Thought experiments and the scarcity of good theories

Sören Holst[1]

Summary

A theory relating to a certain aspect of reality can (minimally) be considered as a set of statements concerning that aspect. But not all sets of statements make sense; not all theories are consistent. In many fields, consistent theories turn out to be surprisingly rare. This fact can indeed be viewed as the basis for many activities of purely theoretical nature, not least theoretical physics. This article investigates the relevance of this perspective for thought experimentation, with a focus on physics. A thought experiment is here considered as an argument that draws upon a specific situation in order to reach a general conclusion. My claim is that the success of many thought experiments of a certain kind – namely deductive ones – can be better understood in light of the rareness of consistent theories. But also the other way around: that the rareness of consistent theories is strikingly illustrated in certain thought experiments.

Keywords: consistent theories, Descartes, epistemology, Huygens, Stevin, thought experiments, twin paradox

1. Introduction

Much has been written on the subject of thought experiments. Their historical role has been investigated (see *e.g.* Naylor, 1989; Norton, 1991; Palmieri, 2005), their use for pedagogical purposes has been discussed (*e.g.* Helm et al, 1985a and 1985b; Velentzas et al, 2007), and the question on whether and how they deliver new knowledge has been debated (*e.g.* Kuhn, 1977; Brown, 1991; Sorensen, 1992; Norton, 1996). In the present article I will focus on a slightly different aspect of this subject, one that I believe will shed further light also on these matters: I will discuss the relation between thought experimentation and possible theories.

Central to the discussion will be a basic claim or observation about theories and the way we think about them. Namely, I will argue that theories which are good in a certain sense are not as abundant as we tend to think; good theories are surprisingly rare. In other words, we tend to underestimate the impact of the constraints that any useful theory has to fulfil. This tendency of ours I dub our Obliviousness to Theoretical Constraints. I will shortly be more specific about the nature of these constraints and about what I mean by a good theory.

There are many thought experiments that make direct use of the “rareness” of theories. That is to say, they take advantage of the constraints on a theory in order to reach some kind of conclusion. Some thought experiments simply show that a proposed theory is not possible. In other cases they make use of the constraints to find an implication of a theory.

There is little new or surprising in these statements. That thought experiments often deal with theories – and that they often expose those theories to various constraints, such that internal consistency – is well known. Still, there are aspects of this relation between constraints on theories

---

[1] Physics Department, Stockholm University, 106 91 Stockholm, Sweden. Email: holst@fysik.su.se

and thought experiments that, in my opinion, have received less attention than they merit in the literature.

First, many thought experiments illustrate the rareness of good theories in particularly striking ways. Since this rareness, as I will argue, is highly non-intuitive, there is a unique and potentially important pedagogical value in such thought experiments: they make us aware of our Obliviousness to Theoretical Constraints.

Second, the perspective adopted here is relevant to the old issue on whether thought experiments actually conway any new knowledge, and, if so, *how* this knowledge is delivered. What is it that makes thought experiments work? I will argue that some thought experiments can be said to owe their success precisely to the fact that good theories are rare.

Before expanding on these claims, I will attempt to be more specific about the concepts involved.

To begin with, what do I mean by "a good theory"? Of course, a good scientific theory should be able to account for certain aspects of reality: it should produce some verifiable predictions. But in addition we expect a good theory to satisfy three other properties – irrespective of whether the theory concerns physics, economics, ethics or anything else. First, the theory should be *consistent*, *i.e.,* free from internal contradictions. Second, the theory should be *general*: the greater number of different phenomena or situations it covers the better. Third, the theory should be *simple*: it should be reducible to a small number of propositions, from which, in the ideal case, all other statements in the theory can be properly derived.

There exists a vast literature discussing the exact meaning of these or similar requirements for scientific theories. I will not here enter into that discussion. For my purposes, it is enough to note that whether or not it is possible to formulate these vague requirements precisely, these properties are among those valued most highly by people who spend their time trying to find new theories or to improve existing ones. That is to say, most researchers would find a suggested theory of little interest if it turned out to be inconsistent, too narrow in its scope of applicability, or too complicated to be of any practical use. Instead, they strive for theories which are consistent, broadly applicable, and whose propositions readily follow from a few basic assumptions.

Obliviousness to Theoretical Constraints is the claim that such good theories are rare, or more precisely, *surprisingly* rare. However, the concept of the rareness of good theories should not be taken too literally – it would probably be impossible to quantify it precisely. Rather, what I mean is that good theories are rare in a *practical* sense: bring together some seemingly sound general statements concerning a certain aspect of reality and these will most probably turn out to be inconsistent, often in some unexpected way. There are many examples of this in the history of physics, some of which we will discuss below.

One should also keep in mind that Obliviousness to Theoretical Constraints is primarily a statement about our psychology, rather than about theories themselves: Good theories are rarer *than we tend to think.* Why should this be true?

When we try to understand the phenomena around us, or when we evaluate various explanations, our first criteria is that a suggested explanation should fit the observational data at hand. More seldom – as long as we are not doing active theoretical research within the subject – do we ask ourselves whether the suggested explanation can be accomodated within a consistent theoretical framework. This may well be due to evolutionary causes: in order to survive – in the jungle or on the savannah, as well as in a big city – it is presumably more important to have working practical skills and strategies for specific situations, than to be able to devise a broad theoretical framework into which these strategies would fit. Thus, as our brains have not primarily evolved to reason abstractly, we often underestimate the difficulties that obtain in finding consistent theoretical frameworks that can house our different beliefs and hypotheses about the world around us. In other words, it is hardly surprising that our intuition for what is required to construct a consistent and general theory is not so well developed. It might well be the case that good theories are rarer than we tend to think. That this is indeed the case, I will attempt to demonstrate in the following.

As noted, besides capturing the relevant experimental data, a good theory should also be consistent, general, and simple. Indeed, it is precisely the rareness of such theories that provides the motivation for and force behind most kinds of theoretical investigations. This is perhaps most obvious in the case of physics. Modern theoretical physics rests on the fact that consistent theories that cover all experimental data are rare. In fact, today *no* such theory is known. There does exist a theory thought to describe all gravitational phenomena (*i.e.*, general relativity) and another thought to describe all microscopic phenomena (*i.e.*, quantum physics). But there is no theory today that can consistently encompass both these types of phenomena.

These surprisingly strong constraints on theories, then, is the reason that it is meaningful for physicists to spend their time searching for an ultimate theory – a theory of everything – even though the relevant energy scales for such a theory are well beyond those probed in real experiments. And this is also the reason that entirely new and unexpected phenomena are sometimes predicted by theoreticians. Did there, in fact, exist lots of good theories – that is, if most sets of statements about the world were internally consistent – there would be very little to gain from theoretical investigations; the only way to make progress in physics would be by carrying out new experiments.

Thus, the above suggestion that many thought experiments work as well as they do because of the paucity of good theories, should be understood as a special case of a more general claim: that the utility of *most* theoretical endeavors depends heavily on the difficulty of finding consistent and general sets of laws. This statement is, I think, uncontroversial. What I want to point out here is merely that the rareness of good theories becomes particularly apparent in the case of thought experimentation. Since we have no intuitive grasp of this rareness – since we are oblivious to the strength of the theoretical constraints – we tend to be surprised at what can be accomplished as a result of it. This has even led some philosophers to argue that the success of thought experimentation requires a particular explanation – an epistemology that goes beyond that of other forms of theoretical reasoning. (See *e.g.* Brown (1991) and Sorensen (1992) – I will return to these issues in the last section.) I contend that we should refrain from giving thought experiments such a distinct epistemological status.

