Optimal Scheduling of Dynamic Transport

Panos Tsimpos   PTSIMPOS@MIT.EDU
Operations Research Center, MIT, Cambridge, MA 02139, USA

Zhi Ren   ZREN@MIT.EDU
Department of Mathematics, MIT, Cambridge, MA 02139, USA

Jakob Zech   JAKOB.ZECH@UNI-HEIDELBERG.DE
Interdisciplinary Center for Scientific Computing, Heidelberg University, Heidelberg 69120, Germany

Youssef Marzouk   YMARZ@MIT.EDU
Laboratory for Information and Decision Systems, MIT, Cambridge, MA 02139, USA

Editors: Nika Haghtalab and Ankur Moitra

Abstract

Flow-based methods for sampling and generative modeling use continuous-time dynamical systems to represent a transport map that pushes forward a source measure to a target measure. The introduction of a time axis provides considerable design freedom, and a central question is how to exploit this freedom. Though many popular methods seek straight line (i.e., zero acceleration) trajectories, we show here that a specific class of "curved" trajectories can significantly improve approximation and learning. In particular, we consider the unit-time interpolation of any given transport map T and seek the schedule $\tau : [0,1] \to [0,1]$ that minimizes the spatial Lipschitz constant of the corresponding velocity field over all times $t \in [0,1]$. This quantity is crucial as it allows for control of the approximation error when the velocity field is learned from data. We show that, for a broad class of source/target measures and transport maps T, the *optimal schedule* can be computed in closed form, and that the resulting optimal Lipschitz constant is *exponentially smaller* than that induced by an identity schedule (corresponding to, for instance, the Wasserstein geodesic). Our proof technique relies on the calculus of variations and Γ-convergence, allowing us to approximate the aforementioned degenerate objective by a family of smooth, tractable problems.

Keywords: Measure transport, optimal transport, calculus of variations, Γ-convergence, approximation theory, flow-based models

1. Introduction

Characterizing complex probability distributions is a central task in computational statistics and machine learning. A typical goal is to draw samples from a distribution of interest, given access to its unnormalized density or to a finite set of training samples. Alternatively, given training samples, one might wish to estimate the density function itself. *Transportation of measure* provides a unifying approach to these goals (Tabak and Vanden-Eijnden, 2010; Moselhy and Marzouk, 2012; Tabak and Turner, 2013; Rezende and Mohamed, 2015; Marzouk et al., 2016; Kingma and Dhariwal, 2018; Papamakarios et al., 2021). The essential idea of transport is to represent ν, the target distribution of interest, as the *pushforward* of a tractable source distribution μ by some map T, i.e., $T_\sharp \mu = \nu$. Sampling then follows by evaluating T on samples from μ, and simple expressions for the density of $T_\sharp \mu$ are available when T is invertible and differentiable (Wang and Marzouk, 2022).

To realize these approaches, we must find representations of transport maps T that are sufficiently expressive and that allow easy computation of Jacobian determinants. To this end, a variety of

parameterizations have been proposed, using classical function approximation schemes (such as polynomials or wavelets) (Baptista et al., 2024), neural networks (Huang et al., 2018; Wehenkel and Louppe, 2019; Durkan et al., 2019), and combinations thereof (Jaini et al., 2019). Generally speaking, we can divide these representations into two categories: *static approaches* that represent the displacement from x to $T(x)$ directly (Rezende and Mohamed, 2015; Huang et al., 2018; Jaini et al., 2019; Wang and Marzouk, 2022), and *dynamic approaches* that employ evolution over some fictitious time (Chen et al., 2018; Grathwohl et al., 2019; Onken et al., 2021; Lipman et al., 2023; Liu et al., 2023; Albergo and Vanden-Eijnden, 2022). The latter is our focus here.

In the dynamical approach, a recurring structure is to represent the transport as the flow map of an ODE system:

$$\begin{cases} \frac{d}{dt}X(x,t) &= v(X(x,t),t), \qquad t \in [0,1], \\ X(x,0) &= x. \end{cases} \tag{1}$$

Under mild assumptions, (1) is solvable and induces trajectories $t \mapsto X(x,t)$ for each x that satisfy

$$X(x,t) = x + \int_0^t v(X(x,s),s)ds, \qquad t \in [0,1]. \tag{2}$$

We will refer to the mapping $x \mapsto X(x,t)$ as the time-t *flow map* of the ODE. The transport map obtained by evaluating this flow map at the terminal time $t = 1$ is denoted by $T^v := X(\cdot,1)$. Such parameterizations of the transport map guarantee invertibility and allow efficient computation of the Jacobian log-determinant $\log \det \nabla_x T^v$ through the *instantaneous-change-of-variables-formula* (Chen et al., 2018).

1.1. Challenges and motivations

Though the dynamic approach to transport has demonstrated considerable success in applications, it is also understood that the extra freedom afforded by the time coordinate raises important challenges. Finlay et al. (2020); Onken et al. (2021) observe that unfettered freedom in the dynamics can be practically detrimental: for ODE models trained through likelihood maximization, or any other criterion that depends only on a divergence between $T_\sharp^v \mu$ and ν, the trajectory of the particles for $t < 1$ does *not* directly impact the training objective. As a result, one often obtains highly irregular paths for $t \in [0,1]$ that produce large approximation and numerical integration errors.

For perspective, we note that for any absolutely continuous pair of measures (μ, ν), there are in general infinitely many transport maps T that achieve $T_\sharp \mu = \nu$; moreover, there are infinitely many velocity fields v whose time-one flow maps (2) realize any given T. It is thus natural to ask what is the "best" velocity field in some relevant sense, e.g., from the perspective of approximation and statistical learning. While this general question remains very challenging, here we focus on its "inner layer" and ask: *Given a transport map T, what is the* optimal *velocity field realizing T as the time-one flow map of the ODE system* (1)*?* We shall define a specific notion of optimality, and a specific class of velocity fields over which we optimize, below.

A widely used strategy to address irregular paths is to seek *straight-line* trajectories interpolating x and $T(x)$ for some transport map T; this can be achieved through regularization of a log-likelihood training objective (Finlay et al., 2020; Onken et al., 2021; Marzouk et al., 2025) or via methods that learn the velocity field directly via least squares (Liu et al., 2023). Straight-line trajectories are desirable in the sense that they minimize numerical integration errors; an explicit Euler scheme with

a step size of one is exact if the trajectories are exactly straight. Yet such constructions essentially do *not* exploit the time axis, and invite the question of why be dynamic in the first place. We also note that straight-line trajectories may not create space-time velocity fields $v(x, t)$ that are easy to approximate: in general, near regions of strong concentration or dilation, the velocity field realizing straight-line trajectories can have very large spatial derivatives. Consider, as a simple example, transport of a diffuse distribution at $t = 0$ to a highly concentrated distribution at $t = 1$: the Lipschitz constant of $v(\cdot, t)$ will be very large near $t = 1$. Figure 1 (left) illustrates this phenomenon for two univariate Gaussians. Simplicity in the Lagrangian picture (e.g., zero acceleration) does not translate to simplicity in the Eulerian sense.
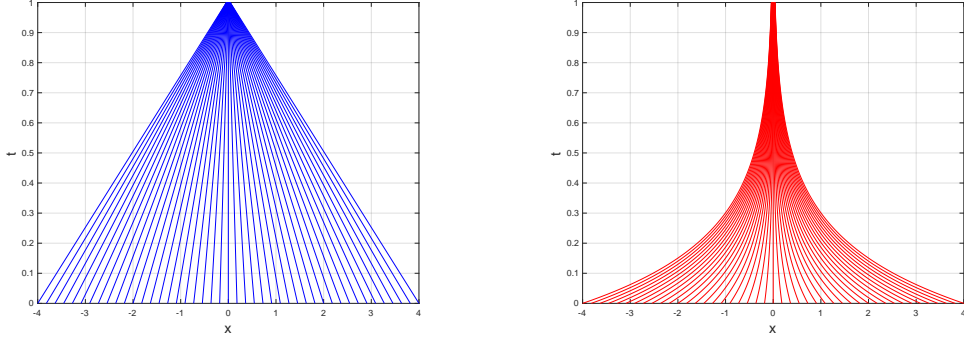

Figure 1: Trajectories $t \mapsto X(x, t)$ transporting $\mu = \mathcal{N}(0, \theta_1^2)$ at $t = 0$ to $\nu = \mathcal{N}(0, \theta_2^2)$ at $t = 1$, with $\theta_2 \ll \theta_1$. *(left)* Straight-line trajectories, with strong spatial derivatives $\partial_x v(x, t)$ apparent near $t = 1$; *(right)* Trajectories produced by the optimal schedule derived in this work.

The (spatial) Lipschitz constant of the velocity field v plays a key role in stability analysis and approximation theory for ODE-based models: Benton et al. (2024); Marzouk et al. (2024) show that distribution error in such models is controlled by the spatial Lipschitz constant of the underlying velocity field. Marzouk et al. (2024, 2025); Gühring and Raslan (2021); Yarotsky (2017) additionally show that the error in a neural network approximation of any target velocity field v is controlled by the spatial Lipschitz constant of the latter.

To make the question highlighted above precise, we will therefore focus on spatial Lipschitz constant of v, seeking to minimize it for any given transport T. We will do so without significantly increasing the small numerical integration error associated with straight-line trajectories (see Remark 9 below). Our idea is intuitive. Given the one-parameter family of flow maps obtained by interpolating T with the identity map, $X(\cdot, t) : x \mapsto tT(x) + (1 - t)x$ for $t \in [0, 1]$, we introduce a *schedule* $\tau : [0, 1] \to [0, 1]$ that is continuously differentiable and monotone increasing, while satisfying $\tau(0) = 0, \tau(1) = 1$. The schedule captures the rate at which we traverse time and naturally defines a new family of flow maps $X_\tau(\cdot, t) : x \mapsto \tau(t) T(x) + (1 - \tau(t)) x$. This family satisfies $X_\tau(\cdot, 0) = \text{id}$ and $X_\tau(\cdot, 1) = T$ and is indeed a dynamic transport scheme. Now we consider velocity fields v that realize $X_\tau(\cdot, t)$ via (1), and search over all admissible τ to find the schedule yielding the smallest spatial Lipschitz constant of v, uniformly over $t \in [0, 1]$. Due to the chain rule, the velocity field that drives the dynamics is now scaled by $\dot{\tau}(t) \coloneqq \frac{d}{dt}\tau(t)$ at time t. As we will show, the largest spatial Lipschitz constant of the original flow occurs at time $t = 0$ and/or $t = 1$, and a good τ slows down

near these endpoints, effectively "averaging out" the spatial derivative over the entire time domain $[0, 1]$. The trajectories traveled are still "straight lines" but with respect to τ.

To the best of our knowledge, the problem of finding schedules that maximize the spatial regularity of a velocity field has not yet been addressed in the literature. For a survey of related work, see Appendix C.

1.2. Our contributions

The main contributions of this work are as follows:

- We formulate the *optimal scheduling* problem: for linear interpolation of given a transport map T and its realization via the flow map of an ODE system, we seek a schedule that minimizes a uniform (in time) bound on the spatial Lipschitz constant of the ODE velocity field v.

- We develop a variational approach to the problem of optimal schedules. Specifically, we show how one can tackle the problem of minimizing the desired Lipschitz bound by using the theory of Γ-convergence (De Giorgi and Franzoni, 1975; Dal Maso, 1993) to cast this non-differentiable problem as a limit of smooth objectives. Then we solve each problem by casting it in Lagrangian form and applying the direct method of the calculus of variations.

- We show that the solution of the optimal scheduling problem satisfies an ODE that can be solved explicitly and in closed form. We then explore the properties of this optimal schedule τ_∞ in a series of examples and observe a universal behavior: τ_∞ has a *sigmoid-like shape* with a unique inflection point $t_0 \in [0, 1]$ and positive curvature in $[0, t_0]$.

- We show that, under the optimal schedule, the uniform Lipschitz bound decreases *exponentially*, and discuss how this can dramatically improve the approximation power of flow models.

While the focus of this paper is on the variational formulation and its analysis, we believe that the optimal schedules found here have many algorithmic implications. In practice, the transport map T is unavailable; otherwise one could directly use T for sampling and there is no need to learn its dynamical representation. As we will show, however, optimal schedules τ_∞ have a certain *universal structure* that is independent of the underlying transport map T. As a result, they can be easily parameterized and made part of the learning problem. Moreover, the optimal schedules found in this work are *independent of any training scheme*: the idea of introducing a time re-parameterized schedule can thus find use in neural ODEs (Chen et al., 2018), flow matching (Lipman et al., 2023), and stochastic interpolants (Albergo and Vanden-Eijnden, 2022).

2. Preliminaries

2.1. Notation

Let $\mathbb{N} = \{1, 2, \dots\}$ be the set of positive integers and let $\mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$. For a closed subset $\Omega \subseteq \mathbb{R}^d$, we denote by $\mathcal{P}(\Omega)_{\text{ac}}$ the set of Borel probability measures on Ω that are absolutely continuous with respect to the Lebesgue measure. For $p \in \mathbb{N}$, we write $L^p(\Omega)$ for the space of measurable functions $f : \Omega \to \mathbb{R}$ such that $\int_\Omega |f|^p \, dx < \infty$, with dx the Lebesgue measure. $L^\infty(\Omega)$ is then the space of measurable functions on Ω with finite essential supremum. For $k \in \mathbb{N}$ and $p \in \mathbb{N}_\infty$, we write $W^{k,p}(\Omega)$ for the space of measurable functions $f : \Omega \to \mathbb{R}$ whose k-th order weak

derivatives are in $L^p(\Omega)$; we also use the shorthand $H^k(\Omega) := W^{k,2}(\Omega)$. Furthermore, for $k \in \mathbb{N}_\infty$, we write $C^k(\Omega)$ for the space of k-times continuously differentiable functions $f : \Omega \to \mathbb{R}$ whose i-th order derivatives, for $i \leq k$, are continuously extendable to the boundary $\partial\Omega$. For $F : \Omega \to \mathbb{R}^d$ we write $F \in C^k(\Omega; \mathbb{R}^d)$ to indicate that each component function F_i of F satisfies $F_i \in C^k(\Omega)$, for $1 \leq i \leq d$. Finally, we introduce a few function spaces particular to our work. For $p \in \mathbb{N}$, let:

$$\mathcal{T}_{2p} := \left\{ \tau \in W^{1,2p}([0,1]) \,:\, \tau(0) = 0\,,\ \tau(1) = 1 \right\} \tag{3}$$

$$\mathcal{T}_{2p}^b := \left\{ \tau \in \mathcal{T}_{2p} \,:\, 0 \leq \tau \leq 1 \right\} \tag{4}$$

$$\mathcal{T}_\infty := \left\{ \tau \in C^1([0,1]) \,:\, \tau(0) = 0\,,\ \tau(1) = 1\,,\ \dot\tau \geq 0 \right\} \tag{5}$$

where we note that (4) is well defined due to the Sobolev embedding theorem; see, e.g., Evans (2022, Theorem 5, Section 5.6). These spaces satisfy $\mathcal{T}_\infty \subset \mathcal{T}_{2p}^b \subset \mathcal{T}_{2p}$ with strict inclusions.

2.2. Sampling with measure transport

Now we present the main problem of our paper. Motivated by the discussion in Section 1, we consider two absolutely continuous measures $\mu, \nu \in \mathcal{P}_{ac}(\Omega)$ supported on a compact, convex domain $\Omega \subset \mathbb{R}^d$ and assume that there exists a measurable map $T : \Omega \to \Omega$ such that the pushforward of μ under T is ν, i.e., for all Borel sets $A \subset \Omega$ we have $\nu(A) = \mu(T^{-1}(A))$. The map additionally satisfies the following conditions:

Assumption 1 *For a compact and convex set $\Omega \subset \mathbb{R}^d$, the transport map $T : \Omega \to \Omega$ is of class $C^1(\Omega; \mathbb{R}^d)$. Moreover, for all $s \in \Omega$, the Jacobian $\nabla T(s)$ is symmetric positive definite.*

For example, T can be the optimal transport (Brenier) map induced by a strictly convex potential.

Now consider a path of measures with endpoints μ and ν, given by $\mu_t = (T_t)_\sharp \mu$ for $t \in [0,1]$, with $T_t(x) := (1-t)x + tT(x)$. The one-parameter family of transport maps $x \mapsto T_t(x)$ is equivalently the time-parameterized flow map $x \mapsto X(x,t)$ of the following ODE:

$$\begin{cases} \frac{d}{dt}X(x,t) &= T(x) - x, \qquad t \in [0,1], \\ X(x,0) &= x. \end{cases} \tag{6}$$

Under the conditions of Assumption 1, Marzouk et al. (2025, Theorem 3.4) establishes the invertibility of T_t for any $t \in [0,1]$. We can then use the chain rule and inverse function theorem to write the above ODE in the more standard form of (1), with a velocity field given by

$$v(\cdot, t) = (T - \mathrm{id}) \circ T_t^{-1}(\cdot).$$

As discussed in Section 1, a key quantity from the perspective of approximation theory is the Lipschitz constant of $x \mapsto v(x,t)$, for each time t. We would like to control this quantity, $\mathrm{Lip}(v(\cdot,t))$, uniformly over $t \in [0,1]$. By Proposition 14 in Appendix A.2, we have:

$$\mathrm{Lip}(v(\cdot,t)) = \sup_{x \in \Omega_t} \|\nabla_x v(x,t)\|_{\mathrm{op}}, \tag{7}$$

where $\Omega_t = \{X(s,t) \,:\, s \in \Omega\}$ is the domain of $v(\cdot,t)$. Hence, we want to control:

$$\Lambda = \sup_{t \in [0,1]} \sup_{x \in \Omega_t} \|\nabla_x v(x,t)\|_{\mathrm{op}}.$$

In the sequel, we will also make the following assumption:

Assumption 2 *Let $(\sigma_j(s))_{j=1}^d$ be the eigenvalues of $\nabla T(s)$ at $s \in \Omega$. There exist $s_0 \in \Omega$ and i such that $\sigma_i(s_0) \neq 1$.*

This condition, albeit technical, makes the problem non-trivial by ruling out "trivial" transports, for example a translation of a distribution. Indeed, if $\sigma_i(s) = 1$ for all i and $s \in \Omega$, then Λ is already zero and there is no benefit in seeking an alternative schedule.

2.3. Schedules

Our goal is to reduce Λ while staying on the trajectories prescribed by T_t. A natural degree of freedom to exploit is the schedule $\tau : [0, 1] \to [0, 1]$ discussed in Section 1, which controls the "rate" at which we traverse time.

Definition 1 *A C^1 function $\tau : [0, 1] \to [0, 1]$ is called a* schedule *if it is non-decreasing, with $\tau(0) = 0$ and $\tau(1) = 1$.*

Recall that the path of measures $(\mu_t)_{t \in [0,1]}$ is induced by flow maps $X(\cdot, t) \equiv T_t(\cdot)$. A schedule naturally modifies the latter by setting $X_\tau(\cdot, t) := X(\cdot, \tau(t))$. The boundary conditions on τ ensure that $X_\tau(\cdot, 0) = \text{id}$ and $X_\tau(\cdot, 1) = T$; hence, transport from $\mu = \mu_0$ to $\nu = \mu_1$ is still achieved. The ODE satisfied by this modified flow is

$$\begin{cases} \frac{d}{dt} X_\tau(x, t) &= \dot{\tau}(t)\, (T(x) - x)\,, \qquad t \in [0, 1], \\ X_\tau(x, 0) &= x, \end{cases} \tag{8}$$

and its associated velocity field is

$$\begin{aligned} v_\tau(\cdot, t) &= \dot{\tau}(t)(T - \text{id}) \circ (X_\tau)^{-1}(\cdot, t) \\ &= \dot{\tau}(t)(T - \text{id}) \circ \left[\, (1 - \tau(t))\, \text{id} + \tau(t) T\, \right]^{-1}, \end{aligned}$$

defined on $\Omega_{\tau(t)}$. We can, therefore, proceed as above to derive an expression for a modified uniform Lipschitz bound $\Lambda[\tau]$:

$$\Lambda[\tau] := \sup_{t \in [0,1]} \sup_{x \in \Omega_{\tau(t)}} \dot{\tau}(t) \left\| \nabla_x v(x, \tau(t)) \right\|_{\text{op}}. \tag{9}$$

The key idea is that the velocity term $\dot{\tau}$ will allow us to suppress the spatial supremum locally in time. The optimization problem that we will investigate is then

$$\inf_{\tau \in \mathcal{T}_\infty} \Lambda[\tau]. \tag{10}$$

In the next section, we will show how to solve problem (10) to optimality.

3. Main Results

Given a transport map T that pushes forward μ to ν, with $\mu, \nu \in \mathcal{P}_{\text{ac}}(\Omega)$ and both T and Ω satisfying Assumption 1, we seek an optimizer τ_∞ of problem (10), called the *optimal schedule*. Our strategy is to put (10) in the form $\inf_{\tau \in \mathcal{T}_2} \int_0^1 \lambda(\tau(t), \dot{\tau}(t))\, dt$ for a search space $\mathcal{T}_2$ (cf. (3)) and a *Lagrangian function* $\lambda : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. The space $\mathcal{T}_2$ strictly contains the space $\mathcal{T}_\infty$ defined in (5). A subsequent argument will establish that the minimizer $\tau_\infty \in \mathcal{T}_2$ is, in fact, in $\mathcal{T}_\infty$.

The first step is to reformulate problem (10) in a more explicit manner, i.e., without appealing to the Lipschitz constant directly.

Theorem 2 *Let T satisfy Assumptions 1 and 2, and let $(\sigma_i(s))_i$ denote the eigenvalues of $\nabla T(s)$, for $s \in \Omega$. Define $f(s) := \max_i \sigma_i(s) - 1$, $g(s) := \min_i \sigma_i(s) - 1$, and $\Omega_{[0,1]} := \Omega \times [0,1]$. Then $\Lambda[\tau]$ can be written as*

$$\Lambda[\tau] = \sup_{(x,t)\in\Omega_{[0,1]}} |\dot{\tau}(t)| \max \left\{ \left| \frac{f(x)}{1 + \tau(t)f(x)} \right| , \left| \frac{g(x)}{1 + \tau(t)g(x)} \right| \right\}. \tag{11}$$

The proof is in Appendix A.2.

To proceed with the variational approach, we wish to cast (11) as an integral of a Lagrangian. Appealing to Lemma 13, we can rewrite (11) as

$$\Lambda[\tau] = \max \left\{ \left\| \frac{\dot{\tau}(t)f(x)}{1 + \tau(t)f(x)} \right\|_{L^\infty(\Omega_{[0,1]})} , \left\| \frac{\dot{\tau}(t)g(x)}{1 + \tau(t)g(x)} \right\|_{L^\infty(\Omega_{[0,1]})} \right\}.$$

The key observation here is that an L^∞ norm is the limit of L^p norms. In fact, this approximation can be iterated in the sense that

$$\max \left\{ \|F\|_{L^\infty}, \|G\|_{L^\infty} \right\} = \lim_{p\to\infty} \left(\|F\|_{L^p}^p + \|G\|_{L^p}^p \right)^{\frac{1}{p}}$$

for functions $F, G \in C^0(\Omega_{[0,1]})$; see Lemma 17 for a proof. It is therefore natural to study the problems

$$\inf_{\tau\in\mathcal{T}_\infty} \int_0^1 \lambda_p(\tau(t), \dot{\tau}(t)) \, dt, \tag{12}$$

where

$$\lambda_p(\tau(t), \dot{\tau}(t)) := \int_\Omega \left(\frac{\dot{\tau}(t)f(x)}{1 + \tau(t)f(x)} \right)^{2p} + \left(\frac{\dot{\tau}(t)g(x)}{1 + \tau(t)g(x)} \right)^{2p} \, dx. \tag{13}$$

Notice that for each $\tau \in \mathcal{T}_\infty$ we have the pointwise limit $\Lambda[\tau] = \lim_{p\to\infty} \left(\int_0^1 \lambda_p(\tau(t), \dot{\tau}(t)) \, dt \right)^{\frac{1}{2p}}$ and the monotonicity of $x \mapsto x^{\frac{1}{2p}}$ allows us to optimize under the exponent.

While this relaxation is intuitive, pointwise convergence of objectives does *not* guarantee convergence of minimizers. Understanding the latter is exactly the aim of Γ-convergence, a powerful tool in the calculus of variations. For a general topological space X, a functional $F : X \to \mathbb{R}$ and a family of functionals $(F_p)_{p\in\mathbb{N}}$ with $F_p : X \to \mathbb{R}$, one says that F_p Γ-converges to F if for any $x \in X$ we have (Braides, 2006)

$$F(x) \le \liminf_{p\to\infty} F_p(x_p) \text{ for any sequence } x_p \to x, \text{ and}$$

$$F(x) \ge \limsup_{p\to\infty} F_p(x_p) \text{ for some sequence } x_p \to x.$$

If Γ-convergence is established, together with a so-called *equi-coercivity*[1] condition, one can deduce that subsequential limits of minimizers of each F_p are minimizers of F. This is the celebrated *fundamental theorem of Γ-convergence* (Braides, 2006). Using this machinery, we prove the following result.

---

1. A family of functionals F_p is said to be equi-coercive (Braides, 2006) if for any $t > 0$ there is a compact set $K_t \subset X$ such that for all p we have $\{x \in X : F_p(x) \le t\} \subset K_t$.

Theorem 3 *Let $\mathcal{T}_2$ be as in (3) with $p = 1$. For arbitrary $p \in \mathbb{N}$, consider the optimization problem*

$$\inf_{\tau \in \mathcal{T}_2} \int_0^1 \lambda_p \left(\tau(t), \dot{\tau}(t) \right) \, dt \,, \tag{14}$$

with λ_p defined in (13). If τ_p is a minimizer of (14), then any subsequential L^2 limit $\tau_{p_j} \to \tau_\infty$ as $j \to \infty$ is a minimizer of (10).

The proof is in Appendix A.3. As discussed above, the key idea is to establish that the functionals

$$\Lambda_p : \mathcal{T}_\infty \to \mathbb{R}$$

$$\tau \mapsto \left(\int_0^1 \lambda_p(\tau(t), \dot{\tau}(t)) \, dt \right)^{\frac{1}{2p}}$$

Γ-converge to Λ, and that the family $(\Lambda_p)_{p \in \mathbb{N}}$ is equi-coercive. Notice that in our previous discussion of Γ-convergence, the topology of the underlying space X was intentionally left vague: too strong of a topology will break the coercivity condition, whereas too weak of a topology will make verifying Γ-convergence impossible. Here, the balance is struck by expanding the space of feasible schedules to a closed subset of the Sobolev space $H^2([0,1])$ and endowing this subset with the topology of weak convergence (or, equivalently, the weak topology). Inside sets of bounded H^1 norm, this topology is equivalent to the strong L^2 topology; hence the conclusion of the theorem.

Having established the validity of our approximation, we now solve the approximating problem (14) by employing calculus of variations.

Theorem 4 *There exists a minimizer τ_p of (14) that satisfies the following ODE:*

$$\frac{d}{dt}\tau_p(t) = \frac{1}{Z_p} \left(\int_\Omega \frac{f(s)^{2p}}{(1 + \tau_p(t)f(s))^{2p}} + \frac{g(s)^{2p}}{(1 + \tau_p(t)g(s))^{2p}} \, ds \right)^{-1/(2p)}, \tag{15}$$

with boundary conditions $\tau_p(0) = 0$ and $\tau_p(1) = 1$ and

$$Z_p = \int_0^1 \left(\int_\Omega \left(\frac{f(s)}{1 + \tau_p(t)f(s)} \right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1 + \tau_p(t)g(s)} \right)^{2p} ds \right)^{-\frac{1}{2p}} dt.$$

Proving the above theorem (see Appendix A.4) is a mostly technical endeavor, i.e., verifying that the Lagrangian λ_p is tame enough so that the minimizers of (14) satisfy the (strong) Euler–Lagrange equations. The difficulty lies in proving that minimizers are, *a priori*, in $C^2([0,1])$; standard regularity theory does not apply. To overcome this, we further approximate λ_p by the family $\lambda_p^\epsilon(x, v) = \lambda_p(x, v) + \epsilon v^2$, which possess smooth minimizers, and take another Γ-limit along $\epsilon \to 0$.

As a result, if we find *some* L^2 subsequential limit of the solutions τ_p of the ODEs (14), we will find a solution to the original problem (10). One is tempted to take the (pointwise) limit $p \to \infty$ of the right hand side of (15) to obtain a new ODE. This ODE in fact characterizes a solution of (10).

