

The Mechanics of n -Player Differentiable Games

David Balduzzi¹ Sébastien Racanière¹ James Martens¹ Jakob Foerster² Karl Tuyls¹ Thore Graepel¹

Abstract

The cornerstone underpinning deep learning is the guarantee that gradient descent on an objective converges to local minima. Unfortunately, this guarantee fails in settings, such as generative adversarial nets, where there are multiple interacting losses. The behavior of gradient-based methods in games is not well understood – and is becoming increasingly important as adversarial and multi-objective architectures proliferate. In this paper, we develop new techniques to understand and control the dynamics in general games. The key result is to decompose the second-order dynamics into two components. The first is related to potential games, which reduce to gradient descent on an implicit function; the second relates to *Hamiltonian games*, a new class of games that obey a conservation law, akin to conservation laws in classical mechanical systems. The decomposition motivates *Symplectic Gradient Adjustment* (SGA), a new algorithm for finding stable fixed points in general games. Basic experiments show SGA is competitive with recently proposed algorithms for finding local Nash equilibria in GANs – whilst at the same time being applicable to – and having guarantees in – much more general games.

1. Introduction

Recent progress in machine learning is heavily dependent on using gradient descent, applied to optimize the parameters of models with respect to a (single) objectives. A basic result is that gradient descent converges to a local minimum of the objective under a broad range of conditions (Lee et al., 2017). However, there is a rapidly growing set of powerful models that do not optimize a single objective, including: generative adversarial networks (Goodfellow et al., 2014), proximal gradient TD learning (Liu et al., 2016), multi-level optimization (Pfau & Vinyals, 2016), synthetic gradients

(Jaderberg et al., 2017), hierarchical reinforcement learning (Wayne & Abbott, 2014; Vezhnevets et al., 2017), curiosity (Pathak et al., 2017), and imaginative agents (Racanière et al., 2017). In effect, the models are trained via games played by cooperating and competing modules.

No-regret algorithms such as gradient descent are guaranteed to converge to coarse correlated equilibria in games (Stoltz & Lugosi, 2007). However, the dynamics do not converge to Nash equilibria – and do not even stabilize – in general (Mertikopoulos et al., 2018). Concretely, cyclic behaviors emerge even in simple cases, see example 1.

This paper presents an analysis of the second-order structure of game dynamics that allows to identify two classes of games, potential and Hamiltonian, that are easy to solve separately. We then derive *symplectic gradient adjustment* (SGA), a method for finding stable fixed points in games. SGA’s performance is evaluated in basic experiments.

1.1. Background and problem description

Procedures that converge to Nash equilibria have been found for restricted game classes: potential games, 2-player zero-sum games and a few others (Hu & Wellman, 2003; Hart & Mas-Colell, 2013). Finding equilibria can be reformulated as a nonlinear complementarity problem, but these are ‘hopelessly impractical to solve’ in general (Shoham & Leyton-Brown, 2008) because the problem is PPAD hard (Daskalakis et al., 2009).

Players are primarily neural nets in our setting. We therefore restrict to methods based on gradient-descent (game-theorists have considered a much broader range of techniques). The losses are not necessarily convex in *any* of their parameters, so the concept of Nash equilibrium does not apply here. Nash equilibria are only guaranteed to exist for convex losses and constraint sets. Even leaving existence aside, finding Nash equilibria is analogous to, but much harder than, finding global minima in neural nets – which is not realistic with gradient-based methods.

There are (at least) three problems with gradient descent in games. Firstly, the potential existence of cycles (recurrent dynamics) implies there are no convergence guarantees, see example 1 and Mertikopoulos et al. (2018). Secondly, even when gradient descent converges, the rate may be

¹DeepMind ²University of Oxford. Correspondence to: David Balduzzi <dbalduzzi@google.com>.

too slow in practice because ‘rotational forces’ necessitate extremely small learning rates (see figure 3). Finally, since there is no single objective, there is no way to measure progress. Application-specific proxies have been proposed, for example the inception score for GANs (Salimans et al., 2016), but these are little help during training – the inception score is no substitute for looking at samples.

1.2. Outline and summary of main contributions

We start with the well-known case of a zero-sum bimatrix game: example 1. It turns out that the dynamics (that is, the dynamics under simultaneous gradient descent) can be reformulated via Hamilton’s equations. The cyclic behavior arises because the dynamics live on the level sets of the Hamiltonian. More directly useful, gradient descent on the Hamiltonian converges to a Nash equilibrium.

Lemma 1 shows that the Hessian of any game decomposes into symmetric and antisymmetric components. There are thus two ‘pure’ cases: when only the symmetric component is present, or only the antisymmetric. The first case, known as potential games (Monderer & Shapley, 1996), have been intensively studied because they are exactly the games where gradient descent *does* converge.

The second case, Hamiltonian¹ games, were not studied previously, probably because they coincide with zero-sum games in the bimatrix case (or constant-sum, depending on the constraints). Zero-sum and Hamiltonian games differ when the losses are not bilinear or there are more than two players. Hamiltonian games are important because (i) they are easy to solve and (ii) general games combine potential-like and Hamiltonian-like dynamics. The concept of a zero-sum game is too loose to be useful when there are many players: any n -player game can be reformulated as a zero-sum $(n + 1)$ -player game where $\ell_{n+1} = -\sum_{i=1}^n \ell_i$. Zero-sum games are as complicated as general-sum games. Theorem 3 shows that Hamiltonian games obey a conservation law – which also provides the key to solving them, by gradient descent on the conserved quantity.

The general case, neither potential nor Hamiltonian, is more difficult and the focus of the remainder of the paper. Section 3 proposes *symplectic gradient adjustment* (SGA), a gradient-based method for finding stable fixed points in general games. Appendix B contains TensorFlow code to compute the adjustment. The algorithm computes two Hessian-vector products, at a cost of two iterations of backprop. SGA satisfies a few natural desiderata: (D1) it is compatible with the original dynamics and (D2, D3) it is guaranteed to find stable equilibria in potential and Hamiltonian games.

For general games, correctly picking the *sign* of the adjustment (whether to add or subtract) is critical since it

determines the behavior near unstable equilibria. Lemma 2 proves a characterization of local Nash equilibria. The characterization yields a local criterion to distinguish stable from unstable equilibria, and a local measure of which direction improves progress, lemma 9. Correctly aligning SGA allows higher learning rates and faster, more robust convergence, see theorem 7 and experiments in section 4.

Section 4 investigates a basic GAN setup from Metz et al. (2017), that tests for mode-collapse and mode hopping. Whereas simultaneous gradient descent completely fails; the symplectic adjustment leads to rapid convergence – slightly improved by correctly choosing the sign of the adjustment.

Finally, section 3.5 applies the same criterion to align consensus optimization (Mescheder et al., 2017), preventing it from converging to unstable equilibria and (slightly) improving performance, figure 9 in the appendix.

Caveat. The behavior of SGA near saddle points is not analysed below. On the one hand, it was only recently shown that gradient descent behaves well near saddles when optimizing a single objective (Lee et al., 2016; 2017). On the other hand, Newton’s method is attracted to saddles, see analysis and proposed remedy in Dauphin et al. (2014). Studying saddles in general games is deferred to future work.

1.3. Related work

Convergence to Nash equilibria in two-player games was studied in Singh et al. (2000). WoLF (Win or Learn Fast) is guaranteed to converge to Nash equilibria in two-player two-action games (Bowling & Veloso, 2002). Extensions include weighted policy learning (Abdallah & Lesser, 2008) and GIGA-WoLF (Bowling, 2004).

There has been almost no work on convergence to fixed points in general games. Optimistic mirror descent (OMD) exhibits approximate convergence in two-player bilinear zero-sum games (Daskalakis et al., 2018), which are a special case of Hamiltonian games. In more general settings it converges to coarse correlated equilibria.

There has been interesting recent work on convergence in GANs. Heusel et al. (2017) propose a two-time scale methods to find Nash. However, it likely scales badly with the number of players. Nagarajan & Kolter (2017) prove convergence for some algorithms, but under very strong assumptions Mescheder (2018). Consensus optimization (Mescheder et al., 2017) is discussed in section 3.

Learning with opponent-learning awareness (LOLA) infinitesimally modifies the objectives of players to take into account their opponents’ goals (Foerster et al., 2018). However, there are no guarantees that LOLA converges and, moreover, in general it may modify the fixed-points of the

¹Lu (1992) defined an unrelated notion of Hamiltonian game.

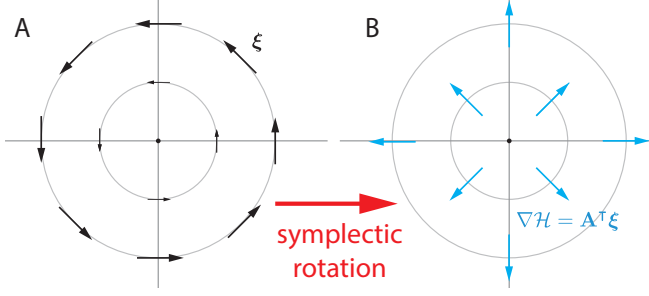


Figure 1. (A) ξ cycles around the origin. (B) gradient descent on \mathcal{H} converges to Nash.

game.

Notation. Dot products are written as $\mathbf{v}^\top \mathbf{w}$ or $\langle \mathbf{v}, \mathbf{w} \rangle$. The angle between vectors is $\theta(\mathbf{v}, \mathbf{w})$. Positive definiteness is denoted $\mathbf{S} \succ 0$. Omitted proofs are in appendix A.

2. The infinitesimal structure of games

In contrast to the classical formulation of games, we do not constrain the parameter sets (e.g. to the probability simplex) or require losses to be convex in the corresponding players' parameters. Players could be interacting neural nets such as GANs (Goodfellow et al., 2014).

Definition 1. A game is a collection of players $[n] = \{1, \dots, n\}$ and twice continuously differentiable losses $\{\ell_i : \mathbb{R}^d \rightarrow \mathbb{R}\}_{i=1}^n$. Parameters are $\mathbf{w} = (\mathbf{w}_1, \dots, \mathbf{w}_n) \in \mathbb{R}^d$ with $\mathbf{w}_i \in \mathbb{R}^{d_i}$ where $\sum_{i=1}^n d_i = d$.

It is sometimes convenient to write $\mathbf{w} = (\mathbf{w}_i, \mathbf{w}_{-i})$ where \mathbf{w}_{-i} concatenates the parameters of all the players other than the i^{th} , which is placed out of order by abuse of notation.

Computing the gradients of the losses with respect to the relevant parameters yields the simultaneous gradient

$$\xi(\mathbf{w}) = (\nabla_{\mathbf{w}_1} \ell_1, \dots, \nabla_{\mathbf{w}_n} \ell_n) \in \mathbb{R}^d.$$

By the **dynamics** of the game, we mean following the negative of the vector field ξ with infinitesimal steps. There is no reason to expect ξ to be the gradient of a *single* function in general, and therefore no reason to expect the dynamics to converge to a fixed point.

2.1. Warmup: Hamiltonian mechanics in games

The following example illustrates the essential problem with gradients in games and a key ingredient of our approach.

Example 1 (Conservation of energy in a zero-sum bimatrix game). Zero-sum games, where $\sum_{i=1}^n \ell_i \equiv 0$, are a well-studied class of games. The zero-sum game

$$\ell_1(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top \mathbf{A} \mathbf{y} \quad \text{and} \quad \ell_2(\mathbf{x}, \mathbf{y}) = -\mathbf{x}^\top \mathbf{A} \mathbf{y}$$

has Nash equilibrium at $(\mathbf{x}, \mathbf{y}) = (\mathbf{0}, \mathbf{0})$. The simultaneous gradient $\xi(\mathbf{x}, \mathbf{y}) = (\mathbf{A} \mathbf{y}, -\mathbf{A}^\top \mathbf{x})$ rotates around the Nash equilibrium, see figure 1.