In the following section, I propose a definition of thought experiments that will facilitate the subsequent discussion, and help to elucidate how and when thought experiments make use of the constraints on theories. In section 3, which constitutes the bulk of this article, the main ideas are explained and illustrated by means of some well-known and some less well-known thought experiments. The closing section 4 contains a short discussion and summary.

2. The thought experiment – a link from the specific to the general

In spite of much activity in this field, there is not yet a generally accepted definition of what counts as a thought experiment. Each author uses his or her own definition, adapted to the particular goals that he or she wishes to attain.[2] There is nothing necessarily wrong with this; if a definition serves the purpose of clarifying a particular point or if it can provide a subject with a new perspective, that could suffice to motivate it.

On the other hand, it would clearly be preferable to have one generally accepted definition. For such a definition to be meaningful, it must somehow capture the intuitive idea of a thought experiment. It should be narrow enough not to include all sorts of reasoning, but wide enough to be applicable in all the different fields in which thought experiments occur. And if there is such a thing as an "essence" of thought experimentation – some aspect that makes thought experiments unique among all different methods of reasoning or among all ways of aquiring knowledge – that trait should be captured by the definition. I will now discuss a definition that I think meets these requirements. The definition I propose is not new; rather, it is a minor elaboration of the criteria for thought experiments suggested by Norton (1991). These criteria, although seldom explicitly stated, tend to hover in the background in most discussions on thought experimentation.

To begin with, let us note that in order to constitute a thought experiment, an argument or piece of reasoning should comprise the following three ingredients. It should

(1) be hypothetical,
(2) involve a specific situation,
(3) lead to a result.

The first requirement above means that a thought experiment is something thought or described, rather than actually performed. Some thought experiments have, in fact, been carried out as real experiments. Other could perhaps be carried out in the future. But the results of such potential

---

[2] For an overview of different definitions, see, *e.g.*, Galili (2009).

realizations should not affect the conclusions to be drawn; a thought experiment does not *depend* upon its possible realization.

There are, however, many hypothetical activities that do not count as thought experiments. A piece of abstract or mathematical reasoning, for example. Thus, the second requirement above tells us that a thought experiment should involve a specific, concrete situation. This situation need not be realistic; it suffices that one is able to imagine the situation. But a purely abstract manipulation of mathematical symbols does not constitute a thought experiment.

The third requirement is that a thought experiment should lead to some kind of result. Mere pondering over a concrete situation does not count. The reasoning must generate some insights, lead to a conclusion, or at least produce a shift in perspective.

So a thought experiment involves a hypothetical consideration of a concrete situation that leads to a result of some kind. But this characterization is clearly far too inclusive. For example, consider the solution of a physics problem. Say that I want to calculate how far I can throw a ball, given my strength, the weight of the ball, and so on. I set up the equations, solve them, and get the result. This process is hypothetical (the answer obtained does not depend on me actually throwing the ball), it concerns a specific and concrete situation (me throwing the ball), and it leads to a result (the distance that I am able to throw the ball). Yet, we would never call this a thought experiment.

Consider another example. I daydream about eating a steak. This daydream makes me decide to visit a restaurant and order a steak. Again, the above requirements are fulfilled. It is hypothetical (as are the contents of all dreams), it concerns a specific situation (me eating a steak), and it leads to a result (my decision to go to a restaurant). But again, this is no thought experiment.

In order to expand the above criteria into a definition that excludes these and similar cases, let us first note that in a true thought experiment, the specific situation in itself leads to some kind of conclusion, which, however, is *not* the same as the result of the thought experiment. This conclusion – which I will call *the specific conclusion* – concerns only the specific situation. The final *result* of a thought experiment is always more ambitious than that: it goes beyond the specific situation and establishes a new principle or a new perspective, applicable to a much wider class of situations. That this is indeed the case will become clear later when we discuss some real thought experiments in detail.

But let me here simply remind the reader of some famous thought experiments. Consider Galileo's falling bodies, Schrödinger's cat or Putnam's twin-earth. These arguments were not invented in order to explain something about two particular falling bodies, a specific cat in a box, or an imaginary planet (almost) identical to the earth, but rather to support general statements about free fall, quantum mechanics, and the meaning of concepts, respectively. I claim that *the most important feature that turns these arguments into thought experiments is the contrast between the specific nature of the concrete situations to which they refer, on the one hand, and the generality of the* *insights they try to mediate, on the other.*

Note that this property is lacking in the two counter-examples mentioned above. A solution to a physics problem does indeed lead to a result. But this result is a specific conclusion – it concerns only the situation to which the solution was applied. There is usually no general insight to be found in a solution to a particular physics problem – and if there is, then the solution deserves to be dubbed a thought experiment. As for my daydream about the steak, it could be claimed that here there *is* a result that differs from the specific conclusion: the specific conclusion could be my insight that I want to eat a steak, the result that I decide to go to the restaurant. But in this case it is hard to maintain that my decision to go to the restaurant is in some way *more general* than the specific conclusion. Rather, it is *different*. On the other hand, suppose that the outcome of my daydream was not a decision to visit a restaurant, but rather to begin to eat meat again after a period of strict vegetarianism. Such a decision is, in a certain sense, more general than the insight that I want to eat steak. This latter version of my daydream should, I think, be classified as a thought experiment. Even though of a rather uninteresting and private character, these mental events together contain the crucial element of thought experimentation: a consideration of a special case that leads to a general conclusion.

I now state what I believe to be a useful definition of thought experiments. It is simply an elaboration of the three criteria above, taking into account the points just made. A thought experiment

(1) has a hypothetical character;
(2) involves a concrete situation, that via (more or less) explicit reasoning leads to a specific conclusion concerning that particular situation;
(3) leads to a result, that in some way is more general than the specific conclusion.

It might be objected that this definition is still too broad, that it still includes too much. One way to narrow it further would be to require that the reasoning leading to the specific conclusion should be purely deductive. But this, I think, would be a mistake. Although this indeed is most often the case for thought experiments within physics, such a requirement would exclude many pieces of reasoning which are usually referred to as thought experiments in other fields. Thus, for example, thought experiments in the field of ethics usually do not obtain their specific conclusion by means of deductive reasoning, but more often by appealing to the intuitions of the reader. This fact should not, I think, disqualify them.

As pointed out – albeit in different words – by Norton (1991) the relation between the specific conclusion (property (2)) and the general result (property (3)) can be of two types. Namely, the step from specific conclusion to general result can be deductive or inductive. Thus we may speak of *deductive thought experiments* and *inductive thought experiments*, although many of them turn out to be somewhere in between; the step to the general result often involves elements of both kinds of reasoning.

The most obvious examples of deductive thought experiments are those that take the form of a *reductio ad absurdum*. In these cases, considerations of the specific situation lead to some sort of impossibility. The final result then is that at least one of the assumptions must be wrong (often a theory or a general hypothesis). But not all deductive thought experiments are *reductios*. Sometimes, once the specific conclusion has been reached, it can be seen that the same deductive steps that led to this conclusion could equally well be applied to a much wider class of situations, thus giving rise to the general result. We will see several examples of such thought experiments in the next section.

