Nesterov's Accelerated Gradient Descent as a Nash Equilibrium Seeking Algorithm for Quadratic Games

Jay Paek

January 17, 2026

Abstract

We investigate the behavior of Nesterov's accelerated gradient descent (NAGD) when applied to Nash equilibrium seeking in N-player quadratic games. Through a projection-based analysis of the associated second-order ordinary differential equation—the Su–Boyd–Candès equation—we establish a complete spectral characterization of stability. Our main result demonstrates that stability is determined by the eigenvalue structure of the pseudo-gradient matrix: trajectories converge to the null space when all eigenvalues are non-negative real numbers, but exhibit exponential divergence when complex or negative eigenvalues are present. These findings delineate the precise conditions under which NAGD succeeds or fails as a Nash equilibrium seeking algorithm.

1 Introduction

Nesterov's accelerated gradient descent (NAGD) [2] is a foundational algorithm in convex optimization, celebrated for its optimal convergence rate among first-order methods. The continuous-time limit of NAGD, introduced by Su, Boyd, and Candès [4], provides a dynamical systems perspective that has deepened our understanding of momentum-based optimization.

This paper examines NAGD in a fundamentally different context: as a *Nash equilibrium seeking algorithm* (NESA) for multi-player games. In game-theoretic settings, each player minimizes their own cost function with respect to their own action, leading to the notion of a *pseudo-gradient* rather than a standard gradient. This distinction has profound implications for the dynamics, as the pseudo-gradient need not derive from a potential function.

Our investigation reveals a dichotomy in the behavior of NAGD for quadratic games. When the pseudo-gradient matrix has favorable spectral properties—specifically, when all eigenvalues are non-negative real numbers—the dynamics exhibit stable behavior, with trajectories converging to the null space of the pseudo-gradient. Conversely, when the spectrum contains complex or negative eigenvalues, the dynamics become unstable, with certain projections of the trajectory growing exponentially. This characterization provides practitioners with a clear criterion for determining when NAGD-based equilibrium seeking is viable.

The paper is organized as follows. Section 2 establishes notation and reviews the NAGD dynamics and quadratic game framework. Section 3 develops the projection-based stability analysis. Section 4 presents our main spectral characterization theorem. Section 5 provides numerical illustrations, and Section 6 concludes with a discussion of implications and future directions.

2 Preliminaries

2.1 Notation

Let $\mathbb{R}_{++}$ denote the positive real numbers. For a symmetric matrix A, the notation $A \succ 0$ indicates positive definiteness. The time variable $t \in \mathbb{R}_{++}$ parameterizes continuous trajectories, with $\dot{X}(t)$ and $\ddot{X}(t)$ denoting first and second time derivatives, respectively.

For a matrix $A \in \mathbb{R}^{m \times n}$, the range space and null space are defined as

$$\mathcal{R}(A) \triangleq \{y \in \mathbb{R}^m : y = Ax \text{ for some } x \in \mathbb{R}^n\}, \qquad \mathbb{N}(A) \triangleq \{x \in \mathbb{R}^n : Ax = 0\}.$$

The orthogonal complement of a set $S \subseteq \mathbb{R}^n$ is $S^\perp \triangleq \{x \in \mathbb{R}^n : x^\top y = 0 \text{ for all } y \in S\}$. We denote the Euclidean norm by $\|\cdot\|_2$ and the imaginary part of a complex number by $\Im(\cdot)$.

2.2 Nesterov Accelerated Gradient Dynamics

The continuous-time model of NAGD, due to Su, Boyd, and Candès [4], is the second-order differential equation

$$\ddot{X}(t) + \frac{3}{t}\dot{X}(t) + \nabla f\big(X(t)\big) = 0, \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is the objective function. The coefficient $3/t$ provides time-varying damping that diminishes as $t \to \infty$, enabling accelerated convergence while maintaining stability.

Introducing position and velocity states $x_1(t) \triangleq X(t)$ and $x_2(t) \triangleq \dot{X}(t)$, we obtain the equivalent first-order system

$$\begin{cases} \dot{x}_1(t) = x_2(t), \\ \dot{x}_2(t) = -\dfrac{3}{t}x_2(t) - \nabla f\big(x_1(t)\big). \end{cases} \tag{2}$$

2.3 Quadratic Games and the Pseudo-Gradient

Consider a game with N players, where player $i \in \{1, \dots, N\}$ controls action $x_i \in \mathbb{R}$ and seeks to minimize a quadratic cost

$$J_i(x) = x^\top Q_i x,$$

where $x = (x_1, \dots, x_N)^\top \in \mathbb{R}^N$ is the joint action profile and $Q_i \in \mathbb{R}^{N \times N}$ is a given matrix.

