The Conserved Quantity Theory of Causation and Closed Systems*

Sungho Choi†‡

Advocates of the conserved quantity (CQ) theory of causation have their own peculiar problem with conservation laws. Since they analyze causal process and interaction in terms of conserved quantities that are in turn defined as physical quantities governed by conservation laws, they must formulate conservation laws in a way that does not invoke causation, or else circularity threatens. In this paper I will propose an adequate formulation of a conservation law that serves CQ theorists' purpose.

1. Causation and Conservation Laws. For advocates of the conserved quantity (CQ) theory of causation, causation is closely related to conservation laws in that they analyze the notions of causal process and interaction in terms of conserved quantities that are in turn defined as physical quantities governed by conservation laws. Hence their CQ theory will reach its complete form only if the exact formulation of a conservation law is offered. Unfortunately, however, it is difficult to find a general formulation of a conservation law of an arbitrary physical quantity in the literature of physics. Instead, we can find formulations of conservation laws of such particular physical quantities as linear momentum, electric charge, etc. For example, in well-known physics textbooks, the linear momentum conservation law and the electric charge conservation law are respectively formulated as follows:

The Linear Momentum Conservation Law. For an isolated system,
   subject only to internal forces (forces between members of the
   system), the total linear momentum is a constant (Kittel, Knight, and
   Ruderman 1973, 83).
The Electric Charge Conservation Law. The total electric charge in an
   isolated system, that is, the algebraic sum of the positive and negative
   charge present at any time, never change (Purcell 1985, 5).

These two formulations suggest that a conservation law of a physical
quantity Q follows the pattern that the total amount of Q possessed by a
closed system[1] does not change through time. Or, to be precise,

CL. For every system that possesses Q, if it is closed from outside, the
   total amount of Q possessed by it remains constant through time.[2]

The requirement to possess the physical quantity Q is indispensable because,
for example, the electric charge conservation law does not apply to such
entities as moving spots and shadows that do not possess electric charge.
   The notion of closed system in (CL) cries out for a definition. And,
(CL) will be available to CQ theorists only if a closed system can be defined
without invoking the notions of causal process and interaction. In fact,
Hitchcock (1995, 315–316) objects to the CQ theory that a closed system
is just a system that does not engage in any causal interactions.[3] On Hitch-
cock's view, a conservation law presupposes the notion of causal inter-
action because a closed system is defined in terms of causal interaction. If
so, the CQ theory comes out as being circular since according to the CQ
theory causal interactions are analyzed in terms of conserved quantities
that are defined as physical quantities governed by conservation laws.
   The crucial question is thus whether or not a closed system in (CL) can
be defined without invoking the notion of causal interaction: whether the
CQ theory comes out as being circular or not depends on the answer to
this question. I believe that we can define a closed system without invoking
the notion of causal interaction. In the following I will propose such a
definition.

2. Hitchcock's Definition. Let me start by pointing out that Hitchcock's
definition that a closed system is just a system that does not engage in any
causal interactions does not work. First of all, it allows an obvious coun-

1. I will use "closed system" interchangeably with "isolated system".

2. In formulating (CL) I am indebted to an anonymous referee.

3. This is not a philosopher's prodigious view. It does not take much effort to find out
something like Hitchcock's view in textbooks of physics (Marion 1970, 214; Arya 1990,
470): a system that does not interact with anything outside the system is called a closed
system.

terexample. The total energy of a system is equal to the sum of its kinetic energy and total potential energy, where the total potential energy is in turn the sum of the external potential energy due to the external forces on it and the internal potential energy due to the internal forces within it. It is clear that physicists will deny that (kinetic energy + internal potential energy)—for later references call it "internal energy"—is a conserved quantity. They will argue that, for example, the internal energy possessed by a free falling rigid body classically described is not conserved during its fall; what is conserved is the total energy of the rigid body, which is equal to the sum of the internal energy and the gravitational potential energy. Let us now consider a system that is closed from outside according to Hitchcock's definition. This system does not engage in any causal interactions with the outside and, therefore, no external forces act on it because force is a species of causal relation (Bigelow, Ellis, and Pargetter 1988). But if no external forces act on a system, then its internal energy will remain constant over time (Symon 1985, 167; Arya 1990, 278). This means that, according to Hitchcock's definition of a closed system, a closed system invariably possesses a constant amount of internal energy through time. Thus, internal energy comes out as a conserved quantity by Hitchcock's definition. But, as noted above, internal energy is not a conserved quantity.[4]

What is worse, I am afraid that Hitchcock's definition has a more serious problem. It is clear that a free falling body that is described classically does engage in continuous interactions with the gravitational field. Does this mean that the free falling body is not closed from outside? In a sense yes, but in a sense no. Notice that the causal interactions between the free falling body and the gravitation field involve exchanges of linear momentum and yet do not involve exchanges of energy; thereby, the free falling body does not possess a constant amount of linear momentum but possesses a constant amount of energy. Hence physicists will say that the free falling body is closed with respect to energy but not with respect to linear momentum.[5] This means that we have another good reason to reject Hitchcock's definition.[6]

4. Here I thank an anonymous referee for pointing out my previous mistake.

5. To be precise, the free-falling body does not possess a constant amount of the component of linear momentum along the direction of the line joining it with the center of the earth; thereby it is not closed with respect to the component of linear momentum in that direction. A similar remark will be pertinent to the following case of the two far distant electrons.

6. Hitchcock might respond to my objection by modifying his definition such that a system is closed with respect to a physical quantity Q iff it engages in no causal interactions involving Q: the free falling body is not closed with respect to linear momentum because it engages in causal interactions with the gravitational field and the causal

I take it that a closed system is always closed with respect to a certain physical quantity. Moreover, a closed system with respect to a physical quantity may not be closed with respect to another physical quantity. To take another example, suppose that the universe has only two far distant free electrons that exert gravitational and electrical forces on each other. In this case, the amounts of linear momentum possessed by them will change, respectively. Therefore, they are not closed systems with respect to linear momentum. But we can say that each of them is a closed system with respect to electric charge since they will not exchange electric charge. Thus each electron is not closed with respect to linear momentum but closed with respect to electric charge.

