

Matters of Time Directionality in Gravitation Theory

Christophe Nickner*

Updated 6 September 2016

Abstract

Several issues related to time directionality as they arise from a semi-classical perspective are discussed which are all directly or indirectly relevant to gravitation theory. Important clarifications are achieved regarding in particular the concept of negative energy and the associated notion of negative mass. Several difficulties associated with the possibility for matter to occupy negative energy states are identified and solutions to the problems they represent are suggested. An alternative interpretation of negative energy states as voids in the positive energy portion of the vacuum is developed. Generalized gravitational field equations which allow to fulfill the requirements set by the axioms derived from this analysis are provided. A revised formulation of the discrete symmetry operations based on the insights gained while addressing the problem of negative energy is then proposed which gives rise to an improved notion of time reversal. A group of symmetry operations which involves a reversal of the sign of action is introduced which allows a mathematically complete set of discrete symmetry operations to be obtained. Those results are then applied to show that no fundamental lopsidedness arises as a consequence of the observed cosmological asymmetry between matter and antimatter. Finally, an explanation for the binary nature of the degrees of freedom associated with the microscopic state of matter trapped by the event horizon of a black hole is provided which allows the derivation of an exact measure for the entropy of elementary black holes.

*Permanent address: c.nickner@gmail.com

Contents

1	Introduction	6
1.1	Motivations	7
1.2	Approach	9
1.3	Historical context	11
1.4	Organizing principle	12
2	Negative Energy	14
2.1	The negative energy hypothesis	14
2.2	The time-direction degree of freedom	21
2.3	Our current understanding	34
2.4	The negative mass concept	41
2.5	The equivalence principle with negative mass	64
2.6	An effect of voids in the matter distribution	80
2.7	Six problems for negative energy matter	101
2.8	The origin of repulsive gravitational forces	105
2.9	No energy out of nothing	111
2.10	The problem of vacuum decay	119
2.11	Energy and momentum conservation	125
2.12	Absolute inertial mass	130
2.13	A few other misconceptions	133
2.14	An axiomatic formulation	141
2.15	Generalized gravitational field equations	147
2.16	Summary	169
3	Discrete Symmetries	180
3.1	The problem of discrete symmetries	180
3.2	The constraint of relational description	182
3.3	The concept of bidirectional time	186

3.4	Alternative definition of C , P , and T	195
3.5	The time reversal operation	201
3.6	The charge conjugation operation	208
3.7	Invariance under combined reversals	212
3.8	The significance of classical equations	216
3.9	Reversal of action	218
3.10	The problem of matter-antimatter asymmetry	227
3.11	Black hole entropy	236
3.12	Negative temperatures	268
3.13	Summary	271
4	Conclusion	284
4.1	Main results	284
4.2	Historical perspective	286
	Acknowledgements	290
	Bibliography	291
	Index	293

List of Figures

3.1	Variation of physical parameters under the proposed alternative definition of P , T , and C as described from the bidirectional time viewpoint	191
3.2	Variation of physical parameters under the proposed alternative definition of P , T , and C as apparent from the unidirectional time viewpoint	194
3.3	Four different outcomes of applying each of the relationally distinct action reversal symmetry operations as described from the bidirectional time viewpoint	221

List of Tables

3.1	Variation of the physical parameters associated with a process transformed by the discrete P , T , and C symmetry operations as they are traditionally defined	197
3.2	Implicitly assumed variation of physical parameters under the discrete P , T , and C symmetry operations as they are traditionally defined	197
3.3	Variation of physical parameters under the redefined discrete P , T , and C symmetry operations as described from the bidirectional time viewpoint	198
3.4	Variation of physical parameters under the redefined discrete P , T , and C symmetry operations as described from the unidirectional time viewpoint	198
3.5	Variations of physical parameters under the four relationally distinct action sign reversing symmetry operations as described from the bidirectional time viewpoint	220

Chapter 1

Introduction

The reflection which gave rise to the developments that will be introduced in this report started with a very simple question: could gravitation be a repulsive force under certain circumstances and what would it mean for gravitational mass to be negative? Even though there appears to be important difficulties associated with the possibility that a gravitationally repulsive body may exist, particularly in the context of a general relativistic theory, the idea of a symmetry which would have to do with the sign of mass or energy is certainly quite appealing aesthetically. Indeed, if the electric charge and all the other charges turning up in particle physics are allowed to be both positive and negative, why should mass or energy be restricted to positive values? What I came to realize through a careful analysis of the assumptions behind the common idea that gravitationally repulsive matter does not exist is that there is actually a general misunderstanding surrounding the whole idea of negative energy in modern physical theory and that this is the single most important stumbling block that is preventing necessary progress to be achieved in several fields of fundamental theoretical physics. The objective of this essay is to clear up the misunderstanding and to provide a detailed account of the most crucial advances which are made possible by adopting a more consistent approach regarding some essential concepts related to time directionality and their relationships with our current classical theory of gravitation.

As a consequence of the relatively long period of gestation during which the mere intuitive insights from which this work originates evolved into a revised classical theory of gravitation, I was able to explore the consequences of some of the most decisive results which were reached in the course of that

process on a rather large number of questions of fundamental interest. I will begin by discussing the general problem of negative energy and explaining the difficulties existing with the current conception of negative mass. Then I will address some related issues concerning our concept of vacuum energy and propose solutions to the problems which may arise in the context of an alternative description of negative energy matter based on those developments. I will then introduce an alternative set of axioms that will allow an appropriate and at last consistent integration of the concept of negative energy to physical theory. I will also provide an outline of the mathematical framework which enables to integrate those ideas into a generalized classical theory of the gravitational field and explore some of the most straightforward consequences of those advances for classical cosmology theory. Then I will revisit the problem of spacetime- and gravitation-related discrete symmetries in the context of this new understanding and I will describe what can thus be learned concerning time directionality. I will finally discuss the relevance of the knowledge here gained for addressing some very specific issues concerning the semi-classical theory of black hole thermodynamics.

1.1 Motivations

It must be mentioned that even though I initially pursued the idea underlying the developments discussed in this report based on largely aesthetic motives, the actual reasons that later fueled my interest in developing a viable model around it were of a more pragmatic nature. In particular, I saw the need that existed, but that few authors recognized, to reformulate the current classical theory of gravitation in a way that would be consistent with the possibility for elementary particles to be found in the negative energy states allowed by special relativistic quantum theories. Indeed, I had come to understand that the current interpretation of those negative energy states as merely being those of antiparticles whose behavior is identical to that of ordinary matter from a gravitational viewpoint, was dependent on the a priori assumption that only some of those energy states were allowed. In other words, we had solved the problem of the puzzling prediction of negative energy states by postulating that those states were not allowed, without justifying this very assumption. But if we recognize that the whole spectrum of energy states predicted to exist by quantum theory can effectively be occupied, even if transitions between positive and negative energy states may not be allowed,

then we need a classical theory of gravitation that is consistent with this requirement. However, further considerations indicated that the general theory of relativity was not entirely compatible with an appropriate notion of negative energy obeying certain consistency requirements which must nevertheless be imposed. Despite those difficulties I believe that the imperative to provide an appropriate description of negative energy matter should prevail over our willingness to leave untouched the current theory of gravitation, because I have recognized the inadequacy of the arguments against the physical nature of negative energy states, while I also understand that quantum theory constitutes a more appropriate basis to decide what states are allowed for elementary particles. This position is made even more appropriate given that I will effectively propose an alternative framework that merely generalizes relativity theory without affecting its basic mathematical structure, while allowing an appropriate description of negative energy matter.

I was also motivated by the desire to obtain a better agreement between theoretical predictions and astronomical observations concerning certain aspects of the gravitational dynamics of the universe. In particular, there was the exceptionally severe disagreement between most theoretical derivations of the expected value of vacuum energy density and observational constraints on the upper (positive or negative) value of the cosmological constant. Very early on I saw that the hypothesis of allowing for the existence of matter in a negative energy state could potentially provide a whole new class of negative contributions to zero-point vacuum energy which would be distinct from those already considered in conventional calculations and which could naturally allow an overall cancellation of all contributions if some level of symmetry exists between positive and negative energy states. Here again I chose not to ignore, as most people did, what seemed to be the necessary conclusion that matter must be allowed to occupy the currently forbidden negative energy states if we are to obtain a compensation for the known positive contributions of ordinary matter to vacuum energy. Despite the apparent difficulties, perceived or real, associated with negative energy as a possible state of matter it had become very clear to me that this was a hypothesis which had become unavoidable.

Finally, I also wanted to bring some much needed clarity to the theoretical context in which we are to address the problem of the elaboration of a theory of the gravitational interaction compatible with the basic principles of quantum theory. Here I will show the essential role played by the discrete spacetime and momentum-energy symmetry operations (appropriately rede-

finer and extended) in characterizing states of matter at the spatial scale and energy level at which we can expect the gravitational interaction among elementary particles to be as strong as the other known interactions. This will be achieved by demonstrating the relevance of those symmetry operations for a definition of the microstates that must be taken into consideration in order to provide an appropriate measure of black hole entropy.

1.2 Approach

Basically the approach I will follow consists in explaining how some specific aspect of the quantum world, namely the ignored possibility for both positive and negative energy states to propagate forward and backward in time, changes our understanding of the classical theory of gravitation and allows to actually improve and simplify its formulation in a way that will have decisive consequences for our description of certain phenomena. The level of this discussion is clearly philosophical, but remains very precise in its reference to quantitative aspects and concepts, unlike most philosophical essays concerning physics. Mathematical developments will be kept to a bare minimum, however, and will be introduced only when absolutely necessary and of utmost significance. This is obviously in contrast with the current tendency observed in the physical sciences to focus on technical aspects and to relegate epistemology to the backseat. Concerning the methodology which is reflected in the style of this report I must emphasize that I have been introduced to quantitative methods very early on, but that I effectively came to rely on them in an increasingly less decisive way as I realized that at the present epoch all the really useful mathematical developments that could be carried out in the field of fundamental theoretical physics have already been performed over and over again by competent people and that what actually matters now is to provide a consistent interpretation of the existing frameworks. Indeed, such a fully consistent interpretation is currently missing, perhaps because the vast majority of competent researchers prefer to dedicate their efforts to more technical aspects, and this is restraining our ability to distinguish between what are viable developments and what is logically and empirically inappropriate. But as I do believe that the objective of a philosophy of science should be to explain and to justify, through logical arguments constrained by observational data, a particular vision of the world, and as I am convinced that it is only when this goal is successfully achieved

that we are allowed to consider some vision of the world as a valid representation of it, then this is the objective toward which I effectively directed my efforts.

Furthermore, it is important to note that if mathematical developments do not dominate the content of this report it is also simply a consequence of the fact that while I have achieved a crucial revision and a necessary improvement of the mathematical framework and of the interpretation of relativity theory in a semi-classical context, I nevertheless ended up confirming the general validity of the basic mathematical structure of the current theory, within a certain limit, so that practically no further mathematical developments were required. The reader must be warned, however, that the density of significant information that is to be found in the text of this report is very high. In some cases, it took me years of dedicated reflection and careful investigation to gain confidence in the validity and inevitability of certain specific results which may be mentioned only once in the main body of this report, as otherwise the length of the treatise would be excessive. Therefore, you must pay attention to every detail of the discussion and be careful not to miss some important information that may be necessary later on for understanding and appreciating the value of other elements of the discussion. I know that this may sound obvious but here the difficulty may be so great that it is nevertheless appropriate to issue such a warning. This, however, does not mean that the present report is actually difficult to read, to the contrary. In fact, I tend to follow a rather educational approach according to which I do not avoid making statements and providing explanations, even when they may appear obvious to some or even most physicists, because I think that it is better to make too many unnecessary statements than to more or less willingly avoid making some which would have been useful. This approach should not be considered as condescending or as an indication that this work is devoted mainly to a beginner audience.

I must mention that I do recognize that the approach I followed to achieve the valuable results that will be described and justified here *is* different from that which is usually followed in theoretical physics. But, even if I would not myself have believed that one could achieve such significant results following that kind of method when I started studying physics, which I did the usual way by learning about the mathematics of quantum theory, statistical mechanics and relativity theory, it is through experience and by force of circumstance (although not as a result of mere incompetence), after having slowly and partly unwillingly deviated from the traditional path, that I began

to understand that there is real value in such an approach which I developed by making systematic a learning process that initially appeared to merely be a faithful but irresponsible time-wasting improvisation. If the reader is willing to immerse herself in the same experience and loosen her grip on more traditional ways of achieving deep understanding, while nevertheless being ready to spend considerable efforts to follow rigorous logical reasoning and analysis, I can assure her that she will not be deceived and will learn useful physics, which is not so bad already by today's standards.

1.3 Historical context

There are many similarities between the current state in which science finds itself and those through which it went at other crucial turning points in its history. Indeed, the situation we have now arrived at is characterized by an accumulation of unanswered questions which creates an impasse that prevents further progress to be achieved. It is my belief that answering just a few key questions among these will release a great deal of pressure that will greatly facilitate future theoretical research. When we examine the present situation in physics it becomes clear in effect that if there are questions which we are justified in not being able to answer right now, because they are related to what may be said of reality under conditions which we cannot yet reproduce in experiments (think of trying to explain the origin of the free parameters of the standard model of particle physics), there are also questions which have to do with known difficulties which we have puzzled about for a long time and which we have no reason to believe further experiments may be particularly useful in helping resolve. But those are problems whose existence is often simply ignored by most people or which are sometimes considered to have already been solved, while careful examination shows that this is not always entirely the case. Most current programs in fundamental theoretical physics are focused on trying to solve the problems raised by questions of the first type and this is unfortunate, because here is precisely the domain in which progress is limited by technological constraints of a practical nature and the cost of achieving the required experiments. Very early on I recognized that if I was going to enable progress to be made in physics I had to concentrate on questions of the latter type, where progress could occur not only in my lifetime, but also as a consequence of the success or failure of my own enterprise.

Among the questions we may have hope to answer using our current knowledge is the question I mentioned earlier on as having being that which launched the reflection process from which this report emerged. It is effectively one of those unsolved questions whose very existence is usually ignored or which is considered to have already been solved, while this is clearly not the case, as I will explain later. You will not see it mentioned in most accounts as being one of today's open questions in physics, but it is one of the most important categories of question regarding classical physics and a field most people currently consider to be free of major difficulties. This problem of negative energy states could actually be called the 'classical gravitation theory problem' or the general relativity problem, because properly answering that question requires introducing slight modifications to that theory, which actually consists in a generalization of its own founding principles. This is the question I will address in this report and satisfactory answers will be provided to the mostly unrecognized issues it currently raises. Doing so will require reconsidering the significance of certain aspects of the problem of vacuum energy and gaining a new understanding of the gravitational effects of homogeneous and inhomogeneous matter distributions that can be extended to our description of the physical vacuum.

1.4 Organizing principle

Every successful venture into unknown territory requires relying on the appropriate beacons and guidelines and this is particularly true when the voyage takes you to the boundaries of traditional certainties and brings you to question some essential aspects of what had previously appeared to constitute a fixed background for scientific exploration. I would therefore like to briefly describe what was the essential principle that guided me on developing the revision of classical and semi-classical theories that is described in this report. It must first of all be understood that this principle was not given as a precondition imposed on any vision of the world, but actually developed alongside improvements in my and other people's knowledge and understanding of the world and was justified by specific observations of that portion of physical reality we effectively experience and through the possibility that this probing allowed of inferring the regularities present in an even larger and more encompassing domain of the same reality.

My awareness of the importance of this principle developed mostly in

conjunction with the elaboration of a more consistent appreciation of the requirements imposed by the classical theory of gravitation. Indeed, it is while tackling the problem of negative energy that I realized the importance (emphasized by others in a different context) of a relational view of the physical properties of objects and that I understood the real significance of the requirement of relativistic invariance. This allowed me to perceive the true meaning of Einstein's insistence that the objects of physics must be conceived of only in relation to the spacetime structure to which they belong, because indeed I saw that the metric properties of space and time must be understood as dependent on the sign of energy of an object (as will be explained later), which is in contrast with traditional expectations. Thus, if a determination of the relationships between physical objects in different spatial locations or states of motion is possible only when we determine the common spacetime structure shared by those objects, then the fact that the spacetime structure itself is dependent on the nature of the objects means that the relationships between them are dependent on their nature and in particular their energy signs. It therefore appeared to me that it is not only the spatial properties of objects which require a relational viewpoint, but that any physical quantity must always be defined or characterized in relation only to similar quantities of other objects in the same universe (the physical properties of a system enable to characterize it merely in relation to the similar properties of other systems and those relationships are determined through the use of reference systems).

When I tried to understand what could logically impose such a requirement I slowly came to realize that it is the very fact that it would be meaningless to relate some physical quantity, in order to define its value, to some reference point not part of the same physical universe. Indeed, in the absence of a well-defined continuous network of causal relationships that would extend to those immaterial reference systems there can be no meaningful definition of the physical quantity involved, because physical relationships are material relationships and an object cannot be put into relation with something that is not part of the same causally related ensemble (the universe) to which it belongs. This requirement of a relational definition of physical quantities will have enormously important consequences on many aspects of the developments to be discussed in the following two chapters.

Chapter 2

Negative Energy

2.1 The negative energy hypothesis

Regarding the question of negative energy, the current situation has much in common with that in which we were at the turn of the previous century with regard to the quantization hypothesis. There was in effect some reluctance initially to recognize the validity of the original suggestion by Max Planck that energy is quantized despite the fact that this proposal would have solved the problem of black body radiation. The trouble was of course that recognizing the validity of the quantization hypothesis would have required abandoning classical physics. There is a similar dilemma with negative energy today because, as I will show, this hypothesis has the potential to solve many important problems facing theoretical physics, but those benefits come at a price which may at first appear to be too high. Indeed, the introduction of negative energy matter as a concept somewhat distinct from that which is currently favored (which I believe is required in order to allow it to be consistent from a basic theoretical viewpoint) seems to imply that general relativity has to be abandoned. But rejecting a theory so well established and so beautifully simple as general relativity is not something that most people would do without very good motives. Yet, if the current assumptions concerning the rules governing negative energy matter (if it was to effectively exist) may appear to better agree with relativity, they actually contradict some of the basic principles on which this theory is founded, therefore making it just as untenable. We must then either abandon the idea that negative energy matter can exist, or else provide a better interpretation of negative energy

which may force a reinterpretation of relativity theory itself. But I will show that the conclusion that the latter alternative is the only viable one is not necessarily as dramatic in terms of its consequences as may seem, because what is required in this context is mainly a reinterpretation of the equivalence principle and not a rejection of the whole mathematical framework of relativity theory.

There is however an additional problem for the negative energy hypothesis which is that there appears to be no observational evidence for matter in such a state. But here also there is an analogy which should teach us a lesson. This is the case of the neutrino as a massive particle. For a long time when I was reading physics papers or any book on the subject of particle physics I could see that it was nearly *always* assumed, more or less implicitly, that the neutrino is massless as if this was a fact, while actually there was absolutely no evidence that this is effectively the case and it was merely the difficulty to prove that the hypothesis is wrong that justified that everyone just assumed that the neutrino is massless. But just as for the idea that negative energy does not exist, I thought that it was incorrect to simply assume that the neutrino is massless when this could not yet be considered a fact. Thus, I always kept an open mind about those issues, because I saw that there were strong arguments (usually not recognized) for rejecting those commonly held assumptions and in the case of the neutrino at least it appears that this position was justified. In fact, I will later explain that there are very good reasons to expect that it should not be easy to confirm the existence of negative energy matter, because, as I have come to understand, it is not even directly observable, just as the more common, hypothetical dark matter. Thus, if I am right, the implicit assumption that negative energy is forbidden would be just one of those ‘reasonable’ assumptions which we should be careful not taking too seriously.

The problem of negative energy has another parallel in a distinct but not entirely unrelated problem which is that of the origin of the arrow of time. Indeed, it was suggested by some eminent figures that the problem of irreversibility could be solved by integrating some fundamental element of irreversibility into the formalism of even the most elementary physical theories. This would seem to be justified by the fact that the problem of time asymmetry has been known to exist for a long time and no acceptable solution of it that would be based on boundary conditions imposed on otherwise time-symmetric evolution has ever been found. But again I think that the difficulty to prove a hypothesis (that time asymmetry can arise from time-symmetric

physical laws) should not be taken as evidence that what may perhaps be its only alternative (that time asymmetry is fundamental) is right. In the case of negative energy, we are also in a situation where we have built into the very formalism of our most fundamental theory of matter (which currently is quantum field theory) the apparently necessary, but clearly unjustified (from a theoretical viewpoint) hypothesis that only positive frequencies (associated with positive energies) are allowed to propagate in the future (the constraint on negative frequencies being merely that they must propagate toward the past).

However, I think that the apparent validity of this artificial restriction does not imply that positive frequencies cannot propagate backward in time or that negative frequencies cannot propagate forward in time, but merely that if there exist two kinds of matter related by their opposite energy signs (the frequency signs relative to the direction of propagation in time) then they can only interact with matter of the same energy sign for some reason (I will eventually explain why such a limitation naturally occurs). This absence of interaction or interference (in the classical sense) is what really justifies that quantum field theory only deals with matter of one energy sign under most circumstances (when gravitation is not involved). But given that I am suggesting that energy sign is a relatively defined physical property, so that there is no absolute (non-relational) distinction between positive and negative energy matter, then it must effectively be concluded that there cannot exist a constraint that would impose that negative energy matter and only matter with such an energy sign does not exist at all if positive energy matter itself is allowed to exist, as required, because it is not even possible to identify the distinguishing property specific to negative energy matter that would justify that its existence be ruled out. Thus, I am allowed to conclude that any attempt at getting rid of the apparently intractable problem of negative energy states by simply imposing a constraint to be applied on the formalism itself is misguided and unnecessary, because indeed once an appropriate understanding of the true nature of negative energy matter is available it becomes apparent that a restriction on allowed frequencies is no longer necessary.

In this context it becomes apparent that one often mentioned argument that must definitely be rejected concerning the nature of the gravitational interaction is the idea that the strength of gravitation on the largest scales is a consequence of the ‘fact’ that this interaction is always attractive, which is usually assumed to follow from the observation that there does not exist

negative gravitational charges (negative energy matter is assumed not to exist). Indeed, what actually explains the fact that gravitation is a dominant force on larger scales (in addition to its long range property) is not the absence of matter in a negative energy state, but the simple observation that gravity is attractive between objects with the same positive gravitational charge, that is, between objects with a positive sign of energy. Thus if gravitation dominates over electrical forces on astronomical scales it is really a consequence of the fact that while identical electric charges tend to disperse under mutual electrostatic repulsion, positive energies have a tendency to coalesce and to gain in strength under mutual gravitational attraction and the fact that electromagnetism is already known to have both positive and negative charges has nothing to do with the fact that those charges do not so readily accumulate, because even if there were only positive electric charges they would not regroup and accumulate, because identical electric charges mutually repel one another and the possibility for a cancellation of such opposite charges actually facilitates an *accumulation* of those charges, but only in neutral configurations and under the influence of gravitation.

It must therefore be understood that there is no requirement for gravitation to always be attractive merely on the basis of the fact that its existence can be felt despite its extreme weakness, as is sometimes suggested. Indeed, if it was found that there effectively exist negative energy particles, the possibility of energy cancellations would not necessarily prevent the accumulation of matter with one or another energy sign, because negative energy matter may also be gravitationally attracted to itself (despite what is usually assumed) and could therefore also be subject to accumulation. To summarize, what makes electrical forces negligible on the large scale is the fact that identical electric charges do not attract one another and therefore do not accumulate as may identical gravitational charges. Instead electric charges of opposite signs are attracted to each other and immediately cancel out, therefore preventing further accumulation, at least under the influence of electric forces. But this does not mean that gravitation would be submitted to the same fate if negative energy particles were found to exist, because it may well be the case that gravitational charges with the same sign always attract one another given that this is already known to be true for positive energy matter and this would not even forbid opposite energy bodies from gravitationally repelling one another. The frequently made remark that gravitation is attractive for all particles should therefore be understood to mean only that it is attractive for all currently known forms of matter.

Thus, again, the observation of large accumulations of positive energy matter is not an argument against the existence of negative energy matter. But it is also true that the apparent absence of large accumulations of *negative* energy matter would not necessarily mean that such matter does not exist, even if we were to assume that this matter gravitationally attracts matter of the same kind. Indeed, it may turn out that this matter is dark and given that it may also be repelled by positive energy matter (even if this is not what is usually assumed) we might be justified to expect that it should be located mainly in regions of the universe where the density of positive energy matter is the lowest. Therefore, negative energy matter would be virtually absent from regions where positive energy matter is more abundant, like that in which we are located, and this would explain that we have never noticed its existence. I will explain later why the assumptions discussed here concerning the nature of negative energy matter should in effect be those which are retained, thus confirming the validity of the above explanations as to why it is that negative energy matter appears to be absent from our universe. It will then be clear that theoretically it is to be expected that if negative energy matter exists it should have the properties which are responsible for our very ignorance of its existence.

I think that what must be recognized above all is that the commonly held view that the occurrence of negative energy in a theory is necessarily always indicative of a problem is not rationally motivated and that it is not true that all traces of negative energy *must* be eradicated at all costs whenever they are encountered. Dirac, at least, understood that the prediction of negative energy states could not be ignored and tried to provide an explanation for the absence of transitions to such states. His solution, based on the idea that negative energy states are already all occupied, was not satisfactory, but at least he did not simply reject the possibility that negative energy matter might have to be considered real. There is no motive to argue, as people often do, that negative energy is totally unacceptable, other than the difficulty to find an appropriate interpretation that would be compatible with empirical facts for this logically unavoidable counterpart to positive energy. In the absence of a theoretical justification for the absence of negative energy matter I think that the only appropriate approach would be to seek to find out why it is that we never observe matter in such states, rather than try to build that assumption into a then necessarily incomplete theory of quantum fields. In this particular sense it is significant that the prediction of

antiparticles was a by-product of Dirac's original interpretation of negative energy states, because this contributed to the belief that the discovery of antiparticles constituted a solution of the negative energy problem. But given that Dirac's interpretation was later found to be inappropriate, I think that it needs to be recognized that in fact antiparticles can only be one particular aspect of a complete solution of the problem of negative energy, which therefore remains unsolved.

In any case it must be understood that even if we were to succeed in justifying that it should be imposed that there cannot be transitions from a positive energy state to a negative energy state, we would not have solved the problem of negative energy. This is because such a restriction would merely impose that no positive energy particle can turn into a negative energy particle (and vice versa maybe), but there would be nothing in that constraint to forbid a particle to already be in a negative energy state, in which case we would still need to provide a consistent description of the properties of matter in such a state and to justify that we do not observe those negative energy particles under most conditions. In fact, I will later provide arguments to the effect that just such a restriction on energy sign shifting transitions is to be expected to occur very naturally, even if negative energy matter must effectively be allowed to exist. Anyhow, the fact is that if there is no reason to assume that some restriction applying to energy sign reversal would forbid *positive* energy matter from existing then there cannot be more justification in assuming that such a restriction forbids negative energy matter from being present in the same way. I must insist again that there is no reason to assume that the concept of negative energy is problematic all by itself and that negative energy must be avoided systematically, because the only requirement, regarding negative energy states, may be that there cannot be transition to such states by a particle in a positive energy state and this only when the transition would be to a state of negative energy propagating forward in time. Such a requirement is necessary (although not entirely sufficient) to keep positive and negative energy matter virtually isolated at the quantum level, so that the experimental constraint of an absence of interference from negative energy matter into the theoretical predictions involving positive energy matter can be observed.

I do understand of course that there are a number of issues associated with the possibility that matter may occupy negative energy states. Of particular concern would be the issue of 'vacuum decay' or the apparent problem that all positive energy particles should fall within a very short interval of

time into the available negative energy states by releasing a compensating amount of positive energy radiation, if those states are not assumed to be forbidden. In fact, this problem would seem to affect negative energy matter itself, even if transitions to negative energy states by positive energy particles were found to be impossible. This is of course the difficulty that motivated Dirac's problematic proposal that those energy states should already be nearly completely filled so that no further decay should occur. But I will show in a later portion of this report that this problem and also some others which may seem to arise in relation to the possibility for negative energy matter to exist in a stable form are merely a consequence of the inappropriateness of the current interpretation of the concept of negative energy. In fact, it will be shown that it is not even necessary to assume that negative energy states cannot be reached by matter in a positive energy state, because even matter already in a negative energy state cannot be assumed to fall to even 'lower' energy states.

I also recognize that the tentative interpretation of negative energy states that came to replace Dirac's solution does effectively provide some level of relief in that it at least allows to take into account those negative energy states that cannot be ignored as they actually interfere with processes involving ordinary matter. This is because we are indeed allowed to consider that antiparticles are negative energy particles propagating backward in time. But even under that particular interpretation, antiparticles can still be conceived as ordinary particles (submitted to normal gravitational interactions) from the forward time perspective relative to which their energy is positive and therefore they cannot be considered to provide an interpretation of negative energy states as would occur from a forward in time viewpoint. Again, the exclusion of true negative energy states may appear to be justified from an observational viewpoint, but it still constitutes an arbitrary rule which would at least require an explanation, as there is no consistency principle behind it. It is therefore certainly amazing that so many otherwise well informed authors suggest that no negative energy, or negative mass particle can exist, as if this was an obvious and unavoidable conclusion. It must be clear that I am not complaining about this situation, I merely want it to be recognized for what it is, because I will take a different course and it should be understood that I am not doing this without good motives or out of a fondness for hopeless, exotic or eccentric ideas.

I must therefore mention that I am aware that the originators of the steady state theory of cosmology once also criticized (based on distinct mo-

tives) the traditional position according to which the existence of negative energy matter is forbidden. But if I do find this criticism to be valid and appropriate I do not, however, find suitable the whole concept of negative energy (which is actually very traditional) proposed by these authors, nor do I agree with the objectives they unsuccessfully (given the failure of steady state cosmology) sought to achieve by using this otherwise interesting idea. I think that the fact that the hypothesis of negative energy matter was historically associated with such failed theoretical models and was also developed into many different inconsistent formulations lacking any epistemological support is more than anything else responsible for the state of suspicion and confusion that currently surrounds the whole idea of negative energy matter. The objective I will try to achieve in this chapter will therefore be to clarify the situation regarding what should be expected regarding the properties of matter in a negative energy state and to demonstrate the validity of the concept itself in the context where it is properly conceived and justified.

2.2 The time-direction degree of freedom

What emerges from my reexamination of the assumptions behind our current understanding regarding the possibility that particles may occupy negative energy states, is that we must first recognize that for any elementary particle there exists a fundamental degree of freedom related to the direction of propagation in time of its charges, including the gravitational charge, that is, including energy. The existence of such a degree of freedom means that a positive charge can effectively be positive either in relation to the positive direction of time, if such a charge propagates in the positive direction of time, or in relation to the negative direction of time, if the same positive charge propagates in the negative direction of time. But the particles so characterized would be physically different from one another. It is not possible therefore to completely specify the physical properties of a particle at a given instant by simply providing the sign of its charges independent from their direction of propagation in time. But given that a particle can actually be identified by the charges (including energy) it carries (it has no other physical properties except for its momentum, position, and spin at a given time) this means that the apparent nature of a particle may depend on whether it propagates its charges in the positive or the negative direction of time, that is, it may depend on whether it is itself propagating forward or

backward in time¹. The physical attributes of a particle can only be unambiguously defined in relation to the direction of time in which this particle propagates and this is true also for energy.

This is what the insights gained by considering the consequences of the relativity of simultaneity for the quantum description of particle interactions should be understood to imply. Indeed, it is the fact that some processes involving the exchange of a virtual particle of interaction cannot be assigned a unique definite direction of occurrence in time that renders the notion of particles propagating backward in time unavoidable. This is because the emission and absorption events of such an exchange process are spacelike separated so that their order of occurrence in time is dependent on the state of motion of the observer. Thus what is viewed by one observer as the emission of some particle carrying a negative charge, can be seen by another observer as the absorption of a similar particle carrying a positive charge, which certainly requires the sign of charge to be dependent on the perceived direction of propagation in time. Given the undeniable validity of this viewpoint, the only argument that could still allow one to reject the reality of a degree of freedom associated with the direction of propagation in time would be one based on the second law of thermodynamics and the apparent impossibility for a macroscopic body to ‘travel’ backward in time. It appears, however, that this argument is not valid, because the thermodynamic constraint only applies to the flow of information as it occurs through the formation of records and in no way forbids individual particles from propagating backward in time as long as they are not involved in processes which (collectively) would allow information to be transferred from the future to the past. It is therefore merely this limitation on the flow of information that explains the fact that our experience of reality has made us suspicious of the possibility that objects themselves (or particles) can propagate backward in time and not the actual impossibility of such an occurrence.

¹I am here considering a particle in a semi-classical way, as if we could always associate with it a definite position and momentum, even though it is clear that actual knowledge of those conjugate attributes cannot be obtained at the same time. This idealization simply allows to gain insight into what would be the properties of an elementary particle if it could be observed at the energy scale of an actual macroscopic body, while still carrying a mere unit of its other charges. We may alternatively consider a real macroscopic body and assume that it has physical properties that evolve in a perfectly coordinated fashion, with all its charges necessarily propagating in the same direction of time at all times (therefore acting as one ‘macroscopic’ charge), but such a viewpoint is actually even less realistic than the former idealization and would change nothing to the following conclusions.

In such a context the possibility to distinguish the sign of a charge, including energy, would depend on the possibility to determine the direction of propagation in time of this charge. Thus, even independently from the argument based on the relativity of simultaneity, we may consider that the sign of charges and in particular the sign of energy is defined only in relation to the state of motion of the particle carrying those charges, where ‘motion’ is here relative to time instead of space. But if we may also assume that the attribution of a direction of propagation in time is merely a matter of convention, because all that can be asserted is whether any two particles are propagating in the same direction of time or in opposite directions, as I will suggest later, then it would appear that the sign of energy itself would become a relative notion dependent on which direction of time is chosen as that in which a given particle propagates. In this particular sense we would have to recognize that associated with the relativity of ‘motion’ in time there is also a relativity of the sign of energy.

Acknowledgement that the sign of energy is a relative property actually allows one to reject the validity of the constraint usually imposed that all energy must be positive, because it means that even what appears to be positive energy according to one particular convention for the direction of propagation in time is actually negative energy according to an alternative choice for the same time-direction parameter. The possibility for particles to propagate backward in time, which is made unavoidable by the fact that backward in time motion is actually required under a consistent understanding of the constraints imposed by a relativistic treatment of quantum processes, as mentioned above, therefore actually implies that negative energies must also be allowed in physical theory, because even what we usually describe as a positive energy particle could be redefined as a negative energy particle if we were to also assume as a matter of convention that the direction of propagation in time of the particle is opposite that which is usually (more or less implicitly) assumed. Negative energies must be considered as possible states of matter even if only for particles propagating in the backward direction of time. This dependence of energy sign on the assumed direction of propagation in time is what actually allows antiparticles to be described as particles propagating backward in time with negative energies and unchanged non-gravitational charges as Feynman once suggested, even if we are also allowed to consider those particles as positive energy particles with reversed non-gravitational charges propagating in the usual forward in time direction.

What is essential to understand here is the dependence of the value of

any charge, including energy, on the direction of time in which this charge is assumed to be propagating. Thus simply saying that a particle has positive electrical charge or positive energy doesn't make sense. We must also always specify the direction of propagation of this energy with respect to the time parameter. What appears to be a positive charge or a positive energy relative to the positive direction of time would be a negative charge or a negative energy relative to the negative direction of time. Thus, all those energy signs are merely established on the basis of practical conventions and can never be asserted in an absolute fashion. It must be recognized, however, that if the energy of an electron is by convention considered positive relative to the future direction of time in which it is, again by convention, assumed to propagate, then the energy of an anti-electron must *necessarily* be considered negative relative to the past direction of time in which it must, under the same convention, be assumed to propagate. It is merely because we ignore the requirement to describe the positron as propagating backward in time that we can attribute to it a positive energy. As a consequence, it would seem that even on the basis of current observations we would not be allowed to assume that particles are forbidden from occupying properly defined negative energy states.

Yet despite the unavoidable character of this conclusion and even in the face of the enormous simplification of our world view that is made possible by the hypothesis of the existence of a fundamental degree of freedom related to time direction, it is still often suggested that the interpretation of antiparticles as particles propagating backward in time with negative energy is merely a mathematical artefact and corresponds to nothing real. But I think that this attitude is similar to that of nineteenth century philosophers and scientists rejecting the hypothesis of the existence of atoms, even in face of the overwhelming evidence in favor of this concept, supposedly because the atoms could not be seen directly, but actually because of an unjustified prejudice in favor of a continuous, macroscopic description of matter. Given the above discussion concerning the relative nature of energy sign, I think that it is clear that there is no basis for assuming, as is often done, that the negative energy of antiparticles as particles propagating backward in time is not real and that those particles are merely 'ordinary' particles which happen to be carrying opposite non-gravitational charges. If we are allowed to describe antiparticles as particles propagating backward in time, then we must recognize the existence of negative energy states.

It must in this context be understood that the commonly met suggestion

that all physical properties are simply reversed for an antiparticle (by comparison with those of the associated particle) is wrong, because the signs of all physical quantities are dependent on the direction of propagation in time and we would at least have to specify with respect to which direction of time the various quantities are to be assumed reversed. Indeed, even from the viewpoint where antiparticles are assumed to propagate in the same direction of time as do regular particles we would have to admit that energy is not reversed for an antiparticle, otherwise a pair annihilation process should release few or even no energy in the form of radiation, contrarily to what is routinely observed. Also, if we do consider instead the viewpoint of an antiparticle's true (when ordinary particles are assumed to propagate forward in time) direction of propagation in time, then energy would effectively be reversed as I already mentioned, but all non-gravitational charges far from being reversed would have to be considered rigorously unchanged given that from the forward in time viewpoint they effectively appear to be reversed while from my perspective the sign of charge is a relative notion dependent on the assumption that is made regarding the direction of propagation in time of a particle.