Theorem 5 *Let τ_∞ be the solution to the following ODE,*

$$\frac{d}{dt}\tau_\infty(t) = \frac{1}{Z} \max \left\{ \left\| \frac{f(s)}{1 + \tau_\infty(t)f(s)} \right\|_{L^\infty}, \left\| \frac{g(s)}{1 + \tau_\infty(t)g(s)} \right\|_{L^\infty} \right\}^{-1}, \tag{16}$$

with boundary conditions $\tau_\infty(0) = 0$ and $\tau_\infty(1) = 1$ and

$$Z = \int_0^1 \max\left\{ \left\| \frac{f(s)}{1 + \tau_\infty(t)f(s)} \right\|_{L^\infty}, \left\| \frac{g(s)}{1 + \tau_\infty(t)g(s)} \right\|_{L^\infty} \right\}^{-1} dt \,.$$

If τ_p are solutions to the ODEs (14), then there is a subsequence $\tau_{p_j} \to \tau_\infty$ in L^2 and therefore, by Theorem 3, τ_∞ is optimal for (10).

To prove the above theorem (see Appendix A.5), one applies a Grönwall inequality (Howard, 1998) to control the difference between solutions to the ODEs (15) and (16). The key idea, and admittedly somewhat overpowered in the context of this proof, is to combine this with the Arzelà–Ascoli theorem, allowing one to obtain uniform convergence of the right hand side of (15) to that of (16). This result, combined with the Grönwall estimate, gives the required convergence.

Somewhat remarkably, the ODE (16) is solvable in closed form:

Theorem 6 *Let $f^* := \sup_{s \in \Omega} f(s)$, $g_* := \inf_{s \in \Omega} g(s)$ and consider*

$$t_0 = \frac{\ln\left[\frac{1}{2}\left(1 - \frac{f^*}{g_*} \right) \right]}{\ln\left[\frac{1}{4}\left(2 - \frac{f^*}{g_*} - \frac{g_*}{f^*} \right) \right] - \ln(g_* + 1)}, \tag{17}$$

$$\tau(t_0) = -\frac{1}{2}\left(\frac{1}{f^*} + \frac{1}{g_*} \right). \tag{18}$$

If $0 \le t_0 \le 1$, then the solution to the ODE (16) is given by:

$$\tau_\infty(t) = \begin{cases} \frac{1}{f^*}\left\{ \frac{1}{4(g_*+1)}\left(2 - \frac{f^*}{g_*} - \frac{g_*}{f^*} \right) \right\}^t - \frac{1}{f^*}, & t \le t_0 \\[2ex] \frac{1}{2}\left(\frac{1}{g_*} - \frac{1}{f^*} \right)\left\{ \frac{2(g_*+1)}{(1 - \frac{g_*}{f^*})} \right\}^t \left\{ \frac{1 - \frac{g_*}{f^*}}{2} \right\}^{1-t} - \frac{1}{g_*}, & t \ge t_0 \,. \end{cases} \tag{19}$$

Otherwise, if $t_0 \notin [0, 1]$, we have two cases: if $f^* \ge -g_*$ then the solution to (16) is given by

$$\tau_\infty(t) = \frac{(f^* + 1)^t - 1}{f^*}, \tag{20}$$

and if $f^* < -g_*$, then the solution to (16) is given by

$$\tau_\infty(t) = \frac{(g_* + 1)^t - 1}{g_*}. \tag{21}$$

The proof of this theorem (see Appendix A.6) is a cumbersome computation. The key observation is that the infinity norms inside the $\max$ in (16) can be computed explicitly in terms of the data of the problem. Once this is done, we notice that the $\max$ has one or zero transition points, in the sense that there is at most one $t_0 \in [0, 1]$ where the two functions inside the $\max$ are equal. We can then solve this problem explicitly by solving two ODEs, one in $[0, t_0]$ and one in $[t_0, 1]$, while matching the solutions at t_0.

Due to Theorem 6, the optimal schedule τ_∞ is available in closed form given tight upper/lower bounds on the largest/smallest eigenvalues of the Jacobian of the transport map T, respectively.

Intuitively, these quantities capture the worst-case dilation and concentration associated with the transport. If T is available numerically, then obtaining these values is relatively straightforward via differentiation and optimization, but that is a trivial case. More interestingly, if T is a Brenier map, then one can estimate the largest and smallest eigenvalues of the Hessian of the convex potential ϕ satisfying $T = \nabla\phi$. It is common to assume there are constants α, β such that

$$\alpha I_d \preceq \nabla^2\phi(s) \preceq \beta I_d, \quad \forall s \in \Omega, \tag{22}$$

i.e., that the potential is strongly convex and smooth. This yields $f^* \leq \beta - 1$ and $g_* \geq \alpha - 1$, leading to an estimate of the optimal schedule; of course, if (22) is tight then the estimate is exact.

Finally, we can substitute our optimal schedule τ_∞ in the objective functional $\Lambda[\,\cdot\,]$ to compute its optimal value. To make notation more transparent, set

$$\sigma^*_{\min} := \inf_{s \in \Omega} \sigma_{\min}(s) \quad \text{and} \quad \sigma^*_{\max} := \sup_{s \in \Omega} \sigma_{\max}(s),$$

noting that that $\sigma^*_{\max} = f^* + 1$ and $\sigma^*_{\min} = g_* + 1$. Then we have:

Theorem 7 *For Λ as in (9) and the trivial schedule $\bar{\tau} : t \mapsto t$,*

$$\Lambda[\bar{\tau}] = \max\left\{\sigma^*_{\max} - 1, \frac{1 - \sigma^*_{\min}}{\sigma^*_{\min}}\right\}.$$

For the optimal schedule τ_∞, i.e., the minimizer of problem (10) given in Theorem 5, we have the following cases. If there exists $t_0 \in [0,1]$ such that $\tau_\infty(t_0) = -\frac{1}{2}\left(\frac{1}{\sigma^*_{\max}-1} + \frac{1}{\sigma^*_{\min}-1}\right)$, then

$$\Lambda[\tau_\infty] = \ln\left[\frac{\sigma^*_{\max} - 1}{\sigma^*_{\min}}\right] + \ln\left[\frac{1}{4}\left(\frac{1}{1 - \sigma^*_{\min}} + \frac{1 - \sigma^*_{\min}}{(\sigma^*_{\max} - 1)^2} + \frac{2}{\sigma^*_{\max} - 1}\right)\right]. \tag{23}$$

Otherwise, we have

$$\Lambda[\tau_\infty] = \begin{cases} \ln\left(\sigma^*_{\max}\right) & \text{if } \sigma^*_{\max} + \sigma^*_{\min} \geq 2 \\ -\ln\left(\sigma^*_{\min}\right) & \text{if } \sigma^*_{\max} + \sigma^*_{\min} < 2. \end{cases}$$

This result (see proof in Appendix A.7) tells us that, for large $\sigma^*_{\max}$ and small $\sigma^*_{\min}$, the optimal uniform Lipschitz bound $\Lambda[\tau_\infty]$ is *exponentially smaller* than the bound corresponding to the trivial schedule $\Lambda[\bar{\tau}]$, up to corrections of order one. More precisely, for real-valued functions $\alpha, \beta : \mathbb{R}^2 \to \mathbb{R}$, let $\alpha \asymp \beta$ denote that there exist constants $c_1, c_2 > 0$ and some $y \in \mathbb{R}^2$ such that $c_1\beta(x) \leq \alpha(x) \leq c_2\beta(x)$ for all $x_i \geq y_i$ and $i \in \{1, 2\}$. The Lipschitz bound then scales as follows:

Corollary 8 *Viewing $\Lambda[\tau_\infty]$ and $\Lambda[\bar{\tau}]$ as functions of $\sigma^*_{\max}$ and $1/\sigma^*_{\min}$, we have*

$$\Lambda[\bar{\tau}] \asymp \max\left\{\sigma^*_{\max}, \left(\sigma^*_{\min}\right)^{-1}\right\}$$

$$\Lambda[\tau_\infty] \asymp \ln\sigma^*_{\max} - \ln\sigma^*_{\min}.$$

Remark 9 *This result has important consequences for control of the error in the target measure ν when the velocity is learned from data or otherwise approximated. In fact, we claim that the standard error bounds become exponentially tighter. Consider an approximate velocity $(x,t) \mapsto \hat{v}(x,t)$ that is ϵ-close to the true velocity field $(x,t) \mapsto v(x,t)$ in a C^0 sense. By Marzouk et al.*

(2025, Corollary 5.2), the resulting distribution approximation error in the Wasserstein-2 metric is bounded by $\epsilon \exp(\sup_{t\in[0,1]} L_t)$ where $L_t = \text{Lip}(v(\cdot,t))$ is the spatial Lipschitz constant of v. Thus, if we re-parameterize time with the optimal schedule τ_∞, the bound on the distribution error becomes $\mathcal{O}\left(\frac{\sigma^*_{\max}}{\sigma^*_{\min}} \epsilon\right)$ instead of $\mathcal{O}\left(\exp\left(\max\left\{\sigma^*_{\max}, (\sigma^*_{\min})^{-1}\right\}\right) \epsilon\right)$ for the trivial schedule. On the other hand, while the straight-line trajectories obtained with the trivial schedule have essentially zero numerical integration error, we claim that the price we pay in numerical error due to non-straight trajectories is small compared to the improvement in the approximation error. Consider approximating the ODE flow map $x \mapsto X_{\tau_\infty}(x,1)$ with a forward Euler discretization of step size h. The numerical integration error is then bounded by $\frac{hM}{2\Lambda[\tau_\infty]}\left(e^{\Lambda[\tau_\infty]} - 1\right)$ (Atkinson, 1989), where $M = \sup_{x\in\Omega} \sup_{t\in[0,1]} \left|\frac{d^2}{dt^2} X_{\tau_\infty}(x,t)\right| = \sup_{x\in\Omega} \sup_{t\in[0,1]} \ddot{\tau}_\infty(t)|T(x) - x| = \mathcal{O}(1)$, if we assume the eigenvalues are uniformly upper and lower bounded. Using the triangle inequality, it is not hard to see that combined distribution approximation error in a Wasserstein-2 sense is bounded by (Sagiv, 2019, Theorem 2)

$$\mathcal{O}\left(\left(\frac{\sigma^*_{\max}}{\sigma^*_{\min}}\right)\epsilon + \left(\frac{\sigma^*_{\max}}{\sigma^*_{\min}\left(\ln\sigma^*_{\max} - \ln\sigma^*_{\min}\right)}\right)h\right),$$

which is still exponentially tighter than the combined error with the trivial schedule. We believe this heuristic argument justifies working with velocity fields that are constant speed with respect to τ, as they exponentially lower the Lipschitz constant while maintaining good control over the integration error.

Remark 10 *In this paper, we have solved the optimization problem (10) by variational calculus. One might instead try treating the problem directly, especially with foresight of the solution. For example, we notice that the optimizer of problem (10), as given in Theorem 6, eliminates time dependence from the argument to the supremum in equation (11). Thus we might have "guessed" a* priori *that the optimal solution satisfies*

$$|\dot{\tau}(t)| \sup_{x\in\Omega} \max\left\{\left|\frac{f(x)}{1 + \tau(t)f(x)}\right|, \left|\frac{g(x)}{1 + \tau(t)g(x)}\right|\right\} = Z,$$

for some constant $Z \in \mathbb{R}$, which anticipates the form of the ODE in Theorem 5. This link can be seen by using Lemma 13 to exchange the sup *and* max *in (11) and then Lemma 45 to compute the* sup, *for any $t \in [0,1]$. Writing down such an ansatz for the minimizer, however, does not establish its optimality.*

4. Examples

In this section, we present analytical examples of optimal schedules τ_∞. In the process, qualitative features of τ_∞ are revealed.

4.1. Univariate Gaussian distribution

Perhaps the simplest example to consider is transport between univariate Gaussian distributions[2] $\mu_1 = \mathcal{N}(\mu_1, \theta_1^2)$ and $\mu_2 = \mathcal{N}(\mu_2, \theta_2^2)$. The optimal transport map between μ_1 and μ_2 is linear

---

2. See Appendix B for a detailed discussion.

([Chewi et al., 2024](#)):

$$T(x) = \mu_2 + \frac{\theta_2}{\theta_1}(x - \mu_1).$$

First, notice that working in one dimension means that for all $s \in \Omega$ we have $\sigma_{\max}(s) = \sigma_{\min}(s) = \theta_2/\theta_1$. Thus we must have $f^* = g_*$, meaning that the transition time is $t_0 = -\infty$. Substituting $f^* = g_* = \theta_2/\theta_1 - 1$ into (20), we obtain

$$\tau_\infty(t) = \frac{\left(\frac{\theta_2}{\theta_1}\right)^t - 1}{\frac{\theta_2}{\theta_1} - 1}.$$

Notice that the boundary conditions are satisfied, i.e., $\tau_\infty(0) = 0$ and $\tau_\infty(1) = 1$, and, moreover, the concavity or convexity of the curve depends on whether $\theta_1 > \theta_2$ or $\theta_1 < \theta_2$, respectively. This is intuitive: if $\theta_1 > \theta_2$ then we are mapping a wider Gaussian to a narrower one, causing particles traveling along the flow to concentrate at $t = 1$, making the spatial Lipschitz constant largest at that time. The optimal schedule *slows* precisely near $t = 1$ to suppress the increased Lipschitz constant. Symmetrically, when $\theta_1 < \theta_2$ particles "explode" at small times $t \approx 0$ and the optimal schedule slows near $t = 0$, tempering this behavior. Finally, for $\theta_1 \nearrow \theta_2$,[3] one can write $\theta_2/\theta_1 = 1 + \epsilon$ with $\epsilon \searrow 0$ and Taylor expand to obtain $\tau_\infty(t) = \frac{(1+\epsilon t + o(\epsilon))-1}{\epsilon} = t + \frac{o(\epsilon)}{\epsilon} \to t$ as $\epsilon \searrow 0$. In other words, as T tends to an isometry (a translation in this case) the optimal schedule tends to the trivial schedule. Finally, one may compute:

$$\Lambda[t \mapsto t] = \max\left\{\frac{\theta_2}{\theta_1} - 1, \frac{\theta_1}{\theta_2}\right\} \quad \text{and} \quad \Lambda[\tau_\infty] = \left|\ln\left(\frac{\theta_2}{\theta_1}\right)\right|,$$

making the exponential improvement manifest. Paths following the trivial schedule and the optimal schedule are illustrated in Figure 1, and an example of the optimal schedule is plotted in Figure 2(left).

4.2. Generic distributions μ and ν and ill-conditioned T

For any $\mu, \nu \in \mathcal{P}_{\mathrm{ac}}(\Omega)$, let T be such that $\sigma_{\max}^* \gg 1$ and $\sigma_{\min}^* \ll 1$. This type of ill-conditioned transport is, in some sense, exactly what we aim to address in this paper. To make notation more explicit, write:

$$M = \sigma_{\max}^* := \sup_{s \in \Omega} \sigma_{\max}(s) \gg 1 \quad \text{and} \quad \epsilon = \sigma_{\min}^* := \inf_{s \in \Omega} \sigma_{\min}(s) \ll 1.$$

First, we will attempt to understand the qualitative behavior of the optimal schedule τ_∞. Looking at (19) and recalling that $f^* = \sigma_{\max}^* - 1$ and $g_* = \sigma_{\min}^* - 1$, we have:

$$\tau_\infty(t) = \begin{cases} \frac{1}{M-1}\left\{\frac{1}{4\epsilon}\left(\frac{M-1}{1-\epsilon} + \frac{1-\epsilon}{M-1} + 2\right)\right\}^t - \frac{1}{M-1}, & t \leq t_0, \\ -\frac{1}{2}\left(\frac{1}{M-1} + \frac{1}{1-\epsilon}\right)\left\{\frac{\epsilon}{\frac{1}{2}(1+\frac{1-\epsilon}{M-1})}\right\}^t \left\{\frac{1+\frac{1-\epsilon}{M-1}}{2}\right\}^{1-t} + \frac{1}{1-\epsilon}, & t > t_0. \end{cases}$$

Now considering the double limit $\frac{1}{\epsilon}, M \to \infty$ and using Landau notation as needed, we have:

$$\tau_\infty(t) = \begin{cases} \frac{1}{M}\left(e^{t \ln\left(\frac{3}{4}\frac{M}{\epsilon}\right)+o(1)}\right)(1 + o(1)), & t \leq t_0, \\ -\frac{1}{2}(1 + o(1))\, e^{t\, \ln(4\epsilon)+\ln\frac{1}{2}+o(1)} + 1 + o(1), & t > t_0. \end{cases}$$

---

3. The limit is needed since the case $\theta_1 = \theta_2$ is not covered by our framework, which assumes that our transformation T is *not* an isometry.

What this first-order analysis reveals is that, for large M and small ϵ, the optimal schedule exhibits a *sigmoid-like shape* with a transition time $t_0 \in (0, 1)$. Specifically, for $t \in [0, t_0]$ the curve will be of the form $Ae^{\alpha t}$, for $A, \alpha \in \mathbb{R}^+$, whereas for $t \in [t_0, 1]$ the curve will behave as $-Be^{-\beta t}$, for $B, \beta \in \mathbb{R}^+$. This behavior is intuitive: the optimal schedule is attempting to slow the dynamics at the beginning *and* at the end of the transport, where the Lipschitz constant is largest. Turning now to the transition time t_0, we use (17) to compute

$$t_0 = \frac{\ln\left[\frac{1}{2}\left(1 + \frac{M-1}{1-\epsilon}\right)\right]}{\ln(\frac{1}{\epsilon}) + \ln\left[\frac{1}{4}\left(\frac{M-1}{1-\epsilon} + \frac{1-\epsilon}{M-1} + 2\right)\right]} = \frac{\ln M + o(1)}{\ln \epsilon^{-1} + \ln M + o(1)}.$$

Notice that if we take $\epsilon^{-1} = \Theta(M^\gamma)$ for some $\gamma \in \mathbb{R}^+$, then we have:

$$t_0 = \frac{\ln M}{\ln (M^\gamma) + \ln M} + o(1) = \frac{1}{\gamma + 1} + o(1).$$

That is, if ϵ decays within a polynomial factor of $1/M$, then *we will always have a transition time* $t_0 \in (0, 1)$ for any positive exponent $\gamma \in \mathbb{R}^+$. Moreover, notice that $t_0 > 1/2 \iff \gamma < 1$, $t_0 < 1/2 \iff \gamma > 1$ and $t_0 = 1/2 \iff \gamma = 1$. This is intuitive: if ϵ decays within the same order of magnitude as $1/M$, then the optimal schedule treats contraction and expansion equally. It therefore allocates the same amount of time to suppressing either type of singularity. If ϵ decays more quickly than $1/M$ (i.e., $\gamma > 1$), then the optimal schedule will allocate more time to suppressing the contractive behavior of T and oppositely when $\gamma < 1$.

4.3. Multi-modal distributions

In this section, we discuss the implications of the optimal schedule on multi-modal distributions, which naturally lead to pathological flows. Fix a compact and convex $\Omega \subset \mathbb{R}^d$ and let μ be localized around some $x_0 \in \Omega$ while also satisfying $\mathrm{supp}\, \mu = \Omega$. By "localized," we mean that there is some $r \ll \mathrm{diam}\, \Omega$ such that $\mu(B_r(x_0)) > 1 - \epsilon$ for some small $\epsilon > 0$. Now consider weights $q_1, \ldots, q_n > 0$ such that $\sum_i q_i = 1$ and let $\nu = \sum_{i=1}^n q_i \nu_i$, where each ν_i is localized around some $x_i \in \Omega$ with all x_i distinct, again meaning that for $r_i \ll \mathrm{diam}\, \Omega$ we have $\nu_i(B_{r_i}(x_i)) > 1 - \epsilon$. Also, $\mathrm{supp}\, \nu_i = \Omega$ for all i. Moreover, let $D = \max_{i,j} \|x_i - x_j\|$ be the maximal modal separation and, without loss of generality, assume that x_1 and x_2 satisfy $\|x_1 - x_2\| = D$. Now suppose T is a transport map coupling μ to ν such that for some $c > 0$, depending only on the domain Ω and the weights $(q_i)_{i=1}^n$, we have that for all $x, y \in \Omega$:

$$\|x - y\| \frac{1}{c} \leq \|T(x) - T(y)\| \leq \|x - y\| \, c \, D. \tag{24}$$

Such inequalities are expected to hold for regular enough μ and ν, e.g., if their densities are uniformly bounded away from zero.[4] Furthermore, assume that $q_1, q_2 \geq 2\epsilon$, i.e., the most well-separated modes contain a non-trivial amount of mass; and that $r_i < D/(2n)$ and $\epsilon < 1/2$, i.e., all modes are localized on a scale smaller than D. Since $T_\sharp \mu = \nu$, there must exist $y_1, y_2 \in B_{r_0}(x_0)$ such that $T(y_1) \in B_{r_1}(x_1)$ and $T(y_2) \in B_{r_2}(x_2)$. In particular, this implies that $\|T(y_1) - T(y_2)\| \geq D - (r_1 + r_2) \geq D/2$. Thus, the Lipschitz constant of T is at least $D/2$ and the objective under the trivial schedule satisfies

$$\Lambda[t \mapsto t] \geq D/2.$$

---

4. Specifically, one does not expect to dilate μ more than the maximal separation D of the modes.

Due to (24), however, we have that the minimal and maximal eigenvalues of the Jacobian of T are bounded by $1/c$ and cD, respectively. Hence via Theorem 7 we have

$$\Lambda[\tau_\infty] \leq \ln\left(\frac{c^2}{4}\,D\right) + o(1) \quad \text{as} \quad D \to \infty\,.$$

This illustration is meant to convey that when the source is unimodal and the target is multi-modal, we expect that under the trivial schedule, the Lipschitz constant of our flow will be at least as bad as the maximal separation of the modes D. Applying the optimal schedule will, most often, bring this down to a logarithmic factor in D, within constants. For a numerical illustration, please refer to Appendix D.

5. Discussion

The present work suggests many avenues for future exploration. First, we believe that the proof techniques developed here can be extended to more general schedules and flows. For instance, we can extend the notion of a schedule, looking for *matrix-valued* maps $\chi : [0,1] \to \mathrm{GL}(d)$ that modify the linear dynamics via $\frac{d}{dt}X_\chi(x,t) = \dot\chi(t)\,(T(x) - x)$, where $\chi(t) = \tau(t)I_d$ recovers the present problem. It is also of interest to solve the optimal scheduling problem for paths that depart from the displacement interpolation of a given transport; this of course touches on the broader problem of identifying optimal *paths*, again from the perspective of velocity field regularity. We also note that the symmetry assumption on ∇T (cf. Assumption 1) can likely be relaxed if one is willing to work with *singular values* instead of eigenvalues; in this case, however, one would loose the straightforward dilation / compression intuition used to understand the optimal schedule.

Second, we suggest that our method for finding optimal schedules, and the form of these schedules, can spark more efficient algorithms for learning dynamical representations of transport. While the optimal τ_∞ in Theorem 6 depends on the eigenvalues of ∇T, its general structure (a sigmoid-like curve with transition time $t_0 \in [0,1]$) is independent of T. It could thus be well represented with a few-parameter function and simply made part of the learning problem. Alternatively, we note that the optimal schedule does not require full information about the underlying map T: only the maximal and minimal eigenvalues of its Jacobian are needed. Thus one could devise an iterative scheme starting, e.g., with the trivial schedule $\tau = \mathrm{Id}$, estimating a velocity field v and the maximal and minimal eigenvalues of the resulting ∇T^v, computing the implied optimal schedule τ_∞, updating the velocity field v_τ accordingly, and so on. We leave the theoretical and empirical exploration of these ideas for future work.

Acknowledgments

References

Michael S. Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. *CoRR*, abs/2209.15571, 2022. doi: 10.48550/ARXIV.2209.15571. URL https://doi.org/10.48550/arXiv.2209.15571.

Santiago Aranguri, Giulio Biroli, Marc Mezard, and Eric Vanden-Eijnden. Optimizing noise schedules of generative models in high dimensions, 2025. URL https://arxiv.org/abs/2501.00988.

Kendall E. Atkinson. *An Introduction to Numerical Analysis*. John Wiley & Sons, New York, second edition, 1989. ISBN 0471500232. URL http://www.worldcat.org/isbn/0471500232.

Ricardo Baptista, Youssef M. Marzouk, and Olivier Zahm. On the representation and learning of monotone triangular transport maps. *Found. Comput. Math.*, 24(6):2063–2108, 2024. doi: 10.1007/S10208-023-09630-X. URL https://doi.org/10.1007/s10208-023-09630-x.

Joe Benton, George Deligiannidis, and Arnaud Doucet. Error bounds for flow matching methods. *Trans. Mach. Learn. Res.*, 2024, 2024. URL https://openreview.net/forum?id=uqQPyWFDhY.

Vladimir Igorevich Bogachev and Maria Aparecida Soares Ruas. *Measure theory*, volume 1. Springer, 2007.

Andrea Braides. A handbook of γ-convergence. In Michel Chipot and Pavel Quittner, editors, *Handbook of Differential Equations: Stationary Partial Differential Equations*, volume 3, pages 101–213. Elsevier, 2006.

H Brezis. Functional analysis, sobolev spaces and partial differential equations, 2011.

Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in Neural Information Processing Systems*, 31, 2018.

Sinho Chewi, Jonathan Niles-Weed, and Philippe Rigollet. Statistical optimal transport. *arXiv preprint arXiv:2407.18163*, 2024.

Bernard Dacorogna. *Direct methods in the calculus of variations*, volume 78. Springer Science & Business Media, 2007.

Gianni Dal Maso. *An Introduction to Γ-Convergence*. Springer International Publishing, 1993.

Ennio De Giorgi and Tullio Franzoni. Su un tipo di convergenza variazionale. *Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti*, 58(6): 842–850, 1975.

Carles Domingo-Enrich, Jiequn Han, Brandon Amos, Joan Bruna, and Ricky T. Q. Chen. Stochastic optimal control matching. *CoRR*, abs/2312.02027, 2023. doi: 10.48550/ARXIV.2312.02027. URL https://doi.org/10.48550/arXiv.2312.02027.

Carles Domingo-Enrich, Michal Drozdzal, Brian Karrer, and Ricky T. Q. Chen. Adjoint matching: Fine-tuning flow and diffusion generative models with memoryless stochastic optimal control. *CoRR*, abs/2409.08861, 2024. doi: 10.48550/ARXIV.2409.08861. URL https://doi.org/10.48550/arXiv.2409.08861.

Conor Durkan, Artur Bekasov, Iain Murray, and George Papamakarios. Neural spline flows. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pages 7509–7520, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/7ac71d433f282034e088473244df8c02-Abstract.html.

Lawrence C Evans. *Partial differential equations*, volume 19. American Mathematical Society, 2022.

Alessio Figalli and Federico Glaudo. *An invitation to optimal transport, Wasserstein distances, and gradient flows*. 2023.

Chris Finlay, Jörn-Henrik Jacobsen, Levon Nurbekyan, and Adam Oberman. How to train your neural ODE: the world of Jacobian and kinetic regularization. In *International conference on machine learning*, pages 3154–3164. PMLR, 2020.

Adriana Garroni, Vincenzo Nesi, and Marcello Ponsiglione. Dielectric breakdown: optimal bounds. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 457(2014):2317–2335, 2001.

Will Grathwohl, Ricky T. Q. Chen, Jesse Bettencourt, Ilya Sutskever, and David Duvenaud. FFJORD: free-form continuous dynamics for scalable reversible generative models. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019. URL https://openreview.net/forum?id=rJxgknCcK7.

Ingo Gühring and Mones Raslan. Approximation rates for neural networks with encodable weights in smoothness spaces. *Neural Networks*, 134:107–130, 2021. doi: 10.1016/J.NEUNET.2020.11.010. URL https://doi.org/10.1016/j.neunet.2020.11.010.

Ralph Howard. The gronwall inequality. *lecture notes*, 1998.

Chin-Wei Huang, David Krueger, Alexandre Lacoste, and Aaron C. Courville. Neural autoregressive flows. In Jennifer G. Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pages 2083–2092. PMLR, 2018. URL http://proceedings.mlr.press/v80/huang18d.html.

Priyank Jaini, Kira A. Selby, and Yaoliang Yu. Sum-of-squares polynomial flow. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, volume 97 of *Proceedings of Machine Learning Research*, pages 3009–3018. PMLR, 2019. URL http://proceedings.mlr.press/v97/jaini19a.html.

Robert C James. Weak compactness and reflexivity. *Israel Journal of Mathematics*, 2:101–119, 1964.