The matrix \mathbf{A} admits singular value decomposition (SVD) $\mathbf{A} = \mathbf{U}^\top \mathbf{D} \mathbf{V}$. Changing to coordinates $\mathbf{u} = \mathbf{D}^{\frac{1}{2}} \mathbf{U} \mathbf{x}$ and $\mathbf{v} = \mathbf{D}^{\frac{1}{2}} \mathbf{V} \mathbf{y}$ gives $\ell_{1/2}(\mathbf{u}, \mathbf{v}) = \pm \mathbf{u}^\top \mathbf{v}$. Introduce the Hamiltonian

$$\mathcal{H} = \frac{1}{2} (\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2) = \frac{1}{2} (\mathbf{x}^\top \mathbf{U}^\top \mathbf{D} \mathbf{U} \mathbf{x} + \mathbf{y}^\top \mathbf{V}^\top \mathbf{D} \mathbf{V} \mathbf{y}).$$

Remarkably, the dynamics can be reformulated via Hamilton's equations in the coordinates given by the SVD of \mathbf{A} :

$$\xi(\mathbf{u}, \mathbf{v}) = \left(\frac{\partial \mathcal{H}}{\partial \mathbf{v}}, -\frac{\partial \mathcal{H}}{\partial \mathbf{u}} \right).$$

The vector field ξ cycles around the equilibrium because ξ conserves the level sets of the Hamiltonian. To see this, check that $\langle \xi, \nabla \mathcal{H} \rangle = 0$. However, **gradient descent on the Hamiltonian converges to the Nash equilibrium.**

Papadimitriou & Piliouras (2016) recently analyzed the dynamics of Matching Pennies (essentially, the above example) and showed with some effort that the cyclic behavior covers the entire parameter space. The Hamiltonian reformulation directly explains the cyclic behavior via a conservation law.

2.2. The generalized Helmholtz decomposition

The **Hessian** of a game is the $(d \times d)$ -matrix of second-derivatives $\mathbf{H}(\mathbf{w}) := \nabla_{\mathbf{w}} \cdot \xi(\mathbf{w})^\top = \left(\frac{\partial \xi_\alpha(\mathbf{w})}{\partial w_\beta} \right)_{\alpha, \beta=1}^d$. Concretely, the Hessian can be written

$$\mathbf{H}(\mathbf{w}) = \begin{pmatrix} \nabla_{\mathbf{w}_1}^2 \ell_1 & \nabla_{\mathbf{w}_1, \mathbf{w}_2}^2 \ell_1 & \cdots & \nabla_{\mathbf{w}_1, \mathbf{w}_n}^2 \ell_1 \\ \nabla_{\mathbf{w}_2, \mathbf{w}_1}^2 \ell_2 & \nabla_{\mathbf{w}_2}^2 \ell_2 & \cdots & \nabla_{\mathbf{w}_2, \mathbf{w}_n}^2 \ell_2 \\ \vdots & \vdots & \ddots & \vdots \\ \nabla_{\mathbf{w}_n, \mathbf{w}_1}^2 \ell_n & \nabla_{\mathbf{w}_n, \mathbf{w}_2}^2 \ell_n & \cdots & \nabla_{\mathbf{w}_n}^2 \ell_n \end{pmatrix},$$

where $\nabla_{\mathbf{w}_i, \mathbf{w}_j}^2 \ell_k$ is the $(d_i \times d_j)$ -block of 2nd-order derivatives. The Hessian of a game is not necessarily symmetric. Note: α, β run over dimensions; i, j run over players.

Lemma 1 (generalized Helmholtz decomposition). *The Hessian of any vector field decomposes uniquely into two components $\mathbf{H}(\mathbf{w}) = \mathbf{S}(\mathbf{w}) + \mathbf{A}(\mathbf{w})$ where $\mathbf{S} \equiv \mathbf{S}^\top$ is symmetric and $\mathbf{A} + \mathbf{A}^\top \equiv 0$ is antisymmetric.*

Proof. Any matrix decomposes uniquely as $\mathbf{M} = \mathbf{S} + \mathbf{A}$ where $\mathbf{S} = \frac{1}{2}(\mathbf{M} + \mathbf{M}^\top)$ and $\mathbf{A} = \frac{1}{2}(\mathbf{M} - \mathbf{M}^\top)$ are symmetric and antisymmetric. Applying the decomposition to the game Hessian yields the result. \square

The decomposition is preserved by orthogonal change-of-coordinates $\mathbf{P}^\top \mathbf{M} \mathbf{P} = \mathbf{P}^\top \mathbf{S} \mathbf{P} + \mathbf{P}^\top \mathbf{A} \mathbf{P}$ since the terms remain symmetric and antisymmetric.

The connection to the classical Helmholtz decomposition in calculus is sketched in appendix D. Two obvious classes of games arise from the decomposition:

Definition 2. A game is a *potential game* if $\mathbf{A}(\mathbf{w}) \equiv 0$. It is a *Hamiltonian game* if $\mathbf{S}(\mathbf{w}) \equiv 0$.

Potential games are well-studied and easy to solve. Hamiltonian games are a new class of games that are also easy to solve. The general case is more delicate, see section 3.

2.3. A characterization of local Nash equilibria

Gradient-based methods can reliably find local – but not global – optima of nonconvex objective functions (Lee et al., 2016; 2017). Similarly, gradient-based methods cannot be expected to find (global) Nash equilibria in games.

Definition 3. A point \mathbf{w}^* is a *local Nash equilibrium* if, for all i , there is a neighborhood U_i of \mathbf{w}_i^* such that $\ell_i(\mathbf{w}_i', \mathbf{w}_{-i}^*) \geq \ell_i(\mathbf{w}_i^*, \mathbf{w}_{-i}^*)$ for $\mathbf{w}_i' \in U_i$.

Lemma 2 (Characterization of local Nash equilibria). *The point \mathbf{w} is a local Nash equilibrium iff $\boldsymbol{\xi}(\mathbf{w}) = 0$ and $\mathbf{S} \succeq 0$ in a neighborhood of \mathbf{w} .*

Proof. If \mathbf{S} is positive semidefinite then so are its $(d_i \times d_i)$ -submatrices $\mathbf{S}_i := \nabla_{\mathbf{w}_i}^2 \ell_i$ for all i . The result follows. \square

We refer to a fixed point, with $\boldsymbol{\xi}(\mathbf{w}) = 0$, as *stable* if $\mathbf{S}(\mathbf{w}) \succeq 0$ and *unstable* if $\mathbf{S}(\mathbf{w}) \prec 0$. Saddle points, neither stable nor unstable, are beyond the scope of the paper.

2.4. Potential games

Potential games were introduced in Monderer & Shapley (1996). A game is a potential game if there is a single potential function $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ and positive numbers $\{\alpha_i > 0\}_{i=1}^n$ such that

$$\phi(\mathbf{w}_i', \mathbf{w}_{-i}) - \phi(\mathbf{w}_i'', \mathbf{w}_{-i}) = \alpha_i (\ell_i(\mathbf{w}_i', \mathbf{w}_{-i}) - \ell_i(\mathbf{w}_i'', \mathbf{w}_{-i}))$$

for all i and all $\mathbf{w}_i', \mathbf{w}_i'', \mathbf{w}_{-i}$. Monderer & Shapley (1996) show a game is a potential game iff $\alpha_i \nabla_{\mathbf{w}_i} \ell_i = \nabla_{\mathbf{w}_i} \phi$ for all i , which is equivalent to

$$\alpha_i \nabla_{\mathbf{w}_i}^2 \ell_i = \alpha_j \nabla_{\mathbf{w}_i}^2 \ell_j = \alpha_j \left(\nabla_{\mathbf{w}_j}^2 \ell_j \right)^\top \quad \forall i, j.$$

It follows that our definition of potential game is the special case where $\alpha_i = 1$ for all i , which Monderer & Shapley (1996) refer to as an *exact* potential game. We use the shorthand ‘potential game’ to refer to exact potential games in what follows.

Potential games have been extensively studied since they are one of the few classes of games for which Nash equilibria can be computed (Rosenthal, 1973). For our purposes,

they are games where simultaneous gradient descent on the losses is gradient descent on a single function. It follows that descent on $\boldsymbol{\xi}$ converges to a fixed point that is a local minimum of ϕ – or saddle, but see Lee et al. (2017).

2.5. Hamiltonian games

A concrete example may help understand antisymmetric matrices. Suppose n competitors play one-on-one and that the probability of player i beating player j is p_{ij} . Then, assuming there are no draws, the probabilities satisfy $p_{ij} + p_{ji} = 1$ and $p_{ii} = \frac{1}{2}$. The matrix $\mathbf{A} = \left(\log \frac{p_{ij}}{1-p_{ij}} \right)_{i,j=1}^n$ of logits is then antisymmetric. Intuitively, antisymmetry reflects a *hyperadversarial* setting where all pairwise interactions between players are zero-sum. In general, Hamiltonian games are related to – but distinct from – zero-sum games.

Example 2 (an unconstrained² bimatrix game is zero-sum iff it is Hamiltonian). *Consider bimatrix game with $\ell_1(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top \mathbf{P} \mathbf{y}$ and $\ell_2(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top \mathbf{Q} \mathbf{y}$. Then $\boldsymbol{\xi} = (\mathbf{P} \mathbf{y}, \mathbf{Q}^\top \mathbf{x})$ and the Hessian components have block structure*

$$2\mathbf{A} = \begin{pmatrix} 0 & \mathbf{P} - \mathbf{Q} \\ (\mathbf{Q} - \mathbf{P})^\top & 0 \end{pmatrix} \quad 2\mathbf{S} = \begin{pmatrix} 0 & \mathbf{P} + \mathbf{Q} \\ (\mathbf{P} + \mathbf{Q})^\top & 0 \end{pmatrix}$$

The game is Hamiltonian iff $\mathbf{S} = 0$ iff $\mathbf{P} + \mathbf{Q} = 0$ iff $\ell_1 + \ell_2 = 0$.

There are examples of Hamiltonian games that are not zero-sum, and vice versa.

Example 3 (Hamiltonian game that is not zero-sum). *Fix constants a and b and suppose players 1 and 2 minimize losses*

$$\ell_1(x, y) = x(y - b) \quad \text{and} \quad \ell_2(x, y) = -(x - a)y$$

with respect to x and y respectively.

Example 4 (zero-sum game that is not Hamiltonian). *Players 1 and 2 minimize*

$$\ell_1(x, y) = x^2 + y^2 \quad \ell_2(x, y) = -(x^2 + y^2).$$

The game actually has potential function $\phi(x, y) = x^2 - y^2$.

Hamiltonian games are quite different from potential games. There is a Hamiltonian function \mathcal{H} that specifies a conserved quantity. Whereas the dynamics *equal* $\nabla \phi$ in potential games; they are *orthogonal* to $\nabla \mathcal{H}$ in Hamiltonian games. The orthogonality implies the conservation law that underlies the cyclic behavior in example 1.

Theorem 3. *Let $\mathcal{H}(\mathbf{w}) := \frac{1}{2} \|\boldsymbol{\xi}(\mathbf{w})\|_2^2$. If the game is Hamiltonian then (i) $\nabla \mathcal{H} = \mathbf{A}^\top \boldsymbol{\xi}$ and (ii) $\boldsymbol{\xi}$ preserves the level*

²The parameters in bimatrix games are usually constrained to the probability simplex.

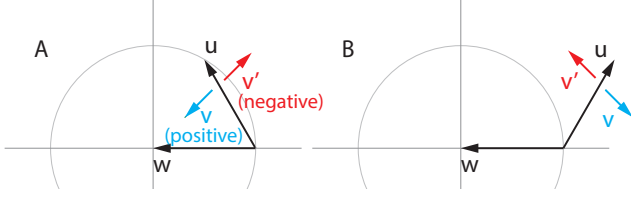


Figure 2. Infinitesimal alignment is positive (cyan) when small positive λ either: (A) pulls \mathbf{u} toward \mathbf{w} , if \mathbf{w} and \mathbf{u} have angle $< 90^\circ$; or (B) pushes \mathbf{u} away from \mathbf{w} if their angle is $> 90^\circ$.

sets of \mathcal{H} since $\langle \xi, \nabla \mathcal{H} \rangle = 0$. If the Hessian is invertible and $\lim_{\|\mathbf{w}\| \rightarrow \infty} \mathcal{H}(\mathbf{w}) = \infty$ then (iii) gradient descent on \mathcal{H} converges to a local Nash equilibrium.

In fact, \mathcal{H} is a Hamiltonian³ function for the game dynamics. We use the notation $\mathcal{H}(\mathbf{w}) = \frac{1}{2} \|\xi(\mathbf{w})\|^2$ throughout the paper. However, \mathcal{H} is only a Hamiltonian function for ξ if the game is Hamiltonian.