Thus, what appears to be a positively charged particle in relation to another particle propagating forward in time would actually appear to be a negatively charged particle in relation to yet another particle propagating backward in time and the same would be true of energy sign. Those relative alterations of the sign of charges occurring as a consequence of a reversal of time are manifested merely in the fact that what is found to be a repulsive non-gravitational interaction between two identical particles propagating in the same direction of time, would upon a reversal of the direction of propagation in time of one of the particles become an attractive interaction, or vice versa, as a result of the equivalent reversal of the sign of charge that occurs when a particle reverses its direction of propagation in time without actually reversing its charge. This is an unavoidable consequence of the fact that the departure of a positively charged particle from a region of space would from a reversed time viewpoint necessarily appear as the arrival of a particle of opposite (negative) charge, therefore implying that there is a relationship between the relative direction of propagation in time and the relative sign of any conserved physical quantity. We do not even have to know what an electric charge is or what energy is from an exact mathematical viewpoint to draw that conclusion. The reversal of charges associated with a reversal of time simply illustrates the subtlety of the relational definition of the sign

of conserved physical quantities in the context where there is a fundamental degree of freedom associated with time direction.

It must be remarked that in the context where there is effectively a dependence of the sign of charges on the direction of propagation in time it follows that there no longer needs to be a mystery regarding why all charges come in two varieties, each having the exact same magnitude but opposite polarities. This is because even if there were only, say, positive electrical charges, the fact that particles are free to propagate either forward or backward in time (under appropriate conditions) means that from a practical viewpoint there would still occur phenomena involving negatively charged, but otherwise identical charges and it would not be possible to say whether it is the positive or the negative charges which constitute the ‘true’ charges. In such a context it seems possible that the requirement imposed by modern unified field theories that the sum of charges of all elementary particles cancel out, so that the overall symmetry is preserved in the context where it is not spontaneously broken, could ultimately be understood to be made possible (if the current elementary particles are actually composed of more fundamental building blocks) by the relativity of the sign of charges with respect to the direction of time, which not only allows, but requires the existence of opposite charges. What I am now suggesting is that we would in fact be justified to consider that the same requirement also applies to energy, which would therefore come in two varieties with opposite signs, not only for particles propagating in opposite directions of time, but even relative to the conventional forward direction of time.

In any case it should be clear that it is no longer possible to consider the sign of charges, including that of energy, independently from their direction of propagation in time. The traditional viewpoint according to which it seems possible to define charge without reference to some direction of time is valid merely because we implicitly always consider the sign of charge with respect to the positive direction of time (conventionally assumed to be the future). The positive definite value of energy under all circumstances is thus an artifact of this implicit choice of the positive direction of time as the direction relative to which energy is measured. It is true though that if it was not for the non-gravitational charges carried by a particle it would effectively be impossible to distinguish between the case of a positive energy propagating forward in time and that of a negative energy propagating backward in time, just as it would be impossible to distinguish between the case of a negative energy propagating forward in time and that of a positive energy propagating

backward in time. But there is no reason to assume that there would be no distinction between positive and negative energies propagating in the same direction of time and therefore the truly significant measure concerning energy is the sign of action, which is obtained by multiplying the sign of energy by the sign of time intervals. If the hypothesis that energy must necessarily be positive has always appeared valid it is merely as a consequence of the fact that we always measure energy relative to the positive or forward direction of time and for all known particles action remains positive. As I suggested above, however, this does not mean that energy really is always positive, but merely that action, or the sign of energy relative to the sign of time intervals, is effectively always positive for all currently known particles, independently from the true sign of energy of those particles.

What I would like to suggest, ultimately, is that in fact it is not only the sign of energy that is to be viewed as a relative quantity, but that the sign of action itself is purely relative, in the sense that there could never exist a generally agreed absolutely defined positive or negative value for the sign of action of a particle. In this context not only would the sign of energy be dependent on the direction of time in which a particle is assumed to propagate, but the sign of action would itself depend on the choice of what direction in time is to be that in which what are assumed to be positive energy particles propagate, or what is the sign of energy of those particles which are considered to propagate forward in time. Here all that matters is that once you define one particle as having positive action, because you assume that it is this particle that propagates positive energy forward in time, then the particles that you must assume to be carrying negative energies forward in time or positive energies backward in time as a consequence of this *choice* are those which will have negative action. But it must be clear that you are always free to describe the first particle as propagating negative energy forward in time and therefore as having negative action, as all by itself this choice is arbitrary, but in this case the other particles would then necessarily have to be assumed to carry positive action instead of negative action, because their *relationships* of time directionality and energy sign with the first particle (the difference or the identity of the signs of time intervals and energy) would remain unchanged.

It must also be remarked that the fact that what we would currently define as negative action particles are related to ordinary matter through a simple convention regarding the direction of propagation in time means that the motive for rejecting the possibility that negative action matter may

actually exist is no stronger than that which would consist in arguing that ordinary matter itself is not allowed to exist. There is absolutely no rational motive for rejecting the viewpoint described here and many reasons to recognize its validity. In any case the fact that the sign of action is a purely relative concept which can vary as a consequence of assumptions regarding the direction of propagation in time means that if a gravitational field depends on the sign of action of its source then it should itself vary as a function of the assumptions made concerning the direction of propagation in time of the objects submitted to it (which determine their own action signs in relation to that of the source) and therefore the gravitational field must itself be considered a relative concept dependent on the conventions used by an observer.

Regarding the relation between the sign of charges in general and the direction of propagation in time it must be noted that energy effectively distinguishes itself from non-gravitational charges by the fact that it is naturally reversed when a particle reverses its direction of propagation in time. Indeed, in the context where a particle-antiparticle annihilation process must be considered as an event during which a particle bifurcates in time to begin propagating the same non-gravitational charges backward in time (which would effect the same kind of change as reversing the charges and keeping the direction of propagation in time unchanged), it must be assumed that the energy of the particle is reversed along with the direction of time intervals when the bifurcation occurs given that the particle now effectively propagates backward in time while its energy remains positive from the conventional forward in time viewpoint. In fact we have no choice but to consider that only non-gravitational charges are left unchanged (relative to the true direction of propagation in time) when the particle begins propagating backward in time during what appears to be a particle-antiparticle annihilation process, because energy is always released by such a process and if the sign of energy had remained unchanged along with that of non-gravitational charges when the direction of propagation in time of the particle reversed, then an antiparticle's energy would be opposite that of its particle with respect to the forward direction of time and therefore the annihilation of such a pair could occur without any energy at all being released, as I previously mentioned. Thus energy must actually reverse along the 'true' direction of propagation in time of a particle, when the particle reverses its direction of propagation in time during a pair annihilation process, just like momentum naturally re-

verses when a particle changes its direction of motion in space. The negative energy of an antiparticle simply propagates backward in time so that relative to the positive or forward direction of time it is left unchanged and from a mathematical viewpoint this interpretation fully agrees with the traditional description.

If this relational interpretation of the energy signs of particles involved in pair annihilation processes is valid then, based on the fact that we also have many reasons to believe that the gravitational properties of antiparticles are the same as those of particles, I can deduce that from a gravitational viewpoint the sign of energy is physically significant merely in relation to the direction in which a particle with that sign of energy is propagating in time. In other words, to produce an anomalous gravitational field or to respond anomalously to a gravitational field a particle would have to propagate its negative energy forward in time rather than backward as does an ordinary antiparticle. This is a simple, but very significant result whose consequences will be developed in the following sections. What must be understood is the fundamental character of the degree of freedom associated with time direction, which in a general relativistic context simply embodies the sum of all relationships of time directionality between a given particle and all the other particles in the universe. This physical property must be considered distinct from any property of time directionality which is merely statistically significant and which is associated with the flow of information, as that which characterizes the irreversible processes obeying the second law of thermodynamics.

Concerning the gravitational properties of antimatter, it appears that it is actually unnecessary to appeal to any independent constraint like the equivalence principle (which seems to require all matter to have the same acceleration in a gravitational field) to justify that antimatter should not ‘fall’ up in the gravitational field of a positive energy planet like the Earth as was often proposed before experiments began to rule out such a possibility. Indeed, any of the arguments traditionally provided to rule out the possibility of an anomalous gravitational behavior of antimatter become unnecessary once it is understood that it is effectively only matter propagating its negative energy forward in time that could experience gravitation distinctively from normal matter, while it is already known that if negative energy is to be associated with antiparticles then this energy would in fact propagate backward in time. There is thus a very good reason to assume that antimatter falls down in the

gravitational field of the Earth, but this is not an argument that we could use to rule out the possibility that some matter that would not be antimatter could perhaps be subject to anomalous gravitational interaction with ordinary matter, because there is no a priori motive for assuming that there cannot exist particles propagating negative energy forward in time. In fact, I will later explain that even the general argument against anomalously gravitating matter based on the necessary application of the equivalence principle is not really unavoidable, because it is possible to better define this principle in a way that allows for the existence of anomalously gravitating matter of the appropriate type, while retaining the general framework of relativity theory which can accommodate such a generalization.

In any case it must be recognized that all those properties of fundamental time directionality discussed above are a reflection of the fact that the sign of charges (including energy) is not only defined in relation to the direction of propagation in time of the particle carrying those charges, but is actually determined completely arbitrarily as being merely significant in relation to the similar physical properties of other particles. From a relational viewpoint it would be incorrect to assume that the direction of propagation in time of a given type of particle, carrying a unit of electric charge with a given, arbitrarily assigned positive or negative sign, is definitely the future direction, say, while the direction of propagation of the antiparticle of the same type is definitely the past, or even that there is a definite character of being an antiparticle by opposition to being a particle. The only property that can be objectively defined without referring to quantitative aspects that are not part of a given universe is the relative direction of propagation in time of two particles. Two particles with the same type of charge may be both propagating in the same direction of time or they may be propagating in opposite directions of time and this is all we can ascertain through physical means.

What must be understood is that while the relationship between the direction of propagation in time and the sign of a given charge, including energy, is a matter of coordinative definition (a definition that must be applied similarly to all processes in the whole universe on the basis of their relationships to one particular process for which an arbitrary choice of properties is assumed), once such a definition is applied the difference between the sign of time intervals and the sign of charges is an objective physical property that is not dependent on a particular viewpoint. But the viewpoint according to which a given particle is propagating a given positive charge (including

energy) forward in time remains totally equivalent to the viewpoint according to which the same particle is propagating the same type of charge with negative sign backward in time. It is never possible to tell if we are observing a positive charge particle propagating forward in time or a negative charge particle propagating backward in time in any specific instance.

It must be clear, however, that once we assume an ordinary electron to be propagating forward in time it is not possible to consider another ordinary electron as perhaps propagating backward in time while carrying a positive electric charge in this direction of time (so that the electron would still appear to be propagating a negative charge relative to the forward direction of time). Indeed, if a certain condition of continuity in time on which I will later elaborate is assumed to apply, such a backward in time propagating ordinary electron could only annihilate with an anti-electron which would be propagating the same positive charge forward in time (instead of propagating a negative charge backward in time). But this would effectively mean that certain positrons cannot annihilate with certain electrons while no constraint of this kind is observed to apply, as all electrons have the same unique probability of annihilating with any positron. Thus, if the constraint of continuity of particle world lines effectively applies, an ordinary electron must be assumed to propagate in one and only one direction of time while its antimatter counterpart must similarly be assumed to always be propagating in the opposite direction of time. Perhaps that this restriction is a consequence of the requirement which we may have to impose that *all* electron and anti-electron world-lines are somehow the manifestation of the existence of one single continuous trajectory in spacetime (which would occur if we were to assume that there is ‘only one electron’ or that all electrons are ‘the same particle’ as John Wheeler argued). The empirical constraint, however, does not specifically require the validity of this hypothesis.

On the basis of those considerations and given the previously reached conclusion that only the sign of energy with respect to a given direction of time has physical significance, it must effectively be recognized that only a particle propagating either negative energy forward in time or positive energy backward in time (in the context where ordinary matter is considered to propagate positive energy forward in time) could potentially respond in an anomalous way to the gravitational interaction. What is important to know about such a particle, which we may call a negative action particle² to

²Despite the ambiguity I still use the term ‘negative energy’ in place of ‘negative ac-

distinguish it from a particle merely propagating negative energy backward in time like an antiparticle, is that the preceding considerations regarding the relational definition of physical quantities would also mean that the particle cannot possibly be considered to have physical properties that would qualify it as responding to the gravitational field of a positive action body in an anomalous fashion that would not also be shared by an ordinary matter particle (propagating positive energy forward in time) submitted to the gravitational field of a negative action body. This must be considered an unavoidable conclusion in the context where one can physically distinguish only a difference or an equality in the signs of action of any two particles and cannot attribute objective meaning to the sign of action itself. That does not mean that there would actually be no anomalous response, only that in a configuration where all ‘anomalously’ gravitating matter is replaced by ordinary matter and all ordinary matter is replaced by anomalously gravitating matter we should observe no difference (for the most part). Thus a particle defined as having negative energy relative to the positive direction of time and which would be located in the gravitational field of a planet having opposite energy relative to the positive direction of time should behave in the same way as a positive energy particle in the gravitational field of a negative energy planet and similarly for any combination of energy signs of particle and planet, because only the relative difference in forward propagated energy signs can be considered significant. Given the preceding discussion this should be crystal clear. But that is not what is usually assumed to occur by people discussing negative energy or making quantitative predictions involving matter in such an energy state.

What is usually assumed is that a positive energy or positive mass body would attract all bodies, regardless of whether those bodies have positive or negative energy or mass, while a negative mass body would repel all bodies, again regardless of whether those bodies have positive or negative mass. It is currently believed that this is the consequence of taking inertial mass to be reversed along with gravitational mass, as would appear to be required by the equivalence principle. It must be clear however that those are not results which are ‘derived’ from relativity theory as is sometimes suggested, but merely the consequence of a choice that is implicitly made regarding

tion’ to identify such anomalously gravitating matter when the context clearly indicates that I mean negative energy propagating forward in time or equivalently positive energy propagating backward in time.

what properties should be associated with negative inertial mass while trying to be as accommodating as possible with the traditional conception of the principle of equivalence. But if I find it appropriate and indeed necessary to consider, as most people do, that inertial mass is reversed along with gravitational mass when we are considering an object with negative energy (so that the equivalence principle can be observed to apply), I cannot agree with the conclusion that is usually drawn from such an assumption. Indeed, for the response of various masses to the presence of a negative mass to be in line with common expectations, it must be possible to determine the sign of mass or the sign of action of particles in an absolute non-relational manner, because we are assigning the attractive or repulsive character of the gravitational field in precisely such an absolute manner (the field is either repulsive for everything or attractive for everything) which I believe could never be justified.

I think that it cannot be assumed that a negative mass is repulsive in an absolute invariant way, because it would not be possible to tell relative to what reference point the distinctiveness of this character is defined given that positive mass cannot be used as a reference if it is itself absolutely defined (not merely in relation to the opposite negative masses). I will explain in a later section of this chapter why it is that the assumption that a negative inertial mass is associated with a reversal of the sign of action, far from having the undesirable consequence of allowing absolutely defined physical properties into physical theory (if there could ever be such a theory) actually has for consequence (when the inertial properties of matter are well understood) to give rise to a description of the gravitational interaction between positive and negative mass bodies that is in perfect agreement with the requirement of relational definition of the sign of mass or energy. All that would then remain to understand is how the equivalence principle can still be satisfied by such a description. For that purpose I will provide arguments to the effect that a simple reconsideration of the true significance of the principle of equivalence, and a better understanding of its motivation in the principle of relativity of accelerated motion, allows its foundations to be preserved while enabling the more consistent relational viewpoint on the sign of mass to be retained and to actually be integrated into the core mathematical framework of relativity theory by introducing a slight modification to this classical theory of gravitation that is actually a simple generalization of it. In order to further justify this approach, I will first try to identify what should be the true properties of negative action matter and why we should

not expect such matter to behave in ways that would make it undesirable not only from the viewpoint of the requirement of a relational description of physical quantities, but with respect to other constraints and other physical principles which we can be confident must also be obeyed.

2.3 Our current understanding

Before addressing the question of how a negative energy particle would actually behave we may first want to explore what the current situation is regarding the notion, or indeed the problem of negative energy. For this purpose, it should first of all be noted that for many reasons no one seems to like the idea that there could exist negative energy particles. Thus it is no surprise that one of the most basic and often implicit assumption that enters our description of physical reality is that energy must always be positive. There are many different formulations of that requirement which impose various degrees of conformity to the hypothesis that matter cannot find itself in a state that would be observed as having negative energy. In its least restrictive form this condition is called the weak energy condition and merely constitute a statement about the positivity of the components of the stress-energy tensor (the most general representation of the energy content of matter). More constraining conditions have also been proposed, among which is the appropriately named strong energy condition which if obeyed under all circumstances would mean that gravity must always be attractive (between all forms of matter which would then be allowed to exist). Those conditions are used as rigorously defined hypotheses in various theorems dealing with the behavior of matter under the influence of the gravitational interaction.

The problem is that it was found at some point that configurations involving negative energy densities are actually allowed to occur in quantum field theory. This does not mean that negative energy particles are explicitly allowed by the current theories, but merely that unlike what we would expect from a classical viewpoint where the vacuum is described as a total absence of matter, quantum field theory allows for the local density of energy to not always be positive definite, even in a context where only positive energy matter is present. A well-known experiment illustrates the kind of phenomena involved. It requires placing two parallel mirrors a very small distance apart in a vacuum so as to forbid some states, which would normally exist in the vacuum, from being present in the space between the mirrors,

as a consequence of the incompatibility of their characteristic wavelengths with the spatial constraints imposed by the presence of the mirrors. The predicted result, which is effectively observed, is that there should arise a small pressure pulling the mirrors together as a consequence of the proportionately larger pressure exerted from the outside, which is actually caused by a decrease in pressure from between the mirrors that can be attributed to the restriction imposed on which virtual particles can be present in this volume. This is of course the phenomenon known as the Casimir effect. It is clear though that we are not directly measuring a negative energy density in such an experiment, but merely the indirect effects of an absence of some positive contribution to vacuum energy, which is then assumed to imply that the energy density is negative in the small volume between the mirrors. But even that kind of manifestation of negative energy is assumed to be so serious a problem by some theorists that they suggested that the description of the vacuum as involving virtual particles coming in and out of existence is actually only a mathematical trick and does not reflect what is really going on in the absence of ‘real’ matter.

However, this aversion for whatever is negative of energy is not shared by all informed authors and some more open-minded specialists have tried to address the issue of negative energies as they occur in quantum field theory and in so doing gained some significant insights into what exactly is allowed by a quantized description of the vacuum. A modified version of the weak energy condition was thus proposed that allows to take into account the fluctuations of energy which arise in the quantum realm. This condition, which is appropriately called the averaged weak energy condition, involves only quantum expectation values of the stress-energy tensor averaged over some period of time during which the observations are assumed to occur, rather than idealized measurements at a point. A feature of the constraint provided by this condition is that it allows for the presence of large negative energies over relatively large regions of space if there is a compensation by the presence of a larger amount of positive energy somewhere during the time period over which the observations are made. It was indeed found out [1, 2, 3, 4] that quantum field theory effectively places strong limits on the values of negative energy density that can be observed over finite periods of time under various conditions. What emerges from those developments is that there appears to be a constraint on the magnitude of negative energy that can be observed and it indicates that negative energy can be merely as large as the time interval during which it is measured is short. I believe

that this is indicative of the fact that while negative energy states cannot be ruled out as strictly forbidden, they should also clearly not be expected to materialize in stable form in the context where we are dealing with ordinary matter configurations for which the particles are already predominantly in positive energy states.

A similar limitation can also be observed to restrain another form of negative energy that occurs in the presence of an attractive force field, even in a classical context. Indeed, the energy contained in the force field between two particles submitted to an attractive interaction must be considered negative. This is because work and positive energy must be provided to separate two particles attracted to one another in such a way and given that it must be assumed that the attractive field responsible for this interaction would contain no energy at all when the particles are separated by a distance that tends to infinity (in the context where the strength of the field associated with a long range interaction decreases in proportion with the square of the distance, so that it must effectively be null when this distance is infinite) then we must conclude that the energy initially contained in the same attractive force field when the particles were near one another was actually negative (so that adding positive energy can produce a null final value). It was effectively observed that the energy of a bound system formed of many interacting particles is lower than the sum of the energies of those particles when they are free. Thus the energy contained in an attractive force field must definitely be considered negative, as this energy is required to provide the negative contribution that reduces the energy of the whole bound system. The additional energy that was present before the formation of such a system is in fact released (through the emission of radiation for example) when the system is created, but except for the additional negative energy contained in the attractive force field the system would be identical, in terms of its matter particle content, to what it was initially and therefore we definitely need the negative energy. This is made more obvious when we consider larger systems like those bound by the gravitational interaction. It was effectively shown that even a system as large as the Earth-Moon system has an asymptotically defined total mass (providing a measure of its total energy) which is smaller than that of its constituent planets and observations confirm this prediction. Therefore, it is clear that the energy contained in the gravitational field maintaining the two planets together must be negative.

What is crucial to understand regarding the situation described above, however, is that even if we must acknowledge the existence of a well-defined

negative contribution to the energy of some physical systems that diminishes their total energy, it is again not possible to measure that energy directly and it can merely be deduced to occur from the behavior of the positive energy subsystems which are submitted to the attractive interaction. Here also the negative energy must be associated with the virtual particles that mediate the interaction and cannot be measured independently from the total energy of the bound systems which usually remains positive. It is simply not possible to isolate the attractive field of a bound system from its positive energy sources and this is true for systems of any size. It would nevertheless certainly be a concern if the negative binding energy of a system made of positive energy components could become so negative as to make the total energy of the bound system itself negative. Once again, however, it was shown that there are unavoidable theoretical constraints on the values that observable total energy can take. It was shown [5, 6, 7, 8, 9, 10, 11, 12, 13], concerning the gravitational interaction in particular, that the energy of matter (everything except gravitation) plus that of gravitation is always positive when the dominant energy condition is assumed to be valid, which effectively amounts to assume that the energy of the component particles is itself positive. If we compress positive energy matter too tightly it simply collapses into a black hole of minimum surface area and maximum energy density before the magnitude of the growing negative gravitational potential energy becomes larger than the positive energy of the matter. Thus positive energy matter cannot turn into negative energy matter through an increase of negative gravitational potential energy. What must be retained from the previous considerations, therefore, is that even though it is often present, negative energy seems to never be measurable. But this conclusion is valid merely under the condition that we are dealing with situations where matter was already in a positive energy configuration to start with. It must be clear, however, that we still have no argument to rule out the possibility that there may exist configurations where the component particles themselves would have negative energies and for which there would exist constraints similar to those unveiled here enforcing the *negativity* of energy.

In a previous section of this chapter I mentioned that it is desirable from a certain viewpoint to consider antiparticles as propagating negative energy backward in time. Indeed, if antiparticles are propagating backward in time, as the reversal of their non-gravitational charges clearly suggests, then they *must* have negative energy relative to the direction of time in which they

are propagating (which is the past), so that relative to the opposite direction of time (which is the future) they would still appear to have positive energy, as required. In fact, it was discovered a long time ago by Paul Dirac (when he achieved his unification of special relativity and quantum theory) that there is a mathematical requirement for the existence of negative energy states. Indeed, it turned out that in order to obtain Lorentz invariant equations for the wave function one had to sacrifice the positivity of energy. After having considered various possible interpretations for what in nature could possibly correspond to those negative energy states Dirac concluded that it required the existence of a new category of particles, the antimatter particles, which would consist of holes in a filled distribution of such negative energy matter. The antiparticles were eventually described by Feynman as particles propagating backward in time, which allowed to fulfill the mathematical requirements imposed by the existence of the negative energy states (by providing an interpretation for those transitions which were predicted to involve a reversal of energy) without requiring the presence of the filled negative energy continuum. But in the process it seems that the discovery that particles could actually occupy negative energy states, which appeared to be implied by the original developments, was somehow forgotten and lost in the details of the proposed solution. This indifference was probably justified by the fact that antiparticles could still be considered to have positive energy for all practical purpose. But what is usually ignored is that while attributing a positive energy to antiparticles may appear more ‘reasonable’ than assuming that those particles propagate negative energy backward in time, such a choice would actually imply that it is the particles themselves (by opposition to antiparticles) which must then be considered to carry negative energy backward in time, because it must be either that or the opposite. This is what the subtleties of the quantum mechanical definition of energy seems to require that was not apparent classically.

The reluctance to recognize the true physical significance of the requirement of negative energy states is probably also in part a consequence of the apparently insurmountable difficulties which would be associated with the possibility for particles to occupy those physically allowed states. First of all it is certainly true that if antiparticles were submitted to anomalous gravitational interaction as a consequence of propagating negative energy backward in time we would run into a number of problems, because it was demonstrated some time ago [14] that if, for any reason, antimatter was to be found experiencing repulsive gravitational interactions with ordinary

matter we would run into a number of problems ranging from violations of the conservation of energy and up to the undesirable and unlikely (from a theoretical perspective) possibility of producing perpetual motion machines. But an analysis of the arguments presented against the possibility for anomalously gravitating antimatter has led me to conclude (for reasons which will be explained later) that the problem effectively has to do merely with the possibility for antimatter ‘as we know it’ to experience what we may call antigravity. It cannot be considered to mean that matter in a true negative energy state (propagating negative energy relative to future directed time intervals) could not exist and experience anomalous gravitational interactions with ordinary matter without violating the principle of conservation of energy or the second law of thermodynamics, because matter in such a negative energy state may also by necessity have different properties from those already known to characterize antimatter, in particular with regards to non-gravitational interactions.

Nevertheless, most people today seem to consider that the developments that followed the introduction of the early theory of relativistic quantum mechanics and which gave rise to modern quantum field theory have eliminated the problem of negative energy states, which can now be considered a mere artifact of the former single particle theory. Thus, the predicted negative energy states would simply be unphysical solutions that must be discarded as irrelevant to physical reality. But it must be clear that this is effectively what we are doing here. We are rejecting the possibility that a particle could be found in a whole set of states that are allowed by the most basic equations without providing any justification as to why those states should be forbidden. Indeed, upon closer examination it becomes clear that if ‘true’ negative energy states do not explicitly arise in quantum field theory it is not because the structure of the theory forbids them, but simply because we *choose* to ignore those solutions to start with and then integrate that choice into the formalism. More specifically it turns out that what prevents negative action particles from showing up in quantum field theory is merely a choice of boundary conditions for the path integrals that provide the probability amplitude for transitions involving particle trajectories in spacetime. There are several possible choices for expanding those integrals which all constitute valid solutions of the equations of the theory, but only those solutions propagating positive frequencies forward in time and negative frequencies backward in time are usually considered to be physically significant, while the solutions propagating negative frequencies forward in time and positive

frequencies backward in time, which are also valid from a mathematical viewpoint, are systematically rejected. But this actually amounts to retain only the positive action portion of the theory, while ignoring all transitions involving negative action particles. There is no other origin for the often mentioned conclusion that quantum field theory does not involve negative energy matter. It is by our very choice that we reject all transitions involving negative action particles.

In order to make the choice of boundary conditions responsible for the absence of negative action particles in quantum field theory more acceptable it is sometimes suggested that the negative energies predicted by the single particle relativistic equations are simply transition energies, or differences between two positive energy states and there is obviously no reason why those variations could not be negative if they can be positive. But no explanation has ever been provided for why the same reasoning could not be applied to the energy states themselves, which are also energy differences given that the energy of a particle is always defined in relation to the zero level of energy associated with the vacuum in which it propagates. There is no justification for this arbitrary distinction between transition energies and particle energies, except for the satisfaction that is obtained by the physicist in having easily disposed of an embarrassing problem. It may of course be argued that there is nothing wrong with those methods, given that they appear to be validated by experimental results. Indeed, we have never observed interferences by negative action particles into the outcome of any experiment conducted at any level of energy and to any degree of precision. But I would like to emphasize that this still doesn't constitute an explanation for the absence of negative action particles.

Thus, the problem I have with the modern approach to quantum field theory is that the formalism is generally introduced in a way that encourages us to believe that after all no particle is actually propagating backward in time with negative energy and that a positron is really just another particle identical to the electron, but with an opposite electrical charge. However, this viewpoint does not only complicate things unnecessarily as a consequence of rejecting the possibility for electrons and positrons and all other particles and their related antiparticles to actually consist in the same particles observed from different perspectives, it is also completely ignorant of the requirement of a relational definition of any physical property dependent on the fundamental time-direction degree of freedom. But if we choose to recognize the validity and the greater value of the viewpoint defended here and according

to which antiparticles are really just ordinary particles propagating backward in time, then we must accept that there definitely exist in nature particles which are known as carrying negative energies and if the arguments provided above concerning the arbitrariness of the current restrictions imposed on the propagation of those negative energy states are valid then we would have to conclude that there should necessarily also exist particles with such energies propagated forward in time and which could be submitted to anomalous gravitational interactions in the presence of ordinary matter.

2.4 The negative mass concept

When discussing the issue of negative mass what must first of all be understood is that if the physical property of mass is to have any polarity associated with it, such that we could attribute to mass either a positive or a negative sign, then this polarity must be directly related to the sign of action, that is, to the sign of energy relative to the positive direction of time. This is because, as I previously emphasized, the sign of action is the only physical property from which the attractive or repulsive character of the gravitational interaction between two bodies could depend. We may thus attribute positive mass to a positive action particle and negative mass to a negative action particle. Mass being a Newtonian concept its polarity must be determined in relation to a particular Newtonian gravitational field. From this viewpoint the sign of mass of a given particle could effectively be understood as determining the response to the gravitational field of a given source, in the sense that it would determine the *direction* of the gravitational force exerted on such a particle. If we may consider the gravitational field of the source (represented by a vector in Newtonian mechanics) to be uniform, then only its own direction or polarity (which we may assume to be dependent merely on the sign of mass of the source when its position is assumed to be fixed) would be decisive in determining the kind of response experienced by a given type of mass submitted to it. Equipped with such a definition we can meaningfully discuss the problem of the gravitational interaction of negative action particles with positive action particles and with themselves as the problem of the gravitational interaction of positive and negative masses. This will allow us to better grasp the significance of the assumptions that will form the basis of the new interpretation of negative energy matter which I shall propose and therefore, also, to gain better confidence in their validity, even

in the more appropriate context of a general relativistic theory.

If we may agree on those requirements, then I think that what must emerge is that if it is indeed important to have a well-defined concept of negative mass then it also seems that such a negative mass must be negative in all respects. That there could be a difference between the sign of gravitational mass and the sign of inertial mass is usually considered to be forbidden merely by the general theory of relativity which is effectively founded on the principle of equivalence which requires the equality of gravitational and inertial masses. However, I think that if this hypothesis is justified it is not because our concept of mass polarity must comply with some perceived requirement from general relativity theory, but because it would not be acceptable to attribute mutually exclusive values to a single unique physical property. Thus, I do believe that the mass of any particle or body should be either definitely positive or definitely negative (but still in a relational way), regardless of whether we are considering gravitational mass or inertial mass, if the concept itself is to have any consistent physical meaning. But unlike most theorists I do not consider that this requirement must be assumed to imply the kind of behavior that is usually attributed to negative mass matter, where gravitational repulsion is an intrinsic property of this type of matter itself, independently from the sign of mass of the matter with which it is interacting. This is indeed the conclusion I was able to draw based on the outcome of the previously discussed analysis of the constraints imposed by a relational definition of the sign of energy, for reasons I will now explain.

The difficulty I originally met when I first began to explore the possibility that inertial mass could be reversed along with gravitational mass when we are dealing with negative mass matter is that if both the gravitational mass and the inertial mass are to be negative at once then it seems that there could occur situations where the principle of inertia would be violated (I will explain what motivates this belief below). I was able to understand, however, that this is merely a consequence of the inappropriateness of current assumptions regarding what we should expect to be the behavior of matter with both a negative gravitational mass and a negative inertial mass. Actually, despite the fact that it is usually taken for granted that we know for sure at least what the behavior of matter with positive mass is, because we routinely observe gravitational phenomena involving this kind of matter and there can be no mistake here, I will explain that this is not entirely the case and that there is still much confusion as to even what we should expect concerning the response of positive mass matter to a concentration of negative mass.

Currently it is assumed that given that positive mass matter gravitationally attracts all matter and resist the action of any force exerted on it, then this must be an intrinsic property of such positive masses. On the other hand, it is usually assumed that two choices exist for what could possibly characterize the behavior of matter with a negative mass. The situation we have right now is thus the following.

First of all, we must assume that *gravitational* mass is effectively negative when mass is reversed. This allows to obtain gravitational repulsion when only the mass of the source (the active gravitational mass) is negative, because it reverses the polarity of the Newtonian gravitational field to which any passive gravitational mass is submitted and therefore should at least reverse the force exerted on positive mass bodies. But once this is recognized it is usually considered that two possibilities effectively exist for a negative mass particle submitted to a given gravitational field, depending on whether *inertial* mass is assumed to remain positive or is itself also negative. Here the inertial mass of a particle is assumed to determine the response of that particle (actually the direction of its acceleration) to any force, including a gravitational force, while the gravitational mass of the same particle is assumed to determine both the polarity of the gravitational field it produces and the response of the particle to a gravitational force. If we agree with those assumptions then we would have to conclude that a negative gravitational mass particle with a negative inertial mass, should actually respond normally to any gravitational force field (because the nature of its response is changed twice, once by the reversal of its inertial mass and once by that of its gravitational mass) while its response to non-gravitational forces would be reversed (same force, opposite acceleration), as current assumptions concerning the reversal of inertial mass would require. But we must also keep in mind that the fact that this kind of matter would respond normally to gravitational force fields would, under the current assumptions, still mean that it is repelled by matter of the same type, because the gravitational field produced by such matter is also assumed to be reversed. Thus such negative masses would repel masses of all signs, be repelled by other negative masses and be attracted to positive masses, still under the hypothesis that the above stated commonly accepted assumptions are valid. Given that it is usually considered that in a general relativistic context all mass (gravitational and inertial) must be negative, this is the choice that is usually retained as defining the behavior of negative mass matter if it could exist.

But despite the support that is usually granted to such a conception of

negative mass or negative energy matter I think that enormous problems would arise if it was retained as a valid proposal. Some of those problems, involving black holes and the second law of thermodynamics, will be discussed later, but even if we remain at the level of classical Newtonian dynamics we can readily identify one very serious problem which is that the existence of such matter would allow violations of the principle of inertia (considered as a generalization of Newton's first law) or the very idea that no physical system can accelerate without work being done on it by an external force. This is because indeed, as stated above, from the current viewpoint a negative mass body would both repel positive mass bodies and be attracted to them. Such a combination of features could then give rise to unlikely phenomena like pairs of opposite mass bodies chasing one another and in the process accelerating to infinite velocities, still without any external energy input. The fact that energy would in principle be conserved under such conditions (because the energy gained by one of the bodies would be opposite that of the other) is no consolation, because we are dealing here with a much more serious and basic violation. Indeed the problem I see is that there would be no equal and opposite force to that applied on a given body that could be attributable to its assumed interaction with the other body and this would be a violation of the principle of action and reaction (Newton's third law), which is one requirement that in all fairness we should recognize as being even more fundamental than that of conservation of energy, because if it does not rigorously apply then absolutely anything could occur and under such conditions we could not give much of even the principle of conservation of energy. However, I think that what those observations show is not the unphysical nature of negative mass, but merely the ineffectiveness of the traditional approach to describe the behavior of this kind of matter. It is important to mention, by the way, that even though this hypothetical situation of accelerating opposite mass pairs has been described by other authors in the past, none of them has ever recognized that what it actually demonstrates is the inconsistency of the current notion of negative mass, which I believe is illustrative of the state of denial in which most people remain concerning the possibility that there could actually exist negative mass matter.

What is also significant concerning the unlikely phenomenon described above is that it would necessarily be the positive mass bodies that would be chased in this way, while the negative mass bodies would inevitably be those trailing them. But isn't it strange indeed that there should be such a clear

and decisive distinction between what constitute the role of positive masses and what constitute that of negative masses? Doesn't it seem like there is something wrong with such a hypothetical phenomenon? Shouldn't we only be allowed to define the property of gravitational attraction and repulsion in such a way that we could not observe such mass-sign-distinguishing behavior? What I have understood is that the unease we may experience in face of the strangeness of such phenomena is in fact justified. Indeed, it does not just seem like there is something wrong here, because what we have just described is actually the perfect example of an attempt to distinguish a physical property (the positivity of mass or the attractiveness of gravitation) despite the absence of any reference in the physical universe to which that arbitrary distinction could be related, which violates the very basic requirement of relational determination of physical properties discussed above. The mistake which is made by assuming the validity of the traditional viewpoint is that we suppose that we can define attraction and repulsion in an absolute (non-relative) manner such that one kind of mass always attracts all kinds of masses regardless of their polarity and another always repels all masses, still regardless of their sign, as if attractiveness and repulsiveness were intrinsic aspects of one and the other type of mass.

However, if the sign of mass is to be considered a meaningful physical property of elementary particles then it must be taken to indicate that there can be a reversed or opposite value to a given mass and this reversed value can be considered to be reversed merely in relation to a non-reversed mass and to nothing else. A mass cannot be considered to be reversed with respect to an absolute point of reference lacking any counterpart in the physical universe. Therefore, if a gravitational field is to be assumed repulsive as a consequence of the reversed (negative) sign of the mass of the matter that is the source of the field then this gravitational field should be repulsive only for an unchanged (positive) mass particle and not with respect to other negative masses. It would be incorrect to assume that the attractive or repulsive nature of gravitational fields depends solely on the sign of mass of the source itself, because no distinction exists for the sign of a mass other than its sameness or oppositeness compared to that of another mass. That does not mean that the field itself must be assumed to change as a consequence of the reversal of the sign of mass of the particle experiencing it (even though that may be one way to describe things if other conventions are adopted for the sign of mass itself as we will see later), but merely that the response of a negative mass particle to a given gravitational field must be reversed in

comparison to the response we would expect from a positive mass particle submitted to the same field, despite the associated reversal of the inertial mass of such a particle. If that was not the case, then I think that we would have to conclude that negative mass is effectively forbidden.