Definition 1 (Pseudo-gradient)**.** *The* pseudo-gradient *(or game gradient) is the vector field $\nabla f : \mathbb{R}^N \to \mathbb{R}^N$ whose i-th component is the partial derivative of player i's cost with respect to player i's own action:*

$$\nabla f(x) \triangleq \begin{bmatrix} \dfrac{\partial J_1}{\partial x_1}(x) & \cdots & \dfrac{\partial J_N}{\partial x_N}(x) \end{bmatrix}^\top. \tag{3}$$

For quadratic costs, the pseudo-gradient is linear:

$$\nabla f(x) = \mathcal{G}x, \tag{4}$$

where $\mathcal{G} \in \mathbb{R}^{N \times N}$ is the *pseudo-gradient matrix*. Unlike standard gradients, $\mathcal{G}$ need not be symmetric, as it encodes the strategic interactions among players rather than derivatives of a single potential function.

A *Nash equilibrium* is a point $x^\star$ satisfying $\mathcal{G}x^\star = 0$, i.e., $x^\star \in \mathbb{N}(\mathcal{G})$. When $\mathcal{G}$ is full rank, the unique equilibrium is $x^\star = 0$.

2.4 The NAGD Game Dynamics

Combining (1) with (4) yields the NAGD game dynamics:

$$\ddot{x}(t) + \frac{3}{t}\dot{x}(t) + \mathcal{G}x(t) = 0, \qquad t \in [t_0, \infty),\ t_0 > 0. \tag{5}$$

We work on $[t_0, \infty)$ with $t_0 > 0$ to avoid the singularity at $t = 0$. In first-order form:

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A(t)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad A(t) \triangleq \begin{bmatrix} \mathbf{0} & I \\ -\mathcal{G} & -\frac{3}{t}I \end{bmatrix}. \tag{6}$$

3 Projection-Based Stability Analysis

Our approach exploits the fact that projecting the dynamics onto eigenvectors of $\mathcal{G}$ yields decoupled scalar equations amenable to explicit solution.

3.1 Stability via Orthogonal Projections

The following proposition establishes that, for linear systems, Lyapunov stability can be analyzed component-wise in any orthonormal basis that diagonalizes the dynamics.

Proposition 1 (Stability via orthogonal projections)**.** *Consider the autonomous linear system*

$$\dot{x}(t) = Ax(t), \qquad x(t) \in \mathbb{R}^n, \tag{7}$$

where $A \in \mathbb{R}^{n\times n}$. Let $\{w_i\}_{i=1}^n \subset \mathbb{R}^n$ be an orthonormal basis, and define scalar coordinates $y_i(t) \triangleq w_i^\top x(t)$. Suppose that for each i, there exists $g_i : \mathbb{R} \to \mathbb{R}$ such that

$$w_i^\top Ax(t) = g_i\big(y_i(t)\big) \qquad \text{for all } t \geq 0. \tag{8}$$

Then the equilibrium $x = 0$ of (7) is Lyapunov stable if and only if $y_i = 0$ is Lyapunov stable for each scalar system $\dot{y}_i = g_i(y_i)$.

Proof. The assumption (8) ensures that $\dot{y}_i = w_i^\top \dot{x} = w_i^\top Ax = g_i(y_i)$, so each scalar system is well-defined.

Necessity. Suppose $x = 0$ is Lyapunov stable. For any $\varepsilon > 0$, there exists $\delta > 0$ such that $\|x(0)\|_2 < \delta$ implies $\|x(t)\|_2 < \varepsilon$ for all $t \geq 0$. Since the basis is orthonormal, $|y_i(t)| = |w_i^\top x(t)| \leq \|x(t)\|_2$, so $|y_i(t)| < \varepsilon$ whenever $|y_i(0)| \leq \|x(0)\|_2 < \delta$. Thus each $y_i = 0$ is stable.

