq T^*$ *and all* $x, y \in O$, *if* $\Phi_{t,v}(x), \Phi_{t,v}(y) \in O$ *for all* $v \in [t, u)$, *then*

$$d(\Phi_{t,u}(x), \Phi_{t,u}(y)) \leq e^{r(u-t)} d(x, y). \tag{3.9}$$

Hence, if $\Phi_{s,t}(p) \in O$ *for all* $s \leq t \leq T^*$ *then, in particular, for all* $s_1, s_2 \leq t \leq u \leq T^*$, *setting* $x := \Phi_{s_1,t}(p)$ *and* $y := \Phi_{s_2,t}(p)$ *gives*

$$d(\Phi_{s_1,u}(p), \Phi_{s_2,u}(p)) \leq e^{r(u-t)} d(\Phi_{s_1,t}(p), \Phi_{s_2,t}(p)). \tag{3.10}$$

(B) *Suppose furthermore that we have a continuous function* $q : (-\infty, T^*] \to O$ *such that for all* $t < T^*$, $\frac{1}{h} d(\Phi_{t,t+h}(q(t)), q(t)) \to 0$ *as* $h \to 0+$. *Suppose also that we have* $C \geq 1$ *such that for all* $s < T^*$, *there exists* $\delta(s) > 0$ *such that for all* $t \in (s, s + \delta(s)]$, *Eq.* (2.3) *holds. For all* $s \leq u \leq T^*$, *if each* $t \in [s, u)$ *has* $\Phi_{s,t}(p) \in O$ *then*

$$d(\Phi_{s,u}(p), p) \leq (C + 1) d(q(u), p) + Cre^{ru} \int_{-\infty}^{u} d(q(t), p) e^{r|t|}\, dt. \tag{3.11}$$

Proof. (A) For each $v \in [t, u)$, since $\Phi_{t,v}(x), \Phi_{t,v}(y) \in O$, we have

$$\liminf_{h \to 0+} \frac{d(\Phi_{t,v+h}(x), \Phi_{t,v+h}(y)) - d(\Phi_{t,v}(x), \Phi_{t,v}(y))}{h} \leq rd(\Phi_{t,v}(x), \Phi_{t,v}(y)).$$

So by [11, Appendix I, Theorem 2.1], for all $v \in [t, u]$,

$$d(\Phi_{t,v}(x), \Phi_{t,v}(y)) \leq d(x, y) + \int_t^v rd(\Phi_{t,\tau}(x), \Phi_{t,\tau}(y))\, d\tau,$$

and therefore a suitable version of Gronwall's Lemma [8, Corollary 3] yields the result. (B) The triangle inequality gives

$$d(\Phi_{s,u}(p), p) \leq d(q(u), \Phi_{s,u}(p)) + d(q(u), p),$$

and obviously $d(q(t), p) e^{r|t|} \geq 0$ for all t; so, for the desired conclusion, it is sufficient that

$$d(q(u), \Phi_{s,u}(p)) \leq C \left(d(q(u), p) + re^{ru} \int_s^u d(q(t), p) e^{r|t|}\, dt \right). \tag{3.12}$$

Writing $d_t := d(q(t), \Phi_{s,t}(p)) - Cd(q(t), p)$, for each $t \in [s, u)$ we have

$$\liminf_{h \to 0+} \frac{d_{t+h} - d_t}{h}$$

$$= \liminf_{h \to 0+} \frac{d(q(t+h), \Phi_{s,t+h}(p)) - d(q(t), \Phi_{s,t}(p)) - C(d(q(t+h), p) - d(q(t), p))}{h}$$

$$\leq \liminf_{h\to 0+} \frac{d(q(t+h),\Phi_{s,t+h}(p)) - d(q(t),\Phi_{s,t}(p)) - d(q(t+h),q(t))}{h}$$

$$\leq \liminf_{h\to 0+} \frac{d(q(t),\Phi_{s,t+h}(p)) - d(q(t),\Phi_{s,t}(p))}{h}$$

$$\leq \liminf_{h\to 0+} \frac{d(q(t),\Phi_{t,t+h}(q(t))) + d(\Phi_{t,t+h}(q(t)),\Phi_{s,t+h}(p)) - d(q(t),\Phi_{s,t}(p))}{h}$$

$$= \liminf_{h\to 0+} \Bigg(\underbrace{\frac{d(q(t),\Phi_{t,t+h}(q(t)))}{h}}_{\to 0 \text{ as } h\to 0+ \text{ by assumption}} + \frac{d(\Phi_{t,t+h}(q(t)),\Phi_{s,t+h}(p)) - d(q(t),\Phi_{s,t}(p))}{h} \Bigg)$$

$$= \liminf_{h\to 0+} \frac{d(\Phi_{t,t+h}(q(t)),\Phi_{s,t+h}(p)) - d(q(t),\Phi_{s,t}(p))}{h}$$

$$\leq rd(q(t),\Phi_{s,t}(p)) \quad \text{since } \Phi_{s,t}(p) \in O,$$

and so by [11, Appendix I, Theorem 2.1], for all $t \in [s,u]$,

$$d_t - d_s \leq \int_s^t rd(q(\tau),\Phi_{s,\tau}(p))\,d\tau,$$

i.e.

$$d(q(t),\Phi_{s,t}(p)) \leq Cd(q(t),p) + \underbrace{(1-C)d(q(s),p)}_{\leq 0} + \int_s^t rd(q(\tau),\Phi_{s,\tau}(p))\,d\tau$$

$$\leq Cd(q(t),p) + \int_s^t rd(q(\tau),\Phi_{s,\tau}(p))\,d\tau.$$

A suitable version of Gronwall's Lemma [8, Theorem 1] then yields Eq. (3.12). □

Lemma 3.2. *Given $r > 0$, $T^* \in \mathbb{R}$ and an increasing function $d\colon(-\infty,T^*] \to [0,\infty)$, if $\int_{-\infty}^{T^*} d(t)e^{r|t|}\,dt < \infty$, then $d(t)e^{r|t|} \to 0$ as $t \to -\infty$.*

(Note that this is not completely obvious, as $d(t)$ is increasing in t but $e^{r|t|}$ is decreasing in t, and so $d(t)e^{r|t|}$ can be non-monotone.)

Proof. For $t \leq T^* - 1$, we have

$$d(t)e^{r|t|} = \frac{r}{1-e^{-r}} d(t) \int_t^{t+1} e^{r|\tau|}\,d\tau$$

$$\leq \frac{r}{1-e^{-r}} \int_t^{t+1} d(\tau)e^{r|\tau|}\,d\tau \to 0 \text{ as } t \to -\infty.$$

□

We can now prove Theorems 2.1 and 2.2.

Proof of Theorem 2.1. By continuity of the mappings $\Phi_{s,t}$, if the limit $p_t := \lim_{s\to-\infty}\Phi_{s,t}(p)$ exists for all $t \le T^*$, then this limit exists for all $t \le 0$ and $(p_t)_{t\le 0}$ is an orbit of $(\Phi_{s,t})$.

We will show that for each $t \le T^*$, $\Phi_{s,t}(p)$ is d-Cauchy as $s \to -\infty$, from which it follows that the limit p_t exists. Fix t, and fix any $\varepsilon > 0$. We need to find $T \le t$ such that every $s_1 \le s_2 \le T$ has $d(\Phi_{s_1,t}(p), \Phi_{s_2,t}(p)) < \varepsilon$. Let C be as in Definition 2.14. As in Remark 2.7, the map $\tau \mapsto d(q(\tau), p)$ is increasing on $(-\infty, T^*]$; so the assumption that $\int_{-\infty}^{T^*} d(q(\tau),p)e^{r|\tau|}\,d\tau < \infty$ together with Lemma 3.2 implies that we can find $T \le t$ such that for all $s \le T$,

$$(C+1)d(q(s),p)e^{r|s|} + Cr\int_{-\infty}^{s} d(q(\tau),p)e^{r|\tau|}\,d\tau < \varepsilon.$$

Lemma 3.1(B) then gives that for all $s_1 \le s_2 \le T$,

$$d(\Phi_{s_1,s_2}(p),p) < \varepsilon e^{rs_2} \le \varepsilon e^{r(s_2-t)}. \tag{3.13}$$

Hence, Eq. (3.10) in Lemma 3.1(A) (with "u" set to t, and "t" and "s_2" both set to s_2) gives that $d(\Phi_{s_1,t}(p), \Phi_{s_2,t}(p)) < \varepsilon$ as required.