If the incorrect hypothesis on which the traditional approach is based regarding the effect of a reversal of inertial mass nevertheless allows to successfully (from my viewpoint) predict that a positive mass would be repelled in the gravitational field of a negative mass, it is simply because we assume the right inertial properties for the positive mass matter submitted to the gravitational force of the negative mass. Thus the positive mass responds in the appropriate way to the gravitational force exerted by the negative mass which is correctly assumed to be a repulsive force given that the gravitational field produced by the negative mass is necessarily opposite that which would be produced by a positive mass of similar magnitude located in the same position. The problem is that given that it seems that we cannot expect the same kind of behavior from a negative mass submitted to the gravitational field of a positive mass, then it would appear that the behavior of both positive and negative masses is the consequence of some predetermined property of absolute attractiveness and repulsiveness (that cannot be related to any property of the source defined with respect to a property of the matter with which it interact) associated with the gravitational fields emanating from positive and negative masses respectively.

The difficulty to which the traditional interpretation gives rise is also made apparent when we consider the case of a negative mass in the gravitational field of another negative mass, given that now the negative mass would be repelled by the same negative mass matter (because the gravitational force is unchanged but the response to this force would be reversed), while on the basis of the relational definition of mass sign there should be no difference between this case and that of a positive mass in the gravitational field of another positive mass (which is symmetric to the other case under exchange of mass signs). The appropriate outcome could only be obtained if in addition to the assumption regarding the nature of the gravitational force between two negative mass bodies it is also assumed that the reversal of the inertial mass of the negative mass body submitted to this force actually changes nothing to the response of that body to the attractive force of the other negative mass body. Thus the problem of the absoluteness of the attractive or repulsive nature of the gravitational field arises as a direct consequence of the current assumptions regarding the effect of a reversal of

inertial mass. It is only in this context that the direction of the Newtonian gravitational field associated with a concentration of matter of positive or negative mass sign acquires an absolute meaning and is not merely dependent on a convention as to what should be the sign of mass of the matter that is the source of this field.

Even if merely as a consequence of the previously discussed considerations regarding the relative nature of the sign of energy (as dependent on the direction of propagation in time of a particle) and the purely conventional (subject to an arbitrary coordinative definition) significance of the sign of action it would appear that a consistent notion of negative mass would require that it is the relative difference or absence of difference between the mass signs of two gravitationally interacting bodies that determines the attractive or repulsive character of this interaction, so that two negative mass bodies should be submitted to the same mutual gravitational attraction experienced by two positive mass bodies and would also repel ordinary positive mass bodies and be repelled by them, unlike is usually assumed. But the fact that it is often not even fully understood that negative mass should effectively be associated with negative action is illustrative of the confusion that surrounds the whole question of negative energy and gravitational repulsion, because there should be no doubt that if it is possible for the sign of mass of a given body to be negative in some way, then this would necessarily have to occur as a consequence of the fact that this body has negative energy, or more precisely negative action. In any case, if the traditional viewpoint allows predictions that violate the expectations of a relational definition of mass sign it is precisely because it allows to assume that there can be an absolute character of attractiveness or repulsiveness associated with a given sign of mass. To be fair, I must acknowledge that some authors did suggest in the past that the gravitational interaction should perhaps be repulsive between bodies of opposite mass signs while it would be attractive between negative mass bodies (just as it is between positive mass bodies), but simply on the basis of the fact that the sign of the gravitational force that is obtained by reversing the sign of one of the masses in Newton's equation for universal gravitation would itself be reversed, while it would be unchanged if the signs of the two masses were together reversed.

But even though it is not necessarily wrong to suggest that the repulsive or attractive nature of the gravitational interaction is determined by the sign of the force in Newton's equation for universal gravitation, it is only when

we realize that the sign of mass must be related to the sign of action that we can begin to understand why it is that there should be a symmetry under exchange of positive and negative masses. This is because, as I previously mentioned, positive action states are related to negative action states by a simple convention regarding the sign of energy and that of time intervals, so that the sign of action is itself a purely relative notion. There must consequently be a symmetry under exchange of positive and negative action matter, which would then require the behavior of positive masses in relation to themselves and in relation to negative masses to be similar to that of negative masses in relation to themselves and in relation to positive masses. I may add that in such a context it appears that the suggestion that perhaps negative mass bodies have never been observed simply because they do not assemble themselves into larger masses (as a consequence of their assumed absolute gravitationally repulsive nature) cannot be valid and if negative mass matter exists then alternative arguments would have to be proposed to explain this absence of observational evidence. In a later portion of this report I will effectively explain how it is possible to reconcile the apparent absence of large scale concentrations of gravitationally repulsive matter with a more consistent notion of negative mass matter.

The contradictions of the traditional conception of negative mass or negative energy matter can be illustrated by using a rarely discussed thought experiment. It has in effect been proposed that the sign of energy of a negative mass particle could be determined by measuring the energy lost or gained while raising or lowering the particle in the gravitational field of some large object. Now according to the traditional conception if we were to raise a negative mass body in the gravitational field of a positive mass object like a planet we would have to produce work and exert a force directed downward because the inertial mass of the body is negative, which according to the traditional viewpoint means that it responds perversely to the applied force. But then it is also the case according to this same viewpoint that the gravitational force exerted by the planet on the body should be attractive, because the planet has positive mass. Thus we would be in the situation where we would have to exert a force downward to raise a negative mass body in the gravitational field of a planet that exerts an *attractive* force on that body. I do not know to what extent people actually believe in the validity of such a conclusion, but I think that faced with such absurdities one has to come to realize that the contradictions involved are a clear indication that the traditional assumptions regarding the behavior of negative mass or

negative action matter are incorrect and that a better interpretation of what such a state of matter may involve is required.

Despite the fact that the question of the validity of the traditional conception of negative mass matter had never been clearly analyzed before, it is no doubt the general feeling that there is something wrong with the possibility of observing phenomena of the type described above (including that where pairs of opposite mass bodies accelerate without any external force being applied on them) which is responsible for having transformed the idea of negative energy or negative mass matter into the synonym of nonsense it has become in the minds of so many physicists. But, is negative mass really to blame here or could it be that we are not considering the right possibility? There is of course, even under the conventional assumptions regarding the response of negative mass particles to applied forces, another possibility which is that when gravitational mass is negative, inertial mass may remain positive for some reason. Of course that would not only appear to contradict the equivalence principle, as is already understood, it would also, if I am right, itself be nonsense, as we would have to assume that one single physical quantity related to one single particle (the mass of that particle) is at once both positive and negative for the same observer. The latter problem has never been discussed, but I think that it is actually the strongest argument against this second possibility. We may nevertheless begin by exploring the consequences of such a choice.

Under the same commonly held assumption to the effect that the response of a particle to any force is dependent on the sign of its inertial mass we would have to conclude that a negative gravitational mass body to which a positive inertial mass would be attributed would respond anomalously (in comparison to the response expected of a positive mass) to any gravitational force field (because the nature of the response is changed only once by the reversal of its gravitational mass), while its response to non-gravitational forces would be unchanged (same force, same acceleration), because the inertial mass remains positive or unchanged in comparison with that of positive mass bodies. Therefore, if material bodies were to exist that would be made of such negative mass matter they should, from the traditional viewpoint, gravitationally attract one another (as do positive masses), repel positive mass bodies and also be repelled by those same positive mass bodies. As a consequence, we would observe no violation of the principle of inertia in this case and also no acceleration without work. If this behavior was to be

observed it would in fact be possible to exchange all positive mass bodies with negative mass bodies and vice versa and no apparent change in the phenomenology of the gravitational interaction would be detectable, because gravitational repulsion would only occur when there is a difference in the signs of the *gravitational* masses which are interacting. Thus from a purely phenomenological viewpoint there would be equivalence between positive and negative mass bodies. Given the previous discussion regarding the necessity of a relational determination of the sign of energy, which would here be a requirement for the relational determination of the sign of mass, this situation would appear more appropriate, because indeed it would be impossible in principle to differentiate any intrinsic property of gravitational attraction or repulsion and only the difference or the equality of the signs of *gravitational* mass of two particles would be physically significant. The problem that most people would have with this possibility, however, is that it would explicitly violate the equivalence principle, because positive and negative gravitational masses would respond differently to a given gravitational field, produced by a given matter distribution, even if they are located in the same local inertial frame of reference.

But I think that even before we consider the issue of the apparent incompatibility with the principle of equivalence we must first of all ask how could it be determined which of the two types of matter would indeed have the *inertial* mass opposite its gravitational mass? And then it is obvious that this question could never be settled (because we could never decide which matter effectively has a negative gravitational mass) and yet it would in such a context be highly meaningful as we do assume a physical difference analogous to an absolute distinction in this respect between positive and negative mass bodies. Indeed, why would the inertial mass remain positive when the gravitational mass is reversed. It is only confusion to pretend that there are multiple aspects of mass and that each of those independent mass properties can have a different sign. An electric charge is either positive or negative and mass clearly defined as the charge associated with the gravitational interaction must also be either positive or negative and this is actually all that the equivalence principle requires I believe. In this context I think that we would be right to object trying to save the principle of inertia by assuming that negative masses could at once also have a positive inertial mass, because this would indeed violate the equivalence principle, not because different masses could accelerate in different directions in a gravitational field, but to the contrary because indeed the *same* inertial masses could effectively respond

differently to a given gravitational field, which would then really mean that there definitely cannot be equivalence between a Newtonian gravitational field and acceleration and this would indeed be a problem for relativity. Clearly there is still something wrong, even with the second alternative that is traditionally considered for the attribution of negative mass.

The preceding discussion should then have made clear the fact that there are two issues regarding negative mass. First, if we accept the requirement for a relational definition of the attractive and repulsive character of a gravitational field, then we must conclude that the currently favored assumption for what would be the behavior of negative mass bodies, having at once negative gravitational mass and negative inertial mass, is incorrect, because as I explained it would involve absolutely defined properties of attractiveness and repulsiveness that would not depend merely on the difference or equality of the signs of the interacting masses. But if we consider the other traditionally considered (but not favored) possibility for the definition of negative *gravitational* mass, we may obtain the required relational definition of gravitational attraction and repulsion, but as I have explained a distinct problem would arise. Indeed the appropriate behavior to be expected of negative mass matter would then have to be that which we currently assume to be possessed by particles with a contradictory definition of their mass sign, which is not only objectionable on the basis of consistency, but which also violates the equivalence principle in a way that cannot possibly be allowed (same mass, different response) if relativity theory is to be retained as a valid theory (if we were to accept this possibility then there would be no reason why relativity itself should still be required). Arguing that the problem here is with the notion that mass is at least in part the same, while this identity of mass signs actually applies merely to a different property of mass which we would call inertial mass, so that the ‘real mass’, which we would call the gravitational mass, could be different, would in my opinion not just be confused, it would be nonsense. What is positive cannot also at the same time be negative if this polarity is to have any meaningful physical significance. Mass is not an abstruse, complicated property with multiple independent and yet interrelated aspects, it is the gravitational charge and even though the stress-energy tensor replaces mass as the source of gravitational fields in a general relativistic context, the lessons learned here are still valid and significant even in the context of the modern theory of gravitation.

It took me some time to realize that the problems we are dealing with here (if we are willing to recognize that the whole question of identifying the

properties of negative energy matter is not itself insignificant) originate from what is usually assumed concerning the observed response to any force field in the case of a body with negative inertial mass. It is only after a rather long process of getting to understand the meaning of the phenomenon of inertia that I was finally able to gain the insight required to solve the problem of identifying the actual properties of negative mass matter in the context where we consider it a consistency requirement to impose on such matter that it should have both a negative gravitational mass and a negative inertial mass. Keep in mind that this explanation will be easier to grasp when the consequences of the integration of such a concept of negative energy matter to the modern theory of gravitation will have been more thoroughly explored. Basically what must be understood is that the direction of the equivalent gravitational field experienced by a given mass in a frame of reference in which it is accelerating, even in the absence of nearby matter inhomogeneities, is in fact dependent on the sign of the mass that is accelerating. The consequence of this hypothesis is that the inertial *force* associated with a given acceleration is left invariant even if the sign of inertial mass is itself reversed along with the gravitational mass for a negative energy particle.

In order to appreciate the following discussion at its true value it is essential to remember that relativity theory effectively implies that there exists a Newtonian gravitational field exerting a gravitational force on a positive mass body which is accelerating relative to a local inertial frame of reference, even far from any large mass. The existence of the inertial force associated with this equivalent gravitational field is what allows a dynamic (by opposition to static) equilibrium to occur when an external force is applied on a body which gives rise to an acceleration. Indeed, in the accelerated frame of reference relative to which a positive mass body submitted to an external force does not accelerate a gravitational force is present which balances the applied external force and this is what explains that there is no acceleration of the body relative to this particular (accelerated) frame of reference. In fact, the equivalent gravitational field is a general feature of acceleration and is present in any accelerated frame of reference, but in the absence of an external force to balance the associated inertial force the equivalent gravitational field only serves to determine the local inertial frame of reference associated with free fall motion. Indeed, given that the force associated with the equivalent gravitational field is a gravitational force we must conclude that when the force responsible for the acceleration is itself gravitational we are effectively in a situation where there would appear to be no force at all. It is therefore pos-

sible to assume that what determines the local inertial frames of reference relative to which a positive mass experiences no gravitational force is the local matter distribution which is the source of the applied gravitational forces which are balanced by the inertial force which would otherwise be present relative to those reference systems (this is the essence of the insight that led to relativity theory). In any case it is clear that the inertial force attributable to an equivalent gravitational field is always directed opposite the direction of the external force which gives rise to the corresponding acceleration for a positive mass body and this means that the direction of the equivalent gravitational field experienced by a positive mass body is opposite the direction of its acceleration, that is, opposite the direction of acceleration of the frame of reference relative to which this equivalent gravitational field exists. But what would occur if we had a negative mass body in place of a positive mass body?

First of all, it must be clear that the gravitational force $\mathbf{F}_g = m\mathbf{g}$ on a particle of mass m attributable to a given matter distribution would be reversed if the mass of the particle was reversed, because the Newtonian gravitational field vector \mathbf{g} at the particle's position would be left unchanged (because the matter distribution that is the source of the field does not change), while the sign of mass of the particle experiencing the field would be reversed. Now the problem usually is that when we want to determine the response of a particle to some gravitational force \mathbf{F} using Newton's second law $\mathbf{F} = m\mathbf{a}$, if the mass of the particle is reversed (negative) then the resulting acceleration \mathbf{a} would appear to have to be opposite that experienced by a positive mass submitted to the same force (the acceleration would be in the direction opposite that of the applied force). This is the traditional conception regarding negative mass. But if we consider things in a more general context, where Newton's second law would be an equation expressing the equilibrium between external forces \mathbf{F}_{ext} and the inertial force $\mathbf{F}_i = m\mathbf{g}_{eq}$ effected by the equivalent gravitational field \mathbf{g}_{eq} associated with a given acceleration, then we may write $\mathbf{F}_{ext} + \mathbf{F}_i = 0$ or $\mathbf{F}_{ext} = -\mathbf{F}_i$ so that for example if the external force is gravitational $\mathbf{F}_{ext} = \mathbf{F}_g = m\mathbf{g}$ then we would have $m\mathbf{g} = -m\mathbf{g}_{eq}$ and this means that the equivalent gravitational field \mathbf{g}_{eq} is usually opposite both the applied gravitational field and the acceleration, because in the present case we also have $\mathbf{F}_{ext} = m\mathbf{a}$, which means that $m\mathbf{g}_{eq} = -m\mathbf{a}$ for the considered positive mass m at least.

But would the equivalent gravitational field experienced by a negative mass particle really be directed opposite the direction of its acceleration as is

the case for a positive mass particle? To that question I think that, contrarily to what is usually assumed implicitly, we would have to answer that this cannot be the case. I will explain that in fact the equivalent gravitational field \mathbf{g}_{eq}^- that would be experienced by a negative mass particle accelerating in a given direction away from any local matter inhomogeneity is the opposite of the equivalent gravitational field \mathbf{g}_{eq}^+ that would be experienced by a similar positive mass particle with the same acceleration under the same conditions, so that we have $\mathbf{g}_{eq}^- = -\mathbf{g}_{eq}^+ = -(-\mathbf{a}) = \mathbf{a}$ for a negative mass particle and given that we still have $\mathbf{F}_{ext} = -\mathbf{F}_i = -m\mathbf{g}_{eq}^-$ it means that $\mathbf{F}_{ext} = -m\mathbf{a}$ when the mass m is negative. If this is correct then it would mean that the acceleration which a negative mass particle would experience as a result of the action of a given force would actually be the same as that which would be experienced by a positive mass particle submitted to the same force (not the same force field but really the same force), even if the mass, including the inertial mass, is effectively negative. The validity of this conclusion depends on only two assumptions. First, the proposed generalized Newton's second law (explicitly involving inertial forces instead of accelerations) must be considered more fundamental than the original formulation involving accelerations, so that the equilibrium it describes is really between forces and not merely between a force and an acceleration. Secondly, it must be assumed that the equivalent gravitational field associated with a given acceleration is reversed when the mass is reversed.

If the preceding conclusions are accurate it would appear that the fact that Newton's second law was always observed to work in its original form, that is, when the equivalent gravitational field is implicitly considered to be opposite the acceleration, is merely a consequence of the fact that it has only ever been verified to apply using positive mass matter. But what is it indeed that might allow one to assume that the equivalent gravitational field would be reversed (would be directed in the same sense as the acceleration) for an accelerating negative mass particle in comparison to what it would be for a similarly accelerating positive mass particle? To understand what is going on we may consider the example provided by Einstein's elevator experiment. Indeed we are allowed by the equivalence principle to assume that the effects observed inside an elevator accelerated in the vacuum away from any local matter inhomogeneity could also be explained by assuming that the elevator is not accelerating in the same vacuum (relative to the local inertial frame of reference which would exist in the absence of any local matter inhomogeneity), but that a large mass, not originally present in

this vacuum, is now located beneath the elevator (in the direction opposite that of the originally assumed acceleration). Thus, it seems that acceleration relative to a local inertial frame of reference always gives rise to an equivalent gravitational field similar to that which we would normally attribute to the presence of a local concentration of matter. We may then define an *equivalent source* to be the matter distribution which would give rise to the equivalent gravitational field experienced by an accelerated body if the presence of this field was not merely the consequence of acceleration.

Now, if we are to assume that the equivalent gravitational *field* associated with the inertial gravitational force is effectively reversed when the mass of the accelerated body is itself reversed, then it seems that even without speculating about what the phenomenon of inertia might actually involve, we should already expect that the sign of mass of the equivalent source associated with the equivalent gravitational field experienced by a negative mass body should itself be reversed. There should be no question in effect that if an accelerating positive mass observer is allowed to assume that the equivalent gravitational field she experiences is actually attributable to the presence of an equivalent source with *positive* mass located in the direction opposite her acceleration, then a similarly accelerating negative mass observer should himself be allowed to attribute the equivalent gravitational field that he would experience to the presence of some equivalent source with *negative* mass also located in the direction opposite his acceleration, otherwise we would have a way to determine in an absolute fashion, the positivity of mass. Indeed, if it was always an equivalent source with positive mass (located in an invariant position relative to the accelerating body) that gave rise to the equivalent gravitational field, we could simply accelerate an observer of any mass and measure the equivalent gravitational field experienced by this observer, which could then be identified as the gravitational field attributable to a positive mass in the assumed position. Therefore, any gravitational field exerting on a given body a force such as that which was observed could be identified as the gravitational field of a positive mass independently from the mere difference or equality between the polarity of the mass producing the field and that of the particle experiencing it. But this is a violation of the above discussed requirement of relational definition of the sign of mass. Thus, the problem with the traditional conception is that it would allow to differentiate between positive and negative mass in an absolute (non-relative) way by referring to the predefined positive mass of the equivalent source whose gravitational field should invariably be observed under otherwise arbitrary motions of acceler-

ation. This is of course the problem I had previously identified based on distinct (although not unrelated) considerations.

As it turns out, an additional difficulty arises when we try to assess the response of negative mass matter to applied forces in the context of the above discussed generalized formulation of Newton's second law. Indeed, under the inappropriate assumption that it is an equivalent source with positive mass that gives rise to the inertial force experienced by a negative mass body it seems that unless we assume that such an object experiences an anomalous response to the equivalent gravitational field itself, then the inertial force would actually have the same direction as the acceleration, which means that the negative mass body would accelerate in the same direction as the *accelerated* frame of reference itself. It is not possible, however, to simply assume that this is what would occur if what we are trying to determine is precisely the response of a negative mass body to some gravitational field. But given that the requirement of relational definition of physical properties also imposes that the properties of attractiveness and repulsiveness of a gravitational force be attributable to the difference or the identity of the signs of mass of the interacting bodies, then we have no choice but to assume that a negative mass would not respond anomalously to the equivalent gravitational field associated with acceleration, because otherwise the properties of gravitational attraction and repulsion would be intrinsic to the masses themselves, as I previously explained.

Thus, even if the equivalent gravitational field experienced by an accelerating negative mass body was the same as that experienced by a similarly accelerating positive mass body it may not even give rise to the kind of motion which is traditionally expected from a negative mass body. If one recognizes the validity of the generalized form of Newton's second law, it is not possible to conclude that a negative mass body with negative inertial mass would experience an acceleration opposite the applied force, even under the assumption that the equivalent gravitational field is not dependent on the sign of mass of the accelerating body. Indeed, when we appropriately assume that the sign of the inertial mass of an object merely determines the direction of the force exerted by an equivalent gravitational field on this object and not its response to the force, then it follows that a negative mass body which would experience the same equivalent gravitational field as does a positive mass body would still accelerate in the direction of the external force, only now its acceleration would be experienced with respect to the accelerated reference system itself. As a consequence, there would no longer be

an equilibrium between the applied force and the inertial force which is experienced by a negative mass body due to its acceleration, which is certainly not a desirable outcome.

What is important to understand in effect is that in the context of a generalized formulation of Newton's second law it must actually be imposed that there is always an equilibrium between the applied forces and the inertial force and under such conditions the acceleration to which a body with a given mass sign is submitted is determined solely by the requirement that the inertial force it experiences effectively balances the applied forces. Thus, once the direction of an applied force is known the acceleration of the body submitted to this force is determined only by the condition that it effectively gives rise to an inertial force which balances the applied force. But if the equivalent gravitational field which gives rise to the inertial force is dependent on both the direction of acceleration and the sign of mass of the accelerated body then the fact that the sign of mass would be reversed would not affect the direction of the acceleration, because the inertial gravitational field would also be reversed, which allows the inertial force associated with this acceleration to remain invariant.

Under such conditions it would not be appropriate to assume that it is the sign of mass itself which determines the direction of the acceleration, because in fact the acceleration of a body submitted to a given force is determined merely by the requirement that the inertial force experienced by such an object balances the applied force in the accelerated reference system relative to which this inertial force is present. There is no a priori justification for considering that a negative mass body with negative inertial mass should experience an acceleration opposite the applied force. This would be an incorrect interpretation of the classical equation between force and acceleration, which must be assumed to be valid only when the mass is positive. What the preceding argument shows in effect is that it would be a mistake to assume that the traditional formulation of Newton's second law applies even when the mass is negative. This equation does not apply when the mass is negative simply because the formula was not derived under the assumption that mass can be negative and was never intended to apply under such circumstances. But in the context of a generalized formulation of Newton's law and when the mass of the equivalent source responsible for the equivalent gravitational field is appropriately reversed for an accelerating negative mass body, it follows that the equivalent gravitational field experienced by such an object must itself be opposite that experienced by a positive mass body,

which means that the inertial force remains unchanged, as does the body's acceleration.

If we are willing to recognize that it would be a serious inconsistency to allow for the same equivalent source (with the same mass sign) to give rise to both the equivalent gravitational field experienced by positive mass particles and that experienced by negative mass particles then we must also recognize that similarly accelerating positive and negative mass bodies would experience opposite equivalent gravitational fields, because those gravitational fields would arise from equivalent sources with opposite mass signs. But given that a negative mass must experience a force opposite that experienced by a positive mass of similar magnitude in response to any gravitational field, it follows that the inertial *force* would effectively have the same direction for both positive and negative mass bodies accelerating in the same direction as a consequence of being submitted to the same external force (which is more constraining than requiring the same applied force *field*), even if we consider inertial mass to be reversed along with gravitational mass, as I previously argued to be necessary. In the present context we would actually be allowed to assume that the requirement to consider that the equivalent gravitational field is reversed for a negative mass body (in comparison with the equivalent gravitational field experienced by a positive mass body with the same acceleration) is justified by the fact that it allows the dynamic equilibrium of forces on such an object to be maintained in the accelerated frame of reference relative to which this equivalent gravitational field is experienced, because if in order to meet this constraint we must consider the same inertial gravitational force to arise from the same acceleration then it means that a negative mass body would necessarily have to experience a reversed equivalent gravitational field given that its mass is indeed reversed. No circular reasoning is involved here, because those results actually follow from the mere requirement of relational definition of the sign of mass applied to the equivalent source that gives rise to the equivalent gravitational field experienced by an accelerating negative mass body.

For this argument to be valid what must be recognized is that the negativity of the inertial mass of a negative 'gravitational' mass is an independent consistency requirement, which actually amounts to assume that mass is mass and that it cannot be both negative and positive at the same time and once this is acknowledged we are allowed to also and independently conclude that just as there is not a unique sign of mass, there is not a unique equivalent gravitational field for bodies with opposite mass signs. In such a context we

have no choice but to recognize that the response of a negative mass body to any applied force would be that which we ordinarily (but inappropriately) attribute to a negative gravitational mass whose inertial mass would remain positive. Yet in the present case it would seem that the validity of the equivalence principle could be preserved to some extent, even while there are two different kinds of response to a given gravitational field, because all mass (gravitational and inertial) is now reversed for a negative mass body and only bodies with opposite mass signs must be assumed to respond differently to a given gravitational field and not bodies with the same inertial mass, which would then have constituted a real violation of the requirement of equivalence of acceleration and gravitation, as I explained before.

It is now possible to understand why it is that the inappropriate choice of a positive inertial mass in association with a negative gravitational mass would seem to agree, from a purely phenomenological viewpoint, with the independently motivated requirement of a relational definition of mass sign (given that it would allow gravitational attraction and repulsion to themselves be features dependent merely on the *difference* between the signs of gravitational mass of any two bodies). It is simply because in such a case instead of appropriately reversing the equivalent gravitational field for a negative mass accelerating in a given direction we would reverse the sign of inertial mass (which must be negative for a negative mass particle) a second time, from negative to positive again (while keeping the gravitational mass negative), which superficially would be equivalent to simply reversing the direction of the equivalent gravitational field while keeping the mass negative as required. But I must emphasize again that if that was the only possible approach to obtain consistent behavior from negative mass bodies we would in fact have to conclude that negative mass is not an appropriate concept in physical theory, because we would have to assume that a single unique physical property (what we may call the gravitational ‘charge’) is required to have at once and *from the exact same viewpoint* (for an observer of unchanged mass sign) two opposite values and this is clearly not acceptable.

It must nevertheless be mentioned that, as later developments will illustrate, it appears that in fact the reversal of the equivalent gravitational field is the trade-off we have to accept for keeping the value of the gravitational field attributable to a local matter inhomogeneity generally invariant while assuming that it is effectively the mass experiencing it that can be reversed. But if instead we considered that the motion of a body is to always be determined from the viewpoint of an observer made of matter with the same

energy sign, then we should naturally assume that the sign of mass of the body (both inertial and gravitational) is an invariant property that may be assumed positive definite, while it is the gravitational field attributable to a given matter inhomogeneity that is variable as a function of the difference between its energy sign and that of the observer. From this viewpoint the equivalent gravitational field due to acceleration far from any local matter inhomogeneities would no longer be dependent on the sign of mass of the accelerating body (because the mass itself would not change), while the gravitational field due to the presence of a local matter inhomogeneity would depend on the perceived sign of energy of its sources which would become an observer dependent property, while the mass or energy of the body experiencing the fields would actually be considered positive definite. In this context there would then still be a practical (although not fundamental) distinction between an equivalent gravitational field due to acceleration far from any local mass concentration and the gravitational field due to the presence of a local matter inhomogeneity (in the absence of forces other than gravity). I will explain below what is the profound origin of this distinction and why it does not constitute an insurmountable difficulty for a consistent general relativistic theory of gravitation based on the equivalence principle.

What must be retained here is that we can still consider the direction of the gravitational field attributable to the presence of a local matter inhomogeneity to remain invariant while it is the mass experiencing it and therefore also the equivalent gravitational field experienced by this mass which may be reversed, but only at the price of changing the equations of motion which will be shown to otherwise describe the trajectories of particles submitted only to the gravitational interaction in a way that is equivalent to considering that the mass experiencing the gravitational field (due to the local matter inhomogeneity) is invariant while it is the field itself which is reversed (in comparison to what it would be if we had considered its effect on a negative mass body). Now, if we do consider the mass (both gravitational and inertial) of the particle experiencing a gravitational field to always be positive definite so that that it is the direction of the gravitational field itself which varies as a function of the *relative* difference between the observer dependent sign of mass of the source (which can still be either positive or negative) and that of the particle experiencing the field (which would always be assumed to be the positive one) then we obtain a framework that is more easily generalizable to a relativistic theory. But it must be clear that the two approaches discussed here are equivalent in the Newtonian context and still require all

mass (gravitational and inertial) to be either positive or negative and when the direction of the gravitational field due to a local matter inhomogeneity is not considered to be an observer dependent property we must effectively consider the equivalent gravitational field to itself be dependent on the sign of the accelerated mass (which is no longer positive definite), otherwise the equivalence between the two viewpoints breaks down.

From the viewpoint where the mass experiencing a gravitational field is considered positive definite, a Newtonian gravitational field experienced by a particle we would normally consider to have positive mass, if it is not the result of an accelerated motion far from any matter inhomogeneity (in which case we would be dealing with an equivalent gravitational field), would be experienced by a particle we would normally consider to have negative mass as an oppositely directed Newtonian gravitational field, while the mass of the particle experiencing this relatively defined gravitational field would not even show up in the equations used to determine its motion. But if the gravitational mass experiencing this reversed gravitational field is kept positive then it must be assumed that the inertial mass is also kept positive and under such conditions the equivalent gravitational field would appear not to be reversed. It is because we do not appropriately keep the mass sign invariant when we try to determine the motion of what we currently describe as a negative mass particle in a given accelerated frame of reference that we need to reverse the experienced equivalent gravitational field. But when the external force applied on what we would currently describe as a negative mass particle is gravitation itself it is possible to assume that this force is reversed (from that which would be experienced by what we currently describe as a positive mass particle), not because the mass of the particle is reversed, but because the local gravitational field itself is reversed. In such a case the inertial *force* would not be reversed, because the mass (both gravitational and inertial) that is experiencing the field is not reversed and it must also be assumed that the equivalent gravitational field is left unchanged (from that which is experienced by what we already consider to be a positive mass particle). Therefore, acceleration still doesn't take place in the direction opposite the applied force and this is all a consequence of the fact that even though the local gravitational field appears to be reversed from such a perspective, the equivalent gravitational field in contrast is left invariant along with the sign of mass of the particle.

It should be clear then that in the context of an approach according to which the particles experiencing a gravitational field are always assumed to

have a positive mass, the crucial assumption is that while the gravitational fields attributable to local matter concentrations are dependent on the nature of the body experiencing their effects, the equivalent gravitational field associated with acceleration away from local masses would for its part remain invariant regardless of how the body experiencing it perceives the gravitational fields attributable to local matter inhomogeneities. This hypothesis can be considered to be equivalent to that which in the above described approach consists in assuming that the equivalent gravitational field must actually be reversed for a negative mass, because this is effectively what allows the inertial properties of an object to be independent from its mass sign. I believe that this observation clearly shows that I am justified in analyzing the problem of negative mass from a conventional perspective according to which the mass experiencing a gravitational field is explicitly assumed to be reversed, because in such a context the underlying assumptions are made more apparent and it is also easier to explain what I am referring to when discussing the case of abnormally gravitating matter. In a Newtonian context I will therefore continue to use the first viewpoint according to which it is possible for the mass experiencing a gravitational field to be negative.

Now, we may want to dig a little deeper and ask why it is exactly that we are allowed to assume that the direction of the equivalent gravitational field is dependent on the sign of mass of the object experiencing it? I have tried very hard to develop a better understanding of the whole phenomenon of inertia and what I have learned has actually helped me to derive the above discussed results. Indeed this investigation has enabled me to realize that the assumption that the equivalent gravitational field is reversed when the mass which is subject to acceleration is itself reversed is not just a requirement of the necessary relational definition of the sign of mass, but must be imposed in order to allow a relational description of the phenomenon of inertia itself, in the sense that inertia should be conceived as arising from purely relative motions between matter particles, as suggested by Ernst Mach a long time ago. In this context I have become convinced that the inertial forces acting on a particle can be understood to arise as a consequence of an imbalance, caused by acceleration relative to the global inertial frame of reference (associated with the distribution of matter on the largest scale), in the sum of forces attributable to the interaction of the accelerating particle with each and every other particle in the universe.

What happens in effect is that there must be a similar imbalance of the gravitational forces exerted on similarly accelerating positive and negative

mass bodies arising from their interaction with the rest of the matter in the universe, because the imbalance responsible for the existence of the inertial gravitational force is similar to a skewed mass distribution and if the actual large scale matter distribution responsible for those effects is roughly the same from the viewpoint of both positive and negative masses in the absence of local matter inhomogeneities then the imbalance should develop in a similar way for both positive and negative masses from the viewpoint of their own mass sign. Thus what must be retained of this investigation is that the equivalent gravitational field which applies on a negative mass body should in fact be the opposite of that which would be experienced by a positive mass body with the same acceleration that is located within the same matter distribution, even if simply as a consequence of the fact that for a reversed mass the same motion relative to the same matter distribution should give rise to a similar imbalance in the sum of *forces* attributable to interaction with all the matter in the universe. Indeed, given that the mass itself is reversed, the invariance of this imbalance would mean that the equivalent gravitational field responsible for the inertial force must also be reversed in the accelerated frame of reference, so that the force existing relative to it can itself be left invariant. But if the equivalent gravitational field associated with the acceleration of a negative mass body is the opposite of that associated with the same acceleration of a positive mass body it follows that the reaction to any applied force is effectively the same for opposite mass particles, despite the fact that there is no distinction between inertial and gravitational mass signs (even for negative mass particles). This may be considered to actually explain why it is appropriate to effectively assume that it is the inertial force itself, instead of merely the product of mass and acceleration, that would be opposite the direction of the applied external force for a negative mass body, as the generalization of Newton's second law that I proposed allows to express.

But it must be clear that if there is a requirement for inertial mass to be reversed along with gravitational mass it does not follow from imposing the validity of the equivalence principle as a condition that all matter should have the same motion in the absence of any interaction other than gravitation, as is usually considered. Indeed, as the previous analysis allows to understand, even a negative mass body for which both the gravitational and the inertial masses are negative should not be expected to follow the same trajectory as a positive mass body in the presence of a local positive or negative mass concentration (despite what is usually assumed). What I have tried to explain

is precisely that even when inertial mass is assumed to be reversed along with gravitational mass it is not possible to preserve the validity of the equivalence principle integrally. Thus, a local inertial frame of reference cannot be defined independently from the sign of mass of the body experiencing it given that the direction of the gravitational force resulting from a particular matter distribution depends on the sign of mass of this body. The less restrictive requirement that all matter with the *same mass sign* in the same location follows the same motion (acceleration) is in fact appropriate and restrictive enough for the equivalence between gravitation and acceleration to apply, precisely when it is considered that both gravitational and inertial masses must always be reversed together, because it is only in such a case that at least all positive (inertial) mass matter follows the same motion (it is usually assumed that a negative gravitational mass with positive inertial mass would not), which is all that is really required by the principle of equivalence (masses with the same sign should have the same acceleration) as I have explained. Thus, it is in this particular sense only that we may assume that the equivalence principle requires inertial mass to be reversed along with gravitational mass.

2.5 The equivalence principle with negative mass

It is not usually recognized that the general theory of relativity is actually based on two postulates, because only the first postulate, which concerns the equivalence between acceleration and a Newtonian gravitational field, is well known and is explicitly taken into account. But actually a second postulate is required to obtain the current formulation of the theory and is implicitly assumed to be valid without justification. It is the hypothesis of absolute significance of the sign of energy. This second assumption appears to be necessary to preserve the validity of the first postulate under conditions where negative energies would effectively have to be taken into account. But even though the postulate of the absolute definiteness of the sign of energy may be considered problematic in the context of the preceding analysis, it remains to be shown whether it is possible to provide a consistent classical theory of the gravitational field where only this second postulate would be rejected. Thus I will try to show in this section and later when discussing the mathematical

aspects of a generalized theory of gravitation that it is perfectly possible and indeed actually necessary to maintain the validity of the equivalence principle in its most essential form while nevertheless rejecting the assumption of an absolute significance of the sign of mass or energy.

Now, I would like to emphasize that the true motivation behind the equivalence principle is to be found in a requirement which we may call the relativity principle and which is actually one particular expression of the requirement of relational definition of all physical quantities. This relativity principle imposes that the state of motion of an object, and in particular its rate of acceleration, is to be determined merely in relation to the state of motion of other physical systems, so that there is no absolute state of acceleration relative to an arbitrarily chosen, unique, metaphysical reference system. The principle that there is an equivalence between a Newtonian gravitational field and an acceleration enables this requirement to be fulfilled, because it allows what might have otherwise appeared to be an acceleration relative to absolute space to merely be a state of rest in the vicinity of a local mass concentration not accelerating relative to the same ‘absolute’ space, as Einstein understood, but as we tend to ignore nowadays in favor of the mere mathematical requirement of general covariance of the field equations. I think that it must be recognized that in fact the only essential implication of the equivalence principle is that indeed there is no longer any motive for arguing that because acceleration is felt (unlike velocity) it must be absolute. Thus it may appear problematic that even if we find generally covariant equations for the gravitational field in the presence of negative energy matter, the fact that according to the previous analysis such matter would not share the same accelerated motion as positive energy matter in the presence of a local matter inhomogeneity (while it should in the absence of such a perturbation for reasons I explained before), would appear to allow the effects of acceleration relative to matter at large to be distinguished from those attributable to the gravitational field of a local mass.

