On the Proper Treatment of Normalized Statistical Moments as Coordinates in Hamiltonian Phase Space

A methodological note for empirical diagnostics

Rachael L. Bond

Currently looking for a new home

Author Note

Correspondence concerning this article should be emailed to Rachael Bond rachael.lk3ba@rlb.me

On the Proper Treatment of Normalized Statistical Moments as Coordinates in Hamiltonian Phase Space

A methodological note for empirical diagnostics

e consider Hamiltonian diagnostics built on the empirical first four moments of a financial return distribution, using mean, standard deviation, skewness, and kurtosis as phase-space coordinates. A common (and seductive) implementation scales *all* coordinates by recent empirical dispersion to produce a “dimensionless” chart. We show this “double-normalisation” is inappropriate for already-normalised shape descriptors (skewness and kurtosis). It induces spurious curvature in the potential energy of an otherwise linear Hamiltonian, leading to explosive energies in quiet windows. We formalise the pathology, provide a principled construction that treats dimensional and dimensionless coordinates differently, and demonstrate empirically that the corrected model yields stable, interpretable energy partitions that cleanly distinguish trend from trap regimes. This methodological note addresses a gap at the intersection of Hamiltonian moment hierarchies (typically applied to raw moments in plasma physics) and statistical diagnostics (which employ pre-normalized shape descriptors).

Introduction

Hamiltonian ideas have been deployed in statistics and finance both as computational tools (e.g. Hamiltonian Monte Carlo [1, 2]) and as diagnostic lenses for complex dynamics. A practical approach builds a low-dimensional phase space from summary statistics of returns, then interprets the energy partition between kinetic (K) and potential (V) as signalling flow (trend) versus trap (mean-reversion) regimes.

A subtle but consequential pitfall arises when one *re-scales* already normalised coordinates. Skewness and kurtosis are dimensionless by construction—they are standardised central moments. If we further divide their displacements by a small empirical “scale,” the tilde displacement and thus V can blow up without any commensurate change in the underlying distribution. This note formalises the issue and gives a principled fix.

Relationship to existing work. Hamiltonian formulations of moment hierarchies exist in plasma physics [3, 4, 5], where *raw* moments (density, momentum density, energy) serve as canonical coordinates. Those approaches derive closed Hamiltonian systems by expressing higher moments as functions of lower moments, preserving the Poisson bracket structure of the parent Vlasov or kinetic equation. However, those formalisms do not encounter the double-normalisation pathology because all coordinates remain dimensional throughout.

We address a distinct problem: the treatment of *pre-normalized* shape descriptors (skewness, kurtosis) that are already dimensionless by construction. This issue arises whenever one builds diagnostic Hamiltonians from empirical summary statistics rather than from first-principles kinetic theory. While the plasma literature provides the mathematical framework (moment closures, Casimir invariants, symplectic structure), the scaling principle we derive—*only dimensional variables require empirical normalisation*—is novel and necessary for statistical diagnostics.

Contributions. (i) We show that re-scaling skewness and kurtosis by vanishing empirical dispersion inevitably produces divergent potential energy in otherwise linear Hamiltonians; (ii) we propose a consistent chart and scaling rule separating *dimensional* from *dimensionless* coordinates; (iii) we demonstrate empirically that the corrected construction yields stable, interpretable diagnostics across FX pairs, with large V now corresponding to genuine mean-reversion traps rather than scaling artefacts.