There is a precise mapping from Hamiltonian games to symplectic geometry (see Appendix D). Symplectic geometry is the modern formulation of classical mechanics. Recall that periodic behaviors (e.g. orbits) often arise in classical mechanics. The orbits lie on the level sets of the Hamiltonian, which expresses the total energy of the system.

3. Algorithms

Fixed points of potential and Hamiltonian games can be found by descent on ξ and $\nabla \mathcal{H}$ respectively. This section explores approaches to finding stable fixed points in general games.

3.1. Finding fixed points

If the Hessian $\mathbf{H}(\mathbf{w})$ is invertible, then $\nabla \mathcal{H} = \mathbf{H}^\top \xi = 0$ iff $\xi = 0$. Thus, gradient descent on \mathcal{H} converges to fixed points of ξ . Unfortunately, there is no guarantee that descent on \mathcal{H} will find a *stable* fixed point.

Mescheder et al. (2017) propose *consensus optimization*, a gradient adjustment of the form

$$\xi + \lambda \cdot \mathbf{H}^\top \xi = \xi + \lambda \cdot \nabla \mathcal{H}.$$

Unfortunately, consensus optimization can converge to unstable fixed points even in simple cases where the ‘game’ is to minimize a single function:

Example 5 (consensus optimization converges to global maximum). Consider a potential game with losses $\ell_1(x, y) = \ell_2(x, y) = -\frac{\kappa}{2}(x^2 + y^2)$ with $\kappa \gg 0$. Then

$$\xi = -\kappa \cdot \begin{pmatrix} x \\ y \end{pmatrix} \text{ and } \mathbf{H} = -\begin{pmatrix} \kappa & 0 \\ 0 & \kappa \end{pmatrix}$$

³Notation: \mathcal{H} for Hamiltonian; \mathbf{H} for Hessian.

Algorithm 1 Symplectic Gradient Adjustment

Input: losses $\mathcal{L} = \{\ell_i\}_{i=1}^n$, weights $\mathcal{W} = \{\mathbf{w}_i\}_{i=1}^n$
 $\xi \leftarrow [\text{gradient}(\ell_i, \mathbf{w}_i) \text{ for } (\ell_i, \mathbf{w}_i) \in (\mathcal{L}, \mathcal{W})]$
 $\mathbf{A}^\top \xi \leftarrow \text{get_sym_adj}(\mathcal{L}, \mathcal{W})$ // appendix B
if align then
 $\nabla \mathcal{H} \leftarrow [\text{gradient}(\frac{1}{2} \|\xi\|^2, \mathbf{w}) \text{ for } \mathbf{w} \in \mathcal{W}]$
 $\lambda \leftarrow \text{sign}\left(\frac{1}{d} \langle \xi, \nabla \mathcal{H} \rangle \langle \mathbf{A}^\top \xi, \nabla \mathcal{H} \rangle + \epsilon\right)$ // $\epsilon = \frac{1}{10}$
else
 $\lambda \leftarrow 1$
end if
Output: $\xi + \lambda \cdot \mathbf{A}^\top \xi$ // plug into any optimizer

Note that $\|\xi\|^2 = \kappa^2(x^2 + y^2)$ and

$$\xi + \lambda \cdot \mathbf{H}^\top \xi = \kappa(\lambda\kappa - 1) \cdot \begin{pmatrix} x \\ y \end{pmatrix}.$$

Descent on $\xi + \lambda \cdot \mathbf{H}^\top \xi$ converges to the global maximum $(x, y) = (0, 0)$ unless $\lambda < \frac{1}{\kappa}$.

Although consensus optimization works well in some special cases; it cannot be considered a candidate algorithm for finding stable fixed points in general games, since it fails in the basic case of potential games.

3.2. Finding stable fixed points

There are two classes of games where we know how to find stable fixed points: potential games where ξ converges to a local minimum and Hamiltonian games where $\nabla \mathcal{H}$, which is orthogonal to ξ , finds stable fixed points. In the general case, the following desiderata provide a set of reasonable properties for an adjustment ξ_λ of the game dynamics:

Desiderata. To find stable fixed points, an adjustment ξ_λ to the game dynamics should satisfy

- D1. compatible with game dynamics: $\langle \xi_\lambda, \xi \rangle = \alpha_1 \cdot \|\xi\|^2$;
- D2. compatible with potential dynamics: if the game is a potential game then $\langle \xi_\lambda, \nabla \phi \rangle = \alpha_2 \cdot \|\nabla \phi\|^2$;
- D3. compatible with Hamiltonian dynamics: If the game is Hamiltonian then $\langle \xi_\lambda, \nabla \mathcal{H} \rangle = \alpha_3 \cdot \|\nabla \mathcal{H}\|^2$;
- D4. attracted to local Nash equilibria: in neighborhoods where $\mathbf{S} \succ 0$, require $\theta(\xi_\lambda, \nabla \mathcal{H}) \leq \theta(\xi, \nabla \mathcal{H})$;
- D5. repelled by unstable equilibria: in neighborhoods where $\mathbf{S} \prec 0$, require $\theta(\xi_\lambda, \nabla \mathcal{H}) \geq \theta(\xi, \nabla \mathcal{H})$;

for some $\alpha_1, \alpha_2, \alpha_3 > 0$.

Note that desideratum D1 does not guarantee that players act in their own self-interest – this requires a stronger

positivity condition on dot-products with subvectors of ξ , see [Balduzzi \(2017\)](#). Desiderata $D4$ and $D5$ are explained in section 3.4. The unadjusted dynamics ξ satisfies all the desiderata except $D3$. Consensus optimization only satisfies $D3$ and $D4$. Ideally, desideratum $D5$ would be strengthened to ‘repelled by saddle points where $\xi(w) = 0$ but $S(w) \not\preceq 0$ ’.

3.3. Symplectic gradient adjustment

Proposition 4. *The symplectic gradient adjustment (SGA)*

$$\xi_\lambda := \xi + \lambda \cdot \mathbf{A}^\top \xi.$$

satisfies $D1$ – $D3$ for $\lambda > 0$, with $\alpha_1 = 1 = \alpha_2$ and $\alpha_3 = \lambda$.

Note that desiderata $D1$ and $D2$ are true even when $\lambda < 0$. This will prove useful, since example 6 and theorem 5 show it is necessary pick negative λ near $S \prec 0$. Section 3.4 shows how to also satisfy desiderata $D4$ and $D5$.

Proof. First claim: $\lambda \cdot \xi^\top \mathbf{A}^\top \xi = 0$ by skew-symmetry of \mathbf{A} . Second claim: $\mathbf{A} \equiv 0$ in a potential game, so $\xi_\lambda = \xi = \nabla \phi$. Third claim: $\langle \xi_\lambda, \nabla \mathcal{H} \rangle = \langle \xi_\lambda, \mathbf{H}^\top \xi \rangle = \langle \xi_\lambda, \mathbf{A}^\top \xi \rangle = \lambda \cdot \xi^\top \mathbf{A} \mathbf{A}^\top \xi = \lambda \cdot \|\nabla \mathcal{H}\|^2$ since $\mathbf{H} = \mathbf{A}$ by assumption and since $\xi^\top \mathbf{A}^\top \xi = 0$ by antisymmetry. \square

In the example below, almost any choice of positive λ results in convergence to an unstable equilibrium. The problem arises from the combination of a weak repeller with a strong rotational force. The next section shows how to pick λ to avoid unstable equilibria.

Example 6 (failure case for $\lambda > 0$). Suppose $\epsilon > 0$ is small and

$$f(x, y) = -\frac{\epsilon}{2}x^2 - xy \text{ and } g(x, y) = -\frac{\epsilon}{2}y^2 + xy$$

with an unstable equilibrium at $(0, 0)$. The dynamics are

$$\xi = \epsilon \cdot \begin{pmatrix} -x \\ -y \end{pmatrix} + \begin{pmatrix} -y \\ x \end{pmatrix} \text{ with } \mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

and

$$\mathbf{A}^\top \xi = \begin{pmatrix} x \\ y \end{pmatrix} + \epsilon \begin{pmatrix} -y \\ x \end{pmatrix}$$

Finally observe that

$$\xi + \lambda \cdot \mathbf{A}^\top \xi = (\lambda - \epsilon) \cdot \begin{pmatrix} x \\ y \end{pmatrix} + (1 + \epsilon\lambda) \cdot \begin{pmatrix} -y \\ x \end{pmatrix}$$

which converges to the unstable equilibrium if $\lambda > \epsilon$.

Lemma 8 in the appendix shows that, if \mathbf{S} and \mathbf{A} commute and $\mathbf{S} \succeq 0$, then $\langle \xi_\lambda, \nabla \mathcal{H} \rangle \geq 0$ for $\lambda > 0$. The proof of theorem 5 introduces the additive condition number to upper-bound the worst-case noncommutativity of \mathbf{S} , which allows to quantify the relationship between ξ_λ and $\nabla \mathcal{H}$.

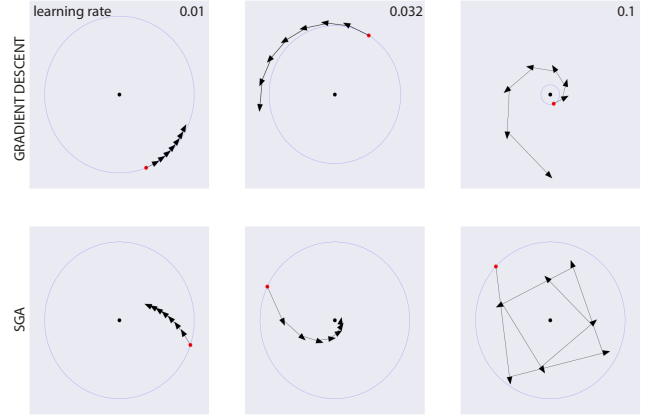


Figure 3. SGA allows faster and more robust convergence to stable fixed points. Note the scale of top-right panel differs from rest.

Theorem 5. Let \mathbf{S} be a symmetric matrix with eigenvalues $\sigma_{\max} \geq \dots \geq \sigma_{\min}$. The **additive condition number**⁴ of \mathbf{S} is $\kappa = \sigma_{\max} - \sigma_{\min}$. Suppose that $\mathbf{S} \succeq 0$ is positive semidefinite with additive condition number κ . Then

$$\langle \xi_\lambda, \nabla \mathcal{H} \rangle \geq 0 \text{ for } 0 < \lambda < \frac{4}{\kappa}.$$

If \mathbf{S} is negative semidefinite, then for λ in the same range

$$\langle \xi_{-\lambda}, \nabla \mathcal{H} \rangle \leq 0.$$

The inequalities are strict if \mathbf{H} is invertible.

If $\kappa = 0$, then $\mathbf{S} = \sigma \cdot \mathbf{I}$ commutes with *all* matrices. The larger κ , the larger the *potential* failure of \mathbf{S} to commute with other matrices. The analysis is pessimistic: \mathbf{S} and \mathbf{A} may approximately commute for large κ (e.g. if $\mathbf{A} \approx 0$).

3.4. How to pick $\text{sign}(\lambda)$

This section explains desiderata $D4$ – $D5$ and shows how to pick $\text{sign}(\lambda)$ to speed up convergence towards stable and away from unstable fixed points.

First, observe that $\langle \xi, \nabla \mathcal{H} \rangle = \xi^\top (\mathbf{S} + \mathbf{A})^\top \xi = \xi^\top \mathbf{S} \xi$. It follows that for $\xi \neq 0$:

$$\begin{cases} \text{if } \mathbf{S} \succeq 0 & \text{then } \langle \xi, \nabla \mathcal{H} \rangle \geq 0; \\ \text{if } \mathbf{S} \prec 0 & \text{then } \langle \xi, \nabla \mathcal{H} \rangle < 0. \end{cases} \quad (1)$$

A criterion to probe the positive/negative definiteness of \mathbf{S} is thus to check the sign of $\langle \xi, \nabla \mathcal{H} \rangle$. The dot product can take any value if \mathbf{S} is not positive nor negative definite. The behavior near saddles is beyond the scope of the paper.