Thought experiments at the other end of the spectrum, the inductive ones, have a different character. As noted above, here the general result follows from the specific conclusion by means of an inductive step. That is, one is led to believe that the specific situation is in some way representative also for other situations, and thus that which seems to be clear under those particular circumstances, must also hold true in the general case. Inductive thought experiments are often used as a kind of illustration of a general principle or idea. In those cases the specific situation usually acts as a reminder, directing the readers attention to earlier experiences that are relevant for whatever the thought experimenter wants to argue.

A typical example of an inductive thought experiment is Galileo's famous assertion of the Relativity Principle. Galileo asks his readers to imagine a ship's cabin containing some unusual furnishings: an aquarium with fish swimming to and fro, a dripping bucket, burning incense among other things. He then notes that observing these artefacts is of no use in determining the speed of the ship; their behaviour does not tell us whether the ship is at anchor in the harbor or on its way over the sea. That is the specific conclusion. A crucial inductive step is supposed to bring out the general result of the thought experiment: that it is not possible in *any* way for an observer inside a ship's cabin – or in similar situations in which the possibility of looking out is lacking – to know whether he or she is actually moving. Only relative speeds matter.

The distinction between inductive and deductive thought experiments is not, however, a sharp one. Most thought experiments make use of both kinds of reasoning, and in some cases they can be considered either as deductive or as inductive depending on the exact phrasing of their assumptions.

Historically, both kinds of thought experiments have proven of great utility for science. But in the present paper, we shall primarily concern ourselves with those on the deductive side of the spectrum. It is, I maintain, mainly deductive thought experiments that illustrate and are a symptom of the scarcity of good theories.

3. Reducing the number of possible theories

The total number of possible statements that can be made about reality is immense. The kind of statements that are of most interest to science are general ones – those that apply to many situations or phenomena. The more general the better. But as a statement becomes more general – or is interpreted as applying to a wider class of situations – it also becomes more vulnerable. In particular, a general statement runs a larger risk of colliding with other statements that purport to apply to the same situation. Not all statements go well together.

A consistent theory should comprise a set of statements that do not contradict each other. Suppose that we are confronted with a set of statements about the world. How can we proceed in order to find out whether this set is consistent or not? One method would be to formalize the statements and then to proceed via mathematical or abstract proofs. Another way is to apply the statements to various concrete situations and investigate whether they then give rise to consistent predictions. In other words, we can perform thought experiments. This second method is less reliable since, no matter the number of concrete situations with which we confront the set of statements, we can never be sure that there are no inconsistencies that we have not yet found. But there are also advantages to this method. Among them is the fact that our brains much more readily grasp concrete reasoning than abstract theory. And due to their ability to reduce the number of possible sets of statements, thought experiments can often pave the way for more formal reasoning and theory building.

Also, as I hope to show, the efficacy of thought experiments in reducing the number of possible theories can often be quite surprising. It may thus be tempting to assign some particular epistemological status to thought experiments. I maintain, however, that we should rather consider this surprise as a symptom of our Obliviousness to Theoretical Constraints.

I will now discuss some examples that I think illustrate, in particularly striking ways, the paucity of good physical theories. These examples will, I hope, also help to clarify how thought experiments actually accomplish the reduction in the number of possible theories.

Stevin's inclined planes

Let us begin with the well-known thought experiment of Simon Stevin, in which he asks how much force is required to balance a body on a frictionless inclined plane (see *e.g.* Mach (1893)).

In order to determine this force, Stevin imagines a triangular block or prism over which a chain of balls are hung, see figure 1. The prism is such that two of its sides form inclined planes of different slopes and lengths. A certain number of the balls in the chain rest on the longer side and a different number on the shorter side. The rest of the balls form a symmetric arc beneath the block. Will the chain of balls thus placed remain at rest, or will it start to rotate? Well, *if* it began to rotate one way or the other, it would have to continue to move around in the same direction perpetually. That, Stevin knows intuitively, will not happen. So the chain of balls remains at rest over the prism. Next, he imagines that the part of the chain that forms the arc beneath the prism is removed. This arc, because of its symmetry, pulls the chain equally in both directions. Hence, after it has been removed, the remaining balls must remain at rest.

In the specific case depicted in figure 1, this means that the three balls resting on the shorter side of the block will balance the five balls on the longer side. That is the specific conclusion in this thought experiment. But the statement is easily generalized by noting that the length of each side is proportional to the number of balls that are resting on it, and therefore to their total mass. Hence, in order to balance a certain mass on an inclined plane one would need a balancing mass on a second plane (of the same height) whose ratio to the first mass is the ratio between the lengths of the two planes. If the second plane is vertical – so that the balancing weight is actually hanging freely – this means that the balancing weight would be (in modern langage) the first mass times the sine of the angle of inclination. This is the general result of Stevin's thought experiment.

Without knowing anything about Newtonian mechanics – which were not to be invented for half a century – Stevin is able to find the correct expression for the balancing force. How is this possible? Clearly, the crucial step in the reasoning is the insight that the chain of balls, when hung over the prism, will remain at rest. This follows from the impossibility of perpetual motion. Thus, what the thought experiment actually shows is *not* that the balancing force is what it is, namely the weight times the sine of the inclination angle. Rather, what it does show is that there is a connection between the following two statements:

(1) There exists no *perpetuum mobile*.
(2) The force required to balance a weight on an inclined plane is the weight times the sine of the angle of inclination.

Stevin's argument shows that in a world where the first statement is true, the second also must be true. In other words, he shows that any theory in which there is no *perpetuum mobile*, but where the balancing force takes another form, is not possible. All such theories are excluded by the thought experiment.[3]

Intuitively, we would not guess that there were any connection between statements (1) and (2). Or, put in another way, we would not expect that the range of possible theories was restricted in just this manner. Stevin's thought experiment wakens our surprise because consistent theories are rarer than we tend to think.

Descartes' laws

Descartes wanted to describe the whole universe as a huge machine, in which all phenomena were to be accounted for in terms of matter that pushes and pulls on other matter. Thus, for example, Descartes could not accept the view that gravity constituted some kind of action at a distance.

---

[3] As pointed out to me by one anonymous referee, Stevin himself did not support his reasoning on the general claim of the impossibility of a *perpetuum mobile*. Rather, he assumed that the particular object envisaged in the argument – a triangular block with chain – would not constitute a *perpetuum mobile* (Kühne, 2005). This does not affect the analysis here in any essential way. Statement (1) can just as well be replaced by Stevin's original less ambitious claim.

Hence he introduced his theory of vortices, which postulated that the universe was filled with a kind of invisible matter constantly whirling around the sun and the planets. In this way he claimed to be able to account for planetary motion as well as for gravitational phenomena near the surface of the earth.

Since a central element in Descartes' picture of the world's machinery is matter pushing on other matter, he was interested in the outcome of collisions between pairs of bodies. He attempted to formulate one law for each possible type of collision. He considered collisions between equal bodies, collisions between non-equal bodies, collisions in which the bodies move towards each other before they collide, and collisions in which one of the bodies starts from rest. He ended up with seven laws (Descartes, 1644). As we will see, these laws provide a particularly clear example of how difficult it can be to find a theory for a given class of phenomena which meet the requirements of consistency and generality.

