In M. C. Galavotti, P. Suppes & D. Costantini, (Eds.),
Stochastic Causality. Stanford, CA: CSLI Publications. 203-220.

12

Weak and Strong Reversibility of Causal Processes*

PATRICK SUPPES
Stanford University

1 Introduction

It is widely known, and often commented upon, that the basic equations of classical physics remain valid under a transformation from time t to $-t$. It is often also remarked that this invariance under a change of direction of time is completely contrary to ordinary experience. So the invariance of well-established classical physics, as well as other parts of physics, for example, relativistic mechanics and the Schrödinger equation in quantum mechanics, creates a natural philosophical tension about the nature of causal processes. The purpose of the present article is not to resolve in any complete way this tension between physics and ordinary experience, but to at least reduce the tension by introducing two concepts of reversibility. The first is that of *weak reversibility,* which is what is exemplified in the invariance of the equations of classical physics under time reversal. This is the *weak sense* of reversibility, because we can usually distinguish by observation whether a system of particle mechanics is running one way or the other. The appearance is not the same when we go to a reversal of time. Elementary examples would be measuring the velocity of a particle accelerating from rest, moving in a straight

*I am indebted to Ubaldo Garibaldi for several useful comments on an earlier draft.

Stochastic Causality.
Maria Carla Galavotti, Patrick Suppes and Domenico Costantini (eds.).
Copyright © 2001, CSLI Publications.

line and hitting an impenetrable wall. The picture of change of velocity, especially, will be very different under time reversal. Under such reversal, the particle has a very large velocity immediately, as motion begins, and will continue to decelerate until it finally reaches the state of rest, quite contrary to what would be observed in the usual direction of time. Of course, this difference is well accepted in discussions of the invariance of classical physics under time reversal.

The concept of *strong reversibility* introduced here, and used in various analyses of a variety of natural phenomena, including stochastic processes, has a much stronger condition for reversibility. A process, deterministic or stochastic, is said to be strongly reversible if we are unable to distinguish whether the process is running forward or backward. To put it in vivid terms often used, the backward movie, that is, running the film in reverse, of the observation of the process, is indistinguishable by observation from the forward movie. In what follows, these ideas are made more precise, especially for stochastic processes. I should mention that the concept of strong reversibility used here corresponds to what is often defined in the stochastic-process literature as simply *reversible*. So the condition given below for a Markov chain to be strongly reversible is the same as the standard concept to be found, for example, in Feller (1950, p. 342) for Markov chains.

Although many of the formal distinctions introduced in what follows are familiar in the literature of probability theory, especially of stochastic processes, it is my impression that the distinction introduced here has not been sufficiently emphasized, as a distinction, in the philosophical discussions of time reversibility. Its introduction is meant to reduce the tension I mentioned at the beginning, in the sense that there is a close alignment between ordinary experience not being strongly reversible, and, similarly, for many trajectories of idealized particles or other processes of importance in physics.

2 Weak Reversibility

As already remarked, it is well known that classical particle mechanics and relativistic particle mechanics have the property of being weakly reversible, i.e., the transformation changing the direction of the time always carries systems of classical particle mechanics into systems of classical particle mechanics and, correspondingly, for the relativistic case. Detailed proofs of these results are to be found in McKinsey and Suppes (1953) for classical particle mechanics and in Rubin and Suppes (1954) for relativistic particle mechanics. Because of their technical character they are summarized in the Appendix.

The proof that first-order Markov chains are carried into first-order Markov chains under time reversal is straightforward. Here is the elementary proof for first-order chains with a finite number of states. In the proof it is assumed that all the probabilities that occur in the denominators of expressions have probability greater than zero.

THEOREM 1. *All first-order Markov chains are weakly reversible.*

Proof (for finite-state, discrete-time processes):

$$\begin{aligned} P(i_{n-1} \mid j_n k_{n+1}) &= \frac{P(i_{n-1} j_n k_{n+1})}{P(k_{n+1} \mid j_n)P(j_n)} \\ &= \frac{P(k_{n+1} \mid j_n)P(j_n \mid i_{n-1})P(i_{n-1})}{P(k_{n+1} \mid j_n)P(j_n)} \\ &= \frac{P(j_n i_{n-1})}{P(j_n)} \\ &= P(i_{n-1} \mid j_n). \end{aligned}$$

The situation in quantum mechanics is somewhat more complicated concerning weak reversibility. Certainly, the Schrödinger equation is so reversible. It is weakly reversible under transformations changing the direction of time. On the other hand, in the standard accounts, this is not true of the measurement process (von Neumann, 1955, Ch. 5).

Finally, from a common-sense standpoint, ordinary experience is certainly not weakly reversible. Time has a fixed direction and the empirical evidence in terms of human experience of this direction is overwhelming.

Forward $\neq$ Backward

People almost never walk up stairs backward. No races are run backward, etc. This is not to claim the "backward movies" violate laws of mechanics, only inductive laws and facts of experience, most especially the universally accepted nature of human memory:

unknown future $\neq$ known past.

