Parrondo's Ratchet I

Analytic Theory of Some Random Walks on a Cyclic Graph

Nicholas Wheeler
January 2014

Background. The origins of the Atomic Hypothesis—the notion that the material world can be understood as a manifestation of the interactions among populations of invisibly tiny "atoms," the nature and properties of which could only be conjectured—has been traced to the Greek philosopher Leucippus (early 5^{th} Century BCE), but may possibly have been entertained as early as the 13^{th} or 12^{th} Century BCE by Mochus of Sidon. Leucippus' idea was elaborated in the better-known work of his student Democritus (c460–c370 BCE). Atomism was embraced by Epicurus (341–270 BCE), whose views were promulgated in *De Rerum Natura* ("On the Nature of Things"), the influential long philosophical poem by Lucretius (c99–c55 BCE). But well into the 19^{th} Century the atomic hypothesis—which had occupied a place in the world of ideas for more than twenty (and perhaps more than thirty) centuries—remained very much that, a mere hypothesis, unsupported by a shred of direct physical evidence.[1]

"Are atoms real?" was for 19^{th} Century chemists and physicists an elusive question of ever greater relevance to the interpretation of developments. In 1803 John Dalton (1766–1844) had endowed elemental "atoms" with a set of five defining properties that provided the foundation for subsequent chemical research, and by 1811 "Avogadro's Law" had been propounded, but it was recognized (and—by such leading figures as Sir Humphry Davy—insistently emphasized) that the success those conceptual devices did not address the ontological issue,[2] did not compel abandonment of the prevailing "it's as if" attitude.

That 19^{th} Century physicists were, by and large (and especially in Britian), quicker than chemists to embrace the idea that atoms/molecules are indeed real

---

[1] The success of Newtonian mechanics provided no such evidence, for the "point particles" contemplated by Newton were idealized abstractions. It served Newton's purposes to think of (say) Mars as a "particle," though it was not at all the kind of object that the atomists had in mind.

[2] For an excellent survey of the complex evolution of thought in this area, see Chapter 5 ("The reality of molecules") of A. Pais, *"Subtle is the Lord...: The Science and the Life of Albert Einstein* (1982).

was due mainly to the development and striking success of the kinetic theory of gases, for which ground had been prepared by the invention of thermodynamics.

In 1824, Sadi Carnot (1796–1832) published his *Reflections on the Motive Power of Fire* which, however, attracted little attention until its importance was finally recognized by Émile Clapeyron (1834), Rudolf Clausius (1850) and William Thomson (Lord Kelvin, 1851), who collectively managed to construct precursors of the entropy concept (though its name and recognizably modern formulation were introduced only in 1865, by Clausius). In 1845 James Joule established the mechanical equivalence of heat, and in 1847 Hermann Helmholtz asserted that thermal energy is a form of mechanical energy and that energy (thermal + mechanical) is conserved. And during the years 1850–1855 Clausius, Helmholtz and Kelvin produced formulations of 1^{st} and 2^{nd}Laws, so that by about 1855 the essential outlines of a classical thermodynamics had been drawn.

These developments motivated Clausius to revisit and to undertake to extend an idea that had been put forward more than a century earlier: Daniel Bernoulli (1700–1782) had in 1738 developed a rudimentary "kinetic theory of gases" on the basis of an assumption that physical gases can be modeled as clouds of flying Newtonian point masses, and Roger Boscovich (1711–1787) published closely related ideas in 1745. But this work seems to have attracted little attention. More elaborate theories—very much in the Bernoulli tradition —were developed in papers by John Herapath (1820) and John Waterston (1845) which, however, the Royal Society declined to publish on grounds that they were "too mathematical," else by authors too obscure, else based upon a hypothesis too implausible to be taken seriously. Clausius' "On the kind of motion we call heat" appeared in 1857, a second paper followed a year later, and it was the publication of Clausius' papers that led the 29-year-old Maxwell to take up that subject,[3] which he developed in three monumental papers (1860, 1867 and 1879, the year of his death). In the first of those papers, after acknowledging his indebtedness to his predecessors and drawing attention to limitations present in their work, Maxwell begins "If we adopt a statistical theory. . . " and within four pages produces the Maxwell distribution.

Reliance upon statistical modes of argument was certainly the most distinctive novel feature of Maxwell's richly detailed kinetic theory of gases. Ludwig Boltzman (1844–1906), whose dissertation (1866) derived from Maxwell's first paper, recognized—as had Maxwell himself—that the swift elegance of Maxwell's derivation of the velocity distribution formula hinged on an assumption (statistical independence of the velocity components) that, while mathematically natural, was physically somewhat dubious. His effort to address this problem led over the years to the invention of the Boltzmann equation

[3] Simultaneously with his development of electrodynamics and work also on many (!) other topics; see W. D. Niven (editor), *The scientific papers of James Clerk Maxwell* (1890), which is available as a Dover reprint. Maxwell's "A dynamical theory of the electromagnetic field" appeared in 1864, and his *A treatise on Electricity & Magnetism* in 1873.

(1872), the H-theorem and by 1884 to the statistical mechanics of thermally equilibrated systems, a subject brought to a state of perfection by Gibbs (1839–1903), whose *Elementary Principles of Statistical Mechanics, developed with especial reference to the Rational Foundation of Thermodynamics* was published in 1902.

It is of interest that Boltzmann was obliged throughout his career to attempt to defend his belief that atoms/molecules are real (or at least that it is legitimate to proceed "as if" they were real); influential German physicists at that time—under the influence principally of Ernst Mach (1838–1919) who, in 1897, after a lecture by Boltzmann, declared "I don't believe that atoms exist"—held that in the absence of direct experimental evidence that such entities actually exist it was improper to allude to them in scientific discourse. So pervasive was that "logical positivist" view that the editors of scientific journals declined to publish Boltzmann papers in which the language appeared to impute physical reality to atoms.

In Britain, Maxwell, though subject to no such constraint, did—beginning in 1866—publish occasional essays bearing titles like "On the dynamical evidence of the molecular constitution of bodies," but in all of his writing made it abundantly clear that, while the evidence for the existence of atoms/molecules —microscopic bodies subject to the laws of mechanics—remained circumstantial, and their detailed properties remained to be discovered, their reality was in his view assured.

Maxwell's demon & Brownian ratchets. Maxwell had come by 1866 to the realization that the 2^{nd} Law of thermodynamics is in essence a *statistical* proposition that derives its seemingly inviolable inevitability not so much from physics (molecular dynamics) as from the Law of Large Numbers—an insight reaffirmed by Boltzmann in 1877.[4] Thus was Maxwell—given his conviction that atoms are real objects—led to describe in a letter to his friend Peter Guthrie Tait (1867) how the 2^{nd} Law could be contravened if a "very observant and neat-fingered being" who "knows the paths and the velocities of all the molecules" and who guards a gate in the partition that separates the two halves of an enclosed initially equilibrated gas were to manipulate the gate in such a way as to allow warm atoms to pass one way, cool atoms to pass the other way. Such a being, using only its intelligence and without the expenditure of any effort, could create a temperature differential that could be used to do work. Maxwell broached the same idea in a letter (1871) to John Strutt (Lord Rayleigh) and made it the subject of a concluding section of his *Theory of Heat* (1871). The first allusion to "Maxwell's Demon" appears in a paper published by Kelvin in 1874. Maxwell's Demon has suffered many instructive deaths and enjoyed many still more instructive reincarnations since brought into the world

[4] Boltzmann's H-theorem (1872) was the fruit of an effort to derive the law of entropy increase from kinetic theory. But Boltzmann's friend Josef Loschmidt (1821–1895) pointed out (as Kelvin was the first to report in an 1874 paper) that reversible mechanics cannot by itself give rise to irreversible phenomena.

by Maxwell, nearly 150 years ago, and is today more productively vigorous than ever.[5] But Maxwell's "neat-fingered being" was for Maxwell himself simply an idea, the principal actor in a thought-experiment which, while predicated on the presumption that atoms are real, provided no direct evidence in support of that presumption.[6]

The "molecular reality question" was settled when (which did not happen immediately) physicists gained an appreciation of the rich theoretical and exprimental implications of Einstein's Brownian motion paper (1905). Einstein once remarked that "Unacquainted with the investigations of Boltzmann and Gibbs which had appeared earlier and which in fact dealt exhaustively with the subject, I [had, during the years 1902–1904, independently rediscovered all the elements of] statistical mechanics and the molecular-kinetic theory of thermodynamics based upon it. My main purpose in doing this was to find facts which would attest to the existence of atoms of definite size."