There is indeed a tension between the principle of relativity and the previously discussed requirements concerning negative mass matter which we may illustrate by once again using Einstein’s elevator experiment. Under circumstances where what I have identified as appropriately behaving negative energy matter would be present it may seem in effect that we could differentiate an acceleration of the elevator occurring far from any local mass from an acceleration of the elevator occurring while it is at rest near such a large mass. This is because near a planet or another large matter inho-

inhomogeneity positive and negative mass bodies would accelerate in opposite directions, one toward the local mass and the other away from it (one upward the other downward), while in the elevator which is simply accelerating far from any large mass, positive and negative energy bodies would share the same acceleration, apparently betraying the fact that the acceleration is ‘real’. We may therefore assume that an observer in the elevator would be able to tell when it is that she is simply standing still in the gravitational field of a planet and when it is that she is actually accelerating far from any big mass. The ‘true’ acceleration would have been revealed to the occupants of the elevator as that for which both the positive and the negative mass bodies have the same acceleration. Consequently, we would seem to be justified to conclude that the notion that acceleration is totally equivalent to a gravitational field (which is the essence of the principle of equivalence) is no longer valid when we introduce negative mass matter with properties otherwise required to make it a consistent concept (according to the preceding analysis).

Indeed, I made it clear before that it is not possible to abandon the principle of inertia or Newton’s third law (action and reaction) in order to accommodate the existence of negative mass matter, because if those rules were not strictly obeyed under all conditions then not much else would make sense. We cannot even tell what a world devoid from this constraint would look like and we would have no reason to assume in particular that the equivalence principle itself would still be obeyed, as is usually assumed, because after all this principle is a reflection of the phenomenon of inertia. Trying to save the principle of equivalence by simply allowing negative mass matter to react abnormally to applied forces (as if that was required when inertial mass is negative), so that it accelerates like positive mass matter in the presence of local matter inhomogeneities, would not make sense, because this would mean that the principle of inertia no longer applies in general and again in such a case there is no guarantee that even the alternative situation we expect to observe under those conditions would effectively occur. I believe that there are reasons why no violations of the principle of inertia have ever been observed despite the fact that the techniques required to reveal such transgressions have long been available. It would not be clever to think that it is by rejecting this principle that we can maintain the requirement of the equivalence between a gravitational field and acceleration. Clearly, there must be something wrong with certain assumptions we take for granted concerning the equivalence principle itself. The fact that this is the principle upon which

relativity theory and our modern concept of gravitation is founded should not prevent us from reexamining some of the implicit assumptions surrounding it. Failing to do so would mean that we have to give up on the idea that negative energy matter could exist, because only so could we then avoid being faced with the annoying and unpredictable consequences of an alternative choice concerning the properties of this matter.

It is important to note at this point that it would be inappropriate to suggest that the requirement that the principle of equivalence also applies in the presence of negative mass matter could perhaps be accommodated if opposite mass bodies were found to always share *opposite* accelerations instead of always sharing the same acceleration as is traditionally assumed. It is certainly true that under such circumstances it would still be impossible to distinguish a true acceleration given that opposite mass bodies would always accelerate in opposite directions, whether those accelerations are the result of the presence of a local concentration of matter or the result of the presence of an equivalent gravitational field far from any large mass. But this situation could only occur if in the context of an appropriate conception of the phenomenon of inertia based on the previously discussed generalized formulation of Newton's second law, the equivalent gravitational field associated with acceleration was not reversed despite the reversal of the mass of the accelerated body experiencing it. From that viewpoint we should effectively expect that one of two opposite mass bodies would fall down while the other would fall up in the accelerating Einstein elevator far from any local mass, even when no force is applied on any of the two masses independently. However, this kind of behavior would constitute an even more severe violation of the principle of inertia than that which would occur in the case of the chasing pair of opposite mass bodies described before, given that in this case there wouldn't even exist any identifiable cause for the upward acceleration of one of the two bodies, because the elevator does not even interact with any of the masses and merely constitutes a reference system. In fact, this situation is so devoid of plausibility that it clearly means that it is not possible to try to salvage the equivalence principle by assuming that the equivalent gravitational field is not reversed for an accelerating negative mass body. The fact that the kind of invariance of the equivalent gravitational field that is involved here would also violate the requirement of relational definition of the sign of mass, as I explained in the previous section, only contributes to confirm the validity of this conclusion. We must therefore accept that while the local inertial frames of reference can differ for positive and negative mass

bodies near some local matter inhomogeneities, they must nevertheless be identical for opposite mass bodies far from local mass concentrations.

I will soon explain why it is exactly that we are allowed to consider that the principle of relativity of motion (concerning acceleration in particular) is not threatened by the conclusion that the free fall state of motion of a negative mass body can be different from that of a positive mass body in the presence of local matter inhomogeneities. But it is important to first point out that in the case of the elevator near a local mass we are effectively considering an inhomogeneous matter distribution for which positive and negative energy matter concentrations are *not superposed* in space (in the classical sense) and therefore do not compensate one another. If such compensations between the effects of *local* matter inhomogeneities were to occur, when for example we would have two superposed gas clouds of opposite energy with the same overall motion, possibly rotating, but in the same direction, then the acceleration of positive and negative energy bodies located near or within those matter distributions would have to be the same despite the presence of local inhomogeneities in the configuration of positive and negative energy matter. This actually means that there couldn't be any effect from the motion relative to such a matter distribution, because whatever gravitational effect positive energy matter would have would be compensated by the opposite effect of the similarly distributed negative energy matter present around the body. This is true also of rotation which according to Einstein's theory induces a frame dragging effect which we may assume to be dependent on the sign of mass like any other gravitational phenomenon.

Now, you may recall this earlier discussion (from the preceding section) in which I suggested that it should be possible to attribute the inertial gravitational forces experienced by positive and negative mass bodies in the accelerating elevator away from local masses to some imbalance in the sum of gravitational forces attributable to interaction with all the matter in the universe arising as a consequence of acceleration relative to the reference system associated with the average state of motion of this large scale matter distribution. However, given what I just mentioned regarding the compensating effects of superposed matter distributions with opposite masses and identical motions, it seems that we would have to assume that no imbalance could arise from the gravitational interaction with positive and negative energy matter if they are similarly distributed in space on the largest scale. Thus we must conclude that if the positive and negative energy matter distributions are effectively mostly identical and are at rest with respect to one another on such

a scale (as appears necessary if the cosmological principle applies equally to both matter distributions) then there should be no effect on both positive and negative mass bodies from the presence of matter on the largest scale. But this means that there could not be any imbalance in the equilibrium of gravitational forces attributable to the large scale matter distribution that would give rise to inertial forces or the equivalent gravitational fields, because one imbalance attributable to motion relative to positive energy matter would be compensated by a similar but opposite one arising from the same motion relative to negative energy matter (all masses would experience two opposite equivalent gravitational fields all at once). It thus appears that there is something wrong with one or more of the implicit assumptions entering this deduction, because inertia does exist and indeed if there was no inertia the world would not be anything even remotely similar to what we observe. Of course the idea that there simply is no negative energy matter in the universe (so that the imbalance due to acceleration relative to the positive energy matter distribution is not compensated by an imbalance due to acceleration relative to the superposed negative energy matter distribution) may be tempting, because after all we do not observe any such matter. But keep in mind that it will later be explained that this hypothesis is not required and that in any case it would again amount to simply reject the possibility that such matter may exist without providing any justification for this very convenient hypothesis.

We may summarize the situation by noting that what we know for sure is that if the identical accelerations of the opposite energy bodies relative to the elevator far from any local mass are due to a similar imbalance in the gravitational forces attributable to the interaction of those bodies with matter on the largest scale then this imbalance must be attributed to motion that takes place relative to what are essentially *identical* matter distributions with the same motion and the same rotation and which *should* therefore have opposite effects of equal magnitude on positive and negative energy bodies with the same motion *relative* to this homogeneous matter distribution. If this is recognized, then we have to admit that in the context where negative energy matter effectively exists it would be difficult to see how a local inertial frame of reference could be determined by the large scale matter distribution through the gravitational interaction. In such a case it would then seem that we have to conclude that there may need to exist something like absolute acceleration relative to an arbitrarily chosen unique reference system lacking any physical underpinning. What I have understood though

(for reasons that will be discussed later) is that the hypothesis that both the large scale positive and negative energy matter distributions have an effect on positive or negative energy bodies considered independently constitutes the incorrect assumption which appears to invalidate the hypothesis that all motion (including accelerated motion) is relative, even in the presence of negative energy matter.

If we drop the assumption that a *negative* energy matter distribution that is uniform on the cosmological scale can exert a force on *positive* energy matter (and vice versa for the effects of positive energy matter on negative energy matter) then it seems that we can explain the imbalance responsible for the force of inertia as being a consequence of the acceleration with respect to the one particular, but relatively defined, reference system which is that relative to which most of the matter in the universe is at rest, because in such a case there would be no canceling of the effects attributable to the positive energy matter distribution by those of the negative energy matter distribution (and vice versa) on the largest scale. Therefore, what I suggest we have to recognize, if only by necessity, is that there is no compensation, for a positive mass body accelerating relative to the average matter distribution on the cosmological scale, between the equivalent gravitational field attributable to positive energy matter and that which we could have attributed to negative energy matter. Similarly, there should be no equivalent gravitational field attributable to acceleration relative to the average distribution of positive energy matter to compensate the equivalent gravitational field attributable to acceleration relative to negative energy matter for a negative mass body. If I am right this is due merely to the fact that on the cosmological scale particles of one energy sign interact only with the matter distribution that has the same energy sign. I am particularly confident in the validity of this proposal given that I had actually understood the requirement of absence of interaction between opposite energy matter distributions on the cosmological scale before I even realized that it was required to solve the problem of the relativity of motion in the context where negative energy matter is effectively allowed to exist. I will explain what independently justifies this conclusion in a following section of the current chapter.

What happens, therefore, is that only the very large scale distribution of *positive* energy matter determines the local inertial frame of reference that is experienced by *positive* energy bodies in the absence of local matter inhomogeneities, while only the overall distribution of *negative* energy matter determines the local inertial frame of reference experienced by *negative* en-

ergy bodies in the absence of local matter inhomogeneities (this language would also be appropriate from a general relativistic viewpoint). Thus, what differentiates the situation of the elevator near a large mass of positive *or* negative sign and the situation we have in the elevator accelerating far from any such local mass is that in the first case the force responsible for the observed acceleration is the result of an imbalance that is caused by unequally distributed inhomogeneities in the positive and negative energy matter distributions and this imbalance is dependent on the sign of energy of the body experiencing it (as there are two possibilities for both the sign of mass of the source and that of the accelerated body), while in the latter case the observed force responsible for the acceleration is the result of an imbalance that is always caused by the motion of a body of given mass sign relative to a uniform matter distribution with the same mass sign (necessarily and invariably) so that it is not dependent on the sign of energy or mass of the body experiencing it (positive and negative energy bodies react in the same way to acceleration relative to matter on the largest scales) as long as the distributions of positive and negative energy matter are similar and are not accelerating or rotating relative to one another on the largest scale.

All accelerations are therefore relative accelerations between well-defined physical points of reference within the universe and no absolute state of rest (more exactly of absence of acceleration) can be identified. This is true even if there does exist a unique particular reference system (actually two unique but corresponding reference systems) which is singled out as that relative to which the motion (state of acceleration) of positive and negative mass bodies is the same in the absence of local disturbances, as a result of the correspondence of the average state of motion of the positive and negative energy matter distributions on the largest scales. But this conclusion applies merely in the context where *globally* any particle is gravitationally influenced only by its interaction with matter of the same energy sign whose state of motion relative to the particle therefore alone determines the local inertial frame of reference in which the particle evolves. Thus, despite the exact correspondence of the positive and negative energy matter distributions on the largest scale (which if those sources were locally concentrated would imply an absence of resulting effect on both positive and negative mass bodies) there nevertheless exists a resulting effect from the presence of this matter on a local mass of any sign that allows to determine a unique frame of reference and this is what explains that there appears to be a difference between acceleration far from any local mass and a gravitational field due to local matter inhomogeneities,

while in fact the difference observed is merely the consequence of the fact that a body with a given mass sign interacts only with the large scale matter distribution with the same sign of mass so that no compensation can exist in this case.

In light of those developments it appears that what the previously discussed insight concerning the nature of the equilibrium involved in determining local inertial frames of references should be understood to mean is that free fall motion, instead of involving a total absence of forces, as is usually assumed in a general relativistic context, must be considered to be a manifestation of the acceleration-dependent equilibrium in the sum of gravitational forces attributable to interaction with both local masses and the large scale matter distribution. This interpretation appears to be required in the context where negative energy matter must be recognized to exist, given that in such a case there cannot even be a unique inertial, or free fall frame of reference dictated by the geometry of spacetime, so that we are forced to consider the reality of the general relativistic gravitational field as being associated with such a physical interaction. Indeed, it is only when we are dealing with a universal force, defined precisely as a force that affects all bodies in the same way, that we can *choose* (as a mere convention) to include this force in our definition of the metric properties of space and time, given that geometry must by definition be shared by all objects in a given space. But this is just a convenient choice which would clearly appear for what it is if the force in question was not affecting all bodies similarly (therefore betraying its material nature).

Einstein himself insisted that if we are to recognize the validity of a principle of general relativity of motion, then the speed of light can no longer be assumed to be constant, given that in the elevator experiment light rays may follow curved paths. But from this viewpoint the curvature of spacetime should naturally be expected to arise as the manifestation of an equilibrium of gravitational forces dependent on acceleration and due to interaction of the bodies experiencing it with all the matter in the universe (except the large scale matter distribution with opposite mass sign), otherwise it would be impossible to determine what affects the trajectory of light in an accelerated reference system far from any local matter inhomogeneity. Indeed, even in a flat space far from any local matter concentration the motion of light in a straight line, which is usually considered to be a consequence of geometry itself, would from my viewpoint be a consequence of the equilibrium of forces arising from the gravitational interaction with the rest of matter in the uni-

verse. This does not mean, however, that the geometrical interpretation of gravitation is incorrect, but merely that the geometrical properties of space must definitely be conceived as arising from those interactions and more precisely from some sort of equilibrium in the sum of gravitational forces that can be altered by the presence of local matter inhomogeneities. Such a viewpoint has the added benefit of being more easily generalizable to a theory where the gravitational interaction must be described as an interaction mediated by quantum particles.

In any case I think that it is clear that statements to the effect that relativity theory has made the concept of gravitational interaction obsolete and replaced it with that of spacetime curvature (so that gravitation is merely a manifestation of the geometry of spacetime) can no longer be assumed meaningful if curvature is itself a relatively defined property which arises as a consequence of an equilibrium of local and inertial gravitational forces which depend on the sign of energy of the objects involved. I think that the situation we have here is similar to that in which electromagnetic theory was before the quantization of light and the photon concept was proposed, because spacetime is now viewed as a continuous medium, not dependent on underlying physical causes, that directly takes part in determining the motion of objects, just like electromagnetism was originally considered a fundamental wavelike phenomenon directly influencing the motion of charged bodies. When it was shown that light is a corpuscular phenomenon the whole notion of electromagnetic wave was not abandoned of course, because there was something real about the wavelike character of electromagnetic phenomena and this is the element which came to be integrated into quantum theory. Similarly, I think that the concept of spacetime curvature cannot and need not be abandoned when gravitation is described as an interaction (which would ultimately be described as a quantum phenomenon) involving some sort of equilibrium which is dependent on the sign of mass of the object submitted to it and from which local inertial frames of reference emerge, only, spacetime curvature can no longer be considered as actually being gravitation itself.

As Hans Reichenbach once emphasized [15], if we choose to integrate the gravitational force into our definition of spacetime we may no longer need to explicitly take the force into consideration to explain the motion of bodies, but we must still invoke a force as the cause of the geometry itself. Thus, it is not gravitation which was replaced by curved geometry, but all of geometry that became a manifestation of the gravitational interaction

and I think that this is particularly relevant in the context of a theory of gravitation that allows to take into account the possibility of the existence of negative energy matter. Actually, the commonly made remark to the effect that relativity allowed to eliminate gravitation as a real force appears to be motivated by the fact that the gravitational force arising from local mass concentrations was given the status of inertial force (similar in kind to the Coriolis force) by relativity and given that inertial forces were never seen as real forces then it is believed that gravitation can now be considered a fictitious force under all circumstances. But I believe that it is rather the contrary that is true and that it is the inertial forces which are allowed by relativity to be considered as real gravitational forces. In such a context the fact that inertial forces are involved in giving rise to the dynamic equilibrium which determines the mass sign-dependent local inertial frames of reference is a further indication that the geometry of spacetime is the product of an equilibrium of real gravitational forces arising from the interaction of any given mass with the rest of matter in the universe.

Having properly identified the origin of the identical response of positive and negative mass bodies to acceleration, I do not want to immediately enter into a discussion as to what are the true elements of justification behind the assumption that particles of one mass sign are not affected from a gravitational viewpoint by the presence of matter of opposite sign on a cosmological scale. But it may nevertheless already be noted that the fact that one particular reference system appears to be singled out as having unique status among all possible states of acceleration is not a unique feature of the approach described here. Actually, in a general relativistic context, even in the absence of negative energy matter, this feature of our description of the motion of objects should appear all the more natural given that all inertial frames of reference are the result of the gravitational interaction and are therefore determined by the surrounding matter distribution. There exists in effect one very particular reference system in our universe which we may call the global inertial frame of reference and which is that which is determined by the average motion of all masses together and relative to which most masses in the universe do not accelerate (or rotate). That there may be such a unique point of reference does not mean that it is not relationally defined. Relativity theory allows to explain the existence of this particular reference system as being a result of the combined gravitational interactions of a local body in any state of motion with all the other masses in the universe (with

the same mass sign) and therefore in relation to the average motion of those masses. Indeed, even far from any big mass there remains the gravitational effect of the universe as a whole, which can never be ignored. Thus the situation we usually refer to as corresponding to an absence of gravitational field and which we expect to be experienced far from any local mass is not different in fact from that occurring in the presence of a local mass, only it is characterized by the fact that the gravitational field is then attributable to uniform distributions of either positive or negative mass matter, which incidentally implies coinciding inertial frames of reference for positive and negative mass bodies.

The fact that inertial frames of reference are always determined by the average state of motion of matter in the universe becomes particularly obvious when we consider the reference system associated with a felt motion of rotation which, as experiments have revealed, must be one that is in rotation relative to the most distant galaxies and therefore relative to the largest visible ensemble of matter in the universe. The reference system relative to which a positive mass observer feels no rotation must then be determined simply by the gravitational field attributable to all matter particles with the same mass sign present in the universe in a way that is dependent on the average state of motion of those particles and as such is definitely unique even though its description involves only relationally defined properties. We may still consider the average matter distribution on the largest scale to be rotating, but then its gravitational field would give rise to a rotating inertial frame of reference which, through relativistic frame dragging, would put the whole matter content of the universe in rotation with it³. Since Einstein there is no longer any mystery with the existence of such a preferred frame of reference and what I am trying to explain is that there is also no problem with the fact that there is a unique reference system relative to which at once positive and negative mass bodies have no acceleration when free from external non-gravitational forces. We are not faced here with a metaphysical reference system associated with absolute acceleration, but merely with an

³It has been mentioned that a (positive mass) observer uniformly rotating with respect to the distant stars and which would choose to consider himself motionless would observe a gravitational field which from a Newtonian viewpoint could not exist, therefore weakening the equivalence principle. But it is interesting to observe that this observation would no longer be significant in the context where a repulsive gravitational field proportional to the distance from a given center could be produced in the presence of an appropriately configured inhomogeneous distribution of negative energy matter.

ordinary reference system relative to which the effects of the gravitational interaction of local masses with matter on the largest scale imposes an absence of acceleration for both positive and negative mass bodies.

Again it must be stressed that even when it may seem that we are dealing with empty space, what the objects actually experience are the effects of the whole surrounding matter distribution conveyed by the gravitational field as an intermediary material entity, which in a general relativistic context effectively determines the possibly distinct local inertial frames of reference affecting positive and negative energy bodies. This aspect of the general relativistic (or physical) space is what allows to conceive of rotation as being purely relative, even when the distance of some objects to the rotation axis of a rotating observer becomes large enough that the objects would actually have to move at faster than light velocities in the reference system tied to the observer. Indeed, it is the rotation of the whole gravitational field, as a material entity (which would also occur in a universe totally devoid of ‘real’ matter), that explains that this motion of the remote objects is possible as a true motion, because locally the objects are not moving (accelerating) relative to the gravitational field (or the local inertial frames of reference), which is then itself rotating, and this is what makes their large velocities and accelerations possible, as is already well understood.

But if acceleration occurs merely relative to the inertial reference systems determined by the gravitational field it must not be forgotten that the state of motion of matter also contributes to determine the gravitational field and therefore it should naturally be expected that there is no acceleration of matter as a whole relative to the global inertial frame of reference determined by the gravitational field produced by this large scale matter distribution. It may also be remarked that the situation we are dealing with here concerning the relativity of acceleration in the presence of negative energy matter is similar to that regarding the relativity of velocity, because there also exists a preferred frame of reference relative to which the temperature of the cosmic microwave background is mostly uniform and which may appear to define a state of absolute rest, but this unique frame of reference is merely that which is not moving relative to the average state of motion (not acceleration) of matter on the largest scale. If there is no conflict with the principle of relativity in such a case, then there need not be a problem in the case of the global inertial frame of reference singled out as being that relative to which there is no difference between the states of acceleration of freely falling positive and negative mass bodies.

There would then be no substance to the argument that the apparent distinction between acceleration and gravitation which appears to be revealed by the distinct motions of positive and negative energy bodies in the standing still elevator near a local mass allows absolute acceleration (or absolute absence of acceleration) to be determined. Indeed the local gravitational fields and the associated local inertial frames of reference are always determined in a relative fashion as dependent on the presence of the local masses that are the source of the fields, while the reference system where the states of acceleration of positive and negative energy bodies are identical is determined as that relative to which the large scale matter distribution (which we may assume to be unique to positive and negative energy matter) is itself not accelerating. This all follows from the fact that positive and negative energy bodies interact only with the homogeneous matter distribution with the same sign of energy as their own on the cosmological scale⁴, so that motions relative to those matter distributions must be treated differently from motions relative to local matter inhomogeneities, although they are still relative motions. It must be noted, however, that if the homogeneous large scale distributions of positive and negative energy matter were in motion relative to one another there would then actually be two different global inertial frames of reference associated with the two types of mass (positive and negative) experiencing them, even away from any local mass. In such a case it would be more difficult to differentiate between the situation of the elevator far from any large mass and that where unequally distributed concentrations of positive and negative mass matter are present locally. It remains, though, that if positive and negative energy matter are produced by similar processes occurring during the big bang it is plausible to assume that this common origin is a strong enough condition to guarantee that the average state of motion of those two types of matter will in fact coincide, at least initially. If we do not expect to discover a relative motion between the large scale distribution of baryonic matter and that of neutrinos, which decoupled from one another not long after the big bang, we should also not expect negative energy matter to be accelerating or even only moving on the average (on the largest scale)

⁴In fact, as I will later explain, the large scale distribution of negative energy matter may exert an influence on positive energy bodies, but only when inhomogeneities are present in this matter distribution. The nature of those interactions is such, however, that there is necessarily a cancellation in the sum of the effects involved on the largest scale, so that there can be no overall effect and the same is true for the effects of positive energy matter on negative energy bodies.

relative to positive energy matter.

Based on the above discussed considerations I have thus come to the conclusion that, after all, the principle of relativity is not really threatened by the introduction of negative energy matter obeying the requirement of relational definition of its mass sign. But clearly the equivalence principle itself (which allows accelerated motion to be treated relativistically) is no longer to be considered valid in the sense it was traditionally believed to be and if it need not and indeed cannot be abandoned it must, however, be generalized or somewhat relativized. In fact, we already know for sure that the equivalence principle always applies only in local frames of reference whose states of motion can be different in various locations. We can effectively tell that a gravitational field is attributable to the presence of local masses instead of being the consequence of an acceleration, even in the total absence of negative energy matter, when we consider a portion of space that is sufficiently large. For example, if we consider two elevators suspended on opposite sides of a planet, instead of a single elevator, it is obvious that even though observers in each of those elevators could assume that they are accelerating far from any local mass, from the global viewpoint where we would be observing oppositely directed gravitational fields and an absence of relative motion of the elevators we would have to conclude that those fields are due to the presence of a local mass and not to acceleration relative to the homogeneous large scale matter distribution, even in the absence of negative mass bodies in the elevators. In fact, even in a single elevator standing still on the surface of a small planet, freely falling positive mass particles would have a tendency to slightly converge toward one another, therefore betraying the fact that the observed acceleration is an effect of the presence of a nearby mass attracting the particles toward its center. Yet we do not consider the equivalence principle to be violated under such conditions.

What I'm suggesting therefore is that instead of assuming that the equivalence of gravitation and acceleration applies only locally, we have to recognize that it really applies only for a single elementary particle, which would be the most localized physical system we may consider. If we assume that no two such particles can be exactly superposed in an elementary volume of space (which ultimately may be true for bosons just as for fermions if there is a maximum local energy level associated with the Planck scale) we could say that the hypothesis that the equivalence of acceleration and gravitation applies merely within a local free fall frame of reference is equivalent to the assumption that the equivalence principle applies only for a single elemen-

tary particle at once. But then such a particle could have either positive or negative mass and the equivalence principle could be considered to apply not merely to one particle at once, but to one particle with one mass or energy sign at once, which would be a simple generalization of the discussed hypothesis and as such should not raise any further issue (of the kind I have considered so far). For one elementary particle with one energy sign there would never be a difference between acceleration and a gravitational field. It is only when we consider two or more particles of *any* mass sign together, or more precisely in relation to one another, in the presence of a gravitational field attributable to a local matter inhomogeneity (when there is no compensation between the gravitational fields attributable to the local positive and negative energy matter distributions) that we can tell the difference between acceleration relative to the large scale matter distribution and such a gravitational field, but this may be assumed irrelevant when we are considering that no two particles (especially two opposite mass particles) can actually be found in the exact same position at the same time.

It is generally recognized, however, that what makes gravitation different from other interactions is the fact that the motion of bodies in a gravitational field does not depend on the physical properties of those bodies. But even though this characteristic would appear to be violated in the presence of negative energy matter obeying the consistency conditions I proposed, this does not make gravitation any less distinct. Indeed, in the context of the previously discussed viewpoint where it is the direction of the gravitational field attributable to a given matter distribution which varies upon a reversal of the mass of the particle submitted to it (which would actually be considered positive definite), the equivalence principle would merely be relativized by the presence of such negative energy matter, because the difference between the motion of positive energy bodies and that of negative energy bodies would actually be a consequence of the different measures of spacetime curvature which (as I will explain later) can be associated with those two measures of the Newtonian gravitational field. But in such a situation it appears natural to expect that opposite mass bodies should not be restricted to share the same local inertial frames of reference, because in fact they do not even evolve in the same space, but in spaces characterized by different metric properties. Thus, the fact that the gravitational field can be conceived in such an observer dependent way means that in the case of gravitation it is not the reaction that varies when the ‘charge’ is reversed, but the field itself, so that it would still be true that, in any given situation, all bodies

(sharing the same measure of the gravitational field) follow the same motion (acceleration does not depend on the detailed characteristics of the bodies experiencing the same gravitational field). The equivalence principle can thus be assumed to still be valid in the presence of negative energy matter, only it would apply separately for positive and negative energy bodies (just as it applies separately for separate portions of space), because each of those two kinds of matter particle is to be attributed its own free fall frame of reference defined in relation to its mass sign. Therefore, all particles with the same energy sign, whether their energy is positive or negative, would still share the same local inertial frame of reference and this is all that is truly required for a general relativistic gravitational field theory to apply.

2.6 An effect of voids in the matter distribution

It is sometimes recognized that there is a kind of equivalence between the presence of a void in an otherwise uniform matter distribution and what would be the presumed effect of the presence of gravitationally repelling matter present in a quantity and with a distribution equivalent to that of the missing matter. In the context of an expanding universe we would indeed observe underdense regions of the cosmos to be producing a local acceleration of the rate of expansion, while overdense regions would produce a local deceleration of it. The acceleration observed in the case of underdense regions would have all the characteristics of a gravitational repulsion originating from those regions, which would force the matter still remaining inside their volume to migrate to the periphery of what would become the observed voids in the matter distribution. The same effect would also cause nearby underdense regions to merge into even larger spherical voids, as if they were attracted to one another by the force of gravity. This is what all authors who have considered the issue agree must occur when underdense regions form in an expanding universe. Thus, in this particular case, it seems that the gravitationally repelling matter formations would actually be submitted to mutual gravitational attraction with similar formations, even while they would repel oppositely configured formations consisting of overdense regions and would presumably also be repelled by them.

But it is usually considered that there is nothing more than an accidental

analogy between the case of those matter formations and any gravitationally repulsive matter, because if the effect occurs as described above then, according to the traditional understanding, such gravitationally repulsive voids would have to have not only negative gravitational mass, but also positive inertial mass and as everyone ‘knows’ this kind of negative mass is forbidden by the equivalence principle and relativity theory, which require the equality of gravitational and inertial masses. Thus what we would *observe* to be happening is not what most people would consider should occur if we are effectively dealing with gravitationally repulsive matter. Indeed, as I previously explained what is usually assumed is that gravitational repulsion is a kind of definite and invariable property of matter of some type and that this kind of matter would therefore itself also be repelled by matter of the same type. This is usually assumed to be the unavoidable consequence of attributing a negative inertial mass to negative energy matter. But, given the previous discussion and the insights I provided concerning what should be a consistent concept of negative mass or negative energy matter, it should be clear that we would not be justified to argue that the observed phenomenon involving voids in a uniform matter distribution does not replicate the behavior we should expect of negative mass matter. In fact, from my viewpoint it rather seems that the described interaction between overdense and underdense regions of an expanding universe would be exactly that which we should expect to occur if positive and negative masses were actually involved. Therefore, we cannot so easily reject the possibility that the discussed phenomenon is actually telling us something important about the nature of negative energy matter.

I do believe that there is actually more than a valid analogy between voids in a uniform positive energy matter distribution and gravitationally repulsive matter and that there is something very profound which we need to understand concerning the phenomenon described here. Indeed, I think that the discussed equivalence should not be restricted to the case of expanding matter, but must be considered valid even in a local context where the rate of universal expansion is a negligible factor. But if the gravitational dynamics of voids in a homogeneous positive energy matter distribution effectively reflects that which we should expect of a phenomenon involving gravitationally repulsive negative energy matter then it may suggest an interpretation of negative energy matter which would have to do with an absence of positive energy of some kind. It must first be explained, however, why it is that we may actually be allowed to consider that the equivalence discussed

above is valid exactly and constitutes a very general feature of the gravitational interaction despite the objections which might be raised against that possibility.

Basically, what we may object concerning the idea that the presence of a void in a uniform positive energy matter distribution could be equivalent to the presence of an excess of negative energy matter is that it is usually assumed that there can actually be no net gravitational force inside a spherical void in a uniform matter distribution that would be attributable to matter outside the void, a conclusion that seems to be supported by Birkhoff's theorem [16]. What Birkhoff's theorem implies is that there can be no net gravitational force from matter outside any spherically symmetric region in a uniform matter distribution that may itself be considered to be spherically symmetric. This is usually assumed to imply that there cannot be any net gravitational force inside a spherical void in a uniform matter distribution given that such a matter distribution is effectively homogeneous and isotropic. This assumption would effectively mean that in the absence of any matter inside a spherical region there can be no gravitational force at the boundary of the region, as any acceleration could only be attributable to matter inside the region considered and there would then be no matter inside that region.

The influence of voids on the local rate of acceleration of cosmic expansion which was discussed above would thus merely be a result of the fact that the rate of growth of the distance between two galaxies located on the boundary of such a void effectively depends on the density of matter *inside* the void and given that this density would be lower than the average then the rate of growth of the distance, or the local rate of expansion would be larger in proportion with the amount of matter missing inside the void. But that does not mean that it is usually assumed that there would actually be a repulsive gravitational field on the surface of the void. In fact there appears to be some confusion surrounding the issue discussed here, as some authors recognize that there cannot be an equilibrium of gravitational forces in the presence of a void in the cosmic matter distribution and yet they fail to recognize that this may actually give rise to repulsive gravitational fields for the surrounding positive energy matter, probably because they assume that the effect of the noted disequilibrium would be that which is observed to affect the local rate of expansion, while actually this is a distinct (but not entirely unrelated) effect associated merely with cosmic expansion. But what I believe must be recognized is that there would effectively be gravitational repulsion in the

presence of an underdensity in an otherwise uniform matter distribution, not only at the boundary of the surface, but everywhere inside the void with a net force that would decrease to reach a null value as we approach the center of the void. This situation would then clearly be different from that we would have in the case of a hollow sphere of finite size inside of which the Newtonian gravitational field should effectively be zero everywhere.

It must in effect be understood that contrarily to what is usually believed, Birkhoff's theorem does not forbid this conclusion, because the decisive condition entering this theorem is that of spherical symmetry, which would effectively be obeyed if we were considering a hollow sphere or a universe that was spherically symmetric around *any* point on any scale, but which I suggest would fail locally for a universe with an actual void in its matter distribution. Indeed, the case of a homogeneous and isotropic universe is equivalent to that of a sphere of finite size only when the universe is considered at the scale at which its matter is uniformly distributed and no significant void is present, which effectively explains why Birkhoff's theorem (which is a necessary element of current cosmological models) is observed to apply on a cosmological scale. But I think that it would only be in the case of a spherical region centered on an actual sphere of matter of finite size located within an otherwise empty universe that the theorem discussed here would effectively remain valid regardless of the distribution of matter inside the spherical region, because only in such a case would we be dealing with a spherical symmetry that is not dependent on the position of the observer. What we usually fail to recognize is that the fact that the matter distribution in the universe would be symmetric around *any* location in the absence of a void in its homogeneous and isotropic matter distribution means that the presence of a void would necessarily alter the equilibrium of forces around that void.

It is clear indeed that in the presence of a uniform matter distribution extending throughout the universe an equilibrium exists locally between the sum of forces attributable to the interaction of a freely-falling body with all the matter in the universe and therefore the removal of a certain quantity of matter in a region of finite volume must have an effect that would be the opposite of that which we would attribute to the presence of an equivalent *additional* quantity of matter in the same region of the same universe (in the absence of the void). This should be expected to occur due to the fact that the removal of a certain amount of positive energy matter to create a void would eliminate the attractive gravitational force which would otherwise be exerted

on positive energy matter by the matter in the void and given that there was no net force before the creation of the void then the other forces which were initially present must now give rise to an acceleration directed away from the void and of similar magnitude to that which would have been produced by the matter that filled the void. Thus, for positive energy matter there would appear to be a repulsive gravitational force originating from the presence of a void in such a uniform matter distribution, which would actually be the consequence of an uncompensated gravitational attraction attributable to the positive energy matter outside the void. But this is a valid conclusion only when we recognize that Birkhoff's theorem is not valid in the sense it is usually assumed to be and that the case of a spherical distribution of matter of finite size with a central cavity is *not* equivalent to the case of a void in a uniform cosmic matter distribution.

What must be understood is that if, in the case of a hollow sphere of finite size, the subtraction of matter to create the cavity does not result in a net force originating from the matter outside the cavity *that is part of the sphere* this does not mean that it would also be the case that there would be no acceleration inside the cavity resulting from the gravitational interaction with the *entire* universe (unless it was effectively assumed that the universe is empty except for the presence of the sphere). What is wrong therefore is the idea that when we are considering a spherical region of the universe the rest of the universe surrounding that region can be considered as a hollow sphere simply on the basis of the fact that according to the cosmological principle matter is distributed uniformly in all directions. It must be understood that such a spherical region in a uniform matter distribution would be free of uncompensated external forces only if it is itself filled with matter as uniformly distributed as the matter found outside the region (which is effectively verified on a cosmological scale in our universe), because it is only in such a case that the spherical symmetry would apply to any point inside the spherical region. Again, it may be noted that, in this context, the fact that the concept of the hollow sphere is nevertheless appropriate to describe the dynamics of the universe on the largest scale is due merely to the fact that we do not actually consider spherical voids in the matter distribution, but really a uniformly filled matter distribution without any spherical regions devoid of matter on the particular scale that is considered (as a requirement of the cosmological principle).

It must be clear that I am not suggesting that there would be uncompensated gravitational forces in the case of the finite size hollow sphere itself

(if it was located in an empty universe for example). In fact, the problem here has to do again with the fact that we fail to apply the requirement of relational definition of physical properties when we are dealing with the resultant effect of the gravitational forces attributable to the universe as a whole. Indeed, from the traditional viewpoint, when we are dealing with a chosen spherical region of the universe we are implicitly assuming that the surrounding matter which may influence the particles located inside that region (through the gravitational interaction, even if there is no net force) is spherically distributed around the center of the spherical region considered, as if the location of the center of mass of the universe was an intrinsic invariable feature of the whole configuration. But the center of a matter distribution in a physical universe without boundary is not an absolute feature (as would be the case for a hollow sphere), but must be defined in a relational manner as any other property, if we are to be able to determine the consequences on a given object of being located in such a position. When we are dealing with the matter distribution in a universe without spatial boundary and in which the local inertial frames of reference are determined by the entire matter distribution (following Mach's principle) the true center of mass defined in terms of the influences exerted on a given body is *always* located right at the position where that body is to be found, wherever this position may be in the matter distribution.

Thus, a particle located *at the center* of a void in a uniform matter distribution could effectively be considered to be in the situation of a particle in a hollow sphere, because for this particle the whole sphere of influence of the universe is centered on the void (in this situation the surrounding matter actually is a hollow sphere centered on the particle's position). Therefore, such a particle would feel no uncompensated gravitational force from the whole universe, as required. But if this particle moves to one side or another in the void, the matter distribution influencing the particle in its new position would be centered on the new position and this means that the void in the previous hollow sphere is shifted to the opposite side, just as the sphere itself is shifted in the direction of the particle's new position. The symmetry of the initial configuration would therefore no longer be present and the equilibrium of forces would no longer apply. In the new configuration a whole layer of matter must be 'removed' on one side of the external surface of the imaginary hollow sphere (in the direction opposite the particle's displacement) and added on the other (this is easier to visualize in a closed universe) which, given the distances involved, means that an enormous amount of mat-

ter has changed position from the viewpoint of the particle. It must therefore be recognized that in the final configuration the void in the imaginary sphere is no longer centered on the center of mass of the sphere, but is actually located away from the center of the sphere. As a consequence, the spherical symmetry from which depended the conclusion that there would be no net gravitational force inside the sphere is no longer to be found in the final configuration experienced by the particle and therefore it must be expected that there would be a net gravitational force on the particle and an acceleration relative to the matter distribution.