Some definitions. Before going further, it is desirable to introduce some standard stochastic concepts, implicitly assumed in Theorem 1. First, let $\mathbf{X}(t)$ be a stochastic process such that for each time t, $\mathbf{X}(t)$ takes values in a finite set, a restriction imposed only for simplicity of exposition. Any finite family $\mathbf{X}(t_1), \mathbf{X}(t_2), \ldots, \mathbf{X}(t_n)$ has a joint probability distribution. Using an obvious simplification of notation, a process is *(first-order) Markov* if and only if for any times $t_1 < t_2, \ldots, t_n$,

(1) $$P(x_n \mid x_{n-1} \ldots, x_1) = P(x_n \mid x_{n-1}),$$

where the simplified notation is defined as follows:

$$P(x_n \mid x_{n-1}) = P(\mathbf{X}(t_n) = j_n \mid \mathbf{X}(t_{n-1}) = i_{n-1}).$$

A Markov process is *homogeneous* if the transition probabilities such as (1) do not depend on time, and it is *irreducible* if any state j can be reached from any other state in a finite number of steps. For a discrete-time homogeneous Markov process, I write the transition probabilities in several different, but useful notations

$$\begin{aligned} P_{ij} = p(i,j) &= P(j_n \mid i_{n-1}) = P(\mathbf{X}_n = j \mid \mathbf{X}_{n-1} = i) \\ &= P(\mathbf{X}(t_n) = j_n \mid \mathbf{X}(t_{n-1}) = i_{n-1}), \end{aligned}$$

all of these often used in the probability literature.

Let $P_{ii}(n)$ be the probability of returning to state i in n transitions. The *period* $r(i)$ of a state i is the greatest common divisor of the set of integers n for which $P_{ii}(n) > 0$. A state is *aperiodic* if it has period 1. Moreover, it is easy to show that if a Markov process is irreducible every state has the same period. So an aperiodic process is one in which every state has period 1.

Hereafter, when I refer to a Markov *chain*, I am assuming a Markov process that is finite-state, discrete-time, homogeneous, irreducible and aperiodic.

The definitions of being Markov, homogeneous and irreducible all apply without change to continuous-time processes, but there are some further conditions to impose to avoid physically unlikely cases. First, we require such a process to remain in any state for a positive length of time, and, second, the process cannot pass through an infinite sequence of states in a finite time. Corresponding to the transition probability for discrete-time processes, the *transition rate* for continuous-time processes is defined as:

$$q_{ij} = \lim_{\tau \to \infty} \frac{P(X(t+\tau) = j \mid X(t) = i)}{\tau}, \quad i \neq j,$$

and for definitional purposes, we set $q_{ii} = 0$. When I refer to a Markov *process*, I will mean a continuous-time one. Such a process remains in each state for a length of time exponentially distributed with the parameter $q(i)$ defined as:

$$q(i) = \sum_j q_{ij},$$

and when it leaves state i it moves to state j with the probability

$$p_{ij} = \frac{q_{ij}}{q_i}.$$

Note that in general

$$\sum_j q_{ij} \neq 1,$$

i.e., transition rates from i to another state need not add up to 1, as must be the case for transition probabilities from state i in the discrete-time processes. (A good discussion of such continuous-time processes may be found in Kelly (1979).)

3 Strong Reversibility

The intuitive idea of strong reversibility is easily characterized in terms of movies. A movie is strongly reversible if an observer cannot tell whether the movie is being run forward or backward. There is no perceptual difference in the two directions. Now, as our ordinary movie experience tells us at once, this is a pretty exceptional situation. On the other hand, there are important physical processes, a few of which we will discuss, that do have strong reversibility. An interesting case to begin with is this. When is a Markov chain strongly reversible? The right setting for this is to restrict ourselves at once to stationary, ergodic Markov chains. A Markov chain is *stationary* if its mean distribution is the same for all times. It is *ergodic* if it has a unique asymptotic distribution independent of the initial distribution. Note that under these definitions, a Markov chain can be ergodic but not stationary if its initial distribution is different from its unique, asymptotic one. But here I assume ergodic chains are stationary.

An ergodic Markov chain is *strongly reversible* if

(2) $$\pi_i p_{ij} = \pi_j p_{ji}$$

for all states i and j, where π is the unique asymptotic distribution. The equations (2) are often called the *detailed balance* conditions for strong reversibility. Bernoulli processes, such as coin flipping, are strongly reversible processes, because there is no memory of past states: $p_{ij} = p_j$ & $p_{ji} = p_i$, so we satisfy (2). Moreover, any 2-state ergodic Markov chain is strongly reversible. Here is the proof. Since the process is ergodic

$$\begin{aligned}\pi_1 &= \pi_1 p_{11} + \pi_2 p_{21} \\ &= \pi_1(1 - p_{12}) + \pi_2 p_{21},\end{aligned}$$

so

$$\pi_1 p_{12} = \pi_2 p_{21}.$$

On the other hand, many 3-state chains are not strongly reversible. Here is an example of a three-color spinner.