Marian Smoluchowski (1872–1917) was a Polish physicist who worked in the tradition of Boltzmann and whose theory (1906) of Brownian motion was developed simultaneously and independently of Einstein's, to which it conformed in all essential particulars. In 1912 he described a hypothetical device which—insofar as it exploited atomicity to achieve seeming violation of the second law of thermodynamics—can be viewed as a mindless variant of Maxwell's Demon: it involved a microscopic paddle wheel that experienced Brownian motion when bombarded by gas molecules. A ratchet and pawl mechanism (analogs of the Demon's "neat fingers") allowed the paddle to rotate in one direction, but prevented rotation in the opposite direction. The uni-directional paddle could, in principle, be arranged to lift a weight, resulting in a device that exploits fluctuations in the random motion of molecules to do work. Since work would be produced without the transfer of heat from one reservoir to another at lower temperature, the device would operate in violation of the 2^{nd} Law. But as Smoluchowski himself pointed out, the pawl is necessarily dissipative, necessarily heats up a bit each time it snaps shut, so that ultimately it experiences its own Brownian motion, and is rendered ineffective.

---

[5] For the classic contributions to this subject (through 2002), with elaborate and very helpful commentary, see Harvey S. Leff & Andrew F. Rex (editors), *Maxwell's Demon 2: Entropy, Classical & Quantum Information, Computing* (2003). Compliance with the 2^{nd} Law was re-established when Leo Szilard ("On the reduction of entropy in a thermodynamic system by the invervention of intelligent beings," Z. fŭr Physik **53**, 840-856 (1929)) assigned entropy-like measure to the information that the Demon requires to do its work, and by Rolf Landauer's argument (1961) to the effect that all information is in fact physical ("Landauer's Principle").

[6] Efforts are presently under way to construct operational physical embodiments of Maxwell's Demon, devices that convert information into energy. See (for example) C. Jarzynski *et al*, "Maxwell's refrigerator: An exactly solvable model," PR Letters **111**, 030602 (2013) and the many relevant websites supplied by Google.

In a lecture presented in 1962 (reproduced as Chapter 46 in Volume I of *The Feynman Lectures on Physics*) Richard Feynman used Smoluchowski's ratchet (without attribution[7]) to make intuitively plausible why "if everything is at the same temperature, heat cannot be converted to work by means of a cyclic process," as Carnot had asserted. And why, more particularly, the kinetic motion of the molecules in an equilibrated gas cannot be made (except briefly) to do work.

Enter: Parrondo. In 1996, Juan M. R. Parrondo and Pep Español published an elaborately detailed criticism of some aspects of Feynman's argument, particularly as it relates to the maximal efficiency of Brownian motors.[8] That work led promptly (by a train of thought that, so far as I am aware, has never been described) to the discovery of "Parrondo's Paradox," which was first reported in a talk on the "Efficiency of Brownian motors" that Parrondo presented at a Workshop on Complexity and Chaos that took place in Torino, Italy in July, 1996.[9] At that workshop Parrondo met Derek Abbott (Director of the Center for Biomedical Engineering at the University of Adelaide). In luncheon conversation over a glass of wine they discovered their mutual interest in "stochastic ratchets" (which they both recognized to be "ubiquitous in physical and biological systems," and of importance in fields as diverse as economics, biogenesis and game theory), and became thereafter frequent collaborators. It fell, in fact, to Abbott to co-author one of the first published accounts of Parrondo's discovery.[10]

Parrondo had called attention to a *mathematical* fact—a counterintuitive manifestation of the not-at-all-surprising circumstance that

> *Composite Markov processes can possess properties that are not shared by their factors*

—that bears no essential relationship to the physical theory of stochastic ratchets from which historically it sprang, and literature relating to the two subjects has diverged. The paradox literature—of which the Parrondo's Paradox

[7] Thus did "Smoluchowski's ratchet" come to be known popularly as "Feynman's ratchet." Such devices are now known most commonly as Brownian or stochastic ratchets.

[8] "Criticism of Feynman's analysis of the ratchet as an engine," AJP **64**, 1125-1130 (1996).

[9] What Parrondo calls " seminal document" is a slide entitled "How to cheat a bad mathematician" that can be found on his website:

http://seneca.fis.ucm.es/parr/

[10] G. P. Harmer & D. Abbott, "Parrondo's paradox," Statistical Science **14**, 206-213 (1999). See also G.P.Harmer, D.Abbott, P.G.Taylor & J.M.R.Parrondo, "Brownian ratchets and Parrondo's games," Chaos **11**, 705-714 (2001).

Group at the University of Adalaide maintains an index[11]—has by now grown to be quite extensive. I have found, however, that many of the papers that I have had an opportunity to examine focus on collateral issues and fail to expose with sharp clarity the mathematical point that lies at the heart of Parrondo's phenomenon, which is my present intent.

Here I look to the Parrondo's phenomenon as it emerges from the theory of random walks on finite graphs, and in a companion essay will look to walks on $\mathbb{Z}$ (the simplest graph of infinite order), which is mathematically more demanding. The present discussion will acquire a ratchet-like flavor, while the discussion of walks on $\mathbb{Z}$ lends itself to the game-theoretic interpretation originally favored by Parrondo.[9]

Essentials of the theory of random walks on finite graphs. A walker strides from vertex to vertex on a finite graph (vertices numbered $1, 2, \ldots, n$). The vertices that are single-step-accessible from any given vertex are indicated by the edges that radiate from that vertex. To each such directed edge is assigned a probability

$$m_{ij} = \text{probability of stepping } i \leftarrow j$$

Let $p_j(k)$ denote the probability that after k steps the walker occupies the j^{th} vertex, and from those assemble the "stochastic vector"[12]

$$\boldsymbol{p}_k = \begin{pmatrix} p_1(k) \\ p_2(k) \\ \vdots \\ p_n(k) \end{pmatrix}$$

Immediately, $p_i(k+1) = \sum_j m_{ij} p_j(k)$ which can be written

$$\boldsymbol{p}_{k+1} = \mathbb{M}\,\boldsymbol{p}_k \quad \text{with} \quad \mathbb{M} = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{pmatrix}$$

Suppose the initial location of the walker is known, given by (say)

---

[11] Go to http://www.eleceng.adalaide.edu.au/groups/parrondo. For recent work relating specifically to stochastic ratchets, see C. Jarzynski & O. Mazonka, "Feynman's ratchet and pawl: An exactly solvable model," Phys. Rev. E **59**, 6448-6459 (1999) and Y. Lee, A. Allison, D. Abbott & H. Eugene Stanley, "Minimal Brownian ratchet: An exactly solvable model," PR Letters **91**, 220601 (2003)—see the first of the ADDENDA to this essay—and papers cited there.