A pictorial description of Descartes' laws is shown in figure 2; see the caption for details.

Descartes didn't invent his seven laws out of nothing. On the contrary, they are all consequences of two principles that he considers to be more fundamental. The first of these is the conservation of what Descartes called the *quantity of motion*. This quantity is defined as the mass of a body times its speed. At first glance, this seems to be identical to the modern concept of momentum. However, there is one crucial difference: Descartes' quantity of motion is a *scalar* rather than a *vector* quantity – it does not take the *direction* of motion into account. Hence, in order to obtain the total quantity of motion for two bodies traveling in opposite directions, Descartes simply adds their individual quantities of motion, rather than subtracting the one from the other, as we would do today when calculating their total momentum.

So the first of the two principles that Descartes uses to construct his collision laws is that the quantity of motion is conserved: the quantities of motion for the colliding objects should add up to the same value before and after the collision. The second principle is that when the two colliding objects have different “strengths” it is always the “strongest” of them that dictates the motion after the collision. An object can be stronger than another for two reasons: either it is heavier or it is moving more quickly. Thus, for example, in the second law, the heavier object continues to move in the same direction (and with the same speed) as before the collision, pushing the lighter object ahead. And in the third law, after the collision both objects move together in the direction of the faster object.

Unfortunately, most of Descartes' laws are wrong. But the question that interests us here is whether they *could* have been right. Is it possible to imagine a world where these laws are true? The answer is affirmative. But in such a world certain other, seemingly plausible, principles must fail to hold.

Leibniz criticized Descartes' laws on the grounds that they violate a kind of Continuity Principle. Here is how Leibniz describes this principle:

... when two instances or data approach each other continuously, so that one at last passes over into the other, it is necessary for their consequences or results ... to do so also. (Leibniz, 1687)

That is: as two causes become more and more equal, their consequences also have to become more and more equal, until they continuously become identical.

It is not difficult to see that Descartes' laws violate this principle. To show this, Leibniz focuses on the first and the second laws. He imagines two equal bodies heading towards each other with the same speed. According to Descartes' first law both bodies will reverse their direction of motion in the collision, but keep their speed. But suppose now that we add a tiny mass to one of the bodies and again allow them to collide in a head-on collision. Now, since one body is slightly larger than the other, it is no longer the first but the second of Descartes' laws that is applicable. But according to this law something completely different will happen: after the collision the two bodies will travel together in the same direction and with the same speed! This simple thought experiment thus shows that Descartes' laws cannot be true in a world where Leibniz' Continuity Principle holds.[4]

More importantly, Descartes' laws also violate the Relativity Principle. We have already met this principle in Galileo's famous formulation: it is not possible for someone inside a ship's cabin to know the speed of the ship without looking out. More modern formulations introduce the concept of *inertial frame*, that is, a non-accelerating coordinate frame. The Relativity Principle then states that *all inertial frames are equivalent*, that is, the laws of physics take the same form in all such frames. The outcomes of experiments do not depend on absolute rectilinear speed, in the same sense that they do not depend on, for example, spatial orientation.

To see how the Relativity Principle conflicts with Descartes' laws, consider a small ball hitting a larger one at rest. According to the forth collision law, the large ball will remain at rest in the collision, while the smaller one will recoil with its original speed. Now, according to the Relativity Principle, we can assume that this collision takes place inside a carriage moving at a certain speed. Suppose that the speed of the carriage is the same as that of the small ball inside the carriage before the collision, but in the opposite direction. This means that, according to someone outside the carriage, the small ball is at rest before the collision. To this person, it is instead the large ball that is moving. But then, it is not the forth law, but rather the fifth, that is to be applied. According to the fifth law our observer outside the carriage will see both balls move together after the collision. As just noted, this was not what would happen according to a passenger in the carriage. The

---

[4] As support for his principle Leibniz refer to nothing less than the immense wisdom of God. The title of Leibniz' criticism of Descartes' laws is characteristic: *Letter of Mr. Leibniz on a General Principle Useful in Explaining the Laws of Nature through a Consideration of the Divine Wisdom.* Despite the wisdom of God, however, Leibniz' principle does not hold in the actual world. An easy way to see this is to consider a stone thrown at a window. As long as the stone is thrown with only slight force, the window may stay unbroken. But as the force is continuously increased, the window will at some point suddenly break. A small difference in the cause does not, in this case, give rise to a small difference in the result. Even more flagrant violations of Leibniz' principle are abundant in modern chaos theory.

contradiction shows that Descartes' laws are not consistent with the Relativity Principle.

In fact, had Descartes been aware of the Relativity Principle and accepted it, he would have found it sufficient to formulate two laws, rather than seven. This is because the initial situations in laws one, three and six all represent one and the same case according to the Relativity Principle, just viewed from different reference frames. Likewise, laws two, four, five, and seven all concern the same situation.

As noted above, Descartes considered his seven laws as expressions of, first, the conservation of his quantity of motion and, second, the idea that it is the strongest of two bodies that determines the outcome in a collision. Our conclusion is that these two principles are inconsistent with both Leibniz' Continuity Principle and the Relativity Principle. These inconsistencies are intuitively far from obvious. But they are made apparent through the thought experiments, namely, by letting the principles confront each other under the particular circumstances envisaged in the specific situations.

Huygens' colliding spheres

Christiaan Huygens (1629 – 1695) turned the above argument based upon the Relativity Principle around: making use of a series of thought experiments he *derived* collision laws consistent with this principle (see *e.g.* Dugas, 1988). His reasoning makes the unexpected strength in the constraints imposed by the Relativity Principle even more clear.

Huygens starts out by considering a symmetric collision between two spheres of the same weight. He assumes that it is possible to arrange things such that if two such spheres move towards one another with equal speeds along a given line, then, after the collision, they will move away from each other with the same speeds and along the same line. Today we would characterize this as a symmetric, one-dimensional, elastic collision. Huygens then asks himself about the outcome of a different type of collision between these spheres: What would happen if one of the spheres was at rest when hit by the other sphere?

The Relativity Principle allows Huygens to consider the first symmetric collision taking place aboard a boat, slowly floating down a river. A man is standing on the shore, observing the collision on the boat. What does he see? Let us assume that the boat is moving with the same speed relative to the shore as the balls are moving relative to the boat. In this case one of the balls (the foremost) will be at rest relative to the shore before the collision, while the other will move at twice the speed of the boat. However, after the collision these speeds will be interchanged: The ball nearest the bow will now move at twice the speed of the boat while the other one will be seen to be at rest relative to the shore. This is the specific conclusion: when the sphere that is moving (as seen from the shore) hits the other sphere, it comes to rest, while the sphere that is hit starts to move at the same speed. Since the specifics of the situation clearly is irrelevant (the boat, the river et cetera) the general result follows immediately: this is the outcome in general of an elastic collision between two equal bodies where one initially is at rest.

The result can be generalized in various ways. By considering cases in which the boat moves at different speeds relative to the observer on the shore, Huygens draws the conclusion that in a (one-dimensional and elastic) collision between two equal bodies, their speeds are *always* interchanged. He then proceeds to consider collisions between bodies of different weight, using the same specific situation in which the collisions take place aboard a boat in a river and with an observer on the shore. Also in these cases, Huygens is able to reach some interesting conclusions, such as that the relative speed between two colliding objects is conserved (provided that the collision is one-dimensional and elastic).