	B	R	Y
B	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
R	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{1}{8}$
Y	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$

We have at once for the mean distributions (π_b, π_r, π_y) of the three states:

$$\begin{aligned}\pi_b &= \tfrac{1}{2}\pi_b + \tfrac{3}{8}\pi_r + \tfrac{1}{8}(1 - \pi_b - \pi_r) \\ \pi_r &= \tfrac{1}{4}\pi_b + \tfrac{1}{2}\pi_r + \tfrac{3}{8}(1 - \pi_b - \pi_r);\end{aligned}$$

solving the two equations, we find:

$$\pi_b = \frac{13}{37}, \quad \pi_r = \frac{14}{37}, \quad \pi_y = \frac{10}{37}.$$

Here is the proof that it is not strongly reversible:

$$\pi_b p_{br} = \frac{13}{17} \cdot \frac{1}{4} \neq \frac{14}{37} \cdot \frac{3}{8} = \pi_r p_{rb}.$$

We can summarize what we have been saying about Markov chains in the following theorem, where in the notation of the theorem, $\overset{d}{=}$ means equal in distribution.

THEOREM 2. *If an ergodic Markov chain is strongly reversible, then it is impossible to distinguish whether a movie of it is running forward or backward, that is, any finite sequence of random variables of the process has the same distribution when the order of the random variables is reversed:*

$$(\mathbf{X}_0, \mathbf{X}_1, \ldots \mathbf{X}_n,) \overset{d}{=} (\mathbf{X}_n, \mathbf{X}_{n-1}, \ldots \mathbf{X}_0,)$$

I give the proof for $n = 2$:

$$\begin{aligned} P(\mathbf{X}_0 = i, \mathbf{X}_1 = j) &= P(\mathbf{X}_1 = j | \mathbf{X}_0 = i) P(\mathbf{X}_0 = i) \\ &= \pi_i p_{ij} \\ &= \pi_j p_{ji} \quad \text{by (2)} \\ &= P(\mathbf{X}_1 = i | \mathbf{X}_0 = j) P(\mathbf{X}_0 = j) \\ &= P(\mathbf{X}_1 = i, \mathbf{X}_0 = j). \end{aligned}$$

The concepts introduced hold for finite-state continuous-time changes by replacing transition probabilities by transition rates $q_{ij}, i \neq j$ with $q_{ij} \geq 0$, as defined above. The detailed balance conditions defining strong reversibility are the same in form as those of (2) with the $\pi_i \geq 0$ and summing to 1, as before:

$$(3) \qquad \pi_i q_{ij} = \pi_j q_{ji}.$$

With a proof just like that of Theorem 2, we have at once:

THEOREM 3. *If an ergodic continuous-time Markov process is strongly reversible, then it is impossible to distinguish whether a movie of it is running forward or backward.*

In the present context, there is nothing special about first-order Markov chains. A second-order ergodic Markov chain is strongly reversible iff

$$(4) \qquad \pi_i \pi_{ij} p_{ijk} = \pi_k \pi_{kj} p_{kji},$$

where a chain is second-order Markov if

$$P(x_n \mid x_{n-1}, \ldots, x_1) = P(x_n \mid x_{n-1}, x_{n-2}),$$

corresponding to (1) for first-order.

THEOREM 4. *If a second-order Markov chain is ergodic and strongly reversible, then it is impossible to distinguish a movie of it running forward or backward, i.e.,*

$$(\mathbf{X}_0, \mathbf{X}_1, \ldots \mathbf{X}_n,) \stackrel{d}{=} (\mathbf{X}_n, \mathbf{X}_{n-1}, \ldots \mathbf{X}_0,).$$

Proof for $n = 3$:

$$\begin{aligned}
P(\mathbf{X}_0 = i, \mathbf{X}_1 = j, \mathbf{X}_2 = k) \\
&= P(\mathbf{X}_2 = k|\mathbf{X}_1 = j, \mathbf{X}_0 = i)P(\mathbf{X}_1 = j|\mathbf{X}_0 = i)P(\mathbf{X}_0 = i) \\
&= p_{ijk}\pi_{ij}\pi_i \\
&= \pi_k\pi_{kj}p_{kji} \quad \text{by (4)} \\
&= P(\mathbf{X}_2 = i|\mathbf{X}_1 = j, \mathbf{X}_0 = k)P(\mathbf{X}_1 = j|\mathbf{X}_0 = k)P(\mathbf{X}_0 = k) \\
&= P(\mathbf{X}_2 = i, \mathbf{X}_1 = j, \mathbf{X}_0 = k).
\end{aligned}$$

As is obvious, this proof also works for continuous-time second-order Markov processes. Without too much effort, this same sort of proof can be extended to chains of infinite order that are ergodic and stationary. The line of reasoning required is developed in Lamperti and Suppes (1959).

Turning again from Markov chains to continuous-time Markov processes, simple examples of strongly reversible processes are ergodic *birth and death* processes, defined by the transition rates being zero except for $q(i, i+1) > 0$, representing a birth, and $q(i, i-1) > 0$ representing a death. The detailed balance conditions (3) then assume the form

(5) $$\pi_i q_{i,i+1} = \pi_{i+1} q_{i+1,i},$$

where the π_i's form the stationary probability distribution of the finite set of states.