Setup: moments as coordinates

Let x_t denote returns over a rolling window. Let $(\mu, \sigma, \gamma, \kappa)$ denote the empirical mean, standard deviation, skewness, and kurtosis (with κ the usual kurtosis, so excess is $\kappa - 3$). We build a phase coordinate $Q = (Q_\mu, Q_\sigma, Q_\gamma, Q_\kappa)$ via the chart

$$Q_\mu = \mu, \qquad Q_\sigma = \log \sigma, \qquad Q_\gamma = \gamma, \qquad Q_\kappa = \operatorname{sign}(\kappa - 3)\, \log\Big(1 + |\kappa - 3|\Big). \tag{1}$$

The chart (1) uses logarithmic coordinates for positivity-constrained quantities ($\sigma > 0$) and a signed-log transform for excess kurtosis to handle both positive and negative excursions while compressing extreme values. This creates a well-behaved

Riemannian structure on the phase space. While the specific choice of chart affects the metric tensor, it is secondary to our main point: *whatever chart is chosen, normalised quantities must not be empirically rescaled.*

Let $Q^\star$ denote an equilibrium (e.g. exponentially weighted baseline). Define tilde displacements $q = \tilde{q} = (Q - Q^\star)/s$, where $s = (s_\mu, s_\sigma, s_\gamma, s_\kappa)$ are positive scales per coordinate used only to make *dimensional* coordinates commensurate.

We adopt a linear, separable Hamiltonian in the tilde chart:

$$K = \frac{1}{2}\sum_i m_i\, \dot{q}_i^2, \qquad V = \frac{1}{2}\sum_i k_i\, q_i^2, \qquad \omega_i^2 = \frac{k_i}{m_i}, \tag{2}$$

with $k_i = m_i\, \omega_i^2$ chosen from a characteristic time-scale and optional gentle hierarchies across coordinates. Masses $m_i > 0$ are estimated from recent motion (quiet coordinates acquire larger m_i).

The double-normalisation pathology

Practitioners often define s_i as the recent RMS of $Q_i - Q_i^\star$ for *all* i. This is harmless for μ and σ (dimensional), but not for γ and κ (already normalised).

Motivating example. Consider a quiet market window where skewness fluctuates in $[-0.15, 0.10]$ around $\gamma^\star = 0.0$. With naive scaling, $s_\gamma = \mathrm{RMS}(\gamma - \gamma^\star) \approx 0.08$. A displacement of $\gamma = 0.10$ yields $q_\gamma = 0.10/0.08 = 1.25$.

Now suppose volatility collapses further and $s_\gamma \to 0.02$ while skewness remains at 0.10. Then $q_\gamma = 0.10/0.02 = 5.0$—a four-fold increase in the normalised coordinate despite no fundamental change in distribution shape. With $k_\gamma \sim 1$, potential energy $V_\gamma \propto q_\gamma^2$ explodes from ~ 0.78 to ~ 12.5. This divergence is purely artifactual.

Proposition 1 (Spurious energy blow-up)**.** *Consider a coordinate Q that is already dimensionless and bounded in practice (e.g. Q_γ or the charted Q_κ). If one defines $s = \mathrm{RMS}(Q - Q^\star)$ on a finite window and constructs $q = (Q - Q^\star)/s$, then as the window variance collapses ($s \to 0$) while $(Q - Q^\star)$ remains bounded, the potential energy contribution diverges:*

$$V_Q = \frac{1}{2}\, k\, q^2 = \frac{1}{2}\, k\, \frac{(Q - Q^\star)^2}{s^2} \xrightarrow[s \to 0]{} \infty. \tag{3}$$

Proof. Immediate, since $k > 0$ is fixed (linear spring) and the numerator is $O(1)$ while $s^2 \to 0$. The divergence is not reflective of the underlying distribution—it is solely induced by the redundant normalisation. $\square$

Remark 1. *The same mechanism contaminates kinetic energy via $\dot{q} = \dot{Q}/s$ if s is time-varying and tiny. In practice, K and V can explode by orders of magnitude in quiet windows despite negligible real changes in shape.*

Remark 2 (Other manifestations)**.** *Similar issues arise in any context where one mixes dimensional and dimensionless variables in a phase-space formulation: (i) using correlation coefficients alongside raw covariances in portfolio optimization; (ii) combining returns with return-on-equity ratios in factor models; (iii) mixing angles and angular momenta in rigid body dynamics without proper chart selection. The common thread is attempting to impose empirical scales on quantities that already possess intrinsic normalisation.*

A principled construction

The resolution is to separate *dimensional* and *dimensionless* coordinates.

