

Minimizing phase-space energies

Michael Updike,^{*} Nicholas Bohlsen[†], Hong Qin,[‡] and Nathaniel J. Fisch[§]

*Princeton Plasma Physics Laboratory, Princeton University, Princeton, New Jersey 08540, USA
and Department of Astrophysical Sciences, Princeton University, Princeton, New Jersey 08540, USA*



(Received 1 May 2025; accepted 30 July 2025; published 5 September 2025)

A primary technical challenge for harnessing fusion energy is to control and extract energy from a nonthermal distribution of charged particles. The fact that phase space evolves by symplectomorphisms fundamentally limits how a distribution may be manipulated. While the constraint of phase-space volume preservation is well understood, other constraints remain to be fully appreciated. To better understand these constraints, we study the problem of extracting energy from a distribution of particles using area-preserving and symplectic linear maps. When a quadratic potential is imposed, we find that the maximal extractable energy can be computed as trace minimization problems. We solve these problems and show that the extractable energy under linear symplectomorphisms may be much smaller than the extractable energy under special linear maps. The method introduced in the present study enables an energy-based proof of the linear Gromov nonsqueezing theorem.

DOI: [10.1103/4rbz-cgxf](https://doi.org/10.1103/4rbz-cgxf)

I. INTRODUCTION

As the deployment of commercial fusion energy accelerates, it becomes increasingly indispensable to develop phase space engineering techniques to energize fusing particles and extract energy from fusion product particles. For standard magnetic confinement D-T fusion, self-sustained burning requires that the energy of alpha particles be transferred to fusing ions, which remains a technical challenge [1]. Phase space engineering techniques have been designed to channel the alpha particle energy to fusing ions directly via electromagnetic waves [2–6]. For advanced fuel fusion using p-B11 or D-He3, the fusion energy released is carried by charged particles [7–16]. It is thus possible to convert the fusion energy directly into electricity by manipulating the charged particles of fusion products using electromagnetic fields. On the other hand, advanced aneutronic fusion operates in nonthermalized conditions, necessitating substantial power circulation to maintain fusing particles in nonequilibrium energy states [9,17–19]. For this purpose, energy needs to be extracted from the thermalized fusing particles and converted into kinetic energy.

Electromagnetic manipulation of charged particles encounters inherent physical constraints. Namely, the particle phase space must evolve by Hamiltonian symplectomorphisms. Liouville's theorem establishes that while phase space volume occupied by particles can be reconfigured, it cannot be reduced. This and other constraints led Qin *et al.* [18,19] to frame a pivotal inquiry: How much energy can we electromagnetically harvest from fusion byproducts (like alpha particles from p-B11, D-He3, or D-T reactions)? Conversely, what is the lowest attainable energy configuration through

electromagnetic processes? Phase-space conservation ensures this baseline state remains above zero, effectively capping the potential energy extraction achievable via radiofrequency waves in plasma systems.

To give rigorous estimates on these questions, we study how much energy can be extracted from a particle distribution undergoing Hamiltonian time evolution. To be general, let \mathcal{P} be a particle phase space, mathematically a symplectic manifold. A collection of many particles can be described by a distribution function $f_t : \mathcal{P} \rightarrow [0, \infty)$, which gives the number of particles in a region \mathcal{V} at time t as $\int_{\mathcal{V}} f_t$. Given a time-dependent Hamiltonian $\mathcal{H}(t) : \mathcal{P} \rightarrow \mathbb{R}$ generating a complete flow $\phi_t : \mathcal{P} \rightarrow \mathcal{P}$, the distribution function evolves in the absence of collisions as $f_t = f_0 \circ \phi_t^{-1}$. The Hamiltonian $\mathcal{H}(t)$ can be taken to include the self-generated fields of the particles, as well as any externally applied fields, in which case the evolution relation $f_t = f_0 \circ \phi_t^{-1}$ is equivalent to the Vlasov equation. Given a reference energy function $\mathcal{E} : \mathcal{P} \rightarrow \mathbb{R}$, representing the energy of particles in the absence of fluctuating fields, we define the energy of a particle distribution as $E[f_t] = \int_{\mathcal{P}} f_t \mathcal{E}$. Our question then translates to finding $\inf_{\phi \in \text{Ham}(\mathcal{P})} E[f_0 \circ \phi^{-1}]$, the minimal energy a particle distribution must maintain under Hamiltonian time evolution.

When the allowed transforms $\{\phi\}$ are relaxed to be merely invertible and area preserving, we recover the problem posed by Gardner in [20]. Under these relaxed assumptions, and with enough decay assumptions on f_0 , the Gardner energy $E_G := \inf_{\{\phi\}} E[f_0 \circ \phi^{-1}]$ can be computed by sequentially permuting equal-measure sets in phase space. This procedure has come to be known as Gardner's restacking algorithm [21–24]. While each permutation in Gardner restacking is noncontinuous, smooth approximations may be found using the theory of Dacorogna and Moser [25]. Thus, E_G defines the minimal energy even when $\{\phi\}$ are restricted to be area-preserving diffeomorphisms.

When $\{\phi\}$ are further restricted to be symplectomorphisms, one must ponder whether symplectic maps behave rigidly or

^{*}Contact author: michaelupdike@princeton.edu

[†]Contact author: nbohlsen@pppl.gov

[‡]Contact author: hongqin@princeton.edu

[§]Contact author: fisch@princeton.edu

flexibly with respect to the problem at hand. Unlike area-preserving maps, symplectic maps can be quite restrictive. Gromov's nonsqueezing theorem [26,27] states there is no symplectic embedding of the ball $B_{2n}(r) = \{\mathbf{z} \in \mathbb{R}^{2n} : |\mathbf{z}|^2 < r^2\}$ into the symplectic cylinder $\mathcal{Z}(R) = \{(\mathbf{x}, \mathbf{p}) \in \mathbb{R}^{2n} : x_1^2 + p_1^2 < R^2\}$ except when $r \leq R$. Yet, symplectic maps can also be quite flexible owing to Darboux's theorem, which implies there are no local invariants of symplectic manifolds [28,29]. This flexibility versus rigidity conundrum can be answered using a result of Katok [30, p. 545]. Namely, for any two equal-measure, compact subsets A, B of a symplectic manifold, and connected open set U , there is a Hamiltonian symplectomorphism ψ supported in U making the symmetric difference between $\psi^{-1}(B)$ and A arbitrarily small. This result allows one to approximate Gardner restacking with Hamiltonian symplectomorphisms, thereby showing $E_G = \inf_{\phi \in \text{Ham}(\mathcal{P})} E[\phi]$. The details of this are given in Appendix 2.

While we have answered our original question, we have done so unsatisfactorily. The symplectic transformations that take an initial distribution f_0 close to its minimal energy must generically include large gradients, and thus be infeasible to implement physically. Consider, for example, the problem of embedding all but an ϵ amount of $B_{2n}(r)$ into the cylinder $\mathcal{Z}_{2n}(R)$. It was shown by Sackel *et al.* [31, p.1116] that for any fixed $r > R$, there is a positive constant C such that the Lipschitz constant $L(\psi)$ of any symplectic embedding $\psi : B_{2n}(r) \rightarrow \mathbb{R}^4$ must satisfy $L(\psi)^2 \geq C\epsilon^{-1}$.

