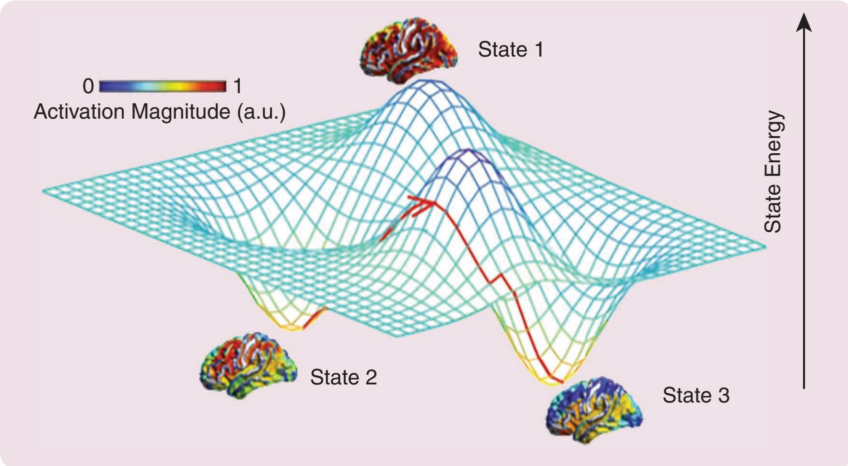
The Potential of a Potential

Behaviors that derive from a potential have huge potential. Their potential has the graphical interpretation of an energy landscape. The gradient behavior is determined by a local update rule that provides the best possible descent for the potential. The potential, a scalar-valued function or functional, provides a complete qualitative characterization of the behavior: an infinite set of trajectories.

The concept of gradient is stunningly simple yet stunningly rich. With the help of a little drawing, most first graders will grasp the concept. Yet, looking at the history of science and engineering, it is hard to find a concept with a stronger explanatory power when it comes to model dynamical behaviors. The gradient concept connects first graders and scientific giants.

A gravity potential led Newton to formulate the general laws of mechanical behaviors. Smale, Brayton-Moser, and Minty modeled electrical circuits and networks as gradient systems. Hopfield modeled neural networks as gradient systems, opening up an entirely novel way to connect artificial neural networks to the computational neuroscience of learning and memory. Entropy potentials rule thermodynamics and information theory. The design of cost functions rules information engineering via the design of gradient algorithms that minimize them. A free energy potential determines Friston's theory of brain and intelligent behavior [1]. Gradient flows on probability measures have renewed our understanding of partial differential equations [9]. The list of successes is much longer, and you will add your own example to it.



A schematic to provide an intuition regarding the nature of an energy landscape for the more general case of continuously-valued brain states. (From [8].)

Gradient systems bridge the discrete and the continuous. Most physical laws are infinitesimal, but algorithmic laws are always incremental. The former must be discretized to be simulated. The latter must be made infinitesimal to be analyzed. When a differential equation is a gradient, it also defines an algorithm. Otherwise, it lacks interpretation. When an algorithm is a gradient, it approximates a differential equation. Otherwise, it lacks interpretation. Potential theory provides key insight in discrete optimization problems that otherwise remain elusive [2]. Without the bridge of a potential, the distance between discrete and continuous behaviors sometimes looks abyssal.

Gradient systems bridge the deterministic and the stochastic. Stochastic gradient systems turn intractable deterministic algorithms into tractable implementations. Mean-field algorithms turn intractable stochastic algorithms into deterministic differential equations that can be analyzed. One of the most brilliant lectures I heard in my life was when Yuri Nesterov contrasted the "Homo economicus" (HE) and "statistical Homo economicus" (SHE) models of human behavior [3]. Without the bridge of a potential, the distance between HE and SHE behaviors sometimes looks abyssal.

But what is a gradient behavior? In particular, what is an *open* gradient system? That is, how do we formalize the inescapable concept of a potential when modeling a behavior that interacts with its environment? As is often in system theory, concepts that seem evident when conceived for closed dynamical systems become surprisingly shaky when generalized to open systems. As is often in the history of our field, mathematics invites us to choose between two distinct routes.