Sufficiency. Suppose each $y_i = 0$ is Lyapunov stable. Given $\varepsilon > 0$, set $\varepsilon_i = \varepsilon/\sqrt{n}$ and let $\delta_i > 0$ be the corresponding stability bound for y_i. Define $\delta = \min_i \delta_i$. If $\|x(0)\|_2 < \delta$, then $|y_i(0)| \leq \|x(0)\|_2 < \delta_i$ for all i, hence $|y_i(t)| < \varepsilon/\sqrt{n}$. By Parseval's identity,

$$\|x(t)\|_2^2 = \sum_{i=1}^n y_i(t)^2 < n \cdot \frac{\varepsilon^2}{n} = \varepsilon^2,$$

establishing Lyapunov stability of $x = 0$. □

Remark 1. *The assumption (8) holds when the basis $\{w_i\}$ consists of eigenvectors of A. For normal matrices (including symmetric matrices), such an orthonormal eigenbasis always exists.*

3.2 Low-Rank Pseudo-Gradient: Null Space Convergence

When $\mathcal{G}$ is rank-deficient, projections onto the left null space decouple from the remaining dynamics.

Lemma 1 (Left null space projection)**.** *Suppose* $\mathrm{rank}(\mathcal{G}) < n$*. For any* $w \in \mathbb{N}(\mathcal{G}^\top) \setminus \{0\}$*, define* $y_1(t) \triangleq w^\top x_1(t)$ *and* $y_2(t) \triangleq w^\top x_2(t)$*. Then* (y_1, y_2) *satisfies the decoupled system*

$$\dot{y}_1 = y_2, \qquad \dot{y}_2 = -\frac{3}{t} y_2. \tag{9}$$

Proof. Since $w^\top \mathcal{G} = 0$, left-multiplying the second equation of (6) by $w^\top$ yields $\dot{y}_2 = -\frac{3}{t} y_2 - w^\top \mathcal{G} x_1 = -\frac{3}{t} y_2$, as claimed. $\square$

The system (9) admits the explicit solution

$$y_2(t) = \left(\frac{t_0}{t}\right)^3 y_2(t_0), \tag{10}$$

$$y_1(t) = y_1(t_0) + \frac{t_0}{2}\left(1 - \frac{t_0^2}{t^2}\right) y_2(t_0). \tag{11}$$

Consequently, $y_2(t) \to 0$ as $t \to \infty$, and $y_1(t)$ converges to the finite limit

$$\lim_{t \to \infty} y_1(t) = y_1(t_0) + \frac{t_0}{2} y_2(t_0). \tag{12}$$

This establishes that projections onto the left null space are uniformly bounded and convergent, consistent with trajectories drifting along (rather than toward) the null space.

3.3 Full-Rank Pseudo-Gradient: Modal Analysis

When $\mathcal{G}$ is full rank, the unique equilibrium is $x = 0$. Stability depends on the eigenvalue structure of $\mathcal{G}$.

Lemma 2 (Eigenvector projection)**.** *Let* $w \neq 0$ *satisfy* $w^\top \mathcal{G} = \lambda w^\top$ *for some* $\lambda \in \mathbb{C}$*. Define* $y(t) \triangleq w^\top x_1(t)$*. Then* y *satisfies*

$$\ddot{y}(t) + \frac{3}{t}\dot{y}(t) + \lambda y(t) = 0. \tag{13}$$

Proof. Projecting (5) onto $w^\top$ and using $w^\top \mathcal{G} = \lambda w^\top$ yields

$$w^\top \ddot{x} + \frac{3}{t} w^\top \dot{x} + w^\top \mathcal{G} x = \ddot{y} + \frac{3}{t}\dot{y} + \lambda y = 0. \qquad \square$$

Equation (13) is a Bessel-type equation. The substitution $y(t) = t^{-1} u(t)$ transforms it into

$$t^2 \ddot{u} + t\dot{u} + (\lambda t^2 - 1) u = 0,$$

which is Bessel's equation of order 1. The general solution is

$$y(t) = \frac{c_1}{t} J_1(\sqrt{\lambda}\, t) + \frac{c_2}{t} Y_1(\sqrt{\lambda}\, t), \tag{14}$$

where J_1 and Y_1 are Bessel functions of the first and second kind, and c_1, c_2 are determined by initial conditions.

