PROJECT DESCRIPTION

INFORMATION GEOMETRY AND MARKOV PROCESSES

JOHN C. BAEZ

1. Introduction

Information geometry [1] applies ideas from differential geometry to probability theory. These ideas can be used to study Markov processes that describe randomly interacting particles. In applications to chemistry, these particles are typically molecules of different kinds [7]. In applications to biology, they could be organisms with different genotypes [6].

In chemistry, these Markov processes obey a version of the Second Law of Thermodynamics, which says that entropy always increases. In biology, the analogous statement is Fisher's Fundamental Theorem of Natural Selection, which says that the mean fitness of a population increases through natural selection. This sets up an fascinating connection between thermodynamics and evolutionary biology. Indeed, already in 1930 Fisher [8] wrote:

> ... the fundamental theorem [of natural selection] bears some remarkable resemblance to the second law of thermodynamics. It is possible that both may ultimately be absorbed by some more general principle.

However, the details are just beginning to be worked out in a mathematically precise way [16, 17]. The vistas opened up by this synthesis are enormous, and worth exploring more deeply.

During a recent 2-year visit to the Centre for Quantum Technologies, the PI began to investigate this subject and explain it in a series of online articles [5]. In the project funded by this grant, the PI will carry this work forward with the help of two graduate students, and publicize the results to a large audience. More specifically, the project will:

1. Study increasing quantities for random many-body interactions. As explained in the section 'Details', there is a disparate collection of results showing that certain quantities increase or decrease with time under Markov processes describing random interactions of particles. Some of these results arose within probability theory, but others first appeared in chemistry or evolutionary biology. The PI and his students will prove theorems that extend and unify the results known so far. They will seek clear statements of these theorems, using the language of information geometry whenever appropriate. The interplay of Riemannian geometry (the Fisher information metric) and convex geometry (the majorization order) on the space of probability distributions will be especially important.

3. Train students working on information theory, Markov processes and applications to biology. This grant will mainly go to provide salary support for two Graduate Student Researchers for 6 academic months per year at 49% for three years. It will also fund these students to attend conferences or visit Marc Harper, an expert on these subjects. These students will help with the projects described above, and go on to do their theses on these subjects. The PI currently has five graduate students who want to begin work with him. Of these, one has already published on stochastic processes (Michael Knap). The PI will use this grant to fund the two students who are best suited to the project.

4. Publicize the results obtained. The PI will publicize the results of this research in papers, lectures, and his blog. His blog, *Azimuth*, is widely read and has attracted an audience of scientists interested in information geometry and evolutionary biology [2]. He is also a widely sought-after speaker, and has been invited to speak on applications of mathematics to biology and ecology at Google, U. C. Berkeley, Sheffield University in England, and the 55th annual conference of the South African Mathematical Society [3]. In the summer of 2012 he was invited to speak on the topic of this proposal at the Exploratory Conference on the Mathematics of Biodiversity at the Centre de Recerca Matemàtica in Barcelona [4]. Thus, he is in a good position to give lectures on the results obtained thanks to this grant.

2. Intellectual Merit and Broader Impacts

The **intellectual merit of the proposed activity** is that it will use ideas from probability theory to better understand evolution in biology. In particular, the relation between 'entropy increase' and 'increase in mean fitness through natural selection' will be clarified.

The **broader impacts resulting from the proposed activity** include developing new bridges between probability theory and evolutionary biology. Through talks and his well-known blog, the PI will publicize the results of this research, explaining the probability theory to biologists and the biology to probability theorists. His two graduate students, after doing theses on this subject, will be in a good position to develop further links between these communities.

3. Details

There is a class of Markov processes that describe collections of abstract particles that randomly interact and turn into particles of different kinds. In chemistry, these have been studied since the 1970's using 'chemical reaction networks' [7, 15]. Any such network gives a continuous-time Markov process where probabilities evolve according to a 'master equation'. However, in the limit of large numbers, the expected number of particles of each kind obeys an equation of its own, called the 'rate equation'.

The proportion of particles of each kind can be viewed as a probability, since these proportions sum to one. Thus, both the master equation and rate equation can be seen as describing the time evolution of a probability distribution. The techniques of information geometry thus apply to both. But the two equations have very different meanings. The master equation describes the time evolution of the probability for having any number of particles of each type. The rate equation describes the time evolution of the *expected* proportion of particles of each type.

The PI and his students will apply techniques of information geometry to both these equations, and synthesize what is known about both. This project is important because different ideas and tools have been applied to the two equations, perhaps due to accidents of history:

The rate equation. In their foundational work on the rate equation for chemical reaction networks, Horn and Jackson [15] showed that for a certain class—the 'weakly reversible, deficiency zero' reaction networks—a quantity related to entropy always increases. This quantity is essentially the Shannon entropy of one probability *relative to another*. In information geometry, the negative of relative Shannon entropy is often called the Kullback–Leibler divergence, so we may also say that the Kullback–Leibler divergence decreases.