To remedy this large gradients problem, we must look for the infimum of $E[f_0 \circ \phi^{-1}]$ over a more suitable family of symplectomorphisms. The simplest case, and the one we study in this paper, is that f_0 is supported on the smallest scale we can manipulate. Such a situation arises in accelerator and plasma physics when one considers beams of particles [32–34]. In such a case, we may safely neglect the topology of \mathcal{P} and assume $\mathcal{P} = \mathbb{R}^{2n}$ with its standard symplectic structure. We may also approximate any allowed symplectomorphism by its linearization $\phi(\mathbf{z}) \approx A(\mathbf{z} - \mathbf{b})$ where $A \in Sp(2n)$ is a symplectic matrix and $\mathbf{b} \in \mathbb{R}^{2n}$. This amounts to ignoring cubic and higher-order terms in the Hamiltonian (see Appendix 4).

In this approximation, our refined question becomes to find

$$E_{Sp(2n)} := \inf_{\mathbf{b} \in \mathbb{R}^{2n}, A \in Sp(2n)} \int \mathcal{E}(\mathbf{z}) f_0(A^{-1}\mathbf{z} + \mathbf{b}) d^{2n}\mathbf{z}, \quad (1)$$

a quantity we call the linear Gromov energy. Since f_0 is assumed to be supported on the smallest scale we can manipulate, it is reasonable to make the further assumption that $\mathcal{E}(\mathbf{z})$ is well approximated by a quadratic polynomial in Eq. (1). To explore the constraints of symplectomorphisms, we will also compute

$$E_{SL(2n)} := \inf_{\mathbf{b} \in \mathbb{R}^{2n}, A \in SL(2n)} \int \mathcal{E}(\mathbf{z}) f_0(A^{-1}\mathbf{z} + \mathbf{b}) d^{2n}\mathbf{z}, \quad (2)$$

which we refer to as the linear Gardner energy. Since every symplectic matrix is area preserving, we will have that $E_{SL(2n)} \leq E_{Sp(2n)}$, but we should not generally expect equality when $n > 1$. Indeed, we will show this to be the case below.

II. ANALYSIS

We now answer the question set forth. We will assume that \mathcal{E} is a quadratic polynomial of the form $\mathcal{E}(\mathbf{z}) = a + \mathbf{b} \cdot \mathbf{z} + \mathbf{z}^T V \mathbf{z}$ where V is a symmetric matrix. To ensure $\mathcal{E}(\mathbf{z})$ has a unique minimum, we assume that V is positive definite. We will show in Appendix 1 how positive semi-definite V can be treated. Since V is invertible, there is a constant vector \mathbf{d} and a scalar V_0 such that $\mathcal{E}(\mathbf{z}) = V_0 + (\mathbf{z} - \mathbf{d})^T V (\mathbf{z} - \mathbf{d})$. For any nonnegative measurable function f_0 such that $\int |\mathbf{z}|^2 f_0 < \infty$, we define the moments

$$\begin{aligned} N &:= \int f_0(\mathbf{z}) d^{2n}\mathbf{z}, \\ \mathbf{c} &:= \frac{1}{N} \int \mathbf{z} f_0(\mathbf{z}) d^{2n}\mathbf{z}, \\ H &:= \int (\mathbf{z} \otimes \mathbf{z}) f_0(\mathbf{z} + \mathbf{c}) d^{2n}\mathbf{z}. \end{aligned} \quad (3)$$

These moments can be computed either directly or by differentiating the Fourier transform of f_0 . It is important to note that H is a positive definite matrix since for any $\xi \neq 0 \in \mathbb{R}^{2n}$, $\xi^T H \xi = \int (\xi \cdot \mathbf{z})^2 f_0(\mathbf{z} + \mathbf{c}) d^{2n}\mathbf{z} > 0$. For any fixed $A \in SL(2n)$, we have that

$$\begin{aligned} &\inf_{\mathbf{b} \in \mathbb{R}^{2n}} \int \mathcal{E}(\mathbf{z}) f_0(A^{-1}\mathbf{z} + \mathbf{b}) \\ &= \inf_{\mathbf{b} \in \mathbb{R}^{2n}} \int \mathcal{E}(\mathbf{z}) f_0(A^{-1}(\mathbf{z} - \mathbf{d} - \mathbf{b}) + \mathbf{c}) d^{2n}\mathbf{z} \\ &= V_0 N + \inf_{\mathbf{b} \in \mathbb{R}^{2n}} \int (\mathbf{z} + \mathbf{b})^T V (\mathbf{z} + \mathbf{b}) f_0(A^{-1}\mathbf{z} + \mathbf{c}) d^{2n}\mathbf{z} \\ &= V_0 N + \int \mathbf{z}^T V \mathbf{z} f_0(A^{-1}\mathbf{z} + \mathbf{c}) d^{2n}\mathbf{z} + N \inf_{\mathbf{b} \in \mathbb{R}^{2n}} \mathbf{b}^T V \mathbf{b} \\ &= V_0 N + \int \mathbf{z}^T V \mathbf{z} f_0(A^{-1}\mathbf{z} + \mathbf{c}) d^{2n}\mathbf{z}, \end{aligned} \quad (4)$$

where we have used that $\int \mathbf{z} f_0(A^{-1}\mathbf{z} + \mathbf{c}) d^{2n}\mathbf{z} = \mathbf{A} \int (\mathbf{z}' - \mathbf{c}) f_0(\mathbf{z}') d^{2n}\mathbf{z}' = 0$ and $\mathbf{b}^T V \mathbf{b} \geq 0$. Equation (4) can be paraphrased as saying that the optimal \mathbf{b} is such that the center of mass of f_0 lies at the potential minimum. We now must find the optimal A . We start with the simplification

$$\int \mathbf{z}^T V \mathbf{z} f_0(A^{-1}\mathbf{z} + \mathbf{c}) d^{2n}\mathbf{z} = \text{tr}(VAHA^T). \quad (5)$$

Thus, to compute either $E_{SL(2n)}$ or $E_{Sp(2n)}$ we must solve a trace minimization problem.

A. Linear Gardner energy

We first analyze the case that $A \in SL(2n)$ since the linear algebra is more familiar. Given any symmetric matrix M there exists a special orthogonal, hence $SL(2n)$, matrix O_M such that $O_M^T M O_M = D_M$, where D_M is diagonal. We may write any $A \in SL(2n)$ as $A = O_V B O_H^T$ with $B \in SL(2n)$. Our trace to be minimized then becomes

$$\begin{aligned} \text{tr}(VAHA^T) &= \text{tr}(D_V B D_H B^T) \\ &= \text{tr}((D_H^{1/2} B^T D_V^{1/2})(D_H^{1/2} B^T D_V^{1/2})^T), \end{aligned} \quad (6)$$

where we have used that D_H and D_V are positive definite in taking their roots. Since $X := (D_H^{1/2} B^T D_V^{1/2})(D_H^{1/2} B^T D_V^{1/2})^T$ is positive definite, X has positive eigenvalues so we may apply the AM-GM inequality to the eigenvalues of X to derive that

$$\text{tr}(X) \geq 2n \det(X)^{1/2n} = 2n \det(VH)^{1/2n}, \quad (7)$$

where we used $\det(B) = 1$. Equality in Eq. (7) is obtainable iff

$B = \det(HV)^{1/4n} D_V^{-1/2} O D_H^{-1/2}$ for some $O \in SO(2n)$. This allows us to conclude that