Ehrenfest model. A simplified model of statistical mechanics nicely exemplifies a birth and death process that is strongly reversible. In the spirit of Maxwell's demon, there are two containers of particles, thought of as ideal gas molecules. Let $\mathbf{X}(t)$ be the number of particles in container I, and, so, $k - \mathbf{X}(t)$ is the number of particles in container II. The "birth and death" process corresponds to increasing or decreasing the number of particles in container I, with the transition rates being:

$$\begin{aligned}
q_{i,i-1} &= i\lambda, & i &= 1, 2, \ldots, k, \\
q_{i,i+1} &= (k-i)\lambda, & i &= 0, 1, \ldots, k-1,
\end{aligned}$$

where λ is the rate parameter. The stationary distribution may be shown to be

$$\pi_i = 2^{-k}\binom{k}{i} \qquad i = 0, 1, \ldots, k,$$

where

$$\binom{k}{i} = \frac{k!}{(k-i)!i!} \text{ and } 0! = 1.$$

That the process is strongly reversible may be seen by checking the detailed balance conditions (5).

$$\begin{aligned}
\pi_i q_{i,i+1} &= 2^{-k}\binom{k}{i}(k-i)\lambda \\
&= \frac{2^{-k}k!(k-i)\lambda}{(k-i)!i!} \\
&= \frac{2^{-k}k!\lambda}{(k-(i+1))!i!} && since \frac{k-i}{(k-i)!} = \frac{1}{(k-(i+1))!} \\
&= \frac{2^{-k}k!(i+1)\lambda}{(k-(i+1))!(i+1)!} && since \frac{1}{i!} = \frac{i+1}{(i+1)!} \\
&= 2^{-k}\binom{k}{i+1}(i+1)\lambda \\
&= \pi_{i+1} q_{i+1,i}.
\end{aligned}$$

Without attempting a detailed discussion of entropy for this process, I note that even though it is stationary and strongly reversible, the mean fluctuation from moment to moment is always stronger toward the equal distribution of particles in the two containers rather than away from this equal distribution, for

$$q_{i,i-1} = i\lambda < (k-i)\lambda = q_{i,i+1}$$

if and only if $i > \frac{k}{2}$, independent of the rate parameter λ. If we compute the entropy by the relative frequency $f(t)$ at time t of the number of particles in container I, then the "momentary" entropy of the process at t is

$$H(t) = -(f(t)\log f(t) + (1 - f(t))\log(1 - f(t))$$

and by the analysis just given the mean rate of change of entropy is always increasing, i.e., is positive, in spite of the strong reversibility of the process. But there is, obviously, no contradiction between these two features of the model, even at equilibrium, i.e., when stationary. An excellent detailed analysis of the Ehrenfest model and also of the Boltzmann equation, and of the associated "paradoxes" of Loschmidt and Zermelo, is to be found in Kac (1959).

Deterministic systems. I now turn to deterministic classical physical systems. It should be evident enough that only very special ones are strongly reversible. Any sort of dissipative system will, in general, not be. A real billiard ball on a real table is not strongly reversible. We start the motion and it comes to a halt somewhere, due to friction on the table. We can easily distinguish the forward from the backward picture. On the other hand, an idealized billiard ball on an idealized table with total conservation of energy and exact satisfaction of the equality of the angle of incidence and the angle of reflection can be put in periodic motion and, once in motion, will continue forever. Moreover, this idealized motion will, of course, be strongly reversible. By "observing" it, we cannot decide whether we are seeing the forward or the backward movie. (I put "observing" in quotes, for there is, in fact, no such billiard case, but we can artificially simulate it well for a finite period of time.)

A simple, but important, example of a classical system that has strong reversibility is an undamped and undriven one-dimensional harmonic oscillator. The differential equation for such an oscillator is:

$$\frac{d^2x}{dt^2} + \omega^2 x = 0,$$

where ω is the natural frequency of the oscillator when undamped and undriven, and the "initial" conditions at time $t = 0$ are:

$$\begin{aligned} x_0 &= \alpha, \\ \frac{dx_0}{dt} &= 0. \end{aligned}$$

(When the model is a pendulum, the initial conditions correspond to the pendulum being at rest at time $t = 0$ with displacement α.) The general solution of (1) is:

$$x(t) = A\cos\omega t + B\sin\omega t,$$

and

$$\frac{dx}{dt} = -A\omega\sin\omega t + B\omega\cos\omega t$$

so at $t = 0$, $B = 0$, whence

$$A\cos\omega t = \alpha \text{ at } t = 0, \text{ and } A = \alpha.$$

Then

$$\frac{d^2x}{dt} = -A\omega^2\cos\omega t$$

and finally

(6) $$x(t) = A\cos\omega t.$$

We have at once that such an oscillator is strongly reversible, since for all t

$$x(-t) = A\cos -\omega t = A\cos\omega t = x(t).$$

Notice that this solution holds for $-\infty < t < \infty$, and the conditions at $t = 0$ are not really initial conditions, but just conditions for some t that yield a solution simple in form.

On the other hand, the damped, undriven harmonic oscillator, whose differential equation is:

$$\frac{d^2x}{dt^2} + 2k\frac{dx}{dt} + \omega^2 x = 0,$$

where k is the damping coefficient, is easily shown not to be strongly reversible. This result is scarcely surprising, since a damped oscillator (in one dimension) is one of the simplest physical examples of a dissipative system. In the standard case where the damping is not too heavy, the oscillating function decreases in amplitude with time. More specifically the envelope of the oscillations is a negative exponential of the form e^{-kt} on the positive side and $-e^{-kt}$ on the negative side.

