International Journal of Theoretical Physics, Vol. 22, No. 9, 1983

Reformulating Classical and Quantum Mechanics in Terms of a Unified Set of Consistency Conditions

Robert F. Bordley[1]

University of Michigan, Dearborn, Dearborn, Michigan

Received January 19, 1982

This paper imposes consistency conditions on the path of a particle and shows that they imply Hamilton's principle in classical contexts and Schrödinger's equation in quantum mechanical contexts. Thus this paper provides a common, intuitive foundation for classical and quantum mechanics. It also provides a very new perspective on quantum mechanics.

INTRODUCTION

It is natural to expect that a particle's choice of trajectory under various conditions is consistent. This paper will formulate the consistency (or rationality) condition in such a way that I can apply recent results in mathematical measurement theory to prove that a particle will choose that trajectory maximizing the expected value of some function.[2]

When I choose that function to be the potential energy minus the kinetic energy and add a single probabilistic constraint on the particle's motion, I can derive the Schrödinger equation. By neglecting this constraint (which corresponds to saying that Planck's constant is negligible), I can

[1]Present address: General Motors Research Laboratories, Warren, Michigan 48090.

[2]Thus we end up with a variational principle. Variational principles are plentiful in the literature. For an attempt to use variational principles as a foundation for physical theory, see Schwinger (1951, 1953, 1953, 1953, 1954, 1954). One consequence of this paper is that there is now an axiomatic basis for variational principles.

0020-7748/83/0900-0803$03.00/0

derive Hamilton's principle. Thus the principle of maximizing the expected value of the potential energy minus the kinetic energy subject to a certain probabilistic constraint leads to classical and quantum mechanics.

Economic theory is similarly based on the idea that an individual's behavior is rational and hence that he acts as if he maximized the expected value of some function (called his utility function). Thus the results of this paper show that there is a common foundation unifying physics and economics. This is hardly surprising inasmuch as both are rational attempts to understand behavior—in the case of physics, the behavior of nature; in the case of economics, the behavior of individuals.

The measurement theory axioms I use concern the behavior of an individual (entity or person) acting under uncertainty. In the special case of certainty, we get a simpler set of axioms. Appendix 3 shows that we can use this simpler set of axioms to derive Hamilton's principle (although not the Schrödinger equation).

The first section of the paper formulates the behavior of the particle in terms of trajectories and applies the conditions of consistent behavior under uncertainty. These conditions, presented in Appendix 1, are adapted from Savage's *Foundation of Statistics*. I then state the result that the particle will follow that trajectory or that probability density function over trajectories maximizing the expected value of some function. That is Proposition 1.

In the second section, I state a time-separability condition which says that specifying the position and momentum at a time t completely specifies the state of the particle. This simplifies the form of the expected utility which the particle maximizes.

I now specify the constraints on that expected utility function. The utility function is defined to be the negative of the Lagrangean. We also constrain the trajectories the particle selects by a condition akin to Heisenberg's uncertainty principle. With these conditions, the calculus of variations allows us to deduce the Schrödinger equation.

When we neglect the uncertainty condition, the maximization of expected utility becomes a maximization of utility which is identical to Hamilton's principle. Thus the derivation is consistent with standard correspondence conditions.

The fact that this derivation consists mainly of rationality axioms is intuitively very satisfying. Since scientific research often implicitly assumes some rationality in natural behavior, making these assumptions explicit and then deducing all the implications of those previously implicit rationality assumptions is theoretically parsimonious. We then add a few physical assumptions to get our final theory.

The first section begins with a *Gedanken experiment*.

2. THE RATIONALITY CONDITIONS

2.1. Trajectories. Let $\mathcal{A} = (a_1, a_2, \ldots)$ be the set of all logically possible trajectories which the particle could follow where a trajectory, a_i, is a complete specification of the position, $q_i(t)$, and the momentum, $p_i(t)$, of the particle at each time t in the interval $[T_0, T_1]$. More formally, $a_i = ((p_i(t), q_i(t)) | t \in [T_0, T_1])$.

2.2. Acts of Nature. We define an act of nature f_i to be a specification of the probability that the particle will take any one of the possible trajectories in T. Thus an act of nature might consist of (1) moving at a speed of 1 m/sec to the right with probability 0.50, (2) moving at a speed of 1 m/sec to the left with probability 0.50. In game theory, these have also been called randomized strategies. Let $F = (f_1, f_2, \ldots)$ be the possible acts of nature.

Let $P_f(a_i)$ be the probability that act of nature, f, leads to trajectory a_i. Then it is clear that specifying an act of nature, f, is equivalent to specifying a probability density function over all trajectories ($P_f(a_i)$ for all $a_i \in \mathcal{A}$) or specifying a probability density function over all possible positions, $q(t)$, and momenta, $p(t)$, for all times t in the interval $[T_0, T_1]$ (designated by $P_f(q, p, t)$ for all t in $[T_0, T_1]$).