Diederik P. Kingma and Prafulla Dhariwal. Glow: Generative flow with invertible 1x1 convolutions. In Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pages 10236–10245, 2018. URL https://proceedings.neurips.cc/paper/2018/hash/d139db6a236200b21cc7f752979132d0-Abstract.html.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023. URL https://openreview.net/forum?id=PqvMRDCJT9t.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023. URL https://openreview.net/forum?id=XVjTT1nw5z.

Youssef Marzouk, Tarek Moselhy, Matthew Parno, and Alessio Spantini. *Sampling via Measure Transport: An Introduction*, pages 1–41. Springer International Publishing, Cham, 2016. ISBN 978-3-319-11259-6. doi: 10.1007/978-3-319-11259-6_23-1. URL https://doi.org/10.1007/978-3-319-11259-6_23-1.

Youssef Marzouk, Zhi Ren, and Jakob Zech. Distribution learning via neural differential equations: minimal energy regularization and approximation theory, 2025. URL https://arxiv.org/abs/2502.03795.

Youssef M. Marzouk, Zhi Ren, Sven Wang, and Jakob Zech. Distribution learning via neural differential equations: A nonparametric statistical perspective. *J. Mach. Learn. Res.*, 25:232:1–232:61, 2024. URL https://jmlr.org/papers/v25/23-1280.html.

Tarek A. Moselhy and Youssef M. Marzouk. Bayesian inference with optimal maps. *J. Comput. Phys.*, 231(23):7815–7850, 2012. doi: 10.1016/J.JCP.2012.07.022. URL https://doi.org/10.1016/j.jcp.2012.07.022.

Derek Onken, Samy Wu Fung, Xingjian Li, and Lars Ruthotto. OT-Flow: Fast and accurate continuous normalizing flows via optimal transport. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 9223–9232, 2021.

George Papamakarios, Eric Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan. Normalizing flows for probabilistic modeling and inference. *Journal of Machine Learning Research*, 22(57):1–64, 2021.

Doanh Pham. Notes on numerical methods. URL https://people.math.wisc.edu/~htran24/note-DoanhPham.pdf.

Danilo Jimenez Rezende and Shakir Mohamed. Variational inference with normalizing flows. In Francis R. Bach and David M. Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6-11 July 2015*, volume 37 of *JMLR Workshop and Conference Proceedings*, pages 1530–1538. JMLR.org, 2015. URL http://proceedings.mlr.press/v37/rezende15.html.

Amirmojtaba Sabour, Sanja Fidler, and Karsten Kreis. Align your steps: Optimizing sampling schedules in diffusion models. In *Forty-first International Conference on Machine Learning, ICML 2024, Vienna, Austria, July 21-27, 2024*. OpenReview.net, 2024. URL https://openreview.net/forum?id=nBGBzV4It3.

Amir Sagiv. The wasserstein distances between pushed-forward measures with applications to uncertainty quantification. *Communications in Mathematical Sciences*, 2019. URL https://api.semanticscholar.org/CorpusID:119129452.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net, 2021. URL https://openreview.net/forum?id=PxTIG12RRHS.

E. G. Tabak and Cristina V. Turner. A family of nonparametric density estimation algorithms. *Communications on Pure and Applied Mathematics*, 66(2):145–164, February 2013. ISSN 0010-3640. doi: 10.1002/cpa.21423.

Esteban G. Tabak and Eric Vanden-Eijnden. Density estimation by dual ascent of the log-likelihood. *Communications in Mathematical Sciences*, 8(1):217–233, 2010. ISSN 1539-6746. doi: 10.4310/CMS.2010.v8.n1.a11.

Sven Wang and Youssef Marzouk. On minimax density estimation via measure transport, 2022. URL https://arxiv.org/abs/2207.10231.

Antoine Wehenkel and Gilles Louppe. Unconstrained monotonic neural networks. *Advances in neural information processing systems*, 32, 2019.

Christopher Williams, Andrew Campbell, Arnaud Doucet, and Saifuddin Syed. Score-optimal diffusion schedules. *CoRR*, abs/2412.07877, 2024. doi: 10.48550/ARXIV.2412.07877. URL https://doi.org/10.48550/arXiv.2412.07877.

Dmitry Yarotsky. Error bounds for approximations with deep relu networks. *Neural Networks*, 94:103–114, 2017. doi: 10.1016/J.NEUNET.2017.07.002. URL https://doi.org/10.1016/j.neunet.2017.07.002.

William P Ziemer. *Weakly differentiable functions: Sobolev spaces and functions of bounded variation*, volume 120. Springer Science & Business Media, 2012.

Appendix A. Proofs of main results

A.1. Preliminary results

We start by proving auxiliary results.

Lemma 11 *For $\theta \in [0,1]$ define the one-parameter family:*

$$\phi_\theta : [-1, \infty] \to \overline{\mathbb{R}}$$

$$x \mapsto \left| \frac{x}{1 + \theta x} \right|$$

where $\overline{\mathbb{R}}$ denotes the extended reals. Then, for any non-empty $S \subset (-1, \infty)$ such that $\sup S < \infty$ and $\inf S > -1$ we have:

$$\sup_{x \in S} \phi_\theta(x) = \max \left\{ \left| \frac{\sup S}{1 + \theta \sup S} \right|, \left| \frac{\inf S}{1 + \theta \inf S} \right| \right\} = \max \left\{ \frac{\sup S}{1 + \theta \sup S}, \frac{-\inf S}{1 + \theta \inf S} \right\}$$

Proof First, we can simplify

$$\phi_\theta(x) = \begin{cases} \frac{x}{1 + \theta x} & x \geq 0 \\ \frac{-x}{1 + \theta x} & x < 0 \end{cases}$$

and then, taking a derivative, we obtain

$$\frac{d}{dx} \phi_\theta(x) = \begin{cases} \frac{1}{(1 + \theta x)^2} & x > 0 \\ \frac{-1}{(1 + \theta x)^2} & x < 0 \,, \end{cases}$$

where, the derivative is ill-defined at 0; however $\phi_\theta(0) = 0$. We notice that for any θ the function decreases monotonically in $(-1, 0)$ and grows monotonically in $(0, \infty)$; hence for $S \subset (-1, \infty)$ we have

$$\sup_{x \in S} \phi_\theta(x) = \max \left\{ \frac{\sup S}{1 + \theta \sup S}, \frac{-\inf S}{1 + \theta \inf S} \right\} \,.$$

For the second part of the claim we want to prove that:

$$\max \left\{ \frac{\sup S}{1 + \theta \sup S}, \frac{-\inf S}{1 + \theta \inf S} \right\} = \left\{ \left| \frac{\sup S}{1 + \theta \sup S} \right|, \left| \frac{\inf S}{1 + \theta \inf S} \right| \right\} \,. \tag{25}$$

First, notice that for all $\theta \in [0,1]$ and any $x \in (-1, \infty)$ we have $1 + \theta x \geq 0$ thus:

$$1 + \theta \sup S \geq 0 \quad \text{and} \quad 1 + \theta \inf S \geq 0$$

for any $\theta \in [0,1]$ using that $-1 < \inf S \leq \sup S < \infty$. Now if $S \subset (-1, 0)$ then by the monotonicity of ϕ_θ we have:

$$\frac{-\sup S}{1 + \theta \sup S} \leq \frac{-\inf S}{1 + \theta \inf S}$$

and both terms are positive so equation (25) is verified. Simiarly, equation (25) can be verified when $S \subset (0, \infty)$. Finally, if $S \cap (0, \infty) \cap (-1, 0) \neq \emptyset$ then $\sup S > 0$ and $\inf S < 0$ and so equation 25 is trivially true. This finishes the proof. ∎

Lemma 12 *Let $(\sigma_i(s))_i$ denote the eigenvalues of $\nabla T(s)$, for $s \in \Omega$, recalling that by Assumption 1 they are real. Moreover, let $\sigma_{\max}(s) := \max_i \sigma_i(s)$ and $\sigma_{\min}(s) := \min_i \sigma_i(s)$. Then, the maps $\Omega \to \mathbb{R}$ given by*

$$s \mapsto \sigma_{\max}(s) \quad and \quad s \mapsto \sigma_{\min}(s)$$

are continuous. Moreover, there exists $c \in \mathbb{R}^+$ such that for all $s \in \Omega$:

$$\sigma_{\min}(s) \geq c > 0. \tag{26}$$

Proof For the first claim, notice that the map $s \mapsto \sigma_{\max}$ coincides with the map $s \mapsto \|\nabla T(s)\|_{\mathrm{op}}$. Now, by Assumption 1 we have that $T : \Omega \subset \mathbb{R}^d \to \Omega$ is in $C^1(\Omega; \Omega)$, meaning that for each $1 \leq i, j \leq d$ the maps $s \mapsto \partial_i \phi_j(s)$ are continuous. As these form the entries of the Jacobian, the map $s \mapsto \nabla T(s)$ is continuous in the operator norm, hence $s \mapsto \sigma_{\max}(s)$ is also continuous as a map $\Omega \to \mathbb{R}$. Similarly, notice that the map $s \mapsto \sigma_{\min}(s)$ coincides with the map $s \mapsto \left\|(\nabla T(s))^{-1}\right\|_{\mathrm{op}}^{-1}$. This can be viewed as a composition of the maps:

$$s \mapsto \nabla T(s) \mapsto (\nabla T(s))^{-1} \mapsto \left\|(\nabla T(s))^{-1}\right\|_{\mathrm{op}} \mapsto \left\|(\nabla T(s))^{-1}\right\|_{\mathrm{op}}^{-1}$$

Since the inversion operator is continuous in $\mathrm{GL}(d)$ and the spectrum of $\nabla T(s)$ contains positive numbers at each $s \in \Omega$, by assumption 1, it follows that the above composition is continuous finishing the proof of the first claim.

The second part is an immediate consequence of assumption 1. Namely, since $\sigma_i(s) > 0$ for all i we can take $c := \inf_{s \in \Omega} \sigma_{\min}(s)$ using the continuity of $\sigma_{\min}$ and the compactness of Ω. ∎

Lemma 13 *Let S be any set and $H, K : S \to \overline{\mathbb{R}}$ be arbitrary functions taking values in the extended real line. Then:*

$$\sup_{s \in S} \max\{H(s), \, K(s)\} = \max\{\sup_{s \in S} H(s), \, \sup_{s \in S} K(s)\} \tag{27}$$

Proof The proof is elementary. Start by considering the case where $\sup_{s \in S} H(s) = +\infty$ or $\sup_{s \in S} K(s) = +\infty$. Then, the right hand side is clearly infinite and the left hand side is bounded below by both $\sup_{s \in S} H(s)$ and $\sup_{s \in S} K(s)$ and so it is also infinite. Now if $\sup_{s \in S} H(s) = -\infty$ then for all $s \in S$ we have $H(s) = -\infty$ and so both the left and right hand side reduce to $\sup_{s \in S} K(s)$. Finally, if $\sup_{s \in S} H(s) = \sup_{s \in S} K(s) = -\infty$ then both sides are equal to $-\infty$. Now assume that H, K are bounded. Start by noticing that for any $s \in S$ we have:

$$H(s) \leq \sup_{s \in S} H(s)$$
$$K(s) \leq \sup_{s \in S} K(s)$$

and so, for any $s \in S$:

$$\max\{H(s), K(s)\} \leq \max\{\sup_{s \in S} H(s), \sup_{s \in S} K(s)\}$$

20

taking the supremum we conclude:

$$\sup_{s \in S} \max\{H(s), K(s)\} \leq \max\{\sup_{s \in S} H(s), \sup_{s \in S} K(s)\}$$

which is the first needed inequality. For the other inequality, start noticing that for any $s \in S$ we have:

$$H(s) \leq \max\{H(s), K(s)\}$$
$$K(s) \leq \max\{H(s), K(s)\}$$

Thus, taking suprema on both sides:

$$\sup_{s \in S} H(s) \leq \sup_{s \in S} \max\{H(s), K(s)\}$$
$$\sup_{s \in S} K(s) \leq \sup_{s \in S} \max\{H(s), K(s)\}$$

and so:

$$\max\{\sup_{s \in S} H(s), \sup_{s \in S} K(s)\} \leq \sup_{s \in S} \max\{H(s), K(s)\}$$

which is the second needed inequality. This concludes the proof. ∎

A.2. Proof of theorem 2

Proposition 14 *Fix $\Omega \subset \mathbb{R}^d$ compact, convex and let $v : \Omega \to \mathbb{R}^d$ be a C^1 vector field. Then the Lipschitz constant of v, denoted by $Lip(v)$, is the smallest number such that for all $x, y \in \Omega$ we have*

$$\|v(x) - v(y)\| \leq Lip(v)\|x - y\|. \tag{28}$$

Recalling the spectral norm of v is defined as

$$\|v\|_{\mathrm{op}} = \sup_{\|x\|=1} \|v(x)\| \tag{29}$$

and denoting the Jacobian of v at x by $\nabla v(x)$, we have the equality

$$Lip(v) = \sup_{x \in \Omega} \|\nabla v(x)\|_{\mathrm{op}}. \tag{30}$$

The proof is an immediate consequence of the following two lemmas:

Lemma 15 *For v as above, we have*

$$\sup_{x \in \Omega} \|\nabla v(x)\|_{\mathrm{op}} \leq Lip(v). \tag{31}$$

21

Proof For the proof of this statement, fix any unit vector v. Now if $x \in \Omega^\circ$ we have:

$$\|\nabla v(x)v\| = \lim_{h \to 0} \frac{\|v(x + hv) - v(x)\|}{h}$$
$$\leq \text{Lip}(v) \lim_{h \to 0} \frac{1}{h}\|x + hv - x\|$$
$$= \text{Lip}(v)$$

Maximizing over all unit vectors v we get

$$\|\nabla v(x)\|_{\text{op}} \leq \text{Lip}(v). \tag{32}$$

Now since v is a C^1 map $\Omega \to \Omega$ we have that $x \mapsto \|\nabla v(x)\|_{\text{op}}$ is a continuous map $\Omega \to \mathbb{R}$. Then, equation (32) can be extended to the boundary $\partial \Omega$ by continuity. We conclude by taking the supremum over Ω. ∎

We also have a converse result:

Lemma 16 *For v as above we have:*

$$Lip(v) \leq \sup_{x \in \Omega} \|\nabla v(x)\|_{\text{op}} \tag{33}$$

Proof To see this, take any two $x, y \in \Omega$. By the convexity of Ω and the Mean Value Theorem, there is a $\theta \in [0, 1]$ such that:

$$v(x) - v(y) = \nabla v(x + \theta(y - x)) \cdot (x - y)$$

Taking norms we have:

$$\|v(x) - v(y)\| \leq \|\nabla v(x + \theta(y - x))\|_{\text{op}} \|x - y\|$$
$$\leq \sup_{z \in \Omega} \|\nabla v(z)\|_{\text{op}} \|x - y\|$$

So we have a Lipschitz condition with constant $\sup_{z \in \Omega} \|\nabla v(z)\|_{\text{op}} < \infty$, where finiteness is guaranteed by the fact that v is C^1 and Ω is compact. By the minimality of the Lipschitz constant, it must be that $\text{Lip}(v) \leq \sup_{z \in \Omega} \|\nabla v(z)\|$. ∎

Theorem *For notational convenience, define $f(s) := \max_i \sigma_i(s) - 1$, $g(s) := \min_i \sigma_i(s) - 1$ and let $\Omega_{[0,1]} := \Omega \times [0, 1]$. Then, $\Lambda[\tau]$ can be written as:*

$$\Lambda[\tau] = \sup_{(s,t) \in \Omega_{[0,1]}} |\dot{\tau}(t)| \max \left\{ \left| \frac{f(s)}{1 + \tau(t)f(s)} \right|, \left| \frac{g(s)}{1 + \tau(t)g(s)} \right| \right\} \tag{34}$$

Proof First, for $t \in [0, 1]$ define the sets

$$\Omega_t = \{X(s, t) \, : \, s \in \Omega\} \subset \Omega.$$

22

and recall the definition of the dynamics

$$v_\tau(\cdot, t) : \Omega_{\tau(t)} \to \Omega$$
$$v_\tau(\cdot, t) : s \mapsto \dot{\tau}(t) \, (T - \mathrm{id}) \circ X_\tau^{-1}(\cdot, t).$$

In Marzouk et al. (2025), theorem 3.4., it is established that for T a C^1 map such that $\nabla T(s)$ has a positive eigen-spectrum for all $s \in \Omega$ (see assumption 1) the flow X_τ is a C^1 diffeomorphism onto its image $\Omega_{\tau(t)}$. Using the chain rule together with the functoriality of the Jacobian we can compute:

$$(\nabla_s v_\tau)(s, t) = \dot{\tau}(t) \, \left[\nabla_s T \left(X_\tau^{-1}(s, t) \right) - I_d \right] \cdot \nabla_s X_\tau^{-1}(s, t)$$
$$= \dot{\tau}(t) \, \left[\nabla_s T \left(X_\tau^{-1}(s, t) \right) - I_d \right] \cdot (\nabla_s X_\tau)^{-1} \left(X_\tau^{-1}(s, t), t \right)$$

Thus, we conclude

$$(\nabla_s v_\tau)(X_\tau(s, t), t) = \dot{\tau}(t) \, \left[\nabla_s T(s) - I_d \right] \cdot (\nabla_s X_\tau)^{-1}(s, t).$$

Now since $s \in \Omega \iff X_\tau(s, t) \in \Omega_{\tau(t)}$ is must be that

$$\sup_{s \in \Omega_{\tau(t)}} \|v_\tau(s, t)\|_{\mathrm{op}} = \sup_{s \in \Omega} \|v_\tau(X_\tau(s, t), t)\|_{\mathrm{op}}$$

and we have, by the definition of $\Lambda[\tau]$:

$$\Lambda[\tau] = \sup_{t \in [0,1]} \sup_{s \in \Omega_{\tau(t)}} \|\nabla_s v_\tau(s, t)\|_{\mathrm{op}}$$
$$= \sup_{t \in [0,1]} \sup_{s \in \Omega} \|\nabla_s v_\tau(X_\tau(s, t), t)\|_{\mathrm{op}}$$
$$= \sup_{t \in [0,1]} \sup_{s \in \Omega} \left\| \dot{\tau}(t) \, \left[\nabla_s T(s) - I_d \right] \cdot (\nabla_s X_\tau)^{-1} \right\|_{\mathrm{op}}.$$

To finish the proof, we will re-write the term inside the operator norm of the last line, above.

Due to assumption 1 we have that $\nabla T(s)$ is diagonalizable by orthogonal transformations, for all $s \in \Omega$. Suppressing now the dependence on $s \in \Omega$, there exist an orthogonal matrix U and a diagonal matrix D such that:

$$\nabla T = U \, D \, U^T \tag{35}$$

Notice that we can then write:

$$\nabla T - I_d = U \, (D - I_d) \, U^T \tag{36}$$

with I_d the identity matrix in $\mathbb{R}^d$. Moreover, we can write:

$$[I_d + \tau(t)(\nabla T - I_d)]^{-1} = U \, [I_d + \tau(t)(D - I_d)]^{-1} \, U^T \tag{37}$$

Since both matrices are simultaneously diagonalizable, it is easy to compute the spectral norm of the product at a fixed $(s, t) \in \Omega \times [0, 1]$, leading to:

$$\|\nabla_s v_\tau(X_\tau(s, t), t)\|_{\mathrm{op}} = |\dot{\tau}(t)| \max_{1 \leq i \leq d} \left| \frac{[D(s) - I_d]_{ii}}{[I_d + \tau(t)(D(s) - I_d)]_{ii}} \right|$$
$$= |\dot{\tau}(t)| \max_{1 \leq i \leq d} \left| \frac{\sigma_i(s) - 1}{1 + \tau(t)(\sigma_i(s) - 1)} \right|$$

writing $\sigma_i(s)$ for the ith eigenvalue of $\nabla T(s)$ at $s \in \Omega$. To proceed, define the one parameter family of extended real valued mappings:

$$\phi_\theta : [-1, \infty] \to \overline{\mathbb{R}}$$

$$\phi_\theta : x \mapsto \left| \frac{x}{1 + \theta x} \right|$$

with $\overline{\mathbb{R}}$ the extended reals. Fixing $s \in \Omega$ and defining $S := \{\sigma_i(s) - 1 \ : \ 1 \leq i \leq d\}$ we have:

$$\|\nabla_s v_\tau(X_\tau(s, t), t)\|_{\mathrm{op}} = |\dot{\tau}(t)| \max_{1 \leq i \leq d} \phi_{\tau(t)}(\sigma_i(s) - 1)$$

$$= |\dot{\tau}(t)| \sup_{x \in S} \phi_{\tau(t)}(x)$$

Due to lemma 12 we have that $\sigma_i(s) - 1 > -1$ for all $1 \leq i \leq d$, $s \in \Omega$ and also $0 \leq \tau(t) \leq 1$ for all $t \in [0, 1]$ due to the monotonicity of τ. Thus, we can apply lemma 11 to conclude:

$$\|(\nabla f)(X_\tau(s, t))\|_{\mathrm{op}} = |\dot{\tau}(t)| \max \left\{ \left| \frac{\sup S}{1 + \tau(t) \sup S} \right|, \left| \frac{\inf S}{1 + \tau(t) \inf S} \right| \right\}$$

$$= |\dot{\tau}(t)| \max \left\{ \left| \frac{f(s)}{1 + \tau(t) f(s)} \right|, \left| \frac{g(s)}{1 + \tau(t) g(s)} \right| \right\} \tag{38}$$

and in the last step we have used the definitions $f = \max_i \sigma_i - 1$ and $g = \min_i \sigma_i - 1$. This finishes the proof. ∎

A.3. Proof of Theorem 3

In this section we prove Theorem 3 from the main text. First, we establish some preliminary results.

Lemma 17 *Let Ω be compact subset of $\mathbb{R}^d$ and $F, G : \Omega \to \mathbb{R}$ be continuous functions. Then, as $p \to \infty$ we have the following limit:*

$$\left[\int_\Omega F^{2p}(s) + G^{2p}(s) \, ds \right]^{\frac{1}{2p}} \to \max \left\{ \|F\|_{L^\infty(\Omega)}, \|G\|_{L^\infty(\Omega)} \right\} \tag{39}$$

Proof Start by noting that since F, G are continuous and Ω is compact then $F, G \in L^p(\Omega)$ for $p \in [1, \infty]$ and so all expressions in the statement are finite. First, we prove an upper bound:

$$\int_\Omega F^{2p}(s) + G^{2p}(s) \, ds \leq$$

$$\|F\|_{L^\infty}^{2p} \mathrm{Leb}(\Omega) + \|G\|_{L^\infty}^{2p} \mathrm{Leb}(\Omega) =$$

$$\mathrm{Leb}(\Omega) \cdot \left(\|F\|_{L^\infty}^{2p} + \|G\|_{L^\infty}^{2p} \right)$$

where $\mathrm{Leb}(\Omega)$ is the Lebesgue measure of Ω. Raising this to the power $\frac{1}{2p}$ and using that for positive reals $x, y > 0$ we have $\left(|x|^{2p} + |y|^{2p} \right)^{1/(2p)} \to \max\{x, y\}$ as $p \to \infty$, we obtain

$$\left[\mathrm{Leb}(\Omega) \cdot \left(\|F\|_{L^\infty}^{2p} + \|G\|_{L^\infty}^{2p} \right) \right]^{\frac{1}{2p}} \to \max\{ \|F\|_{L^\infty}, \|G\|_{L^\infty} \}$$

and thus

$$\lim_{p\to\infty} \left(\int_\Omega F^{2p}(s) + G^{2p}(s)\, ds \right)^{\frac{1}{2p}} \leq \max\{\|F\|_{L^\infty}, \|G\|_{L^\infty}\}. \tag{40}$$

To prove a lower bound, recall that by continuity of F, G and compactness of Ω for any $\epsilon > 0$ there are $\delta_1, \delta_2 > 0$ such that:

$$\text{Leb}\left(\{x : F(x) \geq \|F\|_{L^\infty} - \epsilon\}\right) = \delta_1$$
$$\text{Leb}\left(\{x : G(x) \geq \|G\|_{L^\infty} - \epsilon\}\right) = \delta_2$$

Now we can estimate:

$$\int_\Omega F^{2p}(s) + G^{2p}(s)\, ds \geq$$

$$\int_{\{F \geq \|F\|_{L^\infty} - \epsilon\}} F^{2p}(s)\, ds + \int_{\{G \geq \|G\|_{L^\infty} - \epsilon\}} G^{2p}(s)\, ds \geq$$

$$\delta_1 \left(\|F\|_{L^\infty} - \epsilon\right)^{2p} + \delta_2 \left(\|G\|_{L^\infty} - \epsilon\right)^{2p} =$$

$$\min\{\delta_1, \delta_2\} \left[(\|F\|_{L^\infty} - \epsilon)^{2p} + (\|G\|_{L^\infty} - \epsilon)^{2p} \right]$$

Now raising the last line to the power $\frac{1}{2p}$ and taking the limit as $p \to \infty$ we obtain:

$$\left\{ \min\{\delta_1, \delta_2\} \left[(\|F\|_{L^\infty} - \epsilon)^{2p} + (\|G\|_{L^\infty} - \epsilon)^{2p} \right] \right\}^{\frac{1}{2p}} \to \max\{\|F\|_{L^\infty} - \epsilon, \|G\|_{L^\infty} - \epsilon\}$$

Taking the limit $\epsilon \searrow 0$ we obtain:

$$\lim_{p\to\infty} \left(\int_\Omega F^{2p}(s) + G^{2p}(s)\, ds \right)^{\frac{1}{2p}} \geq \max\{\|F\|_{L^\infty}, \|G\|_{L^\infty}\} \tag{41}$$

Combining (40) and (41) we obtain the desired result.

∎

Another lemma that will be useful:

Lemma 18 *Let Ω be compact subset of $\mathbb{R}^d$ and $F, G : \Omega \to \mathbb{R}$ be continuous functions. Set $s_p := \left[\int_\Omega F^{2p}(s) + G^{2p}(s)\, ds \right]^{\frac{1}{2p}}$. Then, there are constants $c_{p,q} \geq 0$ for each $p \leq q$, with $p, q \in \{1, 2, \dots\} \cup \{\infty\}$, depending only on $\text{Leb}(\Omega)$, such that:*

$$s_p \leq c_{p,q}\, s_q \tag{42}$$

moreover, these constants are uniformly bounded in p and q by $\sqrt{2\, Leb(\Omega)}$.