Without question, if the aim is to find examples of systems in classical physics that are not strongly reversible, and, therefore, the causal analysis is not strongly reversible, the place to look is among the many kinds of dissipative systems. The great prevalence of such systems in nature reinforces our personal experience to support the common-sense view that, as the saying goes, "... of course time is not reversible. Who could ever think otherwise?"

APPENDIX

A Classical Mechanics

A.1 Primitive Notions

The axiomatization of particle mechanics is based on five primitive notions: P, T, m, s, and f. P and T are sets, m is a unary function, s is a binary function, and f is a ternary function. The intended physical interpretation of P is as the set of particles. T is to be interpreted physically as a set of real numbers measuring elapsed times (in terms of some unit of time, and measured from some origin of time). If p is a member of P (that is to say, in the physical interpretation, if p is a particle), then $m(p)$ is to be interpreted physically as the numerical value of the *mass* of p. If p is in P, and t is in T, then $s(p,t)$ is an n-dimensional vector. For $n = 3$ (or for $n < 3$, if we are concerned with plane particle mechanics, or with one-dimensional particle mechanics) $s(p,t)$ is to be interpreted physically as a vector giving the *position* of p at time t. The primitive s fixes the choice of a coordinate system.

A.2 Axioms

Definition A1. *A structure $\mathcal{P} = (P, T, m, s, f)$ is an (n-dimensional) system of classical particle mechanics if and only if it satisfies the following axioms:*

Kinematical Axioms

Axiom P1. *P is a non-empty, finite set.*

Axiom P2. *T is an interval of real numbers.*

Axiom P3. *If p is in P and t is in T, then s(p,t) is an n-dimensional vector such that* $\frac{d^2}{dt^2}s(p,t)$ *exists.*

Dynamical Axioms

Axiom P4. *If p is in P, then m(p) is a positive real number.*

Axiom P5. *If p is in P and t is in T, then f(p,t,1),f(p,t,2), . . . f(p,t,i), . . . are n-dimensional vectors such that the series* $\sum_{i=1}^{\infty} f(p,t,i)$ *is absolutely convergent.*

Axiom P6. *If p is in P and t is in T, then*

$$m(p)\frac{d^2}{dt^2}s(p,t) = \sum_{i=1}^{\infty} f(p,t,i).$$

The omission of Newton's third law is deliberate in this general formulation. I specialize the axioms further below. But the level of generality of Definition A1 is useful for two purposes. First, a more general invariance theorem than is standard can be proved, as can be seen from the statement of the next theorem. Secondly, this general form of the axioms is most directly comparable to the axioms given below for relativistic particle mechanics, for which analogues of Newton's third law do not exist, because of the noninvariance of simultaneous distant events in the relativistic setting.

THEOREM A1. *Let* $\mathcal{P} = (P,T,m,s,f)$ *be an n-dimensional system of classical particle mechanics; let* $\mathcal{A}$ *be a nonsingular square matrix of order n; let* $\mathcal{B}$ *and* $\mathcal{C}$ *be n-dimensional vectors; and let* α, β, *and* γ *be real numbers such that* $\beta \neq 0$ *and* $\gamma > 0$. *Let* T' *be the set of all real numbers* t' *such that, for some t in T,*

(i) $$t' = \alpha + \beta t,$$

and let m', s', *and* f' *be defined by the following equations, for p any element of P,* t', *any element of* T', *and i any member of I:*

(ii) $$m'(p) = \gamma m(p)$$
$$s'(p,t') = \beta^2 s\left(p, \tfrac{t'-\alpha}{\beta}\right) \cdot \mathcal{A} + t'\mathcal{B} + \mathcal{C}$$
$$f'(p,t',i) = \gamma f\left(p, \frac{t'-\alpha}{\beta}, i\right) \cdot \mathcal{A}.$$

Then $\mathcal{P} = (P,T',m',s',f')$ *is also an n-dimensional system of classical particle mechanics. Conversely, if equations (i) and (ii) hold, and* (P,T',m',s',f') *is an n-dimensional system of particle mechanics, then so is* (P,T,m,s,f).

The surprise in Theorem A1 is that the matrix A is not restricted to a uniform change in the unit of measurement of spatial distance by necessarily being a similarity matrix, but is only nonsingular, which makes it a general affine matrix permitting different changes in spatial units along different dimensions.