$$E_{SL(2n)} = NV_0 + 2n \det(VH)^{1/2n}. \quad (8)$$

B. Linear Gromov energy

We now restrict to the trickier case that $A \in Sp(2n)$. Williamson's theorem [35] states that for any symmetric, positive-definite matrix M there exists a symplectic matrix S_M such that

$$S_M^T M S_M = \begin{bmatrix} \mathcal{D}_M & 0 \\ 0 & \mathcal{D}_M \end{bmatrix} = \mathcal{D}_M \oplus \mathcal{D}_M, \quad (9)$$

where $\mathcal{D}_M = \text{diag}(\lambda_1^M, \dots, \lambda_n^M)$ with $\lambda_1^M \geq \dots \geq \lambda_n^M > 0$. The values λ_i^M are called the symplectic eigenvalues of M . For computational purposes, we note that λ_i^M is a symplectic eigenvalue of M iff $\pm i\lambda_i^M$ are eigenvalues of JM where $J = \begin{bmatrix} 0 & \mathbb{I}_n \\ -\mathbb{I}_n & 0 \end{bmatrix}$ is the standard symplectic form [36]. Writing $A = S_V B S_H^T$ with $B \in Sp(2n)$ we find that

$$\text{tr}(VAHA^T) = \text{tr}(B^T (\mathcal{D}_V \oplus \mathcal{D}_V) B (\mathcal{D}_H \oplus \mathcal{D}_H)). \quad (10)$$

If B is chosen to be the symplectic linear map relabeling the canonically conjugate pairs by the formula $x_i \mapsto x_{n+1-i}$ and $p_i \mapsto p_{n+1-i}$, we obtain the inequality

$$\inf_{A \in Sp(2n)} \text{tr}(VAHA^T) \leq 2 \sum_{i=1}^n \lambda_i^H \lambda_{n+1-i}^V. \quad (11)$$

It is a nontrivial fact, which we hold off on proving until Theorem 1, that the inequality in Eq. (11) can be replaced by equality. Hence,

$$E_{Sp(2n)} = NV_0 + 2 \sum_{i=1}^n \lambda_i^H \lambda_{n+1-i}^V. \quad (12)$$

Just as with the linear Gardner energy, there is a continuous family of $Sp(2n)$ matrices minimizing the linear Gromov energy. Notably, for any n tuple of angles $(\theta_1, \dots, \theta_n)$, we can define the symplectic rotation matrix

$$X(\theta_i) : \begin{bmatrix} x_i \\ p_i \end{bmatrix} \mapsto \begin{bmatrix} \cos(\theta_i) & \sin(\theta_i) \\ -\sin(\theta_i) & \cos(\theta_i) \end{bmatrix} \begin{bmatrix} x_i \\ p_i \end{bmatrix}. \quad (13)$$

Replacing B with $X(\theta_i)B$ in Eq. (10) leaves the trace invariant so there is a $SO(2)^n$ family of linear maps taking f_0 to a minimal energy configuration. This family of matrices is enlarged when the symplectic eigenvalues of either V or H are degenerate.

It is interesting to compare Eqs. (8) and (12). Notably, the AM-GM inequality shows that $E_{Sp(2n)} \geq E_{SL(2n)}$ since

$$2 \sum_{i=1}^n \lambda_i^H \lambda_{n+1-i}^V \geq 2n (\det(HV))^{1/2n}, \quad (14)$$

with equality iff $\lambda_i^H \lambda_{n+1-i}^V \equiv \text{const.}$ In particular, $E_{Sp(2)} = E_{SL(2)}$ as expected.

C. Example 1

We first compute an easy example in $n = 2$. For $\epsilon > 0$, suppose that $\mathcal{E}(\mathbf{z}, \epsilon) = x^2 + \epsilon^2 y^2 + p_x^2 + p_y^2$. Suppose further that $f_0(\mathbf{z}, R) = \frac{6}{R^2 |B(R)|} \chi_{B(R)} = \frac{6}{R^2 |B(R)|} \Theta(R^2 - |\mathbf{z}|^2)$ is a rescaled indicator function on the ball. We compute that $N = 1$, $\mathbf{c} = 0$, $H(R) = \mathbb{I}_4$, $V_0 = 0$, $\mathbf{d} = 0$, and $V(\epsilon) = \text{diag}(1, \epsilon^2, 1, 1)$. The initial energy stored in f_0 is $E[f_0] = (3 + \epsilon)$. For ϵ small, we should expect that $E_{SL(4)}$ is small since f_0 can be squeezed onto the y axis via area-preserving maps. Indeed, Eq. (8) gives us that

$$E_{SL(4)} = \epsilon^{1/2}. \quad (15)$$

Dividing this equation by $E[f_0]$, we can alternatively compute the inaccessible energy fraction $F_{SL(2n)} := \frac{E_{SL(2n)}}{E[f_0]}$ to be

$$F_{SL(4)} = \frac{4\epsilon^{1/2}}{3 + \epsilon^2}. \quad (16)$$

The distribution function after an energy minimizing linear map is $f_0 \circ \phi^{-1} = \frac{1}{|B(R)|} \Theta(R^2 - \epsilon^{-1/2}(x^2 + p_x^2 + p_y^2) - \epsilon^{3/2}y^2)$ which looks as expected.

In contrast, the linear Gromov's nonsqueezing theorem prohibits squeezing f_0 onto the y axis via linear symplectomorphisms. This implies we should find $E_{Sp(2n)}$ to be finite in the limit $\epsilon \rightarrow 0^+$. $V(\epsilon)$ can be symplectically diagonalized by $S_V = \text{diag}(1, \epsilon^{-1/2}, 1, \epsilon^{1/2})$ giving $\mathcal{D}_V = \text{diag}(1, \epsilon)$. From Eq. (12) we can then compute

$$E_{Sp(4)} = \frac{1}{2}(1 + \epsilon), \quad (17)$$

which limits to a finite value as expected. We compute the inaccessible energy fraction $F_{Sp(2n)} := \frac{E_{Sp(2n)}}{E[f_0]}$ as

$$F_{Sp(4)} = \frac{2 + 2\epsilon}{3 + \epsilon^2}. \quad (18)$$

Figure 1 shows E_G and F_G against ϵ for $G \in \{Sp(4), SL(4)\}$.

The distribution function after the energy-minimizing mapping is $f_0 \circ \phi^{-1} = \frac{6}{R^2 |B(R)|} \Theta(R^2 - \epsilon y^2 - \epsilon^{-1} p_y^2 - x^2 - p_x^2)$ which looks as expected since one cannot symplectically compress f_0 along the x axis without increasing p_x by a corresponding factor. This addresses a question put forth in [19, p.4]. It is interesting to note that $E_{Sp(4)}$ and $E_{SL(4)}$ can differ, either in difference or ratio, by an arbitrarily large amount. As we show in Sec. II E, Eq. (16) is also the correct formula for the Gardner energy. This refutes a conjecture in [19, p.4] that the linear Gromov energy must be close to the (nonlinear) Gromov energy, $\inf_{\phi \in \text{Ham}(\mathcal{P})} E[f_0 \circ \phi^{-1}]$.