Proof Notice that:

$$s_p = \left[\|F\|_{L^{2p}}^{2p} + \|G\|_{L^{2p}}^{2p} \right]^{\frac{1}{2p}}$$

Now by the Hölder inequality there are constants $C_{r,p}, D_{r,p} \geq 0$ such that:

$$s_p = \left[\|F\|_{L^{2p}}^{2p} + \|G\|_{L^{2p}}^{2p} \right]^{\frac{1}{2p}}$$

$$\leq \left[C_{2p,2q}^{2p} \|F\|_{L^{2q}}^{2p} + C_{2p,2q}^{2p} \|G\|_{L^{2q}}^{2p} \right]^{\frac{1}{2p}}$$

$$\leq C_{2p,2q}\, D_{2p,2q} \left[\|F\|_{L^{2q}}^{2q} + \|G\|_{L^{2q}}^{2q} \right]^{\frac{1}{2q}}$$

$$= c_{p,q}\, s_q$$

where we have taken $c_{p,q} := C_{2p,2q}\, D_{2p,2q}$. In fact, we can take [5]

$$C_{2p,2q} = \text{Leb}\,(\Omega)^{\frac{1}{2p} - \frac{1}{2q}}$$
$$D_{2p,2q} = 2^{\frac{1}{2p} - \frac{1}{2q}}$$

and so

$$c_{p,q} = 2^{\frac{1}{2p} - \frac{1}{2q}}\, \text{Leb}\,(\Omega)^{\frac{1}{2p} - \frac{1}{2q}}\,. \tag{43}$$

Thus, the constants $c_{p,q}$ are uniformly bounded in p and q by $\sqrt{2\,\text{Leb}(\Omega)}$. ∎

Recall, now, the spaces introduced in equations (3) – (5) of the main text:

$$\mathcal{T}_2 = \left\{\tau \in H^1([0,1]) \,:\, \tau(0) = 0,\, \tau(1) = 1\right\}$$
$$\mathcal{T}_2^b = \mathcal{T}_2 \cap \{\tau : 0 \leq \tau \leq 1\}$$

Remark 19 *In the interest of well-definedness recall that, by the Sobolev embedding theorem, elements of H^1 can be identified with continuous functions on $[0,1]$ and as such the constraints $t : \tau(t) \geq 0$ as well as $\tau(0) = 0,\, \tau(1) = 1$ are well posed. Moreover, the set $\{\tau \in C^0\,([0,1]) : 0 \leq \tau \leq 1 \text{ and } \tau(0) = 0,\, \tau(1) = 1\}$ is a closed subset of $C^0\,([0,1])$ and thus will be a closed subset of $H^1([0,1])$.*

Now we prove a lemma that will be used multiple times:

Lemma 20 *Recall the notation $f(s) = \max_i \sigma_i(s) - 1$ and $g(s) = \min_i \sigma_i(s) - 1$ for $(\sigma_i(s))_i$ the eigenvalues of $\nabla T(s)$ at $s \in \Omega$. Moreover, we write $f^* = \sup_{s \in \Omega} f(s)$ and $g_* = \inf_{s \in \Omega} g(s)$. If Assumptions 1 and 2 are satisfied we have:*

$$-1 < g_* < 0 < f^* < \infty, \tag{44}$$

Moreover, there are non-negative reals:

$$M_f = M_f(f, g, \Omega) \quad \text{and} \quad M_g = M_g(g, f, \Omega)$$
$$m_f = m_f(f, g, \Omega) \quad \text{and} \quad m_g = m_g(g, f, \Omega)$$

such that for all $\tau \in \mathcal{T}_2^b$ and all $p \in \mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$ we have:

$$m_f \leq \left[\int_0^1 \int_\Omega \left(\frac{f(s)}{1 + \tau(t)\, f(s)}\right)^{2p} ds\, dt\right]^{\frac{1}{2p}} \leq M_f$$

$$m_g \leq \left[\int_0^1 \int_\Omega \left(\frac{g(s)}{1 + \tau(t)\, g(s)}\right)^{2p} ds\, dt\right]^{\frac{1}{2p}} \leq M_g$$

---

5. For completeness, we conduct the computation here: For a finite measure space (X, Σ, μ) and $p, q \in [1, \infty]$ consider a pair $q > p$ and some function $f \in L^q(X, \Sigma, \mu)$. Then, by a Holder inequality and denoting the identity function by **1** we have:

$$\|f \cdot \mathbf{1}\|_{L^p} \leq \left(\|f^p\|_{L^{\frac{q}{p}}} \|\mathbf{1}\|_{L^{\frac{q}{q-p}}}\right)^{\frac{1}{p}} = \|f\|_{L^q}\, \mu(X)^{\frac{q-p}{qp}} = \|f\|_{L^q}\, \mu(X)^{\frac{1}{p} - \frac{1}{q}}$$

So above we can take X to be either Ω equipped with the Borel sets and the Lebesgue measure or the discrete space $\{1, 2\}$ equipped with its power set and the counting measure.

and moreover, $m_f \cdot m_g > 0$ for all $p \in \mathbb{N}_\infty$. In fact, the bounds hold pointwise in t in the sense that for any $\tau \in \mathcal{T}_2^b$ and any $p \in \mathbb{N}_\infty$ we have:

$$m_f \leq \left[\int_\Omega \left(\frac{f(s)}{1 + \tau(t) f(s)} \right)^{2p} ds \right]^{\frac{1}{2p}} \leq M_f \quad \text{for all} \quad t \in [0, 1]$$

$$m_g \leq \left[\int_\Omega \left(\frac{g(s)}{1 + \tau(t) g(s)} \right)^{2p} ds \right]^{\frac{1}{2p}} \leq M_g \quad \text{for all} \quad t \in [0, 1]$$

Proof As usual, for $s \in \Omega$ denote $\sigma_{\max}(s) = \max_i \sigma_i(s)$ and $\sigma_{\min}(s) = \min_i \sigma_i(s)$. Firstly, recall that by lemma 12 there is a $c > 0$ such that:

$$\sigma_{\max}(s) \geq \sigma_{\min}(s) \geq c > 0 \quad \text{for all} \quad s \in \Omega$$

Moreover, again by lemma 12 we have that $\sigma_{\max}, \sigma_{\min} : \Omega \to \mathbb{R}$ are continous and recalling that Ω is compact there is a $C < \infty$ such that:

$$\sigma_{\min}(s) \leq \sigma_{\max}(s) \leq C < \infty \quad \text{for all} \quad s \in \Omega$$

Therefore, we have shown the first part of the claim:

$$-1 < c - 1 \leq g_* \leq f^* \leq C - 1 < \infty$$

Secondly, for any $\tau \in \mathcal{T}_2^b$ we have $1 + \tau(t)f(s) \geq c > 0$ and $1 + \tau(t)g(s) \geq c > 0$ for all $(s, t) \in \Omega \times [0, 1]$ and so we can estimate:

$$\sup_{s \in \Omega} \left| \frac{f(s)}{1 + \tau(t) f(s)} \right| = \frac{\sup_{s \in \Omega} |f(s)|}{\inf_{(s,t) \in \Omega \times [0,1]} |1 + \tau(t) f(s)|}$$

$$\leq \frac{\max\{|C - 1|, |c - 1|\}}{c}$$

Thus, for any $p \in \mathbb{N}_\infty$ we can take:

$$M_f(f, g, \Omega) := \frac{\max\{|C - 1|, |c - 1|\}}{c} \max\{\text{Leb}(\Omega), 1\} < \infty$$

Replacing f by g and arguing exactly as in the above line, we have proven the first part of the claim.

Thirdly, notice that by assumption 2 there must exist an $s_0 \in \Omega$ and an index $1 \leq k \leq d$ such that:

$$\lambda_k(s_0) \neq 1$$

As a result, either $\sigma_{\max}(s_0) \neq 0$ or $\sigma_{\min}(s_0) \neq 0$ and so $f(s_0) \neq 0$ or $g(s_0) \neq 0$. Assuming the latter and using the continuity of g there is an $\epsilon > 0$ and an open set U, both depending on g, such that $|g(s)| \geq \epsilon$ for all $s \in U$. Now for finite p we can use Jensen's inequality to estimate:

$$\int_\Omega \left(\frac{g(s)}{1 + \tau(t) g(s)} \right)^{2p} ds \geq \text{Leb}(\Omega)^{1-2p} \left(\int_\Omega \left| \frac{g(s)}{1 + \tau(t) g(s)} \right| ds \right)^{2p}$$

$$\geq \text{Leb}(\Omega)^{1-2p} \left(\frac{1}{\sup_{(s,t) \in \Omega \times [0,1]} |1 + \tau(t) g(s)|} \right)^{2p} \left(\int_U |g(s)| ds \right)^{2p}$$

$$\geq \frac{\epsilon^{2p}}{C^{2p}} \frac{\text{Leb}(U)}{\text{Leb}(\Omega)^{2p-1}}$$

Thus, we have:

$$\left[\int_0^1 \int_\Omega \left(\frac{g(s)}{1+\tau(t)\,g(s)}\right)^{2p} ds\,dt\right]^{\frac{1}{2p}} \geq \frac{\epsilon}{C} \frac{\mathrm{Leb}\,(U)^{\frac{1}{2p}}}{\mathrm{Leb}\,(\Omega)^{\frac{2p-1}{2p}}}$$

$$\geq \frac{\epsilon}{C} \frac{\min\{1\,,\,\mathrm{Leb}\,(U)\}}{\max\{1\,,\,\mathrm{Leb}\,(\Omega)\}}$$

$$=: m_g(g,f,\Omega) > 0$$

The argument is similar for $p = \infty$ and the same lower bound holds. Finally, if $f(s_0) \neq 0$ we can use the same argument to obtain $m_f(f,g,\Omega) > 0$. Since we cannot have both $f(s_0) = 0$ and $g(s_0) = 0$ simultaneously, we have proven the claim.

∎

Using lemma 20 can now make the following definitions:

Definition 21 *Recall that $\Omega_{[0,1]} = \Omega \times [0,1]$ and define the operators:*

$$\Psi^f : \mathcal{T}_2^b \to L^\infty\left(\Omega_{[0,1]}\right) \qquad \text{and} \qquad \Psi^g : \mathcal{T}_2^b \to L^\infty\left(\Omega_{[0,1]}\right)$$

$$\Psi^f[\tau](s,t) = \frac{f(s)}{1+\tau(t)f(s)} \quad \text{and} \quad \Psi^g[\tau](s,t) = \frac{g(s)}{1+\tau(t)g(s)}$$

Now for $p \in \mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$ introduce the functionals:

$$\Lambda_p : \mathcal{T}_2 \to \mathbb{R}$$

$$\Lambda_p[\tau] = \begin{cases} \left(\left\|\dot{\tau}\,\Psi^f[\tau]\right\|_{L^{2p}\left(\Omega_{[0,1]}\right)}^{2p} + \left\|\dot{\tau}\,\Psi^g[\tau]\right\|_{L^{2p}\left(\Omega_{[0,1]}\right)}^{2p}\right)^{\frac{1}{2p}} & \text{if } \tau \in \mathcal{T}_2^b \\ +\infty & \text{if } \tau \in \mathcal{T}_2 \setminus \mathcal{T}_2^b \end{cases}$$

Moreover, if $\dot{\tau} \notin L^{2p}\left(\Omega_{[0,1]}\right)$ for some $p \in \mathbb{N}$ we set $\Lambda_p[\tau] = +\infty$.

Remark 22 *For finite $p \in \mathbb{N}$ one may wish to use a more explicit formula for Λ_p:*

$$\Lambda_p[\tau] = \left(\int_0^1 \int_\Omega \left(\frac{\dot{\tau}(t)\,f(s)}{1+\tau(t)\,f(s)}\right)^{2p} ds\,dt + \int_0^1 \int_\Omega \left(\frac{\dot{\tau}(t)\,g(s)}{1+\tau(t)\,g(s)}\right)^{2p} ds\,dt\right)^{\frac{1}{2p}}$$

Remark 23 *Notice that the restriction of Λ_∞ to $\mathcal{T}_\infty$ coincides with Λ, the objective of problem (10). This will play an important role in the proof of the main theorem.*

Now we prove a crucial property of the functionals Λ_p:

Lemma 24 *For each $p \in \mathbb{N}$ and each sequence $\gamma_n \in \mathcal{T}_2$ converging to some $\gamma_\infty \in \mathcal{T}_2$ in the weak-H^1 topology we have:*

$$\Lambda_p[\gamma_\infty] \leq \liminf_{n\to\infty} \Lambda_p[\gamma_n]$$

Remark 25 *Put succinctly, this lemma says that for each $p \in \mathbb{N}$ the functional Λ_p is weakly lower semi-continuous in the affine space $\mathcal{T}_2$ equipped with the H^1 norm.*

Proof We note that our proof technique is based on (Garroni et al., 2001, Lemma 2.4 and Proposition 2.6). The essence of this augment is to show that $\dot{\gamma}_n \rightharpoonup \dot{\gamma}_\infty$ in $L^{2p}([0,1])$. This fact together with the well-behavedness of the operators Ψ^f, Ψ^g and the weak lower semi-continuity of the L^{2p} norm will allow us to conclude the proof.

Fix a p and assume without loss of generality that $\liminf_{n\to\infty} \Lambda_p[\gamma_n] = M < \infty$ else verifying the lower semi-continuity of Λ_p is trivial. Moreover, we can, again without loss of generality, assume that $\lim_{n\to\infty} \Lambda_p[\gamma_n] = \liminf_{n\to\infty} \Lambda_p[\gamma_n]$, by passing to a subsequence that attains the limit inferior. In fact, this allows us to assume that $\gamma_n \in \mathcal{T}_2^b$ for all $n \in \mathbb{N}$, again without loss of generality. By lemma 20 we obtain reals m_f and m_g, noth both zero, such that for all $\tau \in \mathcal{T}_2^b$ and any $p \in \mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$ we have:

$$m_f \leq \left[\int_\Omega \left(\Psi^f[\tau](s,t) \right)^{2p} ds \right]^{\frac{1}{2p}} \quad \text{for all } t \in [0,1]$$

$$m_g \leq \left[\int_\Omega \left(\Psi^g[\tau](s,t) \right)^{2p} ds \right]^{\frac{1}{2p}} \text{for all } t \in [0,1]$$

We use this to estimate:

$$\left(m_f^{2p} + m_g^{2p} \right)^{\frac{1}{2p}} \|\dot{\gamma}_n\|_{L^{2p}([0,1])} \leq \left(\left\| \dot{\gamma}_n \, \Psi^f[\gamma_n] \right\|_{L^{2p}\left(\Omega_{[0,1]}\right)}^{2p} + \left\| \dot{\gamma}_n \, \Psi^g[\gamma_n] \right\|_{L^{2p}\left(\Omega_{[0,1]}\right)}^{2p} \right)^{\frac{1}{2p}} = \Lambda_p[\gamma_n]$$

Noting that $\Lambda_p[\gamma_n] \to M < \infty$ we have that $\dot{\gamma}_n$ is bounded in $L^{2p}([0,1])$ for some $p \in \mathbb{N}$. Since $L^{2p}([0,1])$ is a reflexive Banach space, theorem 7 in James (1964) allows us to extract a subsequence that converges to some $\alpha \in L^{2p}([0,1])$ in the weak-L^{2p} topology.

We now wish to show that $\alpha = \dot{\gamma}_\infty$. To see this, note that since $\dot{\gamma}_n$ is bounded in $L^{2p}([0,1])$ for $p \geq 1$ then, by the Poincaré inequality, it is also bounded in H^1. Now recall that by the Rellich–Kondrachov theorem in $\mathbb{R}$ (theorem 9.16 in Brezis (2011)) we have that H^1 compactly embeds into $C^0([0,1])$. This implies then (example 8.9. in Dal Maso (1993)) that γ_n converges to γ_∞ in $C^0([0,1])$ also establishing that $\gamma_\infty \in \mathcal{T}_2^b$. We can now identify α as the weak derivative $\dot{\gamma}_\infty$ as follows: for any test function $\phi \in C_c^\infty([0,1])$ compute

$$\int_0^1 \alpha(t)\, \phi(t)\, dt = \lim_{i\to\infty} \int_0^1 \dot{\gamma}_{n_i}(t)\, \phi(t)\, dt = -\lim_{i\to\infty} \int_0^1 \gamma_{n_i}(t)\, \dot{\phi}(t)\, dt = -\int_0^1 \gamma_\infty(t)\, \dot{\phi}(t)\, dt$$

where in the first equality we used the subsequential weak-L^{2p} convergence $\dot{\gamma}_{n_i} \rightharpoonup \alpha$ and in the last equality we used the convergence of $\gamma_n \to \gamma_\infty$ in $C^0([0,1])$. By definition, we have that $\alpha = \dot{\gamma}_\infty$ in $L^{2p}([0,1])$. Recall that our original goal was to show:

$$\Lambda_p[\gamma_\infty] \leq \liminf_{n\to\infty} \Lambda_p[\gamma_n] =: M$$

Clearly, the subsequence τ_{n_i} satisfies $\Lambda_p[\gamma_{n_i}] \to M$ and if we verify that:

$$\Lambda_p[\gamma_\infty] \leq \liminf_{i\to\infty} \Lambda_p[\gamma_{n_i}]$$

we are done. For this reason and by abuse of notation, we denote the subsequence $(\gamma_{n_i})_i$ by $(\gamma_n)_n$ for the rest of this proof.

Now we will establish $\Psi^f[\gamma_n]\,\dot\gamma_n \rightharpoonup \Psi^f[\gamma_\infty]\,\dot\gamma_\infty$ in $L^{2p}\left(\Omega_{[0,1]}\right)$ using the notation $\Omega_{[0,1]} = \Omega \times [0,1]$. To see this, notice first that $\Psi^f[\gamma_n] \to \Psi^f[\gamma_\infty]$ in $L^\infty\left(\Omega_{[0,1]}\right)$ since:

$$\sup_{(s,t)\in\Omega_{[0,1]}} \left| \frac{f(s)}{1+f(s)\gamma_n(t)} - \frac{f(s)}{1+f(s)\gamma_\infty(t)} \right| =$$

$$\sup_{(s,t)\in\Omega_{[0,1]}} \left| \frac{f(s)\left(\gamma_n(t)-\gamma_\infty(t)\right)}{\left(1+f(s)\gamma_n(t)\right)\left(1+f(s)\gamma_\infty(t)\right)} \right| \leq$$

$$\sup_{(s,t)\in\Omega_{[0,1]}} \left| \frac{f(s)}{\left(1+f(s)\gamma_n(t)\right)\left(1+f(s)\gamma_\infty(t)\right)} \right| \sup_{t\in[0,1]} |\gamma_n(t)-\gamma_\infty(t)| \leq$$

$$\left| \frac{\sup_s f(s)}{\left(1+\inf_s f(s)\right)^2} \right| \sup_{t\in[0,1]} |\gamma_n(t)-\gamma_\infty(t)| \to 0$$

using that $\gamma_n \to \gamma_\infty$ in L^∞ as well as that all $\tau \in \mathcal{T}_2^b$ are uniformly bounded above by 1 and f is continuous, bounded away from -1 in Ω by lemma 20. Now for the Hölder conjugate q of $2p$ and a test function $\psi \in L^q(\Omega_{[0,1]})$ we have:

$$\int_{\Omega_{[0,1]}} \psi\left(\Psi^f[\gamma_n]\,\dot\gamma_n - \Psi^f[\gamma_\infty]\,\dot\gamma_\infty\right) =$$

$$\int_{\Omega_{[0,1]}} \psi\left(\Psi^f[\gamma_n]-\Psi^f[\gamma_\infty]\right)\dot\gamma_n + \int_{\Omega_{[0,1]}} \psi\,\Psi^f[\gamma_\infty]\left(\dot\gamma_n-\dot\gamma_\infty\right)$$

Since $\Psi^f[\gamma_n] \in L^\infty(\Omega_{[0,1]})$ (see lemma 20) the second term vanishes by the weak convergence of $\dot\gamma_n$ to $\dot\gamma_\infty$ in[6] $L^{2p}([0,1])$. For the first term, we argue as follows:

$$\left| \int_{\Omega_{[0,1]}} (\psi\dot\gamma_n)\left(\Psi^f[\gamma_n]-\Psi^f[\gamma_\infty]\right) \right| \leq \sup_{(s,t)\in\Omega_{[0,1]}} \left| \Psi^f[\gamma_n]-\Psi^f[\gamma_\infty] \right| \left(\int_{\Omega_{[0,1]}} |\psi\dot\gamma_n|\,dt\right) \to 0$$

using that $\Psi^f[\gamma_n] \to \Psi^f[\gamma_\infty]$ in $L^\infty\left(\Omega_{[0,1]}\right)$ and bounding the second integral uniformly over n by a Hölder inequality, noting that $\psi \in L^q\left(\Omega_{[0,1]}\right), \dot\gamma_n \in L^{2p}([0,1])$, with $\frac{1}{q}+\frac{1}{2p}=1$, and we have shown that the $\dot\gamma_n$ are bounded in $L^{2p}([0,1])$. Thus, we have proven that:

$$\Psi^f[\gamma_n]\,\dot\gamma_n \rightharpoonup \Psi^f[\gamma_\infty]\,\dot\gamma_\infty \text{ in } L^{2p}\left(\Omega_{[0,1]}\right)$$

6. More precisely, since $\gamma_n \rightharpoonup \gamma_\infty$ in $L^{2p}([0,1])$ but $\psi \in L^q(\Omega_{[0,1]})$, with $\frac{1}{2p}+\frac{1}{q}=1$, one has to use the Fubini theorem (theorem 3.4.4. in Bogachev and Ruas (2007)) to see that the map:

$$[0,1] \to \mathbb{R}$$
$$t \mapsto \int_\Omega \psi(s,t)\,ds$$

is in $L^q([0,1])$ and so the weak convergence of $\dot\gamma_n$ to $\dot\gamma_\infty$ in $L^{2p}([0,1])$ can be used to conclude. Indeed, by the Fubini theorem and the Jensen inequality we have:

$$\int_0^1 \left(\int_\Omega \psi(s,t)\,ds\right)^q dt \leq \int_0^1 \int_\Omega (\psi(s,t))^q\,ds\,dt = \int_{\Omega_{[0,1]}} (\psi(s,t))^q\,ds\otimes dt < \infty$$

Here, we use the standard fact[7] that the map $L^{2p} \to \mathbb{R}$ given by $h \mapsto \|h\|_{L^{2p}(\Omega_{[0,1]})}$ is lower-semi-continuous in the weak-L^{2p} topology to conclude that:

$$\gamma \mapsto \left\| \Psi^f[\gamma]\,\dot\gamma \right\|_{L^{2p}(\Omega_{[0,1]})}$$

is lower semi-continuous in weak-$L^{2p}\left(\Omega_{[0,1]}\right)$. The exact same argument establishes that the map:

$$\gamma \mapsto \left\| \Psi^g[\gamma]\,\dot\gamma \right\|_{L^{2p}(\Omega_{[0,1]})}$$

is also lower semi-continuous in weak-$L^{2p}\left(\Omega_{[0,1]}\right)$.

To conclude, start by recalling that the limit inferior commutes with continuous, non-decreasing functions. We can thus write:

$$
\begin{aligned}
\liminf_{n\to\infty} \Lambda_p[\gamma_n] &= \liminf_{n\to\infty} \left(\left\| \Psi^f[\gamma_n]\,\dot\gamma_n \right\|^{2p}_{L^{2p}(\Omega_{[0,1]})} + \left\| \Psi^g[\gamma_n]\,\dot\gamma_n \right\|^{2p}_{L^{2p}(\Omega_{[0,1]})} \right)^{\frac{1}{2p}} \\
&= \left(\liminf_{n\to\infty} \left(\left\| \Psi^f[\gamma_n]\,\dot\gamma_n \right\|^{2p}_{L^{2p}(\Omega_{[0,1]})} + \left\| \Psi^g[\gamma_n]\,\dot\gamma_n \right\|^{2p}_{L^{2p}(\Omega_{[0,1]})} \right) \right)^{\frac{1}{2p}} \\
&\geq \left(\liminf_{n\to\infty} \left\| \Psi^f[\gamma_n]\,\dot\gamma_n \right\|^{2p}_{L^{2p}(\Omega_{[0,1]})} + \liminf_{n\to\infty} \left\| \Psi^g[\gamma_n]\,\dot\gamma_n \right\|^{2p}_{L^{2p}(\Omega_{[0,1]})} \right)^{\frac{1}{2p}} \\
&\geq \left(\left\| \Psi^f[\gamma_\infty]\,\dot\gamma_\infty \right\|^{2p}_{L^{2p}(\Omega_{[0,1]})} + \left\| \Psi^g[\gamma_\infty]\,\dot\gamma_\infty \right\|^{2p}_{L^{2p}(\Omega_{[0,1]})} \right)^{\frac{1}{2p}} \\
&= \Lambda_p[\gamma_\infty]
\end{aligned}
$$

and in the third line we are using the boundedness of $\Psi^f[\gamma_n]\,\dot\gamma_n$ and $\Psi^g[\gamma_n]\,\dot\gamma_n$ in $L^{2p}\left(\Omega_{[0,1]}\right)$ that is evident since both terms lower bound $\Lambda_p[\gamma_n]$ which was assumed to be finite, for each $n \in \mathbb{N}$, and converging to $M < \infty$. This completes the proof. ∎

Proposition 26 *The family* $(\Lambda_p)_{p\in\mathbb{N}}$ *is equi-coercive on* $\mathcal{T}_2$ *equipped with the weak-H^1 topology.*

Proof To see this, we first leverage Proposition 7.7. in Dal Maso (1993) allowing us to prove equicoercivity by finding a lower semi-continuous coercive functional $\Phi : \mathcal{T}_2 \to \mathbb{R}$ such that $\Phi \leq \Lambda_p$ for all p. Recall that the functional Φ is coercive if for any $\alpha \in \mathbb{R}$ the sublevel set $\{\tau \in \mathcal{T}_2 : \Phi[\tau] \leq \alpha\}$ has compact closure, see Definitions 1.10, 1.12. and remark 1.11. in Dal Maso

---

7. To see this, one idea presented in Pham is to assume $f_n \rightharpoonup f$ in L^p and use the test function:

$$\psi = \frac{\text{sgn}(f)}{\|f\|_{L^p}^{p/q}}\,|f|^{p/q}$$

where p and q are Hölder conjugate. Note that we have $\|\psi\|_{L^q} = 1$ and $\int f\,\psi = \|f\|_{L^p}$. Now using the weak convergence together with the Hölder inequality, we obtain:

$$\|f\|_{L^p} = \int f\,\psi = \lim_{n\to\infty} \int f_n\,\psi \leq \liminf_{n\to\infty} \|f_n\|_{L^p}\,\|\psi\|_{L^q} = \liminf_{n\to\infty} \|f_n\|_{L^p}$$

(1993). Since $\mathcal{T}_2$ is homeomorphic to a linear subspace of the reflexive Hilbert space $H_0^1([0,1])$ it follows by Theorem 7 in James (1964) that the closure of H^1 bounded sets is weak-H^1 compact. Thus, to show coercivity of such a Φ in weak-H^1 is suffices to show that for any $\alpha \in \mathbb{R}$ the sublevel set $\{\tau \in \mathcal{T}_2 : \Phi[\tau] \leq \alpha\}$ is bounded in H^1.

Start by using that by lemma 18 the number $\sqrt{2 \operatorname{Leb}(\Omega)}$ is a uniform bound for the constants $c_{p,q}$. Define:

$$\Phi = \frac{\Lambda_1}{\sqrt{2 \operatorname{Leb}(\Omega)}}$$

Then, by Lemma 18 we have that

$$\Phi[\tau] \leq \Lambda_p[\tau] \quad \text{for all} \quad p \in \mathbb{N} \tag{45}$$

for all $\tau \in \mathcal{T}_2^b$ whereas for $\tau \in \mathcal{T}_2$ this is inquality reduces to the triviality $+\infty \leq +\infty$. Now Φ is lower semi-continuous by lemma 24 and so it suffices to show it is coercive. To see this, use lemma 20 to obtain reals m_f and m_g, noth both zero, such that for all $\tau \in \mathcal{T}_2^b$ and any $p \in \mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$ we have:

$$m_f \leq \left[\int_\Omega \left(\Psi^f[\tau](s,t) \right)^{2p} ds \right]^{\frac{1}{2p}} \quad \text{for all } t \in [0,1]$$

$$m_g \leq \left[\int_\Omega \left(\Psi^g[\tau](s,t) \right)^{2p} ds \right]^{\frac{1}{2p}} \text{for all } t \in [0,1].$$

We use this to estimate:

$$\left(m_f^2 + m_g^2 \right)^{\frac{1}{2}} \|\dot{\tau}_n\|_{L^2([0,1])} \leq \left(\left\| \dot{\tau}_n \, \Psi^f[\tau_n] \right\|_{L^2\left(\Omega_{[0,1]}\right)}^2 + \left\| \dot{\tau}_n \, \Psi^g[\tau_n] \right\|_{L^2\left(\Omega_{[0,1]}\right)}^2 \right)^{\frac{1}{2}} = \Lambda_1[\tau_n].$$

Combining this with the Poincaré inequality we have:

$$\|\tau\|_{H^1([0,1])} \leq C_{[0,1]} \|\dot{\tau}\|_{L_2([0,1])} \leq \frac{C_{[0,1]} \sqrt{2 \operatorname{Leb}(\Omega)}}{\left(m_f^2 + m_g^2 \right)^{\frac{1}{2}}} \Phi[\tau] \tag{46}$$

and the constant $C_{[0,1]}$ depends only on the domain $[0,1]$. In particular, this implies that $\Phi[\tau_i] \to \infty$ whenever $\tau_i \to \infty$ in H^1 and as a result the sublevel sets $\{\Phi \leq \alpha\}$ are bounded in $H^1([0,1])$, for any $\alpha \in \mathbb{R}$. Combining this with the discussion at the beginning of the proof we can conclude. ∎

This proof leads to a useful lemma:

Lemma 27 *Assume that for some sequence $\gamma_p \in \mathcal{T}_2$ there is a constant $0 \leq C < \infty$ such that:*

$$\Lambda_p[\gamma_p] \leq C \quad \text{for all} \quad p \in \mathbb{N}$$

then, γ_p is bounded in $H^1([0,1])$.