The asymptotic behavior for large $|\sqrt{\lambda}\,t|$ is governed by [3]:

$$J_1(z) \sim \sqrt{\frac{2}{\pi z}} \cos\left(z - \frac{3\pi}{4}\right) \cdot \left(1 + \mathcal{O}(|z|^{-1})\right), \tag{15}$$

$$Y_1(z) \sim \sqrt{\frac{2}{\pi z}} \sin\left(z - \frac{3\pi}{4}\right) \cdot \left(1 + \mathcal{O}(|z|^{-1})\right), \tag{16}$$

valid for $|\arg z| < \pi$. When $\sqrt{\lambda}$ has nonzero imaginary part, the trigonometric functions acquire exponentially growing factors of order $e^{|\Im(\sqrt{\lambda})|t}$.

Lemma 3 (Modal asymptotics)**.** *Let $y(t)$ solve (13) with $\lambda \in \mathbb{C}$.*

(i) If $\lambda > 0$ is real, then $y(t) = \mathcal{O}(t^{-3/2})$ as $t \to \infty$.

(ii) If $\Im(\sqrt{\lambda}) \neq 0$ (equivalently, $\lambda \notin \mathbb{R}_{\geq 0}$), then for generic initial conditions, $|y(t)| \to \infty$ exponentially.

Proof. (i) When $\lambda > 0$, we have $\sqrt{\lambda} \in \mathbb{R}$, so $J_1(\sqrt{\lambda}\,t)$ and $Y_1(\sqrt{\lambda}\,t)$ are bounded oscillatory functions. Combined with the $1/t$ prefactor in (14) and the $t^{-1/2}$ decay from (15)–(16), we obtain $y(t) = \mathcal{O}(t^{-3/2})$.

(ii) When $\sqrt{\lambda}$ has nonzero imaginary part, the asymptotic expansions contain factors $e^{|\Im(\sqrt{\lambda})|t}$, which dominate the polynomial decay, yielding exponential growth for generic initial conditions (those not annihilating the growing component). □

4 Main Results

We now state our main theorem, which provides a complete spectral characterization of stability for normal pseudo-gradient matrices.

Theorem 1 (Spectral characterization of NAGD stability)**.** *Consider the NAGD game dynamics (5) on $[t_0, \infty)$ with $t_0 > 0$. Assume $\mathcal{G}$ is normal, so there exists an orthonormal eigenbasis $\{w_i\}_{i=1}^n$ with eigenvalues $\{\lambda_i\}_{i=1}^n$. Then:*

*(a) **Instability criterion.** If $\lambda_i \notin \mathbb{R}_{\geq 0}$ for some i, then the origin of (6) is unstable. There exist arbitrarily small initial conditions for which $\|x_1(t)\|_2$ grows exponentially.*

*(b) **Stability and convergence.** If $\lambda_i \in \mathbb{R}_{\geq 0}$ for all i, then the origin is Lyapunov stable. Moreover:*

(i) For each $\lambda_i > 0$, the modal projection $y_i(t) = w_i^\top x_1(t)$ satisfies $y_i(t) \to 0$ with rate $\mathcal{O}(t^{-3/2})$.

(ii) For each $\lambda_i = 0$, the modal projection converges to a finite limit:

$$\lim_{t\to\infty} y_i(t) = y_i(t_0) + \frac{t_0}{2}\dot{y}_i(t_0).$$

Consequently, letting $\Pi_{\mathbb{N}(\mathcal{G})}$ denote orthogonal projection onto $\mathbb{N}(\mathcal{G})$,

$$x_1(t) \to \Pi_{\mathbb{N}(\mathcal{G})}\left(x_1(t_0) + \frac{t_0}{2}x_2(t_0)\right), \qquad x_2(t) \to 0.$$

In particular, if $\mathcal{G} \succ 0$, then $(x_1(t), x_2(t)) \to (0, 0)$.

Proof. Since $\mathcal{G}$ is normal, it admits an orthonormal eigenbasis $\{w_i\}$ with $\mathcal{G}w_i = \lambda_i w_i$. Decompose $x_1(t) = \sum_i y_i(t) w_i$, where $y_i(t) = w_i^\top x_1(t)$. By Lemma 2, each y_i satisfies the scalar ODE (13) with parameter λ_i.

Part (a). If some $\lambda_i \notin \mathbb{R}_{\geq 0}$, then $\Im(\sqrt{\lambda_i}) \neq 0$, and Lemma 3(ii) implies exponential growth of $y_i(t)$ for generic initial conditions. Since $|y_i(t)| \leq \|x_1(t)\|_2$, instability of the full system follows.