D. Example 2

Suppose we continue to consider the energy function $\mathcal{E}(\mathbf{z}, \epsilon) = x^2 + \epsilon^2 y^2 + p_x^2 + p_y^2$ but f_0 is instead given by

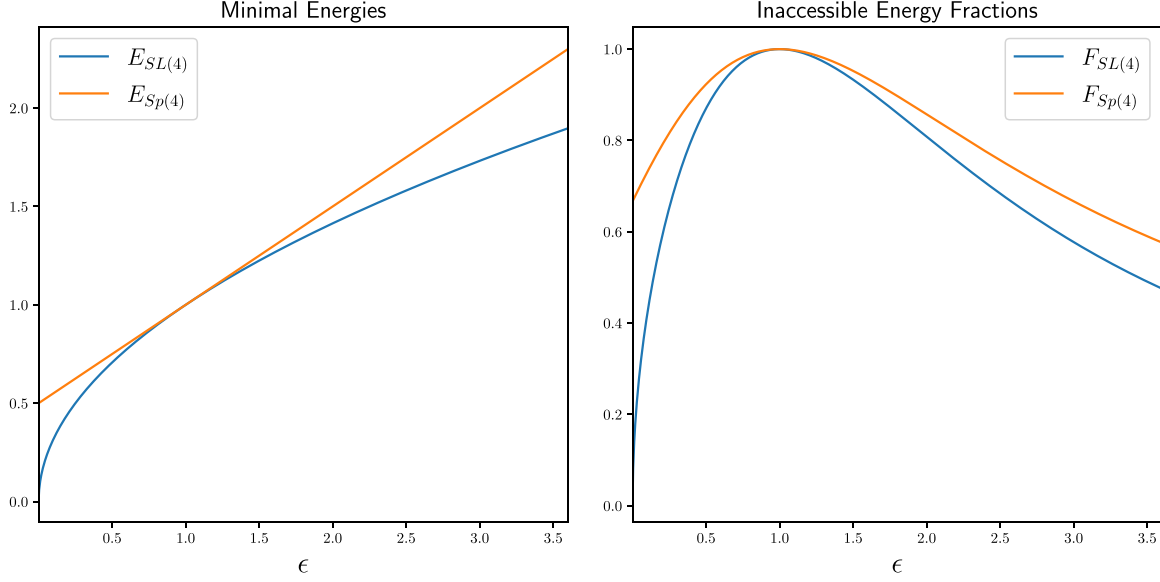


FIG. 1. Minimal energies for example Sec. II C.

$f_0(\mathbf{z}) = \frac{1}{2(2\pi)^2} \exp(-\frac{1}{2}(p_x^2 + p_y^2 + \frac{1}{4}x^2 + y^2))$. Then either by direct integration or by computing the Fourier transform of f_0 , we learn that $H = \text{diag}(4, 1, 1, 1)$. H is symplectically diagonalized by $S_H = \text{diag}(2^{-1/2}, 1, 2^{1/2}, 1)$, implying $D_H = \text{diag}(2, 1)$. The linear Gardner energy is therefore

$$E_{SL(4)}(\epsilon) = 4\sqrt{2\epsilon}, \quad (19)$$

whereas

$$E_{Sp(2n)}(\epsilon) = \begin{cases} 4\epsilon + 2 & 0 < \epsilon < 1 \\ 2\epsilon + 4 & \epsilon \geq 1 \end{cases}. \quad (20)$$

Figure 2 shows E_G and F_G against ϵ for $G \in \{Sp(4), SL(4)\}$.

Equation (20) illustrates that even for smoothly varying H and V , the linear Gromov energy does not need to vary smoothly. A lack of differentiability of $E_{Sp(2n)}(\epsilon)$ at points

where either D_V or D_H has repeated symplectic eigenvalues is to be expected since the symplectic eigenvalue pairing in Eq. (12) is generically reordered. At these points, a sort of saturation occurs and a “large” symplectomorphism exchanging canonical pairs must be applied before further energy can be extracted from f_0 via smoothly varying symplectomorphisms. The lack of such nondifferentiable behavior in the linear Gardner energy exemplifies the flexibility of area-preserving maps in comparison to symplectic maps.

E. Example 3: Symplectic equivalence of ellipsoids

We now consider a subclass of problems for which the Gardner and linear Gardner energy agree. We will use our observations to prove that two ellipsoids are linearly symplec-

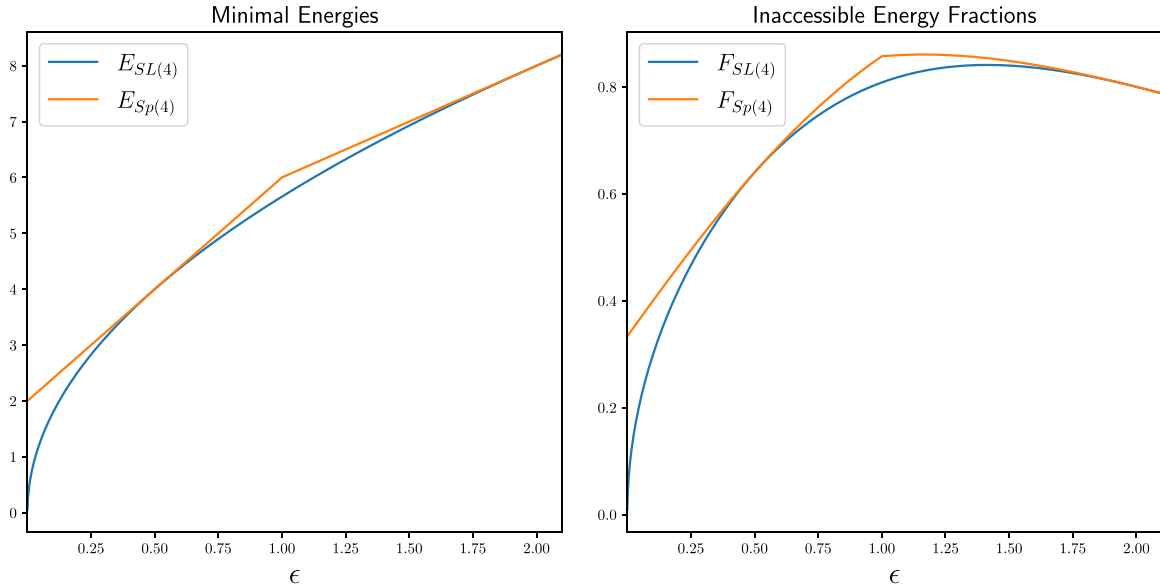


FIG. 2. Minimal energies for example Sec. II D.

tomorphic iff their defining matrices have the same symplectic eigenvalues.

Given $2n \times 2n$ symmetric, positive-definite matrix M , we define the ellipsoid $El(M) := \{\mathbf{z} \in \mathbb{R}^{2n} : \mathbf{z}^T M \mathbf{z} \leq 1\}$. For M and M' fixed symmetric, positive-definite matrices with the same determinant, we take $\mathcal{E}(\mathbf{z}) = \mathbf{z}^T M \mathbf{z}$ and $f_0 = \chi_{El(M')}$. Since $\det(M) = \det(M')$, the areas of $El(M')$ and $El(M)$ agree. It is easy to check that there is a $SL(2n)$ matrix A mapping $El(M')$ to $El(M)$. It is also not hard to see that $E_G = E_{SL(2n)} = \int \mathcal{E}(\mathbf{z}) f_0(A^{-1}\mathbf{z})$. Indeed, suppose $\phi : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ is any area-preserving diffeomorphism not mapping $El(M')$ to $El(M)$. Then part of $\phi(El(M'))$ lies outside of $El(M)$, say the set S . Since $\mathcal{E}|_S > 1$, more energy could be extracted from f_0 by moving S into $El(M)$. Given that the volume of $El(M)$ and $El(M')$ agree, this proves that the optimal amount of energy from f_0 is extracted by mapping $El(M')$ to $El(M)$. Conversely, this argument shows that if a map ϕ minimizes $E[f_0 \circ \phi^{-1}]$, then ϕ necessarily maps $El(M')$ to $El(M)$.