[12] The defining properties of such vectors are (i) that their elements fall within the unit interval $[0, 1]$ and (ii) sum to unity, and so describe a discrete distribution.

$$\boldsymbol{p}_0 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

We can speak only probabilistically about where the walker's first step might take him, but it must take him *somewhere*. We have

$$\boldsymbol{p}_1 = \mathbb{M}\,\boldsymbol{p}_0 = \begin{pmatrix} m_{11} \\ m_{21} \\ \vdots \\ m_{n1} \end{pmatrix}$$

and conclude on this basis that all the columns of the transition matrix $\mathbb{M}$ must be stochastic; *i.e.*, that $\mathbb{M}$ must be a **Markov matrix**. The (non-zero) elements of the j^{th} column of $\mathbb{M}$ decorate the edges that radiate from the j^{th} vertex, and will be considered to be fixed/constant attributes of the graph itself. Individual walks that proceed from a given intitial vertex are highly diverse, but it is easy to describe the evolving statistics of large populations of such walks: one has the **Markov process**

$$\boldsymbol{p}_0 \to \boldsymbol{p}_1 \to \boldsymbol{p}_2 \to \cdots \to \boldsymbol{p}_k = \mathbb{M}^k \boldsymbol{p}_0 \tag{1}$$

To develop the (not necessarily integral) powers of $\mathbb{M}$ we in numerical work might use the *Mathematica* command `MatrixPower[`$\mathbb{M}, k$`]` but for analytical purposes employ the generalized spectral decomposition

$$\mathbb{M} = \lambda_1\mathbb{P}_1 + \lambda_2\mathbb{P}_2 + \cdots + \lambda_n\mathbb{P}_n \tag{2.1}$$

where $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ is the spectrum of $\mathbb{M}$ and where $\{\mathbb{P}_1, \mathbb{P}_2, \ldots, \mathbb{P}_n\}$ is a certain complete set of orthogonal projection matrices

$$\sum_{i=1}^{n} \mathbb{P}_i = \mathbb{I} \quad \text{and} \quad \mathbb{P}_i\mathbb{P}_j = \delta_{ij}\mathbb{P}_j \tag{2.2}$$

constructed as follows:[13]

Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ and $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be right eigenvectors of $\mathbb{M}$ and $\mathbb{M}^{\mathsf{T}}$, respectively:[14]

$$\mathbb{M}\,\boldsymbol{u}_i = \lambda_i\,\boldsymbol{u}_i \quad \text{and} \quad \mathbb{M}^{\mathsf{T}}\boldsymbol{v}_i = \lambda_i\,\boldsymbol{v}_i$$

The $\boldsymbol{u}$ and $\boldsymbol{v}$-vectors are readily seen to be *biorthogonal* in the sense

[13] See "Generalized spectral resolution and some of its applications," (April, 2009).

[14] The row vectors $\boldsymbol{v}_i^{\mathsf{T}}$ are left eigenvectors of $\mathbb{M}$.

$$\boldsymbol{u}_i \perp \boldsymbol{v}_j \quad \text{if} \quad \lambda_i \neq \lambda_j$$

and one can arrange to have biorthogonality in the more general sense $\boldsymbol{u}_i \perp \boldsymbol{v}_{j\neq i}$ even in cases of spectral degeneracy. From biorthogonality it follows that the matrices

$$\mathbb{P}_i = \frac{\boldsymbol{u}_i \cdot \boldsymbol{v}_i^{\mathsf{T}}}{(\boldsymbol{u}_i, \boldsymbol{v}_i)} \quad : \quad n = 1, 2, \ldots, n \tag{3}$$

possess the properties that give rise to (2). These results pertain to *all* real square matrices, and in symmetric cases give back the standard theory of spectral decomposition. Our present interest, however, is in their application to Markov matrices.[15]

The leading eigenvalue of every Markov matrix is unity, and the other eigenvalues, whether real or complex (complex eigenvalues occur in conjugate pairs), fall *within or on the boundary of the unit disk*. Simple examples of the latter sort are

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad : \quad \text{eigenvalues } \{1, -1\}$$

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad : \quad \text{eigenvalues } \{1, \omega, \omega^2\} \text{ with } \omega = e^{i\frac{2\pi}{3}}$$

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad : \quad \text{eigenvalues } \{1, \omega, \omega^2 = -1, \omega^3 = -\omega\} \text{ with } \omega = e^{i\frac{2\pi}{4}}$$

but those Markov matrices are *permutation* matrices: they direct the walker to visit the vertices in a specific cyclic order devoid of any randomness. More interesting is the example

$$\begin{pmatrix} 0 & a & 0 & 1-a \\ 1-a & 0 & a & 0 \\ 0 & 1-a & 0 & a \\ a & 0 & 1-a & 0 \end{pmatrix} \quad : \quad \text{eigenvalues } \{\pm 1, \pm \lambda\} \tag{4}$$

where $\lambda = \sqrt{-1 + 4a - 4a^2}$ is complex except at $a = \frac{1}{2}$ (where λ vanishes). Here the walker executes a nearest-neighbor walk on the perimeter of a square (cyclic graph of order 4); he advances with probability a, retreats with probability $1-a$. Standing-in-place is disallowed by the 0s on the diagonal. The eigenvalue $\lambda = -1$ exerts a profound effect on the asymptotics of the process, as will soon emerge.

From (2) we obtain

$$\mathbb{M}^k = \lambda_1^k \mathbb{P}_1 + \lambda_2^k \mathbb{P}_2 + \cdots + \lambda_n^k \mathbb{P}_n \tag{5}$$

[15] At this point I must be content to make certain assertions without benefit of explicit proof, all of which could be supported by numerical experimentation done with randomly constructed populations of Markov matrices.

For Markov matrices with spectra of the form $\{1, \lambda_2, \lambda_3, \ldots, \lambda_n\} : |\lambda_i| < 1$ we therefore have

$$\lim_{k\to\infty} \mathbb{M}^k = \mathbb{P}_1 \tag{6}$$

where $\mathbb{P}_1$ projects onto the leading "eigenray" of $\mathbb{M}$ (by which I mean the real 1-space of vectors $\boldsymbol{e}$ that satisfy $\mathbb{M}\boldsymbol{e} = \boldsymbol{e}$). We take note now of these additional general properties of Markov matrices:[15]

- If $\boldsymbol{e}$ satisfies $\mathbb{M}\boldsymbol{e} = \boldsymbol{e}$ then the elements of $\boldsymbol{e}$ are real to within a shared complex factor and are all of the same sign. Division by their sum produces a (necessarily real) *stochastic* vector, which will be denoted $\boldsymbol{p}_\infty$.
- If $\boldsymbol{e}$ satisfies $\mathbb{M}\boldsymbol{e} = \lambda\boldsymbol{e}$ with $\lambda \neq 1$ then the elements of $\boldsymbol{e}$, whether real or complex, have mixed signs and sum to zero. Such vectors *cannot be rendered stochastic*.

For Markov processes of type $\{1, \lambda_2, \lambda_3, \ldots, \lambda_n\}$ we therefore have the simple asymptotic statement

$$\boldsymbol{p}_0 \longrightarrow \mathbb{P}_1\boldsymbol{p}_0 = \boldsymbol{p}_\infty \quad : \quad \text{all } \boldsymbol{p}_0$$

Note that all the information that distinguishes one initial state from another is asymptotically lost: the projection matrix $\mathbb{P}_1$ is not invertible. More to the point, non-singular Markov matrices, though invertible—we by (2) have

$$\mathbb{M}^{-1} = \mathbb{P}_1 + \lambda_2^{-1}\mathbb{P}_2 + \cdots + \lambda_n^{-1}\mathbb{P}_n$$

—possess inverses that are non-Markovian, since the numbers λ_i^{-1} fall outside the unit disk.

Look now again, by way of contrast, to (4), which typifies nearest-neighbor walks of type $\{1, -1, \lambda_3, \lambda_4, \ldots, \lambda_n\}$. After many steps we have

$$\boldsymbol{p}_k \sim \big\{\mathbb{P}_1 + (\pm)^k\mathbb{P}_2\big\}\boldsymbol{p}_0 \quad : \quad k \text{ large}$$

By calculation (here I write $\pm$ for ± 1)

$$\mathbb{P}_1 = \tfrac{1}{4}\begin{pmatrix} + & + & + & + \\ + & + & + & + \\ + & + & + & + \\ + & + & + & + \end{pmatrix}, \quad \mathbb{P}_2 = \tfrac{1}{4}\begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}$$

so

$$\mathbb{P}_1 + \mathbb{P}_2 = \tfrac{1}{2}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \quad \mathbb{P}_1 - \mathbb{P}_2 = \tfrac{1}{2}\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

We find, therefore, that after many steps the state "blinks"

$$\cdots \to \boldsymbol{p}_{\text{even}} \to \boldsymbol{p}_{\text{odd}} \to \boldsymbol{p}_{\text{even}} \to \boldsymbol{p}_{\text{odd}} \to \cdots$$

where

$$\boldsymbol{p}_0 = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix}, \quad \boldsymbol{p}_{\text{even}} = \begin{pmatrix} \alpha \\ \beta \\ \alpha \\ \beta \end{pmatrix}, \quad \boldsymbol{p}_{\text{odd}} = \begin{pmatrix} \beta \\ \alpha \\ \beta \\ \alpha \end{pmatrix} \quad \text{with} \quad \begin{cases} \alpha = \frac{1}{2}(p_1 + p_3) \\ \beta = \frac{1}{2}(p_2 + p_4) \end{cases}$$