Today, these conclusions may not strike us as very surprising in themselves. They are, after all, straightforward consequences of the conservation of momentum. What is surprising, though, to most of us is, I think, the close connection between these results and the Relativity Principle. When we consider Galileo's cabin, and the fact that a passenger in there would be unable to determine if the ship was moving or at rest, it does not occur to us that so seemingly innocent an idea should have any consequences for the outcomes of collisions.

In case this does not convince the reader of the surprising predictive power of the constraints imposed on physics by the Relativity Principle and their importance in thought experiments, let me provide a similar example, but one where the final result is less obvious.[5]

Consider two balls, one of them very much heavier than the other. Indeed, we shall assume that the heavier of the two balls is so much heavier that its speed is essentially unaffected when hit by the smaller one. We shall also assume that the two balls collide elastically, and, as above, consider one-dimensional collisions only. This means that if the heavy ball is at rest when hit by the smaller one, then, after the collision, the heavy ball will remain at rest, while the small ball will move with the same speed as before, but in the opposite direction. (This, to a reasonably good approximation, is what happens when, for example, a tennis ball hits a basketball.)

Now, let us ask for the outcome of another type of collision between these two balls, namely one in which the two balls initially move towards each other at the same speed and then hit each other in a head-on collision. What is the speed of the small ball afterwards? Again we make use of Huygens' river and boat. As before, we arrange things such that the collision as seen on the boat is the familiar one, whereas the collision as seen from the shore is the one with unknown outcome. To be more precise, let us say that the boat is moving at speed v down the river. The heavy ball is at rest in the boat, and the small ball is initially moving towards it from the bow at speed $2v$ (relative to the boat). After the collision the heavy ball is still at rest while the small one has reversed its direction of motion, and is now moving with the same speed $2v$ towards the bow in accordance with our assumptions.

---

[5] The following example has been used, for example, by Mermin (2005) to illustrate the power of the Relativity Principle.

As seen from the shore, prior to the collision the balls are moving towards each other at the same speed v. What does someone on the shore see after the collision? To find out we just have to translate the motion aboard the boat after the collision to motion relative to the shore. The heavy ball is just moving with the boat, so its speed relative to the shore is just v, the same speed as before the collision. But since the small ball relative to the boat is moving at $2v$ towards its bow, and since the boat itself is moving at speed v, the resulting speed of the small ball relative to the shore must be $3v$. We can immediately formulate the general conclusion: in a head-on collision between two balls of very different weight which approach each other at the same speed, the smaller of them will reverse its direction and increase its speed by a factor of three![6] As is demonstrated in the thought experiment, this surprising result is a necessary and immediate consequence of the Relativity Principle.

To summarize, then, in a world in which the Relativity Principle holds, phenomena must obey certain rules; not all statements about collisions are compatible with the Relativity Principle. Thought experiments of the kind considered by Huygens provide an efficient means to sort out which statements are compatible and which are not.

Bondi's account of the twin paradox

Nowhere in physics is the deductive force of thought experimentation as explicit and as obvious as in Special Relativity. There are several well-known examples of this: Einstein's thought experiment with the train and the two flashes of lightning to demonstrate the breakdown of absolute simultaneity (Einstein, 1920); the light clock consisting of a light pulse that is reflected to and fro between two mirrors used to derive the time dilation effect (Lewis *et al*, 1909); Einstein's considerations of a photon-emitting box to show the relation between energy content and inertial mass (Einstein, 1905).

The essential assumption in these and in most other thought experiments concerning Special Relativity is what might be called the *Extended Relativity Principle*. This comprises the idea that the Relativity Principle should encompass all electro-magnetic phenomena, and in particular that of light. That is to say, no experiment involving electro-magnetism, light, or any other phenomenon whatsoever, can be used to establish a preferred inertial frame. In particular, any observer measuring the speed of light in vacuum will obtain one and the same value, c, irrespective of his or her own speed.

The Extended Relativity Principle turns out to have unexpected and far-reaching consequences.

---

[6] A vivid demonstration of this fact can be obtained using a basketball and a tennis ball. Drop the basketball and shortly after the tennisball right above it. The basketball will hit the ground first, and bounce upwards. When it meets the tennisball, which is moving downwards, a collision of the kind considered here ensues. The tennisball will bounce up with an unexpected force. (Ideally it will reach a height nine times the height from which the balls initially were dropped, as a simple consideration of the conservation of energy shows.)

These can be found by abstract reasoning, using mathematical language and physical modeling. But an alternative way to discover them is to apply the principle directly to various concrete situations – like those involving flashes of lightning hitting a train, or a pulse of light captured between two mirrors, or a photon-emitting box. In other words, we can determine the consequences of the Extended Relativity Principle by performing thought experiments. This will tell us which previously accepted statements about physics must be abandoned, and what new statements we have to accept – provided that we accept the principle.

Here I turn to one example which, I believe, particularly well illustrates these points. This thought experiment, first discussed by Hermann Bondi (1962), reveals in an unusually lucid way the very intimate connection between the Extended Relativity Principle and the fact that our aging depends on how we move. The whole argument actually comprises two steps, where each step constitutes a thought experiment of its own. In the first step Bondi establishes a basic fact about pulses of light and the intervals between their emission and reception. In the second step he shows how this basic fact leads directly to the fact that two observers in a classical twin paradox situation age at different rates. Since the thought experiments involved are less well known than those referred to so far, my exposition here will be somewhat more detailed.

Consider two observers, receding from each other at some constant speed. One of the observers sends out two pulses of light, the second pulse at a certain time interval t after the first one. The two pulses are then received by the other observer. Since this observer is moving away from the one that emits the signals, she will not receive the two pulses at the same interval t at which they were sent. Let us say that she receives them at an interval of qt. The factor q is the so-called Doppler factor. If the observers are receding from each other, as in this case, then $q>1$. The Doppler factor q can only be a function of the relative speed v between the observers, and not of their individual speeds, since, according to the Extended Relativity Principle, these latter quantities have no meaning. It is not difficult to derive the exact form of the function $q(v)$, but we have no use for that here. The only thing that we need to know is that if one observer sends two signals, separated by an interval t, to another observer, then the signals will be received at an interval qt, where q is a function only of the relative speed of the two observers.

But suppose that the two observers were approaching each other instead. Again the time interval that they measure between the pulses will differ. Say that the observer who sends the signals measures a certain time t' between them. Then the one receiving them will measure time rt' between them, where r is another Doppler factor, again depending only upon the relative speed between the observers. What is the relation between the "receding" Doppler factor q and the "approaching" Doppler factor r? That is, what is the relation between the Doppler factor that is relevant for two observers that are moving away from each other at a given speed v, and that relevant for two observers that are approaching each other at the same speed v?