By our observations, we conclude that $El(M)$ and $El(M')$ are linearly symplectomorphic iff both $\det(M) = \det(M')$ and $E_{Sp(2n)} = E_{SL(2n)}$. For the given f_0 and \mathcal{E} , we have that $V_0 = 0$, $\mathbf{c} = 0$, $V = M$, and $H \propto (M')^{-1}$. The symplectic eigenvalues of H are proportional to the reciprocals of the symplectic eigenvalues of M' so, with a uniform proportionality constant, $\lambda_i^H \propto \frac{1}{\lambda_{n+1-i}^{M'}}$. Hence, by equality condition for Eq. (14), we learn that $El(M)$ and $El(M')$ are linearly symplectomorphic iff $\frac{\lambda_i^M}{\lambda_i^{M'}} \equiv \text{const}$ and $\det(M) = \det(M')$. These two conditions are combined to give $\lambda_i^M = \lambda_i^{M'}$, concluding the proof.

F. Example 4: Gromov's nonsqueezing theorem

In this last example, we prove the affine Gromov nonsqueezing theorem using our energy-minimization theory. Namely, we will show for $R > r$ there is no affine symplectomorphism taking the ball $B(R)$ into the symplectic cylinder $\mathcal{Z}(r)$.

We define the energy $\mathcal{E}(\mathbf{z}) = x_1^2 + p_1^2$ so $V = e_{x_1}e_{x_1} + e_{p_1}e_{p_1}$. For $R > r$ we take $f_0 = \chi_{B(R)}$. We compute $H = \frac{R^{2n+2}}{2n+2(2n)} |S^{2n-1}| \mathbb{I}_{2n}$. Using the result of Appendix 1 to justify the limit swap $\lim_{\epsilon \rightarrow 0^+} \inf_{X \in Sp(2n)} \text{tr}(X(V + \epsilon \mathbb{I}_{2n})X^T H) = \inf_{X \in Sp(2n)} \text{tr}(XVX^T H)$, we compute $E_{Sp(2n)} = \frac{R^{2n+2}}{n+1(2n)} |S^{2n-1}| = \int f_0(\mathbf{z}) \mathcal{E}(\mathbf{z}) d^{2n}\mathbf{z}$. This shows there is no way to reduce the energy of f_0 by symplectomorphisms. Under any affine symplectomorphism, f_0 transforms into the characteristic function of an ellipsoid. We will show for every ellipsoid $El \subset \mathcal{Z}(r)$ linearly symplectomorphic to $B(R)$ that $\int_{El} \mathcal{E} < E_{Sp(2n)}$, proving no such El exists.

If $\mathbf{z}_0 + El(M^{-1})$ is an ellipsoid in $\mathcal{Z}(r)$ then we can move $\mathbf{z}_0 + El(M^{-1})$ to the coordinate origin without increasing its energy and while keeping it confined to $\mathcal{Z}(r)$. Without loss of generality, we thereby set $\mathbf{z}_0 = 0$. We compute that

$$\int_{El(M^{-1})} x_1^2 + p_1^2 = \sqrt{\det(M)} \frac{|S^{2n-1}|}{2n(2n+2)} (M_{1,1} + M_{n+1,n+1}). \quad (21)$$

By assumption, $El(M^{-1}) \subset \mathcal{Z}(r)$, so for every unit vector \mathbf{v} we have $\langle x_1, M^{1/2} \mathbf{v} \rangle^2 + \langle p_1, M^{1/2} \mathbf{v} \rangle^2 \leq r^2$. Since $M^{1/2}$ is self-adjoint, taking $\mathbf{v} = \|M^{1/2} x_1\|^{-1} M^{1/2} x_1$ shows that

$\|M^{1/2} x_1\|^2 \leq r^2$. Similarly, we conclude $\|M^{1/2} p_1\|^2 \leq r^2$. These inequalities imply that $M_{1,1} \leq r^2$ and $M_{n+1,n+1} \leq r^2$. If, for the sake of contradiction, $B(R)$ is linearly symplectomorphic to the ellipsoid $El(M^{-1})$ then

$$\frac{R^{2n+2}}{2n(n+1)} |S^{2n-1}| \leq \frac{r^2 \sqrt{|M|} |S^{2n-1}|}{2n(n+1)}. \quad (22)$$

Since the areas of $B(R)$ and $El(M)$ necessarily agree, it must be that $|M| = R^{4n}$. Simplifying, we have the erroneous inequality $R^2 \leq r^2$, the desired contradiction.

Theorem 1. For any $2n \times 2n$, symmetric, positive-definite matrices V, H with respective symplectic eigenvalues $\lambda_1^V \geq \dots \geq \lambda_{2n}^V > 0$ and $\lambda_1^H \geq \dots \geq \lambda_{2n}^H > 0$,

$$\inf_{S \in Sp(2n)} \text{tr}(SVS^T H) = 2 \sum_{i=1}^n \lambda_i^H \lambda_{n+i-1}^V.$$

Proof. To complete the proof of Theorem 1, we have only to show that

$$\inf_{S \in Sp(2n)} \text{tr}(S^T (\mathcal{D}_V \oplus \mathcal{D}_V) S (\mathcal{D}_H \oplus \mathcal{D}_H)) \geq 2 \sum_{i=1}^n \lambda_i^H \lambda_{n+1-i}^V. \quad (23)$$

We will prove Eq. (23) with a method similar to Son and Stykel [37] using a theorem of Liang *et al.* [38].

Theorem 2. [38, p.489] Let $A, B \in \mathbb{C}^{d \times d}$ and $D_{\pm} \in \mathbb{C}^{k_{\pm} \times k_{\pm}}$ be Hermitian matrices such that $A \neq 0$, B has both positive and negative eigenvalues, $k_+ + k_- = d$, $D_{\pm} \geq 0$, and the matrix pencil $A - \lambda B$ is positive semidefinite. Let

$$J_k = \begin{bmatrix} \mathbb{I}_{k_+} & 0 \\ 0 & -\mathbb{I}_{k_-} \end{bmatrix}, \quad D = \begin{bmatrix} D_+ & 0 \\ 0 & D_- \end{bmatrix},$$

and let $\omega_1^+ \geq \dots \geq \omega_{k_+}^+ \geq 0$ and $\omega_1^- \geq \dots \geq \omega_{k_-}^- \geq 0$ be the eigenvalues of D_+ and D_- , respectively. Let $\lambda_{k_-}^- \leq \dots \leq \lambda_1^- \leq \lambda_1^+ \leq \dots \leq \lambda_{k_+}^+$ be the eigenvalues of the matrix pencil $A - \lambda B$. Then

$$\inf_{X^{\dagger} B X = J_k} \text{tr}(D X^{\dagger} A X) = \sum_{i=1}^{k_+} \lambda_i^+ \omega_i^+ - \sum_{i=1}^{k_-} \lambda_i^- \omega_i^-.$$