2.3. The Gedanken Experiment. Consider the following thought experiment:

We have a set of acts of nature, F. We ask which act of nature will occur given the circumstances of the situation. Call it f_1. Now suppose we add an ideal constraint to the problem which rules out act of nature f_1 without affecting the viability of any other act of nature. Then what act of nature will occur? Call it f_2. We similarly rule out act of nature f_2 with an ideal constraint. What now shall be the act of nature? Call it f_3. We define an ordering, $>$, and say that $f_1 > f_2$, $f_2 > f_3$, and $f_1 > f_3$. More generally we define $f_i > f_j$ if, in this thought experiment, we will never pick f_j to be the act of nature which will occur as long as act of nature f_i has not been ruled out by an ideal constraint. We then define the ordering, $\geqslant$, by saying that $f_i \geqslant f_j$ if it is false that $f_j > f_i$.

We could metaphorically interpret $f_i > f_j$ as meaning that "nature prefers act of nature f_i to act of nature f_j." And we could interpret $f_i \geqslant f_j$ as saying that "nature considers act of nature f_i at least as preferable as act of nature f_j." However, the precise meaning of "$\geqslant$" is given by the *Gedanken Experiment*.

Let $(F, \geqslant)$ denote the elements of F, every pair of which has been ordered by "$\geqslant$." We now impose certain consistency conditions which

Savage postulated for a rational individual's preference among alternative actions. These conditions will describe a rational nature's "preference" among alternative actions.

The arguments are presented in Appendix 1. They require seven consistency assumptions: (A1), (A2),... (A7). Using them, we get the following result:

Proposition 1. The particle will be governed by that probability distribution, $P_f(a)$ maximizing: $\int_{\mathscr{A}} u(a)P_f(a)da$ for some function $u(a)$.

The next section presents a time-separability assumption enabling us to rewrite Proposition 1 in terms of a time integral. We then make certain assumptions about $u(a)$ and $P_f(a)$ which will give us Hamilton's principle and the Schrödinger equation.

3. HAMILTON'S PRINCIPLE

3.1. The Time-Separability Property. Section 2 showed that if $(F, \geqslant)$ satisfies consistency conditions, then the probability density function, $P_f(a)$, describing the particle's motion maximizes:

$$\int_{\mathscr{A}} P_f(a)u(a)da$$

Now the trajectory, a, is $((q(t), p(t))|t \in [T_0, T_1])$, a specification of the position and momentum of the particle at every time in the time interval $[T_0, T_1]$. We are now about to define a new property which describes how $\geqslant$ relates to time. First we need some definitions.

Definition. Let

$$a(t) = ((q(t), p(t))|t \in [T_0, T_1])$$

$$a_1(t) = ((q_1(t), p_1(t))|t \in [T_0, T_1])$$

$$a_2(t) = ((q_2(t), p_2(t))|t \in [T_0, T_1])$$

be three trajectories, let A be a subinterval of $[T_0, T_1]$, and let A^c be the complement of A in $[T_0, T_1]$. We say that $a = (a_1(t), t \in A;\ a_2(t), t \in A^c)$ iff

(1) $q(t) = q_1(t)$ and $p(t) = p_1(t)$ for $t \in A \cap [T_0, T_1]$

(2) $q(t) = q_2(t)$ and $p(t) = p_2(t)$ for $t \in A^c \cap [T_0, T_1]$

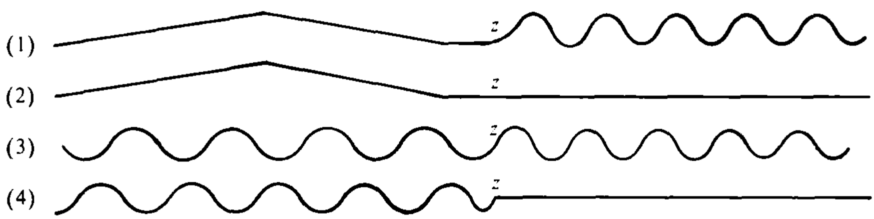
Definition. $(F, \geqslant)$ has time separability iff for any interval A and for any trajectories $a_1(t)$, $a_2(t)$, $a_3(t)$, we have that whenever

$$(a_1(t), t \in A;\ a_2(t), t \in A^c) \geqslant (a_3(t), t \in A;\ a_2(t), t \in A^c)$$

then for any other trajectory $a_4(t)$

$$(a_1(t), t \in A;\ a_4(t), t \in A^c) \geqslant (a_3(t), t \in A;\ a_4(t), t \in A^c)$$

3.2. Example of Time Separability. Consider the following four trajectories:



We assume that at point z, all particles have the same position and momentum. Then if $(F, \geqslant)$ has time separability and trajectory 1 $\geqslant$ trajectory 2 then trajectory 3 $\geqslant$ trajectory 4. In other words, given that all particles have the same position and momentum at point z, the past trajectory they followed becomes irrelevant in determining the future trajectory after point z.