It is important to understand that however large you consider the imaginary sphere encompassing the matter distribution (the size of the universe) to be when dealing with the effects of the gravitational interaction with the whole universe, if the center of the sphere is shifted to one side there would be a non-negligible effect from the displacement of its center of mass. This is true even if the distance to the periphery of the sphere (where the changes occur) is very large and the strength of the gravitational interaction decreases with the square of the distance, because the larger the distances (the larger the sphere) considered, the larger the quantity of matter that is shifted from one side to the other and thus the larger the changes involved in the local gravitational field. We should not be surprised then that even the retarded interaction with matter so distant could have an effect similar in magnitude to the effect that would be exerted by the matter missing from a void located near some particle experiencing those forces. If the center of mass of the universe is always located at the position of the particle experiencing the gravitational effects of all the infinitesimal elements of matter in this universe then the local effect of the absence of gravitational attraction from those portions of matter which would be present if a nearby void in the positive energy matter distribution was absent would necessarily result in a net force on positive energy matter arising from the gravitational attraction of all the portions of matter located on the opposite side of the void. But such a force would be completely equivalent to a repulsive gravitational force arising from the void itself.

The fact that from a practical viewpoint a local void in a uniform positive energy matter distribution would actually have to occur through the expulsion of positive energy matter outside the region that is to become the void and therefore would necessarily produce a compensating overdensity of negative energy matter in the region surrounding the void would not forbid the existence of a net repulsive force on positive energy matter inside the void

even though it effectively means that there would be no resulting force on matter located some distance away from the void. If we consider for example the ideal situation of a spherical void produced through the creation of a surrounding spherical shell of positive energy matter at higher than average density, then as long as a positive energy particle is located outside this shell it would effectively feel no net force because any reduction of attractive force from the void would be compensated by an increased attractive force arising from the presence of the shell. But as soon as the particle would enter the shell it would start to experience the equivalent gravitational repulsion, because the outer layers of the shell would no longer provide any net force on the particle while the void for its part would still exert its net effect, because the equivalent repulsive force it produces is attributable to *all* the surrounding matter (whose distribution is centered on the position of the particle) and not just to the spherical shell. Thus the case of the particle which experiences no gravitational force *at the center* of a void in a uniform matter distribution is merely a particular case of the more general description where there is actually a net force everywhere inside the void, except at the exact location of its center, as would be the case if we were considering the gravitational attraction attributable to an isolated sphere full of matter (like a planet or a spherical gas cloud). This is an important result which will have decisive consequences for a consistent description of the nature and properties of negative energy matter.

Concerning the insight just described it is important to note that even if under certain circumstances there may be an equivalence between an imbalance in the sum of gravitational attractions attributable to all the positive energy matter elements in the universe and what would appear to be a gravitational repulsion exerted on a positive energy body, we are nevertheless always dealing with gravitational attraction. Indeed, there is no question that it is the gravitational attraction of positive energy matter that is responsible for the apparent gravitational repulsion which would be exerted on a positive energy body by a void in the otherwise uniform positive energy matter distribution. It is clearly as a consequence of the fact that positive energy matter is missing in the direction where the void is located, while the matter present in the opposite direction still exerts its gravitational pull, that there exists a net force directed away from the void. Thus what looks like a gravitational repulsion exerted in a given direction by some matter configuration and which could from a certain viewpoint be equivalent to it would actually be the product of

a gravitational attraction arising from an absence of matter exerting a compensating attraction in the opposite direction. This is particularly significant in the context where local inertial frames of reference are to be considered as always arising from a perturbation of the equilibrium of inertial gravitational forces by the gravitational forces attributable to local matter concentrations, as I have emphasized in the preceding section. Yet the fact that we are here dealing only with gravitational attraction does not rule out the validity of the analogy which may exist from a classical viewpoint between the presence of true gravitationally repulsive, negative energy matter and an absence of positive energy of some sort. In fact, it rather seems that what allows an interpretation of negative energy matter as being equivalent to an absence of positive energy to be valid as a general feature of gravitation theory is the possibility that always exists (not only in the case of voids in a uniform matter distribution) of attributing an apparent gravitational repulsion to uncompensated gravitational attraction.

To explain what motivates that conclusion I may recall the previous discussion concerning the occurrence of negative energy in certain experiments described using the methods of quantum field theory. There I pointed out that the absence of some positive energy states from the vacuum in certain limited regions of space (between the plates of two parallel mirrors for example) can actually give rise to a vacuum with negative energy density in the volume considered, because removing positive energy from a vacuum state whose energy is already minimum is like decreasing the energy below its zero point into negative territory. The fact that the vacuum is known to have only a very small energy density should not be considered an obstacle to the occurrence of large negative energies in such a way, because as I will explain later this small energy density appears to be the outcome of very large (actually maximum) but (mostly) compensating opposite energy contributions, which could be reduced to an arbitrarily large extent by the conditions effectively responsible for locally decreasing (under particular circumstances) the energy of the vacuum below the cancellation point. But if we may effectively attribute a negative energy to some configurations in which particular states are missing from the vacuum along with their contribution to the total energy of this vacuum, then there is no reason why we could not consider that negative energy states in general are equivalent in some ways to an absence of positive energy from the vacuum, if from a phenomenological viewpoint there is no distinction between those two situations.

I must again mention in this regard that many physicists have expressed

doubts concerning the concept of vacuum energy as arising from fluctuations involving virtual particles and have suggested that there may be nothing real with the processes so described outside of the context where they are occurring as part of otherwise real processes involving ‘real’ particles. But I think that it is precisely the fact that the existence of those processes would imply the reality of negative energy states that really motivates this mistrust, because it is no secret that for most physicists the theoretical possibility of the existence of negative energy states is not well viewed. However, I believe that this aversion is merely a consequence of the fact that the traditional concept of negative energy matter is effectively not viable and that it has not yet been realized that a better description of negative energy matter is possible and even necessary, as I emphasized before. In any case the idea that virtual processes would only occur as part of otherwise real processes, thus explaining why we must nevertheless consider the effects of such fluctuations when calculating transition probabilities, is meaningless, because in a given universe anything that occurs is related (directly or indirectly) to everything else and even in empty space, far from any ‘real’ matter, the virtual processes of particle creation and annihilation characteristic of the quantum vacuum would occur as an integral part of the surrounding real processes to which they are connected by relationships of local causality (between adjacent portions of this vacuum) and their common origin in the big bang. Therefore, the argument that the negative energy states predicted to occur in the vacuum under the right conditions are not real, because our description of the vacuum is itself not appropriate in general, cannot be retained. Also, the fact that it has been confirmed that the cosmological constant is not absolutely null is a strong motive to conclude that the rejection of the reality of vacuum fluctuations as essential aspects of our description of empty space is not vindicated from the viewpoint of observations and therefore that negative energy states are a real possibility.

I have already explained why we should expect mutual gravitational attraction between two bodies with the same sign of energy and gravitational repulsion between opposite energy bodies. But on the basis of my conclusion concerning the nature of the gravitational force on a positive energy body that would be attributable to voids in a uniform positive energy matter distribution we now also have the possibility to assert what would be the effects of missing positive energy from the vacuum. Indeed, given that the vacuum is to be conceived as involving a constant and uniform density of energy on the largest scale, any negative local variation in its density must share the

features of voids in a uniform matter distribution. It therefore appears that if the presence of voids in an otherwise homogeneous positive energy matter distribution effectively produces an equivalent gravitational repulsion on positive energy bodies, then the absence of positive vacuum energy in localized regions should effectively exert an equivalent gravitational repulsion on the surrounding positive energy matter. This would occur as a result of the fact that an absence of positive energy from a region of the vacuum would result in an uncompensated gravitational attraction from the surrounding positive energy vacuum pulling positive energy matter *away* from the region where the energy is missing. From that viewpoint we can thus deduce that the physical properties (related to the gravitational interaction) that we should expect to be associated with missing positive vacuum energy are the same properties which I explained we should expect to be associated with the presence of negative action matter, which confirms that from a phenomenological viewpoint negative energy matter is gravitationally equivalent to an absence of positive energy from the vacuum.

Given this equivalence between negative energy and absence of positive energy from the vacuum, it follows that if states of negative vacuum energy are allowed by current theories then we must conclude that negative energy matter is itself allowed to exist and may not always be constrained by the limitations observed to apply in the currently considered experiments where it occurs merely as a consequence of the suppression of positive energy from the vacuum consequent to singular configurations of otherwise positive energy matter. It must be recognized, however, that if the presence of negative energy matter in a region of space is equivalent for positive energy matter to an absence of positive energy from the vacuum this is simply because in general for an equilibrium state of any kind the presence of a negative contribution is equivalent to the absence of a positive contribution of the same magnitude and it just happens that the vacuum is a physical system that appears to arise from precisely such an equilibrium state. But we must remember that a void in a uniform matter distribution of a given energy sign (not involving the vacuum) is physically different from a local absence of vacuum energy of the same sign, even if in both of those cases the effects are equivalent to the presence of an excess of matter of opposite energy sign, because in the first case we are dealing with an *absence* of matter of a given energy sign, while in the latter case we are actually dealing with the *presence* of matter (of opposite energy sign).

At this point it is important to mention that there would occur a phenomenon of gravitational repulsion similar to that described above, but which would apply from the viewpoint of negative energy matter in the presence of voids in a negative energy matter distribution or in the negative energy portion of the vacuum. Indeed, using the same logic that allowed me to derive the consequences of the presence of a void in a uniform positive energy distribution it is possible to deduce that the absence of negative energy from an otherwise homogeneous matter distribution would actually be equivalent from a gravitational viewpoint to the presence of a concentration of positive energy matter. One assumption that will be crucial for my derivation of the modified general relativistic gravitational field equations is indeed that the equivalence described here is valid both ways and that positive energy matter can always be considered to actually consists of voids in the *negative* energy portion of the vacuum, which makes the whole situation symmetrical in a way that does not even depend on the viewpoint of the observer. It must be clear, however, that I am not suggesting that positive energy matter is equivalent to voids in a filled distribution of negative energy matter, even if I do suggest that we must assume that an absence of negative energy matter from an otherwise uniform distribution of such matter would effectively have gravitational effects similar to those attributable to the presence of positive energy matter. I must emphasize once again that a void in a uniform *matter* distribution remains clearly distinct from a void in the uniform energy distribution of the *vacuum*. This means that my proposal is distinct from Dirac's failed hole theory (proposed as an attempt to solve the negative energy problem), in particular because what I am suggesting is that *all* positive energy matter particles (and not just antimatter particles) are actually equivalent to voids in the negative energy portion of the vacuum rather than in a filled continuum of negative energy matter.

What Dirac proposed in effect is that all negative energy states are already occupied, so that positive energy fermions at least should not be expected to make transitions to those negative energy states. But even if the existence of such a filled, uniform continuum of negative energy matter had no effect on positive energy matter (perhaps due to its uniformity), the fact that from my viewpoint there would be no reason to assume that positive energy states are not completely filled in the same way means that this hypothesis would not agree with observations. Indeed, it is not possible to assume, in a theory that respects the requirement of a purely relational definition of the sign of energy, that positive energy antiparticles are merely voids in

a completely filled negative energy matter continuum, as Dirac proposed, without also assuming that negative energy antiparticles would be voids in a completely filled positive energy matter continuum. But given that positive energy states are obviously not all occupied by matter particles it appears that this requirement cannot be satisfied. We may then instead assume that *all* positive energy particles are voids in a filled negative energy matter continuum, but again in such a case we would have no reason not to assume that all negative energy particles are also voids in a filled positive energy matter continuum. The problem, however, is that it seems impossible to assume that we could have a completely filled distribution of negative energy matter and at the same time a completely filled distribution of positive energy matter if negative energy matter is to also consists of voids in a filled distribution of positive energy matter, because so many voids in the positive energy matter distribution as would be necessary to describe the filled negative energy matter distribution would leave no possibility for the positive energy matter distribution to itself be nearly filled.

What cannot be assumed therefore is that negative energy states are completely filled and positive energy particles are voids in this negative energy distribution while positive energy states are completely filled and negative energy particles are voids in this positive energy distribution, because those two possibilities are mutually exclusive (cannot occur together). But while it may perhaps appear appropriate from an observational viewpoint to assume that we simply have a filled negative energy matter continuum combined with a nearly empty distribution of positive energy matter, there would also be problems with such a proposal. Indeed, what reason would we have not to assume that it is only the positive energy matter distribution that is filled (even though this assumption would clearly contradict observations)? The problem is that we cannot in effect postulate that both positive and negative energy matter are voids in their respective opposite energy matter distributions if we also postulate that there is no absolute (non-relational) difference between positive and negative energy matter. In other words, it is not possible to assume symmetry under exchange of positive and negative energy particles if matter of a given energy sign is to be conceived as voids in the matter distribution of opposite energy sign and this simply because matter cannot be at once present and absent. The truth is that any description of matter or antimatter as voids in a matter distribution of opposite energy sign would require giving preferred status to negative energy matter as being the matter whose distribution is completely filled (because obviously the positive

energy matter distribution at least is not completely filled) and this would break the requirement that only differences in the energy sign of particles are to be conceived as physically significant.

What must be clear, therefore, is that if we were to make use of such a description we would allow the identification of a preferred sign of energy as being that which would be associated with the filled matter distribution, while from a theoretical viewpoint that should be considered impossible. A theory of matter as voids in a uniform, opposite energy matter distribution would in effect imply that the requirement of symmetry under exchange of positive and negative energy matter is violated in a way that cannot be allowed if the sign of energy is to be conceived as a relationally defined physical property. Thus, it must be recognized as forbidden to consider that the presence of matter with a given energy sign could be explained as resulting from the presence of voids in a *matter* distribution of opposite energy sign, even if there does exist a phenomenological equivalence between the effects of missing positive or negative vacuum energy and the absence of matter from a homogeneous distribution with the same sign of energy, because again those are two distinct phenomena. The contradiction which would occur if we were to assume that positive energy particles are voids in a filled uniform distribution of negative energy matter, while negative energy particles are voids in a filled uniform distribution of positive energy matter is that we would require the presence of a lot of particles of both energy signs to fill the matter distributions and at the same time the presence of a limited number of particles of both energy signs due to the presence of all the voids attributable to the presence of the nearly filled opposite energy matter distributions. According to my proposal by contrast it becomes possible for both positive and negative energy particles to actually exist as real observable particles independently from the presence of one another. Thus, if the voids in the negative energy portion of the vacuum, which I assume to be equivalent to the presence of positive energy matter, are not equivalent to voids in a hypothetical filled distribution of negative energy matter it is simply because in fact voids in the vacuum cannot be equivalent to an absence of voids in the vacuum.

I may add that from the viewpoint of a consistent interpretation of negative energy matter there would also be a problem with Dirac's original proposal that a void in the filled negative energy continuum could be created along with a positive energy particle (as would a particle-antiparticle pair) when photons provide enough energy to raise a negative energy particle to

a positive energy level. Indeed, as I mentioned before and for reasons I will explain later, a consistent theory of negative energy matter would require that negative energy matter be dark, which means that there would be no electromagnetic interactions between opposite energy particles and therefore a positive energy photon could not even interact with a negative energy electron to provide it with the required positive energy. Thus even if we insist on assuming the existence of a filled negative energy continuum we could not use this hypothesis to explain the existence of antimatter.

It is essential to understand, therefore, that the situation we would have if all negative energy states were filled is different from that we would have when dealing with a vacuum in which there would be a very large negative contribution to the energy density of zero point fluctuations. Indeed, in contrast with the vacuum, a negative energy matter distribution which would be filled at one particular epoch would no longer be filled at a later time given that space is expanding. This is reflected in the fact that vacuum energy obeys an equation of state which is different from that of a homogeneous matter distribution. Also, even if there is a large negative contribution to the energy of the fluctuating vacuum there is no reason to expect that it gives rise to a situation similar to that which would occur if space was filled with negative energy matter, because there must also be a large positive contribution to the energy of empty space. A space filled with positive or negative energy matter would be as different from the true vacuum as the primordial soup which existed in the first instants of the big bang is different from the space nearly devoid of particles that currently exists between galaxies. Thus if a theory of voids is to have any relevance in a gravitational context it must involve a description of matter of any energy sign as consisting of voids in the opposite energy portion of the vacuum, so that the presence of matter with a given energy sign does not imply an absence of *matter* with opposite energy sign.

When the energy distribution in which the voids equivalent to the presence of positive energy matter occur is the negative energy portion of the *vacuum* it therefore becomes possible to assume the presence of arbitrarily high or arbitrarily low densities of matter of both energy signs all at once in the same region of space, because in effect the presence of matter of one energy sign in a given location does not preclude the presence of matter with an opposite energy sign in the same location (at least when the matter distributions are smooth enough). Thus we do not need to assume the presence at all times of a nearly filled negative energy matter continuum combined with a

distribution of positive energy matter of arbitrarily low density, which would otherwise be the only (perhaps) observationally acceptable configuration, but which would also have allowed to establish an absolute (non-relational) distinction between positive and negative energy matter, as I just explained. But what makes the vacuum particularly suitable for accommodating the above proposed description of matter as consisting of voids in some uniform energy distribution is the fact that we are effectively allowed to assume that there are both positive and negative contributions to vacuum energy density, even as arise from otherwise identical particles. We can therefore expect a certain level of compensation between the gravitational effects of those two contributions that would enable to obtain an arbitrarily small residual value for the cosmological constant. Indeed, one of the consequences of the assumption that there exists a distinct negative component to vacuum energy is that the natural value of the cosmological constant which we can expect to observe under such conditions is actually as small as the symmetry under exchange of positive and negative energy matter is perfect. This is all by itself an important result which will have an impact on many aspects of cosmology theory.

What is perhaps even more significant, however, is that when we understand that all positive and negative energy particles are actually equivalent to voids in their respective opposite energy portions of the *vacuum*, as I propose, then the unsatisfactory categorical distinction between matter and vacuum actually becomes meaningless. This is because in such a context all matter can effectively be considered to consist in a particular manifestation of some property of the vacuum. It is by building on this insight that I will be able to provide a unified and totally symmetric description of the gravitational dynamics of positive and negative energy matter according to which the measure of energy of matter is significant merely in relation to an energy scale associated with objective properties of the vacuum. I was able to obtain those results only at a relatively late stage of my reflection, because I had initially assumed that only the nearly vanishing total energy density of the vacuum could have an influence on matter of any energy sign and that the positive and negative contributions to vacuum energy could not be considered independently from one another. But once I realize the inappropriateness of this hypothesis the above discussed results emerged as clearly unavoidable and extremely significant. The notion that both positive and negative energy particles are actually voids in their respective opposite energy portions of the vacuum therefore appears to be the ultimate embod-

iment of the requirement of a relational definition of all physical properties understood as a basic consistency condition that must apply to any physical theory.

Before delving deeper into the consequences of the assumption that missing positive vacuum energy is equivalent to the presence of negative energy matter it is important to mention that when we are considering a void in the positive energy portion of the vacuum we must ultimately always be referring to the absence of a specific elementary particle from this vacuum and not merely to the absence of a certain amount of energy. But while the vacuum must be mostly neutral from the viewpoint of its energy content in the absence of voids, as I have just remarked, it is also quite clear that it should be expected to be neutral from the viewpoint of non-gravitational charges as well. If that is effectively the case, then we must conclude that removing a particle with a given charge from the vacuum must be equivalent to creating a charge of opposite sign and this for the same reason that removing positive energy from the vacuum is equivalent to creating negative energy. This conclusion is particularly appropriate in the context where a real particle must be conceived as a many-body phenomenon in the sense that it is at all times surrounded by a cloud of virtual particles, so that removing even just one particle from the vacuum can be expected to have consequence at an observable level despite the fact that the actual particle involved may be idealized as a dimensionless object. Thus if we must consider specific positive energy particles to be absent from a vacuum when ‘real’ negative energy matter is present then this negative energy matter must be assumed to consist of particles with a sign of charge opposite that of the particles missing from the positive energy portion of the vacuum. In this context, even the charges which we currently attribute to positive energy particles must be considered to actually be a manifestation of the absence of opposite charges from the negative energy portion of the vacuum. This is an unavoidable consequence of the description of all matter as consisting of voids in the uniform distribution of zero-point vacuum fluctuations.

Concerning the effects which I am suggesting should be attributed to energy missing either from a homogeneous matter distribution or from the homogeneous vacuum we may ask to what extent a void may effectively be considered as physically significant in the sense of being merely an anomaly in an otherwise uniform distribution of matter or energy. If we examine the

situation carefully it becomes clear in effect that given that for both matter and vacuum it must be the surrounding energy that would exert the *outward* directed gravitational pull that would be experienced as a gravitational repulsion, it follows that as we consider voids of larger sizes there may come a point when there would be no matter left outside the void to produce the uncompensated attraction that must exist to produce the equivalent repulsion. Normally this is not an issue, as any void that forms in a matter distribution which can be assumed to be arbitrarily smooth initially (and this appears to be a feature of our universe at the big bang) will necessarily involve the creation of a surplus of matter in its surroundings, which for a remote observer would have the effect of compensating the equivalent force arising from the presence of the void itself, as I previously mentioned. Such voids, regardless of how large they may become, would therefore leave the universe at large in a state equivalent to that of uniform density which would allow it to continue to exert its influence in the empty regions. But if we are to consider the equivalence between missing positive vacuum energy and the presence of negative energy matter to be generally valid, then the presence of a uniform negative energy matter distribution would imply the existence of a void in the positive energy portion of the vacuum which would effectively extend to the whole universe. This void would have been present in the vacuum from the very beginning of the universe's history and would not have developed through the production of some inhomogeneity. In such a case we would no longer be able to assume the existence of an uncompensated gravitational pull on positive energy bodies from the surrounding positive vacuum energy, because indeed there would be no surrounding higher energy vacuum to effect the attraction. Under such conditions, therefore, I am allowed to conclude that no outward directed gravitational force which we could assimilate with an equivalent gravitational repulsion would exist.

Now, given that I will later argue that the equivalent gravitational repulsion exerted on positive energy matter by voids in the positive energy portion of the vacuum actually constitutes the only form of gravitational interaction between this matter and negative energy matter, it would appear that the preceding conclusion imposes very strong limitations on such an interaction. Indeed, it transpires that the absence of equivalent gravitational repulsion on positive energy matter from a completely homogeneous negative energy matter distribution, is a very general and unavoidable feature of the description of the gravitational interaction between positive and negative energy matter. This is because such a limitation would also be verified in the case

of a uniform distribution of positive energy matter from the viewpoint of negative energy bodies if the gravitational repulsion exerted on those objects by positive energy matter can be attributed to an absence of *negative* energy from the vacuum. Thus, if opposite energy bodies can be shown to interact only through their respective vacuums, we would be allowed to conclude that negative energy matter interacts with positive energy matter only in the presence of inhomogeneities in any of the two matter distributions. But given that only an inhomogeneity that develops over the initially smooth negative energy matter distribution (if we may suppose that negative energy matter is as homogeneously distributed as positive energy matter on the cosmic scale) can contribute to the gravitational dynamics of positive energy matter and given that the formation of such an inhomogeneity would involve the formation of a compensating one involving an opposite variation of density in the surroundings of the first, we must then conclude that the presence of an average density of negative energy matter has absolutely no effect (at least from a gravitational viewpoint) on the gravitational dynamics of positive action matter (and vice versa). This would mean in particular that the rate of universal expansion of positive energy matter cannot be influenced by the presence of negative energy matter and similarly that the expansion of negative energy matter is not affected by the presence of positive energy matter. This, again, is a very significant result which will have important implications for cosmology.

I may add that the conclusion discussed here is the one on which is founded the hypothesis discussed in a preceding section of this report (dealing with the significance of the equivalence principle in the context of the presence of negative energy matter) which allowed a relational explanation of the phenomenon of inertia. There I explained that if both the large scale positive and negative energy matter distributions were to exert an influence on positive energy bodies, then the hypothesis that accelerated motion is relative would be invalidated in the presence of negative energy matter on a cosmological scale. Indeed, under such circumstances there would be equal and opposite imbalances in the sum of gravitational forces (to which we would try to attribute the resultant inertial force) arising from the acceleration of a positive mass body relative to the two opposite energy matter distributions whose average states of motion should correspond with one another on the largest scale. But if only matter of positive energy has a gravitational effect on positive energy bodies on the cosmological scale, then the global inertial frame of reference experienced by a positive energy body could effectively be

determined by the average state of motion of positive energy matter given that the inertial force exerted on such a body would result only from its gravitational interaction with the large scale distribution of positive energy matter. Thus, we can now see why the rejection of the assumption that a uniform negative energy matter distribution can exert a force on positive energy matter (and vice versa), which appears to be required for a relational explanation of the phenomenon of inertia based on the principle of relativity, was effectively justified. The preceding discussion actually shows (when we recognize that positive and negative energy matter can interact only through the effect they exert on the opposite energy portions of the vacuum) that this hypothesis is not only desirable, but actually constitutes an unavoidable consequence of the description of negative energy matter as being equivalent to missing positive vacuum energy.

But in the context where the description of negative energy matter as being equivalent to voids in the positive energy portion of the vacuum is similarly applied to positive energy matter (in the sense that positive energy matter would be equivalent to the presence of voids in the negative energy portion of the vacuum) a further distinction would arise. Indeed, just as negative energy matter would interact with itself independently from the fact that it is equivalent to voids in the positive energy portion of the vacuum, positive energy matter, as voids in the negative energy portion of the vacuum, would have to still interact with itself, which means that such voids do interact with themselves. In fact, even if the missing negative energy was uniformly distributed throughout all of space it would still exert an influence on itself despite the fact that a similar distribution of missing positive vacuum energy would have no effect on *positive* energy matter, that is, on voids in the *negative* energy portion of the vacuum. In other words, the fact that a void in the negative energy portion of the vacuum, which is equivalent to the presence of positive energy matter, could leave no outside surrounding negative energy to affect the behavior of negative energy matter (if this void is uniformly distributed throughout the entire volume of the universe) would not affect the ability for such a void to gravitationally attract *positive* energy matter or other voids in the negative energy portion of the vacuum, because in such a case the interaction is actually occurring between the matter particles themselves (or the voids) and not between a particle and the surrounding vacuum with the same energy sign.

Finally, it may be of interest to mention that if we were to consider the

effect on a positive energy body of a void in a uniform negative action *matter* distribution then based on the above discussed insights we should deduce that the outcome would be a gravitational attraction directed toward the center of the void. This could be predicted to occur in two different ways. First, given that we can now expect negative energy matter to exert a gravitational repulsion on positive energy bodies, then on the basis of what has been learned concerning the effects of voids in a uniform matter distribution we could conclude that the absence of gravitational *repulsion* in the direction of the void consequent to the absence of negative energy matter in this void would give rise to an uncompensated *repulsive* force directed toward the center of the void, which would be equivalent to a gravitational attraction directed toward the center of that same void, but which would effectively arise from the gravitational repulsion of the surrounding negative energy matter. But given that we now also know that a uniform distribution of negative energy matter has no influence on positive energy bodies it would seem preferable to derive the consequences of an absence of such negative energy matter based on an alternative approach which borrows from the results discussed in the preceding paragraphs.

Indeed, what allows me to conclude that a uniform negative energy matter distribution has no effect on positive energy bodies is that the presence of such uniformly distributed matter is equivalent to a void of universal proportion in the positive energy vacuum which therefore leaves no surrounding positive energy to produce uncompensated gravitational forces. But then, if you remove negative energy matter in a portion of this void the resulting configuration would be that of an imperfect void or an imperfect distribution of *absence* of positive energy from the vacuum. But a local absence of *absence* of energy is really just the same as a local presence of energy and if the energy that was absent (when negative energy was present) was positive then the energy that is locally present will itself be positive. This local absence of negative energy matter will thus be totally equivalent to the presence of an equal amount of positive energy matter and should therefore be expected to produce on positive energy bodies a gravitational attraction directed toward the void. This is an effect which may have interesting consequences on the cosmological scale where variations in the density of negative energy matter may have a magnitude comparable with the average density of the matter itself. In any case, the effectiveness of this description is a further confirmation of the existence of a close relationship between vacuum energy and matter energy and the high level of symmetry involved also indicates that

the description of the properties of negative energy matter proposed above fully agrees with the requirement of a relational definition of the physical property of energy sign.

2.7 Six problems for negative energy matter

The preceding discussion may already make us feel more comfortable with the possibility that there could exist negative energy matter, despite the traditional reluctance to accept the reality of negative energy states. But at the current stage of my account this confidence would not yet be totally appropriate. Even in the context of the new understanding unveiled in the previous sections there indeed remains many problems associated with the possibility that negative energy matter may exist in our universe. First of all, we do not observe in the universe any matter or celestial object which would clearly appear to be involved in repulsive gravitational interaction with other material bodies. This is a very basic but also very constraining fact. Associated with this problem is the fact that the current predictions of quantum field theory are based on a systematic rejection of the possibility of a transition to negative action states (as states of negative energy propagated forward in time) and yet they appear to produce results which agree very well with observations in all situations where the nature of the interactions involved is well understood and appropriate use of the associated computational methods can be made. This could provide an additional motive for arguing against the possibility of the existence of negative energy matter. Such pieces of evidence certainly cannot be dismissed without very good reasons. Any theory involving particles propagating negative energies forward in time must explain why it is that we can safely ignore the existence of those particles in formulating a quantum theory of elementary particles and their interactions, even while we would presumably have to take their effects into account in a classical astronomical context where the effects of the gravitational interaction are not negligible.

A second category of difficulty has to do with the possibility that seems to be allowed, in the context where negative energy particles would exist, to observe the annihilation of particle-antiparticle pairs where one of the particles would have negative energy, therefore permitting matter to vanish, leaving absolutely nothing behind. This would of course require the annihilating opposite energy particles to also have opposite electric and other

non-gravitational charges, because charge must still be conserved. We have no reason, however, to assume that negative action matter does not also come in two varieties, one propagating negative energy and all non-gravitational charges forward in time and the other propagating positive energy and the same charges backward in time (so that we have opposite charges from the forward time viewpoint). Therefore, we cannot a priori reject the possibility that such annihilations could take place. But that is a much worse problem than may perhaps appear to be, because if such annihilations were possible there would then be no reason why the time-reverse processes could not also take place. If that was the case it would actually mean that pairs of opposite energy particles could be spontaneously created out of nothing without immediately returning to the vacuum like ordinary pairs given that the process could occur without requiring a violation of energy conservation.

Currently there appears to be no justification for the fact that those matter creation processes occur at a rate that is as low as required by observations which actually provide no evidence at all that such events take place. The fact that opposite energy particles would gravitationally repel one another should not prevent an annihilation process involving such pairs from taking place, as the gravitational interaction is very weak and the fluctuations present in the vacuum could still allow the process to occur at least when charged particles are involved, because the opposite charges carried by the particles would give rise to attractive forces that would counter the gravitational repulsion. Indeed, if the electrostatic attraction between opposite charges does not prevent ordinary particle-antiparticle pair creation process from occurring then there is no reason why such an effect would need to be taken into account in the case of pair annihilation processes involving opposite action particles. In any case the fact that the gravitational repulsion between opposite energy particles would not affect the possibility for the associated *creation* process to occur means that the problem is real. Therefore, in the absence of an appropriate natural constraint, a continuous process of matter creation from all regions of space should occur even long after the big bang. This category of difficulties may then be called the energy out of nothing problem.

A third potential problem has to do with the possibility that appears to be offered as a consequence of the existence of negative energy states for ordinary positive energy matter particles or even any preexisting negative energy matter particles to ‘fall’ into the allowed negative energy states in a continuous unstoppable process during which they would release radiation

and reach ever ‘lower’ energies. This is a difficulty which would also affect negative energy matter as it is traditionally conceived and which is effectively known as the vacuum decay problem. It would arise from the fact that the zero energy level would no longer constitute a minimum level of energy (the ground state) at which there could no longer be any transition to lower energies. Here we appear to have a situation where the existence of negative energy states raises the specter of allowing an arbitrarily large amount of work to be generated out of nearly nothing (by letting matter fall into the negative energy states and using the energy difference to produce work), as if energy conservation alone was not enough to restrict the evolution to negative energy states. This is clearly another issue of incompatibility with observation, because such decays are not observed to occur, even under the previously discussed conditions where negative energy densities are allowed to occur in a limited way by ordinary quantum field theory. In this context we may in effect ask what it is that prevents positive energy particles from falling into the lower negative energy levels predicted to exist under particular circumstances by quantum field theory? This is all by itself a legitimate question which has remained unanswered. Even from the viewpoint of the traditional interpretation of negative energy states this situation looks like a deep mystery.

But what is probably the most difficult problem which one must face upon recognizing the necessity of introducing a notion of negative energy matter obeying the requirements of a relational definition of physical quantities (which imply that opposite energy bodies must gravitationally repel one another) is that the existence of such matter may appear to allow violations of the principle of conservation of energy. This issue arises as a consequence of the fact that it seems possible for energy and momentum to be exchanged between positive and negative energy systems in a way that is similar to that by which positive energy systems exchange energy among themselves. Basically, it appears that the positive energy of a positive energy body can be turned into an equal amount of negative energy belonging to a negative energy body and vice versa when a ‘collision’ between two such opposite energy bodies would occur. For example, positive energy could be lost by a positive energy body colliding with a negative energy body initially at rest, while negative energy would be gained by the negative energy body with which the first body has interacted (or vice versa). This would give rise to a net variation in the total energy of the two bodies that would be equal to twice the individual change of energy (rather than allowing a cancellation

of changes, as is observed when two positive energy bodies collide). The solution of that problem will have to arise from a proper understanding of some remarkable consequences of the insights gained while solving the first category of problems discussed above.

A further difficulty could arise in the context where the inertial force on a negative mass body has the same direction as that which applies on a similarly accelerating positive mass body, despite the reversal of inertial mass which I have argued must occur when gravitational mass itself reverses. Indeed, from the viewpoint of an improved conception of the phenomenon of inertia based on a generalized formulation of Newton's second law it is no longer possible to consider that acceleration would take place in the direction opposite the applied force for a negative mass body and given that the equivalent gravitational field due to acceleration would be reversed for such an object it follows that the inertial force it would experience is identical to that which is experienced by a similar positive mass body. It would therefore appear that while the presence of a negative mass body could contribute to reduce the gravitational mass in a region of space in which positive mass matter is also present, it would still provide the same resistance to acceleration despite the fact that it would also provide a negative contribution to the inertial mass contained in this volume. This may not be a problem when we are dealing with independent physical systems with opposite masses but, as I previously mentioned, when a bound system is involved the energy contained in the field of interaction between its constituent particles would be opposite that of the system as a whole and in such a case it would seem that while the energy of the field should reduce the gravitational mass of the system it should nevertheless contribute to increase its resistance to acceleration. Given that bound systems with various force field configurations are quite common, it would seem that objects made of different materials should experience distinct accelerations when submitted to a gravitational force, but no such variations are observed. Some much needed clarification is required here if the concept of negative mass which I have proposed is to be considered viable from an observation viewpoint.

One last potential category of arguments which one might believe could disprove the validity of the idea of gravitationally repulsive, negative energy matter does not actually have to do with the concept of negative energy matter developed here, but merely with more traditional concepts of 'antigravity' and gravitational repulsion. The problems involved would be difficulties for a theory according to which ordinary antimatter is gravitationally repulsive.

They would also constitute a challenge for the traditionally favored description of negative energy matter according to which gravitational repulsion is an absolute property of negative energy itself, while gravitational attraction is an absolute property of positive energy matter (so that negative energy matter repels positive energy matter and is attracted to it). If such conceptions were to be retained as valid they would allow paradoxical situations such as perpetual motion and time travel to arise. Given that for most people those difficulties are associated with the general concept of negative energy it is important to explain why the issues involved here would not affect a more consistent theory of gravitationally repulsive, negative energy matter such as that which will emerge from the developments I introduced in the preceding sections.

We are then faced with six categories of problems which appear to undermine a conception of physical reality according to which matter would be allowed to occupy energy levels below zero. I have wrestled with the questions raised by those difficulties for a long time and on many occasions I had nearly given up on the possibility to ever be able to find appropriate answers that would perhaps explain why negative energy is not an inappropriate concept for physical theory. But gradually I came to understand that the problems really have to do with some incorrect implicit assumptions we make when considering the expected behavior of matter in a context where those negative energy states are effectively allowed. In the next six sections I will explain the nature of the insights required to appropriately deal with those severe problems.

2.8 The origin of repulsive gravitational forces

When as a young man I first started to contemplate the possibility that there could exist matter in a state of negative energy I soon realized that if such matter was to attract matter of the same type while it would repel ordinary matter and be repelled by it (as I had intuitively assumed should occur, ignorant of the dominant paradigm), then this matter would have to be dark, because nowhere was it mentioned that we observe gravitational repulsion arising from the presence of any planet, star or galaxy. While I was working on improving my understanding of physics in general and trying to develop a theory incorporating the concept of negative mass I simply assumed that negative mass particles were such that they would interact

with ordinary matter only through gravitation. I remember that I had read that Feynman once said that we must not question *why* things are the way they are, but simply try to describe in the most accurate way possible *how* they behave. Thus for a while I was comfortable with the idea that negative energy matter simply does not interact other than through the gravitational force with ordinary matter (although it could interact with itself through the whole spectrum of forces), even if I had no idea why that should be the case and had to assume that this is just the way things are. The only concern I had regarding this situation is that it appeared odd that negative energy matter should not interact with ordinary positive energy matter through the same interactions by which positive energy particles were interacting among themselves, given that negative energy matter could be assumed to actually be composed of the exact same particles as positive energy matter. But then came the shock.

I had for some time tried to figure out what determined the repulsive or attractive nature of an interaction which clearly depends on the signs of the charges of the interacting particles and had slowly come to realize that this property seemed to be related to the sign of energy of the field of interaction, not yet fully aware that it was actually rather the attractive or repulsive nature of an interaction (determined by the sign of the charges involved) that determined the sign of energy of the field and not the opposite. In any case I had understood that the energy of a field associated with a repulsive interaction between positive energy particles, for example the energy of the electromagnetic field between two electrons, is always positive, while the energy of a field associated with an attractive interaction between positive energy particles, for example the energy of the electromagnetic field between an electron and a positron, is always negative. But it also had to be the case (as I will explain below) that the energy of a field associated with a repulsive interaction between negative energy particles is always *negative*, while the energy of a field associated with an attractive interaction between negative energy particles is always *positive*. What that means is that when two negative action particles are attracted toward one another or bound together in a single system, the contribution of the attractive field mediating the interaction to the energy of the whole system should be positive, while for positive action particles it would be negative.