The states $\boldsymbol{p}_{\text{even}}$ and $\boldsymbol{p}_{\text{odd}}$ retain some memory of the initial state, but the *average* of those states

$$\tfrac{1}{2}\big(\boldsymbol{p}_{\text{even}} + \boldsymbol{p}_{\text{odd}}\big) = \mathbb{P}_1\boldsymbol{p}_0 = \tfrac{1}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} \quad : \quad \text{all } \boldsymbol{p}_0$$

retains no such memory. The origin of the blinking phenomenon has to do with the fact that we can color the vertices of the square graph blue and red in such a way that the nearest neighboring vertices of every blue vertex are red and *vice versa*. A walker who departs from (say) a blue vertex will—with certainty—after an even number of steps stand on a blue vertex, and after an odd number of steps stand on a red vertex. Every cyclic graph of even order is in this sense a blinker. So are the cube and the planar projections of hypercubes of all orders. And so, as I have had previous occasion to remark,[16] are all finite rectangular and hexagonal tilings. Blinkers are, in short, commonplace. And it must be borne in mind that the simple asymptotic formula (6) pertains only to non-blinkers.

Random walks on the cyclic graph of order three. The simplest non-blinker graph is the cyclic graph of order 3, which is "complete" in the standard sense that every vertex is a nearest neighbor of (*i.e.*, linked by an edge to) every other vertex. The most general walk on such a graph (stand-in-place forbidden) is generated by the Markov matrix

$$\mathbb{M} = \begin{pmatrix} 0 & Y & z \\ x & 0 & Z \\ X & y & 0 \end{pmatrix} \tag{7}$$

where $X = 1 - x$, $Y = 1 - y$, $Z = 1 - z$. The spectrum of $\mathbb{M}$ reads

$$\{1, \lambda_2, \lambda_3\} \quad \text{with} \quad \begin{cases} \lambda_2 = -\tfrac{1}{2}(1+\sigma) \\ \\ \lambda_3 = -\tfrac{1}{2}(1-\sigma) \end{cases}$$

where

$$\begin{aligned}\sigma &= \sqrt{-3 + 4(x+y+z) - 4(xy+yz+zx)} \\ &= \sqrt{(x-X)(Y-y) + (y-Y)(Z-z) + (z-Z)(X-x)}\end{aligned}$$

We have $\mathbb{M} = \mathbb{P}_1 + \lambda_2\mathbb{P}_2 + \lambda_3\mathbb{P}_3$, and if we were interested in the details of the sequence $\boldsymbol{p}_0 \rightarrow \boldsymbol{p}_1 \rightarrow \boldsymbol{p}_2 \rightarrow \cdots \rightarrow \boldsymbol{p}_k \rightarrow \cdots$ would busy ourselves with construction of the $\mathbb{P}$-matrices that follow from (7). But if—as is the case—we

[16] "Some miscellaneous adventures in experimental mathematical physics," Notes for a Reed College Physics Seminar (9 November 2011). See the "Blinking Graphs" notebook that is stored in that file.

had interest only in the asymptotics of the process we could forego that tedium: we could (as below) proceed directly to the calculation of the stochastic solution $\boldsymbol{p}_\infty$ of $\mathbb{M}\boldsymbol{e} = \boldsymbol{e}$. *Mathematica* supplies

$$\boldsymbol{e} = \mathcal{K} \cdot \begin{pmatrix} 1 - y + yz \\ 1 - z + zx \\ 1 - x + xy \end{pmatrix} \quad : \quad \mathcal{K} \text{ any constant}$$

whence

$$\begin{aligned} \boldsymbol{p}_\infty &= \frac{1}{\mathcal{D}} \begin{pmatrix} 1 - y + yz \\ 1 - z + zx \\ 1 - x + xy \end{pmatrix} \\ \mathcal{D} &= 3 - (x + y + z) + (xy + yz + zx) \\ &= 2 + \lambda_2\lambda_3 = 2 + \det \mathbb{M} \end{aligned}$$

which in the special cases $x = y = z = a \in [0, 1]$, *i.e.*, when the next-step probabilities are site-independent, becomes

$$\boldsymbol{p}_\infty = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$$

The walker—in such cases, but not more generally—is equally likely, after many steps, to be found at any of the three vertices of the graph.

Probability current. The stochastic vector $\boldsymbol{p}_k$ that assigns probabilities to the the possible positions of a walker after he has taken k steps on $\mathbb{Z}$, or possible winnings after k flips of a loaded coin, are infinite dimensional, and evolve by action of an infinite dimensional Markov matrix. It is intuitively evident that such processes do not proceed to a steady asymptotic state; in this fundamental respect random walks on graphs of infinite order differ profoundly from random walks on graphs of finite order. We do, however, expect *velocity* of such a walker (rate of growth of the expected mean position, or expected winnings) to become asymptotically steady (on $\mathbb{Z}$, but not on $\mathbb{Z}^D : D > 1$). The theory of walks on finite graphs—even though on such graphs $\boldsymbol{p}_\infty$ is constant/unchanging—supports an analog of the latter idea, as I proceed now to explain. For expository purposes I work within the context provided by (7),[17] of which I adopt temporarily this notational variant:

$$\mathbb{M} = \begin{pmatrix} 0 & 1 - m_2 & m_3 \\ m_1 & 0 & 1 - m_3 \\ 1 - m_1 & m_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & M_{1\leftarrow 2} & M_{1\leftarrow 3} \\ M_{2\leftarrow 1} & 0 & M_{2\leftarrow 3} \\ M_{3\leftarrow 1} & M_{3\leftarrow 2} & 0 \end{pmatrix}$$

[17] This is a bit like describing Khirchhoff's Laws as they relate to a specific simple circuit. It will become obvious that the "probability current" idea—of which I was made aware by Y. Lee *et al*[11]—pertains quite generally to random walks on finite graphs.

We show that the $\mathbb{M}$-induced adjustments $\boldsymbol{p}_k \to \boldsymbol{p}_{k+1}$ of the probabilities at the respective vertices can be considered to be the result of *probability flow* along the edges of the graph. To that end we define **probability currents** (after the k^{th} step)

$$\begin{aligned} J_{1\to2}(k) &= M_{2\leftarrow1}\,p_1(k) - M_{1\leftarrow2}\,p_2(k) = m_1\,p_1(k) - (1-m_2)p_2(k) \\ J_{2\to3}(k) &= M_{3\leftarrow2}\,p_2(k) - M_{2\leftarrow3}\,p_3(k) = m_2\,p_2(k) - (1-m_3)p_3(k) \\ J_{3\to1}(k) &= M_{1\leftarrow3}\,p_3(k) - M_{3\leftarrow1}\,p_1(k) = m_3\,p_3(k) - (1-m_1)p_1(k) \end{aligned} \tag{8}$$

with $J_{i\leftarrow j}(k) = -J_{i\to j}(k)$. We expect now to have (for example)

$$\begin{aligned} p_1(k+1) = p_1(k) &+ \Delta p_1(k) \\ \Delta p_1(k) &= J_{3\to1}(k) - J_{1\to2}(k) \end{aligned}$$

which—as I now demonstrate—checks out: we have

$$\begin{aligned} \Delta p_1(k) &= J_{3\to1}(k) + J_{2\to1}(k) \\ &= \{M_{1\leftarrow3}\,p_3(k) - M_{3\leftarrow1}\,p_1(k)\} - \{M_{2\leftarrow1}\,p_1(k) - M_{1\leftarrow2}\,p_2(k)\} \\ &= m_3\,p_3(k) - (1-m_1)\,p_1(k) - m_1\,p_1(k) + (1-m_2)\,p_2(k) \\ &= -p_1(k) + (1-m_2)\,p_2(k) + m_3\,p_3(k) \end{aligned}$$

giving

$$\begin{aligned} p_1(k+1) &= p_1(k) + \Delta p_1(k) \\ &= (1-m_2)\,p_2(k) + m_3\,p_3(k) \end{aligned}$$

in precise agreement with the result obtained from $\boldsymbol{p}_{k+1} = \mathbb{M}\boldsymbol{p}_k$. When we assign to the stochastic state its asymptotic value—which in present notation reads

$$\begin{aligned} \boldsymbol{p}_\infty &= \frac{1}{\mathcal{D}}\begin{pmatrix} 1-m_2+m_2m_3 \\ 1-m_3+m_3m_1 \\ 1-m_1+m_1m_2 \end{pmatrix} \\ \mathcal{D} &= 3-(m_1+m_2+m_3)+(m_1m_2+m_2m_3+m_3m_1) \end{aligned}$$