Let ϕ_1 be a function which maps R (the set of all real numbers) onto itself in a one-to-one way; let ϕ_2 be a real-valued function defined over the set of positive real numbers; let ϕ_3 be a function which is defined over $E_n \times R$ (where E_n is the set of all n-dimensional vectors), and whose values are in E_n; and let ϕ_4 be a function which maps E_n onto itself in a one-to-one way. Then we call the ordered quadruple $(\phi_1, \phi_2, \phi_3, \phi_4)$ a *classically eligible transformation.*

Definition A2. *Let* $\Phi = (\phi_1, \phi_2, \phi_3, \phi_4)$ *be a classically eligible transformation, and let* $\mathcal{P} = (P, T, m, s, f)$ *be a system of particle mechanics. Then the* Φ*-transform of* $\mathcal{P}$ *is* the structure (P, T', m', s', f'), *where* T' *is the set of all real numbers* t' *such that, for some* t *in* T,

$$t' = \phi_1(t);$$

and m', s', *and* f' *are defined by the following equations (for* p *any element of* P, t' *any element of* T', *and* i *any element of* I*):*

$$\begin{aligned} m'(p) &= \phi_2[m(p)] \\ s'(p, t') &= \phi_3[s(p, \phi_1^{-1}(t')), t'] \\ f'(p, t', i) &= \phi_4[f(p, \phi_1^{-1}(t'), i)]. \end{aligned}$$

An important, but not necessary, restriction on φ_1 of an eligible transformation is that it holds uniformly as the transform of the measure of time for all particles p in P. Since the general axioms of Definition A1 do not postulate any kind of interaction between the particles, a separate time measure could have been introduced for each particle, so that we would have an indexing of φ_1 on P:

$$t' = \varphi_{1,p}(t),$$

and then the measure of time could be reversed for an arbitrary subset of P. Our subsequent imposition of Newton's third law prevents this, but it raises a problem, discussed below, for relativistic mechanics, for which simultaneous action at a distance, as by gravity or classical electrostatic attraction or repulsion, is not invariant under the Lorentz transformations.

THEOREM A2. *Let* $\Phi = (\phi_1, \phi_2, \phi_3, \phi_4)$ *be an eligible transformation, and suppose that, for every n-dimensional system of classical particle mechanics* $\mathcal{P}$, *the* Φ*-transform of* $\mathcal{P}$ *is again a system of classical particle mechanics. Then there are real numbers* α, β, γ, *with* $\beta \neq 0$, *and* $\gamma > 0$, *n-dimensional vectors* $\mathcal{B}$ *and* C, *and a nonsingular square matrix* $\mathcal{A}$, *of order n, such that, for x and Z any members of R and* E_n *respectively, and for y any positive* *member of R,*

$$\begin{aligned}\phi_1(x) &= \alpha+\beta x\\ \phi_2(y) &= \gamma y\\ \phi_3(Z,x) &= \beta^2 Z\cdot\mathcal{A}+x\mathcal{B}+\mathcal{C}\\ \phi_4(Z) &= \gamma Z\cdot\mathcal{A}.\end{aligned}$$

We now modify the axioms of Definition A1 to introduce Newton's third law—the new axioms are those given in Suppes (1957/1999). New Axiom P6 requires that the internal force $g(p,q,t)$ of particle q on particle p at time t is equal and opposite to $g(q,p,t)$. New Axiom P7 requires that the direction of interaction of the internal forces between particles p and q be the same as the direction of the vector $s(p,t)-s(q,t)$ from the position of p at time t to the position of q at time t. In stating P7 it is standard to use the vector or cross product x of two vectors (not to be confused with the Cartesian product of sets). Thus

$$\begin{aligned}x\times y &= (x_1,x_2,x_3)\times(y_1,y_2,y_3)\\ &= (x_2y_3-x_3y_2, x_3y_1-x_1y_3, x_1y_2-x_2y_1).\end{aligned}$$

Note that the direction of the vector product is perpendicular to the plane formed by the vectors x and y. We also restrict the internal forces (Axiom P8) below to be functions only of the distance between particles. Finally, in addition to introducing the concept of internal force $g(p,q,t)$, we must modify Newton's second law, Axiom P6 of Definition A1, to include internal forces in new Axiom P9 below.

Definition A3. *A structure $\mathcal{P}=(P,T,m,s,f,g)$ is a system of ultraclassical particle mechanics if and only if it satisfies the following axioms:*

Axioms P1-P5. *Same as Definition A1.*
Axiom P6. *For p and q in P and t in T,*

$$g(p,q,t)=-g(q,p,t).$$

Axiom P7. *For p and q in P and t in T,*

$$s(p,t)\times g(p,q,t)=-s(q,t)\times g(q,p,t).$$

Axiom P8. *Any internal force $g(p,q,t)$ is a function only of the Euclidean distance between the particles p and q at time t.*
Axiom P9. *For p in P and t in T,*

$$m(p)\frac{d^2}{dt^2}s(p,t)=\sum_{q\in P}g(p,q,t)+\sum_{i=1}^{\infty}f(p,t,i).$$

The following two theorems are about such systems of ultraclassical particle mechanics. Together they show that once interacting forces satisfying

P6-P8 are assumed, then the standard transformation result, with $\mathcal{A}$ a similarity matrix, obtains.

THEOREM A3. *If the matrix $\mathcal{A}$ of Theorem A1 is a similarity matrix, then $\mathcal{P}'$ is ultraclassical if and only if $\mathcal{P}$ is ultraclassical.*

THEOREM A4. *Let $\Phi = (\phi_1, \phi_2, \phi_3, \phi_4)$ be an eligible transformation in the sense of Definition A2, and for ultraclassical particle mechanics, let φ_4 be extended to the internal forces: $g'(p,q,t') = \varphi_4[f(p,q,\varphi^{-1}(t')]$ and suppose, moreover, that the Φ-transform of every ultraclassical system is ultraclassical. Then the matrix $\mathcal{A}$ of Theorem A2 is a similarity matrix.*

B Relativistic Mechanics

The primitive concepts of relativistic mechanics are the same as those used in Definition A1, with the addition of the positive constant c, standing physically for the speed of light.