In view of this, we have to amend (CL) by inserting "with respect to Q" in the clause, "if it is closed from outside [with respect to Q], the total amount of Q possessed by it remains constant through time." The linear momentum conservation law does not apply to either the free falling body or the two electrons since they are not closed with respect to linear momentum. Yet, the energy conservation law does apply to the free falling body because it is closed with respect to energy and the electric charge conservation law does apply to each of the two electrons because each is closed with respect to electric charge.

3. Definitions of Closed Systems in the Literature of Physics. What is a closed system with respect to a physical quantity? It seems to be a sensible strategy to examine the physics literature to seek out an answer to this question. Unfortunately, physicists seem to have little concern for answering the question. However, we can infer definitions of closed systems with respect to such conserved quantities as energy, linear momentum, etc., from textbook formulations of the corresponding conservation laws (Goldstein 1980, Chapter 1; Marion 1970, 62–75; Purcell 1985, 4–5). Take some examples:

(1) A system is closed at a time t with respect to the component of linear momentum along a direction iff the totality of external forces acting on it has a zero component along the direction at t. And, a system

interactions involve linear momentum; yet it is closed with respect to energy because the causal interactions do not involve energy. But I am afraid that this response does not work. First, it is not at all clear how we can clarify "causal interactions involving Q" in an innocuous way, especially when Q is not a conserved quantity. Furthermore, even if we grant that Hitchcock's possible response is successful, it cannot completely save Hitchcock's definition because, as stated above, Hitchcock's definition suffers from an obvious counterexample. Hitchcock's possible response was brought to my attention by an anonymous referee.

is closed at a time t with respect to linear momentum iff it is closed at t with respect to every component of linear momentum.

(2) A system is closed at a time t with respect to the component of angular momentum along a direction iff the totality of external torques has a zero component along the direction at t, where external torques are torques due to external forces acting on the system. And, a system is closed at a time t with respect to angular momentum iff it is closed at t with respect to every component of angular momentum.

(3) A system is closed at a time t with respect to energy iff there are no non-conservative external forces acting on it at t, where a force $\vec{F}$ is non-conservative iff $\nabla \times \vec{F} \neq 0$.

(4) A system is closed at a time t with respect to electric charge iff no matter with non-zero amount of electric charge goes out of or into it at t.[7]

A closed system with respect to such a conserved quantity as baryon number and lepton number can be defined in the same way as (4). Thus we seem to have definitions of closed systems with respect to conserved quantities.

On the one hand, according to (1), the free falling body is not closed with respect to the component of linear momentum along the direction (call it z-direction) of the line linking it with the center of the earth because of the gravitational force. But, the free falling body is closed with respect to the component of linear momentum along a direction vertical to the z-direction since that component of linear momentum is unaffected by the gravitational force. Moreover, the gravitational force is conservative. Therefore, according to (3), the free falling body is closed with respect to energy. On the other hand, each of the two far distant electrons is closed with respect to electric charge according to (4) since no matter with non-zero amount of electric charge go into or out of it. Yet, according to (1), each of them is not closed with respect to the component of linear momentum along the direction of the line linking them because of the electromagnetic and gravitational forces.

Unfortunately, I take it that the definitions (1) to (4) are not satisfactory. Note that we need to define closed systems with respect to such non-conserved quantities as length, volume, rotational inertia, velocity of center of mass, force, potential energy, etc., as well as those with respect to conserved quantities. Take the example of force. To say that force is not

7. Purcell is one of the rare physicists who explicitly state what a closed system with respect to electric charge is. He (1985, 4) says, "By *isolated* we mean that no matter is allowed to cross the boundary of the system. We could let light pass into or out of the system, since the "particle" of light, called photons, carry no charge at all."

a conserved quantity is roughly to say that "the force conservation law" does not hold true in our actual world. And, according to (CL), "the force conservation law" would go that a closed system from outside with respect to force possesses a constant amount of force through time. Hence we can make sense of the statement that force is not a conserved quantity only if we have gotten an appropriate definition of a closed system with respect to force. However, physicists have had no concern for providing such a definition and so we cannot rely on the literature of physics in searching for it. What is worse, the above-mentioned definitions of closed systems with respect to conserved quantities give us no clue to how a closed system with respect to force can be defined. Thus, given (1) to (4), it is not at all clear how we can define closed systems with respect to non-conserved quantities like force. This suggests that (1) to (4) are not on the right track for providing adequate definitions of closed systems.

More importantly, since the concept of force is closely entwined with the concept of causation, such definitions as (1), (2) and (3) are not available to CQ theorists. Bigelow, Ellis, and Pargetter (1988) argue that force is a species of causal relation between a change in a field and an action on a particle or particles. On their view, "Forces are constituents within causal interactions: they are the causal relations which hold between salient participants in the causal interaction" (Bigelow, Ellis, and Pargetter 1988, 624–625). I do not wish to be committed to all the details of their view on forces, but I agree with them that the concept of force should be analyzed in terms of causation. Hence, in my opinion, definitions of closed systems invoking the concept of force render the CQ theory circular since, according to them, closed systems are ultimately defined in terms of causation—or causal interaction. This means that CQ theorists cannot help themselves to those definitions.

One might respond that from Noether's theorem, we can infer alternative definitions of closed systems with respect to conserved quantities that do not invoke the concept of force.[8] It is said that Noether's theorem enables physicists to get conserved quantities from symmetries of the laws of nature.[9] For example, it follows from the theorem that whenever a physical system whose dynamical behavior can be described by Lagrange's equation of motion has a Lagrangian that is invariant under translations in a direction, the component of its linear momentum along that direction is a constant of motion. In light of this consideration, one might propose that a system is closed with respect to the component of linear momentum

8. I thank Jeayoung Ghim for raising this issue.

9. For a simple and straightforward presentation of the classical version of Noether's theorem see (Desloge and Karch 1977) and (Nilo Bobillo-Ares 1988). For a detailed discussion of Noether's theorem see (Doughty 1990).

along a direction iff its Lagrangian is invariant under translations in that direction.[10] Likewise, it follows from Noether's theorem that whenever a physical system whose dynamical behaviour can be described by Lagrange's equation of motion has a Lagrangian that is invariant under translations in time, the Hamiltonian of the system is a constant of motion. And under natural assumptions, the Hamiltonian is equal to the total energy of the system (Goldstein 1980, 62). Hence it seems natural to suggest that a system is closed with respect to energy iff its Lagrangian is invariant under translations in time.