For our purposes, we will take $d = 4n$, $k_{\pm} = 2n$, $D_+ = D_- = \mathcal{D}_V \oplus \mathcal{D}_V$, and $A = \mathcal{D}_H \oplus \mathcal{D}_H \oplus \mathcal{D}_H \oplus \mathcal{D}_H$. Trivially, $\omega_{2i-1}^+ = \omega_{2i}^+ = \omega_{2i-1}^- = \omega_{2i}^- = \lambda_i^V$ are the symplectic eigenvalues of V . As in [37], for any $S \in Sp(2n)$ we define

$$B = \begin{bmatrix} 0 & J \\ -J & 0 \end{bmatrix}, \quad \mathcal{X}(S) = \frac{1}{\sqrt{2}} \begin{bmatrix} S & S J^T \\ S J^T & S \end{bmatrix}. \quad (24)$$

B is Hermitian and has eigenvalues ± 1 , each with multiplicity $2n$. Considering $\lambda = 0$ shows the matrix pencil $A - \lambda B$ is positive definite. It's also easy to verify that the eigenvalues of the pencil $A - \lambda B$ are $\pm \lambda_i^H$, each with multiplicity two. Since $\mathcal{X}(S)^{\dagger} B \mathcal{X}(S) = J_k$, and since all the hypotheses of the Theorem 2 are satisfied,

$$\begin{aligned} \text{tr}(D \mathcal{X}^{\dagger}(S) A \mathcal{X}(S)) &= 2 \text{tr}(S (\mathcal{D}_V \oplus \mathcal{D}_V) S^T (\mathcal{D}_H \oplus \mathcal{D}_H)) \\ &\geq 4 \sum_{i=1}^n \lambda_i^V \lambda_{n+1-i}^H. \end{aligned} \quad (25)$$

Since Eq. (25) holds for every symplectic matrix, we conclude our proof of Theorem 1.

III. DISCUSSION

Efficiently extracting energy from particle distributions is vital for making fusion energy a reality. It is therefore critical to have useful bounds for the amount of extractable energy in a plasma. Complicating the minimization of energy in a system is the constraint that particles ideally evolve by Hamiltonian symplectomorphisms. With this constraint in mind, we were able to compute the minimal energy of a system in two extremes: one has arbitrarily fine control of phase space, and one has only very coarse control of phase space. The former energy we showed was equivalent to the Gardner energy E_G , while the latter we showed could be computed with Eq. (12). Both of these cases were interesting, highlighting the flexibility vs rigidity of symplectomorphisms. Nonetheless, our analysis seems unsatisfactory in places. The distribution function becomes pathological when minimizing energy using nonlinear symplectomorphisms. On the other hand, in coarse graining the problem, we were left with transformations too blunt to efficiently extract energy from many classical nonthermal distributions (see Appendix 3). Given the rich theory already developed here, it is likely a rich medium ground may be found. There are many nonequivalent approaches to explore, but we leave this for future work.

Our analysis is not meant to conclude the question of energy extraction. We have not addressed thermodynamic or realistic technical constraints. Rather, we view this work as a first step toward understanding phase space engineering and the constraints of symplectomorphisms. To this end, much work remains to be done. Even within the limited problem of exacting energy with a restricted set of symplectomorphisms, many families of transforms may be fruitfully considered. Perhaps the most important family of transformations to consider are those corresponding to a Hamiltonian of the form $\mathcal{H} = \frac{1}{2}(p - qA(x, t))^2 + \phi(x, t)$, or even simply those of the form $\mathcal{H} = \frac{p^2}{2} + \phi(x, t)$. Understanding these flows would lead to a better understanding of electromagnetic phase-space engineering. We leave such a study to a future work.

While not always sharp, the linear Gromov energy gives a good benchmark for energy extraction, even when only the fluid moments of a distribution are known. The linear Gromov energy can also be useful when computing the allowed energy under nonlinear symplectomorphisms. Composing nonlinear maps with optimal linear maps can give a computationally feasible way to compute the Gardner energy. In a future publication, we will present a machine learning inspired numerical approach to computing upper bounds on the nonlinear Gromov energy with both free and periodic spatial boundary conditions. This approach will allow us to investigate the effect of regularizing the Gromov-Gardner problem by constraining the gradient of the symplectomorphism, a question which is functionally intractable within the linear theory presented here.

Perhaps the biggest goal of this work has been to show how physical problems and mathematical theory can fruitfully coalesce. In deriving the ground state energy of a particle distribution, we were led to prove a trace-minimization theorem. Many more results may yet be derived from the problem of phase-space engineering. These results would deepen our

understanding of how particles can be manipulated, and might themselves be of mathematical interest.

ACKNOWLEDGMENT

This research was supported by the DOE Grant No. DE-AC02-09CH1146.

DATA AVAILABILITY

No data were created or analyzed in this study.

APPENDIX

1. Degenerate Potentials

Suppose the potential matrix V is merely positive semidefinite. Then for either $G = Sp(2n)$ or $G = SL(2n)$, we will show that $E_G = NV_0 + \inf_{X \in G} \text{tr}(XVX^TH) = NV_0 + \lim_{\epsilon \rightarrow 0} \inf_{X \in G} \text{tr}(X(V + \epsilon \mathbb{I}_{2n})X^TH)$. Since $V + \epsilon \mathbb{I}_{2n}$ is positive definite for $\epsilon > 0$, this allows us to compute E_G using Eq. (8) or Eq. (12).

Proof. Define for $\epsilon \geq 0$ the nonnegative function $E_G(\epsilon) = NV_0 + \inf_{X \in G} \text{tr}(X(V + \epsilon \mathbb{I})XH)$. We must show $E(\epsilon)$ is continuous at $\epsilon = 0$. It is easy to verify that $E(\epsilon)$ is monotonously increasing, and in particular that $E(0) \leq E(\epsilon)$ for any $\epsilon \geq 0$. Let $\delta > 0$ be fixed. Let X_0 be such that $\text{tr}(X_0VX_0H) - E(0) < \eta/2$. Then for any $\epsilon > 0$ such that $\epsilon \text{tr}(X_0VX_0H) < \delta/2$, the triangle inequality implies

$$\delta > \text{tr}(X_0(V + \epsilon \mathbb{I}_{2n})X_0H) - E(0) \geq E(\epsilon) - E(0) \geq 0. \quad (\text{A1})$$

Hence, $\lim_{\epsilon \rightarrow 0} E(\epsilon) = E(0)$, proving the result. ■

2. Gardner's restacking and Hamiltonian approximations

We elaborate on Gardner's restacking algorithm and Hamiltonian approximations thereof. For simplicity, we will assume $\mathcal{P} = \mathbb{R}^{2n}$, but there is no obstruction in allowing \mathcal{P} to be a manifold. We will also make the simplifying assumption that f_0 is continuous, but this restriction can easily be lifted. As always, we assume \mathcal{E} is bounded below and that $f_0\mathcal{E}$ is integrable.

When $\{\phi\}$ are required to be area preserving, but possibly discontinuous, Gardner's restacking allows us to compute $E_G = \inf_{\{\phi\}} E[f_0 \circ \phi^{-1}]$. To describe the algorithm in a mathematically rigorous manner, let $n \in \mathbb{N}$ be a natural number. Let $h = 2^{-n}$. We define the lattice $\Lambda^h = h\mathbb{Z}^{2n}$. For $\lambda \in \Lambda^h$, we define $S_\lambda^h = \lambda + [0, h]^{2n}$. We break phase space into a collection of disjoint squares of side length h , viz $\mathcal{P} = \bigsqcup_{\lambda \in \Lambda^h} S_\lambda^h$.