The significance of the Kullback–Leibler divergence for chemistry has been nicely elucided by Hangos [11], but only when a further assumption holds: detailed balance. This assumption is *not* typically true in applications to biology [9], so the PI and his students will study how Hangos' argument generalizes to other situations.

On the other hand, in the context of evolutionary biology, Harper has shown that the Kullback–Leibler divergence decreases for another class of rate equations [12]. This work is especially interesting because he makes the connection to information geometry very explicit. He shows that time evolution given by these rate equations can been seen as a gradient flow with respect to a certain Riemannian metric on the space of populations distributions. This Riemannian metric is none other than the Fisher information metric, a fundamental structure in information geometry. He also constructs other functions that always decrease with time for this class of rate equations. These are relative versions of various non-Shannon entropies, such as Rényi entropies.

The master equation. The rate equation is a nonlinear equation derived from the master equation in the limit of large numbers. On the other hand, the master equation directly describes a Markov process, so a large class of 'entropy increase' results immediately apply, at least when this Markov process is reversible. For example, van Erven and Harremoës [18] have shown that not only the relative Shannon entropy but also the relative Rényi entropy increases under any reversible Markov process.

In fact, Gorban, Gorban and Judge have classified *all* entropy functions that increase under reversible Markov processes and obey some subsidiary conditions [10]. As they emphasize, the deepest approach to this subject is not simply to define more and more kinds of entropy, but to work with a partial order on the space of probability distributions: the 'majorization order'. The idea is that time evolution under a reversible Markov process always sends any probability distribution to one that it 'majorizes'. In a sense, this the most general statement of the Second Law of Thermodynamics for Markov processes. The majorization order is another fundamental structure in information geometry, since it arises from fact that the space of probability distributions is convex.

Synthesis. The PI and his students will investigate how the quantities shown to increase under the rate equation are related to those known to increase under the master equation. They will prove theorems that extend and unify the rather disparate results known so far, both for the rate and the master equation. They will seek clear statements of these theorems, using the language of information geometry whenever appropriate. The interplay of Riemannian geometry (the Fisher information metric) and convex geometry (the majorization order) will be especially important. Moreover, they will study the extent to which results for the rate equation can be derived as 'large-number limits' of those for the master equation.

4. Prior NSF Support

The PI has won two prior NSF grants. The most recent and relevant is award number 0653646, entitled 'Feynman Diagrams and the Semantics of Quantum Computation', a grant for $149,938.00 awarded in July 2007 and ending July 2012.

Summary of results, including broader impacts. In 'Physics, Topology, Logic and Computation: A Rosetta Stone', the PI and his graduate student Mike Stay, supported by this grant, worked out and carefully explained how Feynman diagrams, string diagrams in topology, proofs in logic, and processes of computation could all be dealt with in a unified way using symmetric monoidal categories with duals. Stay, who now works at Google, is now becoming an expert on categorical semantics and its applications to computer science. He is running a category theory mailing list at Google and is completing his Ph.D. thesis, which describes a theory of compact closed bicategories with duals in which computations are 2-morphisms.

After Mike Stay took a job at Google, most of the grant money went to supporting another graduate student of the PI, Alexander Hoffnung. Together with Hoffnung and another graduate student, Christopher Walker, the PI found that spans of groupoids are able to do much of what we normally do with linear operators in quantum theory. This was a rather unexpected turn. It turned out one can use this to 'groupoidify' a large portion of the mathematics of quantum theory, shedding light on its combinatorial underpinnings. Alexander Hoffnung has gone on to postdoctoral positions first at the University of Ottawa and now Temple University, and is carrying on this line of work. Christopher Walker has a tenure-track position at Odessa College.

In further work with his graduate students Christopher Rogers and John Huerta, the PI also studied the algebra of grand unified theories and applications of higher category theory to string theory. Thiese students have completed Ph.D.'s on closely related topics, and Christopher Rogers now has a postdoctoral position at the University of Göttingen, while John Huerta obtained postdocs first at Australian National University and now the Instituto Superior Técnico in Lisbon. Both are actively publishing more work along similar lines.

The PI gave several talks on the subject of the 'Rosetta Stone' paper. For example, he gave a one-hour plenary talk about it in 'Algebraic Topological Methods in Computer Science 2008' at University Paris 7 on July 7, 2008. The PI also gave a one-hour plenary talk at the '24th Annual IEEE Symposium on Logic in Computer Science' (LICS 2009) on August 13, 2009, and a colloquium talk at California State University, Fresno on April 9, 2010.