Chart and scales

Use the chart (1). Choose scales

$$s_\mu = \max\Big(\mathrm{RMS}(Q_\mu - Q_\mu^\star),\, \lambda_\sigma\, \sigma^\star,\, \varepsilon\Big), \quad s_\sigma = \max\Big(\mathrm{RMS}(Q_\sigma - Q_\sigma^\star),\, \varepsilon\Big), \quad s_\gamma = 1, \quad s_\kappa = 1, \tag{4}$$

with small $\varepsilon > 0$ and a volatility-anchored floor $\lambda_\sigma \in [0.1, 0.5]$. Then

$$q_\mu = \frac{Q_\mu - Q_\mu^\star}{s_\mu}, \quad q_\sigma = \frac{Q_\sigma - Q_\sigma^\star}{s_\sigma}, \quad q_\gamma = Q_\gamma - Q_\gamma^\star, \quad q_\kappa = Q_\kappa - Q_\kappa^\star. \tag{5}$$

This keeps q_γ and q_κ in intrinsic units; no double-normalisation occurs.

Springs and masses

Pick a characteristic time-scale τ (in days), set a base frequency $\omega_0 = 2\pi/(c\,\tau)$ for some $c \gtrsim 6$, clamp ω_0 into a reasonable band, and define $k_i = m_i\, \omega_i^2$ with optional per-coordinate multipliers $\omega_i = \omega_0 h_i$ (e.g. $h = (1,\, 0.8,\, 0.6,\, 0.45)$). Estimate m_i from recent motion in the same chart; clamp m_i into a modest range to avoid pathologies in ultra-quiet regimes.

Dimensionless momenta

Compute finite-difference velocities in chart space and normalise only the dimensional coordinates:

$$\dot{q}_\mu = \frac{dQ_\mu/dt}{s_\mu}, \quad \dot{q}_\sigma = \frac{dQ_\sigma/dt}{s_\sigma}, \tag{6}$$

$$\dot{q}_\gamma = \frac{dQ_\gamma}{dt}, \quad \dot{q}_\kappa = \frac{dQ_\kappa}{dt}. \tag{7}$$

Time-decay these to current momenta with a horizon τ, and form K, V, H in the tilde chart. This keeps the Hamiltonian linear and interpretable.

Discussion

The point is philosophical as much as technical: a phase-space chart must respect the *intrinsic units* of its coordinates. Standardised moments already encode their own scale; forcing an additional data-driven scale manufactures curvature where none exists. In linear Hamiltonians this appears as spurious energy. The fix restores both interpretability (virial-like diagnostics) and numerical stability (bounded energy oscillations under symplectic integration).

Extensive versus intensive quantities. This distinction parallels thermodynamics: extensive quantities (volume, energy) scale with system size and require context-dependent units, while intensive quantities (temperature, pressure) have intrinsic meaning independent of scale. Skewness and kurtosis are intensive-like: they describe distribution *shape*, not magnitude. Mean and variance are extensive-like: they describe location and scale, which lack intrinsic units. Treating shape descriptors as extensive variables creates the pathology we observe.

Just as one would not “normalise” temperature by its recent RMS to make it commensurate with volume, one should not normalise skewness by its empirical dispersion to make it commensurate with mean.

Gauge-theoretic interpretation. In gauge-theoretic language, our fix amounts to choosing a coordinate chart where some coordinates carry natural units (skewness in its own dimensionless scale) rather than forcing all coordinates into an artificial common gauge. This preserves the geometric structure of the phase space.

The Hamiltonian structure imposes constraints: the symplectic form $\omega = \sum_i dq_i \wedge dp_i$ must be preserved under coordinate transformations. When scales s_i are functions of state (as occurs with empirical RMS on short windows), the transformation is no longer canonical unless carefully constructed. Our prescription—using constant scales for normalised coordinates—avoids this subtlety.