On this proposal, the free falling body described classically is not closed with respect to the component of linear momentum along the z-direction since its Lagrangian depends explicitly on the z-coordinate and, therefore, is not invariant under translations in the z-direction. By contrast, it is closed with respect to energy since its Lagrangian does not depend explicitly on the time and, therefore, is invariant under translations in time. Note that the two definitions of closed systems inferred from Noether's theorem do not invoke the concept of force. Hence they do not render the CQ theory of causation circular.

I am afraid, however, that we cannot draw adequate definitions of closed systems from Noether's theorem. Suppose that we describe a particle near a strong source of gravitation in general relativity. In the curved spacetime interpretation of general relativity, the space around the particle is not "homogeneous" in that particular positions where the particle is located are relevant to its lawlike behavior. To be specific, as the particle moves toward the source of gravitation the Lagrangian varies solely because of qualitative differences of positions in space itself. If so, the Lagrangian would not be invariant under translations in space. But in general relativity gravitation is a kind of fictitious force akin to the centrifugal force that is due to non-uniform motion by the observer's reference frame. Therefore, we would have to say that the particle is closed with respect to linear momentum since no external "real" forces act on it.[11] This suggests that if space is not homogeneous, even a closed system with respect to linear momentum may have a Lagrangian that is not invariant under translations in space. For this reason I believe that the reference to the homogeneity of space is indispensable in Marion's (1970, 216) following

10. On this proposal, Noether's theorem could be taken as proving that the linear momentum conservation law holds for any physical system whose dynamical behaviour can be described by Lagrange's equation of motion.

11. One might insist that the particle is not closed with respect to linear momentum since gravitational force, though fictitious, acts on it. But this has the unacceptable consequence that no physical system is closed with respect to linear momentum; for, every physical system when described in a non-inertial reference frame is subject to a fictitious force.

statement: "Since space is homogeneous in an inertial reference frame, the Lagrangian of a closed system will be unaffected by a translation of the entire system in space."

Thus as long as space is not homogenous, the concept of closed system with respect to linear momentum is extensionally different from the concept of invariance of Lagrangian under translations in space. This means that we cannot define the former in terms of the latter. In my opinion, what Noether's theorem implies is that "if the Lagrangian of a system, *closed or otherwise,* is invariant with respect to a translation in a certain direction, then the linear momentum of the system in that direction is constant in time" (Arya 1990, 471; my italic). If so, it is unpromising to draw adequate definitions of closed systems from Noether's theorem.

4. Dowe's Proposal. In the previous section we have seen that we cannot draw adequate definitions of closed systems from textbook formulations of conservation laws nor from Noether's theorem. Dowe, in his recent book, takes another tack. He says:

> . . . we need to explicate the notion of a closed system in terms only of the quantities concerned. For example, energy is conserved in chemical reactions, on the assumption that there is no net flow of energy into or out of the system. (Dowe 2000, 95)

In this passage, Dowe seems to propose a definition of a closed system with respect to energy: a system is closed with respect to energy at a time t iff there is no net flow of energy into or out of the system at t. But it is not clear what Dowe means by the metaphorical expression "net flow of energy". In fact, Schaffer (2001, 811) complains that he does not see what alternative explication of a closed system Dowe is offering. McDaniel (2002, 261) goes so far as to say "there doesn't seem to be a way to explicate the concept of energy flow without appealing to the concept of causation."

However, I take it that there is one obvious, though ultimately wrong, way to define energy flow without appealing to the concept of causation: we can define it by invoking the notion of identity of physical quantity over time, which is similar to the notion of identity of substance over time. On this view, whenever an amount of energy outside a system before a time t is genidentical with a numerically identical amount of energy inside the system after t,[12] we can say that this amount of energy "flows" from the outside of the system to the inside about t. And there is a net flow of

12. The expression "energy inside the system" means energy possessed by the system. And, the expression "energy outside the system" means energy possessed by the rest of the world outside of the system.

energy into (out of) a system at a time t iff the total amount of energy flow from the outside of the system to the inside is larger (smaller) than the total amount of energy flow from the inside of the system to the outside at t.

In general, whenever an amount of a physical quantity Q outside a system before a time t is genidentical with a numerically identical amount of Q inside the system after t, the amount of Q "flows" from the outside of the system to the inside about t. Similarly, whenever an amount of Q inside a system before t is genidentical with a numerically identical amount of Q outside the system after t, the amount of Q "flows" from the inside of the system to the outside about t. Then Dowe's definition of a closed system with respect to a physical quantity Q will go as follows: a system is closed with respect to Q at a time t iff there is no net flow of Q into or out of the system at t, where there is no net flow of Q into or out of a system at t iff the total amount of Q that flows from the outside of the system to the inside is equal to the total amount of Q that flows from the inside of the system to the outside at t. This definition does not invoke the notions of causal process and interaction. Hence it does not render Dowe's CQ theory circular.

But, as Dowe (1995, 369) correctly points out, we do not think of a quantitative property as having such an identity over time. We do not want to say that the height of this tree at a time is identical to the height of it at a later time in the same sense as this book at a time is identical to it at a later time. In everyday speech, we say that the height of this tree at a time is identical to the height of it at a later time only in the sense that the tree instantiates the one universal at different times. Hence Dowe concludes that, as far as a physical quantity is a property that cannot stand possessed by nothing, the notion of identity of physical quantity over time is not at all clear.

Another problem for the above-mentioned definition of flow of physical quantity is that it is unclear how we can identify a physical quantity over time. It may be argued that we can identify energy and momentum through time by using the energy and momentum conservation laws, respectively (Fair 1979, 234). For example, we can identify the amount of energy possessed by a free particle at a time with the amount of its energy at a later time by relying on the energy conservation law. But in such cases of many body problems as Fair's (1979, 238) and Ehring's (1986, 256), the energy conservation law cannot determine which energy is genidentical to which energy. For this reason, Dowe (1995, 370) says "whenever we move to many body problems the conservation laws leave the question of identity over time of the energy indeterminate." If so, we cannot always identify energy and momentum through time by relying on the energy and momentum conservation laws.