Part (b). If all $\lambda_i \geq 0$, then by Lemmas 1 and 3(i), each scalar mode is either convergent to zero (when $\lambda_i > 0$) or convergent to a finite limit (when $\lambda_i = 0$). Proposition 1 then implies Lyapunov stability of the origin.

The limiting behavior follows by decomposing $x_1(t)$ into its range and null space components. The range space components decay to zero, while the null space components converge to the limits given by (12). Reconstructing $x_1(t)$ from these components yields the stated formula. □

Corollary 1.1. *If $\mathcal{G} = \mathcal{G}^\top$ (symmetric), then the NAGD dynamics are Lyapunov stable if and only if $\mathcal{G}$ is positive semidefinite. In this case, trajectories converge to the null space of $\mathcal{G}$.*

Proof. Symmetric matrices are normal with real eigenvalues. The result follows immediately from Theorem 1. □

4.1 Remarks on Frozen-Time Analysis

A natural approach to analyzing (6) is to study the "frozen" system $\dot{z} = A(\alpha)z$ with $\alpha = 3/t$ held constant, then appeal to results on slowly-varying systems. We explain why this approach fails.

The limiting matrix as $t \to \infty$ (equivalently, $\alpha \to 0$) is

$$A^* = \lim_{t\to\infty} A(t) = \begin{bmatrix} \mathbf{0} & I \\ -\mathcal{G} & \mathbf{0} \end{bmatrix}.$$

The eigenvalues of A^* satisfy $\det(-\mathcal{G} - \mu^2 I) = 0$, so $\mu = \pm\sqrt{-\lambda}$ for each eigenvalue λ of $\mathcal{G}$. Unless $\mathcal{G}$ has only negative real eigenvalues (which contradicts our setting), A^* is not Hurwitz.

Standard results for slowly-varying systems, such as Khalil's Lemma 5.15 [1], require the frozen system to be *uniformly* Hurwitz with a spectral gap bounded away from zero. In our case, even when $A(\alpha)$ is Hurwitz for each fixed $\alpha > 0$, the spectral margin vanishes as $\alpha \to 0$. This precludes exponential stability conclusions via frozen-time analysis and necessitates the direct modal approach developed above.

5 Numerical Simulations

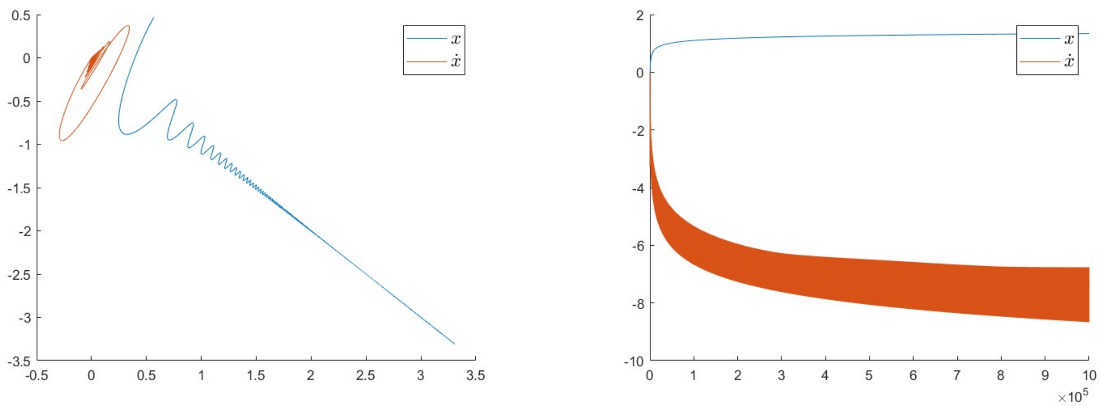
We illustrate Theorem 1 with numerical experiments. All simulations use fourth-order Runge–Kutta integration with step size $\Delta t = 0.01$ over $t \in [1, 1000]$, corresponding to 10^5 iterations.

5.1 Low-Rank Pseudo-Gradient Matrices

Figures 1 and 2 show trajectories for rank-one pseudo-gradient matrices. In both cases, the position $x_1(t)$ converges to the one-dimensional null space, while the velocity $x_2(t)$ decays to zero. The convergence is algebraic, consistent with the $\mathcal{O}(t^{-3})$ decay rate for null space components.