The PI also gave have also given talks on groupoidification. He spoke on this in October 2007 as the keynote speaker at 'Deep Beauty: Mathematical Innovation and the Search for an Underlying Intelligibility of the Quantum World', a workshop in honor of John von Neumann at Princeton University.The PI also spoke about it at the 'Groupoids in Analysis and Geometry' seminar in Berkeley on Tuesday May 20, 2008, at the conference 'Homotopy Theory and Higher Categories' at the Centre de Recerca Matemàtica (CRM) in Barcelona on June 30, and at the 2009 Joint Mathematics Meetings, Washington, D.C. in January 2009. The PI's students have also given many talks on the subjects of this research project.

Publications. The publications arising from grant number 0653646 were:

(1) J. Baez and A. Lauda, A prehistory of n-categorical physics, in *Deep Beauty: Mathematical Innovation and the Search for an Underlying Intelligibility of the Quantum World*, ed. Hans Halvorson, Cambridge U. Press, Cambridge, pp. 13–128.
(2) J. Baez, A. Hoffnung and C. Rogers, Categorified symplectic geometry and the classical string, *Comm. Math. Phys.* **293** (2010), 701–715.
(3) J. Baez and C. Rogers, Categorified symplectic geometry and the string Lie 2-algebra, *Homotopy, Homology and Applications* **12** (2010), 221–236.
(4) J. Baez, A. Hoffnung and C. Walker, Higher-dimensional algebra VII: groupoidification, *Th. Appl. Cat.* **24** (2010), 489–553.
(5) J. Baez and J. Huerta, The algebra of grand unified theories, *Bull. Amer. Math. Soc.* **47** (2010), 483–552.
(6) J. Baez and M. Stay, Physics, topology, logic and computation: a Rosetta Stone, in *New Structures for Physics*, ed. Bob Coecke, Lecture Notes in Physics vol. 813, Springer, Berlin, 2011, pp. 95–174.

References

[1] S. Amari and H. Nagaoka, *Methods of Information Geometry*, American Mathematical Society, Providence, Rhode Island, 2000

[2] J. Baez, *Azimuth* blog. Available at ⟨http://johncarlosbaez.wordpress.com/⟩.

[3] J. Baez, Lectures. Available at ⟨http://math.ucr.edu/home/baez/lectures.html⟩.

[4] J. Baez, Diversity, entropy and thermodynamics, lecture at the Exploratory Conference on the Mathematics of Biodiversity, Centre de Recerca Matemàtica, Barcelona, 2012. Available at ⟨http://math.ucr.edu/home/baez/biodiversity/⟩.

[5] J. Baez, *Information Geometry*. Available at ⟨http://math.ucr.edu/home/baez/information/⟩.

[6] P. Dodd and N. Ferguson, A many-body field theory approach to stochastic models in population biology, *PLoS ONE* **4** (2009). Also available at ⟨http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0006855⟩.

[7] M. Feinberg, *Lectures On Reaction Networks*, 1979. Available at ⟨http://www.che.eng.ohio-state.edu/∼FEINBERG/LecturesOnReactionNetworks/⟩.

[8] R. Fisher, *The Genetical Theory of Natural Selection*, Clarendon Press, Oxford, 1930.

[9] J. Grilli, S. Azaele, J. R. Banavar and A. Maritan, Absence of detailed balance in ecology. Available at ⟨http://arxiv.org/abs/1210.5819⟩.

[10] A. N. Gorban, P. A. Gorban and G. Judge, Entropy: the Markov ordering approach, *Entropy* **12** (2010), 1145–1193. Also available at ⟨http://www.mdpi.com/1099-4300/12/5/1145⟩.

[11] K. M. Hangos, Engineering model reduction and entropy-based Lyapunov functions in chemical reaction kinetics, *Entropy* **12** (2010), 772–797. Also available at ⟨http://www.mdpi.com/1099-4300/12/4/772⟩.

[12] M. Harper, Information geometry and evolutionary game theory. Available at ⟨http://arxiv.org/abs/0911.1383⟩.

[13] M. Harper, The replicator equation as an inference dynamic. Available at ⟨http://arxiv.org/abs/0911.1763⟩.

[14] M. Harper, Escort evolutionary game theory, *Physica D* **240** (2011), 1411-1415. Also available at ⟨http://arxiv.org/abs/0911.1764⟩.

[15] F. Horn and R. Jackson, General mass action kinetics, *Arch. Rational Mech. Analys.* **47** (1972), 81–116.

[16] V. Mustonen and M. Lässig, Fitness flux and ubiquity of adaptive evolution, *Proc. Nat. Acad. Sci. USA* **107** (2010), 4248–4253. Also available at ⟨http://www.pnas.org/content/107/9/4248⟩.

[17] G. Sella and A. E. Hirsh, The application of statistical physics to evolutionary biology, *Proc. Nat. Acad. Sci. USA* **102** (2005), 9541–9546. Also available at ⟨http://www.pnas.org/content/102/27/9541⟩.

[18] T. van Erven and P. Harremoës, Rényi divergence and majorization. Available at ⟨http://arxiv.org/abs/1001.4448⟩.