The question is easily answered by means of a thought experiment. Consider three observers, A, B

and *C*. *A* and *C* are at rest relative to each other, but well separated in space. *B* is moving at speed v from *A* to *C*. The situation is depicted in the spacetime diagram in figure 3. *A* sends out a pair of light pulses to *B*, separated by a time interval t according to himself. Observer *B* receives the signals at the time interval $q_v t$. But as she receives these pulses, she immediately sends out new pulses to *C*. Alternatively, we can consider *B*'s pulses as the same as *A*'s – observer *B* simply notes their arrival times and let them pass her. *B* is approaching *C* at the same speed v that she is receding from *A*. The relevant Doppler factor between *B* and *C* therefore is r_v. Hence the two pulses reach *C* with time interval $r_v q_v t$. On the other hand, *C* is at rest relative to *A*. He must therefore receive the pulses at the same interval at which they were sent out, that is, t. Hence, $r_v q_v t = t$. Which means that

$$r_v = \frac{1}{q_v}$$

This is the specific conclusion in this thought experiment: the Doppler factor relevant for *B* and *C* is the inverse of that relevant for *A* and *B* in this particular situation. But, unless we think that this result somehow depends on the specifics of the situation – for example, the fact that our three observers are lined up in this particular way, or the fact that they are human, or the fact that there are precisely two pulses emitted – this conclusion must hold generally. The result of the thought experiment is that the Doppler factor for a source of light which is approaching at a given speed v is the inverse of that for a source that is receding at the same speed v.

Having established this fact, Bondi now goes on to consider a variant of the classical twin paradox. One of a pair of twins – let us call him Castor – departs on a long space trip, while his brother – Pollux – stays on earth. Castor travels at a certain high and constant speed v straight out into space. After a given time he turns around and returns to earth at the same speed v. During his voyage Castor continuously sends light pulses back to earth, as greetings to his brother. The situation is depicted in a spacetime diagram in figure 4. Castor sends out the light pulses at some regular time interval t. Let us say that before he turns back he has sent ten such pulses to earth. On his return voyage he has sent out ten more pulses, all separated by the same interval t. In total, then, during the whole trip, Castor has sent out twenty pulses all separated in time by the interval t. Thus, he has aged $20t$ during his absence from earth.

Pollux has, of course, received all twenty greetings. But due to the Doppler effect, they have not reached him at equal intervals during his brothers absence. The first ten reached him at intervals of $q_v t$; the last ten reached him at intervals of $r_v t = (1/q_v)t$. By adding up these twenty intervals, we can calculate how much Pollux has aged:

$$10 q_v t + \frac{10t}{q_v}$$

How does the amount Pollux has aged compare to that of his twin Castor? The ratio between these respective amounts is

$$\frac{10q_v t \;+\; 10t/q_v}{20t} = \frac{q_v + 1/q_v}{2}$$

We don't know the value of q_v or exactly how it depends on v. But that doesn't matter. As is easily shown[7], the numerator ($q_v + 1/q_v$) is always larger than 2 (unless q_v=1, which is only the case if v=0). Hence Pollux has aged more than Castor during Castor's trip. This is the specific conclusion of this thought experiment. The general result is that there is nothing absolute or universal about the passage of time. We age differently depending on our motion through spacetime.

At first glance, this result may seem to contain too much. How can we reach such a drastic conclusion concerning time and aging when we hardly have assumed anything at all? The crucial assumption in the thought experiment is, of course, the Extended Relativity Principle.

What this thought experiment reveals, then, is the intimate and unexpected relation between the Extended Relativity Principle and the fact that time runs differently for different observers. To discover this fact – usually referred to as the time dilation effect – we don't have to take the usual detour over the Lorentz transformations. We don't even have to discuss relative simultaneity. We just have to confront a single aspect of the Doppler effect with the usual twin paradox situation. In this thought experiment, we see immediately that as soon as we extend the Principle of Relativity to cover situations involving light, there is no escape from the conclusion that time runs differently depending on how we move.

Note also that the Extended Relativity Principle creeps into the argument very discreetly: it only enters in the assumption that the Doppler factor depends solely upon the relative speed between source and receiver. Most of us do not expect such a seemingly innocent idea about light signaling to have consequences for other parts of physics. And we certainly do not expect it to have far-reaching consequences for something as seemingly unconnected as aging.

This thought experiment then, like the others above, constitutes a powerful illustration of our Obliviousness to Theoretical Constraints. Considerations of how these thought experiments actually work, or how it can be that we learn anything from them, I think, should take this into account: the thought experiments reveal to us constraints on theories that we otherwise have difficulty grasping.

4. Discussion and conclusions

Ever since Mach first discussed the phenomenon of thought experiments (Mach, 1905), many

---

[7] We want to show that ($q_v + 1/q_v$) > 2. But this would be the same as showing that $q_v + 1/q_v - 2 > 0$. Multiplication with q gives $q_v^2 + 1 - 2q_v > 0$, which can be rewritten as $(q_v - 1)^2 > 0$. But this is obviously true (unless q_v=1), since the square of a real number is always non-negative.

suggestions as to how to understand this particular form of reasoning have been put forward. One of the central issues in the discussion has been how it is possible to learn anything about the world by merely imagining various situations. How can knowledge be gained without any new empirical input?

Mach himself argued that thought experiments provide us with a key to a "treasure-store" of accumulated experiences which are "ever close at hand, and of which only the smallest portion is embodied in clear articulate thought" (Mach, 1905). Kuhn stressed the way thought experiments help us to develop our concepts: by confronting our ideas and conceptual scaffoldings with concrete imagined situations we can develop our conceptions of reality (Kuhn, 1977). Others have stressed the importance for thought experiments of biologically acquired knowledge: how our brains during millions of years of natural selection have acquired intuitions about what is possible and what is not possible (Sorensen, 1992). It also has been suggested that some thought experiments may act as a kind of window into a Platonic world of existing laws of nature (Brown, 1991).

As Norton (1996), to my mind convincingly, has shown, one should consider thought experiments as epistemologically equivalent to other forms of scientific arguments. Thought experiments practically never provide us with any information that we could not also have gained from more abstract reasoning, starting out from corresponding premises.[8] This granted, however, many questions still remain to be answered. For example, what is it that makes thought experiments such valuable pedagogical tools? What kind of knowledge can we expect to gain from thought experiments? Are there areas where we should not trust the outcomes of thought experiments? These, as well as many other related issues, have been thoroughly analyzed in the literature, both in the references mentioned above and in other articles.

In the present article I have argued that many thought experiments owe their success partly to the fact that good theories are rare. Not all sets of statements can be true, even in a hypothetical world. And the more general that a statement purports to be, the greater the risk that it will clash with some other statement. Often, at first glance, a given set of statements seems to make sense, but when confronted with the imagined but concrete situation of a thought experiment, inherent contradictions are implacably revealed. In other cases, one or several statements together are found to imply other, seemingly unrelated, statements.

I do not here claim that thought experiments differ from other theoretical methods of reasoning in this respect. Rather, it can be said about many theoretical endeavors – including theoretical physics – that they owe their utility to the scarcity of consistent theories. There is nothing revolutionary or fundamentally new in this claim. But in the case of thought experimentation, I think that this aspect of theoretical reasoning becomes particularly evident. And as we do not possess any intuitive understanding of the potency of this scarcity in theoretical considerations, we are often surprised at what can be accomplished by thought experiments.

---

[8] I say "practically never" since it is at least in principle possible that the specifics of a thought experiment are essential for the result to follow. See also Häggqvist (2007).

Let me put this another way. It has often been asked how thought experiments can accomplish so surprisingly much – where does the knowledge come from? But the question that I am asking here, and which may well be more relevant, is rather this: Why do we find this fact – that thought experiments do accomplish so much – so surprising? My answer is that most of us lack an intuitive understanding of one of the basic elements in all theoretical endeavors, namely the fact that it is difficult to create theories that are both general and consistent – we are, as I have put it, to a large extent oblivious to theoretical constraints. Since many thought experiments make more or less direct use precisely of these constraints, we tend to marvel at their outcomes.