Desiderata $D4$ can be interpreted as follows. If ξ points at a stable equilibrium then we require that ξ_λ points *more*

⁴The condition number of a positive definite matrix is $\frac{\sigma_{\max}}{\sigma_{\min}}$.

towards the equilibrium (i.e. has smaller angle). Conversely if ξ points away then the adjustment should point *further* away. More formally,

Definition 4. Let \mathbf{u} and \mathbf{v} be two vectors. The *infinitesimal alignment* of $\xi_\lambda := \mathbf{u} + \lambda \cdot \mathbf{v}$ with a third vector \mathbf{w} is

$$\text{align}(\xi_\lambda, \mathbf{w}) := \frac{d}{d\lambda} \{ \cos^2 \theta_\lambda \}_{|\lambda=0} \text{ for } \theta_\lambda := \theta(\xi_\lambda, \mathbf{w}).$$

If \mathbf{u} and \mathbf{w} point the same way, $\mathbf{u}^\top \mathbf{w} > 0$, then $\text{align} > 0$ when \mathbf{v} bends \mathbf{u} further toward \mathbf{w} , see figure 2A. Otherwise $\text{align} > 0$ when \mathbf{v} bends \mathbf{u} away from \mathbf{w} , figure 2B.

Proposition 6. Desiderata D4–5 are satisfied for λ such that $\lambda \cdot \langle \xi, \nabla \mathcal{H} \rangle \cdot \langle \mathbf{A}^\top \xi, \nabla \mathcal{H} \rangle \geq 0$.

Computing the sign of $\langle \xi, \nabla \mathcal{H} \rangle$ provides a check for stable and unstable fixed points. Computing the sign of $\langle \mathbf{A}^\top \xi, \nabla \mathcal{H} \rangle$ checks whether the adjustment term points towards or away from the (nearby) fixed point. Putting the two checks together yields a prescription for the sign of λ .

Alignment and convergence rates. Finally, we show that increasing alignment helps speed convergence. Gradient descent is also known as the method of steepest descent. In general games, however, ξ does not follow the steepest path to fixed points due to the ‘rotational force’, which forces lower learning rates and slows down convergence:

Theorem 7. Suppose f is convex and Lipschitz smooth with $\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\| \leq L \cdot \|\mathbf{x} - \mathbf{y}\|$. Let $\mathbf{w}_{t+1} = \mathbf{w}_t - \eta \cdot \mathbf{v}$ where $\|\mathbf{v}\| = \|\nabla f(\mathbf{w}_t)\|$. Then the optimal step size is $\eta^* = \frac{\cos \theta}{L}$ where $\theta := \theta(\nabla f(\mathbf{w}_t), \mathbf{v})$, with

$$f(\mathbf{w}_{t+1}) \leq f(\mathbf{w}_t) - \frac{\cos^2 \theta}{2L} \cdot \|\nabla f(\mathbf{w}_t)\|^2.$$

Increasing the cosine with the steepest direction improves convergence. The alignment computation in algorithm 1 chooses λ to be positive or negative such that ξ_λ is bent towards stable fixed points (increasing the cosine) and away from unstable fixed points. Adding a small $\epsilon > 0$ to the computation introduces a weak bias towards attractors.

3.5. Aligned consensus optimization

The stability criterion in (1) also provides a simple way to prevent consensus optimization from converging to unstable equilibria. **Aligned consensus optimization** is

$$\xi + |\lambda| \cdot \text{sign} \left(\langle \xi, \nabla \mathcal{H} \rangle \right) \cdot \mathbf{H}^\top \xi, \quad (2)$$

where in practice we set $\lambda = 1$. Aligned consensus satisfies desiderata D3–D5. However, it behaves strangely in potential games. Multiplying by the Hessian is the ‘inverse’ of Newton’s method: it increases the gap between small and

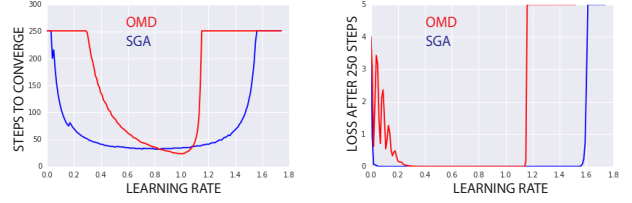


Figure 4. Comparison of SGA with optimistic mirror descent. Sweep over learning rates in range $[0.01, 1.75]$; $\lambda = 1$ throughout. (A) iterations to convergence (up to 250). (B): average absolute value of losses over iterations 240–250, cutoff at 5.

large eigenvalues, increasing the (usual, multiplicative) condition number and slows down convergence. Nevertheless, consensus optimization works well in GANs (Mescheder et al., 2017), and aligned consensus may improve performance, see figure 9 in appendix.

Finally, note that dropping the first term ξ from (2) yields a simpler update that also satisfies D3–D5. However, the resulting algorithm performs poorly in experiments (not shown), perhaps because it is attracted to saddles.

4. Experiments

We compare SGA with simultaneous gradient descent, optimistic mirror descent (Daskalakis et al., 2018) and consensus optimization (Mescheder et al., 2017) in basic settings.

4.1. Learning rates and alignment

We investigate the effect of the SGA when a weak attractor is coupled to a strong rotational force:

$$\ell_1(x, y) = \frac{1}{2}x^2 + 10xy \quad \text{and} \quad \ell_2(x, y) = \frac{1}{2}y^2 - 10xy$$

Gradient descent is extremely sensitive to the choice of learning rate η , top row of figure 3. As η increases through $\{0.01, 0.032, 0.1\}$ gradient descent goes from converging extremely slowly, to diverging slowly, to diverging rapidly. SGA yields faster, more robust convergence. SGA converges faster with learning rates $\eta = 0.01$ and $\eta = 0.032$, and only starts overshooting the fixed point for $\eta = 0.1$.

4.2. Basic adversarial games

Figure 4 compares SGA with optimistic mirror descent on a zero-sum bimatrix game with $\ell_{1/2}(\mathbf{w}_1, \mathbf{w}_2) = \pm \mathbf{w}_1^\top \mathbf{w}_2$. The example is modified from Daskalakis et al. (2018) who also consider a linear offset that makes no difference. A run converges, panel A, if the average absolute value of losses on the last 10 iterations is < 0.01 .

Although OMD’s peak performance is better than SGA, we find that SGA converges – and does so faster – for a wider range of learning rates. OMD diverges for learning rates

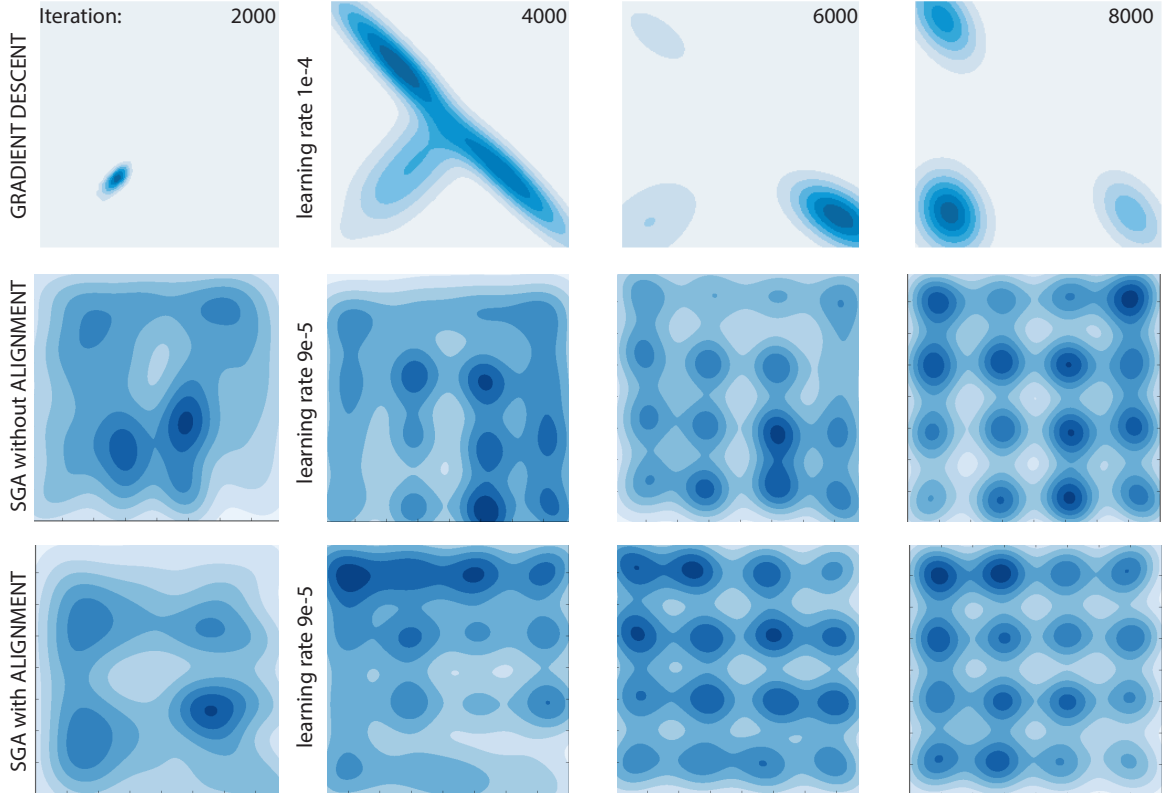


Figure 5. **Top:** Simultaneous gradient descent suffers from mode collapse and in later iterations (not shown) mode hopping. **Middle:** vanilla SGA converges smoothly to the ground truth (figure 8 in appendix). **Bottom:** SGA with alignment converges slightly faster.

not in the range $[0.3, 1.2]$. Simultaneous gradient descent oscillates without converging (not shown). Individual runs are shown in the appendix. The appendix also compares the performance of the algorithms on four-player games.

4.3. Generative adversarial networks

We apply SGA to a basic setup adapted from Metz et al. (2017). Data is sampled from a highly multimodal distribution designed to probe the tendency to collapse onto a subset of modes during training. The distribution is a mixture of 16 Gaussians arranged in a 4×4 grid, see ground truth in figure 8. The generator and discriminator networks both have 6 ReLU layers of 384 neurons. The generator has two output neurons; the discriminator has one.

Figure 5 shows results after $\{2000, 4000, 6000, 8000\}$ iterations. The networks are trained under RMSProp. Learning rates were chosen by visual inspection of grid search results at iteration 8000, see appendix. Simultaneous gradient descent and SGA are shown in the figure. Results for consensus optimization are in the appendix.

Simultaneous gradient descent exhibits mode collapse followed by mode hopping in later iterations (not shown). Mode hopping is analogous to the cycles in example 1. Un-

aligned SGA converges to the correct distribution; alignment speeds up convergence slightly. Consensus optimization performs similarly, see figure 9 in the appendix. However, it can converge to local maxima, recall example 5.

5. Discussion

Modern deep learning treats differentiable modules like plug-and-play lego blocks. For this to work, at the very least, we need to know that gradient descent will find local minima. Unfortunately, gradient descent does *not* necessarily find local minima when optimizing multiple interacting objectives. With the recent proliferation of algorithms that optimize more than one loss, it is becoming increasingly urgent to understand and control the dynamics of interacting losses. Although there is interesting recent work on two-player adversarial games such as GANs, there is essentially no work on finding stable fixed points in general.

The generalized Helmholtz decomposition provides a powerful new perspective on game dynamics. A key feature is that the analysis is indifferent to the number of players. Instead, it is the interplay between the simultaneous gradient ξ on the losses and the symmetric and antisymmetric matrices of second-order terms that guides algorithm design and governs the dynamics under gradient adjustments.

Symplectic gradient adjustment is a straightforward application of the Helmholtz decomposition. It is unlikely that SGA is the only or best approach to finding stable fixed points. A deeper understanding of the interaction between the potential and Hamiltonian components will lead to more effective algorithms. Of particular interest are pure first-order methods that do not use Hessian-vector products.

Finally, it is worth raising a philosophical point. The goal in this paper is to find stable fixed points (e.g. because in GANs they yield pleasing samples). We are not concerned with the losses of the players *per se*. The gradient adjustments may lead to a player acting against its own self-interest by increasing its loss. We consider this acceptable insofar as it encourages convergence to a local Nash equilibrium.

Acknowledgements. We thank Guillaume Desjardins and Csaba Szepesvari for useful feedback.