Proof First, notice that $\gamma_p \in \mathcal{T}_2^b$ for all $p \in \mathbb{N}$ since $\Lambda_p[\gamma_p] < \infty$. Then, apply equations (45) and (46) with $C_{[0,1]}$ the Poincaré constant for the interval $[0,1]$:

$$\|\gamma_p\|_{H^1([0,1])} \leq C_{[0,1]} \|\dot{\gamma}_p\|_{L^2([0,1])} \leq \frac{C_{[0,1]} \sqrt{2 \operatorname{Leb}(\Omega)}}{\left(m_f^2 + m_g^2 \right)^{\frac{1}{2}}} \Lambda_p[\gamma_p] \leq \frac{C_{[0,1]} \sqrt{2 \operatorname{Leb}(\Omega)}}{\left(m_f^2 + m_g^2 \right)^{\frac{1}{2}}} C < \infty$$

for all $p \in \mathbb{N}$. ∎

Proposition 28 *There are constants $b_p \in \mathbb{R}$ such that $\frac{1}{b_p}\Lambda_p \to \Lambda_\infty$ in a Γ-sense, where the Γ-convergence is taken in the space $\mathcal{T}_2$ equipped with the weak H^1 topology.*

Proof Start by recalling that by conditions (e)-(f) of (Dal Maso, 1993, Proposition 8.1) it suffices to show that for any sequence $\gamma_p \to \gamma_\infty$ in $\mathcal{T}_2$ we have:

$$\liminf_{p \to \infty} \Lambda_p[\gamma_p] \geq \Lambda_\infty[\gamma_\infty] \tag{47}$$

as well as that for any $\gamma_\infty \in \mathcal{T}_2$ there exists a sequence $\gamma_p \to \gamma_\infty$ such that:

$$\limsup_{p \to \infty} \Lambda_p[\gamma_p] \leq \Lambda_\infty[\gamma_\infty] \tag{48}$$

We start by proving (47). Fix a sequence $\gamma_p \in \mathcal{T}_2$ converging to some $\gamma_\infty \in \mathcal{T}_2$, i.e. $\gamma_p \rightharpoonup \gamma_\infty$ in H^1. Without loss of generality we can assume that $\liminf_p \Lambda_p[\gamma_p] < \infty$ and moreover that the sequence γ_p attains the infimum in the sense that:

$$\lim_{p \to \infty} \Lambda_p[\gamma_p] = \liminf_p \Lambda_p[\gamma_p] < \infty \tag{49}$$

Moreover, we can $\{\gamma_p\}_p \subset \mathcal{T}_2^b$ by removing the finitely many $\gamma \in \mathcal{T}_2^b \setminus \mathcal{T}_2$ that evaluate to $+\infty$ under Λ_p.

Our first task is to show that $\gamma_\infty \in \mathcal{T}_2^b$. To do this, we appeal to lemma 27 yielding that $\{\gamma_p\}_p$ is a bounded set in $H^1([0,1])$. By Dal Maso (1993, example 8.9.) and the Rellich–Kondrachov theorem (Brezis, 2011, theorem 9.16.) in $\mathbb{R}$ it follows that γ_n converges to γ_∞ strongly in $C^0([0,1])$. It then follows that $\gamma_\infty \in \mathcal{T}_2^b$ since $\gamma_p \in \mathcal{T}_2^b$ for all p.

Now we can apply Lemma 18 to obtain constants $a_p := 2^{\frac{1}{p(p+1)}}$ such that for any $\gamma \in \mathcal{T}_2^b$ we have:

$$\frac{1}{a_p}\Lambda_p[\gamma] \leq \Lambda_{p+1}[\gamma] \tag{50}$$

moreover, notice that $b_p = \prod_{k=p}^{\infty} a_k$ is bounded by $2^{\sum_{i=1}^{\infty} \frac{1}{k^2}}$ uniformly in p. This, together with (50) yields:

$$\frac{1}{b_p}\Lambda_p[\gamma] \leq \frac{1}{b_q}\Lambda_q[\gamma] \leq \Lambda_\infty[\gamma] \tag{51}$$

for any $\gamma \in \mathcal{T}_2^b$ and all $p \leq q$. Using equation (51) together with Lemma 24 we have, for each p:

$$\frac{1}{b_p}\Lambda_p[\gamma_\infty] \leq \liminf_{n \to \infty} \frac{1}{b_p}\Lambda_p[\gamma_n] \leq \liminf_{n \to \infty} \frac{1}{b_n}\Lambda_n[\gamma_n] \tag{52}$$

Finally, using the point-wise convergence established[8] in lemma 17 together with $b_p \to 1$ we obtain:

$$\Lambda_\infty[\gamma_\infty] \leq \liminf_{n \to \infty} \frac{1}{b_n}\Lambda_n[\gamma_n]$$

The second condition (48) is trivially satisfied by taking $\gamma_p = \gamma_\infty$ for all p and using the point-wise convergence of Lemma 17 together with $\frac{1}{b_p} \to 1$. ∎

Now we can proceed with the proof of Theorem 3. We recall the statement first:

---

8. This is where the fact that $\gamma_\infty \in \mathcal{T}_2^b$ is needed.

Theorem *Define $\mathcal{T}_2 = \big\{\tau \in W^{1,2}([0,1]) : \tau(0) = 0, \tau(1) = 1\big\}$ where $W^{1,2}([0,1])$ is the Sobolev space of functions with square integrable first derivative on the unit interval. For each positive integer p, consider the optimization problem:*

$$\inf_{\tau \in \mathcal{T}_2} \int_0^1 \lambda_p(\tau(t), \dot{\tau}(t))\, dt \qquad (53)$$

with:

$$\lambda_p(\tau(t), \dot{\tau}(t)) := \dot{\tau}(t)^{2p} \int_\Omega \frac{f(s)^{2p}}{(1 + \tau(t)f(s))^{2p}}\, ds + \dot{\tau}(t)^{2p} \int_\Omega \frac{g(s)^{2p}}{(1 + \tau(t)g(s))^{2p}}\, ds$$

If τ_p is optimal for (53) then any subsequential L^2 limit $\tau_{p_j} \overset{L^2}{\to} \tau_\infty$ as $j \to \infty$ is optimal for (10).

Proof Start by noticing that:

$$\arg\min_{\tau \in \mathcal{T}_\infty} \int_0^1 \lambda_p(\tau(t), \dot{\tau}(t))\, dt = \arg\min_{\tau \in \mathcal{T}_\infty} \left(\int_0^1 \lambda_p(\tau(t), \dot{\tau}(t))\, dt \right)^{\frac{1}{2p}}$$

$$= \arg\min_{\tau \in \mathcal{T}_\infty} \Lambda_p[\tau]$$

and the equality holds in a sense of sets. This follows from the definition of Λ_p as well as that $\lambda_p \geq 0$ and $x \mapsto x^{1/2p}$ is increasing in $\mathbb{R}^+$. Notice, moreover, that Λ_∞ restricted to $\mathcal{T}_\infty \subset \mathcal{T}_2$ coincides with the objective Λ of problem (53).

Now we apply the *Fundemental Theorem of Γ-convergence*, theorem 2.10. in Braides (2006) appearing also as corollary 7.20. in Dal Maso (1993): by proposition 28 there are constants $b_p \to 1$ such that $(1/b_p)\, \Lambda_p \to \Lambda_\infty$ in a Γ-sense and by proposition 26 it follows that the $(1/b_p)\Lambda_p$ is equi-coercive in the weak H^1 topology, thus it follows that if minimizers $\tau_p \in \arg\min(1/b_p)\,\Lambda_p$ form a pre-compact set in $\mathcal{T}_2^b$ then any weak H^1 subsequential limit of τ_p will converge to a minimizer of $\inf \Lambda_\infty$. Note that since $b_p > 0$ the τ_p will also be minimizers of Λ_p.

To establish pre-compactness in weak H^1 we show H^1 boundedness (theorem 7 in James (1964)). Due to the Γ-convergence and equicoercivity of the $(1/b_p)\,\Lambda_p$ we can apply theorem 7.8. in Dal Maso (1993) to obtain:

$$\lim_{p \to \infty} \inf_{\tau \in \mathcal{T}_2^b} \frac{1}{b_p}\Lambda_p[\tau] = \inf_{\tau \in \mathcal{T}_2^b} \Lambda_\infty[\tau] < \infty$$

so there is a constant $C \geq 0$ such that the sequence of minimizers τ_p satisfies $\Lambda_p[\tau_p] \leq C < \infty$ for all p. Applying lemma 27 we have that the set $\{\tau_p\}_p$ is bounded in $H^1([0,1])$ and so by James (1964, theorem 7) it is pre-compact in the weak H^1 topology.

Thus, we have established that subsequential weak-H^1 limits of the τ_p will be minimizers of Λ_∞. As a final step, note that by corollary 8.8. in Dal Maso (1993) (see also example 8.9.) and the Rellich–Kondrachov theorem (Brezis, 2011, theorem 9.16.), inside any H^1 norm bounded set weak H^1 convergence is equivalent to strong convergence in L^2. We just showed that the set $\{\tau_p\}_p$ is norm bounded, so we can conclude. ∎

A.4. Proof of Theorem 4

In this section, we solve the Lagrangian problem (14) by employing the direct method of the Calculus of Variations. First, some definitions:

Definition 29 *The function*

$$\lambda_p : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

$$\lambda_p : (x, v) \mapsto \int_\Omega \left(\frac{vf(s)}{1 + xf(s)} \right)^{2p} ds + \int_\Omega \left(\frac{vg(s)}{1 + xg(s)} \right)^{2p} ds$$

will be called the Lagrangian of problem (14).

Now recall the spaces introduced in equations (3) – (5) of the main text. For $p \in \mathbb{N}$ we have:

$$\mathcal{T}_{2p} := \left\{ \tau \in W^{1,2p}([0,1]) \ : \ \tau(0) = 0 \,, \ \tau(1) = 1 \right\}$$
$$\mathcal{T}_{2p}^b := \left\{ \tau \in \mathcal{T}_{2p} \ : \ 0 \leq \tau \leq 1 \right\}$$
$$\mathcal{T}_\infty := \left\{ \tau \in C^1([0,1]) \ : \ \tau(0) = 0 \,, \ \tau(1) = 1 \,, \ \dot{\tau} \geq 0 \right\}$$

where $W^{1,2p}([0,1])$ is the Sobolev space of functions with square integrable first derivative on the unit interval.

Remark 30 *Notice that*

$$\mathcal{T}_\infty \subset \mathcal{T}_{2p}^b \subset \mathcal{T}_{2p}$$

for any $p \in \mathbb{N}$.

First, we establish that to solve problem (14) it is sufficient to optimize over the space $\mathcal{T}_{2p}^b$. For brevity, we denote problem (14) by:

$$\inf_{\tau \in \mathcal{T}_2} \int \lambda_p(\tau, \dot{\tau}) := \inf_{\tau \in \mathcal{T}_2} \int_0^1 \lambda_p(\tau(t), \dot{\tau}(t)) \, dt \tag{54}$$

We now have:

Lemma 31 *For each $p \in \mathbb{N}$ we have:*

$$\inf_{\tau \in \mathcal{T}_{2p}^b} \int \lambda_p(\tau, \dot{\tau}) = \inf_{\tau \in \mathcal{T}_2} \int \lambda_p(\tau, \dot{\tau})$$

Proof Since for each $p \in \mathbb{N}$ we have $\mathcal{T}_{2p}^b \subset \mathcal{T}_2$ it suffices to prove that:

$$\inf_{\tau \in \mathcal{T}_{2p}^b} \int \lambda_p(\tau, \dot{\tau}) \leq \inf_{\tau \in \mathcal{T}_2} \int \lambda_p(\tau, \dot{\tau})$$

Now if $p > 1$ recall that for any $\tau \in \mathcal{T}_2 \setminus \mathcal{T}_{2p}$ we have

$$\int \lambda_p(\tau, \dot{\tau}) = \infty \,.$$

35

thus, one has

$$\inf_{\tau \in \mathcal{T}_{2p}} \int \lambda_p(\tau, \dot{\tau}) \leq \inf_{\tau \in \mathcal{T}_2} \int \lambda_p(\tau, \dot{\tau}).$$

It suffices to show that for any $\tau \in \mathcal{T}_{2p}$ there exists a $\tau_b \in \mathcal{T}_{2p}^b$ such that:

$$\int \lambda_p(\tau_b, \dot{\tau}_b) \leq \int \lambda_p(\tau, \dot{\tau})$$

To that end, take any path $\tau \in \mathcal{T}_\infty$ and consider the modified path $\tau_b \in \mathcal{T}_{2p}^b$ defined by:

$$\tau_b(t) := \min\{\max\{0, \tau(t)\}, 1\} \tag{55}$$

Due to theorem 2.1.11. in Ziemer (2012) we have that $\tau_b \in W^{1,q}([0,1])$. Indeed, the maps $x \mapsto \max\{0, x\}$ and $x \mapsto \{x, 1\}$ are Lipschitz and τ_b is in $L^p([0,1])$, for any positive integer p, since the Sobolev embedding guarantees $\tau \in C^0([0,1])$. We now show that τ_b attains a lower value for our objective. Notice that for any t such that $\tau(t) \neq \tau_b(t)$ we have $\tau(t) < 0$ or $\tau(t) > 1$. By continuity, there exists $\epsilon > 0$ such that for any $s \in (t - \epsilon, t + \epsilon)$ we have $\tau(s) < 0$ or $\tau(t) > 1$. This implies that for all $s \in (t - \epsilon, t + \epsilon)$ we have $\tau_b(s) = 0$ or $\tau_b(s) = 1$ and hence $\dot{\tau}_b(s) = 0$ for a.e. $s \in (t - \epsilon, t + \epsilon)$. This last part follows from the fact that functions in $W^{1,2p}([0,1])$, for $p \geq 1$, satisfy the Fundamental Theorem of Calculus and hence for any $0 < \delta \leq \epsilon$ we have:

$$\int_{t-\delta}^{t+\delta} \dot{\tau}_b(s)\, ds = \tau_b(t + \delta) - \tau_b(t - \delta) = 0$$

Taking $\delta \searrow 0$ and using the Lebesgue Differentiation Theorem we obtain $\dot{\tau}_b(s) = 0$ for a.e. $s \in (t - \epsilon, t + \epsilon)$. Finally, notice that $\lambda_p(\tau(t), \dot{\tau}(t)) = 0$ whenever $\dot{\tau}(t) = 0$ so for a set $\mathcal{N}$ of measure zero we have $\{t : \tau_b(t) \neq \tau(t)\} \subset \{t : \lambda_p(\tau(t), \dot{\tau}(t)) = 0\} \cup \mathcal{N}$ and so we can estimate:

$$\int_0^1 \lambda_p(\tau_b(t), \dot{\tau}_b(t))\, dt = \int_{\{t:\tau_b(t)\neq\tau(t)\}} \lambda_p(\tau_b(t), \dot{\tau}_b(t))\, dt + \int_{\{t:\tau_b(t)=\tau(t)\}} \lambda_p(\tau_b(t), \dot{\tau}_b(t))\, dt$$

$$= \int_{\{t:\tau_b(t)=\tau(t)\}} \lambda_p(\tau(t), \dot{\tau}(t))\, dt$$

$$\leq \int_0^1 \lambda_p(\tau(t), \dot{\tau}(t))\, dt$$

and the last step follows since $\lambda_p \geq 0$. This completes the proof. ∎

The above proof establishes that to solve problem (53) we can instead solve:

$$\inf_{\tau \in \mathcal{T}_{2p}^b} \int \lambda_p(\tau, \dot{\tau}) := \inf_{\tau \in \mathcal{T}_{2p}^b} \int_0^1 \lambda_p(\tau(t), \dot{\tau}(t))\, dt \tag{56}$$

Our goal is prove that problem (56) has a minimizer in $C^2([0,1])$ that satisfies the strong Euler–Lagrange equations. To achieve that goal, we will need to inroduce an auxiliary Lagrangian:

Definition 32 *For $p \in \mathbb{N}$ and some $\epsilon > 0$ we define the auxiliary Lagrangian:*

$$\lambda_p^\epsilon : [0,1] \times \mathbb{R} \to \mathbb{R}$$

$$\lambda_p^\epsilon : (x, v) \mapsto \int_\Omega \left(\frac{vf(s)}{1 + xf(s)}\right)^{2p} ds + \int_\Omega \left(\frac{vg(s)}{1 + xg(s)}\right)^{2p} ds + \epsilon\, v^2$$

Remark 33 *The reason we cannot work directly with the original Lagrangian is the following: our goal, here, is to identitfy a minimizer of problem* (53). *One standard approach would be to write the stong Euler–Lagrange equations for the Lagrangian λ_p, solve it, and claim that this solution is optimal for* (53). *This, however, crucially relies on the convexity of the map $(x, v) \mapsto \lambda_p(x, v)$ which does not hold here.*[9]

A second approach would be to take a minimizer τ_p, which can be shown to exist, and characterize it by showing it satisfies the strong Euler–Lagrange equations. Although insightful, this approach cannot be rigorously justified since it assumes the minimizer it at least of class $C^2([0, 1])$. To show this, one would employ standard regularity theory, requiring:

$$\frac{\partial^2 \lambda_p}{\partial v^2}(x, v) > 0 \text{ for all } (x, v) \in [0, 1] \times \mathbb{R}$$

Unfortunately, this condition is violated here but only mildly. Indeed, we have:

$$\frac{\partial^2 \lambda_p}{\partial v^2} \geq 0 \text{ and } \frac{\partial^2 \lambda_p}{\partial v^2} = 0 \text{ if and only if } v = 0$$

Thus, it will be possible to show that the minimizer τ_p of (53) *satisfies the strong Euler–Lagrange (sEL) equations by constructing a perturbed ϵ-Lagrangian with smooth minimizers τ_p^ϵ, satisfying the sEL equations, and recovering τ_p as an L^2 subsequential limit of τ_p^ϵ as $\epsilon \to 0$. This will, again, involve the theory of Γ-convergence.*

We now study the family of problems:

$$\inf_{\tau \in \mathcal{T}_2} \int \lambda_p^\epsilon(\tau, \dot{\tau}) := \inf_{\tau \in \mathcal{T}_2} \int_0^1 \lambda_p^\epsilon(\tau(t), \dot{\tau}(t)) \, dt \tag{57}$$

First, we establish existence of a minimizer:

Lemma 34 *The problem* (57) *has a minimizer in $\mathcal{T}_{2p}^b$.*

Proof We now apply theorem 4.1. in Dacorogna (2007) with a slight modification to account for the additional constraint $\{\tau : 0 \leq \tau \leq 1\}$ – indeed, the theory developed in Dacorogna (2007) applies directly to the problem:

$$\inf_{\tau \in \mathcal{T}_{2p}} \int_0^1 \lambda_p^\epsilon(\tau(t), \dot{\tau}(t)) \, dt$$

However, the proof of theorem 4.1. carries through for $\mathcal{T}_{2p}^b$ identically, since the set $\{\tau : 0 \leq \tau \leq 1\}$ is closed inside $W^{1,2p}([0, 1])$, due to the Sobolev Embedding Theorem. To apply theorem 4.1. in Dacorogna (2007) we need to, first, verify the convexity of the map:

$$v \mapsto \lambda_p^\epsilon(x, v)$$

for all $x \in [0, 1]$. This is immediate since the map in question is just:

$$v \mapsto v^{2p} \left[\int_\Omega \left(\frac{1}{1 + xf(s)} \right)^{2p} ds + \int_\Omega \left(\frac{1}{1 + xg(s)} \right)^{2p} ds \right] + \epsilon \, v^2$$

---

9. In fact, this non-convexity is crucial to the structure of the problem: physically, it is exactly this interaction between the position and velocity variables that allow us to reduce the "cost" at position x by traveling at low speeds v.

and the term in parentheses is finite for all x. The second and final condition to be verified is the coercivity of $(x, v) \mapsto \lambda_p^\epsilon(x, v)$, i.e., we need to find $\alpha_1 > 0$ and $\alpha_2, \alpha_3 \in \mathbb{R}$ such that:

$$\lambda_p^\epsilon(x, v) \geq \alpha_1 |v|^{2p} + \alpha_2 |x|^q + \alpha_3$$

for positive integer q such that $2p > q$. To do so, notice first that for any $\epsilon > 0$ we have:

$$\lambda_p^\epsilon \geq \lambda_p \geq 0$$

Secondly, we use lemma 20 to obtain constants $m_f, m_g \geq 0$ such that $m_f \cdot m_g \neq 0$ satisfying:

$$\int_\Omega \left(\frac{1}{1 + xf(s)}\right)^{2p} ds \geq m_f^{2p}$$

$$\int_\Omega \left(\frac{1}{1 + xg(s)}\right)^{2p} ds \geq m_g^{2p}$$

We can now estimate:

$$\lambda_p^\epsilon(x, v) \geq v^{2p} \min\left\{m_f^{2p}, m_g^{2p}\right\} .$$

Thus, we can take $\alpha_1 = \min\left\{m_f^{2p}, m_g^{2p}\right\}$, $\alpha_2 = 0$ and $\alpha_3 = 0$ to conclude the proof. ∎

The next step is to establish that minimizers of (57) satisfy the strong Euler–Lagrange equations.

Lemma 35 *Any minimizer of problem* (57) *satisfies the strong Euler–Lagrange equations.*

Proof Here, we apply theorem 4.12. in Dacorogna (2007), adapted to the case of a Lagrangian λ_p^ϵ with a bounded first variable (see the discussion in Lemma 34). To claim that a minimizer τ_ϵ of the problem (57) satisfies the Euler–Lagrange equation it suffices to have that the minimizer is at least $C^2([0, 1])$ as well as that the Lagrangian λ_p^ϵ is $C^2([0, 1] \times \mathbb{R})$ and satisfies the following regularity conditions: there are functions $\alpha_1 \in L^1([0, 1])$ and $\alpha_2 \in L^{2p/(2p-1)}([0, 1])$ and a constant $\beta \in \mathbb{R}$ such that for every $(x, v) \in [0, 1] \times \mathbb{R}$ we have:

$$\left|\lambda_p^\epsilon(x, v)\right|, \left|\partial_x \lambda_p^\epsilon(x, v)\right| \leq \alpha_1(x) + \beta|v|^{2p}$$
$$\left|\partial_v \lambda_p^\epsilon(x, v)\right| \leq \alpha_2(x) + \beta|v|^{2p-1}$$

With regards to the regularity of the minimizer we have:

$$\frac{\partial^2 \lambda_p^\epsilon}{\partial v^2}(x, v) \geq \epsilon > 0 \text{ for all } (x, v) \in [0, 1] \times \mathbb{R}$$

which by theorem 4.36. in Dacorogna (2007) guarantees that $\tau_\epsilon \in C^\infty([0, 1])$.

Now to verufy the regularity of the we appeal to lemma 20 to obtain constants $M_f, M_g \geq 0$ such that:

$$\int_\Omega \left(\frac{f(s)}{1 + xf(s)}\right)^{2p} ds \leq M_f^{2p}$$

$$\int_\Omega \left(\frac{g(s)}{1 + xg(s)}\right)^{2p} ds \leq M_g^{2p}$$

Setting $C = \max\left\{M_f^{2p}, M_g^{2p}\right\}$ and recalling that for each $q \leq p$ there is a constant $A_{q,p} > 0$ such that $|v|^q \leq A_{q,p} |v|^p$ for all $v \in \mathbb{R}$, we have:

$$|\lambda_p^\epsilon(x,v)| = v^{2p}\left(\int_\Omega \left(\frac{f(s)}{1+xf(s)}\right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1+xg(s)}\right)^{2p} ds\right) + \epsilon\, v^2$$

$$\leq \left(2^{2p}\, C^{2p} + \epsilon\, A_{2,2p}\right) v^{2p}$$

$$\left|\partial_x \lambda_p^\epsilon(x,v)\right| = 2p\, v^{2p}\left(\int_\Omega \left(\frac{f(s)}{1+xf(s)}\right)^{2p+1} ds + \int_\Omega \left(\frac{g(s)}{1+xg(s)}\right)^{2p+1} ds\right) + \epsilon\, v^2$$

$$\leq (2\,C)^{2p+1}\, v^{2p}$$

$$\left|\partial_v \lambda_p^\epsilon(x,v)\right| = 2p\,|v|^{2p-1}\left(\int_\Omega \left(\frac{f(s)}{1+xf(s)}\right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1+xg(s)}\right)^{2p} ds\right) + 2\epsilon\, v$$

$$\leq \left(2^{2p}\, C^{2p} + 2\,\epsilon\, A_{1,2p-1}\right) |v|^{2p-1}$$

so taking:

$$\beta = \max\left\{2^{2p}\, C^{2p} + \epsilon\, A_{2,2p}\,,\; (2\,C)^{2p+1}\,,\; 2^{2p}\, C^{2p} + 2\,\epsilon\, A_{1,2p-1}\right\}$$

we have that the conditions are satisfied. Finally, since the functions f and g are continuous on the compact set Ω and for each $s \in \Omega$ and $x \in [0,1]$ the maps $x \mapsto \frac{f(s)}{1+xf(s)}$ and $x \mapsto \frac{g(s)}{1+xg(s)}$ are smooth, we can apply the Leibniz Rule to conclude that the Lagrangian λ_p^ϵ is smooth on $[0,1] \times \mathbb{R}$. This completes the proof. ∎

We can now write the strong Euler–Lagrange equations for the Lagrangian λ_p^ϵ:

Lemma 36 *For each $p \in \mathbb{N}$ define:*

$$K_p : [0,1] \to \mathbb{R}$$

$$K_p(x) \mapsto \lambda_p(x,1) = \int_\Omega \left(\frac{f(s)}{1+xf(s)}\right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1+xg(s)}\right)^{2p} ds$$

For each $\epsilon > 0$ the strong Euler–Lagrange equations for the Lagrangian λ_p^ϵ

$$\frac{d}{dt}\left[\frac{\partial \lambda_p^\epsilon}{\partial v}\left(\tau(t), \dot\tau(t)\right)\right] - \frac{\partial \lambda_p^\epsilon}{\partial x}\left(\tau(t), \dot\tau_p(t)\right) = 0. \tag{58}$$

can be re-written as

$$\ddot\tau(t) = \frac{\dot\tau^{2p}(t)\, K_{p+1}(\tau(t))}{\dot\tau^{2p-2}(t)\, K_p(\tau(t)) + \frac{\epsilon}{p(2p-1)}}. \tag{59}$$

Moreover, the strong Euler–Lagrange equations for λ_p are obtained by setting $\epsilon = 0$ in (59).