These claims primarily concern thought experiments on the deductive side of the spectrum, that is, the ones in which the general conclusion follows mainly deductively from the specific conclusion. How about inductive thought experiments? Their specific situations often act as examples or reminders of earlier experienced cases. When we learn how a proposed principle may be applied in a specific case already familiar to us, or how a certain idea nicely fits with a particular situation, we become more likely to accept the principle or the idea as a general one. This is, as noted above, the case with Galileo's famous ship's cabin-argument for the Relativity Principle. This is also the case with “Heisenberg's microscope” (Heisenberg, 1930) and some other thought experiments intended to illustrate the uncertainty principle in quantum mechanics (see *e.g.* Bohr, 1949). Here we are, typically, led through one case where the uncertainty principle just escapes leading to an inconsistency. And we are expected to accept that similar considerations will apply in all cases and under all circumstances. Although pedagogically highly useful, I don't think that such inductive thought experiments have much to do with the rareness of good theories.

All examples discussed in this article are taken from physics. Nevertheless, I believe the perspective put forward here could also be of use in analyzing thought experiments in other fields, such as ethics. At first glance, this may sound strange, as one doesn't tend to think of thought experiments in ethics as being of the deductive kind. But in many cases, I think they are. True – the specific conclusion is seldom reached by deductive means in ethics. More often it is rather a question of intuition, or so-called gut-feeling (as when we are asked how to act when confronted with a bolting car that is certain to kill five people unless we push one other person in front of the car). However, the step from specific conclusion to general result in these thought experiments can be deductive – and this, as we have seen, is what would make us classify them as such. In those thought experiments, pointing at ethical dilemmas, our moral beliefs in different situations are confronted with each other – and often found inconsistent. That is, what such thought experiments illustrate is that not all sets of moral beliefs are internally consistent; good moral theories – like good physical theories – are rare. Needless to say, however, these comments must be viewed as preliminary; this should prove a fruitful area for future studies.

Besides providing us with new insights regarding deductive thought experiments, I also claim that there is a certain pedagogical value in the perspective put forward here. As the examples show, some thought experiments illustrate very clearly the rareness of general and consistent theories.

Thereby, they help us to recognize our Obliviousness to Theoretical Constraints – our lack of awareness that most sets of general statements relating to the world actually are contradictory. Since this insight is essential in understanding the value of thought itself – whether in the form of thought experimentation or in other forms – I suggest that thought experiments could be put to greater use when explaining both the values and the methods of scientific endeavors to students, as well as to politicians and other non-scientists.

Acknowledgements

The author would like to thank two anonymous referees for valuable comments on an earlier version of this article, and Dr. M Henri Day for fruitful discussions.

References

Atkinson, D., Peijnenburg, J. (2004) “Galileo and prior philosophy” *Studies in History and Philosophy of Science* **35**, 115

Bohr, N. (1949) “Discussion with Einstein on epistemological problems in atomic physics”. In Wheeler, J. A., Zurek, W. H. *Quantum theory and measurement* (Princeton University Press, 1983)

Bondi, H. (1962) *Relativity and common sense* (Heinemann Educational Books Ltd)

Brown, J. (1991) *The laboratory of the mind* (London, Routledge)

Descartes, R. (1644) *Principles of Philosophy* (Dortrecht: Kluwer 1991)

Dugas, R. (1988) *A history of mechanics* (Dover Publications)

Einstein, A. (1905) “Ist die Trägheit eines Körpers von seinem Energiegehalt abhängig?” *Annalen der Physik* **17**. English translation (“Does the inertia of a body depend upon its energy-content?”*)* in Einstein et.al. *The principle of relativity* (Dover Publications, 1923)

Einstein, A. (1920) *Relativity. The special and the general theory* (Methuen & Co. Ltd)

Galilei, G. (1638) *Dialogues concerning two new sciences* (Dover Publications, 1954)

Galili, I. (2009) “Thought experiments: determining their meaning” *Science and Education* **18**, 1

Gendler, T. S. (1998) “Galileo and the indispensability of scientific though experiments” *British Journal for the Philosophy of Science* **49**, 397

Heisenberg, W. (1930) *The physical principles of the quantum theory* (University of Chicago Press)

Helm, H., Gilbert, J. (1985a) “Thought experiments and physics education – part 1” *Physics Education* **20**, 124

Helm, H., Gilbert, J., Watts, D. M. (1985b) “Thought experiments and physics education – part 2” *Physics Education* **20**, 211

Häggqvist, S. (2007) “The a priori thesis: a critical assessment” *Croatian Journal of Philosophy* **VII:19**, 47

Kuhn, T. (1977) *The essential tension*, p240-265 (Chicago: University of Chicago)

Kühne, U. (2005) *Die Methode des Gedankenexperiments* (Frankfurt: Suhrkamp)

Leibniz, G. W. (1687) “Letter of Mr. Leibniz on a General Principle Useful in Explaining the Laws of Nature through a Consideration of the Divine Wisdom; to Serve as a Reply to the Response of the Rev. Father Malebranche” in *Philosophical Papers and Letters*, vol 2, ed. L. E. Loemker (Dordrecht: D. Reidel, 1969)

Lewis, G. N., Tolman, R. C. (1909) “The Principle of Relativity and Non-Newtonian Mechanics” *Philosophical Magazine* **18**, 510

Mach, E. (1893) *The science of mechanics: A critical and historical account of its development* (Open Court Publishing Company, 1988)

Mach, E. (1905) *Knowledge and error* (Dortrecht: D. Reidel, 1976)

Mermin, N. D. (2005) *It's about time* (Princeton University Press)

Naylor, R. H. (1989) “Galileo's experimental discourse” in Gooding, D., Schaffer, S., Pinch, T. (eds) *The uses of experiment: Studies in the natural sciences* (Cambrigde University Press, 1989)

Norton, J. D. (1991) “Thought experiments in Einstein's work” i Horowitz, T., Massey, G. J. (eds) *Thought experiments in Science and Philosophy* (Savage MD, Rowman and Littlefield Publ., 1991)

Norton, J. D. (1996) “Are thought experiments just what you thought?” *Canadian Journal of Philosophy* **26**(3), 333

Palmieri, P. (2005) “'Spuntar lo scoglio piú duro': did Galileo ever think the most beautiful thought experiment in the history of science?” *Studies in History and Philosophy of Science* **36**, 223

Sorensen, R. (1992) “Thought experiments and the epistemology of laws” *Canadian Journal of Philosophy* **22**(1), 15

Velentzas, A. Halkia, K. Skordoulis, C. (2007) “Thought experiments in the theory of relativity and in quantum mechanics: their presence in textbooks and in popular science books” *Science & Education* **16**, 353

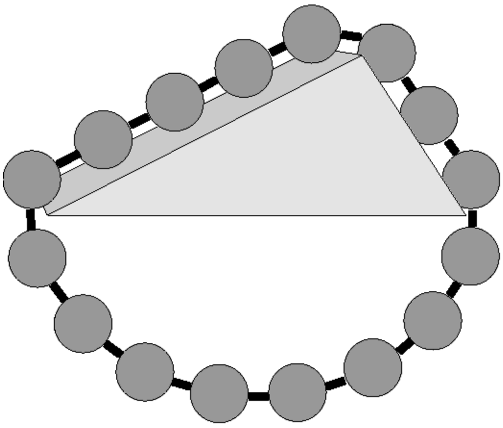

Figure 1 Stevin imagines a chain of balls hung over a triangular block. The chain thus placed will remain at rest, since otherwise it would constitute a *perpetuum mobile*. If the symmetrical arc below the block is cut away, it is therefore found that the three balls on the right will balance the five balls on the left. This is the specific conclusion in this thought experiment.