Relation to information geometry. The issue also connects to information geometry, where the Fisher metric defines distances on the manifold of probability distributions. Skewness and kurtosis are coordinates on this manifold that already incorporate the "natural" metric structure. Re-scaling them by empirical dispersion distorts this geometry, analogous to using a non-Riemannian connection in a space that possesses intrinsic curvature.

The idea sits naturally alongside Hamiltonian moment closures (where moments are canonical variables) and information-geometric HMC (where metrics live on parameter manifolds), but it is distinct: we focus on *empirical moments as state coordinates* for diagnostics, and on the scaling principle that prevents double-normalisation.

Conclusion

When using empirical moments as phase coordinates, only dimensional variables require empirical scaling. Treat shape descriptors (skewness, kurtosis) with intrinsic scales in an appropriate chart. This small change prevents artificial energy blow-ups and yields robust, interpretable diagnostics of market regime in a Hamiltonian framework.

The broader lesson extends beyond finance: *in any Hamiltonian diagnostic built from mixed dimensional and dimensionless observables, empirical normalisation must respect the intrinsic structure of each coordinate.* Ignoring this principle creates spurious dynamics that obscure genuine physical behavior.

Practical recipe.

1. Use the chart (1); compute $Q^\star$ as an equilibrium baseline.

2. Set s_μ, s_σ from robust dispersion with sensible floors; set $s_\gamma = s_\kappa = 1$.

3. Compute $\dot{Q}$ in chart space; form $\dot{q}$ by dividing only where needed.

4. Choose $k_i = m_i\omega_i^2$ in the tilde chart; avoid state-dependent k_i for linear diagnostics.

5. Report raw K, V, H and normalised diagnostics $K_{\text{norm}}, V_{\text{norm}}, \eta = K/(K+V)$.

6. **Diagnose the pathology:** Compute $\max_{t,i} |V_i(t)|/\bar{V}$ where $\bar{V}$ is the median potential energy. Values > 100 indicate likely scaling artifacts. If detected, check whether s_γ or s_κ have become anomalously small (< 0.05).

Acknowledgements

Thanks to collaborators and colleagues for discussions. This work was developed during the development of a mathematical-physics model of human decision-making. All errors remain my own.

References

[1] M. Girolami and B. Calderhead, *Riemann manifold Langevin and Hamiltonian Monte Carlo methods*, J. R. Stat. Soc. B **73**, 123–214 (2011).

[2] M. Betancourt, *A Conceptual Introduction to Hamiltonian Monte Carlo*, arXiv:1701.02434 (2017).

[3] P. J. Morrison, *Hamiltonian description of the ideal fluid*, Rev. Mod. Phys. **70**, 467–521 (1998).

[4] M. Perin and P. J. Morrison, *Hamiltonian closures for fluid models with four moments by dimensional analysis*, J. Phys. A: Math. Theor. **48**, 275501 (2015).

[5] M. Perin, C. Chandre, P. J. Morrison, and E. Tassi, *Hamiltonian fluid closures of the Vlasov-Ampère equations: From water-bags to N moment models*, Phys. Plasmas **22**, 092309 (2015).

[6] C. Chandre and M. Perin, *Hamiltonian reductions of the one-dimensional Vlasov equation using phase-space moments*, J. Math. Phys. **57**, 032701 (2016).

[7] M. Kendall and A. Stuart, *The Advanced Theory of Statistics, Vol. 1: Distribution Theory*, 4th ed. (Griffin, London, 1977).

[8] J.-P. Bouchaud and M. Potters, *Theory of Financial Risk and Derivative Pricing*, 2nd ed. (Cambridge University Press, Cambridge, 2003).

[9] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, (Chapman & Hall/CRC, Boca Raton, 2004).