As I was trying to make sense of this observation in the context where the interaction involved would be that between a positive and a negative energy body I suddenly realized that a catastrophe had just happened. Indeed if

this relation between the sign of energy of the field and the attractive or repulsive nature of the related interaction was right in general it meant that if there was any gravitational interaction between positive and negative energy bodies it should be either repulsive for positive energy matter and attractive for negative energy matter (if the field was attributed positive energy) or repulsive for negative energy matter and attractive for positive energy matter (if the field was attributed negative energy), but never repulsive for both the positive and the negative energy bodies involved in the interaction. This is because a repulsive field would have to have positive energy for a positive energy matter particle, while this same positive energy field would have to exert an attractive force from the viewpoint of a negative energy matter particle for which the same relation would exist in general between the *difference* in the signs of energy of the matter particle and its field on the one hand and the repulsive or attractive nature of the associated interaction on the other (the problem is not restricted to gravitation). This is again a consequence of the requirement of relational definition of the physical properties associated with attraction and repulsion which cannot be considered to be determined by the energy sign of the interaction field only, but must be a consequence of the difference between the energy sign of the field and that of the matter particles submitted to the force associated with this field.

But it was just nonsense to conclude that an interaction could be both attractive and repulsive at the same time and it is even more so now, in the context where we must recognize that the hypothesis of the mutual gravitational repulsion between positive and negative energy matter is also required for a relational description of the gravitational interaction between those two types of objects. The conclusion I had to draw was thus very clear: no definite energy sign could be attributed to the fields of interaction between positive and negative energy particles (as must be the case for any interaction involving particles with the same sign of energy) and therefore there simply cannot be any interaction between those two types of particle, not even gravitational. This appeared to be a fatal blow, because if there are no interactions of any kind between positive and negative energy matter then how could negative energy matter have any relevance for the world we experience?

When I realized the existence of this difficulty for a theory of negative energy matter I had already come to appreciate the many advantages that there would be if such matter was allowed to exist (if it could indeed gravitationally interact with ordinary matter). This is because I had been able

to solve important problems using even the incomplete description I had by then managed to develop and it seemed improbable to me that the whole idea could simply be wrong. I know that this may look like it was more a hopeful wish than a rational conclusion, but in fact it was actually both hope and reason. Indeed, we had struggled with the problems I was able to solve for a very long time and there really appeared to be no viable alternative solutions to those problems, while theoretically the basic idea of negative energy had a lot of appeal. It is as a consequence of the fact that I had so much confidence in the validity of the basic concept of a symmetry between positive and negative energy states that I did not stop working on developing the idea when I encountered the difficulties discussed here. And as it turned out the problems encountered became just another challenge for reason and imagination on the way to a satisfactory solution of the problem of negative energy.

So, I went from having to explain why there would be no electromagnetic interactions between positive and negative action matter to having to explain why there can be any interaction at all between the same two kinds of matter. Of course I was glad that at least I now had an explanation for why there is indeed no electromagnetic or other non-gravitational interactions between opposite energy particles, because it was clear that on the basis of the above discussed observations it had to be recognized that there cannot be any direct quantized interactions (mediated through the exchange of interaction bosons) between such particles. But gravitation is different, because it is not yet described as a quantized field and I had hope that it might be its singular classical character that would allow the existence of *some kind* of interaction. It must be clear, however, that the problem described above is very real and unavoidable and its significance should not be underestimated as it effectively means that there can be no interactions and no exerted force *between positive and negative energy particles*. It must also be understood that this is not a hypothesis, as no consistent theory could describe the interaction of positive and negative mass particles and this must simply be taken as an indication that such interactions are effectively nonexistent.

At this stage you may remember that when I explained that there must be an equivalence (for a positive energy body) between the effects arising from the presence of a void in a uniform positive energy distribution and those which we may identify with a gravitational repulsion directed away from the void, I insisted that this repulsion was really the consequence of an uncompensated gravitational *attraction* directed away from the void. There-

fore, when dealing with matter distributions which are uniform on a cosmic scale we can observe gravitational repulsion to arise from what are actually purely attractive gravitational interactions. But I also insisted that negative energy matter would be equivalent from a classical gravitational viewpoint to the presence of missing positive energy from the vacuum, while the vacuum can itself be considered as being equivalent to some extent (only in this respect) with a uniform matter distribution. This of course means that the gravitational repulsion experienced by a positive energy body and which we would expect to arise from the presence of negative energy matter actually results from an uncompensated gravitational attraction attributable to the surrounding positive energy portion of the vacuum. In other words, we can explain the gravitational repulsion apparently exerted by negative energy matter as really consisting of a gravitational attraction involving only positive energy sources. Therefore, even if we assume an absence of direct interaction between positive and negative energy bodies we can nevertheless expect to obtain an equivalent repulsive gravitational force between these objects⁵. It is in this particular sense that the concept of gravitationally repulsive matter developed here can effectively be assumed to involve effects which are analogous to the situation we have in the case of voids in a uniform matter distribution. But under such circumstances the above discussed problem of the impossibility of direct interactions of either gravitational or non-gravitational kind between positive and negative energy particles is turned into an advantage, because it effectively forbids any interactions to occur between opposite energy particles except for the equivalent gravitational repulsion just described and this is precisely what we needed.

This result should be encouraging, as the category of problems it allows to solve was the most basic and the most serious of those which I identified above as facing a theory of negative energy matter. It effectively allows to explain why it is that we have never observed gravitationally repulsive matter, because indeed such matter, if it exists, should not be visible, as it would not interact with ordinary positive energy matter through the long range electromagnetic interaction. It also allows to explain why it is that the predictions of quantum field theory made under the hypothesis that negative energy states are not allowed in the formalism produce very accurate

⁵The same conclusion cannot be drawn in the case of the other interactions, because it is not opposite charge particles which cannot interact with one another, but really opposite energy or opposite action particles.

results which correspond to observations to a very high degree of precision. Because indeed if only the equivalent repulsive gravitational interaction just described exists as a kind of influence of negative energy matter on the processes involving positive energy particles which are described by quantum field theory, then given the weakness of the gravitational interaction there should only be a marginal impact from the existence of this negative energy matter on estimations of physical observables currently made under the assumption that negative energy particles do not exist. Indeed, if we do not need to take into account the effects of the attractive gravitational interaction between ordinary *positive energy* matter particles in such calculations, then we should certainly not expect to have to take into account any effects from the equivalent repulsive interaction with the very sparse amount of negative energy particles that could perhaps be found to wander around apparatuses located on Earth. Thus, if I am right, we would have here the solutions to two quite serious problems which were never addressed by any of the authors that previously discussed the possibility of gravitationally repulsive matter, because it can now be understood at once why gravitationally repulsive matter is dark and why it nevertheless interacts gravitationally with ordinary matter.

It must be noted however that even in the context where we have to assume an absence of direct interaction between positive and negative energy particles it would be wrong to assume that positive energy matter interacts only with the positive energy portion of the *vacuum* and not with the negative energy portion of it, because, as I explained in a preceding section, positive energy matter must itself be considered to consist of voids in the negative energy portion of the vacuum and as such certainly cannot be considered to behave independently from this negative energy vacuum. Yet it should be clear that we are not really dealing with an interaction between opposite energy particles here, but merely with the gravitational interaction of this negative energy portion of the vacuum with itself. Such a phenomenon is somewhat similar to the gravitational dynamics of a uniform negative energy matter distribution in which voids may also be present that would exert attractive gravitational forces on each other and repulsive forces on the rest of the negative energy matter. In such a case it is clear indeed that even if we could assimilate the voids with the presence of positive energy matter, their effects would actually be the outcome of the interaction of negative energy particles among themselves. We may therefore still consider that there is no *direct* interaction of any kind between positive and negative energy matter

or vacuum, but again that does not mean that positive energy matter does not experience the gravitational effects of the negative energy portion of the vacuum or that negative energy matter does not experience the gravitational effects of the positive energy portion of the vacuum, because if positive energy matter is a manifestation of negative vacuum energy it cannot be expected that this portion of the vacuum does not interact with itself and the same can be said of negative energy matter as a manifestation of positive vacuum energy. This conclusion will obviously have enormous consequences for a description of the cosmological effects of vacuum energy.

Finally, I may add that a further justification for the fact that we do not yet have strong evidence for the existence of negative energy matter is that, given that such matter is submitted to gravitational repulsion with ordinary matter and is also gravitationally attracted to itself, it should be expected to migrate away from concentrations of positive energy matter and to concentrate itself in regions of the universe where there is a lesser density of positive energy matter. It would therefore be difficult to observe anomalous gravitational effects which could arise from the presence of celestial objects composed of gravitationally repulsive matter in a region of the universe like ours, where positive energy matter can be assumed to be the dominant form of matter given its relatively large density. In fact, at this point, the lack of evidence for negative action matter has been so well justified that it appears that if we are to ever obtain direct confirmation for the existence of this matter it will be necessary to use alternative methods of investigation and to concentrate on the ability which may be offered in this context to predict features of the visible very large scale matter distribution with better accuracy than current models which neglect the effects of this invisible gravitationally repulsive matter distribution.

2.9 No energy out of nothing

Before we can conclude that there should effectively be no interference with current predictions made using quantum field theory from allowing the existence of negative energy particles in stable states we must first explain why it is that there should be no creation or annihilation processes involving pairs of opposite energy particles with opposite charges, as such a phenomenon could also disrupt current predictions. This is the second category of problems I previously identified as potentially affecting the viability of the negative

energy matter hypothesis. Given the plausibility of the hypothesis that negative energy particles should be very rare in our region of the universe it may seem that the problem of the annihilation of opposite energy particles does not constitute a decisive issue. But, as I previously mentioned, we cannot avoid having to face the related problem of the creation of pairs of opposite energy particles, because in such a case it would appear that no favorable initial conditions are required for the discussed processes to occur. Thus, an explanation must be provided for why matter is not, under appropriate conditions, being created out of the vacuum in massive amounts, despite the fact that the processes involved can occur without violating the principle of conservation of energy, because this prediction clearly disagrees with observations which indicate a complete absence of such processes, at least under ordinary circumstances.

We may perhaps suggest that given that the opposite energy particles emerging from a creation event in opposite directions would have their momenta both pointing in the same direction (because we must assume that a negative action particle would have momentum opposite the direction of its velocity) could prevent the creation of such pairs when we impose that momentum is to be conserved. But it does not seem that this would constitute a strong enough constraint under appropriate circumstances, because the pairs could be created without much momentum or through an input of momentum from the environment, as is the case for ordinary particle-antiparticle creation processes arising from the disintegration of a single boson. In an upcoming section of this report I will examine the question of momentum and energy conservation more specifically, but for now it suffices to mention that when all contributions are taken into account it becomes clear that it is not the requirement of momentum conservation which prevents pair creation processes involving particles with opposite energies from occurring.

The fact that the kind of creation (or annihilation) processes which would require no energy input (or output) could be described as processes in which a particle reverses its direction of propagation in time while retaining the sign of its energy, may suggest another explanation for why such events would be forbidden. Indeed, we may ask why it is that when a particle changes its direction of propagation in time in the course of all those particle-antiparticle annihilation processes which do occur under the right conditions, the energy is invariably reversed relative to the new direction of propagation in time (so that it appears to be unchanged from the forward time perspective)? Why must it be imposed that a reversal of the direction of propagation in

time be combined with such a reversal of energy which leaves the sign of action invariant and which must therefore be compensated by the emission of photons carrying away the energy? Could it be that it is a requirement of continuity of physical properties along the world-lines of elementary particles that prevents a positive action particle from turning into a negative action particle? Such a change would in effect involve the transformation of a particle experiencing the gravitational interaction in a given way into a particle experiencing it in a different way, but perhaps that a particle cannot change the way it gravitationally interacts with the rest of the universe on a continuous world-line.

I must acknowledge that I once contemplated the possibility that action sign changing reversals of the direction of propagation in time may be forbidden by a requirement of continuity of physical parameters along a particle's world-line. But I later came to understand that what a requirement of continuity imposes is merely an absence of interruption of the flow of the fundamental time direction parameter, which can be satisfied even when energy does not reverse upon a change of the direction of propagation in time of some particle. In the following chapter I will explain what constraint a condition of continuity would impose on the transformation of physical parameters and it will be clear that a reversal of the action is not forbidden by such a requirement. In any case, if the *charge* of a particle can change discontinuously (can reverse) from the forward time viewpoint when the particle reverses its direction of propagation in time in a continuous fashion (during a process perceived as an ordinary particle-antiparticle annihilation process), then there is no a priori reason why the *action* of a particle could not reverse in a similar manner when a particle reverses its direction of propagation in time, if the reversal in time also occurs in a continuous way, which would simply mean that the particle does not actually experience the usual reversal of its energy sign at the bifurcation point when it reverses its direction of propagation in time.

Actually, I believe that the simple fact that two opposite action particles of the same type must be considered to consist in the same particle which simply happens to be in a different energy state (or to propagate in a different direction of time) means that such particles should be allowed to transform into one another on a continuous world-line if their similarity is to ever be explainable in a causal way, but this is precisely what must occur only in rare circumstances. Must one then conclude that there exists an unexplainable decree simply banning negative action particles (carrying positive energy

backward in time) from existing? This would again be the easy way out: there is a difficulty so let's just forget about the whole thing. But if we recognize that the existence of particles carrying positive energies backward in time is theoretically inevitable, then a satisfactory explanation for the absence of spontaneous matter creation is required.

Before dealing with the problem of matter creation I would like to address the related issue of the annihilation of pairs of opposite energy particles whose solution turns out to be much simpler than one could perhaps imagine. To understand what imposes a limit on the annihilation of pairs of opposite action particles we simply need to take into account the results obtained in the preceding section. Indeed, one may ask how it is supposed to occur that a positive action particle with positive charge, say, could annihilate with a negative action particle with negative charge if positive and negative action particles are to be considered as equivalent to voids in opposite energy portions of the vacuum. How could the two particles ever come into contact with one another and annihilate when annihilation is to be considered a kind of interaction and there is absolutely no direct interaction of any kind between opposite action particles? Had I taken the lesson learned while solving the problem of the nature of repulsive gravitational interactions more seriously I would have understood much more readily that what limits the annihilation of particles with opposite energy signs is the absence of any direct interaction between such particles combined with the weakness of the indirect gravitational interaction they experience. Indeed, in the absence of any direct interactions between them, two opposite action particles cannot even come into contact with one another and therefore would not be able to annihilate one another. Even if they were to find themselves arbitrarily close to one another at a given time, two opposite action and opposite charge particles could not merge and combine their physical properties to perhaps produce a final state of no energy, because they do not even experience the presence of one another directly.

It is true though that opposite action particles would, according to the results I introduced in the preceding section, be subject to some indirect gravitational interaction as a consequence of the equivalence between the presence of a particle with a given energy sign and an absence of energy of opposite sign from the vacuum. But given that the gravitational interaction between two elementary particles is negligible under most circumstances, it must be concluded that the probability of observing the annihilation of opposite action particles is very low unless the energies involved are extremely high (of

the order of the Planck energy) or the spatial scale is so short that fluctuations in energy become a decisive factor. Thus, given that under ordinary circumstances opposite action particles are only subject to weak gravitational interactions (of an indirect nature) whose effects become visible only when large amounts of matter are involved (in which case the energy exchanges between individual particles are still negligible), we have to conclude that no annihilation of opposite energy sign particles back to the vacuum would occur at any observationally significant rate, even if negative energy matter was present in our region of the universe with a density comparable to that of positive energy matter. It should be the case, however, that in situations of very high energy density, like those encountered in the very first instants of the big bang, processes involving a gravitational interaction between elementary particles of opposite action signs would be likely to occur and could effectively give rise to the annihilation of pairs of particles carrying positive energies in opposite directions of time.

Given those conclusions one may perhaps be tempted to argue that the problem of the creation of pairs of opposite action particles out of nothing is also one that arises merely when we fail to recognize that there are no direct interactions between the particles forming such a pair. This argument is not valid, however, because in the case of matter creation we do not need the energy to be present beforehand and there is no a priori reason why the opposite energies of the particles which would be created could not be arbitrarily large, therefore allowing the process to occur through the indirect gravitational interaction that is allowed to take place between the two particles involved. It is true that, somewhat paradoxically, pairs of opposite action particles with lower energies would be more difficult to create in the absence of an additional constraint, because such pairs would be subject to weaker indirect gravitational interactions and therefore would be less likely to respond to local perturbations in the gravitational field. But the problem is precisely that there appears to be no limit to the amount of (positive and negative) energy which could be produced in the vacuum by such pair creation processes, so that it seems that particles with very high opposite energies should be continuously produced, even under normal conditions. What I have come to realize, however, is that despite the apparently inescapable nature of the problem of energy out of nothing, the required explanation for the observed absence of processes of matter creation is to be found in a quite familiar constraint.

I think that what imposes a limit on processes of creation out of nothing is

in effect simply the second law of thermodynamics which favors the spreading of energy into a larger number of lower energy particles, as a consequence of the fact that there are more microstates compatible with energy spread into less massive, lower energy particles and ultimately the vacuum. What happens is that given that the undesirable processes must involve the creation of very high (positive and negative) energy particles, the violation of the law of entropy growth involved in going from the vacuum to such states is considerable and could only occur as very short lived fluctuations which would necessarily be followed by the subsequent annihilation of the particles so produced with particles of opposite energy. In other words, the creation of matter out of the vacuum can only occur as a random fluctuation and like any fluctuation it is likely to be immediately followed by an opposite change, which in the present case would be an annihilation of matter particles without energy output. Thus it seems that under ordinary circumstances it is not possible for matter to be permanently created out of nothing, even when energy would be conserved in the process, because when matter particles are produced in such a way they usually annihilate back to the vacuum in a very short time, just as is the case for ordinary particle-antiparticle pairs in the absence of energy input. Only, in the present case the problem is not with energy conservation, but with the fact that the processes involved would allow a reduction of entropy, because a large amount of energy would become concentrated into a single pair of particles which was originally absent in the vacuum.

It is easier to understand what justifies this limitation in the context where we recognize that in the absence of any direct interaction between opposite energy particles energy must be conserved independently for positive and negative energy matter, so that the energy needed for the creation of a positive energy particle through such a pair creation process must actually come from variations of energy associated with the positive energy portion of the vacuum itself, while the energy required to create the negative energy particle would come from variations of energy occurring in the negative energy portion of the vacuum, as I will explain in a following section. But what must also be understood in this context is that there is no tendency in nature to spontaneously create particles with ever larger negative energies and masses that could compensate for the simultaneous creation of particles with large positive energies. Indeed, if the creation of a negative energy particle may allow to compensate for the energy change involved in the creation of a positive energy particle it cannot, however, compensate for the loss of

entropy that is also involved in the creation of this very high positive energy particle under conditions where it would emerge from a vacuum in which no matter or radiation is already present and which is therefore already in a state of maximum thermodynamic stability. I will explain in the next section (dealing with the problem of vacuum decay) what justifies the argument that the creation of matter with ever larger negative energies is not favored from a thermodynamic viewpoint and therefore cannot compensate for the thermodynamically unlikely creation of particles with large positive energies out of the vacuum.

Thus it seems that under normal conditions matter creation can effectively occur only when the energy is already present in some form or another initially, not because the creation of pairs of particles with opposite energies is totally impossible, but as a consequence of the fact that such processes can only occur as short lived fluctuations which we can expect to observe only at the correspondingly short scale of distance characteristic of quantum gravitational phenomena (given the magnitude of the energy variations involved). The related processes of matter annihilation to nothing on the other hand are allowed to occur as permanent outcomes, but they still take place only under conditions where the particles involved can influence one another through the indirect gravitational interaction they experience, which again only occurs when the energies involved are those characteristic of quantum gravitational phenomena. But despite the existence of a constraint limiting the creation of energy out of nothing under normal circumstances, it appears necessary to assume that during the big bang processes involving pairs of opposite action particles would allow matter to be permanently created as a consequence of the rapid expansion of space. This is because when the expansion is very fast, as it must have been in the very first moments of the big bang, two opposite action particles created as a pair can move away from one another rapidly enough that they may no longer be able to annihilate back to the vacuum (given that on the scale of quantum gravitational phenomena the distance would have become too large and the energy of the particles too low), which would mean that the creation process has become permanent. In fact, if matter cannot be considered to simply exist but must be created along with space and time at the big bang, then the occurrence of such processes of creation of pairs of opposite energy particles out of the vacuum would become an absolute requirement.

Actually, even if it was assumed that matter already existed prior to the big bang it seems that the processes of creation out of nothing which are

continuously occurring on a very short time scale would have to be allowed to become permanent under the conditions which existed in the very first moments of the universe's expansion given that such creation processes are required to balance the effects of the related annihilation processes involving opposite energy particles, which would also necessarily occur during the very first instants of the big bang. Indeed, processes of creation out of nothing are necessary if any preexisting matter is to be prevented from vanishing completely, leaving behind a vacuum devoid of any real particles. The conclusion that the existence of negative energy matter does not give rise to creation out of nothing under ordinary circumstances is certainly significant, but I believe that the conclusion that it is nevertheless possible for pairs of opposite action particles to be permanently created without energy input under the most extreme conditions is even more significant, particularly from a cosmological viewpoint.

It must be understood, however, that even when opposite action particles are involved, the processes of pair creation and annihilation would have to involve elementary particles with opposite directions of propagation in time, just as is the case for ordinary particle-antiparticle creation and annihilation processes. This is the true requirement of the condition of continuity which will be introduced in the following chapter and which can therefore be seen not to forbid all processes of the kind that would involve opposite action particle pairs, but merely those among such processes which would effectively involve an interruption of the direction of the flow of time along particle world-lines, as when two forward in time propagating particles with opposite energies would meet and vanish. In this context it is of importance to understand that backward in time propagating negative action particles must effectively be those with charges opposite (from the viewpoint of an observer measuring them in the forward direction of time) those of forward in time propagating positive action particles, if charge is to be conserved during the unlikely processes of creation and annihilation of opposite action particles. This will later be explained to be allowed by the necessary invariance of the sign of charge (relative to its true direction of propagation in time) under both action sign preserving and action sign reversing discrete symmetry operations.

2.10 The problem of vacuum decay

There is an unavoidable question that arises whenever one proposes that negative energy states may be physically allowed. What is it in effect that prevents particles from falling into those ‘lower’ energy states? It has been argued that positive energy matter particles may not be able to do so because they would first have to surmount the limit imposed by the irreducible value of their positive mass. But that would clearly not prevent particles already in a negative energy state from reaching even ‘lower’ energy states and given that I am here working under the assumption that negative energy matter can exist in stable form this would appear to be a serious problem. Under such conditions it would seem that if even a small amount of matter was to ever find itself in one of the available negative energy states this would give rise to a catastrophic process of creation of negative matter energy and positive radiation energy, because the matter would radiate energy in going from the ‘higher’ energy states (with negative values nearer to zero) to the allowed ‘lower’ energy states (with larger negative values) without ever reaching a minimum energy in which it could settle down. Thus, as I mentioned before, it would seem that if negative energy matter can exist we could produce an infinite amount of work by simply harvesting the positive energy radiation produced when negative energy particles fall into lower negative energy states. But given that quantum field theory already allows for states of negative energy to occur in limited portions of space it would seem that we have a very serious problem, even in the current theoretical context, because if negative energy can be made to exist under such conditions (which have already been produced in the laboratory) it should immediately collapse to even lower negative energies and in the process produce an arbitrarily large amount of positive energy radiation, while of course no such phenomenon has ever been observed.

The insights gained while studying the problem of matter creation discussed in the preceding section, however, provide the elements needed to tackle this additional difficulty from a different angle. We may recall in effect that according to the preceding discussion an important consequence of the absence of any direct interaction between opposite action particles is that it is effectively impossible, under ordinary circumstances, for a particle to annihilate with its opposite energy antiparticle counterpart, which is another way to say that an already existing particle cannot reverse its direction of propagation in time without also reversing its energy sign (rel-

ative to its new direction of propagation in time), therefore describing an ordinary particle-antiparticle annihilation process. But another perhaps less obvious consequence of the absence of any direct interactions between opposite action particles is that a negative energy particle cannot emit a real (by opposition to virtual) positive energy interaction boson regardless of what energy changes the original particle goes through, because the positive energy boson could not even have been into contact with the negative energy particle it is assumed to transform. Thus a negative energy particle could not gain negative energy at the expense of the production of a compensating amount of positive radiation energy and the same limitation also implies that a positive energy particle couldn't absorb negative energy radiation and diminish its own positive energy in the process. This constraint must apply even if such processes could occur without violating conservation laws when the energy change of the matter particle involved would be compensated by the emission of an opposite amount of radiation energy. But this means that even the emission of *positive* energy radiation by a positive energy matter particle could not occur in such a way that the positive energy particle could turn into a negative energy particle, given that this would imply that there would have been a direct interaction between the now negative energy matter particle and the positive energy radiation it would have released, while according to my analysis this must be considered impossible.

Thus, the same constraint which allowed me to conclude that a particle cannot change its direction of propagation in time without reversing its energy sign also implies that it is impossible for a particle to reverse its energy without reversing its direction of propagation in time (in which case the particle would not continue to exist with opposite energy in the future). The existence of such a limitation suggests that no interaction vertex involving particles with mixed action signs needs to be taken into account in determining the transition probabilities of quantum processes. This is a valid conclusion even if the merger of certain opposite action particle world-lines may be allowed under conditions where the gravitational field is very strong, as I explained in the preceding section, because such annihilation processes would not occur through the emission of gravitational radiation (especially since they need not release any energy at all) but merely as a consequence of the interaction of the two particles involved with their respective same-energy-sign vacuums. A certain limitation against the possibility of transitions to negative energy states therefore effectively exists, because a positive energy particle cannot 'fall' into a negative energy state by releasing pos-

itive energy radiation. The only reversal of energy which may occur on a continuous particle world-line would effectively have to involve a reversal of the direction of propagation in time, in which case the energy of the particle would no longer be negative relative to the forward direction of time and we would merely observe a conventional antiparticle in a positive energy state annihilating with the ‘original’ particle.

The limitation imposed on vertexes that they cannot involve particles with mixed action signs would therefore actually prevent a particle that is already in a negative energy state from falling into even ‘lower’ energy states by releasing positive energy radiation, because such negative energy matter could never have been in contact with the positive energy radiation it is assumed to emit. In fact, this explanation works both ways, as it is also true that a particle in a negative energy state could not ‘gain’ energy and turn into a positive energy particle by releasing a compensating amount of negative energy radiation, because the bosons so released could not have been emitted by the now positive energy particle with which they can have no contact. What must be understood, again, is that while the requirement of energy conservation may not alone forbid transitions involving a reversal of the sign of energy, the fact that those transitions would involve the emission or the absorption of radiation with an energy sign opposite that of the original particle effectively prevents them from occurring in the context where a negative energy particle (be it matter or radiation) can only interact with a positive energy particle through the very weak indirect gravitational interaction which exists by virtue of the fact that a negative energy particle can be described as a void in the positive energy portion of the vacuum.

Yet it must be remarked that the constraint described here would not prevent the vacuum itself from decaying by creating pairs of very high opposite energy particles, given that when the (positive and negative) energies are high enough indirect gravitational interactions are allowed to occur between opposite energy particles. In the previous section I mentioned that this problem occurs only when we fail to take into account the fact that processes during which negative energy would spontaneously increase (become more negative) are not thermodynamically favored. But given that we are dealing here with processes during which a particle would effectively ‘lose’ energy, it seems that an explanation is needed as to why it is exactly that such decays to ‘lower’ negative energy states are not likely to occur from a thermodynamic viewpoint. In this particular sense it would therefore appear that a certain aspect of the problem of vacuum decay remains unsolved.

I believe that the situation we have here is analogous to that which was faced upon the introduction of the Rutherford atom model, which was initially rejected despite its apparent empirical inevitability, because it was assumed that the electron in orbit around the nucleus should lose energy in the form of electromagnetic radiation and end up collapsing into the nucleus, while no such catastrophe was observed. But just like the Rutherford model it appears that negative energy states are unavoidable and thus a solution of the problem of vacuum decay that does not simply amount to reject the physical nature of those states must be provided. Based on the results achieved in the preceding sections I would like to suggest that the difficulties described here arise again from the fact that we ignore the requirements imposed by the necessary relational definition of physical quantities. Indeed, what is happening is that we are attributing a direction to energy variations without referring to a physical aspect from our universe relative to which that direction could be compared. In other words, we use an absolutely defined direction for the energy scale which we arbitrarily define as ‘lower’ and we attribute distinctive physical properties to this absolute direction of energy variations which actually has no objective significance. This traditional assumption seems to be justified by the fact that, for positive energy states at least, there effectively exists a singled out direction on the energy scale that is related to the natural tendency for matter to disintegrate and to reach thermal equilibrium. This direction can be associated with a well-defined physical aspect of our universe which is the direction of time in which entropy is growing. In the absence of such a relationship we would have no motive to assume the existence of a preferred direction on the energy scale distinct from that which is associated with growing absolute values of energy, which on the negative energy scale is opposite that on the positive energy scale.

However, when I examined what the motives are exactly that allow us to consider the existence of this objectively defined ‘lower’ direction on the positive energy scale, arising in relation to the direction of time in which entropy grows, I realized that there is absolutely no reason to assume that this direction on the energy scale can be extended into negative energy territory without being subjected to a reversal like energy itself. The only assumption necessary to assert the validity of this conclusion is that the initial large scale distribution of negative energy matter at the big bang was similar (regarding its smoothness in particular) to that of positive energy matter, which certainly constitutes a plausible hypothesis. Therefore, it seems that the objectively

defined ‘low’ energy direction on the positive energy scale cannot be extended into negative energy territory, but would actually be effective toward smaller, *less negative* states (toward the zero energy ground state) for negative energy matter.

Basically, what allows me to conclude that the low energy direction for negative energy matter is toward zero energy, as is the case for positive energy matter, is that the singled out, objectively defined direction on the energy scale is simply that relative to which the energy tends to dissociate itself and become less concentrated, so as to spread into a larger number of independent particles which thus necessarily have smaller (nearer to zero) energy as time goes. What explains this tendency is the fact that such a final configuration is associated with a larger number of microscopic degrees of freedom and a higher entropy (when gravitation can be neglected) and therefore is more likely to be reached in this direction of time in which entropy is effectively allowed to grow. But if the direction in time of entropy growth is the same for positive and negative energy systems (which would be unavoidable if the initial distribution of negative energy matter is similar to that of positive energy matter as I mentioned above) then the direction that would emerge as the low direction on the negative energy scale would have to be the opposite of that which constitute the equivalent objectively or relationally defined low direction on the positive energy scale, because the spreading of energy into a larger number of particles with smaller *negative* energies, which is necessarily associated with a higher entropy, occurs in the opposite direction on the energy scale to that in which smaller positive energies are reached. Thus what we traditionally called ‘low’ energies, far below the zero point of the vacuum, are in fact high energies for negative energy matter and what we called ‘higher’ energies, nearer to the zero point on the negative scale, are actually lower energies for negative energy matter. This is in perfect agreement with the previously discussed requirement to the effect that there should be a symmetry under exchange of positive and negative action matter, so that the sign of energy can be defined as a relational property.

Such a conclusion is significant, because it allows one to deduce that it is not to be expected that matter should have a tendency (arising from a thermodynamic necessity) to reach for negative energies past the zero energy level. Negative energy particles must be expected to have the same tendency as positive energy particles to reach energies which from the perspective of an observer made of such matter would be lower energies and therefore to reduce the absolute values of their energies and reach for the vacuum

ground state in the future direction of time. If a particle was found in a negative energy state it would not have a natural tendency to decay in a direction on the energy scale which is actually upward for a negative energy observer. It would be incorrect to assume that negative energy particles have a tendency to spontaneously *gain* even more negative energy as time goes, because such configurations are not thermodynamically favored, but are actually less likely to occur for the same reason that lower positive energy states are more likely to be reached. As a consequence, regardless of the energy level in which a physical system is to be found at a given time, energy can only be lost until the system reaches the energy contained in its rest mass and if it disintegrates and loses its mass it is not to be expected that it would continue to decay into negative energy territory. Thus the vacuum itself should not have a tendency to produce negative energy particles that could allow some thermodynamically unlikely processes of positive energy matter creation to become possible. This conclusion should have been expected even if only as a consequence of the fact that if we do not expect positive energy systems to spontaneously (without work being performed on them) reach for larger positive energies then we should also not expect negative energy systems to spontaneously reach for larger negative energies.

In the context where there already exists a constraint on the release of positive radiation energy by matter entering a negative energy state, it must be recognized that the unavoidable character of the conclusion that there is no preference for ‘lower’ more negative energy states means that there should be no continuous decay to such negative energy states. It must therefore be considered impossible to produce a large amount of work by making use of the processes by which particles would spontaneously decay to ever more negative energies, despite the assumption that matter is effectively allowed to occupy those negative energy states. I should finally mention that the fact that we observe no collapse to larger negative energies under the conditions where small negative energy densities are routinely produced in a limited way (as when a negative pressure is observed between two parallel mirrors in a vacuum) is a confirmation of the validity of the conclusions discussed in this section.

Thus, the outcome of the progress achieved in the last three sections is that it is possible to conceive of a fully consistent interpretation of negative energy states that would allow to at least preserve the validity of the current framework of quantum field theory. Indeed it would appear that what we obtain are two more or less independent frameworks describing the two more

or less independently evolving categories of systems with opposite energies, which interfere with one another only under those special conditions where it is possible for an observer of one energy sign to indirectly deduce the existence of opposite energy densities as they occur in the context where constraints are imposed which forbid the presence of certain states which would otherwise be present in that portion of the vacuum with the same sign of energy as that of the observer. This particularity allows the near perfect agreement between the predictions and the observations related to the small scale realm of quantum theory to naturally be maintained despite the fact that it is possible for matter to occupy the available negative energy states, which is also remarkable.

2.11 Energy and momentum conservation

I would now like to discuss the case of that most difficult of problems, which could have proved fatal to the alternative concept of negative energy developed here and which I have identified above as being that raised by the apparent possibility of a violation of the law of conservation of energy under conditions where interactions (even if merely of the indirect kind envisaged here) are allowed to occur between positive and negative energy matter. The nature of the issue can be illustrated through the use of a simple thought experiment. I briefly discussed in a previous section the problem that would arise in the case where a ‘collision’ would occur between a positive energy body and a negative energy body. I explained that such a collision would involve a loss or gain of positive energy by the positive energy body that would not be compensated, but instead be made worse by the associated gain or loss (respectively) of negative energy by the negative energy body. This is because instead of witnessing a loss of energy by one particle that would be gained by another, as when two particles with the same energy sign collide, we would here seem to have equal variations of energy, either both positive or both negative, depending on which particle accelerates and which decelerates as a result of the collision. For example, a negative action body could lose negative energy, while the positive action body it repels would gain positive energy, resulting in a net overall increase of energy twice as large as the individual changes. It would then seem that energy conservation is not possible under such circumstances.

The problem discussed here is also apparent when we consider the vari-

ations of momentum involved in such a process. Indeed, if action is to be assumed negative for a body propagating negative energy forward in time then it means that the sign of its momentum relative to its direction of propagation in *space* must be negative, that is, momentum must be opposite the direction of the motion for a negative energy particle (because action has the dimension of an energy multiplied by time or that of a momentum multiplied by a distance). In such a context it is easy to deduce that the variation of momentum occurring upon a collision between two opposite energy bodies would be twice as large as the absolute values of the changes in each particle's momentum rather than be zero as when two positive energy bodies collide. This is a problem that does not exist in the context of the traditional conception of negative energy matter according to which positive energy bodies attract negative energy bodies which repel them (if we assume that only gravitational forces exist between opposite energy bodies) and therefore the existence of such a difficulty could be used as an argument in favor of this traditional viewpoint despite the fact that it also raises other problems of its own, as I previously explained.

But given that we now understand that there are no direct interactions between opposite energy particles we have to recognize that the only way a collision between opposite energy bodies could occur would be through the indirect gravitational repulsion that would arise as a consequence of what are actually attractive gravitational forces attributable to a surrounding energy distribution, which are made to exist as a consequence of the equivalence between the presence of matter of one energy sign and an absence of energy of opposite sign in the vacuum. In this context it should in fact appear unlikely that there could occur violations of energy conservation arising from a collision between positive and negative energy bodies, if indeed there are no direct interactions between such objects. Mathematically at least, it certainly seems that a general relativistic theory of negative energy matter which would involve only gravitational interactions should not give rise to violations of the law of conservation of energy, given that energy conservation in such a context is actually a constraint concerning the exchange of energy between matter and the gravitational field. Thus if, as I am proposing, opposite energy bodies effectively interact only through the gravitational interaction then it means that from the viewpoint of a general relativistic description of those interactions any variation of the energy of matter would in effect come from a variation of the energy of the gravitational field. The absence of any direct non-gravitational interaction between positive and negative energy bodies

should effectively allow one to expect that it would be variations in the energy of the gravitational field that would balance the variations of energy occurring in the course of the interaction of such opposite energy bodies. The problem I initially had, however, is that I was not able to figure out how this could come about in the more intuitive context of a Newtonian description of such interactions and I am always suspicious of conclusions drawn solely on the basis of mathematical deductions, which often conceal totally inappropriate assumptions. So, where exactly does the positive energy go which is lost by a fast moving positive energy body colliding with a negative energy body initially at rest and where does the negative energy come from which is gained by the negative energy body that is accelerated during such a collision?

I was allowed to understand what is going on when a positive energy body interacts with a negative energy body only when I became aware of the possibility which is sometimes contemplated that on the cosmological scale the total energy consisting of that of matter combined with that of its associated gravitational field may be null. Indeed, if this hypothesis is right it would mean that the gravitational potential energy of matter might contribute a total amount of energy comparable in magnitude (but opposite in sign) with the total amount of energy contained in the matter which is the source of this gravitational field. But if this is right then it seems that to any parcel of positive energy matter may be associated a negative gravitational potential energy of equal magnitude arising from the gravitational interaction of this parcel of matter with the rest of the positive energy matter in the universe. What I would like to suggest is that in such a context it would be plausible to assume that the required compensation for the energy gained or lost by positive energy matter as a consequence of its indirect gravitational interaction with negative energy matter could arise from a variation in the negative gravitational potential energy associated with the variation of positive vacuum energy that is equivalent to this variation in the energy of negative energy matter.