Descartes' collision laws

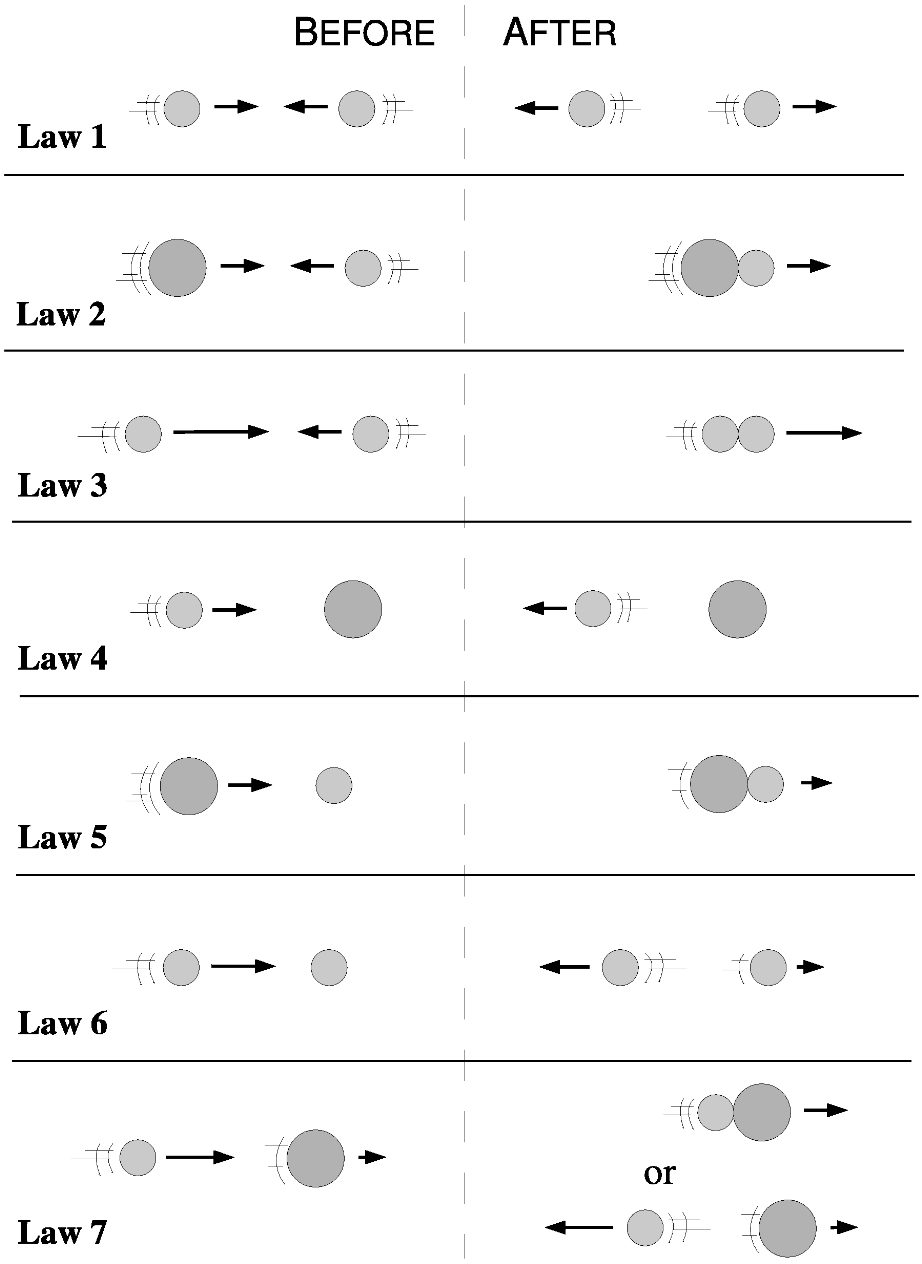

Figure 2 Colliding bodies of different weights are here represented by spheres of different sizes. Arrow lengths represent speeds. Absence of an arrow means that the body concerned is at rest. Note that in laws 1, 2 and 4 the speeds before and after the collision are the same for both objects, although the directions may have changed. In law 3 the common speed after the collision is the mean of the two speeds before the collision. In general, the speeds are such that Decartes' quantity of motion is conserved in each collision. In the last law there are two possible outcomes depending on the relative sizes of the two bodies and their speeds.

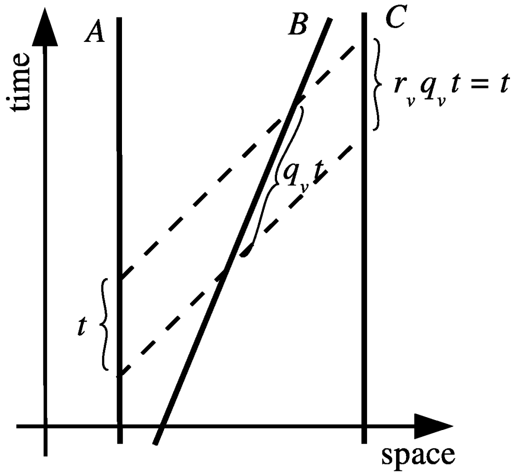

Figure 3 Observer *B* is moving at speed *v* from observer *A* to observer *C*. *A* sends out two pulses of light that pass *B* and arrive at *C*. This simple thought experiment shows that the relevant Doppler factor for *B* and *C* (*i.e.*, r_v) is the inverse of that relevant for *A* and *B* (*i.e.*, q_v).

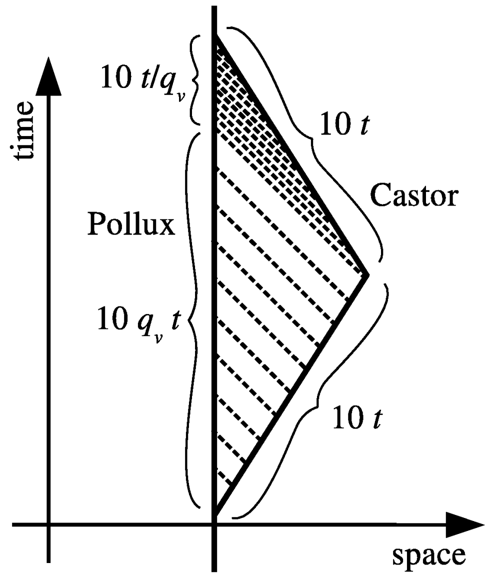

Figure 4 Pollux stays on earth while Castor departs on a long space trip. During his trip Castor sends light pulses to his brother at regular intervals *t*. Ten pulses are dispatched during the first part of the trip, when he is receding from the earth, and ten when he is on his way back again. By adding the twenty time intervals between the pulses for Castor and Pollux respectively, it can be shown that Castor will age less than Pollux during the former's voyage.