We next show that $d(p_t,p)e^{r|t|} \to 0$ as $t \to -\infty$. Fixing ε and taking T as above, we have that for all $s_2 \le T$, Eq. (3.13) holds for all $s_1 \le s_2$, and therefore (by taking $s_1 \to -\infty$), $d(p_{s_2},p) \le \varepsilon e^{rs_2}$. ◻

Proof of Theorem 2.2. First note that since O is closed and $\Phi_{s,t}(p) \in O$ for all $s \le t \le T^*$, it follows that $p_t \in O$ for all $t \le T^*$. Let $T_1, T_2 \ge 0$ be as in Definition 2.7, and let R_1, R_2 be as in Definition 2.8. Due to the continuity of the mappings $\Phi_{s,t}$, it is sufficient to fix $t \le \min(T^*, -T_2)$ and show that $\Phi_{s,t}(x) \to p_t$ as $s \to -\infty$. Fix $\varepsilon > 0$. Since $d(p_s,p)$ is $o(e^{rs})$ as $s \to -\infty$, we can take a value $\tau < \min(t, -R_2(\frac{1}{3}\varepsilon e^{r|t|}))$ sufficiently large in magnitude that $d(p_s,p) < \frac{1}{3}\varepsilon e^{r(s-t)}$ for all $s \le \tau$.

Now take any $s < \tau - \max(T_1, R_1(\frac{1}{3}\varepsilon e^{r(\tau-t)}))$; we will show that $d(\Phi_{s,t}(x), p_t) < \varepsilon$. For all v in the interval $[\tau, t]$, we have that $v \le -T_2$ (since $t \le -T_2$) and $s \le v - T_1$ (since $s \le \tau - T_1$), and therefore $\Phi_{s,v}(x) \in O$. Therefore, by Eq. (3.9) in Lemma 3.1(A), in order to show that $d(\Phi_{s,t}(x), p_t) < \varepsilon$ it is sufficient to show that $d(\Phi_{s,\tau}(x), p_\tau) < \varepsilon e^{r(\tau-t)}$. But we know that $d(p_\tau, p) < \frac{1}{3}\varepsilon e^{r(\tau-t)}$; and Eq. (2.1) with "$t$" set to τ, and with $\varepsilon_1 := \frac{1}{3}\varepsilon e^{r(\tau-t)}$ and $\varepsilon_2 := \frac{1}{3}\varepsilon e^{r|t|}$, gives $d(\Phi_{s,\tau}(x), p) < \frac{2}{3}\varepsilon e^{r(\tau-t)}$. So $d(\Phi_{s,\tau}(x), p_\tau) < \varepsilon e^{r(\tau-t)}$ as required. ◻

The proof of Theorem 2.3 is, unsurprisingly, similar in nature to the proof of Theorem 2.2, but requires a small amount more work.

Lemma 3.3. *For any $\varepsilon \in [0,1]$ and $\mu_1, \mu_2 \in M_X$, letting*

$$\nu_\varepsilon := (1-\varepsilon)\mu_1 + \varepsilon\mu_2 \in M_X,$$

we have

$$d_W(\mu_1, \nu_\varepsilon) \le \varepsilon d_W(\mu_1, \mu_2) \le \varepsilon \mathrm{diam}(X).$$

Proof. Let $\mathbb{P}_1$ be the pushforward measure of μ_1 under the map $x \mapsto (x, x)$. Let $\mathbb{P}_2 \in \mathcal{J}(\mu_1, \mu_2)$ be such that $d_W(\mu_1, \mu_2) = \mathbb{E}_{\mathbb{P}_2}[d(\pi_1, \pi_2)]$. Let $\mathbb{P} = (1-\varepsilon)\mathbb{P}_1 + \varepsilon\mathbb{P}_2$. We will show that $\mathbb{P} \in \mathcal{J}(\mu_1, \nu_\varepsilon)$ and $\mathbb{E}_{\mathbb{P}}[d(\pi_1, \pi_2)] = \varepsilon d_W(\mu_1, \mu_2)$.

We have

$$\pi_{1*}\mathbb{P} = (1-\varepsilon)\pi_{1*}\mathbb{P}_1 + \varepsilon\pi_{1*}\mathbb{P}_2 = (1-\varepsilon)\mu_1 + \varepsilon\mu_1 = \mu_1,$$

$$\pi_{2*}\mathbb{P} = (1-\varepsilon)\pi_{2*}\mathbb{P}_1 + \varepsilon\pi_{2*}\mathbb{P}_2 = (1-\varepsilon)\mu_1 + \varepsilon\mu_2 = \nu_\varepsilon,$$

so $\mathbb{P} \in \mathcal{J}(\mu_1, \nu_\varepsilon)$. Now

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}[d(\pi_1, \pi_2)] &= (1-\varepsilon)\mathbb{E}_{\mathbb{P}_1}[d(\pi_1, \pi_2)] + \varepsilon\mathbb{E}_{\mathbb{P}_2}[d(\pi_1, \pi_2)] \\ &= (1-\varepsilon)\mathbb{E}_{\mu_1}[d(\mathrm{id}_X, \mathrm{id}_X)] + \varepsilon d_W(\mu_1, \mu_2) \\ &= \varepsilon d_W(\mu_1, \mu_2).\end{aligned}$$

So we are done. □

Proof of Theorem 2.3. Just as at the start of the proof of Theorem 2.2, $\mu_t \in \mathcal{O}$ for all $t \le T^*$. Let $T_1, T_2 \ge 0$ be as in Definition 2.10, and let R_1, R_2 be as in Definition 2.11. Again due to the continuity of the mappings $\Phi_{s,t}$, it is sufficient to fix $t \le \min(T^*, -T_2)$ and show that $\frac{1}{T}\int_{t-T}^{t} \Phi_{s,t*}\lambda\, ds \to \mu_t$ as $T \to \infty$. Fix $\varepsilon > 0$. Since $d_W(\mu_s, \mu)$ is $o(e^{rs})$ as $s \to -\infty$, we can take a value $\tau < \min(t, -R_2(\frac{1}{4}\varepsilon e^{r|t|}))$ sufficiently large in magnitude that $d_W(\mu_s, \mu) < \frac{1}{4}\varepsilon e^{r(s-t)}$ for all $s \le \tau$.

Now, take any T sufficiently large that

- $\tau - (t-T) > \max(T_1, R_1(\frac{1}{4}\varepsilon e^{r(\tau-t)}))$, and
- $\frac{t-\tau}{T} < \min(\frac{\varepsilon}{4\mathrm{diam}(X)}, 1)$;

we will show that

$$d_W\left(\frac{1}{T}\int_{t-T}^{t} \Phi_{s,t*}\lambda\, ds, \mu_t\right) < \varepsilon.$$

Let

$$\xi = \frac{1}{T}\int_{t-T}^{t} \Phi_{s,t*}\lambda\, ds$$

$$\tilde{\xi} = \frac{1}{\tau-(t-T)}\int_{t-T}^{\tau} \Phi_{s,\tau*}\lambda\, ds$$

$$\xi_1 = \Phi_{\tau,t*}\tilde{\xi} = \frac{1}{\tau-(t-T)}\int_{t-T}^{\tau} \Phi_{s,t*}\lambda\, ds$$

$$\xi_2 = \frac{1}{t-\tau}\int_{\tau}^{t} \Phi_{s,t*}\lambda\, ds.$$

So, we want $d_W(\xi, \mu_t) < \varepsilon$. Note that

$$\xi = \frac{1}{T}\left(\int_{t-T}^{\tau} \Phi_{s,t*}\lambda\, ds + \int_{\tau}^{t} \Phi_{s,t*}\lambda\, ds\right) = \frac{\tau-(t-T)}{T}\xi_1 + \frac{t-\tau}{T}\xi_2$$

and so Lemma 3.3 gives that

$$d_W(\xi, \xi_1) \le \frac{t-\tau}{T}\mathrm{diam}(X) < \frac{\varepsilon}{4}.$$

So, it remains to show that $d_W(\xi_1, \mu_t) < \frac{3\varepsilon}{4}$. For each v in the interval $[\tau, t]$, we have

$$\Phi_{\tau,v*}\tilde{\xi} = \frac{1}{\tau-(t-T)}\int_{t-T}^{\tau} \Phi_{s,v*}\lambda ds,$$

and so, since $v \le t \le -T_2$ and $\tau-(t-T) > T_1$, it follows that $\Phi_{\tau,v*}\tilde{\xi} \in \mathcal{O}$. Therefore, by Eq. (3.9) in Lemma 3.1(A), in order to show that $d_W(\xi_1, \mu_t) < \frac{3}{4}\varepsilon$ it is sufficient to show that $d_W(\tilde{\xi}, \mu_\tau) < \frac{3}{4}\varepsilon e^{r(\tau-t)}$. But we know that $d_W(\mu_\tau, \mu) < \frac{1}{4}\varepsilon e^{r(\tau-t)}$; and Eq. (2.2) with "$t$" set to τ and "T" set to $\tau-(t-T)$, and with $\varepsilon_1 := \frac{1}{4}\varepsilon e^{r(\tau-t)}$ and $\varepsilon_2 := \frac{1}{4}\varepsilon e^{r|t|}$, gives $d_W(\tilde{\xi}, \mu) < \frac{1}{2}\varepsilon e^{r(\tau-t)}$. So $d_W(\tilde{\xi}, \mu_\tau) < \frac{3}{4}\varepsilon e^{r(\tau-t)}$ as required. □

3.2. *Proofs of the results for differential equations*

Assume the setting presented in Secs. 2.3.1 and 2.3.2. For each $\gamma \in \mathcal{I}$, write $(\Psi_\gamma^t)_{t\ge 0}$ for the semiflow generated by Eq. (2.5), and for every non-empty bounded $O \subset \mathbb{R}^N$ let

$$L(O;\gamma) := \inf\{r \in \mathbb{R} : \forall\, x, y \in O,\ (x-y)\cdot(f_\gamma(x) - f_\gamma(y)) \le r|x-y|^2\}.$$

Recall that K_X denotes the set of non-empty closed subsets of X; likewise, for any closed $G \subset X$, let K_G be the set of non-empty closed subsets of G. For any $\gamma \in \mathcal{I}$, we say that $Q \in K_X$ is an *invariant set of Eq.* (2.5) if $\Psi_\gamma^t(Q) = Q$ for all $t \ge 0$. Note that if $\nu \in M_X$ is an invariant measure of Eq. (2.5), then $\mathrm{supp}\,\nu$ is an invariant set of Eq. (2.5).