References

- Abdallah, Sherief and Lesser, Victor R. A multiagent reinforcement learning algorithm with non-linear dynamics. *J. Artif. Intell. Res.*, 33:521–549, 2008.
- Balduzzi, D. Strongly-Typed Agents are Guaranteed to Interact Safely. In *ICML*, 2017.
- Bottacin, Francesco. A Marsden-Weinstein Reduction Theorem for Presymplectic Manifolds. 2005.
- Bowling, Michael and Veloso, Manuela. Multiagent learning using a variable learning rate. *Artificial Intelligence*, 136:215–250, 2002.
- Bowling, Michael H. Convergence and no-regret in multiagent learning. In *NIPS*, pp. 209–216, 2004.
- Daskalakis, C, Goldberg, P W, and Papadimitriou, C. The Complexity of Computing a Nash Equilibrium. *SIAM J. Computing*, 39(1):195–259, 2009.
- Daskalakis, C, Ilyas, A, Syrgkanis, V, and Zeng, H. Training GANs with Optimism. In *ICLR*, 2018.
- Dauphin, Y, Pascanu, R, Gulcehre, C, Cho, K, Ganguli, S, and Bengio, Y. Identifying and attacking the saddle point problem in high-dimensional non-convex optimization. In *NIPS*, 2014.
- Foerster, J N, Chen, R Y, Al-Shedivat, M, Whiteson, S, Abbeel, P, and Mordatch, I. Learning with Opponent-Learning Awareness. In *AAMAS*, 2018.
- Goodfellow, I J, Pouget-Abadie, J, Mirza, M, Xu, B, Warde-Farley, D, Ozair, S, Courville, A, and Bengio, Y. Generative Adversarial Nets. In *NIPS*, 2014.
- Hart, Sergiu and Mas-Colell, Andreu. *Simple Adaptive Strategies: From Regret-Matching to Uncoupled Dynamics*. World Scientific, 2013.
- Heusel, M, Ramsauer, H, Unterthiner, T, Nessler, B, Klambauer, G, and Hochreiter, S. GANs Trained by a Two Time-Scale Update Rule Converge to a Nash Equilibrium. In *NIPS*, 2017.
- Hu, Junling and Wellman, Michael P. Nash Q-Learning for General-Sum Stochastic Games. *Journal of Machine Learning Research*, 4:1039–1069, 2003.
- Jaderberg, M, Czarnecki, W M, Osindero, S, Vinyals, O, Graves, A, and Kavukcuoglu, K. Decoupled Neural Interfaces using Synthetic Gradients. In *ICML*, 2017.
- Lee, Jason D, Simchowitz, Max, Jordan, Michael I, and Recht, Benjamin. Gradient Descent Converges to Minimizers. In *COLT*, 2016.
- Lee, JD, Panageas, I, Piliouras, G, Simchowitz, M, Jordan, MI, and Recht, B. First-order Methods Almost Always Avoid Saddle Points. In *arXiv:1710.07406*, 2017.
- Liu, Bo, Liu, Ji, Ghavamzadeh, Mohammad, Mahadevan, Sridhar, and Petrik, Marek. Proximal Gradient Temporal Difference Learning Algorithms. In *IJCAI*, 2016.
- Lu, Xiaoyun. Hamiltonian games. *Journal of Combinatorial Theory, Series B*, 55:18–32, 1992.
- Mertikopoulos, Panayotis, Papadimitriou, Christos, and Piliouras, Georgios. Cycles in adversarial regularized learning. In *SODA*, 2018.
- Mescheder, Lars. On the convergence properties of GAN training. In *ArXiv:1801.04406*, 2018.
- Mescheder, Lars, Nowozin, Sebastian, and Geiger, Andreas. The Numerics of GANs. In *NIPS*. 2017.
- Metz, L, Poole, B, Pfau, D, and Sohl-Dickstein, J. Unrolled generative adversarial networks. In *ICLR*, 2017.
- Monderer, Dov and Shapley, Lloyd S. Potential Games. *Games and Economic Behavior*, 14:124–143, 1996.
- Nagarajan, Vaishnavh and Kolter, J Zico. Gradient descent GAN optimization is locally stable. In *NIPS*, 2017.
- Nesterov, Yurii. *Introductory Lectures on Convex Optimization: A Basic Course*. Kluwer, 2004.
- Papadimitriou, Christos and Piliouras, Georgios. From Nash Equilibria to Chain Recurrent Sets: Solution Concepts and Topology. In *ITCS*, 2016.

Pathak, D, Agrawal, P, Efros, A A, and Darrell, T. Curiosity-driven Exploration by Self-supervised Prediction. In *ICML*, 2017.

Pfau, David and Vinyals, Oriol. Connecting Generative Adversarial Networks and Actor-Critic Methods. In *arXiv:1610.01945*, 2016.

Racanière, S, Weber, T, Reichert, D P, Buesing, L, Guez, A, Rezende, D J, Badia, A P, Vinyals, O, Heess, N, Li, Y, Pascanu, R, Battaglia, P, Hassabis, D, Silver, D, and Wierstra, D. Imagination-Augmented Agents for Deep Reinforcement Learning. In *NIPS*, 2017.

Rosenthal, R W. A Class of Games Possessing Pure-Strategy Nash Equilibria. *Int J Game Theory*, 2:65–67, 1973.

Salimans, Tim, Goodfellow, Ian, Zaremba, Wojciech, Cheung, Vicki, Radford, Alec, and Chen, Xi. Improved Techniques for Training GANs. In *NIPS*, 2016.

Shoham, Yoav and Leyton-Brown, Kevin. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2008.

Singh, S, Kearns, M, and Mansour, Y. Nash Convergence of Gradient Dynamics in General-Sum Games. In *UAI*, 2000.

Stoltz, G and Lugosi, G. Learning correlated equilibria in games with compact sets of strategies. *Games and Economic Behavior*, 59:187–208, 2007.

Vezhnevets, A, Osindero, S, Schaul, T, Heess, N, Jaderberg, M, Silver, D, and Kavukcuoglu, K. FeUdal Networks for Hierarchical Reinforcement Learning. In *ICML*, 2017.

Wayne, G and Abbott, L F. Hierarchical Control Using Networks Trained with Higher-Level Forward Models. *Neural Computation*, (26), 2014.

APPENDIX

A. Proofs

Proof of theorem 3.

Proof. Direct computation shows $\nabla \mathcal{H} = \mathbf{H}^\top \xi$ for any game. The first statement follows since $\mathbf{H} = \mathbf{A}$ in Hamiltonian games. For the second statement, the directional derivative is $D_\xi \mathcal{H} = \langle \xi, \nabla \mathcal{H} \rangle = \xi^\top \mathbf{A}^\top \xi$ where $\xi^\top \mathbf{A}^\top \xi = (\xi^\top \mathbf{A}^\top \xi)^\top = \xi^\top \mathbf{A} \xi = -(\xi^\top \mathbf{A}^\top \xi)$ since $\mathbf{A} = -\mathbf{A}^\top$ by skew-symmetry. It follows that $\xi^\top \mathbf{A}^\top \xi = 0$.

For the third, gradient descent on \mathcal{H} will converge to a point where $\nabla \mathcal{H} = \mathbf{H}^\top \xi(\mathbf{w}) = 0$. If the Hessian is invertible then clearly $\xi(\mathbf{w}) = 0$. The fixed-point is a local Nash

equilibrium by Lemma 2 since $0 \equiv \mathbf{S} \succeq 0$ in a Hamiltonian game. \square

Proof of theorem 5. The following lemma, whilst not used in the proof of theorem 5, nevertheless provides a useful intuition about the role of commutativity.

Lemma 8. *If $\mathbf{S} \succeq 0$ is symmetric positive semidefinite and \mathbf{S} commutes with \mathbf{A} then*

$$\langle \xi_\lambda, \nabla \mathcal{H} \rangle \geq 0 \text{ for all } \lambda \geq 0.$$

Proof. First observe that $\xi^\top \mathbf{A} \mathbf{S} \xi = \xi^\top \mathbf{S}^\top \mathbf{A}^\top \xi = -\xi^\top \mathbf{S} \mathbf{A} \xi$, where the first equality holds since the expression is a scalar, and the second holds since $\mathbf{S} = \mathbf{S}^\top$ and $\mathbf{A} = -\mathbf{A}^\top$. It follows that $\xi^\top \mathbf{A} \mathbf{S} \xi = 0$ if $\mathbf{S} \mathbf{A} = \mathbf{A} \mathbf{S}$.

Rewrite the inequality as

$$\begin{aligned} \langle \xi_\lambda \cdot \mathbf{A}^\top \xi, \nabla \mathcal{H} \rangle &= \langle \xi + \lambda \cdot \mathbf{A}^\top \xi, \mathbf{S} \xi + \mathbf{A}^\top \xi \rangle \\ &= \xi^\top \mathbf{S} \xi + \lambda \xi^\top \mathbf{A} \mathbf{A}^\top \xi \geq 0 \end{aligned}$$

since $\xi^\top \mathbf{A} \mathbf{S} \xi = 0$ and by positivity of \mathbf{S} , λ and $\mathbf{A} \mathbf{A}^\top$. \square

The proof of the theorem follows.

Proof. We prove the case $\mathbf{S} \succeq 0$; the case $\mathbf{S} \preceq 0$ is similar. Rewrite the inequality as

$$\begin{aligned} \langle \xi + \lambda \cdot \mathbf{A}^\top \xi, \nabla \mathcal{H} \rangle &= (\xi + \lambda \cdot \mathbf{A}^\top \xi)^\top \cdot (\mathbf{S} + \mathbf{A}^\top) \xi \\ &= \xi^\top \mathbf{S} \xi + \lambda \xi^\top \mathbf{A} \mathbf{S} \xi + \lambda \xi^\top \mathbf{A} \mathbf{A}^\top \xi \end{aligned}$$

Let $\beta = \|\mathbf{A}^\top \xi\|$ and $\tilde{\mathbf{S}} = \mathbf{S} - \sigma_{\min} \cdot \mathbf{I}$, where \mathbf{I} is the identity matrix. Then

$$\xi^\top \mathbf{S} \xi + \lambda \xi^\top \mathbf{A} \mathbf{S} \xi + \lambda \cdot \beta^2 \geq \xi^\top \tilde{\mathbf{S}} \xi + \lambda \xi^\top \mathbf{A} \tilde{\mathbf{S}} \xi + \lambda \cdot \beta^2$$

since $\xi^\top \mathbf{S} \xi \geq \xi^\top \tilde{\mathbf{S}} \xi$ by construction and $\xi^\top \mathbf{A} \tilde{\mathbf{S}} \xi = \xi^\top \mathbf{A} \mathbf{S} \xi - \sigma_{\min} \xi^\top \mathbf{A} \xi = \xi^\top \mathbf{A} \mathbf{S} \xi$ because $\xi^\top \mathbf{A} \xi = 0$ by the skew-symmetry of \mathbf{A} . It therefore suffices to show that the inequality holds when $\sigma_{\min} = 0$ and $\kappa = \sigma_{\max}$.