Digital Object Identifier 10.1109/MCS.2023.3329911
Date of current version: 10 January 2024

The question of how to bridge the two is left to control science.

The route grounded in the tradition of mechanics uses differential geometry. Differential geometry provides us with a concept of gradient vector field, which in turn defines the concept of (closed) gradient dynamical system. How to extend this approach to systems with inputs and outputs was originally asked by Brockett [4]. It pretty much defined the research program of Port–Hamiltonian systems and control. The specific question of when a linear time-invariant system can be regarded as a gradient system was solved by Arjan van der Schaft [5]. It did not make the news (yet). The nonlinear extension of that result is still a work in progress.

The route grounded in the tradition of electricity uses operator theory. Minty defined the concept of a monotone operator to characterize nonlinear resistors with positive conductance. Rockafellar then showed that a maximal monotone operator derives from a potential if and only if it is cyclically monotone [6]. An input–output operator can thus be said to derive from a potential if it can be expressed as a difference of cyclically monotone operators. In recent work, we expressed the solutions of mixed feedback systems (such as Van der Pol oscillators) as zeros of difference of maximally monotone operators [7]. It is still an open question whether such behaviors can also be regarded as critical points of a difference of convex potentials. That would make them "open gradient behaviors."

Neural networks are intractable if they are not gradient neural networks. Learning rules are intractable when they are not gradient rules. I am sometimes led to think that we will not make nonlinear control algorithmic until we make nonlinear behaviors deriving from a potential.

Please share your feedback at r.sepulchre@eng.cam.ac.uk.

Rodolphe Sepulchre

REFERENCES

[1] K. Friston, "The free-energy principle: A unified brain theory?" *Nature Rev. Neurosci.*, vol. 11, no. 2, pp. 127–138, 2010, doi: 10.1038/nrn2787.

[2] S. Borodachov, D. Hardin, and E. Saff, *Discrete Energy on Rectifiable Sets* (Springer Monographs in Mathematics), New York, NY, USA: Springer-Verlag, 2010, pp. 1–666.

[3] Y. Nesterov, "Algorithmic models of human behaviors," Francqui lecture, Univ. of Liege, Liège, Belgium, 2012. [Online]. Available: https://people.montefiore.uliege.be/francqui/slides/Lect5_Econ_Boadilla.pdf

[4] R. W. Brockett, "Control theory and analytical mechanics," in *Geometric Control Theory*, C. Martin and R. Hermann, Eds., Brookline, MA, USA: Math. Sci. Press, 1977, pp. 1–48.

[5] A. Van der Schaft, "On the relation between port-Hamiltonian and gradient systems," *IFAC Proc. Vol.*, vol. 44, no. 1, pp. 3321–3326, 2011, doi: 10.3182/20110828-6-IT-1002.00555.

[6] R. T. Rockafellar, "Characterization of the subdifferentials of convex functions," *Pacific J. Math.*, vol. 17, no. 3, pp. 497–510, 1966, doi: 10.2140/pjm.1966.17.497.

[7] A. Das, T. Chaffey, and R. Sepulchre, "Oscillations in mixed-feedback systems," *Syst. Control Lett.*, vol. 166, Aug. 2022, Art. no. 105289, doi: 10.1016/j.sysconle.2022.105289.

[8] S. Gu et al., "The energy landscape of neurophysiological activity implicit in brain network structure," *Scientific Rep.*, vol. 8, Feb. 2018, Art. no. 2507, doi: 10.1038/s41598-018-20123-8.

[9] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient Flows: In Metric Spaces and in the Space of Probability Measures* (Lectures in Mathematics). Basel, Switzerland: Birkhäuser, 2005.

Exploring ChatGPT

"I don't quite agree with it, but "a calculator for words" is an interesting framing for ChatGPT"

—Sam Altman

Unveiling Cybernetics

Besides electrical engineering theory of the transmission of messages, there is a larger field [cybernetics] which includes not only the study of language but the study of messages as a means of controlling machinery and society, the development of computing machines and other such automata, certain reflections upon psychology and the nervous system, and a tentative new theory of scientific method.

—Norbert Wiener