Proof For notational convenience, we supress the p and ϵ notation on λ_p^ϵ, for the rest of this proof. Now, suppoose τ satisfies the strong Euler–Lagrange equations (58) for λ. We have

$$\frac{d}{dt}\left(\frac{\partial \lambda_p}{\partial v}\left(\tau_p(t), \dot\tau_p(t)\right)\right) - \frac{\partial \lambda_p}{\partial x}\left(\tau_p(t), \dot\tau_p(t)\right) = 0$$

with boundary conditions $\tau_p(0) = 0$ and $\tau_p(1) = 1$. This is equivalent to:

$$\ddot{\tau}_p(t)\frac{\partial^2 \lambda_p}{\partial v^2}(\tau_p(t),\dot{\tau}_p(t)) + \dot{\tau}_p(t)\frac{\partial^2 \lambda_p}{\partial v\,\partial x}(\tau_p(t),\dot{\tau}_p(t)) = \frac{\partial \lambda_p}{\partial x}(\tau_p(t),\dot{\tau}_p(t)) \tag{60}$$

Now to make computations less cumbersome, notice the decomposition:

$$\lambda(x,v) = \lambda_f(x,v) + \lambda_g(x,v) + \epsilon\, v^2$$

$$\lambda_f(x,v) := \int_\Omega \left(\frac{vf(s)}{1+xf(s)}\right)^{2p} ds$$

$$\lambda_g(x,v) := \int_\Omega \left(\frac{vg(s)}{1+xg(s)}\right)^{2p} ds$$

We conduct computations on λ_f since those for λ_g are identical. Now, once again, the continuity of $f:\Omega \to \mathbb{R}$ on the compact set Ω and the smoothness of the maps $x \mapsto \frac{f(s)}{1+xf(s)}$ and $x \mapsto \frac{g(s)}{1+xg(s)}$ for $x \in [0,1]$ allows us to differentiate under the integral sign to obtain

$$\frac{\partial \lambda_f}{\partial v}(x,v) = \int_\Omega \frac{2p\, v^{2p-1} f(s)^{2p}}{(1+x\,f(s))^{2p}}\, ds$$

$$\frac{\partial \lambda_f}{\partial x}(x,v) = \int_\Omega \frac{-2p\, v^{2p} f(s)^{2p+1}}{(1+x\,f(s))^{2p+1}}\, ds\,.$$

Differentiating once more we have

$$\frac{\partial^2 \lambda_f}{\partial v^2}(x,v) = \int_\Omega \frac{2p(2p-1)\, v^{2p-2} f(s)^{2p}}{(1+x\,f(s))^{2p}}\, ds$$

$$\frac{\partial^2 \lambda_f}{\partial x\,\partial v}(x,v) = \int_\Omega \frac{-(2p)^2\, v^{2p-1} f(s)^{2p+1}}{(1+x\,f(s))^{2p+1}}\, ds\,,$$

thus, combining with the corresponding expressions for λ_g we can re-write (60) as:

$$\ddot{\tau}(t)\int_\Omega \frac{2p(2p-1)\,\dot{\tau}(t)^{2p-2} f(s)^{2p}}{(1+\tau(t)f(s))^{2p}}\, ds + \ddot{\tau}(t)\int_\Omega \frac{2p(2p-1)\,\dot{\tau}(t)^{2p-2} g(s)^{2p}}{(1+\tau(t)g(s))^{2p}}\, ds + 2\,\epsilon\,\ddot{\tau}(t) +$$

$$\dot{\tau}(t)\int_\Omega \frac{-(2p)^2\,\dot{\tau}(t)^{2p-1} f(s)^{2p+1}}{(1+\tau(t)\,f(s))^{2p+1}}\, ds + \dot{\tau}(t)\int_\Omega \frac{-(2p)^2\,\dot{\tau}(t)^{2p-1} g(s)^{2p+1}}{(1+\tau(t)\,g(s))^{2p+1}}\, ds =$$

$$\int_\Omega \frac{-2p\,\dot{\tau}(t)^{2p} f(s)^{2p+1}}{(1+\tau(t)f(s))^{2p+1}}\, ds + \int_\Omega \frac{-2p\,\dot{\tau}(t)^{2p} g(s)^{2p+1}}{(1+\tau(t)g(s))^{2p+1}}\, ds\,.$$

Dividing[10] both sides by $2p\,\dot{\tau}(t)^{2p}$ and re-arranging we obtain:

$$\frac{\ddot{\tau}_p(t)}{\dot{\tau}(t)^2}\left(\int_\Omega \frac{(2p-1)f(s)^{2p}}{(1+\tau(t)f(s))^{2p}}\, ds + \int_\Omega \frac{(2p-1)g(s)^{2p}}{(1+\tau(t)g(s))^{2p}}\, ds\right) + \frac{\epsilon\,\ddot{\tau}(t)}{p\,\dot{\tau}(t)^{2p}} =$$

$$\int_\Omega \frac{2p\, f(s)^{2p+1}}{(1+\tau(t)\,f(s))^{2p+1}}\, ds + \int_\Omega \frac{2p\, g(s)^{2p+1}}{(1+\tau(t)\,g(s))^{2p+1}}\, ds +$$

$$-\int_\Omega \frac{f(s)^{2p+1}}{(1+\tau(t)\,f(s))^{2p+1}}\, ds - \int_\Omega \frac{g(s)^{2p+1}}{(1+\tau(t)\,g(s))^{2p+1}}\, ds$$

---

10. Note that if $\dot{\tau}(t) = 0$ for some t then the equation is trivially satisfied.

Combining the last four terms and dividing both sides by $2p - 1$ times the factor multiplying $\frac{\ddot{\tau}_p(t)}{\dot{\tau}_p(t)^2}$ and recalling the definition of K_p we obtain:

$$\frac{\ddot{\tau}_p(t)}{\dot{\tau}_p(t)^2} + \frac{\epsilon}{p(2p-1)} \frac{\ddot{\tau}_\epsilon(t)}{\dot{\tau}^{2p}(t) \, K_p(\tau(t))} = \frac{K_{p+1}(\tau(t))}{K_p(\tau(t))}$$

Notice that K_p is never zero since by lemma 20 we can lower bound it by $m_f^{2p} + m_g^{2p} > 0$. We can now re-arange the equation to obtain (59):

$$\ddot{\tau}(t) = \frac{\dot{\tau}^{2p}(t) \, K_{p+1}(\tau(t))}{\dot{\tau}^{2p-2}(t) \, K_p(\tau(t)) + \frac{\epsilon}{p(2p-1)}} \, . \tag{61}$$

Finally, notice that for $\epsilon = 0$ the Lagrangian λ_p can be decomposed as:

$$\lambda_p(x, v) = \lambda_f(x, v) + \lambda_g(x, v)$$

and following the same steps yields (59) with $\epsilon = 0$, that is,

$$\ddot{\tau}(t) = \frac{\dot{\tau}^2(t) \, K_{p+1}(\tau(t))}{K_p(\tau(t))} \, . \tag{62}$$

This completes the proof. ∎

Now problem (57) is related to the original problem (56) in a natural way. To see this, start by defining:

Definition 37 *For $p \in \mathbb{N}$ and $\epsilon > 0$ we define the family of functionals*

$$\mathcal{L}_p^\epsilon : \mathcal{T}_2 \to \mathbb{R}$$
$$\mathcal{L}_p^\epsilon = \begin{cases} \int_0^1 \lambda_p^\epsilon \left(\tau(t), \dot{\tau}(t) \right) \, dt & \text{if } \tau \in \mathcal{T}_{2p}^b \\ +\infty & \text{if } \tau \in \mathcal{T}_2 \setminus \mathcal{T}_{2p}^b \end{cases}$$

and for $\epsilon = 0$ we let

$$\mathcal{L}_p^0 : \mathcal{T}_2 \to \mathbb{R}$$
$$\mathcal{L}_p^0 = \begin{cases} \int_0^1 \lambda_p \left(\tau(t), \dot{\tau}(t) \right) \, dt & \text{if } \tau \in \mathcal{T}_{2p}^b \\ +\infty & \text{if } \tau \in \mathcal{T}_2 \setminus \mathcal{T}_{2p}^b \, . \end{cases}$$

Remark 38 *Notice that by lemma 31 we have that*

$$\inf_{\tau \in \mathcal{T}_2} \mathcal{L}_p^0[\tau] = \inf_{\tau \in \mathcal{T}_2} \int \lambda_p(\tau, \dot{\tau}) \, dt \, ,$$

that is, we can solve problem (53) by solving problem (56) which is captured by the functional $\mathcal{L}_p^0$.

Lemma 39 *For each $p \in \mathbb{N}$, the family of functionals $\left(\mathcal{L}_p^\epsilon \right)_{\epsilon > 0}$ is equicoercive and, moreover:*

$$\mathcal{L}_p^\epsilon \to \mathcal{L}_p^0 \quad \text{as } \epsilon \searrow 0 \quad \text{in the } \Gamma\text{-sense,}$$

where the Γ-convergence is understood over the space $\mathcal{T}_2$ equipped with the weak H^1 topology.

41

Proof Notice that we have the point-wise, monotone convergence

$$\mathcal{L}_p^\epsilon[\tau] \searrow \mathcal{L}_p^0[\tau] \text{ as } \epsilon \searrow 0$$

for all $\tau \in \mathcal{T}_2$, thus for any $\tau \in \mathcal{T}_2$ we have

$$\limsup_{\epsilon \to 0} \mathcal{L}_p^\epsilon[\tau] = \mathcal{L}_p^0[\tau].$$

Now fix a sequence $\gamma_n \rightharpoonup \gamma_0$ in $H^1([0,1])$. Again due to the monotonicity of $\mathcal{L}_p^\epsilon$ we have that for each $n \in \mathbb{N}$:

$$\mathcal{L}_p^{1/n}[\gamma_n] \geq \mathcal{L}_p^0[\gamma_n]$$

Taking the limit inferior as $n \to \infty$ we obtain:

$$\liminf_{n \to \infty} \mathcal{L}_p^{1/n}[\gamma_n] \geq \liminf_{n \to \infty} \mathcal{L}_p^0[\gamma_n] \geq \mathcal{L}_p^0[\gamma_0]$$

And the last step follows by lemma 24: there, it is proven that $\mathcal{L}_p^0$ is weakly lower semi-continuous in the H^1 topology, albeit in the different notation $\Lambda_p \equiv \mathcal{L}_p^0$. By the sequential characterizarion of Γ limits in metric spaces – see Braides (2006, theorem 2.1.) part (c) – we have established that the $\mathcal{L}_p^\epsilon$ converge to $\mathcal{L}_p^0$ in the Γ-sense.

To see equicoercivity, we use the monotonicity of $\mathcal{L}_p^\epsilon$ together with lemma 18 to obtain:

$$\mathcal{L}_p^\epsilon \geq \mathcal{L}_p^0 \gtrsim \mathcal{L}_1^0,$$

with the hidden constant only depending on Ω, and $\mathcal{L}_1^0 \equiv \Lambda_1$ is lower-semi continous and coercive, as seen in lemmas 24 and 26. We can now conlude by applying Dal Maso (1993, proposition 7.7.) ∎

Before showing that there is a minimizer of (56) that satisfies the strong Euler–Lagrange equations, we need one more technical lemma:

Lemma 40 *Fix a $p \in \mathbb{N}$ and let τ_ϵ be a minimizer of problem* (57)*, guaranteed to exsist by lemma* 34*. There is a constant $M > 0$ such that for all $\epsilon > 0$ we have:*

$$\|\dot{\tau}_\epsilon\|_{L^\infty([0,1])} \leq M.$$

Proof To prove this claim, we start by bounding the second derivative of τ_ϵ by constants independent of ϵ. Using the notation of lemma 36 we have

$$|\ddot{\tau}_\epsilon(t)| \leq \left| \frac{\dot{\tau}_\epsilon^{2p}(t) K_{p+1}(\tau_\epsilon(t))}{\dot{\tau}_\epsilon^{2p-2}(t) K_p(\tau_\epsilon(t))} \right| \leq \dot{\tau}_\epsilon^2(t) \frac{\sup_{x \in [0,1]} |K_{p+1}(x)|}{\inf_{x \in [0,1]} |K_p(x)|} \tag{63}$$

using that for all ϵ the map $t \mapsto \tau_\epsilon(t)$ is a bijection on $[0,1]$. Now we apply lemma 20 to obtain constants $M_f, M_g, m_f, m_g \geq 0$ depending on f, g and Ω, with $m_f \cdot m_g \neq 0$, such that for each $p \in \mathbb{N}$ we have

$$m_f^{2p} + m_g^{2p} \leq K_p(x) \leq M_f^{2p} + M_g^{2p} \quad \text{for all } x \in [0,1]$$

Thus, equation (63) becomes:

$$|\ddot{\tau}_\epsilon(t)| \leq D \, \dot{\tau}_\epsilon^2(t) \tag{64}$$

for a constant $D = D(f, g, \Omega, p, p+1) > 0$ depending only on f, g, Ω, p and $p+1$.

Now, for a fixed $p \in \mathbb{N}$, recall that the family $(\tau_\epsilon)_\epsilon$ is a family of minimizers for the problems:

$$\inf_{\tau \in \mathcal{T}_{2p}^b} \mathcal{L}_p^\epsilon[\tau]$$

By[11] the Γ-convergence of $\mathcal{L}_p^\epsilon$ to $\mathcal{L}_p^0$ as well as the equi-coercivity of $(\mathcal{L}_p^\epsilon)_{\epsilon>0}$ we have (Dal Maso, 1993, theorem 7.8.) that

$$\lim_{\epsilon \to 0} \mathcal{L}_p^\epsilon[\tau_\epsilon] = \min_{\tau \in \mathcal{T}_2} \mathcal{L}_p^0[\tau] < \infty \,.$$

Thus, there is a constant $B > 0$ such that $\mathcal{L}_p^\epsilon[\tau_\epsilon] \leq B$ for all $\epsilon > 0$. Moreover, looking into the proof of lemma 39 we see that:

$$\mathcal{L}_1^0 \leq C_\Omega \, \mathcal{L}_p^\epsilon$$

and the constant $C_\Omega > 0$ only depends on Ω. Moreover, looking at equation (46) of proposition 26 and noting that $\mathcal{L}_1^0 \equiv \Lambda_1$ we find another constant $C_{f,g,[0,1]} > 0$, depending only on f, g and the domain $[0, 1]$ such that

$$\|\tau\|_{H^1([0,1])}^2 \leq C_{f,g,[0,1]} \, \mathcal{L}_1^0[\tau]$$

obtaining

$$\|\tau\|_{H^1([0,1])}^2 \leq C_{\Omega,f,g,[0,1]} \, \mathcal{L}_p^\epsilon[\tau] \tag{65}$$

for all $\tau \in \mathcal{T}_2$. Combining with the above we get

$$\|\dot{\tau}_\epsilon\|_{L^2([0,1])}^2 \leq \|\tau_\epsilon\|_{H^1([0,1])}^2 \leq B \, C_{\Omega,f,g,[0,1]} =: C < \infty \,, \tag{66}$$

that is, the family $(\dot{\tau}_\epsilon)_\epsilon$ is bounded in $L^2([0,1])$ uniformly in ϵ and p. Now by the fundamental theorem of calculus we have:

$$\dot{\tau}_\epsilon(t) = \dot{\tau}_\epsilon(0) + \int_0^t \ddot{\tau}_\epsilon(s) \, ds \tag{67}$$

Thus, combining (67) with (64) and (66) we have

$$|\dot{\tau}_\epsilon(t)| \leq \int_0^t |\ddot{\tau}_\epsilon(s)| \, ds + |\dot{\tau}_\epsilon(0)| \leq D \int_0^1 |\dot{\tau}_\epsilon(s)|^2 \, ds + |\dot{\tau}_\epsilon(0)| \leq C \, D + |\dot{\tau}_\epsilon(0)|$$

for all $\epsilon > 0$. If we can show that $|\dot{\tau}_\epsilon(0)|$ is uniformly bounded in ϵ we are done. To do this, we integrate (67) once more and using the boundary conditions $\tau_\epsilon(0) = 0$ and $\tau_\epsilon(1) = 1$ we obtain:

$$1 = \dot{\tau}_\epsilon(0) + \int_0^1 \int_0^t \ddot{\tau}_\epsilon(s) \, ds \, dt$$

Applying (64) and (66) one again we obtain

$$|\dot{\tau}_\epsilon(0)| \leq 1 + \int_0^1 \int_0^t |\ddot{\tau}_\epsilon(s)| \, ds \, dt \leq 1 + \int_0^1 \int_0^1 D |\dot{\tau}_\epsilon(s)|^2 \, ds \, dt \leq 1 + C \, D$$

for all $\epsilon > 0$. This completes the proof. ∎

We can now show:

---

11. In effect, this is a weaker version of the Fundamental Theorem of Γ-convergence.

43

Proposition 41 *Problem* (54) *has a minimizer* τ_p *that satisfies the strong Euler–Lagrange equations:*

$$\frac{d}{dt}\left[\frac{\partial \lambda_p}{\partial v}\left(\tau_p(t), \dot{\tau}_p(t)\right)\right] - \frac{\partial \lambda_p}{\partial x}\left(\tau_p(t), \dot{\tau}_p(t)\right) = 0$$

Proof By lemma 36 the strong Euler–Lagrange equations for the problem (56) are given by the ODE

$$\ddot{\tau}(t) = \dot{\tau}^2(t)\frac{K_{p+1}(\tau(t))}{K_p(\tau(t))} \tag{68}$$

with K_p is defined by the relation

$$K_p(x) = \int_\Omega \left(\frac{f(s)}{1+xf(s)}\right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1+xg(s)}\right)^{2p} ds\,.$$

Now by lemma 34 each optimization problem (57) has a minimizer τ_ϵ. By lemma 35 the τ_ϵ satisfy the strong Euler–Lagrange equations for the Lagrangian λ_p^ϵ that, by lemma 36, can be written as:

$$\ddot{\tau}_\epsilon(t) = \frac{\dot{\tau}_\epsilon^{2p}(t)\,K_{p+1}(\tau_\epsilon(t))}{\dot{\tau}_\epsilon^{2p-2}(t)\,K_p(\tau_\epsilon(t)) + \frac{\epsilon}{p(2p-1)}}\,. \tag{69}$$

Let τ_0 be a solution to the ODE (68). We wish to show that it will also be a minimizer of the problem (54). To do this, we will establish that for a subseqence $\tau_{\epsilon_i} \to \tau_0$ in $L^2([0,1])$ and, by the Γ-convergence of the family $\mathcal{L}_p^\epsilon$ to $\mathcal{L}_p^0$, τ_0 will be a minimizer of the problem (54). More explicitly, recall that the τ_ϵ are, by definition, solutions of the following problems

$$\tau_\epsilon \in \arg\min_{\tau \in \mathcal{T}_2^b} \mathcal{L}_p^\epsilon[\tau]$$

and, moreover, problem (54) is equivalent to

$$\inf_{\tau \in \mathcal{T}_2^b} \mathcal{L}_p^0[\tau]\,.$$

In lemma 39 we showed that the family $(\mathcal{L}_p^\epsilon)_{\epsilon>0}$ is equicoercive and Γ-converges to $\mathcal{L}_p^0$. Thus, by the *Fundamental Theorem of Γ-convergence* (Braides, 2006, theorem 2.10.) we have that cluster points of pre-compact sequences of minimizers of $\mathcal{L}_p^\epsilon$ are minimizers of $\mathcal{L}_p^0$. By Dal Maso (1993, theorem 7.8.) we also have

$$\lim_{\epsilon\to 0}\mathcal{L}_p^\epsilon[\tau_\epsilon] = \inf_{\tau\in\mathcal{T}_2}\mathcal{L}_p^0[\tau] =: C < \infty$$

and using equation (65) from lemma 40 we have

$$\|\tau_\epsilon\|_{H^1([0,1])}^2 \leq C_{\Omega,f,g,[0,1]}\,\mathcal{L}_p^\epsilon[\tau_\epsilon] \leq C_{\Omega,f,g,[0,1]}\,C$$

for some constant $C_{f,g,\Omega,[0,1]} > 0$ depending only on f, g, Ω and the domain $[0,1]$. Thus, we have shown that the family $(\tau_\epsilon)_\epsilon$ is uniformly bounded in $H^1([0,1])$ and so it is pre-compact in the weak-H^1 topology. Thus, a weak-H^1 cluster point of the τ_ϵ will be a minimizer of (54). However, the Rellich–Kondrachov theorem (Brezis, 2011, theorem 9.16.), the H^1 boundedness of the τ_ϵ and Dal Maso (1993, example 8.9.) show that weak-H^1 cluster points of the τ_ϵ are strong L^2 cluster

points and vice-versa. Therefore, to conclude, it suffices to identify τ_0 as a strong L^2 cluster point of the τ_ϵ.

As done above, the key is to apply a Gronwall estimate together with the Arzelà-Ascoli Theorem. First, for $\epsilon \geq 0$ we write the dynamics as:

$$\frac{d}{dt}\begin{pmatrix} \dot{\tau}_\epsilon(t) \\ \tau_\epsilon(t) \end{pmatrix} = \begin{pmatrix} f_\epsilon(\tau_\epsilon(t), \dot{\tau}_\epsilon(t)) \\ \tau_\epsilon(t) \end{pmatrix}$$

where:

$$f_\epsilon(x, v) := \frac{v^{2p} K_{p+1}(x)}{v^{2p-2} K_p(x) + \frac{\epsilon}{p(2p-1)}}$$

Now, by Lemma 40 there is an $M > 0$ such that for all $\epsilon > 0$ we have $\|\dot{\tau}_\epsilon\|_{L^\infty([0,1])} < M$ and so we can restrict the above dynamics to the phase space box $[0,1] \times [0, M]$. Owing to the smoothness of f_ϵ we now have uniform estimates for any derivative on $[0,1] \times [0, M]$ showing, in particular, that the family $(f_\epsilon)_{\epsilon>0}$ is uniformly bounded and uniformly equi-continuous. [12] Thus, applying the Arzelà-Ascoli Theorem we obtain that there is a sequence $\epsilon_i \to 0$ such that $f_{\epsilon_i} \to f_0$ uniformly on $[0,1] \times [0, M]$ and then applying a Gronwall estimate (see theorem 2.1. in Howard (1998)) we obtain:

$$\|\tau_{\epsilon_i} - \tau_0\|_{L^\infty([0,1])} \to 0$$

In particular, this implies that $\tau_{\epsilon_i} \to \tau_0$ in $L^2([0,1])$ and we can conclude. ∎

Theorem *Problem* (53) *has a solution τ_p which satisfies*

$$\frac{d}{dt}\tau_p(t) = \frac{1}{Z_p}\left(\int_\Omega \frac{f(s)^{2p}}{(1 + \tau_p(t)f(s))^{2p}} + \frac{g(s)^{2p}}{(1 + \tau_p(t)g(s))^{2p}}\, ds \right)^{-\frac{1}{2p}} \quad \text{for all } t \in (0,1) \qquad (70)$$

---

12. For example, equi-continuity follows by differentiating first:

$$\partial_x f_\epsilon(x,v) = \frac{v^{2p} K'_{p+1}(x)}{v^{2p-2} K_p(x) + \frac{\epsilon}{p(2p-1)}} - \frac{v^{2p} K_{p+1}(x)\, v^{2p-2} K'_p(x)}{(v^{2p-2} K_p(x) + \frac{\epsilon}{p(2p-1)})^2}$$

$$\partial_v f_\epsilon(x,v) = \frac{2p v^{2p-1} K_{p+1}(x)}{v^{2p-2} K_p(x) + \frac{\epsilon}{p(2p-1)}} + \frac{v^{2p} K_{p+1}(x)\,(2p-2)v^{2p-3} K_p(x)}{(v^{2p-2} K_p(x) + \frac{\epsilon}{p(2p-1)})^2}$$

and now noticing that there is a constant $B > 0$ such that for all $(x,v) \in [0,1] \times [0, M]$ and all $\epsilon > 0$ we have:

$$|\partial_x f_\epsilon(x,v)|\,,\ |\partial_v f_\epsilon(x,v)| \leq B$$

Thus, for all $(x_1, v_1), (x_2, v_2) \in [0,1] \times [0, M]$ and all $\epsilon > 0$, by the Mean Value Theorem and a Cauchy-Schwarz estimate we have:

$$\|f_\epsilon(x_1, v_1) - f_\epsilon(x_2, v_2)\| \leq \sqrt{2}\, B\, \|(x_1, v_1) - (x_2, v_2)\|$$

on $[0,1] \times [0, M]$ as well as that f_0 is Lipschitz continuous.

together with the boundary conditions $\tau_p(0) = 0$ and $\tau_p(1) = 1$, and where

$$Z_p = \int_0^1 \left(\int_\Omega \left(\frac{f(s)}{1 + \tau_p(t)f(s)} \right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1 + \tau_p(t)g(s)} \right)^{2p} ds \right)^{-\frac{1}{2p}} dt.$$

Proof By proposition 41 we have that the problem (54) has a minimizer τ_p that satisfies the strong Euler–Lagrange equations which, by lemma 36 can be written as:

$$\ddot{\tau}(t) = \dot{\tau}^2(t) \frac{K_{p+1}(\tau(t))}{K_p(\tau(t))}$$

with

$$K_p(x) = \int_\Omega \left(\frac{f(s)}{1 + xf(s)} \right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1 + xg(s)} \right)^{2p} ds.$$

Notice that this can be re-written as

$$\frac{\ddot{\tau}(t)}{\dot{\tau}(t)} = \dot{\tau}(t) \left(-\frac{1}{2p} \right) \frac{\frac{d}{dx}K_p(\tau(t))}{K_p(\tau(t))},$$

or, equivalently:

$$\frac{d}{dt} \ln \dot{\tau}_p(t) = \frac{d}{dt} \left[-\frac{1}{2p} \ln K_p(\tau_p(t)) \right] \iff \dot{\tau}_p(t) = \frac{1}{Z_p} K_p(\tau_p(t))^{-\frac{1}{2p}}$$

where $Z_p \in \mathbb{R} \setminus \{0\}$ is a constant. Writing K_p out explicitly we obtain:

$$\dot{\tau}_p(t) = \frac{1}{Z_p} \left(\int_\Omega \left(\frac{f(s)}{1 + \tau_p(t)f(s)} \right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1 + \tau_p(t)g(s)} \right)^{2p} ds \right)^{-\frac{1}{2p}}$$

Finally, we integrate both sides from $t = 0$ to $t = 1$ and use the boundary conditions $\tau_p(0) = 0$ and $\tau_p(1) = 1$ to obtain:

$$1 = \frac{1}{Z_p} \int_0^1 \left(\int_\Omega \left(\frac{f(s)}{1 + \tau_p(t)f(s)} \right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1 + \tau_p(t)g(s)} \right)^{2p} ds \right)^{-\frac{1}{2p}} dt.$$

Solving for Z_p completes the proof. ∎

A.5. Proof of Theorem 5

In this subsection, we establish the subsequential L^2 convergence of the τ_p to τ_∞. Throughout the rest of this section we denote for all $p \in \mathbb{N}$ and all $x \in [0, 1]$

$$F_p(x) := \frac{1}{Z_p} \left[\int_\Omega \left(\frac{f(s)}{1 + xf(s)} \right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1 + xf(s)} \right)^{2p} ds \right]^{-\frac{1}{2p}},$$

$$F_\infty(x) := \max \left\{ \sup_{s \in \Omega} \left| \frac{f(s)}{1 + xf(s)} \right|, \sup_{s \in \Omega} \left| \frac{g(s)}{1 + xg(s)} \right| \right\}^{-1}$$

and additionally

$$f^* = \sup_{s \in \Omega} f(s) = \sup_{s \in \Omega} \sigma_{\max}(s) - 1, \qquad g_* = \inf_{s \in \Omega} g(s) = \inf_{s \in \Omega} \sigma_{\min}(s) - 1.$$

Lastly, we let τ_p and τ_∞ denote the solutions to the following ODEs:

$$\dot{\tau}_p(t) = F_p(\tau_p(t)) \text{ for all } t \in (0,1) \tag{71}$$

$$\dot{\tau}_\infty(t) = F_\infty(\tau_\infty(t)) \text{ for all } t \in (0,1). \tag{72}$$

Notice, now, that for any $x \in [0,1]$ the map $y \mapsto \frac{y}{1+xy}$ is monotonically increasing for $y \in [-1, \infty)$, by (44)

$$F_\infty(x) = \max\left\{ \frac{f^*}{1+xf^*}, \frac{-g_*}{1+xg_*} \right\}^{-1} = \min\left\{ \frac{1}{f^*} + x, -\frac{1}{g_*} - x \right\}.$$

As a minimum of two linear function with slope one, this function has a Lipschitz constant of 1. We now prove two lemmas that will be useful in this section.