We start by proving Proposition 2.1. We need the following elementary lemma.

Lemma 3.4. *For any $A \in \mathbb{R}^{N\times N}$, the maximum of the real parts of the eigenvectors of A is bounded above by $\mathfrak{L}(A)$.*

Proof. Define $\Theta : \mathbb{C}^N \to \mathbb{R}$ by

$$\Theta(v) := \mathrm{Re}(v)^{\mathrm{T}}\mathrm{Im}(v).$$

For each $v \in \mathbb{C}^N$, the sets

$$\{\theta \in \mathbb{R} : \Theta(e^{i\theta}v) \ge 0\}$$
$$\{\theta \in \mathbb{R} : \Theta(e^{i\theta}v) \le 0\}$$

must both have non-empty interior, since taking $\theta = \frac{\pi}{2}$ gives

$$\Theta(iv) = (-\mathrm{Im}(v))^{\mathrm{T}}\mathrm{Re}(v) = -\Theta(v).$$

Furthermore, if $v \neq \mathbf{0}$ then there are at most two values $\theta \in [0, 2\pi)$ for which $e^{i\theta}v$ is purely imaginary. Hence, for $v \neq \mathbf{0}$, the sets

$$\{\theta \in \mathbb{R} : \Theta(e^{i\theta}v) \geq 0, \mathrm{Re}(v) \neq \mathbf{0}\}$$

$$\{\theta \in \mathbb{R} : \Theta(e^{i\theta}v) \leq 0, \mathrm{Re}(v) \neq \mathbf{0}\}$$

are both non-empty. Therefore, if we fix any eigenvalue $\lambda = \lambda_{\mathrm{Re}} + i\lambda_{\mathrm{Im}}$ of A, we can find a corresponding eigenvector $v = v_{\mathrm{Re}} + iv_{\mathrm{Im}}$ such that $v_{\mathrm{Re}} \in S_N$ and $\lambda_{\mathrm{Im}} v_{\mathrm{Re}}^{\mathrm{T}} v_{\mathrm{Im}} \leq 0$. We have

$$\begin{aligned}
\mathfrak{L}(A) &\geq v_{\mathrm{Re}}^{\mathrm{T}} A v_{\mathrm{Re}} \\
&= v_{\mathrm{Re}}^{\mathrm{T}} \mathrm{Re}\big(Av_{\mathrm{Re}} + iAv_{\mathrm{Im}}\big) \\
&= v_{\mathrm{Re}}^{\mathrm{T}} \mathrm{Re}(Av) \\
&= v_{\mathrm{Re}}^{\mathrm{T}} \mathrm{Re}(\lambda v) \\
&= v_{\mathrm{Re}}^{\mathrm{T}} (\lambda_{\mathrm{Re}} v_{\mathrm{Re}} - \lambda_{\mathrm{Im}} v_{\mathrm{Im}}) \\
&= \lambda_{\mathrm{Re}} - \lambda_{\mathrm{Im}} v_{\mathrm{Re}}^{\mathrm{T}} v_{\mathrm{Im}} \\
&\geq \lambda_{\mathrm{Re}}.
\end{aligned}$$

□

Proof of Proposition 2.1. For each $x, y \in P$, writing $x_t = \Psi^t(x)$, $y_t = \Psi^t(y)$ and $d_t = x_t - y_t$ for all $t > 0$, we have

$$\lim_{h \to 0} \frac{d_{t+h} - d_t}{h} = f_{\gamma_0}(x_t) - f_{\gamma_0}(y_t)$$

and so

$$\begin{aligned}
\lim_{h \to 0} \frac{|d_{t+h}| - |d_t|}{h} &= \begin{cases} \dfrac{d_t}{|d_t|} \cdot (f_{\gamma_0}(x_t) - f_{\gamma_0}(y_t)) & x_t \neq y_t \\ 0 & x_t = y_t \end{cases} \\
&\leq \begin{cases} L(P)|d_t| & x_t \neq y_t \\ 0 & x_t = y_t \end{cases} \\
&\leq r_0 |d_t|.
\end{aligned}$$

Hence, Gronwall's Lemma (the basic version for linear differential inequalities) gives that $|d_t| \leq e^{r_0 t}|x - y|$ for all $t \geq 0$. Thus

$$\mathrm{diam}(P) = \mathrm{diam}(\Psi^t(P)) \leq e^{r_0 t}\mathrm{diam}(P)$$

for all $t \geq 0$. But since $r_0 < 0$, it follows that $\mathrm{diam}(P) = 0$, i.e. P is a singleton $\{p\}$. Again since $r_0 < 0$, Lemma 3.4 gives that all the eigenvectors of $(Jf_{\gamma_0})(p)$ have negative real part, and so p is a hyperbolic stable fixed point of f_{γ_0}. Furthermore, it is well known that hyperbolic stable fixed points are robust in the following sense: there is a neighborhood U of p and a C^1-neighborhood $\mathcal{V}$ of f_{γ_0} such that every $f \in \mathcal{V}$ has a hyperbolic stable equilibrium $q_f \in U$ with the property that every $x_0 \in U$ is the initial condition of an infinite-time solution $x\colon[0,\infty) \to \mathbb{R}^N$ of $\dot{x} = f(x)$ with $x(t) \to q_f$ as $t \to \infty$. Since $\tilde{Q}(\gamma)$ depends d_H-continuously on γ and f_γ depends continuously in the C^1 topology on γ, every γ sufficiently close to γ_0 has $\tilde{Q}(\gamma) \subset U$ and $f_\gamma \in \mathcal{V}$; but then, since every initial condition in U is attracted to the fixed point q_{f_γ} under f_γ, the only invariant probability measure of f_γ supported on U is the Dirac mass at q_{f_γ}, and so $\tilde{Q}(\gamma) = \{q_{f_\gamma}\}$. □

It remains to prove Theorem 2.4, according to the strategy laid out at the end of Sec. 2.3.3. We start with the construction of the set O.

Lemma 3.5. *For each $r > r_0$, there exists a neighborhood O of P and a value $\tilde{T}^* \le 0$ such that for all $t \in [-\infty, \tilde{T}^*]$, $L(O;\Gamma(t)) \le r$.*

Proof. For each $\gamma \in \mathcal{I}$ and each $x, y \in \mathbb{R}^N$ with $x \neq y$, define

$$\mathfrak{f}_\gamma(x,y) = \frac{(x-y)\cdot(f_\gamma(x) - f_\gamma(y))}{|x-y|^2}.$$

It is clear that the map $(\gamma, x, y) \mapsto \mathfrak{f}_\gamma(x,y)$ is continuous on $\mathcal{I} \times \{(x,y) : x \neq y\}$.

Now since f_γ has continuous dependence on γ in the C^1-topology, the map $(\gamma, x) \mapsto (Jf_\gamma)(x)$ is continuous, and hence the map $(v, \gamma, x) \mapsto v^{\mathrm{T}}(Jf_\gamma)(x)v$ is continuous; and since S_N is compact, it follows that the map $(\gamma, x) \mapsto \mathfrak{L}((Jf_\gamma)(x))$ is continuous. So, since $\{\gamma_0\} \times P$ is compact, let $\delta > 0$ (taken sufficiently small that $B_\delta(P)$ is in the interior of X relative to $\mathbb{R}^N$) be sufficiently small that for all $\gamma \in \mathcal{I}$, $x \in P$ and $y \in X$, if $\max(\gamma - \gamma_0, |y-x|) \le \delta$ then

$$|\mathfrak{L}((Jf_\gamma)(y)) - \mathfrak{L}((Jf_{\gamma_0})(x))| \le r - r_0.$$

Since $\mathfrak{L}((Jf_{\gamma_0})(x)) \le r_0$, it then follows that $\mathfrak{L}((Jf_\gamma)(y)) \le r$. So then,

$$\sup_{\gamma \in [\gamma_0, \gamma_0+\delta]} \sup_{x \in P} L(B_\delta(x);\gamma) \le r. \tag{3.14}$$