5.2 Full-Rank Positive Definite Matrices

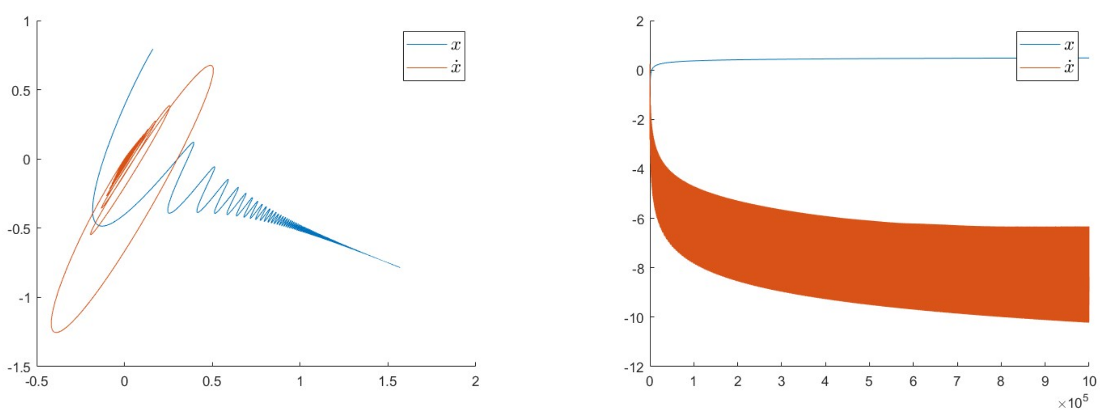
Figures 3 and 4 show dynamics for positive definite $\mathcal{G}$. The trajectories spiral inward toward the origin with the characteristic $\mathcal{O}(t^{-3/2})$ decay rate predicted by Theorem 1.



(a) Phase portrait

(b) $\log_{10}\|x(t)\|_2$ vs. time

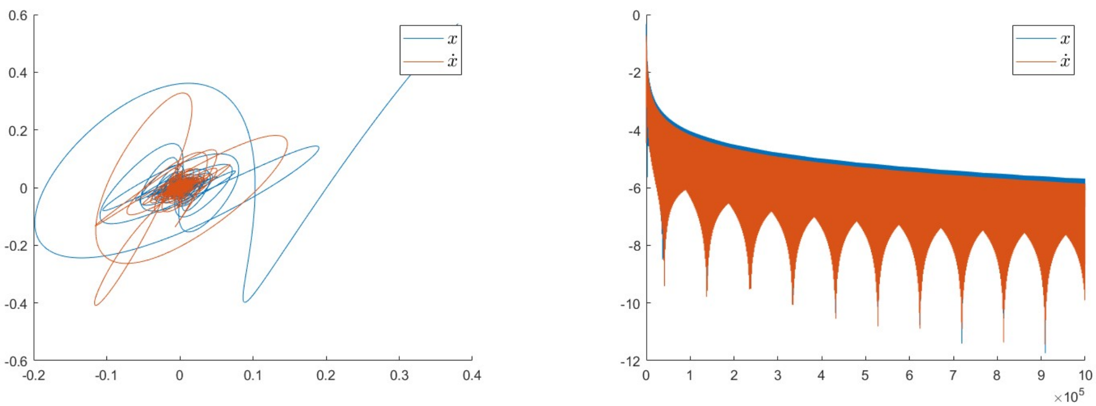
Figure 1: NAGD dynamics with $\mathcal{G} = \left[\begin{smallmatrix}1 & 1\\ 2 & 2\end{smallmatrix}\right]$. Eigenvalues: $\lambda_1 = 3$, $\lambda_2 = 0$. The trajectory converges to $\mathbb{N}(\mathcal{G}) = \mathrm{span}\{(1, -1)^\top\}$.