Lemma 42 *The sequence $F_p : [0,1] \to \mathbb{R}$ is equi-bounded in p.*

Proof We use lemma 20 to obtain constants $M_f, M_g, m_f, m_g \geq 0$ such that $m_f \cdot m_g \neq 0$ satisfying:

$$m_f^{2p} \leq \int_\Omega \left(\frac{1}{1+xf(s)} \right)^{2p} ds \leq M_f^{2p}$$

$$m_f^{2p} \leq \int_\Omega \left(\frac{1}{1+xg(s)} \right)^{2p} ds \leq M_g^{2p}$$

This yields the estimate:

$$F_p \leq \frac{1}{Z_p} \left(\frac{1}{m_f^{2p} + m_g^{2p}} \right)^{\frac{1}{2p}}$$

Now to bound $\frac{1}{Z_p}$ we use the implicit expression given to us by Theorem A.4:

$$Z_p = \int_0^1 \left(\int_\Omega \left(\frac{f(s)}{1+\tau_p(t)f(s)} \right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1+\tau_p(t)g(s)} \right)^{2p} ds \right)^{-\frac{1}{2p}} dt$$

where τ_p solves ODE (70) with boundary conditions $\tau_p(0) = 0$ and $\tau_p(1) = 1$. Combining with our above estimate we obtain:

$$\frac{1}{Z_p} \leq \left(M_f^{2p} + M_g^{2p} \right)^{\frac{1}{2p}} \tag{73}$$

Thus, we conclude that:

$$0 \leq \liminf_{p \to \infty} F_p(x) \leq \limsup_{p \to \infty} F_p(x) \leq \frac{\max\{M_f, M_g\}}{\max\{m_f, m_g\}} \quad \text{for all } x \in [0,1]$$

This proves the lemma. ∎

Lemma 43 *The sequence $F_p : [0,1] \to \mathbb{R}$ is uniformly equi-continuous.*

Proof Notice first that for each p, F_p is $C^1([0,1])$. To see this note that since $f, g : \Omega \to \mathbb{R}$ are continuous, Ω is compact and the maps $x \mapsto \frac{f(s)}{1+xf(s)}$ and $x \mapsto \frac{g(s)}{1+xg(s)}$ are smooth in $[0,1]$, we can exchange integration over s with differentiation over x, to obtain:

$$\frac{d}{dx}F_p(x) = \left(-\frac{1}{2p\,Z_p}\right)\frac{\int_\Omega \frac{d}{dx}\left(\frac{f(s)}{1+xf(s)}\right)^{2p} ds + \int_\Omega \frac{d}{dx}\left(\frac{g(s)}{1+xg(s)}\right)^{2p} ds}{\left(\int_\Omega \left(\frac{f(s)}{1+xf(s)}\right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1+xg(s)}\right)^{2p} ds\right)^{\frac{1}{2p}+1}} \tag{74}$$

$$= \left(\frac{1}{Z_p}\right)\frac{\int_\Omega \left(\frac{f(s)}{1+xf(s)}\right)^{2p+1} ds + \int_\Omega \left(\frac{g(s)}{1+xg(s)}\right)^{2p+1} ds}{\left(\int_\Omega \left(\frac{f(s)}{1+xf(s)}\right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1+xg(s)}\right)^{2p} ds\right)^{\frac{1}{2p}+1}} \tag{75}$$

Notice that since the integrands are all continuous functions of x (they are bounded away from the singularity, again due to $\min\{f, g\} \geq c > -1$ on Ω by lemma 20) the map $x \mapsto F_p(x)$ is indeed C^1. Recall that to show uniform equi-continuity is suffices to find a number $K > 0$ such that for all p, if any $x, y \in [0,1]$ satisfy $|x - y| < \delta$ we have $|F_p(x) - F_p(y)| < K|x - y|$. By the Mean Value Theorem applied to the convex domain Ω, this K can be chosen as:

$$K = \sup_{p\in\{1,2,\dots\}} \sup_{x\in[0,1]} \left|\frac{dF_p(x)}{dx}\right| \tag{76}$$

provided this quantity is finite. We now show this is indeed the case. The key observation is to bound:

$$\int_\Omega \left(\frac{f(s)}{1+xf(s)}\right)^{2p+1} ds + \int_\Omega \left(\frac{g(s)}{1+xg(s)}\right)^{2p+1} ds$$

$$\leq \max\left\{\sup_{s\in\Omega}\left|\frac{f(s)}{1+xf(s)}\right|, \sup_{s\in\Omega}\left|\frac{g(s)}{1+xg(s)}\right|\right\}\left(\int_\Omega \left(\frac{f(s)}{1+xf(s)}\right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1+xg(s)}\right)^{2p} ds\right)$$

Thus, equation (75) yields:

$$\left|\frac{dF_p(x)}{dx}\right| \leq \frac{1}{Z_p}\frac{\max\left\{\sup_{s\in\Omega}\left|\frac{f(s)}{1+xf(s)}\right|, \sup_{s\in\Omega}\left|\frac{g(s)}{1+xg(s)}\right|\right\}}{\left(\int_\Omega \left(\frac{f(s)}{1+xf(s)}\right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1+xg(s)}\right)^{2p} ds\right)^{\frac{1}{2p}}}$$

Now using (73) from Lemma 42 we obtain:

$$\left|\frac{dF_p(x)}{dx}\right| \leq \left(M_f^{2p} + M_g^{2p}\right)^{\frac{1}{2p}}\frac{\max\left\{\sup_{s\in\Omega}\left|\frac{f(s)}{1+xf(s)}\right|, \sup_{s\in\Omega}\left|\frac{g(s)}{1+xg(s)}\right|\right\}}{\left(\int_\Omega \left(\frac{f(s)}{1+xf(s)}\right)^{2p} ds + \int_\Omega \left(\frac{g(s)}{1+xg(s)}\right)^{2p} ds\right)^{\frac{1}{2p}}}$$

and taking the limit $p \to \infty$ we see that:

$$\limsup_{p\to\infty}\left|\frac{dF_p(x)}{dx}\right| \leq \max\{M_f, M_g\}\frac{\max\left\{\sup_{s\in\Omega}\left|\frac{f(s)}{1+xf(s)}\right|, \sup_{s\in\Omega}\left|\frac{g(s)}{1+xg(s)}\right|\right\}}{\max\left\{\sup_{s\in\Omega}\left|\frac{f(s)}{1+xf(s)}\right|, \sup_{s\in\Omega}\left|\frac{g(s)}{1+xg(s)}\right|\right\}}$$

$$= \max\{M_f, M_g\}$$

48

using Lemma 17 to handle the limit in the denominator. Thus, we conclude that K as defined in (76) is finite and the lemma has been proven. $\blacksquare$

Lemma 44 *There is a subsequence $p_j \to \infty$ such that for constants $\epsilon_{p_j} \searrow 0$ we have:*

$$|F_{p_j}(\tau_{p_j}(t)) - F_\infty(\tau_{p_j}(t))| \leq \epsilon_{p_j}$$

Proof Note that by Lemma 17, for any $x \in [0,1]$ we have $F_p(x) \to F_\infty(x)$. By Lemma 42 the family $F_p : [0,1] \to \mathbb{R}$ is equi-bounded and by Lemma 43 it is also uniformly equi-continuous. Then, by an application of the Arzela-Ascoli theorem, we obtain the existence of a subsequence $p_j \to \infty$ such that $F_{p_j} \to F_\infty$ uniformly in $[0,1]$. Since $\tau_p(t) \in [0,1]$ for all t we then obtain the desired result. $\blacksquare$

Theorem *Let τ_∞ be the solution to the following ODE:*

$$\frac{d}{dt}\tau_\infty(t) = \frac{1}{Z} \max\left\{ \left\|\frac{f(s)}{1+\tau_\infty(t)f(s)}\right\|_{L^\infty}, \left\|\frac{g(s)}{1+\tau_\infty(t)g(s)}\right\|_{L^\infty} \right\}^{-1}$$

with boundary conditions $\tau_\infty(0) = 0$ and $\tau_\infty(1) = 1$ and

$$Z = \int_0^1 \max\left\{ \left\|\frac{f(s)}{1+\tau_\infty(t)f(s)}\right\|_{L^\infty}, \left\|\frac{g(s)}{1+\tau_\infty(t)g(s)}\right\|_{L^\infty} \right\}^{-1} dt.$$

If τ_p are solutions to the parametric ODEs (14) then there is a subsequence $\tau_{p_j} \to \tau_\infty$ in L^2 and therefore, by theorem 3, τ_∞ is optimal for (10).

Proof The proof is based on a Gronwall estimate as well as the the Arzela-Ascoli theorem. Consider the ODEs:

$$\dot{\tau}_p = F_p(\tau_p)$$
$$\dot{\tau}_\infty = F_\infty(\tau_\infty)$$

with initial conditions $\tau_p(0) = \tau_\infty(0) = 0$ and $\tau_p(1) = \tau_\infty(1) = 1$. Now, by the discussion in the beginning of the section the function F_∞ is 1-Lipschitz thus for all $x, y \in [0,1]$ we have

$$|F_\infty(x) - F_\infty(y)| \leq |x - y|.$$

Moreover, by lemma 44 there is a subsequence $p_j \to \infty$ such that for constants $\epsilon_{p_j} \searrow 0$ we have:

$$|F_{p_j}(\tau_{p_j}(t)) - F_\infty(\tau_{p_j}(t))| \leq \epsilon_{p_j}$$

It then follows by an application of Gronwall's inequality (Howard, 1998, theorem 2.1.) that for all $t \in [0,1]$ we have

$$|\tau_{p_j}(t) - \tau_\infty(t)| \leq e^t \int_0^t \epsilon_{p_j} e^{-s} \, ds$$

where we have used the initial conditions $\tau_p(0) = \tau_\infty(0) = 0$ for all $p \in \mathbb{N}$. Thus, we get:

$$\|\tau_{p_j} - \tau_\infty\|_{L^\infty([0,1])} \leq \epsilon_{p_j}$$

and so

$$\|\tau_{p_j} - \tau_\infty\|_{L^2} \to 0 \quad \text{as } j \to \infty$$

which completes the proof. $\blacksquare$

A.6. Proof of Theorem 6

In this subsection, we prove theorem 6. Let $(\sigma_i(s))_i$ denote the spectrum of $\nabla T(s)$ at $s \in \Omega \subset \mathbb{R}^d$ that is real by assumption 1. For each $s \in \Omega$ we write $\sigma_{\max}(s) = \max_i \sigma_i(s)$ and $\sigma_{\min}(s) = \min_i \sigma_i(s)$. Moreover, we introduce we the notation:

$$f^* := \sup_{s \in \Omega} f(s) = \sup_{s \in \Omega} \sigma_{\max}(s) - 1$$

$$f_* := \inf_{s \in \Omega} f(s) = \inf_{s \in \Omega} \sigma_{\max}(s) - 1$$

$$g^* := \sup_{s \in \Omega} g(s) = \sup_{s \in \Omega} \sigma_{\min}(s) - 1$$

$$g_* := \inf_{s \in \Omega} g(s) = \inf_{s \in \Omega} \sigma_{\min}(s) - 1$$

We start by establishing a critical lemma:

Lemma 45 *Suppose $\tau \in C^1([0,1])$ with $\tau(0) = 0$ and $\tau(1) = 1$. Then, for any $t \in [0,1]$ we have:*

$$\max \left\{ \sup_{s \in \Omega} \left| \frac{f(s)}{1 + \tau(t)f(s)} \right| , \sup_{s \in \Omega} \left| \frac{g(s)}{1 + \tau(t)g(s)} \right| \right\} = \max \left\{ \frac{f^*}{1 + \tau(t)f^*} , \frac{-g_*}{1 + \tau(t)g_*} \right\}$$

Proof Recall that for all t we have $0 \le \tau(t) \le 1$. This can be seen by noting that $\tau(0) = 0$, $\tau(1) = 1$ and $\dot\tau(t) \ge 0$ for all t, as seen manifestly by the ODE. We start by computing the left term inside the $\max$. Define the parametric real valued map:

$$\phi_\theta : [-1, \infty] \to \overline{\mathbb{R}}$$

$$\phi_\theta : x \mapsto \left| \frac{x}{1 + \theta x} \right|$$

And notice that:

$$\sup_{s \in \Omega} \left| \frac{f(s)}{1 + \tau(t)f(s)} \right| = \sup_{s \in \Omega} \phi_{\tau(t)}(f(s)) \,.$$

Letting $S = \{f(s) : s \in \Omega\}$ we can apply lemma 11 to obtain:

$$\sup_{s \in \Omega} \phi_{\tau(t)}(f(s)) = \max \left\{ \frac{\sup S}{1 + \tau(t)\sup S}, \frac{-\inf S}{1 + \tau(t)\inf S} \right\}$$

and since $\sup S = f^*$ and $\inf S = f_*$ we conclude that

$$\sup_{s \in \Omega} \left| \frac{f(s)}{1 + \tau(t)f(s)} \right| = \max \left\{ \frac{f^*}{1 + \tau(t)f^*} , \frac{-f_*}{1 + \tau(t)f_*} \right\} \,.$$

Arguing similarly for g we obtain:

$$\sup_{s \in \Omega} \left| \frac{g(s)}{1 + \tau(t)g(s)} \right| = \max \left\{ \frac{g^*}{1 + \tau(t)g^*} , \frac{-g_*}{1 + \tau(t)g_*} \right\}$$

and so we have:

$$\max\left\{\sup_{s\in\Omega}\left|\frac{f(s)}{1+\tau(t)f(s)}\right|,\ \sup_{s\in\Omega}\left|\frac{g(s)}{1+\tau(t)g(s)}\right|\right\}=$$
$$\max\left\{\max\left\{\frac{f^*}{1+\tau(t)f^*},\ \frac{-f_*}{1+\tau(t)f_*}\right\},\ \max\left\{\frac{g^*}{1+\tau(t)g^*},\ \frac{-g_*}{1+\tau(t)g_*}\right\}\right\}=:M$$

Notice the definition of M above, for brevity.

Case I: $g^* \geq 0$ and $f_* \leq 0$.

Now, since for all $s \in \Omega$ we have $\sigma_{\max}(s) \geq \sigma_{\min}(s) \implies f(s) \geq g(s)$ and so:

$$f^* \geq g^* \geq 0 \tag{77}$$
$$0 \geq f_* \geq g_* \tag{78}$$

Consider, now, the following sub-case:

$$\frac{f^*}{1+\tau(t)f^*} \leq \frac{-f_*}{1+\tau(t)f_*} \tag{79}$$

which is equivalent to $\frac{1}{2}\left(-\frac{1}{f_*}-\frac{1}{f^*}\right) \leq \tau(t)$. Now due to inequalities (77) - (78) we have $-\frac{1}{f^*} \geq -\frac{1}{g^*}$ and $-\frac{1}{f_*} \geq -\frac{1}{g_*}$ which imply that $\frac{1}{2}\left(-\frac{1}{g^*}-\frac{1}{g_*}\right) \leq \frac{1}{2}\left(-\frac{1}{f^*}-\frac{1}{f_*}\right) \leq \tau(t)$. Therefore, again by direct calculation it must be that $\frac{g^*}{1+\tau(t)g^*} \leq \frac{-g_*}{1+\tau(t)g_*}$ Thus, we have computed:

$$M = \max\left\{-\frac{f_*}{1+\tau(t)f_*},\ \frac{-g_*}{1+\tau(t)g_*}\right\}$$

However, by (77) and without loss of generality assuming $f_* \neq 0$ we have $-\frac{1}{f_*} \geq -\frac{1}{g_*}$ so for any $t \in [0,1]$ we have $0 \geq -\frac{1}{f_*}-\tau(t) \geq -\frac{1}{g_*}-\tau(t)$, as $f_* \geq -1$, leading to $-\frac{1}{\frac{1}{f_*}+\tau(t)} \leq -\frac{1}{\frac{1}{g_*}+\tau(t)}$ that is:

$$\frac{-f_*}{1+\tau(t)f_*} \leq \frac{-g_*}{1+\tau(t)g_*} \tag{80}$$

In other words:

$$M = \frac{-g_*}{1+\tau(t)g_*} = \max\left\{\frac{f^*}{1+\tau(t)f^*},\ \frac{-g_*}{1+\tau(t)g_*}\right\}$$

and the second equality holds by combining (79) and (80). Entering in the next sub-case, suppose that:

$$\frac{f^*}{1+\tau(t)f^*} \geq \frac{-f_*}{1+\tau(t)f_*} \tag{81}$$

which is equivalent to $\tau(t) \leq \frac{1}{2}\left(-\frac{1}{f_*}-\frac{1}{f^*}\right)$. Now if, in addition, $\frac{1}{2}\left(-\frac{1}{g^*}-\frac{1}{g_*}\right) \leq \tau(t) \leq \frac{1}{2}\left(-\frac{1}{f^*}-\frac{1}{f_*}\right)$ then $\frac{g^*}{1+\tau(t)g^*} \leq \frac{-g_*}{1+\tau(t)g_*}$ and so:

$$M = \max\left\{\frac{f^*}{1+\tau(t)f^*},\ \frac{-g_*}{1+\tau(t)g_*}\right\}$$

On the other hand, if $\tau(t) \leq \frac{1}{2}\left(-\frac{1}{g^*} - \frac{1}{g_*}\right) \leq \frac{1}{2}\left(-\frac{1}{f^*} - \frac{1}{f_*}\right)$ then $\frac{-g_*}{1+\tau(t)g_*} \leq \frac{g^*}{1+\tau(t)g^*}$ and so:

$$M = \max\left\{\frac{f^*}{1+\tau(t)f^*}, \frac{g^*}{1+\tau(t)g^*}\right\}$$

In this we can simplify further: due to (77) we have that $\frac{1}{f^*} \leq \frac{1}{g^*}$ and so for any $t \in [0,1]$ we have $0 \leq \frac{1}{f^*} + \tau(t) \leq \frac{1}{g^*} + \tau(t)$, leading to:

$$\frac{f^*}{1+\tau(t)f^*} \geq \frac{g^*}{1+\tau(t)g^*} \tag{82}$$

Thus:

$$M = \frac{f^*}{1+\tau(t)f^*} = \max\left\{\frac{f^*}{1+\tau(t)f^*}, \frac{-g_*}{1+\tau(t)g_*}\right\}$$

and the second equality holds by using $\frac{-g_*}{1+\tau(t)g_*} \leq \frac{g^*}{1+\tau(t)g^*}$.

We have proved the desired equality in an exhaustive collection of sub-cases, given $g^* \geq 0$ and $f_* \leq 0$. We can now examine the other cases.

Case II: $g^* < 0$ and $f_* > 0$.

Notice that of $\sup_{s\in\Omega} g =: g^* < 0$ then for all $s \in \Omega$ we have $g(s) < 0$. Similarly, $f_* > 0$ implies that $f(s) > 0$ for all $s \in \Omega$. Thus, recalling the function ϕ_θ we have that for all $t \in [0,1]$:

$$\sup_{s\in\Omega} \phi_{\tau(t)}(f(s)) = \left|\frac{\sup_{s\in\Omega} f(s)}{1+\tau(t)\sup_{s\in\Omega} f(s)}\right| = \frac{f^*}{1+\tau(t)f^*}$$

$$\sup_{s\in\Omega} \phi_{\tau(t)}(g(s)) = \left|\frac{\inf_{s\in\Omega} g(s)}{1+\tau(t)\inf_{s\in\Omega} g(s)}\right| = \frac{-g_*}{1+\tau(t)g_*}$$

since in the first case we will always be in the first (non-negative argument) branch of $\phi_{\tau(t)}$ and in the second case we will always be in the second (non-positive argument) branch.

Case III: $g^* \geq 0$ and $f_* > 0$.

As above, we leverage $f_* > 0$ to compute:

$$\sup_{s\in\Omega} \phi_{\tau(t)}(f(s)) = \frac{f^*}{1+\tau(t)f^*}$$

On the other hand, have established that:

$$\sup_{s\in\Omega} \phi_{\tau(t)}(g(s)) = \max\left\{\frac{g^*}{1+\tau(t)g^*}, \frac{-g_*}{1+\tau(t)g_*}\right\}$$

However, in Case I we computed that $\frac{g^*}{1+\tau(t)g^*} \leq \frac{f^*}{1+\tau(t)f^*}$ and thus we have:

$$M = \max\left\{\sup_{s\in\Omega} \phi_{\tau(t)}(f(s)), \sup_{s\in\Omega} \phi_{\tau(t)}(g(s))\right\} = \max\left\{\frac{f^*}{1+\tau(t)f^*}, \frac{-g_*}{1+\tau(t)g_*}\right\}$$

Case IV: $g^* < 0$ **and** $f_* \leq 0$**.**

As in Case III, one notes that:

$$\sup_{s \in \Omega} \phi_{\tau(t)} \left(f(s) \right) = \max \left\{ \frac{f^*}{1 + \tau(t)f^*} , \frac{-f_*}{1 + \tau(t)f_*} \right\}$$

and:

$$\sup_{s \in \Omega} \phi_{\tau(t)} \left(g(s) \right) = \frac{-g_*}{1 + \tau(t)g_*}$$

which leads, in the same way, to:

$$M = \max \left\{ \frac{f^*}{1 + \tau(t)f^*} , \frac{-g_*}{1 + \tau(t)g_*} \right\}$$

This concludes the proof of the lemma. ∎

Now using Lemma 45 we can recast our ODE as:

$$\dot{\tau}(t) = \frac{1}{Z} \max \left\{ \frac{f^*}{1 + \tau(t)f^*} , -\frac{g_*}{1 + \tau(t)g_*} \right\}^{-1} . \tag{83}$$

Before proceeding with the proof of the main theorem of this section, we need one last lemma:

Lemma 46 *If there is some $t_0 \in [0, 1]$ such that $\tau(t_0) = \tau_0 := \frac{1}{2} \left(-\frac{1}{f_*} - \frac{1}{g_*} \right)$ then the ODE* (83) *is equivalent to the system:*

$$\begin{cases} \dot{\tau}(t) = \frac{1}{Z} \left(\frac{1}{f^*} + \tau(t) \right) & \text{if } \tau(t) \leq \tau_0 \\ \dot{\tau}(t) = \frac{1}{Z} \left(-\frac{1}{g_*} - \tau(t) \right) & \text{if } \tau(t) \geq \tau_0 \end{cases} \tag{84}$$

If such t_0 does not exist then we have two cases: if $f^* \geq -g_*$ the solution to (83) *is*

$$\tau(t) = \frac{(f^* + 1)^t - 1}{f^*} \quad \text{for all } t \in [0, 1], \tag{85}$$

else if $f^* < -g_*$ then the solution to (83) *is*

$$\tau(t) = \frac{(g_* + 1)^t - 1}{g_*} \quad \text{for all } t \in [0, 1]. \tag{86}$$

Remark 47 *Note that the constant Z is fully determined by the equation*

$$Z = \int_0^1 \min \left\{ -\frac{1}{g_*} - \tau(t), \frac{1}{f^*} + \tau(t) \right\} dt,$$

which follows by integrating both sides of the ODE and using the boundary conditions $\tau(0) = 0$ and $\tau(1) = 1$.

Proof For any $c \in \mathbb{R}$, note that the the function $x \mapsto c/(x+c)$ is monotonic on any (possibly unbounded) interval that does not not contain the solution to $x + c = 0$. By lemma 20 we have that $f^* \geq g_* > -1$ and so applying the above obervation, togetther with $\dot{\tau} \geq 0$ and $\tau(0) = 0$ and $\tau(1) = 1$, we have that the $\max$ in the right hand side of the ODE will transition between its two arguments at most once, at some $t_0 \in [0, 1]$ satisfying:

$$\frac{f^*}{1 + \tau(t_0) f^*} = \frac{-g_*}{1 + \tau(t_0) g_*} \iff \tau(t_0) = \frac{1}{2}\left(-\frac{1}{f^*} - \frac{1}{g_*}\right) \tag{87}$$

Now since $0 \leq \tau(t) \leq 1$ for all $t \in [0, 1]$ and we have the ordering $-1 < g_* \leq f^*$, for the above equation to have a solution it must be that $g_* \leq 0 \leq f^*$. This implies that for $t \leq t_0$ the $\max$ chooses the left term and for $t \geq t_0$ the $\max$ chooses the right term.

If there does not exist $t_0 \in [0, 1]$ satisfying (87) then the form of the ODE (83) is determined by the relation of its two arguements at $t = 0$. If $f^* \geq -g_*$ then we have

$$\dot{\tau}(t) = \frac{1}{Z}\left(\frac{1}{f^*} + \tau(t)\right) \quad \text{for all } t \in [0, 1], \tag{88}$$

which, together with the boundary conditions $\tau(0) = 0$ and $\tau(1) = 1$ has the solution

$$\tau(t) = \frac{(f^* + 1)^t - 1}{f^*}. \tag{89}$$

One can check this explicitly, noting that in this case the constant Z is determined by the equation

$$Z = \int_0^1 \left(\frac{1}{f^*} + \tau(t)\right) dt.$$

Similarly, if $f^* < -g_*$ then we have

$$\dot{\tau}(t) = \frac{(g_* + 1)^t - 1}{g_*} \quad \text{for all } t \in [0, 1]. \tag{90}$$

∎

Theorem *Let:*

$$t_0 = \frac{\ln\left[\frac{1}{2}\left(1 - \frac{f^*}{g_*}\right)\right]}{-\ln(g_* + 1) + \ln\left[\frac{1}{4}\left(\frac{f^*}{-g_*} + \frac{-g_*}{f^*} + 2\right)\right]}$$

$$\tau(t_0) = -\frac{1}{2}\left(\frac{1}{f^*} + \frac{1}{g_*}\right)$$

Now if $0 \leq t_0 \leq 1$ then the solution to the ODE (16) is given by:

$$\tau_\infty(t) = \begin{cases} \frac{1}{f^*}\left\{\frac{1}{4(g_*+1)}\left(\frac{f^*}{-g_*} + \frac{-g_*}{f^*} + 2\right)\right\}^t - \frac{1}{f^*}, & \text{if } t \leq t_0 \\ \frac{1}{2}\left(\frac{1}{g_*} - \frac{1}{f^*}\right)\left\{\frac{g_*+1}{\frac{1}{2}\left(1-\frac{g_*}{f^*}\right)}\right\}^t \left\{\frac{1}{2}\left(1 - \frac{f^*}{g_*}\right)\right\}^{1-t} - \frac{1}{g_*}, & \text{if } t \geq t_0. \end{cases}$$

Otherwise, if $t_0 \notin [0,1]$ we have two cases: if $f^* \geq -g_*$ then the solution to (16) is given by

$$\tau_\infty(t) = \frac{(f^* + 1)^t - 1}{f^*},$$

and if $f^* < -g_*$ then the solution to (16) is given by

$$\tau_\infty(t) = \frac{(g_* + 1)^t - 1}{g_*}.$$

Proof To avoid overloading notation, let τ be a solution to ODE (16) (as opposed to τ_∞). Now, assuming there is a $t_0 \in [0,1]$ such that $\tau(t_0) = \tau_0$ we can solve the ODE (16) by solving the system of ODEs (84) together with the consistency condition

$$\lim_{\delta \searrow 0} \tau(t_0 + \delta) = \lim_{\delta \searrow 0} \tau(t_0 - \delta) = \tau_0.$$

We note that that t_0 will be uniquely determined to be the expression in the statement of the theorem. For $t \in [0, t_0]$ we can use an integrating factor to recast the ODE as:

$$\frac{d}{dt}\left(\tau(t)e^{-t/Z}\right) = \frac{e^{-t/Z}}{Z f^*}$$

Integrating both sides and using $\tau(0) = 0$ we have:

$$\tau(t) = \frac{e^{\frac{t}{Z}} - 1}{f^*} \tag{91}$$

Finally, using $\tau(t_0) = \tau_0$ we can compute Z:

$$\frac{1}{Z} = \frac{1}{t_0} \ln\left(\tau_0 f^* + 1\right) \tag{92}$$

This yields the path:

$$\tau(t) = \frac{e^{\frac{t}{t_0} \ln(\tau_0 f^* + 1)} - 1}{f^*} \tag{93}$$

Repeating the integrating factor approach for the case $t \in [t_0, 1]$ and using $\tau(t_0) = \tau_0$ we obtain:

$$\tau(t) = \frac{e^{\frac{t_0 - t}{Z}} - 1}{g_*} + e^{\frac{t_0 - t}{Z}} \tau_0 \tag{94}$$

Using that $\tau(1) = 1$ we can compute Z:

$$\frac{1}{Z} = \frac{1}{t_0 - 1} \ln\left(\frac{g_* + 1}{\tau_0 g_* + 1}\right) \tag{95}$$

This yields the path:

$$\tau(t) = \frac{e^{\frac{t_0 - t}{t_0 - 1} \ln\left(\frac{g_* + 1}{\tau_0 g_* + 1}\right)} - 1}{g_*} + e^{\frac{t_0 - t}{t_0 - 1} \ln\left(\frac{g_* + 1}{\tau_0 g_* + 1}\right)} \tau_0 \tag{96}$$

Now equating (92) and (95) we can compute t_0:

$$\frac{t_0 - 1}{t_0} = \frac{\ln\left(\frac{g_* + 1}{\tau_0 g_* + 1}\right)}{\ln\left(\tau_0 f^* + 1\right)} \tag{97}$$

$$\frac{1}{t_0} = 1 - \frac{\ln\left(\frac{g_* + 1}{\tau_0 g_* + 1}\right)}{\ln\left(\tau_0 f^* + 1\right)} \tag{98}$$

$$t_0 = \frac{1}{1 - \frac{\ln(g_* + 1) - (\tau_0 g_* + 1)}{\ln(\tau_0 f^* + 1)}} \tag{99}$$

$$t_0 = \frac{\ln\left(\tau_0 f^* + 1\right)}{-\ln\left(g_* + 1\right) + \ln\left(\tau_0 f^* + 1\right) + \ln\left(\tau_0 g_* + 1\right)} \tag{100}$$

Recalling that $\tau_0 = -\frac{1}{2}\left(1/f^* + 1/g_*\right)$ we can compute:

$$\tau_0 f^* + 1 = -\frac{1}{2}\left(1 + \frac{f^*}{g_*}\right) + 1 = \frac{1}{2}\left(1 - \frac{f^*}{g_*}\right) \tag{101}$$

$$\tau_0 g_* + 1 = -\frac{1}{2}\left(1 + \frac{g_*}{f^*}\right) + 1 = \frac{1}{2}\left(1 - \frac{g_*}{f^*}\right) \tag{102}$$