I take it that Dowe's objection to genidentity of physical quantity is absolutely plausible. Then it is misdirected to define the flow of physical quantity by invoking genidentity of physical quantity. Thus one obvious answer to how to define a net flow of physical quantity goes nowhere. However, this failure itself does not pose a conclusive problem for Dowe's definition that a system is closed with respect to a physical quantity Q iff there is no net flow of Q into or out of the system. In my opinion, the conclusive problem for Dowe's definition is that the notion of flow of physical quantity seems to make sense only if the quantity is a conserved quantity. Note that physicists speak of "flow of physical quantity" only if the quantity is a conserved quantity: they speak of "energy flow" or "electric charge flow (i.e., electric current)" but do not speak of "flow of speed" nor "flow of force". In fact, it is very unclear what "flow of speed" or "flow of force" may mean. I believe that this is because the metaphorical expression "flow" implies a kind of conservation. If energy could be created ex nihilo, then we would no longer say that an amount of energy flows from here to there. This suggests that the notion of flow of physical quantity assumes that the quantity in question is a conserved quantity. If so, the notion of flow of physical quantity will be conceptually dependent on the notion of closed system in the end because a conserved quantity is a quantity governed by a conservation law that is in turn formulated in terms of a closed system. Therefore, I conclude that Dowe's definition faces a blatant circularity.

In my opinion, Dowe is wrong that a closed system with respect to a physical quantity can be defined in terms of net flow of the physical quantity. Nevertheless I believe that Dowe is right that the notion of closed system should be defined only in terms only of the quantities concerned. I will now provide such a definition.

5. What Is a Closed System? Let us first consider the following definition of a closed system with respect to a scalar quantity:

DC. A system is closed with respect to a physical quantity Q at a time t iff (a) $dQ_{in}/dt = dQ_{out}/dt = 0$ at t or, (b) $dQ_{in}/dt \neq -dQ_{out}/dt$ at t,

where Q_{in} is the amount of Q inside the system and Q_{out} is the amount of Q outside the system.[13] In case a physical quantity Q is a vector or a tensor,

13. Readers should beware of misinterpreting dQ_{in}/dt and dQ_{out}/dt. What I have in mind is that dQ_{in}/dt represents the rate at which the amount of Q inside the system changes and dQ_{out}/dt the rate at which the amount of Q outside the system changes. Hence it is incorrect to interpret, for example, dQ_{in}/dt as representing the rate at which Q "enters" the system from the outside and dQ_{out}/dt the rate at which Q "leaves" the system. I thank an anonymous referee for pointing out that there is a need to take precautions.

we can say that a system is closed with respect to Q iff all the components of Q satisfy the definiens of (DC).

We can state what (DC) means in a more succinct way by considering what is not a closed system with respect to Q according to (DC):

(5) A system is not closed with respect to Q at a time t iff $dQ_{in}/dt = -dQ_{out}/dt \neq 0$ at t.

When Q_{in} increases and Q_{out} decreases at the same non-zero rate, or when Q_{in} decreases and Q_{out} increases at the same non-zero rate, the system is not closed with respect to Q according to (5). Otherwise, the system is closed with respect to Q according to (DC).

Take some examples. A free falling body possesses a constant amount of energy. According to (DC), this means that it is closed with respect to energy regardless of whether the energy outside it remains constant or not. Let Pz denote the component of linear momentum along the direction of the line linking the free falling body with the center of the earth. During the fall, the Pzs possessed by the free falling body and possessed by the earth increase in opposite directions at the same rate. Hence, the free falling body does not count as being closed with respect to Pz by (DC); so it is not closed with respect to linear momentum. According to (DC), two far distant free electrons are closed with respect to electric charge, respectively, because each of them possesses a constant amount of electric charge. Let Px denote the component of linear momentum in the direction of the line joining the two electrons. Since they are free, the Pxs possessed by them increase in opposite directions at the same rate. Hence, each electron does not count as being closed with respect to Px by (DC); therefore, it is not closed with respect to linear momentum. An object sliding on a surface with friction present possesses a decreasing amount of energy. And, the energy decrease of the object is exactly equal in magnitude to the energy increase of molecules constituting the surface (Feynman, Leighton and Sands 1989, 14–7). This means that, according to (DC), the object is not closed with respect to energy. Thus (DC) does justice to physicists' common-sense understanding of closed systems. Hence I believe that (DC)—and its generalization to vector or tensor quantities—is an adequate definition of a closed system with respect to a physical quantity.[14]

In the following I will provide a motivating argument for (DC). Before

14. Physicists believe that physical quantities like energy that are known as conserved quantities remain constant in the world as a whole. For those quantities, (b) of (DC) cannot be fulfilled since it require that $Q_{in} + Q_{out}$, which is the total amount of Q in the world, should not remain constant. Accordingly, a system is closed with respect to such a quantity iff (a) of (DC) is fulfilled. But since the quantity remains constant in the world as a whole, we can state that a system is closed with respect to it iff the system possesses a constant amount of it.

going into the matters, let me introduce some terminology that will help express my idea more clearly. When a system is such that if it is closed from outside with respect to Q then it possesses a constant amount of Q, I will say "Q is conserved in the system". It is clear that (CL) is true iff Q is conserved in every Q-possessing system.

Suppose that two systems, R_1 and R_2, each of which is closed with respect to energy, possess certain amounts of energy at a time t_1. Suppose further that energy is conserved in each of R_1 and R_2. Then we can predict the total amount of energy possessed by the two systems at a later time t_2—for short, E-amount—by applying the energy conservation law to each of R_1 and R_2 and then summing their respective amounts of energy at t_2. In this case, it seems clear that we can also predict the E-amount by applying the energy conservation law to a large system R_t composed of the two systems and then getting the amount of energy possessed by R_t at t_2. This is because if energy is conserved in each of R_1 and R_2, then it is also conserved in R_t. This line of thought suggests the following principle:

CP. Suppose a system S is composed of subsystems S_1, S_2, . . . S_n. Then energy is conserved in S iff, for every i, energy is conserved in S_i.

(CP), together with the supposition that energy is conserved in each of R_1 and R_2, implies that it is also conserved in R_t.