(a) Phase portrait

(b) $\log_{10}\|x(t)\|_2$ vs. time

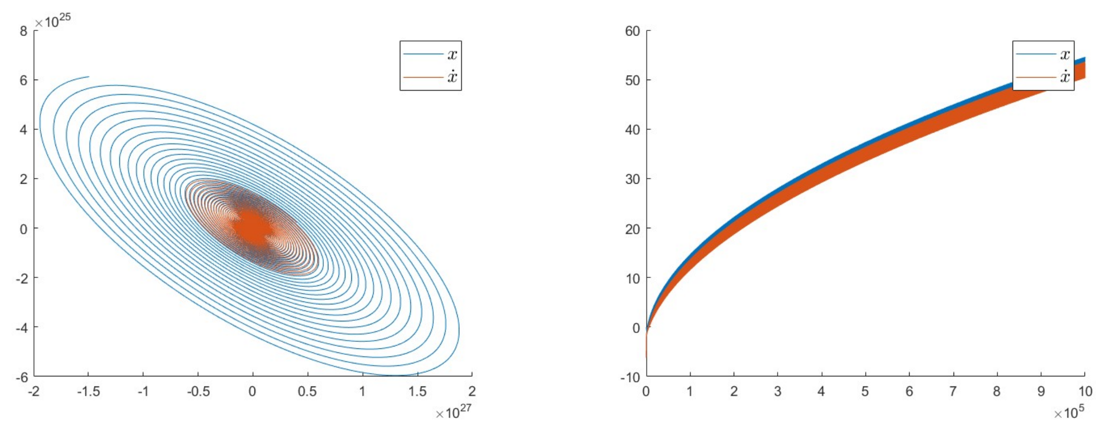
Figure 2: NAGD dynamics with $\mathcal{G} = \left[\begin{smallmatrix}1 & 2\\ 2 & 4\end{smallmatrix}\right]$. Eigenvalues: $\lambda_1 = 5$, $\lambda_2 = 0$. The trajectory converges to $\mathbb{N}(\mathcal{G}) = \mathrm{span}\{(2, -1)^\top\}$.



(a) Phase portrait

(b) $\log_{10} \|x(t)\|_2$ vs. time

Figure 3: NAGD dynamics with $\mathcal{G} = \left[\begin{smallmatrix} 0.4 & 0.2 \\ 0.2 & 0.8 \end{smallmatrix}\right]$ (symmetric positive definite). Eigenvalues: $\lambda \approx 0.27, 0.93$. Convergence to the origin.



(a) Phase portrait

(b) $\log_{10} \|x(t)\|_2$ vs. time

Figure 4: NAGD dynamics with $\mathcal{G} = \left[\begin{smallmatrix} 0.1 & -0.1 \\ 0 & 0.1 \end{smallmatrix}\right]$ (upper triangular, positive eigenvalues). Eigenvalue: $\lambda = 0.1$ (multiplicity 2). Convergence to the origin.

5.3 Unstable Cases: Complex and Negative Eigenvalues

Figures 5 through 7 illustrate instability arising from unfavorable spectra.

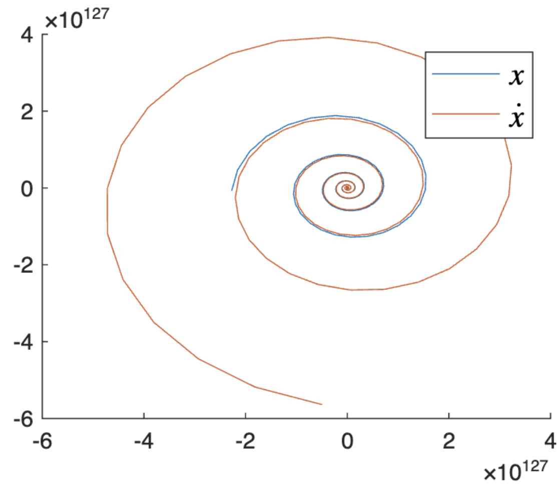


Figure 5: NAGD dynamics with $\mathcal{G} = \left[\begin{smallmatrix} 6 & 1.5 \\ -1.5 & 6 \end{smallmatrix}\right]$. Eigenvalues: $\lambda = 6 \pm 1.5i$ (complex). Despite positive real parts, the complex eigenvalues induce exponential growth through the Bessel function asymptotics.

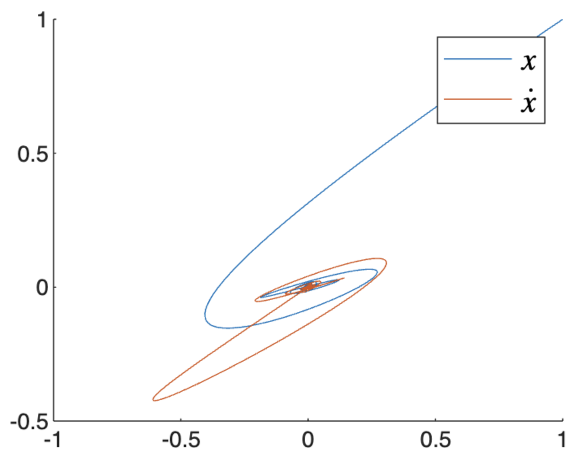


Figure 6: NAGD dynamics with $\mathcal{G} = \left[\begin{smallmatrix} 0.6557 & 0.8491 \\ 0.0357 & 0.9340 \end{smallmatrix}\right]$. Eigenvalues: $\lambda \approx 0.572, 1.018$ (both positive real). Stable convergence to the origin.