Indeed, from my perspective, what is happening when a moving positive energy body indirectly communicates energy to a negative energy body is that while the positive energy body effectively loses positive energy, the consequent gain in negative energy by the negative energy body is equivalent to a decrease in the amount of positive energy from the vacuum. But associated with this positive energy was a negative gravitational potential energy arising as a consequence of the interaction of this vacuum energy with the rest of matter and energy in the universe and if the above suggestion is right then

this negative potential energy could be as large in magnitude as the positive vacuum energy which was present initially. Thus the loss of positive energy by the positive energy body would be compensated by the loss of negative gravitational potential energy (which is a positive change) consequent to the reduction in positive vacuum energy which is equivalent for a positive energy observer to the energy increase (toward more negative values) experienced by the negative energy body. A similar reasoning also allows to conclude that the gain in negative energy experienced by the negative energy body is itself balanced by the gain in positive gravitational potential energy which follows from the increase in negative vacuum energy which is equivalent for a negative energy observer to the loss of positive energy experienced by positive energy matter (because a lesser amount of positive matter energy means a smaller void in the negative energy portion of the vacuum).

What must be understood here is that the reduction in positive vacuum energy which is equivalent to the gain in negative matter energy is effectively a negative energy phenomenon and therefore does not have to be compensated by any change in positive matter energy or negative gravitational potential energy, which are positive energy phenomena (in the sense that they are associated with changes occurring in positive energy matter or in the gravitational field between positive energy particles and between positive energy particles and the positive energy portion of the vacuum). Similarly the gain in negative vacuum energy which is equivalent to the loss of positive matter energy is to be considered a positive energy phenomenon that need not be compensated by a variation in negative matter energy or positive gravitational potential energy, which are effectively negative energy phenomena (in the sense that they involve changes occurring in negative energy matter or in the gravitational field between negative energy particles and between negative energy particles and the negative energy portion of the vacuum). In any case if the above description is accurate then the energy that is lost or gained by a positive energy body as a result of its indirect gravitational interaction with a negative energy body could always be considered to be compensated by an opposite change in the gravitational potential energy associated with the variation of positive vacuum energy occurring as a consequence of the associated gain or loss of energy by the negative energy body.

I do recognize, of course, that under most circumstances the energy contained in the gravitational field associated with the interaction of a positive mass body with that portion of the surrounding vacuum with the same en-

ergy sign whose energy varies as a consequence of the equivalent variation of energy of the negative energy body with which the positive energy body interacts is much smaller than the energy change observed in the matter itself. But this does not mean that there is something wrong with the suggestion that the discussed variation in matter energy is compensated by some opposite change in gravitational potential energy, because the change in gravitational potential energy which I am referring to here has to do with the interaction of this same portion of vacuum energy with the *entire* matter and energy content of the universe. Yet the fact that under all circumstances only as much energy as is present in a field of interaction can actually be exchanged between the particles interacting through that force field means that the energies exchanged during the process of indirect gravitational interaction between a positive and a negative energy body are relatively small and thus it is plausible that they could be compensated by a variation in some measure of gravitational potential energy associated with the changes involved. It must be clear, however, that we are not dealing here with the gravitational potential energy that could be associated with a repulsive force field mediating an interaction between the positive and negative energy bodies themselves, which in fact cannot exist as I explained before, but merely with *independent* measures of gravitational potential energy associated with the interactions occurring between those portions of the vacuum affected by the changes involved and the rest of matter and energy in the universe.

Thus, what must be understood is that following any interaction between a positive energy body and a negative energy body there effectively occurs a variation in the total energy associated with positive and negative energy matter considered together, but this is only half of the equation, as to any such change there must be a related compensating change in the gravitational potential energies associated with the equivalent variations in the positive and negative portions of vacuum energy. In the case of an interaction during which velocity is lost by a positive energy body and gained by a negative energy body, positive energy could actually be considered to flow from positive kinetic energy to positive gravitational potential energy, while negative energy flows from negative gravitational potential energy to negative kinetic energy. But it must be clear that this is only a reflection of the compensating opposite energy changes occurring in positive energy matter and its associated gravitational field on the one hand and in negative energy matter and its associated gravitational field on the other, because there is no actual exchange of energy between those two kinds of matter. It must also

be mentioned that the variation in the momentum of matter which would be observed during such an indirect interaction is also compensated by the opposite variation in the momentum associated with the gravitational fields which occurs as a consequence of the changes in vacuum energy which are equivalent to the changes in the energy of matter. The fact that the gravitational interaction is very weak means that this energy flow between matter and gravitational fields is relatively small, but it nevertheless exists and it appears to be what allows energy to be conserved during such interaction processes.

2.12 Absolute inertial mass

One last objection which could be raised against the interpretation of negative energy states which I proposed has to do with the fact that from my viewpoint negative energy matter would offer the same resistance to acceleration as would positive energy matter. This would traditionally be described as being a consequence of the alternative assumption that inertial mass is positive even for negative energy matter otherwise characterized by a negative gravitational mass. Of course, as I already explained, the inertial mass must be considered to actually be reversed along with the gravitational mass from the viewpoint of a consistent description of the gravitational dynamics of negative energy matter. But in the context of the previously discussed improved conception of the phenomenon of inertia that emerged from a generalization of Newton's second law it was shown that acceleration would not occur in the direction opposite the applied force for a negative mass body. In fact, once it is recognized that the equivalent gravitational field experienced by such an object must be opposite that experienced by a positive mass body, it is necessary to conclude that negative mass matter would effectively experience the same resistance to acceleration as positive mass matter when submitted to the same forces, despite the reversal of its inertial mass. Thus negative mass or negative energy matter would appear to violate the principle of equivalence as it is traditionally conceived.

Now, there could be situations where the gravitational mass in a volume of space would be relatively small or even zero despite the presence of a potentially large amount of matter in this volume, as when two opposite mass bodies are present at once in the same location (which would be allowed in the absence of strong interactions between them). Such configurations

would not be equivalent from an inertial viewpoint to the case of a system with nearly vanishing total mass, because the matter that is present would be more difficult to accelerate than if it actually had such a small mass. To better describe such vanishing energy configurations, which are clearly different physically from the vacuum, we may define a measure of inertial mass that would be related to the physically significant properties with which it is traditionally associated and that would correspond to the true amount of matter present under such circumstances, independently from the total mass which may partially or totally cancel out. The *absolute inertial mass* obtained by adding the absolute values of the masses of all material bodies present in some volume of space (or by adding all masses as negative from the viewpoint of a negative mass observer) would constitute such a measure of the true amount of matter present.

It is clear that the motion of negative energy matter in a gravitational field attributable to a local matter inhomogeneity (such as the gravitational field which exists on the surface of the Earth) would not be that which is shared by all objects made of positive energy matter. Yet experiments provide very strong constraints on the degree of violation of the equivalence principle and to date there is in fact no evidence at all that any such violations have ever occurred when systems of various different compositions are utilized. However, I did say in a previous section that negative energy was as common as bound systems of particles such as atomic nuclei and molecules, due to the negative energy of their attractive force field. Why then do we never observe an altered level of resistance to gravitational acceleration? We may for example consider atomic nuclei formed of many protons and neutrons bound together by the strong nuclear interaction, with various measures of negative energy of the force field associated with various configurations involving a variable number of component particles. It would then appear that the gravitational acceleration of such bound systems should be reduced by the negative value of the energy of the field while the inertial resistance would be proportionately larger, as the absolute inertial masses attributable to both the component particles and the force field would not cancel out like the gravitational masses. If we measured the acceleration of a whole body composed of one such type of nucleus on the surface of the Earth and compared it with the acceleration of another body made of another kind of nucleus containing a lesser proportion of such negative energy, we may then perhaps expect to discern a difference. But it appears that this is precisely what the experiments discussed above rule out to a very good degree of precision. Shall

we then once again abandon everything and conclude that negative energy, even though it is definitely present in bound systems, must be described in a non-relational manner (so that the phenomenological aspects associated with inertial mass cancel out like those associated with gravitational mass)?

It must be understood that in fact this conclusion would constitute a theoretical problem as grave as apparently is the empirical difficulty revealed by the absence of differences in the acceleration of various bound systems. Can we indeed ever hope to solve a problem by creating a ‘new’ one and assume that despite all indications to the contrary the latter difficulty is not real, simply because it only affects consistency on a more general level? This is not the path I chose to follow, because I realized that despite what is often suggested there is simply no reason to expect the kind of violations of the principle of equivalence which are described here, even if inertial forces do not cancel out when we consider two masses with opposite signs. What is wrong, I believe, with traditional assumptions is that when we are considering a bound system and its force field we assume that we have two masses with opposite signs, while what we really have is one single mass with one overall magnitude and one polarity, both from the viewpoint of inertia and from that of the response to local gravitational fields. Indeed, what motive would we have for considering that there could be independent contributions to the mass of a bound system (inertial or otherwise) when in fact the energy of the subsystems forming it (in particular the particles mediating the attractive force fields) could not be measured independently, given that they may be implicated in virtual processes which do not even have classically well-defined physical properties?

It is a fact that the particles mediating an interaction are virtual and as such exist merely by virtue of quantum uncertainty, which allows them to carry energy but only for a time that is short enough that this energy cannot be determined. The virtual particles involved in giving rise to interactions must then be considered unobservable, if only because to effectively establish their presence in any one particular instance would require a time length greater than the duration of the exchange process. But under such circumstances how could we be talking about an *independent* contribution of those particles to the energy or the mass of the bound systems in which they materialize? I think that this would effectively be non-sense and that it must be recognized that any component of an entangled system whose physical properties cannot be directly and independently observed does not contribute independently to any of the properties associated with the mass of

the entangled system as a whole, when those are effectively measured. Failure to understand this very decisive requirement would mean that we again allow one more inconsistency to obscure our conception of negative energy in a way that could only be made acceptable by rejecting one or another of the fundamental constraints identified above. In the present context this could not even be avoided by assuming that negative energy does not exist at all, because the issue is no longer merely about deciding if negative energy exists, but about determining its properties in a context where we must definitely accept that it is occurring.

There is no contradiction here, because there is definitely a negative contribution to the energy of bound systems, only this energy contribution cannot be independently measured in any specific case and this is the crucial distinction we must take into account when estimating the absolute inertial mass of such a system. Thus the difference between the situation described above of the two superposed opposite mass objects with large absolute inertial masses and that of a composite system with absolute inertial mass smaller than that of its constituent particles is that in the former case we are effectively dealing with two independent systems which may be interacting only negligibly with one another, while in the latter case we have a single entangled system which is physically different from the sum of its parts and to which must therefore be associated one single combined measure of mass, gravitational and inertial. In any case the fact that we do not observe violations of the principle of equivalence for bound systems whose observable total energy is positive confirms that this conclusion is appropriate.

2.13 A few other misconceptions

Before finishing this discussion concerning the potential problems facing a theory of negative energy matter I would like to provide arguments to the effect that a few other problems which are often associated with the possibility that there could exist gravitationally repulsive matter are actually of no concern, because they are significant only in the context of the traditional interpretations of negative energy and gravitational repulsion. It is nevertheless important for me to discuss those issues, because I have come to realize that the perception of negative energy as being associated with all sorts of strange phenomena that defy common sense is responsible more than anything else for making the perfectly acceptable idea of negative energy matter

look like a pseudo-scientific concept without any relevance to physical reality. I will thus try to make clear that what is wrong is not the hypothesis of matter in a negative energy state, but merely the current conception regarding what would be the properties of such matter.

One of the problems I would like to discuss arose as an outcome of the first attempts at finding an interpretation for the negative energy states which were predicted to occur by relativistic quantum theories. Indeed, when the existence of antimatter was experimentally confirmed it was suggested that this kind of matter may perhaps actually gives rise to ‘antigravity’, in the sense that antimatter would experience repulsive gravitational forces in the presence of ordinary matter. But only theoretical arguments could be given to disprove this possibility when it was first suggested, because no experiment had already been performed to demonstrate that antimatter would not fall upward in the gravitational field of the Earth. One of those arguments was based on the recognition that if antimatter was to repel or be repelled by ordinary matter this would allow perpetual motion machines to be build that would extract more energy from a process than was initially available. Indeed, under such circumstances it would take no energy to slowly raise a particle-antiparticle pair in the gravitational field of our planet (because there would be as much gravitational repulsion as attraction). But when this would be accomplished the pair could be made to annihilate and the positive energy of the photons so produced could fall back to a detector on the ground where they would be measured as carrying more energy than the pair initially had (this would be allowed in the context where the energy of the gravitationally repelled antiparticle is assumed to be positive relative to the forward direction of time) as a consequence of the frequency increase to which the positive energy photons would be submitted on their way down. It would then seem that energy can be freely produced if antimatter ‘falls’ up.

I think that this argument is perfectly valid, only it cannot be used to justify the rejection of anomalous gravitational interactions in general, but rather simply means that given that antimatter does not have negative energy (as observed in the forward direction of time) then it should not be expected to be submitted to anomalous gravitational forces. Now, could the same experiment be performed with negative energy (actually negative action) antimatter and then what would it mean for energy conservation? The answer to that question is to be found in the developments achieved by solving the problems discussed in the previous sections. First of all, it must

be understood that given that there are no interactions between positive and negative energy matter other than the indirect repulsive gravitational interaction which I have already described, it seems that it would be much more difficult to raise a pair of opposite energy particles together in the gravitational field of a planet without doing work on at least one of them. Yet this may not constitute an insurmountable difficulty, because it is possible to imagine arrangements which would allow a negative energy body to achieve the task of raising a positive energy body in the gravitational field of a positive energy planet by making use of the indirect repulsive gravitational forces existing between the two bodies (which could also be composed of matter with opposite charges). But in fact the same limitation concerning the absence of any direct interaction between opposite energy particles would also imply that it is not possible to make such a pair to annihilate under normal circumstances, although again it is possible to imagine that the appropriate conditions to achieve this (a very high energy collision between two opposite energy particles) could perhaps be met when the appropriate technology would become available. However, other means would probably exist for harvesting the energy contained in each particle (or in each of the two bodies) so that this limitation does not really constitute a decisive constraint that would allow to rule out the kind of processes discussed here.

The real difficulty for any incipient free energy harvesters would actually arise from the fact that in the context of a concept of gravitationally repulsive matter such as the one I proposed, even if a pair of opposite energy and opposite charge bodies could be raised together in the gravitational field of our planet without applying any external force on them, when the two bodies would annihilate they would release no energy at all. Indeed, if the objects have equal but opposite energies initially, they would not gain or lose any kinetic energy as a result of their ascension and this means that their respective final energies would still be equal in magnitude. As a consequence, even if their component particles could annihilate, no energy would be released, so that there would be no photons to fall back toward the surface of the planet with a net gain of energy. Of course we could arrange things so that the positive energy particles annihilate with other positive energy anti-particles already in place at the destination point, while the negative energy antiparticles would annihilate with negative energy particles already in place. But if the positive energy photons produced by the annihilation of the positive energy particles could effectively gain positive energy while falling back to a detector on the ground, the negative energy photons for their part would

lose negative energy while reaching the same detector and would therefore end up with less negative energy than they would have had if the negative energy matter had been submitted to annihilation before rising to a higher altitude. Thus, while positive matter (and radiation) energy is gained during such a process, negative matter (and radiation) energy is lost and this means that no useful energy can be produced in such a way.

In order to better understand the significance of the changes involved we can consider the variations occurring in the potential energy of the two bodies as they are raised in the gravitational field of the planet. From this more general perspective what would be observed in effect is that any potential energy that would be gained by one of the two bodies (the one that was actually lifted by the other) would necessarily be lost by the other body, thereby preventing any useful energy to be produced as a result of such a process. Indeed, while the positive energy body would gain positive potential energy (due to a loss of negative gravitational potential energy) the negative energy body would lose negative potential energy (due to an equivalent gain of positive gravitational potential energy). Now, this may seem to imply that a forbidden net increase of (positive) energy can be obtained despite the fact that no work would have been done to take the system to its final state. Yet, as I have explained in a preceding section this variation is not significant, because any change in the energy of matter resulting from an interaction between positive and negative energy bodies is compensated by an opposite change in the energy of the gravitational fields associated with the equivalent variations in the positive and negative portions of vacuum energy.

What must be understood here is that even if there may occur changes in the energy of matter this would not mean that we have gained the ability to perform more work, as would be required to produce perpetual motion, because what the loss of negative potential energy by the negative energy body means is precisely that there was a loss of useful energy (energy that could be used to do work) by that object during the process by which it would have performed work to raise the positive energy body and increase the ability of this positive energy body to perform work. In other words, despite the net energy gain for the pair as a whole, the ability to do work would not have increased, because the negative energy body having been raised by the repulsive gravitational field it experiences would now have a decreased potential to perform work (even though its kinetic energy would remain unchanged), which is precisely what its loss of *negative* potential energy implies, because indeed the object would have lost energy of the

same sign as its own and therefore would actually end up with less energy available to perform work after the process has occurred. The gain in useful energy by the positive energy body would actually have been provided by the negative energy body which would have lost its own useful energy and in fact, if the usual friction and other degradation of energy had been taken into consideration, it should be observed that the positive energy body would have gained less useful energy than the negative energy body would have lost. The fact that positive energy seems to have been created on the other hand is a simple consequence of the fact that the process discussed involves an indirect gravitational interaction between the two bodies and between the negative energy body and the positive energy planet during which the total energy of matter may effectively vary, as I remarked above, given that it is compensated by an opposite variation of the gravitational potential energy associated with the equivalent changes occurring in the energy of the vacuum. No additional difficulty is involved here and therefore it seems that the perpetual motion argument against gravitational repulsion cannot be considered significant other than as an argument against the possibility of an anomalous gravitational interaction between ordinary matter and ordinary antimatter.

A more exotic and hypothetical phenomenon which according to certain accounts could have interesting practical applications, but which would raise serious problems from a theoretical viewpoint, given that it may provide the means of achieving faster than light space travel and therefore also time travel, is that of wormholes. It is often thought that wormholes would naturally occur in the presence of some types of black hole singularities and may allow remote regions of space to be directly connected in some way, so that traveling through such wormholes would enable to bypass the limitations associated with the passage of time experienced under normal circumstances when traveling over such long distances at slower than light velocity. It is not clear exactly what regions of space could be connected in such a way or if we are effectively talking about connecting regions of our own universe, but if we leave aside those uncertainties then it would seem that all that is required for unlocking the potential of faster than light space travel is the existence of traversable versions of such hypothetical shortcuts through space and time. What is required therefore is a means to maintain the ‘throat’ of a wormhole open for a long enough period of time that space travelers can safely traverse it despite the tendency for the matter configurations involved here to collapse under the effect of the gravitational attraction exerted by

the singularity. The idea is that gravitationally repulsive negative energy matter (often called exotic matter) may allow to achieve that goal given that it could be used to exert a gravitational repulsion that would compensate the attraction exerted by the spacetime singularity at the center of the black hole. But again, when we look at the details of such proposals, it becomes clear that the conditions necessary for achieving the desired results are incompatible with a consistent notion of negative energy matter. That may not be good news for science fiction lovers, but if I am right negative energy matter could never be used to achieve such a goal.

To help identify what's wrong with current expectations I would suggest that we ask how it is exactly that negative energy matter could be brought not just inside some black hole, but toward the point of maximum density of positive energy matter (the singularity), despite the enormous gravitational repulsion that this positive energy matter would exert on the exotic matter? It should be clear that it is merely because we traditionally assume that negative energy matter would be attracted by a positive energy black hole and its singularity, even while it would repel it, that this appears to constitute an achievable goal. But the truth is that any negative energy matter approaching a large concentration of positive energy matter such as an ordinary black hole would be submitted to repulsive forces as large as those maintaining positive energy matter trapped inside the same black hole. In this context the only way by which negative energy matter could find itself inside the event horizon of a positive energy black hole would be by having already been present inside the region destined to collapse into that positive mass black hole before it formed. But even if that was to happen there is no way that the negative energy matter could be made to remain near the black hole singularity where repulsive forces would be the largest. This situation is simply unstable and given that stability is precisely what is required for a traversable wormhole to exist, we must recognize that negative energy matter could not provide the necessary element for allowing spacetime singularities to be used for faster than light space travel and time travel. The possibility that the kind of phenomenon discussed here could effectively have been used for achieving theoretically problematic, causality violating processes may seem far-fetched, but I think that it is nevertheless important to show that even under such extreme conditions there is no reason to expect that the hypothesis of the existence of negative energy matter could facilitate the occurrence of such self-contradictory phenomena.

The same argument I have used to rule out the possibility of engineering

traversable wormholes can also be utilized to solve a more down-to-earth problem that is not often discussed, but which would contradict one of the most unavoidable constraint applying to the evolution of physical systems with a large number of degrees of freedom such as black holes. The problem is that negative energy matter, as it is traditionally conceived, could be used to reduce the mass of a black hole and therefore also the area of its event horizon. This could be achieved by simply throwing negative energy matter into a black hole, which would effectively absorb it given that negative energy matter is usually assumed to be gravitationally attracted by a positive energy black hole. This would be possible even if negative energy matter repels a positive mass black hole, because we could throw negative energy particles in small amounts and their gravitational fields would be too small to resist the much larger gravitational attraction of the black hole. But the surface area of a black hole has been shown to constitute a measure of the entropy of such an object, so that reducing the area of the black hole is similar to reducing its entropy. Again, however, if we reject the traditional conception of negative energy matter the problem does not exist, because a negative energy particle cannot even get near a black hole without experiencing extreme gravitational repulsion, so that it certainly cannot be absorbed by the object, as would be necessary for reducing its mass and the area of its event horizon. If negative energy states are to be considered a true possibility then the fact that the traditional concept of negative energy matter would allow such violations of the second law of thermodynamics, while the alternative approach proposed in this report would forbid them, constitutes a strong indication to the effect that this latter description is more appropriate.

In fact, we are dealing with a much more general problem in this case, because from the traditional viewpoint it is actually assumed that when negative energy radiation would come into contact with positive energy matter (not necessarily a black hole) it could be used to withdraw positive thermal energy from this matter (as if it was providing negative heat), therefore again raising the possibility of allowing entropy to decrease as a consequence of the existence of negative energy matter. Of course given that from my viewpoint negative energy radiation cannot even come into contact with positive energy matter, the possibility raised here appears to be mostly irrelevant from a practical viewpoint. We may nevertheless examine the situation which would arise following an exchange of energy between positive and negative energy systems occurring as a consequence of the indirect repulsive gravitational forces they exert on one another. The conclusion we must draw in

such a case is that negative energy is not equivalent to negative heat for a positive energy system. Indeed, according to my conception of negative energy matter, kinetic energy is exchanged between opposite energy particles as if it was a positive definite quantity, which is allowed by the fact that the energy of matter is not conserved independently from certain contributions to gravitational potential energy associated with variations in the energy of the vacuum, as I explained before. But the fact that only the absolute value of the kinetic energy of matter is conserved means that thermal energy itself can only be exchanged as a positive definite quantity (or equivalently as a negative definite quantity from the viewpoint of negative energy observers) between opposite energy systems. Thus when heat is provided by a negative energy system it can only raise the temperature of a positive energy system (as if positive thermal energy was provided) and the same is true for the heat provided by a positive energy system to a negative energy system which can only raise the temperature of the negative energy system (as if negative thermal energy was provided by the positive energy system).

Thus, we have no reason to expect that even the indirect gravitational interactions between opposite energy systems could be used to transform useless forms of energy into more useful forms and in such a way reduce entropy. Negative energy cannot reduce the temperature of a positive energy system any more than positive energy could diminish (into positive territory) the thermal energy of a negative energy system, except under conditions where the temperature of one or another of two opposite energy systems (considered as a positive definite quantity) is larger than that of the other system, in which case it is necessarily the higher temperature system, regardless of its energy sign, that would lose heat (considered as a positive definite quantity) and thereby raise the temperature of the other system by an amount proportional to that which is lost by the cooled system, as when only positive energy systems are involved. What must be understood is that adding thermal energy from a negative energy source to a positive energy system is not equivalent to removing positive energy from the same system. In fact, it turns out that adding energy from a negative energy system to a gas of positive energy matter can actually raise its temperature (unlike most people considering the possibility of negative energy matter would assume) instead of decreasing it. This is all a consequence of the fact that negative kinetic energy can be turned into positive kinetic energy and vice versa, even when energy is assumed to be conserved, as I mentioned above. Thus the positive thermal energy of a gas of positive energy matter can actually be raised

through contact with a gas of negative energy matter at a higher temperature (the temperature that would be measured by a negative energy observer) because thermal energy is a measure of the average kinetic energy of such a gas and this energy would become more evenly distributed (independently from energy sign) between the two gases if they could be put into contact through the indirect gravitational interaction⁶.

Therefore, once again, the traditional expectation can be seen to arise from a misconception. You should take note, however, that I am not just trying to debunk myths here. The opposite conclusion, that a low temperature gas would be cooled even further upon contact with negative energy matter or radiation, regardless of the actual temperatures of the interacting systems, and the above discussed assumption that a black hole's mass could be reduced by absorbing negative energy matter, would constitute serious problems for a theory of negative energy matter. There are very strong motives behind my desire to demonstrate that the possibility of such entropy decreasing processes can be rejected and they are actually related to those which one might raise against the above discussed possibility of causality violating processes.

2.14 An axiomatic formulation

Before I complete the process of integration of negative energy to classical gravitation theory I would like to provide formal statements of each of the significant rules I have derived in relation with this issue and which were discussed in the previous sections of the current chapter. Basically there are ten fundamentally decisive results which clarify the situation regarding the nature and the behavior of negative energy matter itself as well as the behavior of positive energy matter in the presence of negative energy matter. Those results actually provide the axioms or the rules on which a generalized classical theory of gravitation can be based. The axioms are legitimized by the fact that they have been shown to be necessary on the basis of both

⁶In this context it would appear that despite the fact that temperature is proportional to energy, it must be considered a positive definite quantity under all situations where we are not dealing with the thermodynamics of the gravitational field itself, as would become necessary when black holes are involved and the gravitational field itself is a measure of temperature. I will address the implications of attributing a negative temperature to negative energy matter configurations in the presence of strong gravitational fields in a latter portion of this report dealing with black hole entropy.

theoretical consistency and agreement with experimental facts and thus we may effectively refer to them as principles. The first principle is the most fundamental and a recognition of its validity opens the way for a derivation of all the other results. The formal statement of this principle goes like this:

Principle 1: The distinction between a positive energy particle and a negative energy particle (propagating negative energy forward in time) can only be defined by referring to the difference or the identity of the energy sign of one particle in comparison with that of another, so that the sign of energy or mass has no absolute meaning.

From a gravitational viewpoint this principle is satisfied when positive energy particles are submitted to mutual gravitational attraction among themselves (as we observe), while negative energy particles (actually negative action particles) also attract one another gravitationally and positive and negative energy particles repel one another as a consequence of the indirect gravitational interaction which actually originates from an uncompensated gravitational attraction between matter of one energy sign and that portion of vacuum energy with the same energy sign. Compliance with this rule means that for a positive energy particle a negative energy particle should be physically equivalent to what a positive energy particle is for a negative energy particle. This property will be decisive for deriving the observer dependent generalized gravitational field equations that will be introduced later.

Another rule applies only in the classical Newtonian context where mass is a significant concept, but given that it allows to derive the rules which must also be obeyed in a general relativistic context it is necessary to mention it as a basic result. It simply amounts to recognize that:

Principle 2: When mass is reversed from its conventional positive value both gravitational mass and inertial mass are reversed and together become negative.

This is actually equivalent to assume that there is indeed only one physical property to which we may refer to as being that of mass and that there cannot be any arbitrary distinction between gravitational and inertial mass.

While principles 1 and 2 are for the most part theoretically motivated the next principle is both theoretically and observationally motivated. Indeed, principle 3 arose as the unavoidable consequence of an analysis of the

relationship between the attractive or repulsive nature of a field of interaction and the sign of the energy classically contained in this field, but it is also a necessary requirement of the fact that we do not observe any negative energy matter despite the fact that the existence of such matter appears to be allowed from a theoretical viewpoint. The third principle therefore is the following requirement:

Principle 3: There are no direct interactions of any type (either gravitational, electromagnetic or any other), mediated by the exchange of bosons of interaction, between positive and negative action particles (respectively propagating positive and negative energies forward in time).

Compliance with this principle means that negative energy observers would also be prevented from directly observing positive energy matter.

Another important result was discussed at length in a previous section of this chapter where its validity was shown to be unavoidable despite the fact that it appears to contradict some assumptions which are usually considered to be irrefutable. This result simply states that:

Principle 4: A void of limited size that develops in an otherwise uniform matter or energy distribution gives rises to uncompensated gravitational forces which are the opposite of those which would otherwise be produced by the matter or energy that is missing.

The effect it describes is the consequence of an alteration (caused by the presence of some local void) of the equilibrium of gravitational forces applying on any particle and due to its interaction with all the other particles in the universe (with which this particle effectively interacts). The importance of this principle becomes clear when we consider its significance in the context where the uniform energy distribution is actually the distribution of vacuum energy and it is recognized that principle 5 below applies.

The following principle is probably the most decisive after principle 1 given that it is the result that allows the whole concept of negative energy matter to have a significance despite the validity of principle 3 and the absence of direct interactions between positive and negative energy particles. It states that:

Principle 5: Locally, the presence of negative energy matter is equivalent to the absence of an equal amount of positive energy from the vacuum, while the presence of positive energy matter is equivalent to the absence of an equal amount of negative energy from the vacuum.

As explained in a preceding section those equivalences constitute the particularity that allows opposite energy bodies to exert gravitational forces on one another despite the absence of direct interactions between them, simply because according to principle 4 voids in a uniform positive energy distribution do have an indirect influence on positive energy matter despite the fact that those voids are actually equivalent to the presence of negative energy matter with which positive energy matter does not directly interact.

But even in the context where we assume the existence of a symmetry between positive and negative energy matter principle 5 would require that it is in fact only the inhomogeneities (either overdensities or underdensities) present in the negative energy matter distribution which can affect the gravitational dynamics of positive energy matter, while it is only the inhomogeneities present in the positive energy matter distribution which can affect negative energy matter. This is because, as previously discussed, the void in the positive energy vacuum that is equivalent to a totally homogeneous distribution of negative energy matter would leave no surrounding positive vacuum energy to produce an uncompensated gravitational attraction that would be equivalent (according to principle 4) to the gravitational repulsion otherwise attributable to the negative energy matter and the same is true concerning a homogeneous distribution of positive energy from the viewpoint of negative energy matter. An additional principle thus emerges that expresses this limitation applying on principle 5. It amounts to assume that:

Principle 6: Only (positive and negative) density variations in an overall homogeneous cosmic scale distribution of negative energy matter can be assumed to exert gravitational forces on positive energy matter.

Of course a similar limitation would also apply which would actually express the absence of gravitational forces on negative energy matter from a totally smooth and uniform cosmic scale distribution of positive energy matter.

A further particularity could be derived from the already stated principles but I will provide it as an additional specific rule because it may not

be obvious that it applies in the context where principle 3 is assumed to constrain the interaction between positive and negative energy matter. This ordinance states that:

Principle 7: Despite its energy sign and its assumed uniformity the negative energy portion of the vacuum does exert the traditionally expected gravitational influence it should have on positive energy matter.

As I previously explained this deduction (which would also apply to the positive energy portion of the vacuum from the viewpoint of negative energy matter) follows from the fact that the restriction that applies on the interaction of positive and negative energy matter does not prevent positive energy matter, when it is conceived as voids in the negative energy portion of the vacuum, from having an influence on that very portion of the vacuum in which the voids occur, just as voids in a matter distribution do exert an influence on this matter. Also, the fact that the energy of the vacuum may be expected to be uniformly distributed does not restrict the influence of the negative portion of it from influencing positive energy matter, simply because we are not dealing in this case with negative energy matter and the negative energy of the vacuum itself cannot be considered as being equivalent to a void in this very vacuum, so that whatever the extent of the distribution of negative energy involved it would still exert its influence on both positive and negative energy matter, unlike a uniform distribution of negative energy matter.

In a previous section I have explained that a consequence of principle 1 in the context where principle 2 (regarding the negativity of the inertial mass of a negative gravitational mass) is considered to apply is that the usual assumption that reversing all mass (gravitational and inertial) would allow to maintain agreement with the equivalence principle (as it is traditionally conceived) is wrong. Therefore, only an altered condition of equivalence between acceleration and a Newtonian gravitational field can remain valid. This condition would be that:

Principle 8: The equivalence of gravitation and acceleration does not apply merely locally, but merely for one single elementary particle (in a given location with a given sign of mass or energy) at once.

What remains true in this context is that the motion of bodies in a gravitational field does not depend on any physical properties of those bodies other than the sign of their mass or energy and this is what will allow the essence of the current theory of the gravitational field to be retained while accommodating a consistent concept of negative energy matter.

Another rule is observed in the context where negative energy matter is governed by principle 1 above and where the appropriate inertial behavior of this type of matter is assumed as a consequence of the validity of principle 2 and principle 6 (which actually imply that the inertial response of negative mass or negative energy bodies to a given force is the same as that of positive energy bodies, as I explained before). This rule would not be required if the traditional assumptions regarding the inertial response of negative energy or negative mass bodies were valid, but given that I have argued that those assumptions are problematic and cannot be justified then it seems that even traditionally we would have a problem if we were not taking the following experimentally motivated principle into account.

Principle 9: When the negative contribution of a field of interaction to the energy of a bound physical system with overall positive energy cannot be independently and directly observed, only the diminished total energy of the bound system contributes to its (previously defined) absolute inertial mass.

Again this is also valid for bound physical systems with overall negative energy for which we may say that when the positive contribution of a field of interaction to the energy of the bound system cannot be independently and directly observed only the diminished (less negative) total energy of the bound system contributes to its absolute inertial mass (which would here be obtained by adding the inertial masses of all positive and negative mass systems as if they were negative). It must be remarked that the validity of this rule does not mean that the opposite contribution to the total energy of a bound system by the attractive field of interaction of its component particles cannot be well defined, only that if it cannot be isolated and independently measured then it also does not independently contribute to the inertial properties of the whole system.

One last constraint is observed to apply when negative energy states are allowed to be occupied (can be propagated forward in time). I have shown in a previous section that this rule can be considered to be theoretically

motivated even though I initially deduced that it was necessary from purely phenomenological arguments. It is the following:

Principle 10: In the absence of an appropriately strong local perturbation from the gravitational field a particle cannot reverse its direction of propagation in time without also reversing its energy and equivalently a particle cannot reverse its energy without also reversing its direction of propagation in time.

Here by ‘negative energy’ I mean negative energy relative to the true (even though conventionally defined) direction of propagation in time, as in the case of the positron as a negative energy electron propagating backward in time. The ten principles enunciated above embody the essence of the insights I gained through an analysis of the problem of negative energy in light of the requirement of relational definition of the physical properties of mass and energy sign. They will now be used to help derive a generalized formulation of the gravitational field equations that will allow to describe the motion of particles with a given sign of energy in the gravitational field of an object with opposite mass or energy.

2.15 Generalized gravitational field equations

I previously indicated that equations would be scarce in this report. But the point has now been reached where it is absolutely necessary to provide some level of quantitative detail regarding the manner by which the concept of negative energy that was developed in the preceding sections is to be integrated into a classical theory of gravitation. The objective I am seeking here though is not to provide a complete treatise on the subject, but merely to introduce the modified gravitational field equations which constitute the core mathematical structure of the generalized theory that emerges from the alternative set of axioms introduced in the preceding section. The essential requirement that must be imposed on a formulation of the gravitational field equations in the context where the principles enunciated in the preceding section are to govern the behavior of negative energy matter is that the gravitational field attributable to a given local source is not to be considered attractive or repulsive depending only on the sign of energy of the source. This can be satisfied by assuming that the gravitational field experienced by a negative energy particle and attributable to a given matter distribution is

actually different from the one experienced by a positive energy particle. In such a context only the difference or the identity between the energy signs of two masses would be physically significant to determine the character of their gravitational interaction, so that any one mass could be considered to have positive energy while masses of opposite energy sign would then have to be the ones to which a negative energy is to be attributed. But the choice of which of two opposite energy bodies has positive energy is itself completely arbitrary.

Thus an observer formed of matter with a given energy sign is free to attribute positive energy to particles with the same sign of energy, even though an observer formed of matter of opposite energy sign may attribute a negative energy to the exact same matter. The only requirement is that the value of the gravitational field (which in a general relativistic theory is associated with the metric properties of space and time) always be adjusted as a consequence of the arbitrary choice regarding attribution of energy signs to various objects. There is, however, a natural choice for the attribution of energy signs by a given observer, which consists in assuming that matter with the same sign of energy as that of the observer itself is always to be considered positive by this type of observer. The viewpoint under which what we traditionally call positive energy matter actually has positive energy is therefore the natural viewpoint of what we traditionally consider to be a positive energy observer, while the viewpoint under which what we traditionally call positive energy matter actually has negative energy is the natural viewpoint of what we would traditionally consider to be a negative energy observer. When this convention is adopted we can write observer dependent gravitational field equations which replace the traditional equations. According to this alternative formulation the motion of matter with a given energy sign is determined by the gravitational field associated with observers of the same energy sign. The gravitational field therefore varies as a function of both the energy sign of the sources and the energy sign of the particles submitted to it, so that only the difference or the identity between the energy sign of the source and that of the matter submitted to the observer dependent gravitational field determines the repulsive or attractive nature of the interaction.