Since \mathbf{S} is positive semidefinite, there exists an upper-triangular square-root matrix T such that $T^\top T = \mathbf{S}$ and so $\xi^\top \mathbf{S} \xi = \|T\xi\|^2$. Further,

$$|\xi^\top \mathbf{A} \mathbf{S} \xi| \leq \|\mathbf{A}^\top \xi\| \cdot \|T^\top T \xi\| \leq \sqrt{\sigma_{\max}} \cdot \|\mathbf{A}^\top \xi\| \cdot \|T\xi\|.$$

since $\|T\|_2 = \sqrt{\sigma_{\max}}$. Putting the observations together obtains

$$\begin{aligned} &\|T\xi\|^2 + \lambda(\|\mathbf{A}\xi\|^2 - \langle \mathbf{A}\xi, \mathbf{S}\xi \rangle) \\ &\geq \|T\xi\|^2 + \lambda(\|\mathbf{A}\xi\|^2 - \|\mathbf{A}\xi\| \|\mathbf{S}\xi\|) \\ &\geq \|T\xi\|^2 + \lambda\|\mathbf{A}\xi\|(\|\mathbf{A}\xi\| - \|\mathbf{S}\xi\|) \\ &\geq \|T\xi\|^2 + \lambda\|\mathbf{A}\xi\|(\|\mathbf{A}\xi\| - \sqrt{\sigma_{\max}} \|T\xi\|) \end{aligned}$$

Set $\alpha = \sqrt{\lambda}$ and $\eta = \sqrt{\sigma_{max}}$. We can continue the above computation

$$\begin{aligned}
 &= \|\mathbf{T}\xi\|^2 + \alpha^2 \|\mathbf{A}\xi\|(\|\mathbf{A}\xi\| - \eta \|\mathbf{T}\xi\|) \\
 &= \|\mathbf{T}\xi\|^2 + \alpha^2 \|\mathbf{A}\xi\|^2 - \alpha^2 \|\mathbf{A}\xi\| \eta \|\mathbf{T}\xi\| \\
 &= (\|\mathbf{T}\xi\| - \alpha \|\mathbf{A}\xi\|)^2 + 2\alpha \|\mathbf{A}\xi\| \|\mathbf{T}\xi\| - \alpha^2 \eta \|\mathbf{A}\xi\| \|\mathbf{T}\xi\| \\
 &= (\|\mathbf{T}\xi\| - \alpha \|\mathbf{A}\xi\|)^2 + \|\mathbf{A}\xi\| \|\mathbf{T}\xi\| (2\alpha - \alpha^2 \eta)
 \end{aligned}$$

Finally, $2\alpha - \alpha^2 \eta > 0$ for any α in the range $(0, \frac{2}{\eta})$, which is to say, for any $0 < \lambda < \frac{4}{\sigma_{max}}$. The kernel of \mathbf{S} and the kernel of \mathbf{T} coincide. If ξ is in the kernel of \mathbf{A} , resp. \mathbf{T} , it cannot be in the kernel of \mathbf{T} , resp. \mathbf{A} and the term $(\|\mathbf{T}\xi\| - \alpha \|\mathbf{A}\xi\|)^2$ is positive. Otherwise, the term $\|\mathbf{A}\xi\| \|\mathbf{T}\xi\|$ is positive. \square

Proof of proposition 6.

Lemma 9. When ξ_λ is the symplectic gradient adjustment,

$$\text{sign} \left(\text{align}(\xi_\lambda, \nabla \mathcal{H}) \right) = \text{sign} \left(\langle \xi, \nabla \mathcal{H} \rangle \cdot \langle \mathbf{A}^\top \xi, \nabla \mathcal{H} \rangle \right).$$

Proof. Observe that

$$\begin{aligned}
 \cos^2 \theta_\lambda &= \left(\frac{\langle \xi_\lambda, \nabla \mathcal{H} \rangle}{\|\xi_\lambda\| \cdot \|\nabla \mathcal{H}\|} \right)^2 \\
 &= \frac{\langle \xi, \nabla \mathcal{H} \rangle + 2\lambda \langle \xi, \nabla \mathcal{H} \rangle \langle \mathbf{A}^\top \xi, \nabla \mathcal{H} \rangle + O(\lambda^2)}{(\|\xi\|^2 + O(\lambda^2)) \cdot \|\nabla \mathcal{H}\|^2}
 \end{aligned}$$

where the denominator has no linear term in λ because $\xi \perp \mathbf{A}^\top \xi$. It follows that the sign of the infinitesimal alignment is

$$\text{sign} \left\{ \frac{d}{d\lambda} \cos^2 \theta_\lambda \right\} = \text{sign} \left\{ \langle \xi, \nabla \mathcal{H} \rangle \langle \mathbf{A}^\top \xi, \nabla \mathcal{H} \rangle \right\}$$

as required. \square

The proposition follows easily.

Proof. If we are in a neighborhood of a stable fixed point then $\langle \xi, \nabla \mathcal{H} \rangle \geq 0$. It follows by lemma 9 that $\text{sign} \left(\text{align}(\xi_\lambda, \nabla \mathcal{H}) \right) = \text{sign} \left(\langle \mathbf{A}^\top \xi, \nabla \mathcal{H} \rangle \right)$ and so choosing $\text{sign}(\lambda) = \text{sign} \left(\langle \mathbf{A}^\top \xi, \nabla \mathcal{H} \rangle \right)$ leads to the angle between ξ_λ and $\nabla \mathcal{H}$ being smaller than the angle between ξ and $\nabla \mathcal{H}$, satisfying desideratum D4.

The proof for the unstable case is similar. \square

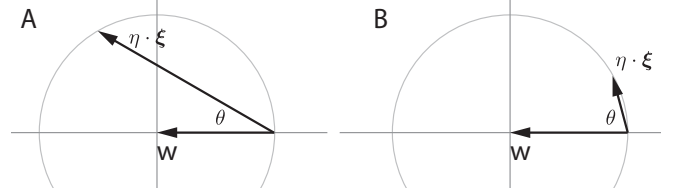


Figure 6. Alignment and learning rates. The larger $\cos \theta$, the larger the learning rate η that can be applied to unit vector ξ without $w + \eta \cdot \xi$ leaving the unit circle.

Proof of theorem 7. The following lemma provides some intuition for theorem 7. The idea is that, the smaller the cosine between the ‘correct direction’ w and the ‘update direction’ ξ , the smaller the learning rate needs to be for the update to stay in a unit ball, see figure 6.

Lemma 10 (alignment lemma). If w and ξ are unit vectors with $0 < w^\top \xi$ then $\|w - \eta \cdot \xi\| \leq 1$ for $0 \leq \eta \leq 2w^\top \xi$.

Proof. Check $\|w - \eta \cdot \xi\|^2 = 1 + \eta^2 - 2\eta \cdot w^\top \xi \leq 1$ iff $\eta^2 \leq 2\eta \cdot w^\top \xi$. The result follows. \square

The next lemma is a standard technical result (e.g. Nesterov, 2004).

Lemma 11. Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a convex Lipschitz smooth function satisfying $\|\nabla f(y) - \nabla f(x)\| \leq L \cdot \|y - x\|$ for all $x, y \in \mathbb{R}^d$. Then

$$|f(y) - f(x) - \langle \nabla f(x), y - x \rangle| \leq \frac{L}{2} \cdot \|y - x\|^2$$

for all $x, y \in \mathbb{R}^d$.

Finally the proof of the theorem is adapted from a standard result in Nesterov (2004), where there is no cosine.

Proof. By the above lemma,

$$\begin{aligned}
 f(y) &\leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|y - x\|^2 \\
 &= f(x) - \eta \cdot \langle \nabla f, \xi \rangle + \eta^2 \frac{L}{2} \cdot \|\xi\|^2 \\
 &= f(x) - \eta \cdot \langle \nabla f, \xi \rangle + \eta^2 \frac{L}{2} \cdot \|\nabla f\|^2 \\
 &= f(x) - \eta(\alpha - \frac{\eta}{2}L) \cdot \|\nabla f\|^2
 \end{aligned}$$

where $\alpha := \cos \theta$. Solve

$$\min_{\eta} \Delta(\eta) = \min_{\eta} \left\{ -\eta(\alpha - \frac{\eta}{2}L) \right\}$$

to obtain $\eta^* = \frac{\alpha}{L}$ and $\Delta(\eta^*) = -\frac{\alpha^2}{2}L$ as required. \square

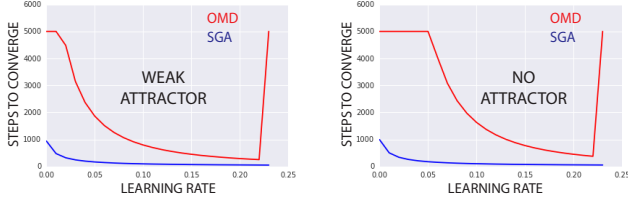


Figure 7. Time to convergence of OMD and SGA on two 4-player games. Times are cutoff after 5000 iterations. **Left panel:** Weak attractor corresponding to potential function $\phi(w, x, y, z) = \frac{1}{200}(w^2 + x^2 + y^2 + z^2)$. **Right panel:** No potential function.

B. TensorFlow code to compute $A^T \xi$:

The code requires a list of n losses, Ls , and a list of variables for the n players, xs . The function `fwd_gradients` is in the module `tf.contrib.kfac.utils`.

```
def jac_vec(ys, xs, vs):
    return fwd_gradients(ys, xs,
                        gradxs = vs, stop_gradients = xs)

def jac_tran_vec(ys, xs, vs):
    dydxs = tf.gradients(ys, xs, grad_ys = vs,
                        stop_gradients = xs)
    return [tf.zeros_like(x) if dydx is None
            else dydx
            for (x, dydx) in zip(xs, dydxs)]

def get_sym_adj(Ls, xs):
    xi = [tf.gradients(l, x)[0]
           for (l, x) in zip(Ls, xs)]
    H_xi = jac_vec(xi, xs, xi)
    Ht_xi = jac_tran_vec(xi, xs, xi)
    At_xi = [ht-h
              for (h, ht) in zip(H_xi, Ht_xi)]
    return At_xi
```

C. Further experiments

OMD and SGA on a four-player game. Figure 7 shows time to convergence (same criterion as section 4.2 for optimistic mirror descent and SGA). The games are constructed with four players, each of which controls one parameter. The game was setup to have antisymmetric and symmetric components of the Hessian as follows. The antisymmetric component of the game Hessian is

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & 0 \end{pmatrix}$$

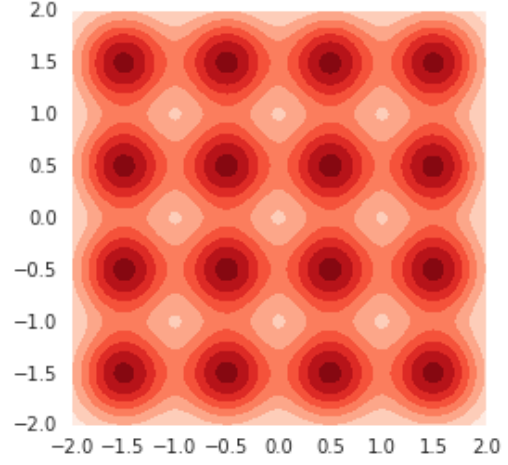


Figure 8. Ground truth for GAN experiments. A mixture of 16 Gaussians.

in both panels. The symmetric component is

$$S = \frac{1}{100} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

in the left panel and $S \equiv 0$ in the right panel.

OMD converges considerably slower than SGA across the full range of learning rates. It also diverges for learning rates > 0.22 . SGA converges more quickly and robustly.

Ground truth for GAN experiments. Figure 8 shows the probability distribution that is sampled to train the generator and discriminator in the GAN example.

More details on GAN experiments. Figure 9 shows the performance of consensus optimization without and with alignment. Introducing alignment slightly improves speed of convergence (second column) and final result (fourth column), although intermediate results in third column are ambiguous.

Grid search was over learning rates $\{1e-5, 2e-5, 5e-5, 8e-5, 1e-4, 2e-4, 5e-4\}$ and then a more refined linear search over $[8e-5, 2e-4]$.

D. Helmholtz and Hamilton

This section explains the mathematical connections with the Helmholtz decomposition and symplectic geometry. The discussion is *not* necessary to understand the main text. It is also not self-contained. The details can be found in textbooks covering differential and symplectic geometry.