We now state our next assumption:

(A8) $(F, \geqslant)$ has time separability.

This leads us to the following proposition:

Proposition 2. Suppose that $(F, \geqslant)$ satisfies (A8). Then there exist functions $P_f(a(t), t)$, $u(a(t), t)$ such that

$$\int_{\mathcal{A}} P_f(a)\, u(a)\, da = \int_{T_1}^{T_2} \int_{\mathcal{A}} P_f(a(t), t)\, u(a(t), t)\, da(t)\, dt$$

For a sketch of a proof, see Koopmans (1960) and Debreu (1960).

We recall that $a(t)$ is just $(p(t), q(t))$, the momentum and position at time t. Hence we have the following:

Proposition 3. Suppose that $(F, \geqslant)$ satisfies (A8). Then the particle's probability density at time t, $P_f(q, p, t)$ maximizes:

$$\int_{T_0}^{T_1} \int_p \int_q P_f(q, p, t) u(q, p, t)\, dq\, dp\, dt$$

subject to any constraints.

We now make the last assumption needed to get Hamilton's principle:

(A9) $u(q, p, t)$, which could be metaphorically called the particle's utility function, equals the potential energy, $V(q, p, t)$, minus the kinetic energy, $T(q, p, t)$, i.e. $u(q, p, t) = V(q, p, t) - T(q, p, t)$.

This leads us to the following:

Proposition 4. Suppose that $(F, \geqslant)$ satisfies (A8) and (A9). Then the particle's probability density function, $P_f(q, p, t)$, maximizes:

$$\int_{T_0}^{T_1} \int_p \int_q P_f(q, p, t)(V(q, p, t) - T(q, p, t))\, dq\, dp\, dt$$

subject to any constraints.

There are two types of constraints: constraints on the form of the probability density function and all other constraints. An example of a constraint on the form of the probability density function would be the constraint: "the variance of x must exceed Planck's constant." If we assume all such constraints on the form of the probability density function can be neglected, then the solution to the problem of Proposition 4 is a δ function with all mass concentrated on that trajectory $((q(t), p(t)) \mid t \in [T_0, T_1])$ which maximizes:

$$\int_{T_0}^{T_1} [V(q, p, t) - T(q, p, t)]\, dt$$

subject to any constraints. More formally:

Proposition 5. If there are no constraints on the form of the probability density function and if $(F, \geqslant)$ satisfies (A8) and (A9), then the particle's trajectory $(q(t), p(t) \mid t \in [T_0, T_1])$ is deterministic and maximizes:

$$\int_{T_0}^{T_1} [V(q, p, t) - T(q, \dot{p}, t)]\, dt$$

subject to any constraints.

And Proposition 5 is obviously Hamilton's principle which, as Goldstein (1950) has emphasized, can be viewed as an alternate formulation of the laws of classical mechanics more fundamental than Newton's.

In the following section, we will add a constraint on the form of the probability density function which will give us Schrödinger's equation.

4. THE SCHRÖDINGER EQUATION

4.1. Derivation. Define a function $\phi(q, p, t)$ so that $P_f(q, p, t) = \phi^2(q, p, t)$. This function $\phi(q, p, t)$ will be discussed in detail further.

The first constraint on the form of the probability function which I want to add is the following:

(A10) $p\phi(q, p, t) = -i\hbar[\partial\phi(q, p, t)/\partial q]$.

Hence the optimization problem is as follows:

Proposition 6. The particle will have that $\phi(q, p, t)$ maximizing:

$$\int_{T_0}^{T_1}\int_p\int_q \phi^2(q, p, t)[V(q, p, t) - T(q, p, t)]\, dq\, dp\, dt$$

subject to

$$\text{(4a)} \qquad \int_p\int_q \phi^2(q, p, t) = 1 \qquad \text{for } T_0 \leqslant t \leqslant T_1$$

$$\text{(4b)} \qquad p\phi(q, p, t) = -i\hbar\frac{\partial\phi(q, p, t)}{\partial q}$$

We can now derive the Schrödinger equation.

Assumption. $T(q, p, t) = p^2/2m$ and $V(q, p, t) = V(q, t)$, i.e., the kinetic energy is purely velocity dependent and the potential energy is velocity independent.