Since Λ^h is countable, we may index every element $\lambda \in \Lambda^h$ by a natural number, denoted by λ_i . We choose two such indexings. The first denoted by a subscript f is such that $f_0(\lambda_{1,f}) \geq f_0(\lambda_{2,f}) \geq \dots \geq 0$. The second indexing of Λ^h , denoted with a subscript \mathcal{E} , is such that $\mathcal{E}(\lambda_{1,\mathcal{E}}) \leq \mathcal{E}(\lambda_{2,\mathcal{E}}) \leq \dots$. We then define a bijective, area-preserving map $\phi^{(n)}$ sending $S_{\lambda_{i,f}}^h \mapsto S_{\lambda_{i,\mathcal{E}}}^h$, say the identity map on squares. Essentially, $\phi^{(n)}$ permutes equal-area sets in phase space until we have minimized the energy of the approximate distribution function $f_0^{(n)} = \sum_{\lambda \in \Lambda^h} f_0(\lambda^h) \chi_{S_\lambda^h}$. The ground state energy of $f_0^{(n)}$ is computed to be $E^{(n)} = \sum_{i=1}^{\infty} f(\lambda_{i,f}) \mathcal{E}(\lambda_{i,\mathcal{E}})$. E_G is computed as the limit of $E^{(n)}$ as n tends toward infinity.

We note that the distribution function after the energy minimizing mapping, formally denoted $f_\infty = \lim_{n \rightarrow \infty} f^{(n)} \circ \phi^{(n)}$, may or may not uniquely exist. If \mathcal{E} has a level set of nonzero measure, then one generically has an infinite number of possible final states, f_∞ . If \mathcal{E} is nondegenerate except around its minimum, then it is likely that f_∞ is a well-posed measure. In 1D for example, if $f_0(p)$ and $\mathcal{E}(p) = \frac{p^2}{2}$, then $f_\infty(p)$ is the symmetric decreasing rearrangement of f_0 [39,40].

To show that $\inf_{\phi \in \text{Ham}(\mathcal{P})} E[f_0 \circ \phi^{-1}] = E_G$, we show that it is possible to approximate the steps in Gardner's restacking with Hamiltonian maps.

For convenience, we will assume that f_0 is compactly supported. Let h be as before. Let R be a square containing the support of f_0 . We can apply Gardner's restacking to R as described above. Let S_1, \dots, S_N be squares of side length h covering R . Gardner's algorithm permutes these squares, mapping $S_i \mapsto S_{\sigma(i)}$ in some area-preserving manner. In order to do the same with Hamiltonian maps, let $\delta > 0$. We remove a $\frac{\delta}{2N}$ amount of area from the edges of each square S_i making each trimmed square \tilde{S}_i closed and disjoint. Then Katok [30, p.545] shows there is a Hamiltonian diffeomorphism $\psi^{(n)}$ supported in a neighborhood of R almost mapping $\tilde{S}_i \mapsto \tilde{S}_{\sigma(i)}$ in the precise sense that $|(\psi^{(n)})^{-1}(\tilde{S}_{\sigma(i)}) \cap \tilde{S}_i| < \frac{\delta}{2N}$. As we send $\delta \rightarrow 0$, we closely approximate a map sending S_i to $S_{\sigma(i)}$. The region of measure δ not mapped according to Gardner's algorithm contributes an error of at most $\delta \|\mathcal{E}\|_{L^\infty(R)} \|f_0\|_{L^\infty(R)}$ to the energy integral. Taking $\delta = \exp(-1/h)$ and taking the limit $h \rightarrow 0$ therefore shows that $E_G = \inf_{\phi \in \text{Ham}(\mathcal{P})} E[f_0 \circ \phi^{-1}]$.

It is interesting to compare this result to that of Kolmes and Fisch [21]. They showed that Gardner's restacking could be arbitrarily approximated by diffusive operations. Given that many fundamental processes can lead to apparent diffusion, it is perhaps not surprising that Gardner's restacking can be approximated by Hamiltonian diffeomorphisms.

3. Linear vs nonlinear operations

While the "bluntness" of linear maps avoided the problem of overcontrolling phase space, this same bluntness prevents energy from being extracted in many classical scenarios. Consider, for example, a 1D idealized bump on tail distribution

$$f_0(p) = \frac{n_0}{\sqrt{2\pi T}} \exp(-p^2/2T) + n_1 \delta(p - p_0) = f_{\text{eq}} + f_{\text{bump}},$$

where T, n_1 , and n_0 are constants. The energy function is the classical energy $\mathcal{E} = \frac{p^2}{2}$. f_0 is spatially homogeneous, but it still makes sense to speak of the energy density. Using R-F waves, one can cause f_0 to form a quasilinear plateau around $p = p_0$, extracting energy in the process. Optimally, one could move f_{bump} to $p = 0$ by Gardner restacking. The resulting distribution function has the lowest possible energy since f_{eq} is symmetric and decreasing. This extracts an energy per unit volume of $p_0^2 n_1$ giving a Gardner energy density of $\frac{n_0 T}{2}$.

With linear operations, none of the previous operations are allowed. The only x independent, affine, area-preserving map is the shift map $p \mapsto p + \delta p$. To extract energy from f_0 , all we can do is shift f_0 until there is no net momentum, i.e., $\int p f_0(p - \delta p) dp = 0$. The energy density of

$f_0(p - \delta p)$ is

$$\frac{n_0}{2}(\delta p^2 + T) + \frac{n_1(p_0 + \delta p)^2}{2}, \quad (\text{A2})$$

which is minimized when $\delta p = -\frac{n_1 p_1}{n_0 + n_1}$. In this case, the energy extractable by linear maps is a pittance compared to the energy extractable by nonlinear maps.

We can play similar games in multiple dimensions with slightly more interesting results. For example, suppose we have a 3D bump on-tail

$$\begin{aligned} f_0(\mathbf{p}) &= n_0 \left(\frac{1}{2\pi T} \right)^{3/2} \exp(-p^2/2T) + n_1 \delta(\mathbf{p} - \mathbf{p}_0) \\ &= f_{\text{eq}} + f_{\text{bump}}. \end{aligned}$$

Then the Gardner energy is again obtained by moving f_{bump} to $\mathbf{p} = 0$. Restricting to linear maps, we can again shift momentum space in the direction of \mathbf{p}_0 space until $\int \mathbf{p} f_0(\mathbf{p} - \delta \mathbf{p}_0) d^3 \mathbf{p} = 0$. Now with more dimensions, we are additionally able to squeeze momentum space in the \mathbf{p}_0 direction while uniformly expanding the orthogonal plane. While still far from saturating the nonlinear energy bound, the linear energy is nonetheless closer due to the increased flexibility of higher dimensions.

Area-preserving linear maps cannot alter the internal structure of f_0 . It is no surprise that the linear Gardner and Gromov energies often depend only on the first few moments of f_0 . This is both a blessing and a curse. At least in many cases, $E_{Sp(2n)}$ and $E_{SL(2n)}$ can be computed directly from a fluid theory. No information about the kinetics is needed. The linear energies can therefore serve as a useful upper bound for the nonlinear energies in the absence of an exactly known distribution f_0 . However, when f_0 is known to some accuracy, and manipulations can be performed on scales smaller than the support of f_0 , the linear theory may fail to yield useful results.