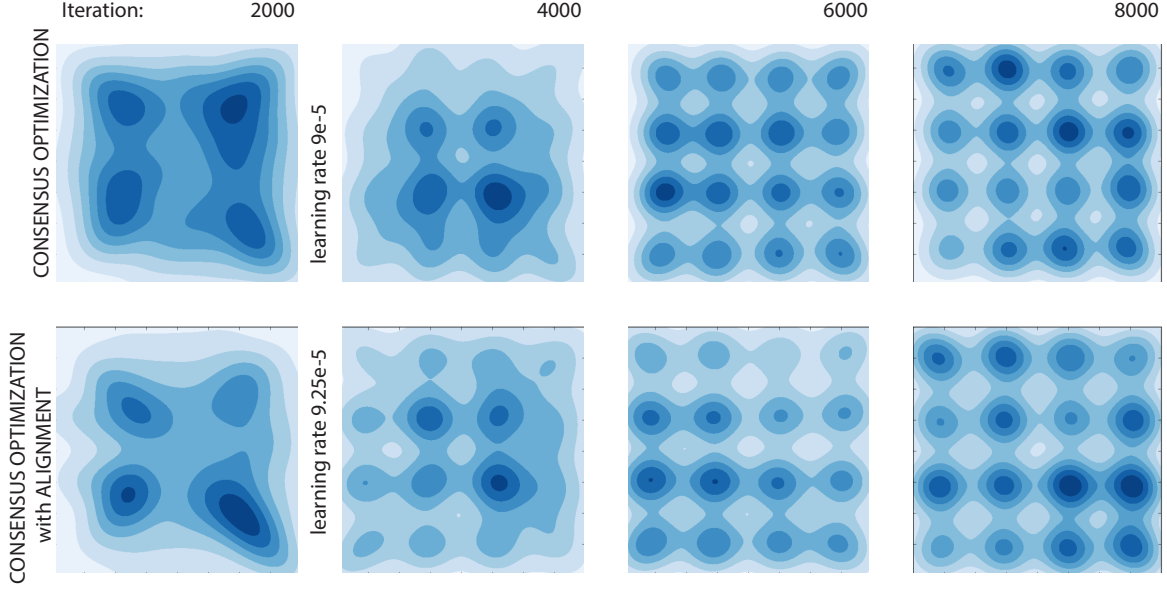


Figure 9. **Top:** Consensus optimization on GAN example. **Bottom:** Consensus optimization with alignment.

D.1. The Helmholtz decomposition

The classical Helmholtz decomposition states that any vector field ξ in 3-dimensions is the sum of curl-free (gradient) and divergence-free (infinitesimal rotation) components:

$$\xi = \underbrace{\nabla\phi}_{\text{gradient component}} + \underbrace{\text{curl}(\mathbf{B})}_{\text{rotational component}} \quad \left[\text{curl}(\bullet) := \nabla \times (\bullet) \right]$$

We explain the link between curl and the antisymmetric component of the game Hessian. Recall that gradients of functions are actually differential 1-forms, not vector fields. Differential 1-forms and vector fields on a manifold are canonically isomorphic once a Riemannian metric has been chosen. In our case, we are implicitly using the Euclidean metric. The antisymmetric matrix \mathbf{A} is the differential 2-form obtained by applying the exterior derivative d to the 1-form ξ .

In 3-dimensions, the Hodge star operator is an isomorphism from differential 2-forms to vector fields, and the curl can be reformulated as $\text{curl}(\bullet) = *d(\bullet)$. In claiming \mathbf{A} is analogous to curl, we are simply dropping the Hodge-star operator.

Finally, recall that the Lie algebra of infinitesimal rotations in d -dimensions is given by antisymmetric matrices. When $d = 3$, the Lie algebra can be represented as vectors (three numbers specify a 3×3 antisymmetric matrix) with the \times -product as Lie bracket. In general, the antisymmetric matrix \mathbf{A} captures the infinitesimal tendency of ξ to rotate at each point in the parameter space.

D.2. Hamiltonian mechanics

A symplectic form ω is a closed nondegenerate differential 2-form. Given a manifold with a symplectic form, a vector field ξ is **Hamiltonian vector field** if there exists a function $\mathcal{H} : M \rightarrow \mathbb{R}$ satisfying

$$\omega(\xi, \bullet) = d\mathcal{H}(\bullet) = \langle \nabla \mathcal{H}, \bullet \rangle. \quad (3)$$

The function is then referred to as the Hamiltonian function of the vector field. In our case, the antisymmetric matrix \mathbf{A} is a closed 2-form because $\mathbf{A} = d\xi$ and $d \circ d = 0$. It may however be degenerate. It is therefore a presymplectic form (Bottacin, 2005).

Setting $\omega = \mathbf{A}$, equation (3) can be rewritten in our notation as

$$\underbrace{\omega(\xi, \bullet)}_{\mathbf{A}^\top \xi} = \underbrace{d\mathcal{H}(\bullet)}_{\nabla \mathcal{H}},$$

justifying the terminology ‘Hamiltonian’ and ‘symplectic’.

E. Type consistency

The next two sections carefully work through the units in classical mechanics and two-player games respectively. The third section briefly describes a use-case for type consistency.

E.1. Units in classical mechanics

Consider the well-known Hamiltonian

$$\mathcal{H}(p, q) = \frac{1}{2} \left(\kappa \cdot q^2 + \frac{1}{\mu} \cdot p^2 \right)$$

where q is position, $p = \mu \cdot \dot{q}$ is momentum, μ is mass, κ is surface tension and \mathcal{H} measures energy. The units (denoted by τ) are

$$\begin{aligned}\tau(q) &= m & \tau(p) &= \frac{kg \cdot m}{s} \\ \tau(\kappa) &= \frac{kg}{s^2} & \tau(\mu) &= kg\end{aligned}$$

where m is meters, kg is kilograms and s is seconds. Energy is measured in joules, and indeed it is easy to check that $\tau(\mathcal{H}) = \frac{kg \cdot m^2}{s^2}$.

Note that the units for differentiation by x are $\tau(\frac{\partial}{\partial x}) = \frac{1}{\tau(x)}$. For example, differentiating by time has units $\frac{1}{s}$. Hamilton's equations state that $\dot{q} = \frac{\partial \mathcal{H}}{\partial p} = \frac{1}{\mu} \cdot p$ and $\dot{p} = -\frac{\partial \mathcal{H}}{\partial q} = -\kappa \cdot q$ where

$$\begin{aligned}\tau(\dot{q}) &= \frac{m}{s} & \tau(\dot{p}) &= \frac{kg \cdot m}{s^2} \\ \tau\left(\frac{\partial}{\partial p}\right) &= \frac{1}{m} & \tau\left(\frac{\partial}{\partial q}\right) &= \frac{s}{kg \cdot m}\end{aligned}$$

The resulting flow describing the dynamics of the system is

$$\xi = \dot{q} \cdot \frac{\partial}{\partial q} + \dot{p} \cdot \frac{\partial}{\partial p} = \frac{1}{\mu} p \cdot \frac{\partial}{\partial q} - \kappa q \cdot \frac{\partial}{\partial p}$$

with units $\tau(\xi) = \frac{1}{s}$. Hamilton's equations can be reformulated more abstractly via symplectic geometry. Introduce the symplectic form

$$\omega = dq \wedge dp \quad \text{with units} \quad \tau(\omega) = \frac{kg \cdot m^2}{s}.$$

Observe that contracting the flow with the Hamiltonian obtains

$$\iota_{\xi} \omega = \omega(\xi, \bullet) = d\mathcal{H} = \frac{\partial \mathcal{H}}{\partial q} \cdot dq + \frac{\partial \mathcal{H}}{\partial p} \cdot dp$$

with units $\tau(d\mathcal{H}) = \tau(\mathcal{H}) = \frac{kg \cdot m^2}{s^2}$.

Losses in classical mechanics. Although there is no notion of “loss” in classical mechanics, it is useful (for the next section) to keep pushing the formal analogy. Define the “losses”

$$\ell_1(q, p) = \frac{1}{\mu} \cdot qp \quad \text{and} \quad \ell_2(q, p) = -\kappa \cdot qp \quad (4)$$

with units $\tau(\ell_1) = \frac{m^2}{s}$ and $\tau(\ell_2) = \frac{kg^2 \cdot m^2}{s^3}$. The Hamiltonian dynamics can then be recovered game-theoretically by differentiating ℓ_1 and ℓ_2 with respect to q and p respectively. It is easy to check that

$$\xi = \frac{\partial \mathcal{H}}{\partial p} \frac{\partial}{\partial q} - \frac{\partial \mathcal{H}}{\partial q} \frac{\partial}{\partial p} = \frac{\partial \ell_1}{\partial q} \frac{\partial}{\partial q} + \frac{\partial \ell_2}{\partial p} \frac{\partial}{\partial p}.$$

The duality between vector fields and differential forms.

Finally recall that the symplectic form in games was not “pulled out of thin air” as $\omega = dq \wedge dp$, but rather derived as $\omega = d\xi^b$, where ξ^b is the differential form corresponding to the vector field ξ under the musical isomorphism $\flat : TM \rightarrow T^*M$.

It is instructive to compute ξ^b in the case of a classical mechanical system and see what happens. Naively, we would guess that the musical isomorphism is $\left(\frac{\partial}{\partial q}\right)^b = dq$ and $\left(\frac{\partial}{\partial p}\right)^b = dp$. However, applying the naive musical isomorphism to ξ to get

$$\xi^b = \frac{\partial \ell_1}{\partial q} \cdot dq + \frac{\partial \ell_2}{\partial p} \cdot dp$$

results in a *type violation* because

$$\tau\left(\frac{\partial \ell_1}{\partial q} \cdot dq\right) = \tau(\ell_1) = \frac{m^2}{s}$$

whereas

$$\tau\left(\frac{\partial \ell_2}{\partial p} \cdot dp\right) = \tau(\ell_2) = \frac{kg^2 \cdot m^2}{s^3}$$

and we cannot add objects with different types.

To correct the type inconsistency, define the musical isomorphism as

$$\left(\frac{\partial}{\partial q}\right)^b = \frac{\mu}{2} \cdot dq \quad \text{and} \quad \left(\frac{\partial}{\partial p}\right)^b = \frac{1}{2\kappa} \cdot dp$$

with inverse

$$(dq)^\sharp = \frac{2}{\mu} \cdot \frac{\partial}{\partial q} \quad \text{and} \quad (dp)^\sharp = 2\kappa \cdot \frac{\partial}{\partial p}.$$

The correction terms in the direction $\flat : TM \rightarrow T^*M$ invert the coupling terms κ and $\frac{1}{\mu}$ that were originally introduced into the Hamiltonian for physical reasons. Applying the corrected musical isomorphism to ξ yields

$$\xi^b = \frac{\mu}{2} \cdot \frac{\partial f}{\partial q} \cdot dq + \frac{1}{2\kappa} \cdot \frac{\partial g}{\partial p} \cdot dp = \frac{1}{2} (p \cdot dq - q \cdot dp).$$

The two terms of ξ^b then have coherent types

$$\begin{aligned}\tau\left(\frac{\partial \ell_1}{\partial q} \cdot \mu \cdot dq\right) &= \frac{m}{s} \cdot kg \cdot m = \frac{kg \cdot m^2}{s} \\ \tau\left(\frac{\partial \ell_2}{\partial p} \cdot \frac{1}{\kappa} \cdot dp\right) &= \frac{kg \cdot m}{s^2} \cdot \frac{s^2}{kg} \cdot \frac{kg \cdot m}{s} = \frac{kg \cdot m^2}{s}\end{aligned}$$

as required. The associated two form is

$$\omega := d\xi^b = -\left(\mu \cdot \frac{\partial^2 f}{\partial q \partial p} - \frac{1}{\kappa} \cdot \frac{\partial^2 g}{\partial q \partial p}\right) dq \wedge dp = -dq \wedge dp$$

which recovers the symplectic form (up to sign) with units $\tau(\omega) = \frac{kg \cdot m^2}{s}$ as required. Finally, observe that

$$\begin{aligned} \langle \xi, \xi^b \rangle &= \frac{1}{2} \left\langle \frac{p}{\mu} \cdot \frac{\partial}{\partial q} - \kappa q \cdot \frac{\partial}{\partial p}, p \cdot dq - q \cdot dp \right\rangle \\ &= \frac{1}{2} \left(\kappa \cdot q^2 + \frac{1}{\mu} \cdot p^2 \right) = \mathcal{H}(p, q) \end{aligned}$$

recovering the Hamiltonian.

E.2. Units in two-player games

Without loss of generality let $\mathbf{w} = (\mathbf{x}; \mathbf{y})$ where we refer to \mathbf{x} as position and \mathbf{y} as momentum so that $\tau(\mathbf{x}) = m$ and $\tau(\mathbf{y}) = \frac{kg \cdot m}{s}$. The aim of this section is to check type-consistency under these, rather arbitrarily assigned, units. Since we are considering a game, we do not require that \mathbf{x} and \mathbf{y} have the same dimension – even though this would necessarily be the case for a physical system. The goal is to verify that units can be consistently assigned to games.