Then

$$\begin{aligned} T(q, p, t)\phi^2(q, p, t) &= \frac{1}{2m}[p\phi(q, p, t)]^2 \\ &= \frac{1}{2m}\left[-i\hbar\frac{\partial\phi(q, p, t)}{\partial q}\right]^2 \\ &= -\frac{1}{2m}\hbar^2\left[\frac{\partial\phi(q, p, t)}{\partial q}\right]^2 \end{aligned}$$

Since $V(q, p, t)$ is independent of p, we have gotten rid of constraint (4b) by substitution. Thus we now have the following:

Proposition 7. The particle will have that $\phi(q, p, t)$ maximizing:

$$\int_{T_0}^{T_1} \int_p \int_q \frac{\hbar^2}{2m} \left[\frac{\partial \phi(q, p, t)}{\partial q} \right]^2 + V(q, p, t) \phi^2(q, p, t)\, dp\, dq\, dt$$

subject to

$$\int_p \int_q \phi^2(q, p, t) = 1 \qquad \text{for } T_0 \leqslant t \leqslant T_1$$

Letting $\lambda(t)$ be the Lagrangian multiplier for the constraint (for each $t \in [T_0, T_1]$), we now use the calculus of variations:
Define

$$G(q, p, t) = \left\{ \frac{\hbar^2}{2m} \left[\frac{\partial \phi(q, p, t)}{\partial q} \right]^2 + V(q, p, t) \phi^2(q, p, t) - \lambda(t) \phi^2(q, p, t) \right\}$$

We know that the optimal $\phi(q, p, t)$ satisfies

$$\frac{\partial G(q, p, t)}{\partial \phi} - \frac{\partial}{\partial q} \left[\frac{\partial G(q, p, t)}{\partial \phi_q} \right] = 0$$

where

$$\phi_q = \frac{\partial \phi(q, p, t)}{\partial q}$$

This gives us the equation

$$(*) \quad -\frac{\hbar^2}{2m} \frac{\partial^2 \phi(q, p, t)}{\partial q^2} + V(q, t) \phi(q, p, t) = \lambda(t) \phi(q, p, t)$$

Now suppose that we consider $\phi(q, p, t)$ to be time independent. Then the normalization constraint will hold for all $t \in [T_0, T_1]$ if it holds for any $t \in [T_0, T_1]$. In this case, $\lambda(t)$ is a constant for all $t \in [T_0, T_1]$. If we call it E, then (*) becomes exactly the time-independent Schrödinger wave equation.

Suppose that $\phi(q, p, t)$ is not time independent. We then make the following further assumption on the form of $\phi(q, p, t)$:

(A11) (a) $\phi(q, p, t) = Y(q, p) Z(t)$.

Part (a) of the assumption says that $\phi(q, p, t)$ is time separable. Then

$$\frac{\partial \phi(q, p, t)}{\partial t} = Y(q, p) \frac{\partial Z(t)}{\partial t} = Y(q, p) Z(t) \left[\frac{1}{Z(t)} \frac{\partial Z(t)}{\partial t} \right]$$

so that

$$i\hbar\frac{\partial\phi(q,p,t)}{\partial t}=\phi(q,p,t)\left[+i\hbar\frac{1}{Z(t)}\,\frac{\partial Z(t)}{\partial t}\right]$$

The second part of the assumption is then
(b)

$$\lambda(t)=+i\hbar\frac{1}{Z(t)}\,\frac{\partial Z(t)}{\partial t}$$

Together (A11a) and (A11b) tell us that

$$(**)\quad \lambda(t)\phi(q,p,t)=i\hbar\frac{\partial\phi(q,p,t)}{\partial t}$$

(**) when substituted into (*) gives us the time-dependent Schrödinger wave equation. We note that (A10) and (**) are analogous to standard operator-notation expressions in quantum mechanics.

4.2. Implications of (A10) and (A11). The equation $p\phi(q,p,t)=-i\hbar[\partial\phi(q,p,t)/\partial q]$ means that $\phi(q,p,t)=X(p,t)e^{ipq/\hbar}$. Then the customary wave function, $\underline{\psi}(q,t)$ is just

$$\underline{\psi}(q,t)=\int_p\phi(q,p,t)$$

$$=\int_p X(p,t)e^{ipq/\hbar}$$

As Merzbacher (1970, p. 19) noted, this is just the standard expression of the wave function as a wave packet.

We also have the customary momenta wave function, $\varphi(p,t)$ as just

$$\varphi(p,t)=\int_q\phi(q,p,t)$$

$$=\int_q X(p,t)e^{ipq/\hbar}$$

We note that both $\underline{\psi}(p,t)$ and $\varphi(q,t)$ are real functions.

Now my analysis has not made use of complex conjugation. Hence

$$P_f(q,p,t)=A^2(p,t)e^{2ipq/\hbar}$$

is, in general, a complex number although

$$P_1(q,t)=\int_p A^2(p,t)e^{2ipq/\hbar}$$

and

$$P_2(p,t)=\int_q A^2(p,t)e^{2ipq/\hbar}$$

are real probability density functions. I will discuss the implications of this later.