Now let $\{x_1, \ldots, x_n\} \subset P$ be a finite set such that $P \subset B_\delta(\{x_1, \ldots, x_n\})$, and let

$$D_0 = (P \times P) \setminus \bigcup_{i=1}^{n} (B_\delta(x_i) \times B_\delta(x_i))$$

$$\mathfrak{r} = \sup_{(x,y) \in D_0} \mathfrak{f}_{\gamma_0}(x,y).$$

Note that $\mathfrak{r} \le r_0$. For each $\varepsilon \in (0, \delta)$, let

$$D(\varepsilon) = \left(\overline{B_\varepsilon(P)} \times \overline{B_\varepsilon(P)}\right) \setminus \bigcup_{i=1}^{n} (B_\delta(x_i) \times B_\delta(x_i))$$

$$\mathfrak{r}_{\gamma,\varepsilon} = \sup_{(x,y) \in D(\varepsilon)} \mathfrak{f}_\gamma(x,y) \quad (\gamma \in \mathcal{I}).$$

Since $D(\varepsilon)$ decreases to the compact set D_0 as $\varepsilon \to 0$ and the map $(\gamma, x, y) \mapsto \mathfrak{f}_\gamma(x, y)$ is continuous, we can find $\delta' \in (0, \delta)$ such that for all $\gamma \in [\gamma_0, \gamma_0 + \delta']$, $\mathfrak{r}_{\gamma,\delta'} - \mathfrak{r} \le r - r_0$. So since $\mathfrak{r} \le r_0$,

$$\sup_{\gamma \in [\gamma_0, \gamma_0+\delta']} \mathfrak{r}_{\gamma,\delta'} \le r. \tag{3.15}$$

Now, let $O = \overline{B_{\delta'}(P)}$. Since

$$P \times P \subset D(\delta') \cup \bigcup_{x \in P} (B_\delta(x) \times B_\delta(x)),$$

Equations (3.14) and (3.15) imply that every $\gamma \in [\gamma_0, \gamma_0 + \delta']$ has $L(O; \gamma) \le r$. So take $\tilde{T}^*$ such that every $t \in [-\infty, \tilde{T}^*]$ has $\Gamma(t) \in [\gamma_0, \gamma_0 + \delta']$. □

To verify condition (i) of Definition 2.13 applied to $(\mathcal{O}, T^*, \nu)$ (where T^* will be constructed later), we have the following lemma.

Lemma 3.6. *For each* $\tau \in (-\infty, 0)$, *if* $\nu \in M_X$ *is an invariant measure of Eq.* (2.5) *with* $\gamma := \Gamma(\tau)$ *then* $\frac{1}{h} d_W(\Phi_{\tau,\tau+h*}\nu, \nu) \to 0$ *as* $h \to 0+$.

Proof. For each small $h > 0$, since $\Psi^h_{\gamma*}\nu = \nu$, the pushforward measure of ν under the map $x \mapsto (\Phi_{\tau,\tau+h}(x), \Psi^t_\gamma(x))$ is an element of $\mathcal{J}(\Phi_{\tau,\tau+h*}\nu, \nu)$, and so we have

$$d_W(\Phi_{\tau,\tau+h*}\nu, \nu) \le \int_X |\Phi_{\tau,\tau+h}(x) - \Psi^h_\gamma(x)| \, \nu(dx).$$

Now for each $x \in X$,

$$\begin{aligned}\lim_{h\to 0+} \frac{\Phi_{\tau,\tau+h}(x) - \Psi^h_\gamma(x)}{h} &= \lim_{h\to 0+} \frac{\Phi_{\tau,\tau+h}(x) - x}{h} - \lim_{h\to 0+} \frac{\Psi^h_\gamma(x) - x}{h} \\ &= f_\gamma(x) - f_\gamma(x) = 0.\end{aligned}$$

We wish to apply the dominated convergence theorem to this; for the integrability condition of the dominated convergence theorem, it is sufficient that $\frac{\Phi_{\tau,\tau+h}(x) - \Psi^h_\gamma(x)}{h}$ is bounded in (x, h). For each $x \in X$ and each $h \in (0, |\tau|]$, we have that

$$\frac{|\Phi_{\tau,\tau+h}(x) - \Psi^h_\gamma(x)|}{h} \le \frac{|\Phi_{\tau,\tau+h}(x) - x|}{h} + \frac{|\Psi^h_\gamma(x) - x|}{h} \le \max_{y \in X, t \in [\tau,0]} 2|f_{\Gamma(t)}(y)|.$$

Hence, the dominated convergence theorem can be applied to give that

$$\lim_{h\to 0+} \frac{1}{h} \int_X |\Phi_{\tau,\tau+h}(x) - \Psi^h_\gamma(x)| \, \nu(dx) = \int_X \lim_{h\to 0+} \frac{|\Phi_{\tau,\tau+h}(x) - \Psi^h_\gamma(x)|}{h} \, \nu(dx) = 0,$$

and so

$$\lim_{h\to 0+} \frac{d_W(\Phi_{\tau,\tau+h*}\nu, \nu)}{h} = 0.$$ □

We next consider how to verify the "growth-control" condition in Definition 2.15.

Lemma 3.7. *Suppose we have* $r > 0$, $t \in (-\infty, 0)$ *and* $\lambda_1, \lambda_2 \in M_X$ *with supports contained in a set* $O \subset X$ *with* $L(O; \Gamma(t)) \leq r$. *Then* $\limsup_{h\to 0+} \frac{1}{h}[d_W(\Phi_{t,t+h*}\lambda_1, \Phi_{t,t+h*}\lambda_2) - d_W(\lambda_1, \lambda_2)] \leq r d_W(\lambda_1, \lambda_2)$.

We will only need the result with lim inf in place of lim sup, but it is no harder to obtain the stronger result about lim sup. The proof is similar in nature to the proof of Lemma 3.6.

Proof of Lemma 3.7. Let $\mathbb{P}$ be a minimizer of $\tilde{\mathbb{P}} \mapsto \mathbb{E}_{\tilde{\mathbb{P}}}[|\pi_1 - \pi_2|]$ on $\mathcal{J}(\lambda_1, \lambda_2)$. For each $h > 0$, the pushforward measure of $\mathbb{P}$ under the map $(x, y) \mapsto (\Phi_{t,t+h}(x), \Phi_{t,t+h}(y))$ is an element of $\mathcal{J}(\Phi_{t,t+h*}\lambda_1, \Phi_{t,t+h*}\lambda_2)$, and so we have

$$d_W(\Phi_{t,t+h*}\lambda_1, \Phi_{t,t+h*}\lambda_2) \leq \int_{X\times X} |\Phi_{t,t+h}(x) - \Phi_{t,t+h}(y)| \, \mathbb{P}(d(x, y)),$$

and hence

$$\begin{aligned} &d_W(\Phi_{t,t+h*}\lambda_1, \Phi_{t,t+h*}\lambda_2) - d_W(\lambda_1, \lambda_2) \\ &\quad\leq \int_{X\times X} |\Phi_{t,t+h}(x) - \Phi_{t,t+h}(y)| - |x - y| \, \mathbb{P}(d(x, y)) \\ &\quad= \int_{O\times O} |\Phi_{t,t+h}(x) - \Phi_{t,t+h}(y)| - |x - y| \, \mathbb{P}(d(x, y)). \end{aligned}$$

Now for each $(x, y) \in O \times O$, similarly to in the proof of Proposition 2.1, we have

$$\lim_{h\to 0+} \frac{|\Phi_{t,t+h}(x) - \Phi_{t,t+h}(y)| - |x - y|}{h} \leq r|x - y|.$$

As in the proof of Lemma 3.6, we wish to apply the dominated convergence theorem. For each $x, y \in X$ and each $h \in (0, |t|]$, we have that

$$\begin{aligned} \left| \frac{|\Phi_{t,t+h}(x) - \Phi_{t,t+h}(y)| - |x - y|}{h} \right| &\leq \frac{|\Phi_{t,t+h}(x) - x|}{h} + \frac{|\Phi_{t,t+h}(y) - y|}{h} \\ &\leq \max_{z\in X, \tau\in[t,0]} 2|f_{\Gamma(\tau)}(z)|. \end{aligned}$$

Hence the dominated convergence theorem can be applied to give that

$$\begin{aligned} &\lim_{h\to 0+} \frac{1}{h} \int_{O\times O} |\Phi_{t,t+h}(x) - \Phi_{t,t+h}(y)| - |x - y| \, \mathbb{P}(d(x, y)) \\ &\quad= \int_{O\times O} \lim_{h\to 0+} \frac{|\Phi_{t,t+h}(x) - \Phi_{t,t+h}(y)| - |x - y|}{h} \, \mathbb{P}(d(x, y)) \\ &\quad\leq \int_{O\times O} r|x - y| \, \mathbb{P}(d(x, y)) \\ &\quad= \int_{X\times X} r|x - y| \, \mathbb{P}(d(x, y)) \\ &\quad= r d_W(\lambda_1, \lambda_2), \end{aligned}$$

and so

$$\limsup_{h\to 0+} \frac{d_W(\Phi_{t,t+h*}\lambda_1, \Phi_{t,t+h*}\lambda_2) - d_W(\lambda_1,\lambda_2)}{h} \leq r d_W(\lambda_1,\lambda_2). \qquad \square$$

The main remaining step towards being able to prove Theorem 2.4(A) is finding T^* such that $\Phi_{s,t*}\mu \in \mathcal{O}$ for all $s \leq t \leq T^*$, as in condition (iii) of Definition 2.13. It is clearly sufficient to find T^* such that $\Phi_{s,t}(P) \subset O$ for all $s \leq t \leq T^*$.