4. Quadratic Hamiltonian and linear maps

To understand how linear maps may arise naturally, consider Hamiltonian's equations on \mathbb{R}^{2n} ,

$$\dot{\mathbf{z}} = J \nabla \mathcal{H}(\mathbf{z}). \quad (\text{A3})$$

If, at least to an approximation, $\mathcal{H}(\mathbf{z}, t) = \frac{1}{2} \mathbf{z}' \mathbb{H} \mathbf{z}$ is a quadratic polynomial, then

$$\dot{\mathbf{z}} = J \mathbb{H} \mathbf{z}. \quad (\text{A4})$$

The flowmap ϕ_t is easily computed as $\phi_t(\mathbf{z}_0) = \exp(J \mathbb{H} t) \mathbf{z}_0$. The matrix $\exp(J \mathbb{H} t)$ is symplectic, and thus ϕ_t is of the allowed linear form. Alone, however, matrices of the form $\exp(J \mathbb{H} t)$ do not give the full symplectic group. To get $Sp(2n)$, we must consider time-dependent Hamiltonian flows. If \mathbb{H} is allowed to be time dependent then the formal solution to Eq. (A4) is $\phi_t(\mathbf{z}_0) = T \exp(\int_0^t \mathbb{H}(s) ds) \mathbf{z}_0$, where $T \exp$ is the time-ordered exponential. It can be shown that matrices of the form $T \exp(\int_0^t \mathbb{H}(s) ds)$ are precisely the symplectic matrices. In fact, only using that $Sp(2n)$ is connected, it was shown by Wüstner that every symplectic matrix can be written as a product of the form $\exp(J \mathbb{H}_1) \exp(J \mathbb{H}_2)$ [41].

Affine linear maps are achieved by additionally considering Hamiltonians of the form $\mathcal{H} = -J \mathbf{b} \cdot \mathbf{z}$, in which case

Hamilton's equations read

$$\dot{\mathbf{z}} = \mathbf{b}. \quad (\text{A5})$$

Trivially, $\phi_t(\mathbf{z}_0) = \mathbf{z}_0 + \mathbf{b}t$. Composing this flow with a linear flow gives all the affine linear flows. Thus, affine symplectomorphisms arise naturally when the Hamiltonian of a system is well approximated by a quadratic polynomial.

-
- [1] R. Budny, *Nucl. Fusion* **58**, 096011 (2018).
 - [2] N. J. Fisch and J.-M. Rax, *Phys. Rev. Lett.* **69**, 612 (1992).
 - [3] N. J. Fisch, *Phys. Plasmas* **2**, 2375 (1995).
 - [4] N. Fisch and M. Herrmann, *Nucl. Fusion* **35**, 1753 (1995).
 - [5] M. C. Herrmann and N. J. Fisch, *Phys. Rev. Lett.* **79**, 1495 (1997).
 - [6] I. E. Ochs, N. Bertelli, and N. J. Fisch, *Phys. Plasmas* **22**, 082119 (2015).
 - [7] R. W. Bussard and N. A. Krall, *Fusion Technol.* **26**, 1326 (1994).
 - [8] N. Rostoker, M. W. Binderbauer, and H. J. Monkhorst, *Science* **278**, 1419 (1997).
 - [9] W. M. Nevins, *J. Fusion Energy* **17**, 25 (1998).
 - [10] V. Volosov, *Nucl. Fusion* **46**, 820 (2006).
 - [11] I. E. Ochs and N. J. Fisch, *Phys. Rev. Lett.* **127**, 025003 (2021).
 - [12] E. J. Kolmes, I. E. Ochs, and N. J. Fisch, *Phys. Plasmas* **29**, 110701 (2022).
 - [13] I. E. Ochs, E. J. Kolmes, M. E. Mlodik, T. Rubin, and N. J. Fisch, *Phys. Rev. E* **106**, 055215 (2022).
 - [14] V. R. Munirov and N. J. Fisch, *Phys. Rev. E* **107**, 065205 (2023).
 - [15] M. E. Mlodik, V. R. Munirov, T. Rubin, and N. J. Fisch, *Phys. Plasmas* **30**, 043301 (2023).
 - [16] I. E. Ochs and N. J. Fisch, *Phys. Plasmas* **31**, 012503 (2024).
 - [17] W. M. Nevins, *Science* **281**, 307 (1998).
 - [18] H. Qin, *Phys. Plasmas* **31**, 050601 (2024).
 - [19] H. Qin, E. J. Kolmes, M. Updike, N. Bohlsen, and N. J. Fisch, *Phys. Rev. E* **111**, 025205 (2025).
 - [20] C. S. Gardner, *Phys. Fluids* **6**, 839 (1963).
 - [21] E. J. Kolmes and N. J. Fisch, *Phys. Rev. E* **102**, 063209 (2020).
 - [22] E. J. Kolmes and N. J. Fisch, *Phys. Plasmas* **31**, 042109 (2024).
 - [23] E. J. Kolmes and N. J. Fisch, *Phys. Rev. E* **106**, 055209 (2022).
 - [24] I. Y. Dodin and N. J. Fisch, *Phys. Lett. A* **341**, 187 (2005).
 - [25] B. Dacorogna and J. Moser, *Annales de l'I.H.P. Analyse non Linéaire* **7**, 1 (1990).
 - [26] M. Gromov, *Invent. Math.* **82**, 307 (1985).
 - [27] H. Hofer and E. Zehnder, *Symplectic Invariants and Hamiltonian Dynamics*, Birkhäuser Advanced Texts Basler Lehrbücher (Birkhäuser Basel, Basel, 1994).
 - [28] G. Darboux, *Bulletin des Sciences Mathématiques et Astronomiques* **6**, 49 (1882).
 - [29] D. McDuff and D. Salamon, *Introduction to Symplectic Topology* (Oxford University Press, Oxford, 2017).
 - [30] A. B. Katok, *Math. USSR Izv.* **7**, 535 (1973).
 - [31] K. Sackel, A. Song, U. Varolgunes, J. Zhu, and J. Brendel, *Geom. Topol.* **28**, 1113 (2024).
 - [32] R. C. Davidson and H. Qin, *Physics of Intense Charged Particle Beams in High Energy Accelerators* (Imperial College Press and World Scientific, Singapore, 2001).
 - [33] H. Qin and R. C. Davidson, *Phys. Plasmas* **18**, 056708 (2011).
 - [34] H. Qin, R. C. Davidson, and B. G. Logan, *Phys. Rev. Lett.* **104**, 254801 (2010).
 - [35] J. Williamson, *Am. J. Math.* **58**, 141 (1936).
 - [36] K. Ikramov, *Moscow Univ. Comput. Math. Cybern.* **42**, 1 (2018).
 - [37] N. T. Son and T. Stykel, *Electron. J. Linear Algebra* **38**, 607 (2022).
 - [38] X. Liang, L. Wang, L.-H. Zhang, and R.-C. Li, *Linear Algebra Appl.* **656**, 483 (2023).
 - [39] A. Baernstein II, *Symmetrization in Analysis* (Cambridge University Press, Cambridge, 2019).
 - [40] P. W. Day, *Can. J. Math.* **24**, 930 (1972).
 - [41] M. Wüstner, *J. Lie Theory* **13**, 307 (2003).