Consider a quadratic two player game of the form

$$\ell_1(\mathbf{w}) = \frac{1}{2} (\mathbf{x}^\top \ \mathbf{y}^\top) \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} + (\mathbf{x}^\top \ \mathbf{y}^\top) \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix}$$

and

$$\ell_2(\mathbf{w}) = \frac{1}{2} (\mathbf{x}^\top \ \mathbf{y}^\top) \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} + (\mathbf{x}^\top \ \mathbf{y}^\top) \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{pmatrix}$$

We restrict to quadratic games since our methods only involve first and second derivatives. We assume the matrices \mathbf{A} and \mathbf{C} are symmetric without loss of generality so that, for example, $\mathbf{A}_{12} = \mathbf{A}_{21}^\top$. Adding constant terms to ℓ_1 and ℓ_2 makes no difference to the analysis so they are omitted.

By (4), the units for ℓ_1 and ℓ_2 should be $\frac{m^2}{s}$ and $\frac{kg \cdot m^2}{s^3}$ respectively. We can therefore derive the correct units for each of the components of the quadratic losses as

$$\underbrace{\left(m \frac{kg \cdot m}{s} \right) \begin{pmatrix} \frac{1}{s} & \frac{1}{kg} \\ \frac{1}{kg} & \frac{s}{kg^2} \end{pmatrix} \begin{pmatrix} m \\ \frac{kg \cdot m}{s} \end{pmatrix}}_{\mathbf{w}^\top \mathbf{A} \mathbf{w}} + \underbrace{\left(m \frac{kg \cdot m}{s} \right) \begin{pmatrix} \frac{m}{s} \\ \frac{m}{kg} \end{pmatrix}}_{\mathbf{w}^\top \mathbf{b}}$$

for ℓ_1 and

$$\underbrace{\left(m \frac{kg \cdot m}{s} \right) \begin{pmatrix} \frac{kg^2}{s^3} & \frac{kg}{s^2} \\ \frac{kg}{s^2} & \frac{1}{s} \end{pmatrix} \begin{pmatrix} m \\ \frac{kg \cdot m}{s} \end{pmatrix}}_{\mathbf{w}^\top \mathbf{C} \mathbf{w}} + \underbrace{\left(m \frac{kg \cdot m}{s} \right) \begin{pmatrix} \frac{kg^2 \cdot m}{s^3} \\ \frac{kg \cdot m}{s^2} \end{pmatrix}}_{\mathbf{w}^\top \mathbf{d}}$$

for ℓ_2 . It follows from a straightforward computation that the vector field $\xi = \frac{\partial \ell_1}{\partial \mathbf{x}} \frac{\partial}{\partial \mathbf{x}} + \frac{\partial \ell_2}{\partial \mathbf{y}} \frac{\partial}{\partial \mathbf{y}}$ has type $\tau(\xi) = \frac{1}{s}$ as required.

The presymplectic form $\omega = d\xi^b$ makes use of the musical isomorphism $\flat : T^M \rightarrow T^*M$. As in section E.1, if we

naively define $(\frac{\partial}{\partial \mathbf{x}})^\flat = d\mathbf{x}$ and $(\frac{\partial}{\partial \mathbf{y}})^\flat = d\mathbf{y}$ then

$$\xi^b = \frac{\partial \ell_1}{\partial \mathbf{x}} \cdot d\mathbf{x} + \frac{\partial \ell_2}{\partial \mathbf{y}} \cdot d\mathbf{y}$$

which is type inconsistent because $\tau(\frac{\partial \ell_1}{\partial \mathbf{x}} \cdot d\mathbf{x}) = \frac{m^2}{s}$ and $\tau(\frac{\partial \ell_2}{\partial \mathbf{y}} \cdot d\mathbf{y}) = \frac{kg^2 \cdot m^2}{s^3}$.

Type-consistency via SVD. It is necessary, as in section E.1, to correct the naive musical isomorphism by taking into account the coupling constants for the mixed position-momentum terms. In the classical setup the coupling constants were the scalars $\frac{1}{\mu}$ and κ , whereas in a game they are the off-diagonal blocks \mathbf{A}_{12} and \mathbf{C}_{12} .

Apply singular value decomposition to factorize

$$\mathbf{A}_{12} = \mathbf{U}_A^\top \mathbf{D}_A \mathbf{V}_A \quad \text{and} \quad \mathbf{C}_{12} = \mathbf{U}_C^\top \mathbf{D}_C \mathbf{V}_C$$

where the entries of the diagonal matrices have types $\tau(\mathbf{D}_A) = \frac{1}{kg}$ and $\tau(\mathbf{D}_C) = \frac{kg}{s^2}$, and the types of the orthogonal matrices \mathbf{U} and \mathbf{V} are pure scalars. The diagonal matrices \mathbf{D}_A and \mathbf{D}_C have the same types as $\frac{1}{\mu}$ and κ in the classical system since they play the same coupling role.

Extending the procedure adopted in the section E.1, fix the type-inconsistency by defining the musical isomorphisms as

$$\left(\frac{\partial}{\partial \mathbf{x}} \right)^\flat = \mathbf{U}_A^\top \mathbf{D}_A^{-1} \mathbf{U}_A \cdot d\mathbf{x}$$

and

$$\left(\frac{\partial}{\partial \mathbf{y}} \right)^\flat = \mathbf{V}_C^\top \mathbf{D}_C^{-1} \mathbf{V}_C \cdot d\mathbf{y}.$$

Alternatively, the isomorphisms can be computed by noting that $\mathbf{U}_A^\top \mathbf{D}_A^{-1} \mathbf{U}_A = (\sqrt{\mathbf{A}_{12} \mathbf{A}_{21}})^{-1}$ and $\mathbf{V}_C^\top \mathbf{D}_C^{-1} \mathbf{V}_C = (\sqrt{\mathbf{C}_{21} \mathbf{C}_{12}})^{-1}$.

The dual isomorphism $\sharp : T^*M \rightarrow TM$ is then

$$(d\mathbf{x})^\sharp = \mathbf{U}_A^\top \mathbf{D}_A \mathbf{U}_A \cdot \frac{\partial}{\partial \mathbf{x}} \quad \text{and} \quad (d\mathbf{y})^\sharp = \mathbf{V}_C^\top \mathbf{D}_C \mathbf{V}_C \cdot \frac{\partial}{\partial \mathbf{y}}$$

If

$$\xi = \begin{pmatrix} \mathbf{A}_{12} \mathbf{y} + \mathbf{b}_1 \\ \mathbf{C}_{21} \mathbf{x} + \mathbf{d}_2 \end{pmatrix}$$

then it follows that

$$\xi^b = \begin{pmatrix} \mathbf{U}_A^\top \mathbf{D}_A^{-1} \mathbf{U}_A \mathbf{b}_1 + \mathbf{U}_A^\top \mathbf{U}_A \mathbf{y} \\ \mathbf{V}_C^\top \mathbf{D}_C^{-1} \mathbf{V}_C \mathbf{d}_2 + \mathbf{V}_C^\top \mathbf{V}_C \mathbf{x} \end{pmatrix}$$

with associated closed two form

$$\omega_\tau = d\xi^b = -(\mathbf{U}_A^\top \mathbf{V}_A - \mathbf{U}_C^\top \mathbf{V}_C) d\mathbf{x} \wedge d\mathbf{y}.$$

where the notation ω_τ emphasizes that the two-form is type-consistent.

E.3. What does type-consistency buy?

Example 7. Consider the loss functions

$$f(x, y) = xy \quad \text{and} \quad g(x, y) = 2xy,$$

with $\xi = (y, 2x)$. There is no function $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ such that $\nabla \phi = \xi$. However, there is a family of functions $\phi_\alpha(x, y) = \alpha \cdot xy$ which satisfies

$$\langle \xi, \nabla \phi_\alpha \rangle = \alpha \cdot (x^2 + 2y^2) \geq 0 \quad \text{for all } \alpha > 0.$$

Although ξ is not a potential field, there is a family of functions on which ξ performs gradient descent – albeit with coordinate-wise learning rates that may not be optimal. The vector field ξ arguably does not require adjustment. This kind of situation often arises when the learning rates of different parameters are set adaptively during training of neural nets, by rescaling them by positive numbers.

The vanilla and type-consistent 1-forms corresponding to ξ are, respectively,

$$\xi^b = y \cdot dx + 2x \cdot dy \quad \text{and} \quad \xi_\tau^b = y \cdot dx + x \cdot dy$$

with

$$\omega = d\xi_{\text{non}}^b = dx \wedge dy \quad \text{and} \quad \omega_\tau = d\xi_\tau^b = 0.$$

It follows that the *type-consistent* symplectic gradient adjustment is zero. Type-consistency ‘detects’ that no gradient adjustment is needed in example 7.

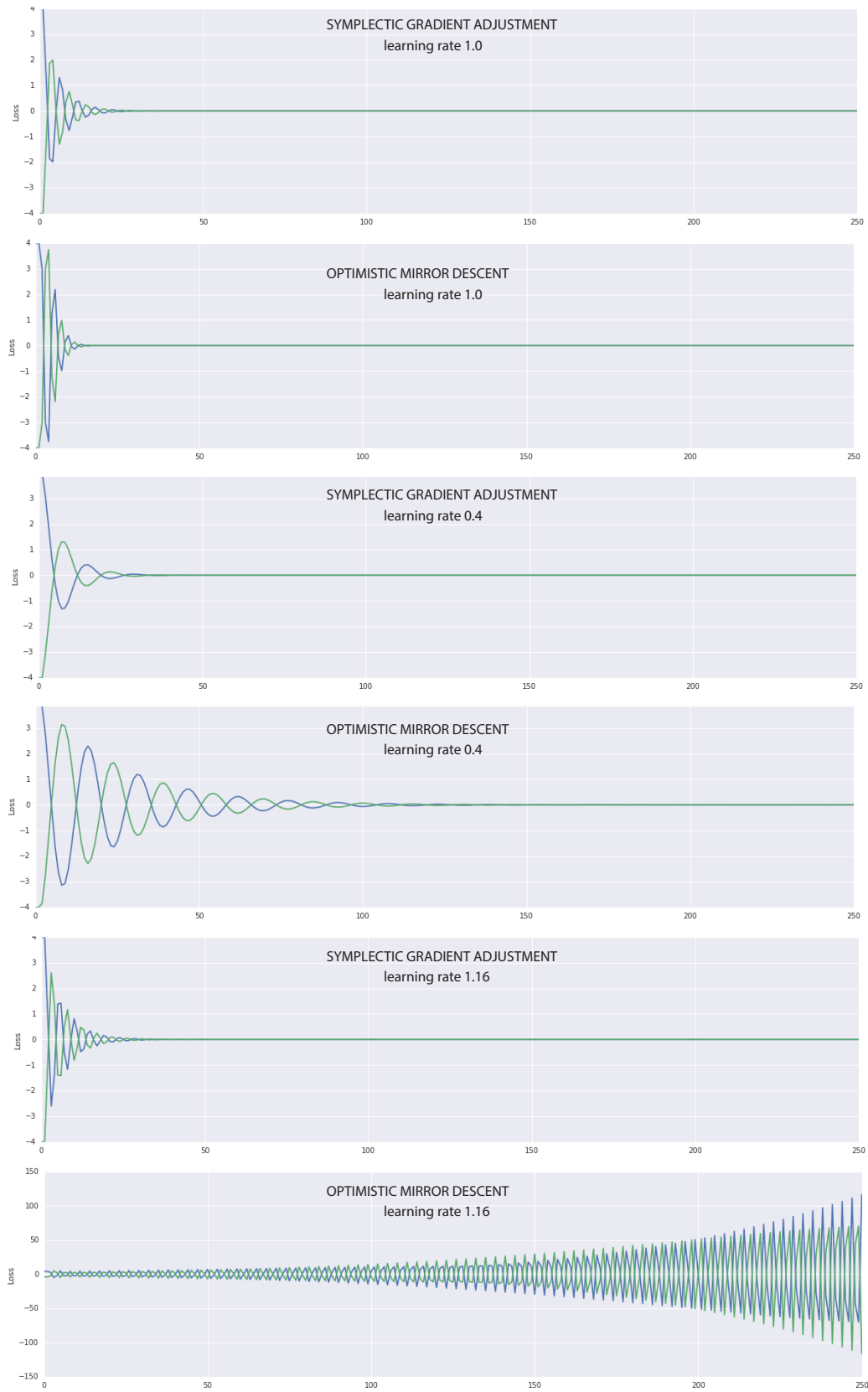


Figure 10. Individual runs on zero-sum bimatrix game in section 4.2.