—we find that

$$\begin{aligned} J_{1\to2}(\infty) = J_{2\to3}(\infty) = J_{3\to1}(\infty) &= \frac{m_1m_2m_3-(1-m_1)(1-m_2)(1-m_3)}{\mathcal{D}} \\ &\equiv \frac{\mathcal{N}(m_1,m_2,m_3)}{\mathcal{D}(m_1,m_2,m_3)} \equiv J_\infty(m_1,m_2,m_3) \end{aligned} \tag{9}$$

The probability current has become the same on every edge; the steady state is maintained because inflow balances outflow at every vertex. And the argument can be reversed: when asked to solve the system

$$\begin{gathered} m_1q_1-(1-m_2)q_2 = m_2q_2-(1-m_3)q_3 = m_3q_3-(1-m_1)q_1 \\ q_1+q_2+q_3 = 1 \end{gathered}$$

Mathematica returns precisely the definition of $\boldsymbol{p}_\infty$.

In cases of the type $m_1 = m_2 = m_3 = a$, where the transition probabilities are site-independent, we have

$$J_{1\to2}(\infty) = J_{2\to3}(\infty) = J_{3\to1}(\infty) = \frac{a^3 - (1-a)^3}{\mathcal{D}(a,a,a)}$$

$\mathcal{D}(a,a,a) = 3(1 - a + a^2)$ is positive for all a, so in such cases the probability current is prograde or retrograde according as a is greater than or less than $\frac{1}{2}$.

More generally, we find[18]

$$2 \leqslant \mathcal{D}(m_1, m_2, m_3) \leqslant 3 \quad : \quad \{m_1, m_2, m_3\} \in \text{unit cube}$$

(which is to say: $\mathcal{D}$ is always positive) so the asymptotic circulation is prograde or retrograde according as $\mathcal{N}(m_1, m_2, m_3)$ is positive or negative. We acquire interest therefore in—see PLATE 1—the null surface defined

$$\mathcal{N}(m_1, m_2, m_3) = 0$$

Edge currents can be associated with random walks on graphs—whether finite or infinite—of *any* design. The "probability current" concept provides an alternative way to think about the "conservation of probability" (persistence of stochasticity) that is built into the design of Markov matrices, and brings to mind the random walk approach to the analysis of resistive circuits that is illustrated in some notes already cited.[16] It is, however, only on non-blinker finite graphs (and particularly on cyclic graphs of odd order) that the vertex probabilities and edge currents inevitably become steady. The remarkable fact is that the asymptotic steady currents *need not vanish*. They endow walks on such graphs with an analog of "velocity" concept (drifting to the right/left, winning/losing) that emerges naturally from the theory of random walks on $\mathbb{Z}$.

Composite Markov processes. It is obvious on persistence of stochasticity grounds that

- Products $\mathbb{M}_1\mathbb{M}_2\cdots\mathbb{M}_n$ of Markov matrices are Markovian.
- Stochastically weighted linear combinations $q_1\mathbb{M}_1 + q_2\mathbb{M}_2 + \cdots + q_n\mathbb{M}_n$ are Markovian.

The latter fact is exploited in "Minimal Brownian ratchet: An exactly solvable model" by Stanley, Abbott and colleagues.[11] We look here to a simple instance of the former fact. From

$$\mathbb{A} = \begin{pmatrix} 0 & 1-a & a \\ a & 0 & 1-a \\ 1-a & a & 0 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} 0 & 1-y & z \\ x & 0 & 1-z \\ 1-x & y & 0 \end{pmatrix}$$

[18] Most easily by graphic demonstration: plot $\mathcal{D}(m_1, m_2, m_3)$ for assorted values of m_3.

we assemble the composite Markov matrix

$$\mathbb{C} = \mathbb{A}\mathbb{B} = \begin{pmatrix} (1-a)x + a(1-x) & ay & (1-a)(1-z) \\ (1-a)(1-x) & (1-a)y + a(1-y) & az \\ ax & (1-a)(1-y) & (1-a)z + a(1-z) \end{pmatrix}$$

It is readily verified that (for $\{a, x, y, z\} \in [0, 1]$) the columns of $\mathbb{C}$ are stochastic; *i.e.*, that $\mathbb{C}$ is Markovian. Note the non-zero elements on the diagonal of $\mathbb{C}$. Those arise from the circumstance that after two steps—each subject to the "no stand in place" rule—a walker can *return* to place (in either of two ways).

By "Parrondo ratchet" I refer to the Markov process generated by $\mathbb{C}$. Already in hand are descriptions of the asymptotic states and currents to which $\mathbb{A}$ and $\mathbb{B}$ give rise. We look now (with *Mathematica*'s assistance) to those that arise from $\mathbb{C}$. To that end, let $\mathbb{C}$ be notated

$$\mathbb{C} = \begin{pmatrix} 1-u-U & v & W \\ U & 1-v-V & w \\ u & V & 1-w-W \end{pmatrix} \tag{10}$$

We verify that the leading eigenvalue is unity, and—looking to the leading eigenvector—find that

$$\boldsymbol{p}_\infty = \frac{1}{\mathcal{D}} \begin{pmatrix} vw + (v + V)W \\ wu + (w + W)U \\ uv + (u + U)V \end{pmatrix} \tag{11}$$
$$\mathcal{D} = (uv + vw + wu) + (uV + vW + wU) + (UV + VW + WU)$$

The diagonal elements of $\mathbb{C}$ make no contribution to the edge currents. Working from (8) by (10) and (11), we obtain

$$\left.\begin{aligned} J_{1\to2}(\infty) &= U\,p_1(\infty) - v\,p_2(\infty) \\ J_{2\to3}(\infty) &= V\,p_2(\infty) - w\,p_3(\infty) \\ J_{3\to1}(\infty) &= W\,p_3(\infty) - u\,p_1(\infty) \end{aligned}\right\} = \frac{\mathcal{N}(u, v, w, U, V, W)}{\mathcal{D}(u, v, w, U, V, W)}$$

with

$$\mathcal{N}(u, v, w, U, V, W) = UVW - uvw$$

This *apart from a reversed sign* is structurally identical to (9), and when spelled out in detail becomes

$$\mathcal{N}(a, x, y, z) = (1-a)^3(1-x)(1-y)(1-z) - a^3xyz$$

The Parrondo phenomenon. The Markov matrix $\mathbb{A}$ generates an a-parameterized family of random walks on the cyclic graph of order 3, the matrix $\mathbb{B}$ generates a $\{x, y, z\}$-parameterized family, and $\mathbb{C} = \mathbb{A}\mathbb{B}$ generates a $\{a, x, y, z\}$-parameterized family. We are in possession now of descriptions of the probability currents that persist in the asymptotic states of those respective walks:

$$
\begin{aligned}
J_A(a) &= \frac{\mathcal{N}(a)}{\mathcal{D}(a)}\\
J_B(x,y,z) &= \frac{\mathcal{N}(x,y,z)}{\mathcal{D}(x,y,z)}\\
J_C(a,x,y,z) &= \frac{\mathcal{N}(a,x,y,z)}{\mathcal{D}(a,x,y,z)}
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{N}(a) &= a^3-(1-a)^3\\
\mathcal{D}(a) &= 3(1-a+a^2)
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{N}(x,y,z) &= xyz-(1-x)(1-y)(1-z)\\
\mathcal{D}(x,y,z) &= 3-(x+y+z)+(xy+yz+zx)
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{N}(a,x,y,z) &= (1-a)^3(1-x)(1-y)(1-z)-a^3xyz\\
\mathcal{D}(a,x,y,z) &= a^2(xy+yz+zx)\\
&\quad +a(1-a)\big[x(1-y)+y(1-z)+z(1-x)\big]\\
&\quad +(1-a)^2\big[(1-x)(1-y)+(1-y)(1-z)+(1-z)(1-x)\big]
\end{aligned}
$$