Definition B1. *A structure $\mathcal{P} = (P, T, m, s, f, c)$ is an (n-dimensional) system of relativistic particle mechanics if and only if it satisfies the following axioms:*

Kinematical Axioms

Axioms P1-P3. *Same as Definition A1.*
Axiom P4. *The constant c is a positive real number such that for every p in P and t in T*

$$\left|\frac{d}{dt}s_p(t)\right| < c.$$

Axioms P5 and P6. *Same as axioms P4 and P5 of Definition A1.*

Dynamical Axiom

Axiom P7. *If $p \in P$ and $t \in T$, then*

$$m(p)\frac{d}{dt}\left[\frac{\frac{ds_p(t)}{dt}}{(1-\frac{|v_p(t)|^2}{c^2})^{\frac{1}{2}}}\right] = \left(1-\frac{|v_p(t)|^2}{c^2}\right)^{\frac{1}{2}}\sum_{i=1}^{\infty} f(p,t,i).$$

As Axiom P7 makes clear, the concept of rest mass, $m(p)$, is taken as primitive, and the relativistic mass is then definable.

We next introduce the concept of a generalized Lorentz matrix, corresponding to the affine matrix $\mathcal{A}$ in the transformation theorems for classical mechanics. The term "generalized" means that changes in units of measurement can be made in transforming from one relativistic mechanical system to another, and thus both c and c', for the speed of light are needed. The positive number λ used with c and c' represents a uniform stretch, if $\lambda > 1$, or shrinking, if $\lambda < 1$, of space and time. As a later theorem shows this holds in general for eligible transformations of the relativistic systems, a more restricted result than that for classical systems in the sense of Definition A1. The n-dimensional vector U is to be interpreted as the velocity of the second frame of reference with respect to the first one, in the notation used below.

Let c, c', and λ be positive real numbers. Then a matrix $\mathcal{A}$ of order $n+1$ is said to be a *generalized Lorentz matrix with respect to* (c, c', λ) if and only if there exist numbers δ and β, an n-dimensional vector U, and an orthogonal matrix $\mathcal{E}$ of order n, such that

$$\delta^2 = 1, \qquad \beta^2\left(1 - \frac{U^2}{c'^2}\right) = 1,$$

and

$$(i) \qquad \mathcal{A} = \lambda \begin{pmatrix} \mathcal{I} & 0 \\ 0 & \frac{c}{c'} \end{pmatrix} \begin{pmatrix} \mathcal{E} & 0 \\ 0 & \delta \end{pmatrix} \begin{pmatrix} \mathcal{I} + \frac{\beta-1}{U^2} U^*U & -\frac{\beta U^*}{c'^2} \\ -\beta U & \beta \end{pmatrix}.$$

The following two lemmas simplify the statement and proof of Theorem B1.

Lemma B1. *Let $(\{1\}, T, m, s, f, c)$ be a system of relativistic particle mechanics, let c' and λ be positive real numbers, and let $\mathcal{A}$ be a generalized Lorentz matrix with respect to (c, c', λ). Let the function h be defined by the equation (for every t in T):*

$$h(t) = [(s_1(t), t)\mathcal{A}]_{n+1}.$$

Then the function $\frac{dh}{dt}$ exists; its values are either always positive or always negative; and the function h is one-to-one.

The following theorem of matrix theory provides a more elegant characterization of generalized Lorentz matrices.

Lemma B2. *Let c, c', and λ be positive real numbers. Then a matrix $\mathcal{A}$ of order $n+1$ is a generalized Lorentz matrix with respect to (c, c', λ) if and only if*

$$(i) \qquad \mathcal{A}\begin{pmatrix} \mathcal{I} & 0 \\ 0 & -c'^2 \end{pmatrix}\mathcal{A}^* = \lambda^2 \begin{pmatrix} \mathcal{I} & 0 \\ 0 & -c^2 \end{pmatrix}.$$

THEOREM B1. *Let (P, T, m, s, f, c) be an n-dimensional system of relativistic particle mechanics. Let c', γ, and λ be positive real numbers, let B be an $(n+1)$-dimensional vector, and let $\mathcal{A}$ be a generalized Lorentz matrix with respect to (c, c', λ). Let the function h be defined as follows (for all t in T):*

$$h(t) = [(s_p(t), t)\mathcal{A} + B]_{n+1}.$$

(By Lemma B1 the inverse function h^{-1} exists.) Let the function T' be defined as follows: T' is the range of the function h; and let the functions m', s', and f' be defined by the following equations (for p in P, t' in T and i in I):

$$m'(p) = \gamma m(p),$$
$$s'(p,t') = [(s(p, h^{-1}(t')), h^{-1}(t'))\mathcal{A} + B]_{1,\ldots,n},$$
$$f'(p,t',i) = \frac{\gamma c'^2}{\lambda^2 c^2}\left[\left(f(p, h^{-1}(t'), i), \frac{f(p, h^{-1}(t'), i)\cdot v_p(h^{-1}(t'))}{c^2}\right)\mathcal{A}\right]_{1,\ldots,n}.$$

Then $\mathcal{P}' = (P, T', m', s', f', c')$ is an n-dimensional system of relativistic particle mechanics.