To do this, we will first need to obtain results for the set-valued dynamics that are analogous to Lemmas 3.6 and 3.7 for the measure-valued dynamics.

Lemma 3.8. *For each $\tau \in (-\infty, 0)$, if $Q \in K_X$ is an invariant set of Eq.* (2.5) *with $\gamma := \Gamma(\tau)$ then $\frac{1}{h} d_H(\Phi_{\tau,\tau+h}(Q), Q) \to 0$ as $h \to 0+$.*

Proof. Fix $\varepsilon > 0$, and let $\delta_1 > 0$ be such that for each $t \in [\tau, \tau+\delta_1]$ and $x \in X$, $|f_{\Gamma(t)}(x) - f_\gamma(x)| < \frac{\varepsilon}{2}$. Let R be a Lipschitz constant for f_γ. The map $t \mapsto \Phi_{\tau,t}(x) - \Psi_\gamma^{t-\tau}(x)$ has derivative $f_{\Gamma(t)}(\Phi_{\tau,t}(x)) - f_\gamma(\Psi_\gamma^{t-\tau}(x))$; and so for each $x \in X$ and $h \in (0, \delta_1)$, we have that

$$\begin{aligned}
&\frac{|\Phi_{\tau,\tau+h}(x) - \Psi_\gamma^h(x)|}{h} \\
&\quad \leq \max_{t\in[\tau,\tau+h]} |f_{\Gamma(t)}(\Phi_{\tau,t}(x)) - f_\gamma(\Psi_\gamma^{t-\tau}(x))| \\
&\quad \leq \left(\max_{t\in[\tau,\tau+h]} |f_{\Gamma(t)}(\Phi_{\tau,t}(x)) - f_\gamma(\Phi_{\tau,t}(x))|\right) \\
&\qquad + \left(\max_{t\in[\tau,\tau+h]} |f_\gamma(\Phi_{\tau,t}(x)) - f_\gamma(\Psi_\gamma^{t-\tau}(x))|\right) \\
&\quad < \frac{\varepsilon}{2} + R\left(\max_{t\in[\tau,\tau+h]} |\Phi_{\tau,t}(x) - \Psi_\gamma^{t-\tau}(x)|\right) \\
&\quad \leq \frac{\varepsilon}{2} + R\left(\max_{t\in[\tau,\tau+h]} |\Phi_{\tau,t}(x) - x| + |\Psi_\gamma^{t-\tau}(x) - x|\right) \\
&\quad \leq \frac{\varepsilon}{2} + h \,.\, \underbrace{\left(2R \max_{y\in X,\, t\in[\tau,\tau+\delta_1]} |f_{\Gamma(t)}(y)|\right)}_{=:L_1}.
\end{aligned}$$

So taking $0 < \delta \leq \min(\delta_1, \frac{\varepsilon}{2L_1})$, for every $h \in (0,\delta)$ and $x \in X$, we have

$$\frac{|\Phi_{\tau,\tau+h}(x) - \Psi_\gamma^h(x)|}{h} < \varepsilon.$$

Since $\Psi_\gamma^h(Q) = Q$ for all $h > 0$, it follows that for $h \in (0,\delta)$, we have

$$\frac{d_H(\Phi_{\tau,\tau+h}(Q), Q)}{h} < \varepsilon. \qquad \square$$

Lemma 3.9. *Suppose we have $r > 0$, $\tau \in (-\infty, 0)$, and sets $A_1, A_2 \in K_X$ contained in the interior of a set $O \subset X$ with $\sup_{t\in[\tau,\tau+\delta]} L(O;\Gamma(t)) \leq r$ for some $\delta > 0$. Then $\limsup_{h\to 0+} \frac{1}{h}[d_H(\Phi_{\tau,\tau+h}(A_1), \Phi_{\tau,\tau+h}(A_2)) - d_H(A_1, A_2)] \leq r d_H(A_1, A_2)$.*

(Again, we will only need the result with lim inf in place of lim sup, but it is no harder to obtain the stronger result about lim sup.)

Proof of Lemma 3.9. Let $\delta_1 \in (0, \delta]$ be such that $\Phi_{\tau,\tau+h}(A_1 \cup A_2) \subset O$ for all $h \in (0, \delta_1]$. For each $x \in A_1$ and $y \in A_2$, writing $x_t = \Phi_{\tau,t}(x)$, $y_t = \Phi_{\tau,t}(y)$ and $d_t = x_t - y_t$, we have

$$\lim_{h\to 0} \frac{d_{t+h} - d_t}{h} = f_{\Gamma(t)}(x_t) - f_{\Gamma(t)}(y_t)$$

for all $t \in (\tau, \tau + \delta_1)$. Hence, as in the proof of Proposition 2.1, for all $h \in (0, \delta_1]$, $x \in A_1$ and $y \in A_2$, we have

$$|\Phi_{\tau,\tau+h}(x) - \Phi_{\tau,\tau+h}(y)| \leq e^{rh}|x - y|.$$

For every $x \in A_1$, there exists $y \in A_2$ with $|x - y| \leq d_H(A_1, A_2)$; and taking any such y, we then have that for all $h \in (0, \delta_1]$,

$$|\Phi_{\tau,\tau+h}(x) - \Phi_{\tau,\tau+h}(y)| \leq e^{rh} d_H(A_1, A_2).$$

So then, for all $h \in (0, \delta_1]$,

$$\max_{x\in A_1} d(\Phi_{\tau,\tau+h}(x), \Phi_{\tau,\tau+h}(A_2)) \leq e^{rh} d_H(A_1, A_2).$$

The same likewise holds with A_1 and A_2 switched round, and so

$$d_H(\Phi_{\tau,\tau+h}(A_1), \Phi_{\tau,\tau+h}(A_2)) \leq e^{rh} d_H(A_1, A_2).$$

In particular,

$$\frac{d_H(\Phi_{\tau,\tau+h}(A_1), \Phi_{\tau,\tau+h}(A_2)) - d_H(A_1, A_2)}{h} \leq \frac{e^{rh} - 1}{h} d_H(A_1, A_2),$$

and taking the superior limit as $h \to 0+$ gives the desired result. □

Lemma 3.10. *Suppose we have $r > 0$, a neighborhood O of P, and a value $\tilde{T}^* \leq 0$ such that for all $t \leq \tilde{T}^*$, $L(O;\Gamma(t)) \leq r$. Suppose furthermore that $\int_{-\infty}^0 d_H(Q(t), P)e^{r|t|}\, dt < \infty$. Then there exists $T^* \leq 0$ such that for all $s \leq t \leq T^*$, $\Phi_{s,t}(P) \subset O$.*

Proof. Take $\Delta > 0$ such that $\overline{B_\Delta(P)} \subset O^\circ$; hence in particular, every $G \in K_X$ with $d_H(G, P) < \Delta$ is contained in O°. Writing

$$\mathfrak{d}(u) := (C + 1)d_H(Q(u), P) + Cre^{ru} \int_{-\infty}^u d_H(Q(t), P)e^{r|t|},$$

it is clear that $\mathfrak{d}(u) \to 0$ as $u \to -\infty$, so take $T^* \leq \tilde{T}^*$ sufficiently large in magnitude that for all $u \leq T^*$,

$$\mathfrak{d}(u) < \Delta. \tag{3.16}$$

It obviously follows in particular that $d_H(Q(u), P) < \Delta$ and hence $Q(u) \subset B_\Delta(P)$. Now for each $s \in (-\infty, \tilde{T}^*)$, since Γ is continuous let $\delta(s) > 0$ be such that for all $t \in [s, s + \delta(s))$, $\Gamma(t) - \Gamma(s) < \tilde{\delta}(\Gamma(s))$. Defining the nonautonomous dynamical system $(\Phi^{\mathrm{set}}_{s,t})$ on K_X by $\Phi^{\mathrm{set}}_{s,t}(G) := \Phi_{s,t}(G)$, Lemmas 3.8 and 3.9 imply that $(\Phi^{\mathrm{set}}_{s,t})$ fulfills the conditions of Lemma 3.1(B) with P in place of p, $K_{\overline{B_\Delta(P)}}$ in place of O, $Q(\cdot)$ in place of $q(\cdot)$ and d_H in place of d. Hence, Lemma 3.1(B) gives that for any $s \leq u \leq \tilde{T}^*$, if $\Phi_{s,t}(P) \subset \overline{B_\Delta(P)}$ for all $t \in [s, u)$ then

$$d_H(\Phi_{s,u}(P), P) \leq \mathfrak{d}(u). \tag{3.17}$$

Since the map $u \mapsto \Phi_{s,u}(P)$ is continuous for each s, it follows that for every $s \leq u \leq T^*$, $\Phi_{s,u}(P) \subset B_\Delta(P) \subset O$: for, otherwise, if we let $u' := \min\{t \in (s, u] : \Phi_{s,t}(P) \not\subset B_\Delta(P)\}$, then $\Phi_{s,t}(P) \subset \overline{B_\Delta(P)}$ for all $t \in [s, u']$ and so Eqs. (3.16) and (3.17) give that $d_H(\Phi_{s,u'}(P), P) < \Delta$, contradicting that $\Phi_{s,u'}(P) \not\subset B_\Delta(P)$. □