Similarly assuming $\phi(q,p,t)=Y(q,p)Z(t)$ leads [as Merzbacher (1970), p. 43 and 44) demonstrates] to

$$\begin{aligned}\phi(q,p,t)&=e^{-iE(t)/\hbar}Y(q,p)\\&=e^{i(pq-E(t))/\hbar}A(p)\end{aligned}$$

Note that $E(t)=Et$ when $\lambda(t)$ is a constant. Thus we can write Proposition 6 as follows:

Proposition 8. Choose $A(p)$ to maximize:

$$\int_{T_0}^{T_1}\int_p\int_q A^2(p)e^{2i(pq-E(t))/\hbar}(V(q,p,t)-T(q,p,t))\,dq\,dp\,dt$$

subject to

$$\int_p\int_q A^2(p)e^{2i(pq-E(t))/\hbar}dp\,dq=1 \qquad \text{for } T_0\leqslant t\leqslant T_1$$

We cannot be completely comfortable with this proposition, as stated, because of the complex numbers. However, we use the fact that $V(q,p,t)=V(q,t)$ and $T(q,p,t)=T(p,t)$ to write

$$\int_{T_0}^{T_1}\int_p\int_q A^2(p)e^{2i(pq-E(t))/\hbar}(V(q,t)-T(p,t))\,dq\,dp\,dt$$

as equal to

$$\begin{aligned}&\int_{T_0}^{T_1}\int_q V(q,t)\int_p A^2(p)e^{2i(pq-E(t))/\hbar}dq\,dp\,dt\\&\quad-\int_{T_0}^{T_1}\int_p T(p,t)\int_q A^2(p)e^{2i(pq-E(t))/\hbar}dq\,dp\,dt\end{aligned}$$

Using the definitions of $P_1(q,t)$ and $P_2(p,t)$, we then can rewrite the proposition as follows:

Proposition 9. Choose $A(p)$ to maximize:

$$\int_{T_0}^{T_1}\int_q V(q,t)P_1(q,t)\,dq\,dt - \int_{T_0}^{T_1}\int_p T(p,t)P_2(p,t)\,dp\,dt$$

subject to

$$\int_q P_1(q,t) = 1 \qquad \text{for } T_0 \leqslant t \leqslant T_1$$

[which also implies $\int_p P_2(p\,t) = 1$].

And this is a well-defined problem.

This paper has viewed the particle as selecting that position and momentum which minimizes the expected value of the Lagrangean. However, we also added the assumption that it is impossible for the particle to select some particular combination of position and momentum as its solution (i.e., the probability distribution for any combination of position and momentum is a complex number). Given these constraints, we find that the probability the particle follows a given path is described by the Schrödinger equation.

5. CONCLUSIONS

The guiding idea in this paper is that physics and economics can be viewed as having a common theoretical foundation. This theoretical foundation can be formalized in certain consistency conditions (see Appendix 1) which lead to the maximization of expected utility.

At this point, one must specify the utility function being considered and the range of alternatives which the entity in question can use to maximize that expected utility. In physics, we add a time-separability condition (A8) which allows us to write our integral over expected utility as the time integral of an integral over position and momentum. We add a condition specifying that the utility function is the negative of the Lagrangean. Finally we allow a particle to choose any position and momentum combination to maximize expected utility subject to a certain uncertainty principle—the particle cannot specify any deterministic solution, it must specify a probability distribution with a certain minimal variance.

Given these specifications, we can derive the basic principles of classical and quantum mechanics.

The situation in economics is similar. In economics, we are concerned with an individual choosing decisions d to maximize his expected utility. The utility function expresses his empirical preferences. He is only allowed to make those decisions which are physically possible for him.

This parallelism indicates the possibility of much fruitful cross-fertilization between economics and physics.

APPENDIX 1: THE CONSISTENCY CONDITIONS

We have defined the ordering, $\geqslant$, over the acts of nature, $f \in F$. Before applying the Savage axioms, we first define the notion of a state. After deducing Savage's result which is expressed in terms of states, we then convert the result into statements about acts of nature and probability density functions, $P_f(a)$.