It should be noted that, given that both R_1 and R_2 are closed with respect to energy, R_t is also closed with respect to energy. It is obvious that this is the case according to (3) or Dowe's definition. Moreover, this is the case according to (DC). Let E_{R1} and E^*_{R1} be the amount of energy inside R_1 and the amount of energy outside R_1, respectively. Likewise, let E_{R2}, E^*_{R2}, E_{Rt} and E^*_{Rt} be the amount of energy inside R_2, the amount of energy outside R_2, the amount of energy inside R_t and the amount of energy outside R_t, respectively. Suppose that R_t is not closed with respect to energy. Then $dE_{Rt}/dt = -dE^*_{Rt}/dt \neq 0$. Since $E_{R1} + E_{R2} = E_{Rt}$, we have $dE_{R1}/dt = dE_{Rt}/dt - dE_{R2}/dt$. Therefore, we have $dE_{R1}/dt = -dE^*_{Rt}/dt - dE_{R2}/dt = -d(E^*_{Rt} + E_{R2})/dt$. But, $(E^*_{Rt} + E_{R2})$ is exactly the amount of energy outside R_1, that is, E^*_{R1}. Hence we have the result that $dE_{R1}/dt = -dE^*_{R1}/dt$. We can get a similar result for R_2: $dE_{R2}/dt = -dE^*_{R2}/dt$. Notice that at least one of dE_{R1}/dt and dE_{R2}/dt is non-zero because dE_{Rt}/dt is supposed to be non-zero and $dE_{Rt}/dt = dE_{R1}/dt + dE_{R2}/dt$. This means that, according to (DC), at least one of R_1 and R_2 is not closed with respect to energy. Therefore, we can conclude that, according to (DC), if both of the two subsystems are closed with respect to energy, then the large system is also closed with respect to energy.

As stated above, (CP) implies that energy is conserved in R_t. Since both R_1 and R_2 are supposed to be closed with respect to energy, R_t is also closed with respect to energy. If so, the amount of energy possessed by R_t

remains constant over time. Thus (CP) ensures that whenever we can pre-dict the E-amount by applying the energy conservation law to each of R_1 and R_2, we can also predict the E-amount by applying the energy conser-vation law to R_r. I take it that (CP) is reasonably plausible and so I will assume it in the following.

It is clear that our world S_t as a whole is a closed system with respect to energy. If not, what else could be? In fact, S_t counts as being closed with respect to energy by (3) stated in Section 3 because there are no non-conservative *external* forces acting on it; and, it counts as such by Dowe's definition because there is no net flow of energy into or out of it; and, it counts as such by (DC) because the amount of energy outside S_t is always zero. Note that we can construe a physical system S_1 and the rest of the world S_2 outside of S_1 as two mutually exclusive and jointly exhaustive subsystems of the world S_r. Let E_{S1} be the amount of energy possessed by S_1 and E_{S2} be the amount of energy possessed by S_2. Now imagine the three following cases that are mutually exclusive and jointly exhaustive:

Case 1. Suppose that $dE_{S1}/dt = dE_{S2}/dt = 0$ during a time interval I. Then both E_{S1} and E_{S2} remains constant over the time interval I; thereby, the total amount of energy possessed by S_t, i.e., $(E_{S1} + E_{S2})$ also remains con-stant over the time interval I. Since S_t is closed with respect to energy, energy is conserved in S_r. Therefore, according to (CP), energy must be conserved in S_1 and S_2, respectively. This is the case according to (DC). Since each subsystem is supposed to possess a constant amount of energy, energy will be conserved in it if it is closed with respect to energy. And, according to (DC), each subsystem is closed with respect to energy during the time interval I since both the amount of energy inside it and the amount of energy outside it remain constant.

Case 2. Suppose that $dE_{S1}/dt = -dE_{S2}/dt > 0$ during a time interval I. Then E_{S1} increases and E_{S2} decreases at the same rate during the time interval I; thereby, the total amount of energy possessed by S_t, i.e., $<E_{S1} + E_{S2}>$ remains constant over the time interval I. Since S_t is closed with respect to energy, energy is conserved in S_r. According to (CP), it follows from this that energy is also conserved in S_1 and S_2, respectively. But if either of the two subsystems, say, S_1 is closed with respect to energy, then energy would not be conserved in S_1 because E_{S1} does not remain constant over the time interval I. This means that, for (CP) not to be violated by S_1, S_1 must not be a closed system with respect to energy during the time interval I. This is the case according to (5) since the energy inside it and the energy outside it change in opposite ways at the same rate. Likewise, S_2 does not count as being closed with respect to energy during the time interval I by (5), wherefore (CP) is not violated by S_2. We can get the same result for the case where $dE_{S1}/dt = -dE_{S2}/dt < 0$.

Case 3. Suppose that $dE_{S1}/dt \neq -dE_{S2}/dt$ during a time interval I. Then the total amount of energy possessed by S_t does not remain constant over the time interval I. Therefore, energy is not conserved in S_t because S_t is a closed system with respect to energy. According to (CP), it follows from this that energy is not conserved in at least one of the two subsystems. That is, (CP) implies that at least one of the two subsystems is closed with respect to energy and yet does not possess a constant amount of energy during the time interval I, which is the case according to (DC). Notice that the two subsystems, S_1 and S_2, count as being closed with respect to energy by (DC). Moreover, since dE_{S1}/dt and $-dE_{S2}/dt$ are not equal, at least one of them is non-zero; thereby, at least one of E_{S1} and E_{S2} does not remain constant; that is, at least one of the two subsystems does not possess a constant amount of energy during the time interval I. As a result, energy is not conserved in at least one of the two subsystems.

To sum up, (DC) ensures that (CP) is not violated in the three mutually exclusive and jointly exhaustive cases. Given that (CP) is assumed as a reasonably plausible principle, this gives us a good reason to accept (DC) as an adequate definition of a closed system with respect to energy.[15]

6. Further Advertisements. (DC) has a number of advantages as a definition of a closed system. First, (DC) implies that a system is closed with respect to a physical quantity iff the outside of it is closed with respect to that quantity, which is in accordance with physicists' common sense. Second, (DC) explains how it is possible that scientists justifiably idealize many systems that are not in fact closed as approximately closed systems.[16] According to (DC), when a system is such that $dQ_{in}/dt = -dQ_{out}/dt \neq 0$, the system is not closed with respect to a physical quantity Q. However, if $|Q_{in}| \gg |dQ_{in}/dt|$, then we are justified in ignoring dQ_{in}/dt and dQ_{out}/dt and idealizing the system as an approximately closed system. For example, a heavy body that falls in air is strictly not closed with respect to energy since during its fall it exchanges energy with air molecules. However, its energy is much larger than the energy that it exchanges with air molecules and scientists can justifiably idealize the body as being approximately closed with respect to energy. By contrast, in case such a light body as a feather falls in air, scientists do not idealize it as being approximately closed with respect to energy because its energy is comparable to the energy that it exchanges with air molecules. Thus (DC) explains one of well-established scientific practices.