In each instance, the asymptotic circulation is prograde or retrograde according as $\mathcal{N}>0$ or $\mathcal{N}<0$. The boundaries between the two regimes are set by the null conditions

$$
\begin{aligned}
\mathcal{N}(a)=0 &\quad:\quad \text{marks a point } (a=\tfrac{1}{2}) \text{ on the unit interval}\\
\mathcal{N}(x,y,z)=0 &\quad:\quad \text{inscribes a surface within the unit cube}\\
\mathcal{N}(a,x,y,z)=0 &\quad:\quad \text{inscribes a surface within the unit 4-cube}
\end{aligned}
$$

Graphic display of surfaces in 4-space necessarily proceeds section by section. With Parrondo, we look specifically to the section that arises from setting $z=y$, on which

$$
\begin{aligned}
\mathcal{N}(x,y,y) &= xy^2-(1-x)(1-y)^2\\
\mathcal{D}(x,y,y) &= 3-(x+2y)+(2xy+y^2)
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{N}(a) = \mathcal{N}(a,a,a) &= a^3-(1-a)^3\\
\mathcal{D}(a) = \mathcal{D}(a,a,a) &= 3-3a+3a^2
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{N}(a,x,y,y) &= (1-a)^3(1-x)(1-y)^2-a^3xy^2\\
\mathcal{D}(a,x,y,y) &= a^2(2xy+y^2)\\
&\quad +a(1-a)\big[x+2y-2xy-y^2\big]\\
&\quad +(1-a)^2\big[3-2x-4y+2xy+y^2\big]
\end{aligned}
$$

The null surfaces $\mathcal{N}(a) = \mathcal{N}(x,y,y) = \mathcal{N}(a,x,y,y) = 0$ are seen (PLATE 2) to partition the unit cube into six regions. The helicities that arise in those respective regions are tabulated below:

Region	$\mathbb{A}$	$\mathbb{B}$	$\mathbb{C}$
1	$\circlearrowleft$	$\circlearrowleft$	$\circlearrowright$
2	$\circlearrowleft$	$\circlearrowright$	$\circlearrowright$
3	$\circlearrowleft$	$\circlearrowright$	$\circlearrowleft$
4	$\circlearrowright$	$\circlearrowleft$	$\circlearrowright$
5	$\circlearrowright$	$\circlearrowleft$	$\circlearrowleft$
6	$\circlearrowright$	$\circlearrowright$	$\circlearrowleft$

I give the name "Parrondo's phenomenon" to the striking fact that in every case one walk circulates in the sense contrary to that of the other two. Parrondo himself drew attention to the "paradoxes" that arise when $\{a, x, y\}$ live in Region **1** ("$\mathbb{A}$ and $\mathbb{B}$ lose, but $\mathbb{C} = \mathbb{A}\mathbb{B}$ wins") else in Region **6** ("$\mathbb{A}$ and $\mathbb{B}$ win, but $\mathbb{C}$ loses"). The other four cases are quite unparadoxical, since the observed $\mathbb{C}$-helicity can be attributed to the predominance of one of the others.

To make clear the identity of the respective regions we look to 2-dimensional a-sections of the surfaces shown in PLATE 2. To that end, we observe that according to *Mathematica*

$$\mathcal{N}(x,y,y) = 0 \quad \text{can be written} \quad y = f(x) \equiv \frac{x + \sqrt{x(1-x)} - 1}{2x - 1}$$

while $\mathcal{N}(a,x,y,y) = 0$ becomes

$$y = g(x\,;a) \equiv \frac{u(x\,;a) + \sqrt{v(x\,;a)}}{u(x\,;a)}$$

with

$$\begin{aligned} u(x\,;a) &= (a-1)^3 + x(1 - 3a + 3a^2) \\ v(x\,;a) &= a^3(a-1)^3 x(x-1) \end{aligned}$$

That information was used to consruct PLATE 3. Tabulated below are the currents associated with the parameters identified by intersecting grid lines:

Region	a	x	y	J_A	J_B	J_C
1	0.30	0.15	0.60	-0.1333	-0.0374	$+0.0717$
2	0.30	0.55	0.60	-0.1333	$+0.0555$	$+0.0387$
3	0.30	0.93	0.60	-0.1333	$+0.1379$	-0.0138
4	0.70	0.15	0.30	$+0.1333$	-0.1658	$+0.0184$
5	0.70	0.55	0.30	$+0.1333$	-0.0753	-0.0240
6	0.70	0.93	0.30	$+0.1333$	$+0.0233$	-0.0500

It is seen that the sign pattern conforms to the helicity pattern reported in the table at the top of the page.

PLATE 1 (page 13): The null surface $\mathcal{N}(m_1, m_2, m_3) = 0$ that arises from the Markov matrix

$$\mathbb{M} = \begin{pmatrix} 0 & 1-m_2 & m_3 \\ m_1 & 0 & 1-m_3 \\ 1-m_1 & m_2 & 0 \end{pmatrix}$$

that was introduced on page 11. Unit points on the m_1, m_2 and m_3 axes are indicated by red/green/blue ••• spheres, respectively. A gray sphere • marks the origin.

PLATE 2 (page 16): Superimposed null surfaces

$$\mathcal{N}(a) = \mathcal{N}(x, y, y) = \mathcal{N}(a, x, y, y) = 0$$

The red/green/blue spheres ••• now refer to unit points on the a, x and y axes. Current is negative on the origin side of the $\mathbb{A}$ and $\mathbb{B}$ surfaces, but (recall the *reversed sign* noted on page 14) positive on the origin side of the $\mathbb{C}$ surface. It is that sign reversal that in Regions 1 and 6 gives rise to "Parrondo's Paradox."

PLATE 3 (page 16): The null surfaces shown in PLATE 2 are sectioned at $a = 0.3$ (above) and $a = 0.7$ (below). The $\mathbb{A}$ current is negative at all points shown in the upper figure, positive at all points shown in the lower figure. Intersecting grid lines mark the parameter values that were used to constuct the second of the tables that appear on page 16.