States. For illustrative purposes, consider both $\mathcal{Q}$ and F to be finite. Let Card($\mathcal{Q}$) and Card(F) denote the number of elements in $\mathcal{Q}$ and F, respectively. (The results apply to the infinite $\mathcal{Q}$ and F case.) We define the state $s(i_1, i_2, \ldots, i_{\mathrm{Card}(F)})$ to be the following vector:

$$s(i_1, i_2, \ldots, i_{\mathrm{Card}(F)}) = (a_{i_1}, a_{i_2}, \ldots, a_{i_{\mathrm{Card}(F)}})$$

We say that state $s(i_1, i_2, \ldots, i_{\mathrm{Card}(F)})$ obtains if and only if

(1) Act f_1, if it were the true act of nature, would have led to trajectory a_{i_1}

(2) Act f_2, if it were the true act of nature, would have led to trajectory a_{i_2}

(3) Act f_3, if it were the true act of nature, would have led to trajectory a_{i_3}

$\vdots \quad \vdots \quad \vdots$

(Card(F)) Act $f_{\mathrm{Card}(F)}$, if it were the true act of nature, would have led to trajectory $a_{i_{\mathrm{Card}(F)}}$

Thus a state is an element of the Card(F)-fold Cartesian product of $\mathcal{Q}$. We use the notation, s, to denote a state. Let S be the set of all such states.

(S has $\text{Card}(\mathcal{A})^{\text{Card}(F)}$ elements.) The probability of a state, s, is the probability of all Card(F) acts leading to the trajectories specified for them by the state. Thus the probability of $s(i_1, i_2, \ldots, i_{\text{Card}(F)})$ is

$$P\big(s(i_1, i_2, \ldots, i_{\text{Card}(F)})\big) = \prod_{k=1}^{\text{Card}(F)} P_{f_k}(a_{i_k})$$

Let S^+ be the set of all states with positive probability. (See Appendix 2 for further discussion of this notion of state.)

With this formulation, we can view the acts of nature, f, as functions mapping each state, s, into a trajectory, a. More formally, we could write $f(s) = a$.

We can now begin stating the Savage axioms.

The Savage Axioms:

(A1) (F, $\geqslant$) is a simple order. In other words, for any f_1, f_2 and f_3 in F: (a) either $f_1 \geqslant f_2$ or $f_2 \geqslant f_1$ but not both; (b) if $f_1 \geqslant f_2$ and $f_2 \geqslant f_3$ then $f_1 \geqslant f_3$. This is just a matter of logic. In the thought experiment, we will either choose f_1 to be true before we choose f_2 or vice versa. This is (a). If we will choose f_1 to be true before we choose f_2 and if we will choose f_2 to be true before we choose f_3, then we will choose f_1 to be true before we choose f_3. This is (b).

We now make a definition:

Definition. We say that for $f, g \in F$, $f \geqslant g$ given B if and only if $f' \geqslant g'$ for every f', g' such that

$$\begin{aligned} &\text{(a)} \quad f'(s) = f(s) && \text{for } s \in B \\ &\text{(b)} \quad g'(s) = g(s) && \text{for } s \in B \end{aligned}$$

and

$$\text{(c)} \quad f'(s) = g'(s) \qquad \text{for } s \notin B$$

And then our second postulate is as follows:

(A2) For every $f, g \in F$ and every $B \subset S$, either $f \geqslant g$ given B or $g \geqslant f$ given B.

Now this postulate, often called Savage's sure thing principle, says that if we are trying to decide whether $f \geqslant g$ or not, and if f and g lead to the same trajectories for states not in B, then to decide whether $f \geqslant g$ or not, we can ignore what trajectories f and g led to for states not in B.

We now make another definition.

Definition. Suppose that we have acts of nature, f and g, such that $f(s) = a_f$ and $g(s) = a_g$ for every $s \in S^+$. Then $a_f \geqslant a_g$ iff $f \geqslant g$ and $a_f > a_g$ iff $f > g$.

This definition extends the $\geqslant$ ordering of acts of nature to an ordering of trajectories. We now state our third postulate:

(A3) If $f(s) = a_f$ and $f'(s) = a_g$ for $s \in B$ and if $B \subset S^+ \neq \varnothing$ then $f \geqslant f'$ given B iff $a_f \geqslant a_g$.
In other words, suppose one trajectory is preferred to another (in the sense of $\geqslant$). Then if one act is the same as another act except for states in B and if that one act leads to the more preferred trajectory while the other leads to the less preferred trajectory for states in B, then the first act is preferred to the second.

This postulate again is fairly reasonable. We now make another definition:

Definition. $A \geqslant B$ iff $f_A \geqslant f_B$ whenever $g \geqslant g'$, $f_A(s) = g$ for $s \notin A$ and $f_A(s) = g'$ for $s \in A$, and $f_B(s) = g$ for $s \in B$ and $f_B(s) = g'$ for $s \notin B$.

In other words, A is more probable than B if that act which gives the more preferred trajectory in A (and the less preferred in A^c) is preferred to that act which gives the more preferred trajectory in B (and the less preferred in B^c).

Savage's next postulate is then as follows:

(A4) For every A, B (subsets of S), either $A \geqslant B$ or $B \geqslant A$.
The probability $P(s(i_1, i_2, \ldots, i_{\text{Card}(F)}))$ which I defined earlier will satisfy the definition and axiom (A4).

