Comments on Jill North's "Symmetry and Probability"

Branden Fitelson

July 31, 2006 @ BSPC

First, I'd like to thank Shieva, Ned, Andy, and the other organizers for putting together such a terrific conference, and for allowing me to participate in my first BSPC. Second, I want to congratulate Jill on a terrific paper. It's been a real pleasure thinking about this stuff and discussing it with Jill.

Jill's paper contains several distinct threads and arguments. I will focus only on what I see as the main theses of the paper, which involve the justification or grounding of the microcanonical probability distribution of classical statistical mechanics (MCD). I'll begin by telling the "canonical" story of the MCD (as I see it). Then I will discuss Jill's proposal. I will describe one worry that I have regarding her proposal, and I will offer a friendly amendment which seems to allay my worry.

Consider a finite, perfectly elastic box containing n perfectly elastic particles ("billiard balls"). Assume the particles obey Newton's laws, and are not being acted upon by any outside forces. Assume further that the particles all have unit mass, they all start out with finite velocities, and that the system (which I will call S) has been evolving for a very long time (*i.e.*, that S is at "equilibrium" in the sense that its macroscopic properties are *time-invariant*). Classical mechanics allows us (*in principle*) to determine the precise trajectories for each of the particles (*i.e.*, the *micro*-dynamics of S). But, if the number of particles is very large (*e.g.*, a *gas*), then the best we can *practically* hope for are (*i*) predictions about *macroscopic* (*i.e.*, thermodynamic) features of S, such as temperature, pressure, *etc.*, and (*ii*) *statistical* predictions about the state of the system S (or an *ensemble* of systems $\mathfrak{S}$ of which S is a member — see *fn.* 2). Our focus today is the latter, *i.e.*, equilibrium statistical mechanics for isolated systems (note: it can be used to *calculate* the former as "averages").

Traditionally, this branch of statistical mechanics has the following set-up. First, we write-down the "canonical" *Hamiltonian phase space* description of S. The traditional Hamiltonian phase space $\mathcal{H}$ for S has $6n$ coordinates or dimensions: three canonical spatial coordinates p_1, p_2, p_3, and three conjugate momentum coordinates q_1, q_2, q_3 for each of the n particles. Thus, the state of S (at a time) can be represented as a $6n$-vector (*i.e.*, a point in the $6n$-dimensional phase space $\mathcal{H}$), which specifies numerical values for the six "canonically conjugate coordinates" for each of the n particles of S. [I'll say more about these "canonically conjugate coordinates" and how they relate to Newtonian position and momentum coordinates, below.] Our traditional Hamiltonian description plays two roles. First, it allows us (*in principle* — *via* Hamilton's equations) to characterize the Newtonian trajectories of each of the n particles (*i.e.*, S's micro-dynamics). Second, it allows us to give the traditional *statistical* description of S. If we put a *uniform probability density* ρ on the (Lebesgue[1]) volume elements of our Hamiltonian phase space $\mathcal{H}$, then this yields the traditional statistical description of S. Formally, we have the following standard recipe: $\Pr(s \in R) \propto \int_R \rho \, d\mathbf{p} d\mathbf{q}$.[2] Informally, according to the standard theory, the probability that the state s of S (in equilibrium) is in some region R of Hamiltonian phase space $\mathcal{H}$ is proportional to the volume of R. In slogan form: "Assign equal probability to (s falling in) regions of phase space with equal volume." *Why*? Jill's proposal:

> This . . . gives us two reasons for [ρ]. (1) It yields empirically successful predictions. (2) It is uniform over the structure required by the dynamics; it is the simplest, most natural distribution, given the dynamics . . . [it] requires no further structure over and above the structure that is "already there" for the dynamics.

---

[1]Lebesgue $6n$-integration over R is the standard measure of "volume" for R's in $\mathcal{H}$. Alternatives include (1) non-standard measure theory (see *fn.* 4), and/or (2) "side length" rather than "volume" (as in van Fraassen's cube example).

[2]The *interpretation* of $\Pr(\cdot)$ is orthogonal to my remarks today, but (at least) a footnote on this is in order. Some [3, 281–88] say $\Pr(\cdot)$ attaches to an *ensemble* $\mathfrak{S}$ of systems having the same energy as S. On this view, s is the state of a "representative" member of $\mathfrak{S}$, and $\Pr(\cdot)$ is a "measure of our ignorance of S's *precise* initial state" (in some sense) or a physical property of the ensemble $\mathfrak{S}$ (*e.g.*, a limiting frequency). Others [3, 288–93] say $\Pr(\cdot)$ attaches to the *token* system S, again, either as a "measure of ignorance of S's initial state" (in some sense) or as a physical property of S.