PLATE 1

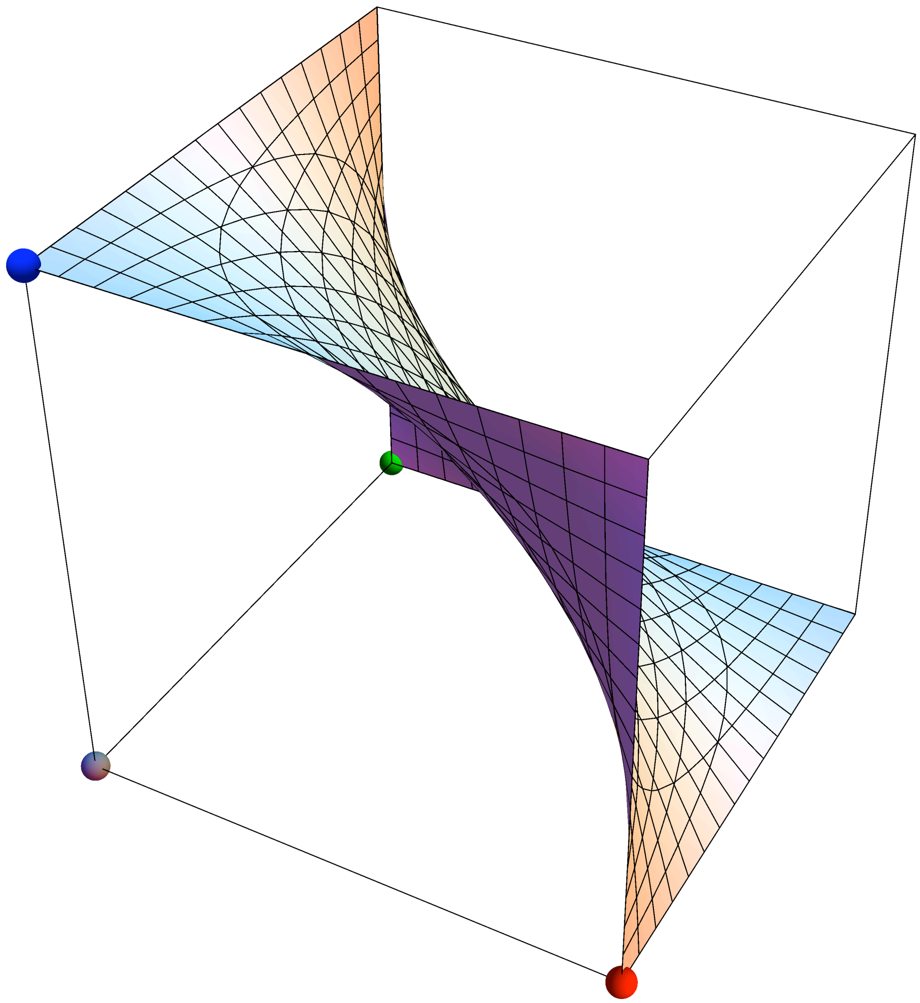


PLATE 2

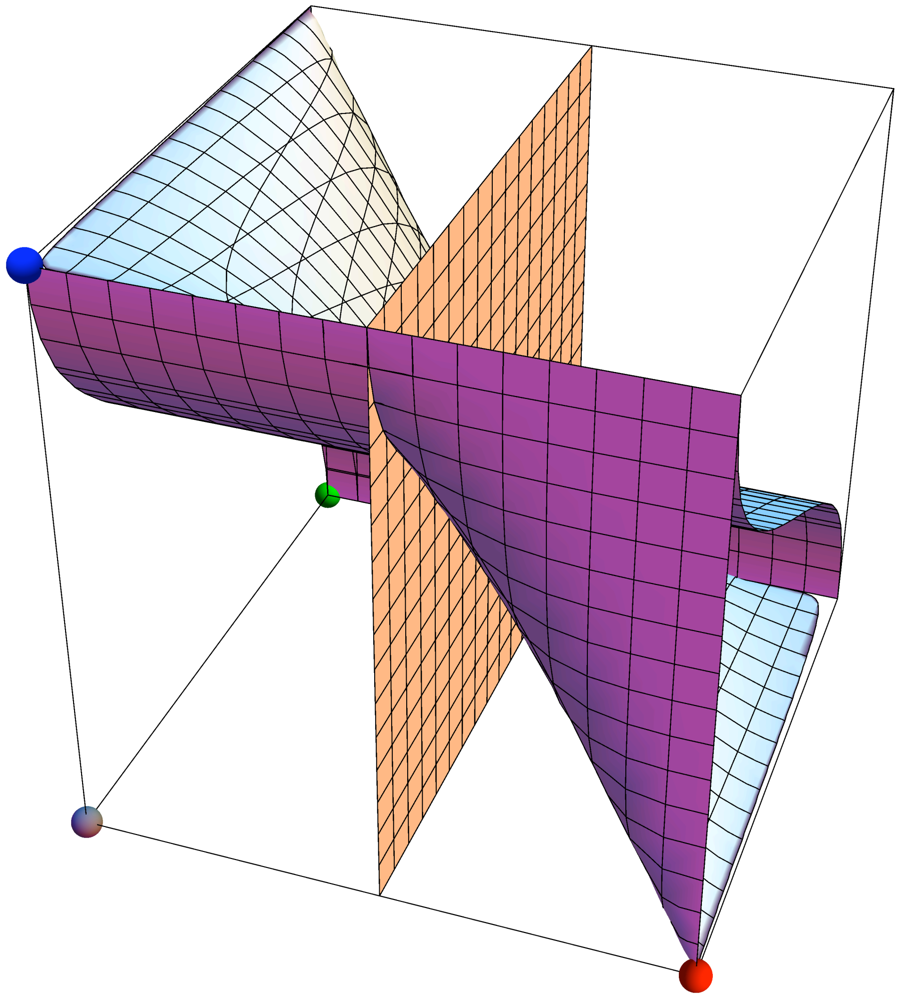
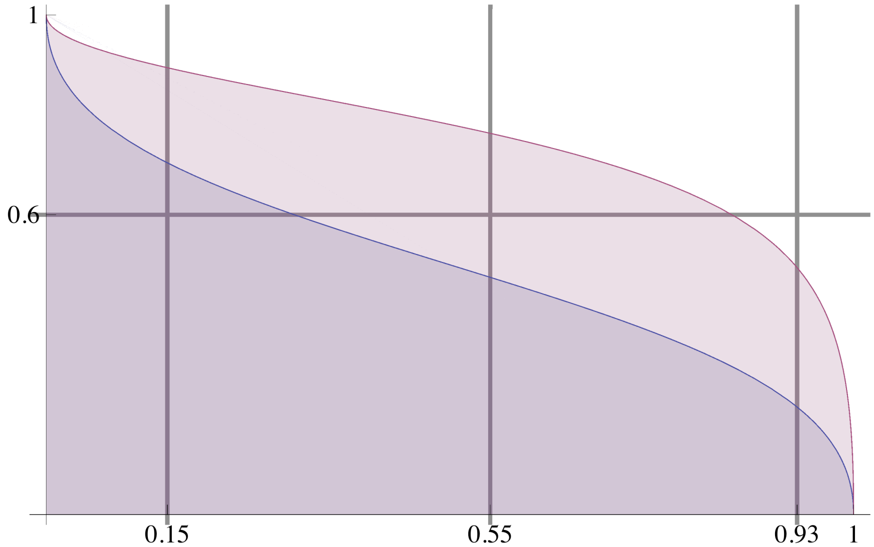
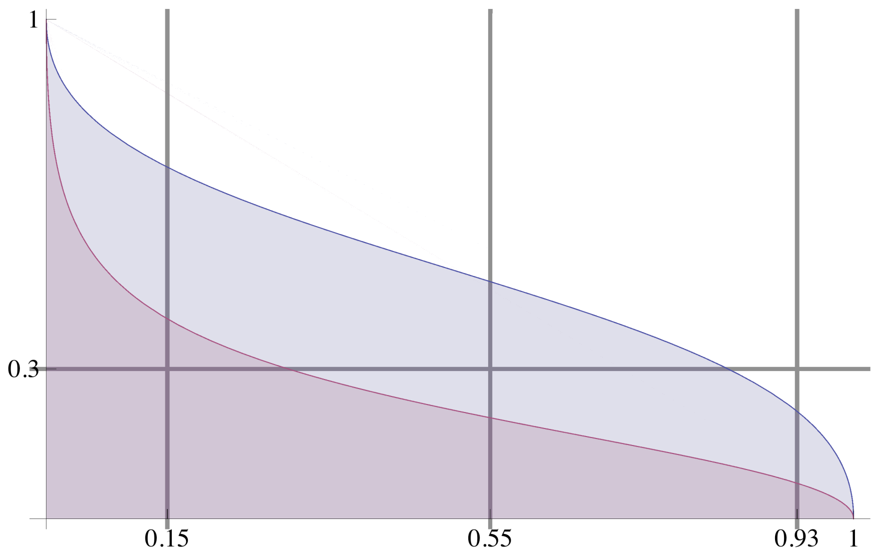


PLATE 3



Out[307]=



"Minimal Brownian ratchet". Reference was made[11] to "an exactly solvable model" devised collaboratively by members of H. Eugene Stanley's group at the Center for Polymer Studies & Department of Physics (Boston University) and Derek Abbott's group at the Centre for Biomedical Engineering & School of Electrical and Electronic Engineering (The University of Adelaide) in 2003. I sketch the essentials of the model because it employs the same material—Markov matrices $\mathbb{A}$ and $\mathbb{B}$ that generate random walks on a cyclic graph of order 3 and the (asymptotic) "probability current" concept—as were central to the preceding discussion, but makes novel use of that material.