The next postulate makes the rather innocuous assertion that there is at least one act which is better than some other act.

(A5) It is false that for every f, f', we have $f \geqslant f'$.
This is clearly true in our case since even without the *Gedanken Experiment*, the particle does follow one trajectory over all others. Hence there is at least one act of nature which is better than others.

The next postulate states the following:

(A6) Suppose that $g > h$. Then for every a_f, there is a finite partition of S such that if g' gives the same trajectories as g and if h' gives the same trajectories as h except on an arbitrary element of the partition (each element of the partition having equal probability) at which both g' and h' lead to a_f, then $g' > h$ and $g > h'$.

This postulate says that no matter how preferred or not preferred trajectory a_i is, if it is assigned a small enough probability and made one possible consequence of g (thus turning g into g') and made one possible consequence of h (thus turning h into h'), then g will only be negligibly different from g' and h will only be negligibly different from h'. We could view this axiom as saying that there is no infinitely preferable trajectory and that we can partition S as finely as desired. (Partitioning S can be thought of as taking subsets of S, possibly flipping coins to subdivide into even smaller sets.)

The final axiom is as follows:

(A7) If $f \geqslant g$ given s for each $s \in B$, then $f \geqslant g$ given B.

These are Savage's seven axioms of rational behavior. Given the basic formulation of an ordering of acts of nature, $\geqslant$, they are fairly plausible. Using them, we can deduce Savage's theorem:

Theorem 1. If $(F, \geqslant)$ satisfies (A1) through (A7), then there exist functions $p(s)$ and $u(f/s)$ such that

$$f \geqslant g \text{ iff } \sum_{s \in S} p(s) u(f/s)\, ds \geqslant \sum_{s \in S} p(s) u(g/s)\, ds$$

where $p(s)$ might be viewed as the probability of state s obtaining and $u(f/s)$ is called the utility of act f in state s (see Savage, 1975, for proof).

But the utility of act f in state s is the utility of $f(s)$, the trajectory act f leads to in state s. Suppose that $f_1(s) = a_1, f_2(s) = a_2, \ldots, f_{\mathrm{Card}(F)}(s) = a_{\mathrm{Card}(F)}$. Then by the definition of state s, $p(s)$ is the probability that act f_1 leads to trajectory a_1 and act f_2 leads to a_2 and act f_3 leads to a_3 and $\cdots$ and act $f_{\mathrm{Card}(F)}$ leads to $a_{\mathrm{Card}(F)}$.

Then we can change the variable of summation and write

$$\sum_{s \in S} p(s) u(f/s) = \sum_{a \in \mathcal{A}} P_f(a) u(a)$$

where $P_f(a)$ is the probability that act f leads to trajectory a and $u(a)$ is the utility of trajectory a.

I can now state my second Theorem:

Theorem 2. If $(F, \geqslant)$ satisfies (A1) through (A7), then there exist functions $P_f(a)$ and $u(a)$ such that

$$f \geqslant g \qquad \text{iff } \sum_{a \in \mathcal{A}} P_f(a) u(a) \geqslant \sum_{a \in \mathcal{A}} P_g(a) u(a)$$

And from that, we get the following corollary:

Corollary 1. If $(F, \geqslant)$ satisfies (A1) through (A7), then the particle will follow that act of nature, f, maximizing:

$$\sum_{a \in \mathcal{A}} P_f(a) u(a)$$

We note that choosing f is equivalent to choosing the probability distribution, $P_f(a)$ over all trajectories a in $\mathcal{A}$. We also note that we can let $\mathcal{A}$ become an uncountable set so that the summation is replaced by an integral. With these constraints, we have the following corollary:

Corollary 2. If $(F, \geqslant)$ satisfies (A1) through (A7), then the particle will be governed by that probability distribution, $P_f(a)$, maximizing:

$$\int_{\mathcal{A}} P_f(a) u(a)\, da$$

subject to any constraints on the particle.

Corollary 2 is Proposition 1 in the text.

APPENDIX 2

Suppose there are two possible trajectories: (1) standing still and (2) moving to the right at the speed of light. Suppose there are only three allowable probabilities: 0, 0.5, and 1.0. Then there are three possible acts of nature: act 1: standing still with probability 0; act 2: standing still with probability 0.5; act 3: standing still with probability 1.0.

There are 2^3 or eight possible states:

(1) The state in which act 1 leads to the particle standing still
act 2 leads to the particle standing still
act 3 leads to the particle standing still

(2) The state in which act 1 leads to the particle standing still
act 2 leads to the particle standing still
act 3 leads to the particle moving to the right

$\vdots \quad \vdots \quad \vdots$

(8) The state in which act 1 leads to the particle moving to the right
act 2 leads to the particle moving to the right
act 3 leads to the particle moving to the right

Most of these states have probability zero. In fact, there are only two which have nonzero probability:

(1) The state in which act 1 leads to the particle moving to the right
act 2 leads to the particle standing still
act 3 leads to the particle standing still

and

(2) The state in which act 1 leads to the particle moving to the right
act 2 leads to the particle moving to the right
act 3 leads to the particle standing still.