Proof of Theorem 2.4(A). Take O and $\tilde{T}^*$ as in Lemma 3.5, with O closed; and then take T^* as in Lemma 3.10, sufficiently large in magnitude that $T^* \leq \tilde{T}^*$ and $Q(t) \subset O^\circ$ for all $t \leq T^*$. It follows that $\nu(t) \in \mathcal{O}$ for all $t \in [-\infty, T^*]$. The fact that $(\mathcal{O}, T^*)$ is a GC of $(M_X, d_W, (\Phi_{s,t*}))$ of rate r is given by Lemma 3.7. We now verify the conditions of Definition 2.13 for $(\mathcal{O}, T^*, \nu)$. First note that, by assumption, ν is continuous; condition (i) is given by Lemma 3.6; condition (ii) is assumed with 0 in place of T^*, but T^* is already less than or equal to 0; condition (iii) follows immediately from the definition of T^* as given by Lemma 3.10. So $(\mathcal{O}, T^*, \nu)$ is a NAC of $(M_X, d_W, (\Phi_{s,t*}))$ near μ of rate r. Furthermore, it is monotone-like, since one can take $\delta(s)$ as in the proof of Lemma 3.10. □

To prove part (B) of Theorem 2.4, we need the following lemma.

Lemma 3.11. *Suppose we have $r > 0$ and a neighborhood $O \subset X$ of P such that $L(O) \leq r$. Then for any $x \in O$ and $s \leq u \leq 0$, if every $t \in [s, u)$ has $\Phi_{s,t}(x) \in O$ and $\Psi^{t-s}(x) \in O$, then*

$$|\Phi_{s,u}(x) - \Psi^{u-s}(x)| \leq e^{ru} \int_{-\infty}^{u} d_{X,C^0}(f_{\Gamma(t)}, f_{\gamma_0}) e^{r|t|} dt. \tag{3.18}$$

Proof. Since $d_{X,C^0}(f_{\Gamma(t)}, f_{\gamma_0}) e^{r|t|} \geq 0$, it is sufficient to show Eq. (3.18) with $\int_s^u$ in place of $\int_{-\infty}^u$. For each $t \in (s, u)$, we have

$$\lim_{h \to 0} \frac{\Phi_{s,t+h}(x) - \Psi^h(\Phi_{s,t}(x))}{h} = f_{\Gamma(t)}(\Phi_{s,t}(x)) - f_{\gamma_0}(\Phi_{s,t}(x)),$$

and similarly to in the proof of Proposition 2.1, we have

$$\left|\lim_{h\to 0}\frac{\Psi^h(\Phi_{s,t}(x))-\Psi^{h+t-s}(x)}{h}\right|\le r|\Phi_{s,t}(x)-\Psi^{t-s}(x)|.$$

Combining these, we have

$$\begin{aligned}&\left|\lim_{h\to 0}\frac{\Phi_{s,t+h}(x)-\Psi^{h+t-s}(x)}{h}\right|\\ &\quad\le\left|\lim_{h\to 0}\frac{\Phi_{s,t+h}(x)-\Psi^h(\Phi_{s,t}(x))}{h}\right|+\left|\lim_{h\to 0}\frac{\Psi^h(\Phi_{s,t}(x))-\Psi^{h+t-s}(x)}{h}\right|\\ &\quad\le|f_{\Gamma(t)}(\Phi_{s,t}(x))-f_{\gamma_0}(\Phi_{s,t}(x))|+r|\Phi_{s,t}(x)-\Psi^{t-s}(x)|\\ &\quad\le d_{X,C^0}(f_{\Gamma(t)},f_{\gamma_0})+r|\Phi_{s,t}(x)-\Psi^{t-s}(x)|.\end{aligned}$$

Hence, for all $t\in[s,u]$,

$$|\Phi_{s,t}(x)-\Psi^{t-s}(x)|\le\int_s^t d_{X,C^0}(f_{\Gamma(\tau)},f_{\gamma_0})+r|\Phi_{s,\tau}(x)-\Psi^{\tau-s}(x)|\,d\tau.$$

A suitable version of Gronwall's Lemma [8, Corollary 2] then yields the result. □

Corollary 3.1. *Suppose we have $r>0$, a neighborhood $U_0\subset X$ of P, and a neighborhood $O\subset X$ of $\bar{U}_0$, such that $L(O)\le r$. Suppose furthermore that $\int_{-\infty}^0 d_{X,C^0}(f_{\Gamma(t)},f_{\gamma_0})e^{r|t|}\,dt<\infty$. Then there exists $\tilde{T}\le 0$ such that for any $x\in U_0$ and $s\le u\le\tilde{T}$, if every $t\in[s,u)$ has $\Psi^{t-s}(x)\in U_0$ then $\Phi_{s,u}(x)\in O$.*

Proof. Take $\Delta>0$ such that $\overline{B_\Delta(U_0)}\subset O^\circ$, and let $\tilde{T}$ be such that for all $u\le\tilde{T}$,

$$e^{ru}\int_{-\infty}^u d_{X,C^0}(f_{\Gamma(t)},f_{\gamma_0})e^{r|t|}\,dt<\Delta. \tag{3.19}$$

Now, similarly to in the proof of Lemma 3.10, suppose for a contradiction that we have $x\in U_0$ and $s\le u\le\tilde{T}$ such that every $t\in[s,u)$ has $\Psi^{t-s}(x)\in U_0$ but $\Phi_{s,u}(x)\notin B_\Delta(U_0)$. Let $u':=\min\{t\in(s,u]:\Phi_{s,t}(x)\notin B_\Delta(U_0)\}$. The map $t\mapsto\Phi_{s,t}(x)$ is continuous, and so every $t\in[s,u']$ has $\Phi_{s,t}(x)\in O$ (as well as, by assumption, $\Psi^{t-s}(x)\in U_0\subset O$), and so Lemma 3.11 together with Eq. (3.19) gives that

$$|\Phi_{s,u'}(x)-\Psi^{u'-s}(x)|<\Delta.$$

But $\Psi^{u'-s}(x)\in U_0$, and hence $\Phi_{s,u'}(x)\in B_\Delta(U_0)$, contradicting the definition of u'. □

Proof of Theorem 2.4(B). Let O and $\mathcal{O}$ be as in the proof of Theorem 2.4(A). Fix a neighborhood U_0 of P such that $\overline{B_\Delta(U_0)}\subset O^\circ$, and since $\int_{-\infty}^0 d_{X,C^0}(f_{\Gamma(t)},f_{\gamma_0})e^{r'|t|}\,dt<\infty$, let $\tilde{T}$ be as given by Corollary 3.1. Since P is Lyapunov-stable under the past-limit system, let $U\subset U_0$ be a neighborhood of P such that for all $t\ge 0$, $\Psi^t(U)\subset U_0$. Fix any $\lambda\in M_X$ with $\lambda(U)=1$. We first show that λ is pullback-eventually in $\mathcal{O}$ under $(\Phi_{s,t*})$. We verify Definition 2.7 with $T_1 = 0$ and $T_2 = \tilde{T}$: every $x \in U$ has $\Psi^t(x) \subset U_0$ for all $t \geq 0$, and therefore by definition of $\tilde{T}$, $\Phi_{s,u}(x) \in O$ for all $s \leq u \leq \tilde{T}$; hence in particular, $\Phi_{s,u*}\lambda \in \mathcal{O}$ for all $s \leq u \leq \tilde{T}$. Now, suppose furthermore that λ is, respectively, attracted or Cesàro-attracted to μ under (Ψ_*^t). For each $\varepsilon > 0$, let $R_1(\varepsilon) \geq 0$ be such that, respectively, $d_W(\Psi_*^t\lambda, \mu) < \varepsilon$ for all $t \geq R_1(\varepsilon)$, or $d_W(\frac{1}{T}\int_0^T \Psi_*^\tau\lambda\, d\tau, \mu) < \varepsilon$ for all $T > R_1(\varepsilon)$. For each $\varepsilon > 0$, let $R_2(\varepsilon) \geq |\tilde{T}|$ be such that

$$\int_{-\infty}^{-R_2(\varepsilon)} d_{X,C^0}(f_{\Gamma(t)}, f_{\gamma_0})e^{r|t|}\, dt < \varepsilon.$$

Now fix $\varepsilon_1, \varepsilon_2 > 0$. First, consider the case that λ is attracted to μ under (Ψ_*^t). Fix $t \leq -R_2(\varepsilon_2)$ and $s \leq t - R_1(\varepsilon_1)$; we need to show that

$$d_W(\Phi_{s,t*}\lambda, \mu) < \varepsilon_1 + \varepsilon_2 e^{rt}.$$

By the triangle inequality,

$$d_W(\Phi_{s,t*}\lambda, \mu) \leq d_W(\Psi_*^{t-s}\lambda, \mu) + d_W(\Phi_{s,t*}\lambda, \Psi_*^{t-s}\lambda).$$