15. I admit that I have not provided a strong enough argument for (CP). Those who do not think (CP) plausible may not take my motivation for (DC) seriously.

16. Here I am indebted to an anonymous referee.

Third, and more importantly, (DC) invokes only such concepts as numerical equality and inequality between physical quantities so that it does not presuppose that the quantity Q in question is a conserved quantity. Hence we can say what are closed systems with respect to such nonconserved quantities as force and velocity of center of mass by relying on (DC). Moreover, (DC) defines a closed system with respect to a physical quantity in terms only of the physical quantity without invoking the notions of causal process and interaction. This means that (DC) does not render the CQ theory of causation circular.

According to (CL), a conservation law of a physical quantity Q is roughly that every Q-possessing system that is closed with respect to Q possesses a constant amount of Q through time. According to (DC), this conservation law is equivalent to the statement that, for every Q-possessing system, if $dQ_{in}/dt \neq 0$ then $dQ_{in}/dt = -dQ_{out}/dt$. On this view, the electric charge conservation law has been supported by such an experimental result as reports that when the electric charge inside a system increases at a non-zero rate, the electric charge outside it invariably decreases at the same rate.

By contrast, "the velocity of center of mass—for short, V_{cm}—conservation law" has been refuted by such an experimental result as reports that, when a system's V_{cm} increases in a direction at a non-zero rate, the V_{cm} of the rest of the world outside of it does not regularly increase in the opposite direction at the same rate. For example, suppose that the universe has only two free (electrically neutral) particles with different masses that exert gravitational forces on each other. Then the light particle will move with a greater acceleration than the heavy one. In this case, the V_{cm} of the light particle increases in a direction at a non-zero rate and yet the V_{cm} of its outside, i.e., the V_{cm} of the heavy particle does not increase in the opposite direction at the same rate. To put another way, the light particle that, according to (DC), is closed from outside with respect to V_{cm} does not possess a constant amount of V_{cm}. Hence, "the V_{cm} conservation law" is refuted, wherefore velocity of center of mass is not a conserved quantity.

In Section 2 I argued that internal energy serves as a counterexample to Hitchcock's definition. It will be instructive to see that (DC) overcomes the counterexample. Let us reconsider the above-mentioned case of two particles with different masses. When we assume that the particles are point-particles, their internal potential energies vanish. Therefore, their internal energies are equal to their kinetic energies, respectively. It is clear that, as the two particles exert gravitational force on each other, both of the kinetic energies possessed by them increases because they are free. It follows from this that the internal energy of the light particle increases at a non-zero rate and yet the internal energy of the heavy particle does not decrease at the same rate. To put another way, the light particle that,

according to (DC), is closed from outside with respect to internal energy does not possess a constant amount of internal energy. Hence, "the internal energy conservation law" is refuted. In consequence, internal energy does not come out as a conserved quantity by (DC).

(DC) has the interesting consequence that, as far as classical mechanics is concerned, force is conserved. Let us assume the weak version of Newton's third law of action and reaction that forces between two particles are equal and opposite (Goldstein 1980, 5).[17] Then the total vector sum of internal forces inside a system always vanishes since a force on a subsystem i due to another subsystem j is equal in magnitude and opposite in direction to a force on the subsystem j due to the subsystem $i.$ Hence the amount of force acting on a system is equal to the total vector sum of external forces exerted on the system. Let us reconsider the system S_1 and the rest of the world S_2 outside of S_1. Suppose now that the amount of force acting on S_1 increases in a direction at a non-zero rate. Given the weak version of Newton's third law, this means that the total vector sum of forces exerted by S_2 on S_1 increases in that direction at the same rate; then the total vector sum of forces exerted by S_1 on S_2 would increase in the opposite direction at the same rate. Note that, according to the weak version of Newton's third law, the amount of force acting on S_2 is equal to the total vector sum of forces exerted by S_1 on S_2. Therefore, we come to the conclusion that when the amount of force acting on S_1 increases in a direction at a non-zero rate, then the amount of force acting on $S_2,$ i.e., the outside of S_1 increases in the opposite direction at the same rate. By a similar reasoning, we can draw the conclusion that when the amount of force acting on S_1 decreases in a direction at a non-zero rate, then the amount of force acting on the outside of S_1 decreases in the opposite direction at the same rate. In consequence, as long as the weak version of Newton's third law is valid, "the force conservation law" holds good: the total amount of force inside a closed system with respect to force remains constant through time.

When we take electromagnetism into account, however, the weak version of Newton's third law does no longer hold good; and so neither does the force conservation law. Consider two charged particles moving (instantaneously) so as to "cross the T", i.e., one charged particle moving directly at the other, which in turn is moving at right angles to the first (Goldstein 1980, 7–8). According to Biot-Savart law, the magnetic field due to the first particle vanishes at the location of the second; thereby, the first particle exerts no magnetic force on the second. By contrast, the sec-

17. The strong version of Newton's third law goes that the forces between the two particles, in addition to being equal and opposite, also lie along the line joining the particles (Goldstein 1980, 7).

ond particle exerts a non-zero amount of magnetic force on the first since the magnetic field due to the second does not vanish at the location of the first. Thus, as the two particles instantaneously move, the total force exerted on the first particle change at a non-zero rate and yet the total force exerted on the second does not. Thus the first particle that is closed with respect to force according to (DC) is not subject to a constant amount of force. Therefore, "the force conservation law" is violated, wherefore force is not a conserved quantity after all.

I think that (DC) makes it clear how a conservation law of a physical quantity Q is connected with a "continuity equation for Q". Let us reconsider the system S_1 and the rest of the world S_2 outside of S_1. Suppose that Q is conserved in each of S_1 and S_2. Then, for each system, if it is closed from outside with respect to Q then it possesses a constant amount of Q through time. According to (DC), this means that if $dQ_{S1}/dt \neq 0$ then $dQ_{S1}/dt = -dQ_{S2}/dt$ and if $dQ_{S2}/dt \neq 0$ then $dQ_{S1}/dt = -dQ_{S2}/dt$, where Q_{S1} is the amount of Q inside S_1 and Q_{S2} is the amount of Q inside S_2, i.e., outside S_1. Therefore, we have the result that Q is conserved in each of S_1 and S_2 iff, if $dQ_{S1}/dt \neq 0$ or $dQ_{S2}/dt \neq 0$, then $dQ_{S1}/dt = -dQ_{S2}/dt$. The right hand side of the biconditional is logically equivalent to "$dQ_{S1}/dt = -dQ_{S2}/dt$" since when both dQ_{S1}/dt and dQ_{S2}/dt are vanishing, $dQ_{S1}/dt = -dQ_{S1}/dt$. In short, Q is conserved in both S_1 and S_2 iff $dQ_{S1}/dt = -dQ_{S2}/dt$.