Combining the logarithm terms in the denominator of (100) we conclude:

$$t_0 = \frac{\ln\left[\frac{1}{2}\left(1 - \frac{f^*}{g_*}\right)\right]}{-\ln\left(g_* + 1\right) + \ln\left[\frac{1}{2}\left(1 - \frac{f^*}{g_*}\right)\frac{1}{2}\left(1 - \frac{g_*}{f^*}\right)\right]} \tag{103}$$

$$\tag{104}$$

Which is the same as:

$$t_0 = \frac{\ln\left[\frac{1}{2}\left(1 - \frac{f^*}{g_*}\right)\right]}{-\ln\left(g_* + 1\right) + \ln\left[\frac{1}{4}\left(\frac{f^*}{-g_*} + \frac{-g_*}{f^*}\right) + \frac{1}{2}\right]} \tag{105}$$

We can now substitute (105) into the expression for (93) to obtain:

$$\tau(t) = \frac{e^{\left\{-\ln(g_* + 1) + \ln\left[\frac{1}{4}\left(\frac{f^*}{-g_*} + \frac{-g_*}{f^*} + 2\right)\right]\right\}t} - 1}{f^*} \tag{106}$$

$$= \frac{\left\{\frac{1}{4(g_* + 1)}\left(\frac{f^*}{-g_*} + \frac{-g_*}{f^*} + 2\right)\right\}^t - 1}{f^*} \tag{107}$$

Now for (96) start by computing, for $t_0 \le t \le 1$, the quantity:

$$t_0 - t = \frac{t\,\ln\left(\frac{g_* + 1}{\tau_0 g_* + 1}\right) + (1 - t)\,\ln\left(\tau_0 f^* + 1\right)}{\ln\left[\frac{(\tau_0 f^* + 1)(\tau_0 g_* + 1)}{g_* + 1}\right]}$$

Proceed by substituting into (96) and refactoring to get:

$$\tau(t) = \left(\frac{1}{g_*} + \tau_0\right) e^{t\,\ln\left(\frac{g_* + 1}{\tau_0 g_* + 1}\right) + (1 - t)\,\ln(\tau_0 f^* + 1)} - \frac{1}{g_*}$$

Finally, using the expressions (101) together with the known form of τ_0 and combining the logarithms in the exponential we get:

$$\tau(t) = \frac{1}{2}\left(\frac{1}{g_*} - \frac{1}{f^*}\right)\left\{\frac{g_*+1}{\frac{1}{2}\left(1-\frac{g_*}{f^*}\right)}\right\}^t \left\{\frac{1}{2}\left(1-\frac{f^*}{g_*}\right)\right\}^{1-t} - \frac{1}{g_*}$$

Thus, the solution to the ODE (84) is:

$$\tau(t) = \begin{cases} \frac{1}{f^*}\left\{\frac{1}{4(g_*+1)}\left(\frac{f^*}{-g_*} + \frac{-g_*}{f^*} + 2\right)\right\}^t - \frac{1}{f^*}, & \text{if } t \leq t_0 \\ \frac{1}{2}\left(\frac{1}{g_*} - \frac{1}{f^*}\right)\left\{\frac{g_*+1}{\frac{1}{2}\left(1-\frac{g_*}{f^*}\right)}\right\}^t \left\{\frac{1}{2}\left(1-\frac{f^*}{g_*}\right)\right\}^{1-t} - \frac{1}{g_*}, & \text{if } t \geq t_0 \end{cases} \tag{108}$$

Lastly, if there is no $t_0 \in [0,1]$ such that $\tau(t_0) = \tau_0$ then, by lemma 46 we can conclude. ∎

A.7. Proof of Theorem 7 and corollary 8

In this subsection, we prove theorem 7 and corollary 8. Let $(\sigma_i(s))_i$ denote the spectrum of $\nabla T(s)$ at $s \in \Omega \subset \mathbb{R}^d$ that is real by assumption 1. For each $s \in \Omega$ we write $\sigma_{\max}(s) = \max_i \sigma_i(s)$ and $\sigma_{\min}(s) = \min_i \sigma_i(s)$ as well as $f(s) = \sigma_{\max}(s) - 1$ and $g(s) = \sigma_{\min}(s) - 1$. Moreover, we introduce the notation:

$$\sigma_{\max}^* := \sup_{s \in \Omega} \sigma_{\max}(s)$$
$$\sigma_{\min}^* := \inf_{s \in \Omega} \sigma_{\min}(s)$$

as well as

$$f^* := \sigma_{\max}^* - 1$$
$$g_* := \sigma_{\min}^* - 1 .$$

Lastly, for real valued functions $\alpha, \beta : \mathbb{R}^2 \to \mathbb{R}$ write $\alpha \asymp \beta$ if there are constants $c_1, c_2 > 0$ and some $y \in \mathbb{R}^2$ such that $c_1 \beta(x) \leq \alpha(x) \leq c_2 \beta(x)$ for all $x_i \geq y_i$ and $i \in \{1, 2\}$. Now we can proceed with the proof of theorem 7.

Theorem *For Λ as in equation (9) and $\bar{\tau}$ the trivial schedule $\bar{\tau} : t \mapsto t$ we have*

$$\Lambda[\bar{\tau}] = \max\left\{\sigma_{\max}^* - 1, \frac{1 - \sigma_{\min}^*}{\sigma_{\min}^*}\right\} .$$

Now if τ_∞ is the optimal schedule, i.e., the minimizer of problem (10) obtained as a solution of ODE (16), we have the following cases: If there exists $t_0 \in [0,1]$ such that $\tau_\infty(t_0) = -\frac{1}{2}\left(\frac{1}{\sigma_{\max}^*-1} + \frac{1}{\sigma_{\min}^*-1}\right)$ then

$$\Lambda[\tau_\infty] = \ln\left[\frac{\sigma_{\max}^* - 1}{\sigma_{\min}^*}\right] + \ln\left[\frac{1}{4}\left(\frac{1}{1-\sigma_{\min}^*} + \frac{1-\sigma_{\min}^*}{(\sigma_{\max}^* - 1)^2} + \frac{2}{\sigma_{\max}^* - 1}\right)\right] .$$

Otherwise, we have

$$\Lambda[\tau_\infty] = \begin{cases} \ln(\sigma_{\max}^*) & \text{if } \sigma_{\max}^* + \sigma_{\min}^* \geq 2 \\ -\ln(\sigma_{\min}^*) & \text{if } \sigma_{\max}^* + \sigma_{\min}^* < 2 . \end{cases}$$

Proof Recall the notation $\Omega_{[0,1]} = \Omega \times [0,1]$. By theorem (2) for any $\tau \in \mathcal{T}_\infty$ we can re-write $\Lambda[\tau]$ as:

$$\Lambda[\tau] = \sup_{(x,t)\in\Omega_{[0,1]}} |\dot{\tau}(t)| \max \left\{ \left| \frac{f(x)}{1+\tau(t)f(x)} \right|, \left| \frac{g(x)}{1+\tau(t)g(x)} \right| \right\}.$$

Using lemma 13 we can push the sup over $s \in \Omega$ inside the max and then apply lemma 45 to conclude:

$$\Lambda[\tau] = \sup_{t\in[0,1]} \dot{\tau}(t) \max \left\{ \frac{f^*}{1+\tau(t)f^*}, \frac{-g_*}{1+\tau(t)g_*} \right\} \tag{109}$$

where we have used that all $\tau \in \mathcal{T}_\infty$ are non-decreasing.

Now take $\tau = \bar{\tau}$ where $\bar{\tau}(t) = t$ is the trivial schedule. Then, we have:

$$\begin{aligned}
\Lambda[\bar{\tau}] &= \sup_{t\in[0,1]} \max \left\{ \frac{f^*}{1+t\,f^*}, \frac{-g_*}{1+t\,g_*} \right\} \\
&= \max \left\{ \max \left\{ f^*, \frac{f^*}{1+f^*} \right\}, \max \left\{ -g_*, \frac{-g_*}{1+g_*} \right\} \right\} \\
&= \max \left\{ f^*, \frac{-g_*}{1+g_*} \right\} \\
&= \max \left\{ \sigma^*_{\max} - 1, \frac{1-\sigma^*_{\min}}{\sigma^*_{\min}} \right\}
\end{aligned}$$

where in the second equality we have used the obvious monotonicity of the function $t \mapsto a/(t+a)$ for $a \in \mathbb{R}$ and in the third equality we argue as follows: recall that

$$f^* \geq g_* > -1$$

by lemma 20. If $f^* \geq 0$ then we clearly have $f^* \geq f^*/(1+f^*)$ and if $f^* \leq 0$ then $f^*/(1+f^*) \leq f^*$ since $1/(1+f^*) > 1$. A similar argument shows that the second nested max reduces to $-g/(1+g)$.

Now for $\tau = \tau_\infty$ the optimal path, recall that by theorem 6 the form of τ_∞ depends on the existence of a solution t_0 to the equation:

$$\tau_\infty(t_0) = -\frac{1}{2} \left(\frac{1}{f^*} + \frac{1}{g_*} \right) \quad \text{for} \quad t_0 \in [0,1] \tag{110}$$

This solution will be called the *transition time*.

Existence of a transition time t_0

Assume there is a transition time $t_0 \in [0,1]$ as stated above. We claim that (110) implies that $f^* \geq 0 \geq g_*$. Recall that by lemma 20 we have that $f^* \geq g_* > -1$. Now if $f^* \geq g_* \geq 0$ then (110) cannot have a solution since the left hand side is non-negative and the right hand side is negative. On the other hand, if $0 \geq f^* \geq g_* > -1$ then $-1 > 1/f^*, 1/g_*$ and so the right hand side of (110) is greater then one whereas $\tau_\infty(t) \leq 1$ for all $t \in [0,1]$.

With this information we can use the monotonicity of τ_∞ to write:

$$\dot{\tau}_\infty(t) = \begin{cases} \frac{1}{Z} \frac{1+f^*\tau_\infty(t)}{f^*} & \text{if } t \leq t_0 \\ \frac{1}{Z} \frac{1+g_*\tau_\infty(t)}{-g_*} & \text{if } t \geq t_0 \end{cases} \tag{111}$$

and Z is defined by the integral expression in remark 47. We can now combine (109) with (111) to obtain:

$$\sup_{s \in \Omega} \Lambda[\tau_\infty](s,t) = \begin{cases} \frac{\dot{\tau}_\infty(t)\,f^*}{1+\tau(t)\,f^*} & \text{if } t \leq t_0 \\ \frac{\dot{\tau}_\infty(t)\,(-g_*)}{1+\tau(t)\,g_*} & \text{if } t \geq t_0 \end{cases}$$

$$= \begin{cases} \frac{1}{Z}\,\frac{1+f^*\tau(t)}{f^*}\,\frac{f^*}{1+\tau(t)\,f^*} & \text{if } t \leq t_0 \\ \frac{1}{Z}\,\frac{1+g_*\tau(t)}{-g_*}\,\frac{-g_*}{1+\tau(t)\,g_*} & \text{if } t \geq t_0 \end{cases}$$

$$= \frac{1}{Z}$$

Now combining equation (92) with equation (100), both part of the proof of theorem 6, we obtain an explicit expression for Z:

$$\frac{1}{Z} = \frac{1}{t_0}\ln\left(\tau_0 f^* + 1\right)$$

$$= -\ln(g_* + 1) + \ln\left[\frac{1}{4}\left(\frac{f^*}{-g_*} + \frac{-g_*}{f^*} + 2\right)\right]$$

Re-writing this in terms of $\sigma^*_{\min}$ and $\sigma^*_{\max}$ we can conclude.

Non-existence of a transition time t_0

Assume, first, that $f^* \geq -g_*$. By lemma 46 the solution to the ODE is given by:

$$\tau(t) = \frac{(f^* + 1)^t - 1}{f^*} \tag{112}$$

and we can compute:

$$\dot{\tau}(t) = \ln(f^* + 1)\frac{(f^* + 1)^t}{f^*} \tag{113}$$

$$= \ln(f^* + 1)\left(\tau(t) + \frac{1}{f^*}\right) \tag{114}$$

$$= \ln(f^* + 1)\frac{1 + \tau(t)f^*}{f^*} \tag{115}$$

Moreover:

$$\max\left\{\frac{f^*}{1 + \tau(t)f^*}, \frac{-g_*}{1 + \tau(t)g_*}\right\} = \frac{f^*}{1 + \tau(t)f^*} \tag{116}$$

as the two expressions in the $\max$ can never match, given that $t_0 \notin [0,1]$, and so their relation for all t is determined by their relation at $t = 0$. Thus:

$$\sup_{s \in \Omega} \Lambda[\tau_\infty](s,t) = \ln(f^* + 1) = \ln\left(\sigma^*_{\max}\right)$$

Similarly, if $f^* < -g_*$, we obtain:

$$\sup_{s \in \Omega} \Lambda[\tau_\infty](s,t) = -\ln(g_* + 1) = -\ln\left(\sigma^*_{\min}\right).$$

Finally, writing:

$$f^* \geq -g_* \iff \sigma^*_{\max} - 1 \geq 1 - \sigma^*_{\min} \iff \sigma^*_{\max} + \sigma^*_{\min} \geq 2$$

we can conclude. ∎

Corollary *Viewing $\Lambda[\tau_\infty]$ and $\Lambda[\bar{\tau}]$ as functions of $\sigma^*_{\max}$ and $1/\sigma^*_{\min}$ we have:*

$$\Lambda[\tau_\infty] \asymp \ln \frac{\sigma^*_{\max}}{\sigma^*_{\min}}$$

$$\Lambda[\bar{\tau}] \asymp \max\left\{\sigma^*_{\max}, (\sigma^*_{\min})^{-1}\right\}$$

Proof The proof is a computation. Let us start with $\Lambda[\bar{\tau}]$. For small enough $\sigma^*_{\min}$ and large enough $\sigma^*_{\max}$ we have:

$$1/2 \leq \frac{\sigma^*_{\max} - 1}{\sigma^*_{\max}} \leq 2 \tag{117}$$

$$1/2 \leq 1 - \sigma^*_{\min} \leq 2 \tag{118}$$

so:

$$\Lambda[\bar{\tau}] = \max\left\{\frac{\sigma^*_{\max} - 1}{\sigma^*_{\max}}\sigma^*_{\max}, \frac{1 - \sigma^*_{\min}}{\sigma^*_{\min}}\right\}$$

which allows us to conclude:

$$\frac{1}{4} \max\left\{\sigma^*_{\max}, (\sigma^*_{\min})^{-1}\right\} \leq \Lambda[\bar{\tau}] \leq 4 \max\left\{\sigma^*_{\max}, (\sigma^*_{\min})^{-1}\right\}$$

For $\Lambda[\tau_\infty]$, start by observing that under the double limit $\sigma^*_{\min} \to 0$ and $\sigma^*_{\max} \to \infty$, due to the continuity of τ, the following equation has a solution:

$$\tau(t) = -\frac{1}{2}\left(\frac{1}{f^*} + \frac{1}{g_*}\right) \quad \text{some} \quad t \in [0, 1]$$

Thus, we can focus out attention to equation (23) of theorem 7. Under equations (117) - (118) and by a similar algebraic manipulation as above we can find a real number $C > 0$ such that

$$\ln \frac{\sigma^*_{\max}}{\sigma^*_{\min}} - \ln C \leq \Lambda[\tau_\infty] \leq \ln \frac{\sigma^*_{\max}}{\sigma^*_{\min}} + \ln C,$$

and noting that $\sigma^*_{\max}/\sigma^*_{\min} \to \infty$ we can conclude

$$\ln \frac{\sigma^*_{\max}}{\sigma^*_{\min}}(1 - \ln C) \leq \Lambda[\tau_\infty] \leq \ln \frac{\sigma^*_{\max}}{\sigma^*_{\min}}(1 + \ln C).$$

∎

Appendix B. Example 4.1 and unbounded domains

In this appendix, we discuss example 4.1, particularly with regards to the difficulty presented by working on $\mathbb{R}$, an unbounded domain. To apply the results of section 3 we need to work in a compact, convex $\Omega \subset \mathbb{R}$. As such, we instead choose to solve an approximate problem. To that end consider:

Definition 48 *For a subset $\Omega \subset \mathbb{R}^d$, we say that a measurable map $T : \Omega \to \mathbb{R}$ is an ϵ-transport of the measure $\mu \in \mathcal{P}(\Omega)$ to the measure $\nu \in \mathcal{P}(\Omega)$ if for all Borel sets $A \subset \Omega$ we have:*

$$|T_\sharp \mu(A) - \nu(A)| < \epsilon$$

Note that this is definitionally equivalent to saying that the *total variation distance* of $T_\sharp \mu$ and ν is less than ϵ.

Let μ_1 and μ_2 be as in example 4.1. First, we choose $\Omega = [-M, M]$ for some $M > 0$ large enough. Instead of transporting μ_1 to μ_2 we will ϵ-transport $\hat{\mu}_1 = C_1 \, \mu_1 \, \mathbf{1}_\Omega$ to $\hat{\mu}_2 = C_2 \, \mu_2 \, \mathbf{1}_\Omega$ and the C_i are normalizing constants. Since $\hat{\mu}_i$ can be made arbitrarily close to μ_i, for $i \in \{1, 2\}$, this will be good enough for our purposes. In proposition 49 we show that this can be achieved by the optimal transport map T between μ_1 and μ_2, given by the formula $T(x) = \mu_2 + \frac{\theta_2}{\theta_1}(x - \mu_1)$. This justifies optimizing the schedule of the linear linear flow $X(\cdot, t) = (1 - t)\,\mathrm{id} + t\,T$ and T is as above.

Proposition 49 *For $i \in \{1, 2\}$ let $\mu_i = \mathcal{N}(\mu_i, \sigma_i^2)$ be Gaussian distributions $\mathbb{R}$ with $\sigma_i > 0$ and $\mu_i \in \mathbb{R}$ and let T be the (unique) optimal transport map coupling μ to ν. For any $\epsilon > 0$ there exists some $M > 0$ such that the truncations $\hat{\mu}_i = C_i \, \mu_i \, \mathbf{1}_{[-M,M]}$ are such that T is an ϵ-transport (see definition 48) from $\hat{\mu}_1$ to $\hat{\mu}_2$.*

Proof Start by noting that C_i ensures that $\hat{\mu}_i$ integrates to one and so:

$$C_i = \frac{1}{\Phi\left(\frac{M - \mu_i}{\sigma_i}\right) - \Phi\left(\frac{-M - \mu_i}{\sigma_i}\right)}$$

and Φ is the cumulative distribution function of the standard normal distribution. Recall that the transport map T is affine:

$$T(x) = \mu_2 + \frac{\theta_2}{\theta_1}(x - \mu_1)$$

and choose M large enough such that:

$$|C_1 - C_2| < \epsilon/3 \quad \text{and}$$

$$\int_{\mathbb{R}\setminus[-M,M]} \frac{C_1}{\sqrt{2\pi}\theta_2} \exp\left(-\frac{(x - \mu_2)^2}{2\theta_2^2}\right) dx < \epsilon/3 \quad \text{and}$$

$$\int_{\mathbb{R}\setminus[T(-M),T(M)]} \frac{C_1}{\sqrt{2\pi}\theta_2} \exp\left(-\frac{(x - \mu_2)^2}{2\theta_2^2}\right) dx < \epsilon/3\,.$$

We claim that the density of $T_\sharp \hat{\mu}_1$ satisfies:

$$\frac{d\left(T_\sharp \hat{\mu}_1\right)}{dx} = C_1 \frac{1}{\sqrt{2\pi}\theta_2} \exp\left(-\frac{(x - \mu_2)^2}{2\theta_2^2}\right) \mathbf{1}_{[T(-M),T(M)]}$$

which follows by[13] the change of variables formula:

$$\frac{d\left(T_\sharp \hat{\mu}_1\right)}{dx} = \frac{1}{\det \nabla T\left(T^{-1}(x)\right)} \frac{d\hat{\mu}_1}{dx}\left(T^{-1}(x)\right)$$

as seen for example in section 1.6. of Figalli and Glaudo (2023). Thus, for any Borel $A \subset [-M, M]$ we have:

$$
\begin{aligned}
|T_\sharp \hat{\mu}_1(A) - \hat{\mu}_2(A)| &\leq \int_A \left| \frac{d\left(T_\sharp \hat{\mu}_1\right)}{dx} - \frac{d\hat{\mu}_2}{dx} \right| dx \\
&= \int_A \left| C_1 \mathbf{1}_{[T(-M), T(M)]} - C_2 \mathbf{1}_{[-M,M]} \right| \frac{1}{\sqrt{2\pi}\theta_2} \exp\left(-\frac{(x - \mu_2)^2}{2\theta_2^2} \right) dx \\
&\leq \int_{\mathbb{R}} |C_1 - C_2| \frac{1}{\sqrt{2\pi}\theta_2} \exp\left(-\frac{(x - \mu_2)^2}{2\theta_2^2} \right) dx \\
&\quad + \int_{\mathbb{R}\setminus[-T(M), T(M)]} \frac{C_1}{\sqrt{2\pi}\theta_2} \exp\left(-\frac{(x - \mu_2)^2}{2\theta_2^2} \right) dx \\
&\quad + \int_{\mathbb{R}\setminus[-M,M]} \frac{C_1}{\sqrt{2\pi}\theta_2} \exp\left(-\frac{(x - \mu_2)^2}{2\theta_2^2} \right) dx \\
&< \epsilon/3 + \epsilon/3 + \epsilon/3 \\
&= \epsilon
\end{aligned}
$$

∎

---

13. Note that we also used the monotonicity of T to write:

$$\mathbf{1}_{[-M,M]}\left(T^{-1}(x)\right) = \mathbf{1}_{[T(-M), T(M)]}(x)$$

Appendix C. Related work

The problem of finding optimal schedules has drawn much attention lately, mostly in the context of *diffusion models* (Song et al., 2021). Here we provide a brief review of relevant concurrent work.

Aranguri et al. (2025) propose a time-dilated stochastic interpolant that effectively alters the noising schedule of a diffusion-based generative model; they show that the associated probability flow ODE can then be discretized in time with a non-uniform grid whose size is asymptotically independent of the dimension of the data-generating distribution. The notion of a noising schedule, while natural for diffusion paths, does not appear explicitly in our work. Moreover, the analysis in Aranguri et al. (2025) focuses on two-mode Gaussian mixture targets, in contrast with the much more general target distributions we consider here.

Again in the context of diffusion models, Williams et al. (2024) propose an algorithm for adapting the discretization of the backwards diffusion process, based on a notion of cost associated with each step. This cost is derived from a hypothetical predictor-corrector scheme for simulating the backwards diffusion equation; a discretization schedule that minimizes the sum of these costs, over all steps, is found iteratively. While the discretization found by this scheme is analogous to our τ applied to a uniform discretization of the unit time interval, the notion of cost is quite different from our Lipschitz objective. Sabour et al. (2024) also consider the discretization of backwards diffusion paths, using numerical optimization over discretization points to minimize a certain upper bound on the Kullback–Leibler divergences between marginals of the continuous and discretized paths.

With the goal of solving stochastic optimal control problems, Domingo-Enrich et al. (2023) propose a least squares-type objective that is analogous to a diffusion model loss function. To make the loss tractable, the authors introduce a family of matrix-valued time reparameterizations M_t that are learned along with the control. The functions $t \mapsto M_t$ have the flavor of a generalized schedule, but no explicit form of the optimal choice is given, except to establish that it satisfies a complex integral equation. Domingo-Enrich et al. (2024) consider the reward fine-tuning of flow-based generative models, where, given a generative diffusion model that produces samples from some base distribution π on $\mathbb{R}^d$, one wishes to modify it so that it samples from the tilted distribution $\pi \exp(r)$, for some reward function $r : \mathbb{R}^d \to \mathbb{R}$. The fine-tuning problem is recast as a stochastic optimal control problem, and the authors argue that a so-called "memoryless" noise schedule is necessary to achieve fine-tuning in this setting. Memorylessness turns out to be equivalent to a specific form of noise schedule for the forward process. While this schedule is given explicitly, the setting of diffusion models and objective of fine-tuning are very different from those considered in this work.

Broadly, we comment that all these works have different notions of optimality, and none of them studies optimal schedules from the viewpoint of velocity approximation, as we do in this work.

Appendix D. Numerical Illustrations

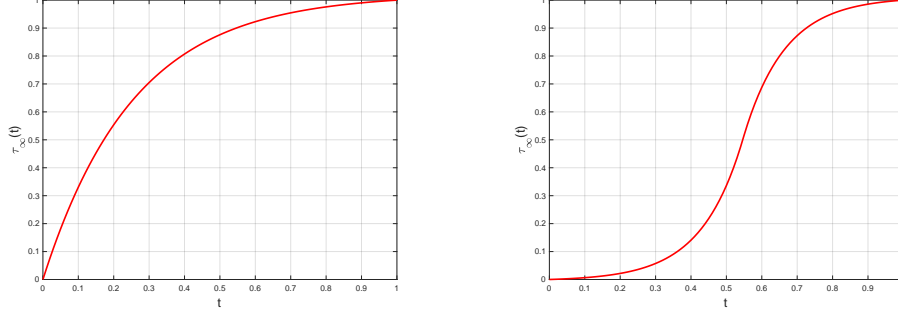### D.1. Optimal schedules for unimodal Gaussian source and target measures.

Figure 2: Optimal schedules τ_∞: *(left)* for the Gaussian example of Figure 1 and Section 4.1, where $\theta_2 \ll \theta_1$; and *(right)* for the Gaussian mixture example given at the end of Section 4.3.

D.2. Optimal schedules for unimodal Gaussian source and multimodal Gaussian mixture target.

Figures 3 and 4 illustrate this scaling for a standard Gaussian source measure and a bimodal Gaussian mixture target. Figure 3 shows that trajectories under the optimal schedule τ_∞ slow down near both $t = 0$ and $t = 1$, consistent with the discussion in Section 4.2. The optimal schedule for this example is shown in the right panel of Figure 2. Figure 4 shows the spatial Lipschitz constant of the velocity field v as a function time, for both the trivial and optimal schedules. The exponential improvement in the worst-case Lipschitz constant is manifest, and we see that the optimal schedule indeed makes the Lipschitz constant itself become constant over time, consistent with the observation made in Remark 10.
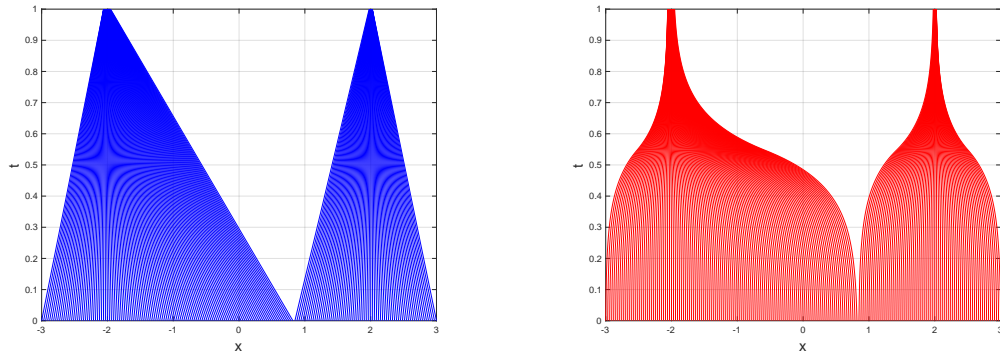

Figure 3: Trajectories $t \mapsto X(x, t)$ transporting a standard Gaussian μ at $t = 0$ to a bimodal Gaussian mixture $\nu = 0.8\mathcal{N}(-2, 0.02^2) + 0.2\mathcal{N}(2, 0.01^2)$ at $t = 1$: *(left)* straight-line trajectories; *(right)* trajectories produced by the optimal schedule τ_∞.
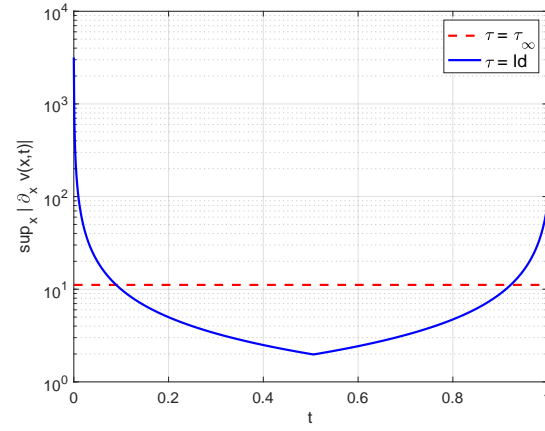
64

Figure 4: Spatial Lipschitz constant of the velocity field versus time t, for the Gaussian mixture example described in Figure 3: under the trivial schedule, $\mathrm{Lip}(v(\cdot, t))$ (blue/solid line); and under the optimal schedule, $\mathrm{Lip}(v_{\tau_\infty}(\cdot, t))$ (red/dashed line).