Since $t - s \geq R_1(\varepsilon_1)$, $d_W(\Psi_*^{t-s}\lambda, \mu) < \varepsilon_1$. So, it is sufficient to show that $d_W(\Phi_{s,t*}\lambda, \Psi_*^{t-s}\lambda) < \varepsilon_2 e^{rt}$. The pushforward measure of λ under the map $x \mapsto (\Phi_{s,t}(x), \Psi^{t-s}(x))$ is an element of $\mathcal{J}(\Phi_{s,t*}\lambda, \Psi_*^{t-s}\lambda)$, and so

$$\begin{aligned} d_W(\Phi_{s,t*}\lambda, \Psi_*^{t-s}\lambda) &\leq \int_X |\Phi_{s,t}(x) - \Psi^{t-s}(x)|\, \lambda(dx) \\ &= \int_U |\Phi_{s,t}(x) - \Psi^{t-s}(x)|\, \lambda(dx). \end{aligned}$$

For all $x \in U$ and $\tau \in [s,t]$, we have (by definition of U) that $\Psi^{\tau-t}(x) \in U_0 \subset O$, and since $\tau \leq \tilde{T}$, we have that $\Phi_{s,\tau}(x) \in O$. Hence, by Lemma 3.11, every $x \in U$ has

$$|\Phi_{s,t}(x) - \Psi^{t-s}(x)| \leq e^{rt} \int_{-\infty}^t d_{X,C^0}(f_{\Gamma(\tau)}, f_{\gamma_0})e^{r|\tau|}\, d\tau < \varepsilon_2 e^{rt}.$$

So $d_W(\Phi_{s,t*}\lambda, \Psi_*^{t-s}\lambda) < \varepsilon_2 e^{rt}$ as required. Now consider the case that λ is Cesàro-attracted to μ under (Ψ_*^t). Fix $t \leq -R_2(\varepsilon_2)$ and $T > R_1(\varepsilon_1)$; we need to show Eq. (2.2). By the triangle inequality

$$\begin{aligned} d_W\left(\frac{1}{T}\int_{t-T}^t \Phi_{s,t*}\lambda\, ds, \mu\right) &\leq d_W\left(\frac{1}{T}\int_{t-T}^t \Psi_*^{t-s}\lambda\, ds, \mu\right) \\ &\quad + d_W\left(\frac{1}{T}\int_{t-T}^t \Phi_{s,t*}\lambda\, ds, \frac{1}{T}\int_{t-T}^t \Psi_*^{t-s}\lambda\, ds\right). \end{aligned}$$

Again, we wish to show that the two terms on the right-hand side are bounded above by ε_1 and $\varepsilon_2 e^{rt}$, respectively. Since $T > R_1(\varepsilon_1)$,

$$d_W\left(\frac{1}{T}\int_{t-T}^{t} \Psi_*^{t-s}\lambda\, ds, \mu\right) = d_W\left(\frac{1}{T}\int_0^T \Psi_*^{\tau}\lambda\, d\tau, \mu\right) < \varepsilon_1.$$

Letting $\mathfrak{u}_{t-T,t}$ be the normalized Lebesgue measure on $[t-T, t]$, the pushforward measure of $\lambda \otimes \mathfrak{u}_{t-T,t}$ under the map $(x, s) \mapsto (\Phi_{s,t}(x), \Psi^{t-s}(x))$ is an element of

$$\mathcal{J}\left(\frac{1}{T}\int_{t-T}^{t} \Phi_{s,t*}\lambda\, ds, \frac{1}{T}\int_{t-T}^{t} \Psi_*^{t-s}\lambda\, ds\right),$$

and so

$$d_W\left(\frac{1}{T}\int_{t-T}^{t} \Phi_{s,t*}\lambda\, ds, \frac{1}{T}\int_{t-T}^{t} \Psi_*^{t-s}\lambda\, ds\right) \leq \frac{1}{T}\int_{t-T}^{T}\int_U |\Phi_{s,t}(x) - \Psi^{t-s}(x)|\,\lambda(dx)\, ds.$$

By the same argument as before, every $x \in U$ and $s \leq t$ has $|\Phi_{s,t}(x) - \Psi^{t-s}(x)| < \varepsilon_2 e^{rt}$, and so

$$d_W\left(\frac{1}{T}\int_{t-T}^{t} \Phi_{s,t*}\lambda\, ds, \frac{1}{T}\int_{t-T}^{t} \Psi_*^{t-s}\lambda\, ds\right) < \varepsilon_2 e^{rt}.$$ □

4. Final Remarks

We have been able to demonstrate existence of nonautonomous physical measures in the context of a class of nonautonomous continuous-time systems with asymptotically autonomous past limits that have physical measures in the usual sense; still, there remain a number of challenges.

Firstly, we expect there are connections between the "monotone-like" hypotheses used to obtain Theorems 2.1 and 2.4 and more general assumptions on linear response; it would be helpful to clarify whether this is the case and/or whether the existence of a nonautonomous physical measure can be usefully proven under weaker assumptions. In particular, we anticipate that there may be ways in which more standard linear response assumptions can give rise to the conclusions of Theorem 2.4 in place of having to verify a "monotone-like" assumption. (As in Remark 2.9 and Appendix A, we have already seen that it is possible to obtain an alternative condition to the "monotone-like" hypothesis in order to obtain the conclusions of Theorem 2.1; but, as detailed in Appendix A, this alternative condition will generally not be applicable to obtaining the existence of nonautonomous physical measures other than time-dependent Dirac measures.)

Secondly, Theorem 2.4 relies on the existence of an autonomous past-limiting system: the natural next step is to consider other important cases, such as nonautonomous random dynamical systems with an autonomous random past limit. However, we do not expect it to be possible to formulate a useful notion of nonautonomous physical measures without making at least some assumption on past-limiting behavior.

Acknowledgments

This is TiPES contribution #149. This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 820970 (TiPES). The authors are grateful to Ian Melbourne for useful and interesting discussions, and to Iosif Pinelis for help (via MathOverflow) in obtaining an efficient proof of Lemma 3.2.

Appendix A. Alternative Conditions for the Conclusion of Theorems 2.1 and 2.2

Definition A.1. Given $r > 0$ and a nonautonomous dynamical system $(\Phi_{s,t})$ on X, an *equilibrated growth-controller* (*EGC*) *of* $(X, d, (\Phi_{s,t}))$ *of rate* r is a triple (O, T^*, q) consisting of a GC (O, T^*) of $(X, d, (\Phi_{s,t}))$ of rate r and a continuous function $q\colon(-\infty, T^*] \to O$ such that for all $t < T^*$ and $x \in O$, $\liminf_{h\to 0+} \frac{1}{h} d(\Phi_{t,t+h}(x), x) \le r d(x, q(t))$.

Heuristically, Definition A.1 says that in O, not only is the *speed of separation* of two trajectories controlled by their distance from each other, but also the *absolute speed* of each trajectory is controlled by its distance from the "equilibrium $q(t)$ of the time-t instantaneous dynamics". It is reasonable to expect this behavior in finite-dimensional settings: e.g., if X is a compact subset of $\mathbb{R}^N$ and $(\Phi_{s,t})$ is the solution flow of a smooth nonautonomous differential equation $\dot{x}(t) = f_t(x(t))$, then the condition $\liminf_{h\to 0+} \frac{1}{h} d(\Phi_{t,t+h}(x), x) \le r d(x, q(t))$ is equivalent to the condition that $|f_t(x)| \le r|x - q(t)|$. However, in infinite-dimensional settings, one cannot so readily expect this kind of "continuously differentiable" behavior, even for autonomous systems. We now give an example of this, for dynamical systems on the space M_X of probability measures on a compact space X.

Proposition A.1. *Let f be a locally Lipschitz vector field on $\mathbb{R}^N$, and let X be a compact subset of $\mathbb{R}^N$ such that every $x_0 \in X$ is the initial condition of an infinite-time solution $(x(t))_{t\ge 0}$ of $\dot{x}(t) = f(x(t))$ with $x(t) \in X$ for all $t \ge 0$. Let $(\Psi^t)_{t\ge 0}$ be the semiflow on X generated by f. Let $\mu \in M_X$ be an invariant measure of $(\Psi^t)_{t\ge 0}$, i.e. a fixed point of the autonomous dynamical system $(\Psi^t_*)_{t\ge 0}$ on M_X. Then for any sequence $\mu_n \in M_X$ of finitely supported probability measures converging weakly to μ, we have that*

$$\lim_{h\to 0+} \frac{d_W(\Psi^h_* \mu_n, \mu_n)}{h} = \int_X |f| \, d\mu_n \tag{A.1}$$

for each n, and hence

$$\lim_{n\to\infty} \lim_{h\to 0+} \frac{d_W(\Psi^h_* \mu_n, \mu_n)}{h} = \int_X |f| \, d\mu. \tag{A.2}$$

In particular, if $\int_X |f| \, d\mu \ne 0$ then

$$\frac{\lim_{h\to 0+} \frac{1}{h} d_W(\Psi^h_* \mu_n, \mu_n)}{d_W(\mu_n, \mu)} \to \infty \qquad \text{as } n \to \infty. \tag{A.3}$$

Proof. Fixing n, let $P_n = \operatorname{supp}\mu_n$, let $\varepsilon_n > 0$ be such that the family of ε_n-balls $(B_{\varepsilon_n}(x))_{x\in P_n}$ is mutually disjoint, and let $\delta_n > 0$ be such that for all $x \in P_n$, $\{\Psi^h(x)\}_{0\le h\le\delta_n} \subset B_{\varepsilon_n}(x)$. Then for each $h \in (0,\delta_n]$, for each $\mathbb{P} \in \mathcal{J}(\Psi^h_*\mu_n, \mu_n)$, $\mathbb{P}$-almost every $(x,y) \in X \times X$ has either

- $x = \Psi^h(y)$ and $|x-y| < \varepsilon_n$, or
- $x \neq \Psi^h(y)$ and $|x-y| > \varepsilon_n$.