Suppose now that (CL) is true. Then Q is conserved in every Q-possessing system and, therefore, we can apply the above line of reasoning to every Q-possessing system. So we have the result that, for every Q-possessing system S, $dQ_{in}/dt = -dQ_{out}/dt$, where Q_{in} is the amount of Q inside S and Q_{out} is the amount of Q outside S. Hence we can say that if (CL) is true, then for every Q-possessing system, $dQ_{in}/dt = -dQ_{out}/dt$. On the other hand, suppose that for every Q-possessing system, $dQ_{in}/dt = -dQ_{out}/dt$. Then, for every Q-possessing system, if $dQ_{in}/dt \neq 0$ then $dQ_{in}/dt = -dQ_{out}/dt$. According to (DC), this means that Q is conserved in every Q-possessing system. Hence we can say that if for every Q-possessing system, $dQ_{in}/dt = -dQ_{out}/dt$, then (CL) is true. In consequence, (CL), i.e., the conservation law of Q, is true iff, for every Q-possessing system, $dQ_{in}/dt = -dQ_{out}/dt$.

Another claim I want to make is that "$dQ_{in}/dt = -dQ_{out}/dt$" can be taken as a "continuity equation for Q". Let us consider a well-known continuity equation, say, the continuity equation for electric charge density: when a system occupies a volume V enclosed by a surface A,

$$\int_A \mathbf{J} \cdot d\mathbf{a} = -\frac{d}{dt}\int_V \rho\, dv, \tag{6}$$

where $\mathbf{J}$ is the current density vector—the amount of electric charge moving through a unit volume in a unit time—and ρ is the charge density. The

volume integral of the charge density is the total amount of electric charge inside the system in question, C_{in}, for short. And, the left hand side of the equation means the instantaneous rate at which electric charge is leaving the enclosed volume; therefore, it means the instantaneous rate at which the electric charge outside the system changes. When we let C_{out} be the total amount of electric charge outside the system, we have,

$$\int_A \mathbf{J} \cdot d\mathbf{a} = dC_{out}/dt \tag{7}$$

Therefore, we have "$dC_{out}/dt = -dC_{in}/dt$". Thus we can understand the continuity equation for electric charge density as saying that $dC_{out}/dt = -dC_{in}/dt$. This naturally suggests that we can take "$dQ_{in}/dt = -dQ_{out}/dt$" as a "continuity equation for Q". So I conclude that (CL) is true iff, for every Q-possessing system, the continuity equation for Q holds true.

It should be noted that the continuity equation is equivalent to the statement that $<Q_{in} + Q_{out}>$ remains constant over time and that, whatever the system in question is, $<Q_{in} + Q_{out}>$ is the total amount of Q in the world as a whole. This means that for every Q-possessing system the continuity equation for Q holds true iff the total amount of Q in the world remains constant over time as a whole. In consequence, (CL) is true iff the total amount of Q in the world remains constant over time as a whole.

7. Reconsideration of Hitchcock's and Dowe's Definitions. According to the CQ theory of causation, causal interaction is analyzed in the following way (Dowe 2000, 90; Salmon 1997, 462):

CQI. A causal interaction is an intersection of world lines that involves exchange of a conserved quantity.

Elsewhere I have argued that, to block a serious counterexample to (CQI), "exchange" in (CQI) should be understood to be governed by a conservation law (Choi 2002, 115). Suppose now that a system causally interacts with the outside of it. Then, according to (CQI), there exists such a conserved quantity Q that $dQ_{in}/dt = -dQ_{out}/dt \neq 0$ since, as noted above, a conservation law of Q implies that if $dQ_{in}/dt \neq 0$ then $dQ_{in}/dt = -dQ_{out}/dt$. Therefore, there exists such a conserved quantity Q that the system is not closed from outside with respect to Q according to (DC). As a result, we can say that if a system causally interacts with the outside of it, then there is such a conserved quantity Q that the system is not closed from outside with respect to Q.

Suppose now that there is such a conserved quantity Q that a system is not closed with respect to Q. Then, according to (DC), the system is

such that $dQ_{in}/dt = -dQ_{out}/dt \neq 0$. Given the natural assumption that a system does not possess a constant amount of a conserved quantity through time only if the system intersects with the outside of it, we can infer that the system intersects with the outside of it involving exchange of Q. According to (CQI), this means that the system causally interacts with the outside of it. Therefore, we can say that, if there is such a conserved quantity Q that a system is not closed with respect to Q, then the system causally interacts with the outside of it.

In short, a system is closed from outside with respect to every conserved quantity according to (DC) iff it does not causally interact with the outside of it. In my opinion, when physicists say that a system is closed from outside simpliciter, what they have in mind is that the system is closed from outside with respect to every conserved quantity. This leads to the conclusion that a system is closed from outside simpliciter iff it does not engage in any causal interactions with the outside of it. This is exactly Hitchcock's definition! Thus (DC), together with the CQ theory of causation, explains a prima facie plausibility of Hitchcock's definition: it is extensionally correct as long as "a closed system simpliciter" means a closed system with respect to every conserved quantity.[18]

Let us now turn to Dowe's definition. We have seen that a system is not closed with respect to energy at a time t iff $dE_{in}/dt = -dE_{out}/dt \neq 0$ at t. I believe that the right hand side of this biconditional proposition can be taken as defining the notion of net flow of energy. To be specific, there is a net flow of energy into (out of) a system at a time t iff $dE_{in}/dt = -dE_{out}/dt > (<) 0$ at t. In general,

(8) For a conserved quantity Q, there is a net flow of Q into (out of) a system at a time t iff $dQ_{in}/dt = -dQ_{out}/dt > (<) 0$ at t.