Let ϕ_1 be a function mapping R^+ into R^+; let ϕ_2 be a function which is a one-to-one mapping of E_{n+1} into itself; and let ϕ_3 be a function mapping E_{2n} into E_n. Then we call the ordered triple (ϕ_1, ϕ_2, ϕ_3) a *relativistically eligible transformation.*

Definition B2. *Let $\Phi = (\phi_1, \phi_2, \phi_3)$ be a relativistically eligible transformation; let $\mathcal{P} = (P, T, m, s, f, c)$ be a system of relativistic particle mechanics; and let the function H be defined as follows (for every t in T):*

$$H(t) = [\phi_2(s(p,t), t)]_{n+1}.$$

Then by the Φ-transform *of P (which we also write: $\Phi(P)$), we mean the structure $(P, \mathcal{T}', m', s', f')$, where for p in P:*

$$m'(p) = \phi_1(m(p));$$

T is the range of the function H; and s' and f' are defined by the following equations for t' in T', if the pre-image $H^{-1}(t')$ of t' under H is unique, and otherwise they are undefined:

$$s'(p,t') = [\phi_2(s(p, H^{-1}(t'))), H^{-1}(t')]_{1,\ldots,n}$$

$$f'(p,t',i) = \phi_3(f(p, H^{-1}(t'), i)v(p, H^{-1}(t'))),$$

for $i \geq 1$.

THEOREM B2. *Let $\Phi = (\phi_1, \phi_2, \phi_3)$ be an eligible transformation, and let c and c' be positive real numbers such that (i) for every n-dimensional system of relativistic particle mechanics $\mathcal{P}_c$, $(\Phi(\mathcal{P}_c), c')$ is a system of relativistic particle mechanics, and (ii) ϕ_2 carries no c-line into a c'-particle path. Then there exist positive real numbers γ and λ, an $(n+1)$-dimensional vector B, and a generalized Lorentz matrix $\mathcal{A}$ with respect to (c, c', λ), such that, for* *any vectors* Z_1 *and* Z_2 *in* E_n *with* $|Z_2| < c$, *every x in R, and y in* R^+,

$$\begin{aligned}\phi_1(y) &= \gamma y,\\ \phi_2(Z_1, x) &= (Z_1, x)\mathcal{A} + B,\\ \phi_3(Z_1, Z_2) &= \frac{\gamma c'^2}{\lambda^2 c^2}\left[\left(Z_1, \frac{Z_1 \cdot Z_2}{c^2}\right)\mathcal{A}\right]_{1,\ldots,n}.\end{aligned}$$

Condition (ii) of the theorem just requires that φ_2 can map no light line in the first frame of reference into a possible particle path in the second frame of reference. The lengthy proof of this theorem, which is given in Rubin and Suppes (1954), arises from the fact that no assumption about the continuity of φ_2 and φ_3 of an eligible transformation is made.

For the framework of the main part of this article, three theorems in the appendix state fundamental results on weak, but not strong, reversibility of mechanical systems. Theorem A2 says that every classical system of particle mechanics is weakly reversible, i.e., can be transformed into a system of classical mechanics with time reversed ($\beta < 0$ in the notation of the theorem). Theorem A4 says the same thing for ultraclassical systems. Finally, in the definition of a generalized Lorentz matrix, just before Lemma A1, $\delta = -1$ represents time reversal, which is not as evident in the formulation of Theorem B2, as it was in the classical case, but this is just a matter of the technical details of formulation. On the matter of time reversal considered here, there is no conceptual difference between classical and relativistic systems of particle mechanics. In both cases the time reversal itself is restricted to a linear transformation.

References

Feller, W. (1950) *Probability Theory and its Applications*. New York: John Wiley & Sons.

Kac, M. (1959) *Probability and Related Topics in Physical Sciences*. New York: Interscience Publishers.

Kelly, F. P. (1979) *Reversibility and Stochastic Networks*. New York: John Wiley & Sons.

Lamperti, J. and Suppes, P. (1959) Chains of infinite order and their application to learning theory. *Pacific Journal of Mathematics*, **9**, 739-143.

McKinsey, J. C. C., Sugar, A. C. and Suppes, P. (1953) Axiomatic foundations of classical particle mechanics. *Journal of Rational Mechanics and Analysis*, **2**, 253-272.

McKinsey, J. C. C. and Suppes, P. (1953) Transformations of systems of classical particle mechanics. *Journal of Rational Mechanics and Analysis*, **2**, 273-289.

Rubin, H. and Suppes, P. (1954) Transformations of systems of relativistic particle mechanics. *Pacific Journal of Mathematics*, **4**, 563-601.

Suppes, P. (1957) *Introduction to Logic*. Princeton: Van Nostrand Co. Reprinted 1999 by Dover Publications, New York.

Von Neumann, J. (1955) *Mathematical Foundations of Quantum Mechanics*. Princeton: Princeton University Press. Translated from the German edition of 1932 by Robert T. Beyer.