Hence the minimizer of the map $\mathbb{P} \mapsto \mathbb{E}_{\mathbb{P}}[|\pi_1 - \pi_2|]$ over $\mathcal{J}(\Psi^h_*\mu_n, \mu_n)$ is precisely the pushforward measure of μ_n under the map $x \mapsto (\Psi^h(x), x)$, and

$$d_W(\Psi^h_*\mu_n, \mu_n) = \int_X |\Psi^h(x) - x|\, d\mu_n.$$

Dividing by h and taking the limit as $h \to 0+$ then gives Eq. (A.1). Since $|f|$ is continuous, Eq. (A.2) then follows by definition from the fact that μ_n converges weakly to μ. If $\int_X |f|\, d\mu \neq 0$, then for all sufficiently large n, $\int_X |f|\, d\mu_n \neq 0$, i.e. $\lim_{h\to 0+} \frac{1}{h} d_W(\Psi^h_*\mu_n, \mu_n) \neq 0$, but $d_W(\Psi^h_*\mu, \mu) = d_W(\mu,\mu) = 0$ for all h, and so $\mu_n \neq \mu$. Thus, the ratio on the left-hand side of Eq. (A.3) is well-defined, with a numerator that converges to a positive value while the denominator converges to 0, and so the ratio tends to ∞. □

Accordingly, the following result may be a useful alternative to Theorem 2.1 for finite-dimensional dynamical systems, but the analogous corollary to Corollary 2.1 for measure-valued dynamics is generically inapplicable (except trivially when just considering Dirac masses).

Theorem A.1. *Suppose we have $p \in X$ and a nonautonomous dynamical system $(\Phi_{s,t})$ on X. Suppose we have $r > 0$ and a triple (O, T^*, q) that is both a NAC of $(X, d, (\Phi_{s,t}))$ near p of rate r and an EGC of $(X, d, (\Phi_{s,t}))$ of rate r. Then the conclusions of Theorem 2.1 hold, namely there exists an orbit (p_t) of $(\Phi_{s,t})$ such that p is pullback-attracted to (p_t) under $(\Phi_{s,t})$, and $d(p_t, p)$ is $o(e^{rt})$ as $t \to -\infty$. Additionally, in this setting, any $x \in X$ fulfilling the hypotheses of Theorem 2.2 fulfills the conclusion, namely that x is pullback-attracted to (p_t) under $(\Phi_{s,t})$.*

As in Remark 2.9, the key difference between Theorems A.1 and 2.1 is that we have removed the "monotone-like" condition, and instead strengthened the requirement that (O, T^*) is a GC to requiring that (O, T^*, q) is an EGC.

Proof of Theorem A.1

We have the following analogous lemma to Lemma 3.1(B).

Lemma A.1. *Suppose we have a nonautonomous dynamical system $(\Phi_{s,t})$ on X, a point $p \in X$, a set $O \subset X$, and values $T^* \le 0$ and $r > 0$. Suppose we have a continuous function $q\colon (-\infty, T^*] \to O$ such that for all $t < T^*$ and $x \in O$,*

$\liminf_{h\to 0+}\frac{1}{h}d(\Phi_{t,t+h}(x),x) \le rd(x,q(t))$. *Then for all* $s \le u \le T^*$, *if each* $t \in [s,u)$ *has* $\Phi_{s,t}(p) \in O$ *then*

$$d(\Phi_{s,u}(p),p) \le re^{ru}\int_{-\infty}^{u} d(q(t),p)e^{r|t|}\,dt. \tag{A.4}$$

Proof. Since $d(q(t),p)e^{r|t|} \ge 0$ for all t, it is sufficient to show Eq. (A.4) with $\int_s^u$ in place of $\int_{-\infty}^u$. For each $t \in [s,u)$,

$$\begin{aligned}\liminf_{h\to 0+}\frac{d(\Phi_{s,t+h}(p),p)-d(\Phi_{s,t}(p),p)}{h} &\le \liminf_{h\to 0+}\frac{d(\Phi_{s,t+h}(p),\Phi_{s,t}(p))}{h}\\ &\le rd(q(t),\Phi_{s,t}(p))\\ &\le rd(q(t),p)+rd(\Phi_{s,t}(p),p),\end{aligned}$$

and so by [11, Appendix I, Theorem 2.1], for all $t \in [s,u]$,

$$d(\Phi_{s,t}(p),p) \le \int_s^t rd(q(\tau),p)+rd(\Phi_{s,\tau}(p),p)\,d\tau,$$

and therefore a suitable version of Gronwall's Lemma [8, Corollary 2] yields the result. □

The proof of the first part of Theorem A.1 proceeds exactly as in the proof of Theorem 2.1, except that instead of choosing T so that every $s \le T$ has

$$(C+1)d(q(s),p)e^{r|s|} + Cr\int_{-\infty}^{s} d(q(\tau),p)e^{r|\tau|}\,d\tau < \varepsilon,$$

we choose T so that every $s \le T$ has

$$r\int_{-\infty}^{s} d(q(\tau),p)e^{r|\tau|}\,d\tau < \varepsilon.$$

(Of course, this is equivalent to saying more simply that $r\int_{-\infty}^{T} d(q(\tau),p)e^{r|\tau|}\,d\tau < \varepsilon$.) Then, applying Lemma A.1 in place of Lemma 3.1(B) then recovers Eq. (3.13). The proof of the second part of Theorem A.1 is then identical to the proof of Theorem 2.2, since, given the *conclusion* of Theorem 2.1, the proof of Theorem 2.2 makes no further reference to the fact that the NAC (O,T^*,q) is monotone-like.

References

1. H. M. Alkhayuon and P. Ashwin, Rate-induced tipping from periodic attractors: Partial tipping and connecting orbits, *Chaos: Interdiscip. J. Nonlinear Sci.* **28** (2018) 033608.
2. V. Araújo and E. Trindade, Robust exponential mixing and convergence to equilibrium for singular-hyperbolic attracting sets, *J. Dyn. Diff. Equ.* (2021), doi:10.1007/s10884-021-10100-
3. P. Ashwin and J. Newman, Physical invariant measures and tipping probabilities for chaotic attractors of asymptotically autonomous systems, *Eur. Phys. J. Special Top.* **230** (2021) 3235–3248.

4. P. Ashwin, C. Perryman and S. Wieczorek, Parameter shifts for nonautonomous systems in low dimension: Bifurcation-and rate-induced tipping, *Nonlinearity* **30** (2017) 2185.
5. R. Bowen and D. Ruelle, The ergodic theory of Axiom A flows, *Invent. Math.* **29** (1975) 181–202.
6. M. D. Chekroun, M. Ghil and J. D. Neelin, Pullback attractor crisis in a delay differential ENSO model, in *Advances in Nonlinear Geosciences* (Springer, 2018), pp. 1–33.
7. M. D. Chekroun, E. Simonnet and M. Ghil, Stochastic climate dynamics: Random attractors and time-dependent invariant measures, *Phys. D: Nonlinear Phenomena* **240** (2011) 1685–1700.
8. S. S. Dragomir, *Some Gronwall Type Inequalities and Applications* (Nova Science Publishers, 2003).
9. P. E. Kloeden and M. Rasmussen, *Nonautonomous Dynamical Systems*, Mathematical Surveys and Monographs, Vol. 176 (Amer. Math. Soc., 2011).
10. S. Pierini, M. Ghil and M. D. Chekroun, Exploring the pullback attractors of a low-order quasigeostrophic ocean model: The deterministic case, *J. Climate* **29** (2016) 4185–4202.
11. N. Rouche, P. Habets and M. Laloy, *Stability Theory by Liapunov's Direct Method*, Vol. 4 (Springer, 1977).
12. W. Tucker, A rigorous ODE solver and Smale's 14th problem, *Found. Comput. Math.* **2** (2002) 53–117.
13. L.-S. Young, What are SRB measures, and which dynamical systems have them? *J. Stat. Phys.* **108** (2002) 733–754.