Since the quantity Q is a conserved quantity, for every Q-possessing system, the continuity equation for Q holds true. This means that the right hand side of (8) can be shortened into "$dQ_{in}/dt > (<) 0$".

In my opinion, whatever details of the correct definition of flow of conserved quantity may be, the following statement must be true:

(9) For a conserved quantity Q, $dQ_{in}/dt =$ (the total amount of Q that flows from the outside of the system in question to the inside per a unit time) $-$ (the total amount of Q that flows from the inside of the system to the outside per a unit time).

(9) joins with (8) to imply that for a conserved quantity Q, there is a net flow of Q into (out of) a system at a time t iff the total amount of Q that

18. This idea was pointed out by Inkyo Chung.

flows from the outside of the system to the inside is larger (smaller) than the total amount of Q that flows from the inside of the system to the outside at t. This is just what I have proposed in developing Dowe's proposal in Section 4.

It is important to note that the notion of net flow of physical quantity presupposes that the quantity is a conserved quantity. Even though S_1's V_{cm} and S_2's V_{cm} increases in opposite directions at the same rate, we would not say that there is a net flow of V_{cm} from the outside of S_1 into the inside. This is because the metaphorical expression "flow" implies a kind of conservation. If so, (8) cannot be extended to physical quantities that are not conserved quantities.

It is remarkable that, according to (8) and (DC), for a conserved quantity Q, a system is closed with respect to Q at a time t iff there is no net flow of Q into or out of the system at t. This is exactly Dowe's definition. Thus, I admit that, as far as conserved quantities are concerned, Dowe's definition works. This explains why Dowe's definition, at first sight, seems to be plausible. Dowe's definition, however, does not amount to a definition of a closed system since, as noted above, it cannot be generalized to non-conserved quantities.

8. Conclusion. We have seen that we can provide such a definition of a closed system that CQ theorists can help themselves to (CL). Let us now consider exactly how CQ theorists can define the notion of conserved quantity by using (CL). Dowe (2000, 91) defines a conserved quantity as "any quantity that is governed by a conservation law." Given that (CL) is the exact formulation of conservation law, we can put Dowe's idea in a more precise way: a physical quantity Q is a conserved quantity iff it is a law of nature that for every Q-possessing system, if it is closed from outside with respect to Q, the total amount of Q possessed by it remains constant through time.[19]

Note that the definiens has subjunctive force because it is a law-statement. Hence, the definiens implies that a Q-possessing system that is not actually closed is such that if it were to be closed from outside with respect to Q, then the amount of Q possessed by it would remain constant. Therefore, it does not follow from the definition of conserved quantity that if no Q-possessing systems are actually closed with respect to Q, then Q is a conserved quantity; even though (CL) is true of Q, it does not follow

19. In Section 6, we found out that, according to (DC), (CL) is true iff the total amount of Q in the world remains constant over time as a whole. This suggests a simple and straightforward alternative way of defining a conserved quantity: a physical quantity is a conserved quantity iff it is a law of nature that the total amount of the quantity in the world remains constant over time as a whole. This point has been brought to my attention by an anonymous referee.

that Q is a conserved quantity. A physical quantity Q is a conserved quantity only if it is not only a true generalization but also a law of nature that every Q-possessing system that is closed with respect to Q possesses a constant amount of Q through time.

The CQ theory of causation analyzes causal process and interaction in terms of conserved quantities such as energy, electric charge, etc., which raises the question of exactly what such notions as conserved quantity, conservation law and closed system are. In the foregoing, we have seen that (DC) is an adequate definition of a closed system; and that (CL), together with (DC), provides the exact formulation of a conservation law that serves CQ theorist's purpose; and finally that we can answer what a conserved quantity is by relying on (CL). Thus it turns out that a number of problems that advocates of CQ theory of causation have been facing can be solved.

REFERENCES

Arya, Atam P. (1990), *Introduction to Classical Mechanics.* Boston: Allyn and Bacon.
Bigelow, John, Brian Ellis, and Richard Pargetter (1988), "Forces", *Philosophy of Science* 55: 614–630.
Choi, Sungho (2002), "Causation and Gerrymandered World Lines", *Philosophy of Science* 69: 105–117.
Desloge, Edward, and Robert Karch (1977), "Noether's Theorem in Classical Mechanics", *American Journal of Physics* 45: 336–339.
Dowe, Phil (1995), "What's Right and What's Wrong with Transference Theories", *Erkenntnis* 42: 363–374.
——— (2000), *Physical Causation.* Cambridge: Cambridge University Press.
Doughty, Noel (1988), *Lagrangian Interaction: An Introduction to the Relativistic Symmetry in Electrodynamics and Gravitation.* Reading, MA: Addison-Wesley.
Ehring, Douglas (1986), "The Transference Theory of Causation", *Synthese* 67: 249–258.
Fair, David (1979), "Causation and the Flow of Energy", *Erkenntnis* 14: 219–250.
Feynman, Richard, Robert Leighton, and Matthew Sands (1989), *The Feynman Lectures on Physics,* Vol. 1. Reading, MA: Addison-Wesley.
Goldstein, Herbert (1980), *Classical Mechanics.* Reading, MA: Addison-Wisely.
Hitchcock, Christopher R. (1995), "Discussion: Salmon on Explanatory Relevance", *Philosophy of Science* 62: 304–320.
Kittel, Charles, Water Knight, and Malvin Ruderman (1973), *Berkeley Physics Course Vol. 1: Mechanics,* 2d ed. New York: McGraw-Hill.
Marion, Jerry B. (1970), *Classical Dynamics of Particles and Systems,* 2d ed. New York: Academic Press.
McDaniel, Kristopher (2002), Review of *Physical Causation,* by Phil Dowe, *Erkenntnis* 56: 258–263.
Nilo, Bobillo-Ares (1988), "Noether's Theorem in Discrete Classical Mechanics", American Journal of Physics 56: 174–177.
Purcell, Edward (1985), *Berkeley Physics Course Vol. 2: Electricity and Magnetism,* 2d ed. New York: McGraw-Hill.
Schaffer, Jonathan (2001), Review of *Physical Causation,* by Phil Dowe, *British Journal for the Philosophy of Science* 52: 809–813.
Salmon, Wesley (1997), "Causality and Explanation: A Reply to Two Critiques", *Philosophy of Science* 64: 461–477.
Symon, Keith (1985), *Mechanics,* 3d ed. Reading, MA: Addison-Wesley.