Each of these two states has probability 0.5.

APPENDIX 3: DETERMINISTIC DERIVATION

This Appendix shows that it is possible to derive Hamilton's principle using only (A8), the time-separability assumption and (A9) which defines utility to be the negative Lagrangian and two other assumptions. We proceed as follows.

First suppose we are considering trajectories the particle could follow. We call the trajectory the particle will, in fact, follows a_1. Now suppose we rule out a_1 with an ideal constraint. Then we call the trajectory which it will follow instead a_2. And so we continue. We define the ordering, $>$, and say that $a_i > a_j$ if we would never choose a_j to be the trajectory as long as trajectory a_i has not been ruled out by an ideal constraint. We define $\geqslant$ by $a_i \geqslant a_j$ if it is false that $a_j > a_i$.

Thus we have constructed a deterministic $\geqslant$ ordering over trajectories similar to the one we constructed in the first section. We now assume the following:

(B.1) $(T, \geqslant)$ is a simple order.

This is analogous to axiom A.1 of Appendix 1.

We now define the following set of trajectories:

Definition. A trajectory belongs in set C if it is possible to describe the trajectory by a polynomial with rational coefficients and exponents with, at most, countably many terms.

We now make another definition:

Definition. A set B is said to be a countable order dense set of T if whenever $a_i > a_j$ for $a_i, a_j \in T$, there exists a $b \in B$ such that $a_i \geqslant b \geqslant a_j$. Furthermore B is countable.

We now make our second assumption:

(B.2) For any two trajectories $a_i, a_j \in T$, we can always find a trajectory, c in C such that if $a_i > a_j$, then $a_i \geqslant c \geqslant a_j$. In other words, in some sense, we can approximate any trajectory in T arbitrarily well by a trajectory in C.

Since C is countable, (B.2) says that T has a countable order dense set. We now cite the theorem from Krantz et al. (1971):

> *Theorem.* Suppose that $(T, \geqslant)$ is a simple order and $(T, \geqslant)$ has a countable order dense set, then there exists a function $u(\ )$ such that (1) if $a_i \geqslant a_j$, then $u(a_i) \geqslant u(a_j)$, (2) the particle will follow that trajectory maximizing $u(a^*)$ and conversely.

Thus axioms (B.1) and (B.2) imply the existence of a utility function for the particle. The time-separability assumption allows us to write that utility function as

$$u(a) = \int_{T_0}^{T_1} u(a(t), t)\, dt$$

Defining utility to equal potential energy minus kinetic energy then gives us Hamilton's principle. Thus Hamilton's principle can be derived from some very simple and very intuitive conditions.

REFERENCES

Debreu, G. (1960). Topological methods in cardinal group utility theory, in *Mathematical Models in the Social Sciences*. Stanford University Press, Stanford, California.

Dugas, R. (1955). *A History of Mechanics*. Central Book Company, New York.

Goldstein, H. (1950). *Classical Mechanics*. Addison-Wesley, Reading, Massachusetts.

Koopmans, T. C. (1960). Stationary ordinal utility and impatience, *Econometrica*, **28**, 2 (April 1960).

Krantz, D., Luce, R. D., Suppes, P., and Tversky, A. (1971). *The Foundations of Measurement*. Academic Press, New York.

Marion, J. (1970). *Classical Dynamics*. Academic Press, New York.

Marschak, J. (1950). Rational behavior, uncertain prospects and measurable utility, *Econometrica*, **18**, 111.

Merzbacher, E. (1970). *Quantum Mechanics*. John Wiley & Sons, New York.

Nelson, E. (1966). Derivation of the Schrödinger equation from Newtonian mechanics, *Physical Review* **28**, 1079.

Savage, L. J. (1975). *The Foundations of Statistics*. Dover Publications, New York.

Schwinger, J. (1951). The theory of quantized fields. I, *Physical Review*, **82**, 914, Part I; The theory of quantized fields. II, *Physical Review*, **91**, 713 (1953); The theory of quantized fields. III, *Physical Review*, **91**, 728 (1953); The theory of quantized fields. IV, *Physical Review*, **92**, 1283 (1953); The theory of quantized fields. V, *Physical Review*, **93**, 615 (1954); The theory of quantized fields. VI, *Physical Review*, **94**, 1362 (1954).

Tolman, R. C. (1948). *The Principles of Statistical Mechanics*. Oxford University Press, London.

Weinstock, R. (1970). *Calculus of Variations*. McGraw-Hill Book Company, New York.