To $\mathbb{A}$ and $\mathbb{B}$—defined as they were on page 13

$$\mathbb{A} = \begin{pmatrix} 0 & A & a \\ a & 0 & A \\ A & a & 0 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} 0 & Y & z \\ x & 0 & Z \\ X & y & 0 \end{pmatrix}$$

—we bring the requirement that both give rise to asympototic currents that vanish, which by (9) entails

$$\begin{aligned} a^3 = A^3 \qquad &\Longrightarrow \quad a = A = \tfrac{1}{2} \\ xyz = (1-x)YZ \qquad &\Longrightarrow \quad x = \frac{YZ}{yz+YZ},\ X = \frac{yz}{yz+YZ} \end{aligned} \tag{12}$$

The resulting matrices are assembled now not multiplicatively but (see again page 13) by stochastically weighted linear combination, producing

$$\begin{aligned} \mathbb{R} &= \gamma\mathbb{A} + (1-\gamma)\mathbb{B} \quad : \quad \gamma \in [0,1] \\ &= \begin{pmatrix} 0 & \bar\gamma Y + \frac{1}{2}\gamma & \bar\gamma\, z + \frac{1}{2}\gamma \\ \bar\gamma\, x + \frac{1}{2}\gamma & 0 & \bar\gamma Z + \frac{1}{2}\gamma \\ \bar\gamma X + \frac{1}{2}\gamma & \bar\gamma\, y + \frac{1}{2}\gamma & 0 \end{pmatrix} \quad : \quad \bar\gamma = 1-\gamma \end{aligned}$$

The matrix

$$\mathbb{A} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

is considered to introduce "noise" into the walk generated by $\mathbb{B}$, and γ to control the "volume" of the noise.

Looking again to (9), we find that the asymptotic $\mathbb{R}$-current is given by

$$J(y,z;\gamma) = \frac{\mathcal{N}(y,z;\gamma)}{\mathcal{D}(y,z;\gamma)}$$

with

$$\begin{aligned}\mathcal{N}(y,z;\gamma) &= \big(\bar\gamma x+\tfrac{1}{2}\gamma\big)\big(\bar\gamma y+\tfrac{1}{2}\gamma\big)\big(\bar\gamma z+\tfrac{1}{2}\gamma\big)\\ &\quad -\big(\bar\gamma X+\tfrac{1}{2}\gamma\big)\big(\bar\gamma Y+\tfrac{1}{2}\gamma\big)\big(\bar\gamma Z+\tfrac{1}{2}\gamma\big)\\ \mathcal{D}(y,z;\gamma) &= 3-\big(\bar\gamma x+\tfrac{1}{2}\gamma\big)-\big(\bar\gamma y+\tfrac{1}{2}\gamma\big)-\big(\bar\gamma z+\tfrac{1}{2}\gamma\big)\\ &\quad +\big(\bar\gamma x+\tfrac{1}{2}\gamma\big)\big(\bar\gamma y+\tfrac{1}{2}\gamma\big)\\ &\quad +\big(\bar\gamma x+\tfrac{1}{2}\gamma\big)\big(\bar\gamma z+\tfrac{1}{2}\gamma\big)\\ &\quad +\big(\bar\gamma z+\tfrac{1}{2}\gamma\big)\big(\bar\gamma x+\tfrac{1}{2}\gamma\big)\end{aligned}$$

where $\bar\gamma = 1-\gamma$ and where $\{x, X\}$ are understood to be given by (12). Entrusting the algebra to *Mathematica*, we obtain

$$\begin{aligned}\mathcal{N}(y,z;\gamma) &= \tfrac{1}{4(1-y-z+2yz)}\Big\{\ \gamma\,(-2+6y-4y^2+6z-16yz+8y^2z-4z^2+8yz^2)\\ &\qquad +\gamma^2(3-9y+6y^2-9z+24yz-12y^2z+6z^2-12yz^2)\\ &\qquad +\gamma^3(-1+3y-2y^2+3z-8yz+4y^2z-2z^2+4yz^2)\Big\}\\ \mathcal{D}(y,z;\gamma) &= \tfrac{1}{4(1-y-z+2yz)}\Big\{(8-8y-8z+24yz-8y^2z-8yz^2+8y^2z^2)\\ &\qquad +(2\gamma+\gamma^2)(-1+y+z+6yz-8y^2z-8yz^2+8)\Big\}\end{aligned}$$

The denominators are identical, so make no contribution to $J(y,z;\gamma)$. Graphic experimentation[19] serves to establish that $\mathcal{D}(y,z;\gamma)$ is positive for all admissible values of its arguments, so the sign of $J(y,z;\gamma)$ is fixed by the sign of $\mathcal{N}(y,z;\gamma)$. It is obvious that $\mathcal{N}(y,z;0)=0$ and evident by inspection that $\mathcal{N}(y,z;1)=0$; these are simply the conditions posited at the outset (both $\mathbb{A}$ and $\mathbb{B}$ generate walks with zero asymptotic current). Graphs showing the value assumed by $J(y,z;\gamma)$ at $\{x,y\}$ for other specified values of γ can be constructed in the manner just described,[19] and show—this being the point of the model!—that the asymptotic current (which can have either sign) is *typically not zero*:

Noise $\mathbb{A}$ can stimulate the ratchet $\mathbb{B}$ to rotate.

The essence of this striking result—which illustrates in a fresh way the elementary fact (see again page 5) that

> *Compound Markov processes can possess properties not shared by any of the contributory components*

—is most easily exposed if (with Parrondo) one sets $z=y$ and looks to graphs of $J(y,y;\gamma)$; *i.e.*, to J *vs.* y for fixed γ, else J *vs.* γ for fixed y.

One expects to see a similar phenomenon if $\mathbb{B}$ describes a nearest-neighbor walk on a cyclic graph of higher odd order, but with more parameters the analysis becomes more complicated (it is in this respect that the ratchet described above is "minimal") and graphic display becomes more awkward. Cyclic graphs of even order are "blinkers" in the sense discussed on pages 9 and 10. The noise-stimulated asymptotic motion of such ratchets is expected to be jiggly.

---

[19] Use `Plot3D[`$\mathcal{D}(y,z;\gamma)$`],`$\{y,0,1\},\{z,0,1\}$`]` for assorted values of γ.

Cubic soapfilm and the relativistic velocity addition rule . Use wires of unit length to connect the six vertices of a unit cube that remain after the vertices $(0,0,0)$ and $(1,1,1)$ have been deleted. Dip the resulting framework into soap solution and produce a film that very closely resembles the null surface shown in PLATE 1, which arose from an equation of the form

$$\begin{aligned} &xyz - (1-x)(1-y)(1-z) \\ &\quad = -1 + (x+y+z) - (xy+yz+zx) + 2xyz = 0 \end{aligned} \tag{13}$$

I am reminded that that same construction came to my attention once in quite another connection.[20] The "relativistic velocity addition formula"—which refers more properly to the rule for composing colinear Lorentz transformations

$$\mathbb{L}(\beta_2)\mathbb{L}(\beta_1) = \mathbb{L}(\beta_3) \quad : \quad \beta_3 = \frac{\beta_1 + \beta_2}{1 + \beta_1\beta_2}$$

—can be written

$$(\beta_1 + \beta_2 + \beta) + \beta_1\beta_2\beta = 0 \quad \text{with} \quad \beta = -\beta_3 \tag{14}$$

where $\{\beta, \beta_1, \beta_2, \beta_3\}$ range on $[-1,+1]$, as would result if we wrote

$$\begin{aligned} \beta_1 &= 2x - 1 \\ \beta_2 &= 2y - 1 \\ \beta\ &= 2z - 1 \end{aligned}$$

and allowed $\{x, y, z\}$ to range (as above) on $[0,1]$. With those substitutions we find that

$$\frac{\text{LHS of (14)}}{4} = \text{LHS of (13)}$$

which accounts for the similarity of the figures. Note, however, that (13) entails

$$z(x,y) = \frac{1 - x - y + xy}{1 - x - y + 2xy} = 1 - \frac{xy}{1 - x - y + 2xy}$$

which does *not* satisfy Lagrange's minimal surface equation

$$(1 + z_x^2)z_{yy} - 2z_x z_y z_{xy} + (1 + z_y^2)z_{xx} = 0$$

It is remarked in *Parrondo's Ratchet II* that the null surface is scribed by straight lines joining midpoints of opposite sides, so resolves into two sets of three congruent surface segments, bounded by line segments (a "folded kite") of lengths $\frac{1}{2}$ and $\frac{1}{2}\sqrt{2}$. Ray Mayer informed me in 1974 that the *minimal* surface bounded by such a folded quadrilateral was first described by Riemann.

[20] See "How Einstein might have been led to relativity already in 1895" (August 1999), page 26. This material is based upon a lecture I gave to mark the centennial of Einstein's birthday (14 March 1979).