There is no real controversy about (1). I will take for granted the empirical adequacy of the $\Pr(\cdot)$ that results from "going uniform over (Lebesgue) volume elements of the traditional phase space $\mathcal{H}$". But, (2) is not as straightforward, and here is where I would like to offer a friendly amendment. As it stands, the main problem I have with (2) is that there are *many* (*prima facie*) dynamically equivalent descriptions of S, which lead to *different* probability functions being generated by "going uniform over volume elements of phase space". For instance, instead of a Hamiltonian description $\mathcal{H}$ (plus Hamilton's equations of motion), we could have used a Newtonian (*a.k.a.*, Lagrangian) description $\mathcal{N}$ (plus Newton's/Lagrange's equations of motion). The Newtonian/Lagrangian approach uses classical *position* and *momentum* coordinates $\langle x, y, z, \dot{x}, \dot{y}, \dot{z} \rangle$ for each particle rather than the "canonically conjugate coordinates" $\langle p_1, p_2, p_3, q_1, q_2, q_3 \rangle$ of $\mathcal{H}$. The Hamiltonian "canonically conjugate coordinates" are related to the Newtonian position and momentum coordinates by a non-measure-preserving (Legendre) transformation. As a result, a uniform distribution ρ over (Lebesgue) volume elements of $\mathcal{H}$ will transform into a *non*-uniform distribution ρ' over (Lebesgue) volume elements of $\mathcal{N}$. Moreover, it has recently been shown [2] that there are alternative (*prima facie* dynamically/thermodynamically equivalent) *Hamiltonian* descriptions $\mathcal{H}'$ of S, which are "non-canonical" (*i.e.*, $\mathcal{H}'$ is related to $\mathcal{H}$ by a non-measure-preserving transformation). Thus, if we take $\mathcal{H}'$ as "the structure required by the dynamics", then generating $\Pr(\cdot)$ *via* the standard recipe will require a *non*-uniform ρ'' over $\mathcal{H}''$'s volume elements. Because of this ambiguity, it is unclear precisely what it means to say that $\Pr(\cdot)$ results from a uniform underlying density over "the structure required by the dynamics". That is, the existence of such (*prima facie* dynamically equivalent) "non-canonical" descriptions suggests that some additional postulate is needed here to *single out* the "canonical" Hamiltonian phase space description $\mathcal{H}$ as "*the* structure" in the desired sense. This brings us to my friendly amendment. Consider the following *desideratum*, which has both dynamical and probabilistic content and which is widely accepted in the literature.[3]

(†) The probability density ρ should be a *conserved quantity* of the dynamics.

> [Rationale: the macroscopic properties of S are *time-invariant*, and ρ *determines* these as "averages".]

It turns out that $\mathcal{H}$ is the *only* description $\mathcal{D}$ that is compatible with *all three* of the following:

(a) (Lebesgue[4]) phase volume of $\mathcal{D}$ is a *conserved quantity* of the dynamics of $\mathcal{D}$.

(b) ρ is *uniform* over (Lebesgue) volume elements of $\mathcal{D}$. [Note: (a) & (b) $\Longrightarrow$ (†).]

(c) $\Pr(s \in R) \propto \int_R \rho \, d\Gamma$ is *empirically adequate* [where $d\Gamma$ is a (Lebesgue) volume element of $\mathcal{D}$].

In other words — assuming (†) — the *only* way to get an empirically adequate $\Pr(\cdot)$ using the standard recipe is by *going uniform over the (Lebesgue) volume elements of $\mathcal{H}$*. Thus, I think (†) bolster's Jill's approach. Indeed, given (†), I sympathize with almost everything in Jill's paper.[5]

REFERENCES

[1] Farquhar, Robert. 1964. *Ergodic Theory in Statistical Mechanics*. London: Interscience.

[2] E. Ercolessi, G. Morandi and G. Marmo. 2002. Alternative Hamiltonian Descriptions and Statistical Mechanics. *International Journal of Modern Physics A* 17(26): 3779-3788.

[3] Sklar, Lawrence. 1993. *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. Cambridge: CUP.

---

[3]Here, I am indebted to Steve Leeds for helpful discussions about the theoretical role of (†) in this context.

[4]We could go *non*-Lebesgue for our underlying volume measure. *E.g.*, we could use a non-standard *measure theory*, and thereby ensure conservation of phase volume [(a)] in *Newtonian* space $\mathcal{N}$ [1, p. 58]. But, l will bracket this "non-standard measure theory" option for 3 reasons: (i) it is beyond the scope of this commentary, (ii) it is *weird*, and (iii) while it *can* allow $\mathcal{N}$ to satisfy (a) & (b), I *think* it *can't also* ensure (c), and so it doesn't ultimately "save" $\mathcal{N}$ anyway. As for $\mathcal{H}'$, it can satisfy *either* (a) & (b) *or* (a) & (c), but *not all three* (and I *think* going non-Lebesgue won't "save" $\mathcal{H}'$).

[5]*Modulo* a couple of minor qualms about Jill's remarks on subjective probability and indifference (which don't bear on her central arguments), some puzzlement over her claim that there can be non-extreme *physical* probabilities in deterministic worlds, and a nagging feeling that we still need *some* explanation of *why* the $\mathcal{H} + \rho$ combination works.