6 Discussion and Conclusion

We have established a complete spectral characterization of the stability of Nesterov accelerated gradient dynamics when applied to Nash equilibrium seeking in quadratic games. The key finding is that stability depends solely on whether the eigenvalues of the pseudo-gradient matrix lie in $\mathbb{R}_{\geq 0}$: nonnegative real eigenvalues yield stable dynamics converging to the null space, while complex or negative eigenvalues produce exponential instability.

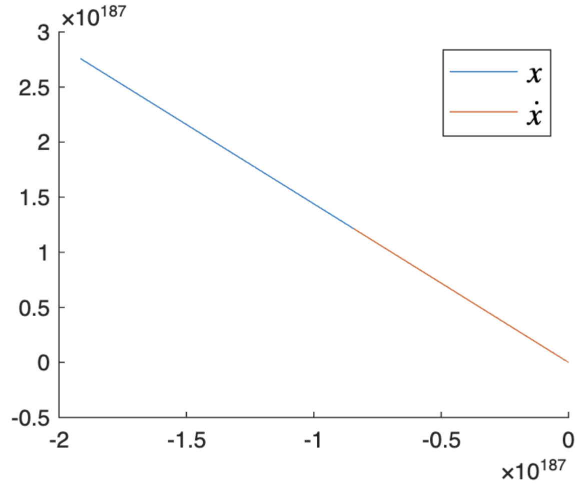


Figure 7: NAGD dynamics with $\mathcal{G} = \left[\begin{smallmatrix} 0.9572 & 0.8003 \\ 0.4854 & 0.1419 \end{smallmatrix}\right]$. Eigenvalues: $\lambda \approx 1.294, -0.195$ (one negative). The negative eigenvalue induces instability.

Several aspects of this characterization merit emphasis:

The role of the pseudo-gradient. Unlike standard optimization where the gradient derives from a single objective, the pseudo-gradient in game settings reflects the strategic interactions among players. This distinction manifests in two ways: (i) the pseudo-gradient matrix $\mathcal{G}$ need not be symmetric, and (ii) even when $\mathcal{G}$ has positive real eigenvalues, complex eigenvalues can arise in the non-symmetric case, leading to instability.

Comparison with standard NAGD. In convex optimization with a strongly convex objective, the Hessian is positive definite, guaranteeing convergence. Our results show that this guarantee does not extend to games: even with "well-behaved" cost functions, the interaction structure encoded in $\mathcal{G}$ can produce divergent dynamics.

Implications for algorithm design. Practitioners seeking to apply NAGD-type methods for equilibrium computation should verify that the pseudo-gradient matrix has non-negative real spectrum. For games where this condition fails, alternative algorithms—such as gradient descent-ascent with appropriate step sizes, or methods designed specifically for non-cooperative settings—may be more appropriate.

Extensions. Several directions for future work emerge naturally: (i) extending the analysis to non-quadratic games via linearization around equilibria, (ii) investigating discrete-time analogs of our stability characterization, and (iii) exploring modifications to the NAGD dynamics (e.g., different damping profiles) that might restore stability for broader classes of games.

Acknowledgments

[To be added.]

References

[1] Hassan K. Khalil. *Nonlinear Systems*. Prentice Hall, Upper Saddle River, NJ, 3rd edition, 2002.

[2] Yurii E. Nesterov. A method for solving the convex programming problem with convergence rate $O(1/k^2)$. *Doklady Akademii Nauk SSSR*, 269(3):543–547, 1983. In Russian. English translation in Soviet Mathematics Doklady.

[3] Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert, and Charles W. Clark, editors. *NIST Handbook of Mathematical Functions*. Cambridge University Press, New York, NY, 2010. Also available online at https://dlmf.nist.gov/.

[4] Weijie Su, Stephen Boyd, and Emmanuel J. Candès. A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights. *Journal of Machine Learning Research*, 17(153):1–43, 2016.