In a relativistic context the observer dependence of the gravitational field would imply that observers of opposite energy signs would actually experience space and time in a different way. But despite the awkwardness of this possibility from the perspective of our conventional perception of spatial

relationships, from a mathematical viewpoint this requirement does not constitute an insurmountable difficulty. We merely have to assume two spaces, related to one another by the fact that the same unique set of events is taking place in both of them, but which may nevertheless have distinct metric properties, in the sense that the events which are taking place in the universe are separated by space and time intervals which are dependent on the energy sign of the observer. Indeed, as I mentioned before, the equations which will be proposed here merely constitute a generalization of the existing mathematical framework of relativity theory and we will therefore be in familiar territory. I am thus in effect assuming that the reader already has a proper understanding of the current general relativistic theory of gravitation and of the physical significance of the various mathematical objects which are relevant to the conventional formulation of this theory. Also, given that attempts at formulating a relativistic theory of gravitation that would allow for the existence of observer dependent gravitational fields were the subject of earlier publications by various authors and since it would be pointless to simply reproduce what has already been discussed elsewhere, I will leave to experts the task of introducing the general framework in which the developments I will propose are to be formulated and concentrate instead on describing the essential, distinctive mathematical features unique to the theory I am proposing.

This choice is appropriate despite the fact that the approach I favor involves several distinctive aspects, because the most general features of the kind of framework involved are not dependent on the specific assumptions of the model considered. The reader may refer in particular to a relatively recent paper [17] in which were introduced meaningful developments essential to any theory in which the gravitational field is assumed to be dependent on the nature of the matter experiencing it. But keep in mind that even the most suitable of the currently available mathematical frameworks still involves theoretical constructs and assumptions which I would consider inappropriate for the formulation of a fully consistent generalized classical theory of gravitation integrating the concept of negative energy matter and therefore only the general structure arising from those developments must be retained. I will here provide an interpretation of such bi-metric theories that is different from those which were tentatively proposed by the few authors that preceded me and this will have significant consequences which will be reflected in the fact that the final equations at which I have arrived are actually distinct from those which had been proposed until now.

In any case it must be mentioned that the gravitational field equations which appear in the above cited paper were not the first equations of that kind to have been developed. Gravitational field equations involving conjugate metrics had already been proposed that simply amounted to allow for negative contributions to the stress-energy tensor of matter⁷, while implicitly conforming to the requirement of symmetry under an exchange of positive and negative energy signs. But even in the more recent publications no justification has ever been provided for the assumptions on which are based the emerging theories and the only practical consequences that were derived from those developments actually appeared to be incorrect or were again unjustified on the basis of the hypotheses which were assumed to characterize the behavior of the gravitationally repulsive matter. In no case did the authors of those developments clearly recognized the exact nature of the anomalously gravitating matter they sought to describe, or attempted to explain how the various problems related to the existence of such matter could be solved. In fact, none of them succeeded in justifying the validity or the superiority of an approach to classical gravitation based on the requirement of exchange symmetry in comparison with the traditional viewpoint according to which gravitational attraction and repulsion are absolutely defined properties of matter.

Meaningful equations were nevertheless derived which happened to be compatible with the simplest of the conditions I have identified above as characterizing a consistent theory of negative energy matter. Those equations therefore constituted a step forward in deriving a quantitative model for the gravitational dynamics of negative energy matter, even if they failed to provide a totally appropriate framework and had to be assumed to apply only under particular circumstances, as they were clearly inappropriate to describe the early phases of cosmic evolution. In any case the equations which were initially proposed were of the following form:

$$\begin{aligned} R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R &= -\frac{8\pi G}{c^4}(T_{\mu\nu} - T_{\mu\nu}^-) \\ R_{\mu\nu}^- - \frac{1}{2}g_{\mu\nu}R^- &= -\frac{8\pi G}{c^4}(T_{\mu\nu}^- - T_{\mu\nu}) \end{aligned} \quad (2.1)$$

⁷I became aware of those developments mainly through the writings of a Frenchman named Jean-Pierre Petit, but given that no official publication from him exists that would contain the set of equations to be discussed here then I cannot provide any relevant reference to his work on the subject.

Here and in what follows G is Newton's constant, c is Einstein's constant (the speed of light in a vacuum) and the Greek indices μ and ν run over the four general coordinate system labels (assuming a metric with diagonal elements +1, +1, +1, -1 in an inertial coordinate system) and the usual notation is used for the curvature tensors $R_{\mu\nu}$ and R experienced by positive energy observers and for the stress-energy tensor $T_{\mu\nu}$ of what we conventionally consider to be positive energy matter. The curvature tensors experienced by negative energy observers are for their part denoted as $R_{\mu\nu}^-$ and R^- , while the stress-energy tensor of what we would conventionally consider to be negative energy matter is here denoted as $T_{\mu\nu}^-$. The first of those two equations can thus be used to determine the geodesics followed by positive energy particles, while the second determines the geodesics followed by negative energy particles. Here all stress-energy tensors would have to be assumed to correspond with positive definite energy densities if it was not for the negative sign in front of the second stress-energy tensor on the right-hand side of each equation which allows a negative contribution to the total stress-energy tensor of matter that is dependent on the particular measure of the sign of energy associated with one or the other type of observer. The negative sign for stress-energies can thus be attributed alternatively to what we would usually consider to be negative energy matter and to what we usually consider to be positive energy matter.

This effectively means that what appears to be negative energy matter to a conventional positive energy observer would really be positive energy matter for an observer we would normally consider to be a negative energy observer, while what appears to be positive energy to a positive energy observer would really be negative energy for an observer usually considered to be made of negative energy matter. Therefore, all energy signs must now be assumed to depend on the energy sign of the observer, which is itself merely a matter of convention. The viewpoint I previously identified as equivalent to a reversal of the sign of mass and according to which it is the gravitational field itself (represented here by the curvature tensors) which actually varies, while the sign of mass (replaced here by the sign of energy) of the observer which experiences that gravitational field is to be considered positive definite, is thus applied and this is certainly appropriate given that it gives rise to equations of the simplest form. It is because there are two different measures for the gravitational field, associated with the two different ways by which the positive and negative contributions to the total energy of matter can be attributed, that there are two equations for the gravitational field instead of

the single one that is usually considered. Otherwise, however, those equations are fairly conventional and were certainly the most straightforward that one could derive for a bi-metric theory, as they were the closest to Einstein's own equation that one could propose.

The fact that, in the context of those equations, the sign of energy contributed by a given mass must now be assumed to depend on the sign of energy which we would normally attribute to the observer determining the associated gravitational field has important consequences. Indeed, if variations in the gravitational field (which is represented by the curvature tensors) are to compensate variations in the stress-energy of matter (as the general covariance of the equations require) then the field attributed to some matter can effectively be either attractive or repulsive depending on the observer that measures the energy of this matter. Four situations may therefore arise when we limit ourselves to merely permute the energy signs of a pair of interacting bodies. First, the source of the field could have what we traditionally consider to be positive energy and the field be attractive, because the particle submitted to it also has positive energy. Next, the source of the field could have what we traditionally consider to be negative energy and the field be repulsive, because again the particle submitted to it also has positive energy. Another possibility is that the source of the field could have what we would traditionally consider to be positive energy and the field nevertheless be repulsive, because we consider its effects on what we would traditionally consider to be a negative energy particle and from which viewpoint the source actually has negative energy. Finally, the source of the field could have what we traditionally consider to be negative energy and the field nevertheless be attractive, again because we consider its effects on what we would traditionally consider to be a negative energy particle and from which viewpoint the source actually has positive energy. This is certainly appropriate from the viewpoint of the principles identified in the preceding section. But given the insights I had already arrived at when I first learned about the mathematical developments which can be used to articulate those requirements, it appeared to me that what the available framework provided was at best an incomplete formulation of the gravitational field equations to be associated with a theory of negative energy matter.

To try to address those shortcomings I thus proposed (in a preprint available online) the following equations which allowed to express the particularities of the indirect gravitational interaction of positive and negative energy

mater that I had come to consider as unavoidable:

$$\begin{aligned} R_{\mu\nu}^+ - \frac{1}{2}g_{\mu\nu}R^+ &= -\frac{8\pi G}{c^4}T_{\mu\nu}^+ \\ R_{\mu\nu}^- - \frac{1}{2}g_{\mu\nu}R^- &= -\frac{8\pi G}{c^4}T_{\mu\nu}^- \end{aligned} \quad (2.2)$$

Here $R_{\mu\nu}^+$ and R^+ are simply the curvature tensors experienced by positive energy observers while $R_{\mu\nu}^-$ and R^- are the curvature tensors experienced by negative energy observers. But the stress-energy tensors figuring in the equations I proposed are actually different from those entering the previously mentioned set of equations despite the similar notation I adopted here, because the $T_{\mu\nu}^+$ tensor encompasses all contributions to the energy and momentum experienced by positive energy observers while the $T_{\mu\nu}^-$ tensor encompasses all contributions to the energy and momentum experienced by negative energy observers and I did assume contributions to those stress-energy tensors which were different from those which had previously been considered in the literature. Thus, when written in a more explicit form with all the components actually entering the stress-energy tensors on the right hand side, the equations I proposed are the following:

$$\begin{aligned} R_{\mu\nu}^+ - \frac{1}{2}g_{\mu\nu}R^+ &= -\frac{8\pi G}{c^4}(T_{\mu\nu}^+ + \check{T}_{\mu\nu}^- - \hat{T}_{\mu\nu}^-) \\ R_{\mu\nu}^- - \frac{1}{2}g_{\mu\nu}R^- &= -\frac{8\pi G}{c^4}(T_{\mu\nu}^- + \check{T}_{\mu\nu}^+ - \hat{T}_{\mu\nu}^+) \end{aligned} \quad (2.3)$$

In this notation all energy-momentum tensors are assumed to be given in their positive definite form and now $T_{\mu\nu}^+$ is the stress-energy tensor of what is usually considered to be positive energy matter while $\check{T}_{\mu\nu}^-$ is the stress-energy tensor associated with the measure of energy of negative energy matter below its average cosmic density (toward the zero energy level) and $\hat{T}_{\mu\nu}^-$ is the stress-energy tensor associated with the measure of energy of negative energy matter above its average cosmic density (toward more negative energies). Similarly, $T_{\mu\nu}^-$ is the stress-energy tensor of what we would usually consider to be negative energy matter while $\check{T}_{\mu\nu}^+$ is the stress-energy tensor associated with the measure of energy of positive energy matter below its average cosmic density and $\hat{T}_{\mu\nu}^+$ is the stress-energy tensor associated with the measure of energy of positive energy matter above its average cosmic density.

This formulation of the generalized gravitational field equations allows me to take into account the fact that there are two distinct categories of

contributions to the total energy density experienced by positive energy observers, one positive definite for all densities of positive energy matter and one that can be either positive or negative depending on the value of energy density of negative energy matter relative not to the ground state but to the density of this negative energy matter averaged over the entire universe. Basically what that means is that the energy measures of the second category of contributions experienced by a positive energy observer are shifted from the traditional zero point of energy to a lower (more negative) energy level below which energies are negative and above which energies are positive up to a maximum value which is reached when no negative energy matter is present at all in the considered location. This redefinition of the measures of energy associated with what we conventionally assume to be negative energy matter simply amounts to subtract the (time dependent) true, negative, average density of energy of this matter (add the absolute value of this density) from every measure of its energy density that contributes to determine the gravitational field experienced by what we conventionally assume to be positive energy matter, that is, the gravitational field observed by positive energy observers. I may add, however, that the required shift in the origin of energy measures for matter with an energy sign opposite that of the observer becomes significant only on the cosmological scale, because in the case of stars and planets it doesn't make much difference if we instead simply consider the true density of positive or negative action matter given that the typical densities which are then involved are much larger than the mean cosmic energy density, which can thus be neglected.

The refinement discussed here is justified (theoretically) by the fact that from the viewpoint of positive energy observers the description of negative energy matter as voids in the positive energy portion of the vacuum requires considering the contribution of negative energy matter as being merely relative to the average density of this matter distribution (and therefore to actually be positive in the presence of underdensities in the average cosmic distribution of negative energy matter) as a consequence of the absence of effects of a uniform negative energy matter distribution on positive energy matter which needs to be assumed for reasons I have explained in a previous section. The equations I proposed also allowed to express the fact that a similar requirement exists for the contributions of positive energy matter to the total stress-energy tensor experienced by negative energy matter. But still I did not find the set of equations I had proposed completely satisfactory. I thought that the right solution should bring a simplification of the

gravitational field equations, while visibly the equations I had derived were even less simple than the equations originally proposed by Einstein, despite the fact that in their compact form they were similar.

As I now understand, however, the equations I had proposed also fell short of meeting a certain mathematical requirement which I have come to appreciate as being essential to a consistent bi-metric theory of gravitation of the kind I sought to develop. This became clear when the paper [17] I mentioned above was published and new equations were proposed, apparently based in part on those I had developed, and which introduced a further refinement to bi-metric theories by not assuming that there is a unique predefined relationship between the metric properties associated with the measurements of positive energy observers and those associated with the measurements of negative energy observers (even though for some reason the author of this paper preferred not to consider that the matter contributing a negative measure to the total stress-energy tensor experienced by positive energy matter actually constitutes negative energy matter). As a consequence of this revised assumption additional variables had to be considered that affected the contribution of negative energy matter to the total stress-energy tensor experienced by positive energy observers or the contribution of what we usually consider to be positive energy matter to the total stress-energy tensor experienced by negative energy observers. The equations proposed were the following, in which the additional factors are written in their explicit form, using my notation⁸, and the quantities are now expressed in units where

⁸From now on I will use a notation that allows to better represent the relative nature of the physical properties associated with spacetime and gravitation. In this notation tensors which refer to positive or negative stress-energies as determined from the viewpoint of positive energy observers will be given a plus or minus upper right indice respectively. Tensors which refer to measures of spacetime curvature or metric properties as observed by positive energy observers will also be given an upper right plus indice, while tensors which refer to the same kind of measures as observed by negative energy observers will be given an upper right minus indice. Also when the distinct ordinary or underlined Greek letter indices used in [17] are not explicitly present to show the nature of the tensor considered, I will simply add another plus or minus indice to the right of that which already characterizes this tensor to define it as an object associated with physical properties as they are experienced by positive or negative energy observers respectively and associated with their own specific metric. For all such tensors, therefore, the first plus or minus indice refers to the matter or gravitational field that is observed while the second plus or minus indice refers to the matter that is observing. The underline which otherwise appears under some letter indices can thus be considered as a shorthand for what should be additional plus or minus indices over the letter indices themselves.

Einstein's constant $c = 1$ and Newton's constant $G = 1/8\pi$:

$$\begin{aligned} R_{\mu\nu}^+ - \frac{1}{2}g_{\mu\nu}R^+ &= -(T_{\mu\nu}^+ - \sqrt{\frac{g^{-+}}{g^{++}}}a_{\nu}^{\underline{\nu}}a_{\mu}^{\underline{\mu}}T_{\underline{\nu}\underline{\mu}}^-) \\ R_{\nu\mu}^- - \frac{1}{2}g_{\nu\mu}R^- &= -(T_{\nu\mu}^- - \sqrt{\frac{g^{+-}}{g^{--}}}a_{\underline{\mu}}^{\mu}a_{\underline{\nu}}^{\nu}T_{\mu\nu}^+) \end{aligned} \quad (2.4)$$

The decisive additional factors are the determinants of what the author calls the pull-overs which are the maps $g_{\nu\mu}^-$ and $g_{\mu\nu}^+$ (originally denoted $h_{\nu\mu}$ and $g_{\mu\nu}$) which we may also write as \mathbf{g}^{-+} and \mathbf{g}^{+-} in tensor form. Those determinants are written here as $g^{-+} = \det(g_{\nu\mu}^-)$ and $g^{+-} = \det(g_{\mu\nu}^+)$ while $g^{++} = \det(g_{\mu\nu}^+)$ is the determinant of the usual metric tensor related to properties of positive energy matter as observed by positive energy observers and $g^{--} = \det(g_{\mu\nu}^-)$ is the determinant of the metric tensor related to properties of negative energy matter as observed by negative energy observers (the map \mathbf{a} is simply used as a means to transform the metric \mathbf{g}^{++} into the \mathbf{g}^{-+} pull-over or the metric \mathbf{g}^{--} into the \mathbf{g}^{+-} pull-over). It is clear therefore that the pull-over \mathbf{g}^{-+} is the map which allows to describe the metric properties obeyed by negative energy matter as they are observed by positive energy observers, while the pull-over \mathbf{g}^{+-} is the map which allows to describe the metric properties obeyed by positive energy matter as they are observed by negative energy observers (which justifies my notation). To better illustrate the relationships involved we may rewrite those equations as:

$$\begin{aligned} R_{\mu\nu}^+ - \frac{1}{2}g_{\mu\nu}R^+ &= -(T_{\mu\nu}^+ - \gamma^{-+}\sqrt{\frac{g^{--}}{g^{++}}}a_{\nu}^{\underline{\nu}}a_{\mu}^{\underline{\mu}}T_{\underline{\nu}\underline{\mu}}^-) \\ R_{\nu\mu}^- - \frac{1}{2}g_{\nu\mu}R^- &= -(T_{\nu\mu}^- - \gamma^{+-}\sqrt{\frac{g^{++}}{g^{--}}}a_{\underline{\mu}}^{\mu}a_{\underline{\nu}}^{\nu}T_{\mu\nu}^+) \end{aligned} \quad (2.5)$$

where γ^{-+} is the absolute value of the determinant of the previously considered map of the metric properties of negative energy matter as negative energy observers measure them to the metric properties of negative energy matter as positive energy observers measure them and vice versa for γ^{+-} . We can then rewrite those equations in compact tensor form by making use of those *metric conversion factors* as:

$$\begin{aligned} \mathbf{G}^+ &= -(\mathbf{T}^{++} - \gamma^{-+}\mathbf{T}^{-+}) \\ \mathbf{G}^- &= -(\mathbf{T}^{--} - \gamma^{+-}\mathbf{T}^{+-}) \end{aligned} \quad (2.6)$$

where \mathbf{G}^+ is the Einstein tensor $G_{\mu\nu}^+ = R_{\mu\nu}^+ - \frac{1}{2}g_{\mu\nu}R^+$ related to positive energy observers, \mathbf{G}^- is the similar Einstein tensor related to negative energy observers, \mathbf{T}^{++} is the stress-energy tensor of positive energy matter as measured by positive energy observers, $\gamma^{-+}\mathbf{T}^{-+}$ is the stress-energy tensor of negative energy matter as measured by positive energy observers, \mathbf{T}^{--} is the stress-energy tensor of negative energy matter as measured by negative energy observers and finally $\gamma^{+-}\mathbf{T}^{+-}$ is the stress-energy tensor of positive energy matter as measured by negative energy observers.

As is apparent, however, the proposed equations were still of the traditional kind, in the sense that they did not allow to take into account the fact that negative energy matter is experienced as voids in the positive energy portion of the vacuum (and vice versa for positive energy matter from the viewpoint of negative energy observers). The complexity of those equations can be made more apparent by explicitly adding a term for the observed positive vacuum energy density related to the cosmological constant and taking into account the observer dependence of the measures of positive energy:

$$\begin{aligned}\mathbf{G}^+ &= -(\mathbf{T}^{++} + \mathbf{T}_{\Lambda}^{++} - \gamma^{-+}\mathbf{T}^{-+}) \\ \mathbf{G}^- &= -(\mathbf{T}^{--} - \gamma^{+-}\mathbf{T}^{+-} - \gamma^{+-}\mathbf{T}_{\Lambda}^{+-})\end{aligned}\tag{2.7}$$

In those equations $\mathbf{T}_{\Lambda}^{++} = -\Lambda\mathbf{g}^{++}$ would be the stress-energy tensor associated with the usual measure of energy density of vacuum fluctuations (with Λ as the positive cosmological constant to which is associated a *positive* energy density) and $\gamma^{+-}\mathbf{T}_{\Lambda}^{+-}$ would be the stress-energy tensor associated with the same positive vacuum energy density as measured by negative energy observers, which would involve the same mapping or conversion factors as apply to measures of the properties of ordinary positive energy matter performed by negative energy matter observers. In such a form and when all relevant contributions to the stress-energy tensors are effectively considered, the apparent symmetry of the original equations is manifestly lost, as their form becomes dependent on the actual value of the sign of the average energy density of vacuum fluctuations. To me at least, it is obvious that those equations cannot be considered to embody a simplification of Einstein's theory that could be considered a definite improvement of it.

In order that such a formulation of bi-metric theory be allowed to at least meet the requirements I had already identified and which were not taken into account in this later proposal I would first suggest that we consider the limitations imposed on the interaction of positive and negative energy matter

by the fact that the void in the positive energy portion of the vacuum that is equivalent to the presence of a homogeneous distribution of negative energy matter has no gravitational effect on positive energy matter (and vice versa when we consider the void in the negative energy portion of the vacuum). In such a case we would simply have to replace the usual stress-energy tensors associated with the measures of energy of positive and negative energy matter made by observers of opposite energies with the following *irregular stress-energy tensors* which provide measures of the observed variations of energy density of positive and negative energy matter above and below their average cosmic densities:

$$\begin{aligned}\gamma^{-+}\tilde{\mathbf{T}}^{-+} &= \gamma^{-+}(\mathbf{T}^{-+} - \bar{\mathbf{T}}^{-+}) \\ \gamma^{+-}\tilde{\mathbf{T}}^{+-} &= \gamma^{+-}(\mathbf{T}^{+-} - \bar{\mathbf{T}}^{+-})\end{aligned}\tag{2.8}$$

where $\gamma^{-+}\mathbf{T}^{-+}$ and $\gamma^{+-}\mathbf{T}^{+-}$ can be assumed to be the usual measures of stress-energy of negative and positive energy matter respectively (as experienced by observers of opposite energy signs) relative to the traditional zero level of energy and $\gamma^{-+}\bar{\mathbf{T}}^{-+}$ and $\gamma^{+-}\bar{\mathbf{T}}^{+-}$ are the measures of average stress-energy of negative and positive energy matter (as experienced by observers of opposite energy signs) observed on a cosmic scale. In such a context it appears that negative energy matter would now contribute negatively to the total measure of stress-energy experienced by a positive energy observer only when the magnitude of its local energy density (as measured by this positive energy observer) is larger than the magnitude of its average energy density (as measured by the same positive energy observer). Otherwise negative energy matter would actually contribute positively to the total measure of stress-energy experienced by a positive energy observer up to a maximum level fixed by the average density of negative energy matter (as measured by this positive energy observer). The same remark would apply for the contribution of what is usually considered to be positive energy matter to the total measure of stress-energy experienced by a negative energy observer, which would be opposite the energy contribution of negative energy matter only when the magnitude of the local density of positive energy matter (as measured by the negative energy observer) is larger than the magnitude of its average cosmic density.

It must be noted, however, that even though positive contributions to the energy density measured by positive energy observers may occur which would be attributable to the presence of underdensities in the negative energy

matter distribution, we must nevertheless apply the conversion factor γ^{-+} to such energy measures, because they still relate to measurements regarding the density of negative energy matter which are subject to the same mapping relationships as apply to other (truly negative) measures of energy related to negative energy matter made by a positive energy observer. Of course this is also true concerning below average measures of the energy density of what we would usually consider to be positive energy matter made by negative energy observers. Indeed, even when the second contribution to the energy density of matter is of the same sign as the energy of the matter experiencing the gravitational field it is still undetermined to the same extent as negative contributions, because what is unknown (due to the impossibility to directly compare measures of distances related to positive and negative energy observers) is the exact true density of negative energy matter (in comparison with that of positive energy matter) and this indefiniteness also affects the positive value of such contributions. Therefore, positive energy contributions from underdensities of negative energy matter are contained in the same tensor as negative energy contributions.

A more appropriate set of gravitational field equations would therefore take into account the shifted origin of the measures of stress-energy related to positive and negative energy matter as they are experienced by observers of opposite energy signs:

$$\begin{aligned} \mathbf{G}^+ &= -(\mathbf{T}^{++} + \mathbf{T}_{\Lambda}^{++} - \gamma^{-+}\tilde{\mathbf{T}}^{-+}) \\ \mathbf{G}^- &= -(\mathbf{T}^{--} - \gamma^{+-}\tilde{\mathbf{T}}^{+-} - \gamma^{+-}\mathbf{T}_{\Lambda}^{+-}) \end{aligned} \quad (2.9)$$

But clearly, for what regards simplicity, we appear to be no better off than with the previous set of equations. Something is still missing from those equations. At this point I suggest that we take a bold step forward and instead of trying to derive the gravitational field equations from a variational principle, as is usually done, we rather follow Einstein's way and simply guess what the final form of the equations should be that would generalize the set of equations (2.9) I have just proposed, which would otherwise constitute the most accurate description of the gravitational dynamics of positive and negative energy matter. As I have been able to understand, the crucial step in this process consists in reconsidering the meaning of the vacuum energy terms whose contributions I had long suspected were inappropriately attributed in the context of bi-metric theories. Indeed, I always thought that the cosmological term should arise from an asymmetry between some positive

contribution and some negative contribution to the energy budget, while in the current set of equations it occurs only as an additional term which must merely be given the appropriate relative measure depending on whether it is observed by a positive energy observer or a negative energy observer, which I do not find satisfactory.

It is only when I recognized the profound significance of my description of positive and negative energy matter as voids in their respective opposite energy vacuums that I was able to achieve the breakthrough that allowed me to guess what the appropriate generalized gravitational field equations are that allow the concept of negative energy matter to be integrated into a general relativistic framework in a way that actually simplifies Einstein's theory rather than further complicate things. What I realized, basically, is that if the results of my previously described analysis is right then all energy is vacuum energy, either present or missing. An additional insight was then necessary which consists in recognizing that the natural value of positive or negative contributions to vacuum energy density is actually provided by the Planck scale. What must be understood is that when we remove energy from the vacuum we decrease its energy density from a maximum (positive or negative) value which is fluctuating quantum mechanically (upon measurement) in just the same measure as does the energy of matter itself. Therefore, if the presence of negative energy matter is to be considered as equivalent to the presence of a void in the positive energy portion of the vacuum, then locally we should observe a value of fluctuating vacuum energy density that would be decreased from its natural maximum value in just the same measure as that of the energy of the matter that is present. Given that the level of fluctuation of vacuum energy involved would be as large as the void considered is small it is possible to assume that there is an exact correspondence between the missing vacuum energy and the energy of the matter ordinarily expected to be present, which is known to be fluctuating (even if it is effectively the measure of momentum that is involved) in proportion with the level of spatial confinement to which the matter is submitted. The natural energy level involved would thus correspond to that which is known to be associated with the highest level of fluctuation, which is effectively the Planck energy⁹.

⁹The validity of this assumption could be the subject of controversy, but given that the most advanced and least speculative theoretical developments toward a theory of quantum gravitation indicate that this is an appropriate and unavoidable constraint I will nevertheless consider it to be universally valid. However, even if the existence of such a limit to the energy associated with quantum fluctuations was to be found irrelevant there

Therefore, any missing vacuum energy attributable to the presence of matter with an energy sign opposite that of the portion of vacuum in which it arises could be considered as a local decrease over the maximum energy density determined by the Planck scale.

Let me thus introduce the generalized gravitational field equations which allow to fulfill all the requirements I have identified as essential aspects of a classical theory of gravitation that solves the problem of negative energies. The formula in all its beauty and simplicity is the following:

$$\mathbf{G}^{\pm} = -\mathbf{T}_v^{\pm} \quad (2.10)$$

where \mathbf{G}^{\pm} is the Einstein tensor associated with the metric properties experienced by what we would usually consider to be positive and negative energy observers and \mathbf{T}_v^{\pm} is the *vacuum stress-energy tensor* associated with the measures of vacuum energy effected by the same positive and negative energy observers. The similarity with the compact form of Einstein's own equation is very clear, but it is also somewhat misleading, as the right hand side of the equation proposed here is a much more general object than the stress-energy tensor of matter which appeared in the original theory. I will now define it with various levels of precision and generality. If we first consider the significance of the equation for a positive energy observer we would obtain the following equation:

$$\mathbf{G}^+ = -(\mathbf{T}_v^{++} - \mathbf{T}_v^{-+}) \quad (2.11)$$

in which \mathbf{G}^+ is again the Einstein tensor associated with the gravitational field experienced by positive energy observers, but now the vacuum stress-energy tensor is decomposed into its positive and negative energy portions \mathbf{T}_v^{++} and \mathbf{T}_v^{-+} as they are measured or experienced by such positive energy observers. In accordance with what was explained above we would then obtain the next level of decomposition of the equations in which the two opposite energy portions of vacuum fluctuations (as they are experienced by positive energy observers) are given their explicit form:

$$-\mathbf{G}^+ = (\mathbf{T}_P^+ - \gamma^{-+}\mathbf{T}^{-+}) - (\gamma^{-+}\mathbf{T}_P^- - \mathbf{T}^{++}) \quad (2.12)$$

is no a priori reason why the following results would have to be considered invalid. I believe that the situation we have here is similar to that which existed at the turn of the twentieth century concerning the hypothesis of the existence of atoms which was often rejected on the basis of an absence of direct observational evidence despite the fact that this assumption had actually become unavoidable theoretically.

where \mathbf{T}_P^+ and $\gamma^{-+}\mathbf{T}_P^-$ are the *natural vacuum stress-energy tensors* associated with the maximum positive and negative contributions to the energy density of zero-point vacuum fluctuations as they are experienced by a positive energy observer. Here, the previously introduced metric conversion factors associated with the absence of fixed relationships between the metric properties experienced by negative energy matter and those experienced by positive energy matter appropriately apply only to the negative energy contributions of either matter or vacuum, independently. The γ^{-+} factors therefore occur only at the level of decomposition of equation (2.12) and not in equation (2.11), because they must effectively be attributed independently to the actual positive and negative energy contributions (whether they are vacuum or matter) and the stress-energy tensors of matter provide opposite energy contributions to those of the portion of vacuum in which they occur (given that matter of a given energy sign is to be conceived as voids in the opposite energy portion of the vacuum).

The preceding equation can then be rewritten under the following form when we take into account the previously introduced definition of the measure of stress-energy associated with negative energy matter as it would actually be experienced by positive energy observers, which are only affected by *variations* in the density of negative energy matter:

$$-\mathbf{G}^+ = \mathbf{T}^{++} - \gamma^{-+}\tilde{\mathbf{T}}^{-+} + (\mathbf{T}_P^+ - \gamma^{-+}\mathbf{T}_P^-) \quad (2.13)$$

which allows to identify the observed value of vacuum energy density associated with the cosmological constant as that which is provided by the following tensor:

$$\mathbf{T}_\Lambda^+ = \mathbf{T}_P^+ - \gamma^{-+}\mathbf{T}_P^- \quad (2.14)$$

where the plus indice attributed to this stress-energy tensor (associated with the energy of the vacuum in the absence of matter) now merely denotes the conventional energy sign of the observer experiencing it without referring to an actual energy sign of vacuum fluctuations themselves, which could in principle be either positive or negative and which is determined solely by the conversion factor provided by the previously discussed map of the metric properties associated with negative energy matter as they are experienced by positive energy observers. Indeed, given the invariant nature of the maximum density of vacuum energy associated with the Planck scale for an observer having the same energy sign as that of the contribution considered,

the above equation means that the net value of vacuum energy density observed by positive energy observers arises as a consequence of a very small but non-trivial difference in the metric properties associated with the motion of positive energy matter as experienced by positive energy observers and the metric properties associated with the motion of negative energy matter as experienced by those same positive energy observers. In any case it is now possible to write the generalized gravitational field equation associated with positive energy observers in its most explicit form as:

$$\mathbf{G}^+ = -(\mathbf{T}^{++} - \gamma^{-+}\tilde{\mathbf{T}}^{-+} + \mathbf{T}_\Lambda^+) \quad (2.15)$$

which confirms its equivalence with the previously proposed equation (2.9) at which I had arrived on the basis of considerations of a physical nature. It may be added that if we are considering this equation in a cosmological context then the $\gamma^{-+}\tilde{\mathbf{T}}^{-+}$ tensor would presumably reduce to zero on average (as the overdensities of negative energy matter would cancel out underdensities present in the same matter distribution) so that the relevant equations for positive energy observers would now be of the traditional form:

$$\mathbf{G}^+ = -(\mathbf{T}^{++} + \mathbf{T}_\Lambda^+) \quad (2.16)$$

as is known to be appropriate given the success of current cosmological models for predicting the relevant features of our universe's history. We may also write the following set of equations which would provide the various levels of decomposition of the general equation (2.10) that apply from the viewpoint of negative energy observers:

$$\begin{aligned} \mathbf{G}^- &= -(\mathbf{T}_v^{--} - \mathbf{T}_v^{+-}) \\ -\mathbf{G}^- &= (\mathbf{T}_P^- - \gamma^{+-}\mathbf{T}^{+-}) - (\gamma^{+-}\mathbf{T}_P^+ - \mathbf{T}^{--}) \\ \mathbf{G}^- &= -(\mathbf{T}^{--} - \gamma^{+-}\tilde{\mathbf{T}}^{+-} + \mathbf{T}_\Lambda^-) \end{aligned} \quad (2.17)$$

where $\mathbf{T}_\Lambda^- = \mathbf{T}_P^- - \gamma^{+-}\mathbf{T}_P^+$ would provide the (positive or negative) value of vacuum energy density observed by such a negative energy observer. The last equation, as well the other two, are now manifestly symmetric with the corresponding equations associated with positive energy observers, as I have argued should be required. But the most remarkable feature of those equations (and the related equations for the gravitational field experienced by a positive energy observer) is that they are effectively obtained from a

very simple expression (the first of the three equations) that determines the gravitational field merely as a function of the relatively defined measures of positive and negative vacuum energy and which alone allows to embody the essence of the emerging framework.

It must be noted that both the value of total vacuum energy density (associated with an absence of matter) related to positive energy observers and that related to negative energy observers (which are here given by the vacuum energy terms \mathbf{T}_Λ^+ and \mathbf{T}_Λ^- respectively) could in principle vary with position (and incidentally also with time) given that they involve the variable metric conversion factors γ^{-+} and γ^{+-} respectively. Thus, the measure of vacuum energy density associated with the cosmological constant and applying on the largest spatial scale would effectively be an average quantity and there is no a priori reason why it could not give rise to local effects that would deviate from those associated with the cosmic scale. But given that we know at least that on the cosmic scale $\mathbf{T}_\Lambda^+ = \mathbf{T}_P^+ - \gamma^{-+}\mathbf{T}_P^-$ is very small compared with the natural energy scale provided by the Planck energy, then it is possible to conclude that the correction provided by the γ^{-+} conversion factor is itself actually very small on such a scale. This observation therefore indicates that there is a near perfect level of symmetry between the metric properties associated with positive energy matter and those associated with negative energy matter, as experienced by positive energy observers.

The quantitative aspects of the proposed integration of negative energy states to classical gravitation theory having being properly introduced, it is now possible to look back and examine whether the equations obtained can effectively provide the structure of an alternative model which would conform to all of the principles enunciated in the preceding section. As I previously remarked the basic structure of the proposed bi-metric theory was adopted precisely because it allows the kind of arbitrariness of the attribution of the sign of energy that is required for this physical property to be defined in a relational manner. But the ultimate confirmation that the proposed framework is compatible with the fundamental requirement expressed by principle 1 is the fact that even in the presence of a non-vanishing value for the cosmological constant, the set of equations (2.17) describing the motion of negative energy matter is now symmetric with the corresponding set of equations describing the motion of positive energy matter. Furthermore, the requirement set by principle 2 that inertial mass be reversed along with gravitational mass is also fulfilled by the proposed gravitational field equations given that my

analysis of the physical property of inertia has shown that imposing such a condition should give rise to gravitational attraction between masses of the same sign (whatever this sign is assumed to be) and to gravitational repulsion between masses of opposite signs and this is precisely what we obtain with the proposed equations, even if the sign of energy that replaces the sign of mass is here arbitrary and the gravitational field is a variable property dependent on the nature of the matter submitted to it.

On the other hand, the validity of principle 3 and the absence of direct interaction between positive and negative energy matter particles may seem to be threatened by the fact that the stress-energy tensor associated with negative energy matter contributes to determine the gravitational field experienced by positive energy matter. But again, in the context of the more refined set of equations I have proposed, it is explicit that the negative contribution that enters the total measure of the stress-energy of matter that determines a gravitational field and which we associate with the presence of negative energy matter is actually a measure of the amount of stress-energy missing from the positive portion of vacuum energy. The effect on positive energy matter which must be taken into account in the presence of negative energy matter cannot therefore be attributable to an interaction with negative energy matter (whose presence is not directly felt by a positive energy observer), but must necessarily come from an interaction between positive energy matter and the surrounding positive energy vacuum. The equations thus naturally require that there are no direct interactions between particles with opposite energy signs.

The new equations are also an expression of principle 4 and principle 5, because they allow the voids in the positive energy portion of the vacuum to actually provide a negative contribution to the total stress-energy tensor of matter and in a general relativistic context, a negative contribution to the stress-energy of matter must be matched by a contribution to the gravitational field that is opposite to that produced by positive stress-energy, so that if positive energy produces an attractive field from the viewpoint of positive energy matter, negative energy must produce a repulsive field from the same viewpoint. The presence of voids in an otherwise uniform distribution of positive vacuum energy should therefore give rise to uncompensated forces opposite those attributable to the presence of an equivalent amount of positive energy matter and by analogy the same should also be true for voids in a uniform positive energy matter distribution. As a result, not only is it implied by the equations that there are no gravitational interactions between

particles with opposite energy signs, but it is also allowed that all gravitational forces be the result of interactions between particles with identical energy signs, whether they are actual matter particles or virtual particles from the vacuum (which may actually be more real than matter particles). The gravitational field equations I proposed are therefore the perfect embodiment of the requirements set by principles 3, 4 and 5, which I had previously identified as necessary conditions of the emerging theory.

We can now understand why it would be inappropriate to assume, as some authors do, that the energy of the gravitationally repulsive matter whose behavior is described by conventional bi-metric theories is positive even for an observer that measures a negative contribution from it to the total stress-energy of matter. Indeed, according to the above proposed equations such matter would produce a gravitational field that would itself have an energy content (to the extent that a definite energy could effectively be associated with the gravitational field) opposite that which is produced by particles contributing positively to the total stress-energy of matter. But this means that if matter was assumed to always have positive energy then when energy is exchanged between the two types of matter the variation of total gravitational energy (which would occur because opposite variations of *opposite* gravitational energies are involved) would not be compensated by a variation of the energy of matter (which would involve opposite variations of *positive* energies). Therefore in the case of our two colliding bodies exerting a gravitational repulsion on one another it would be impossible for the variation of energy of the decelerating body to be compensated by a variation of energy of the gravitational field attributable to the changes occurring in the related portion of vacuum energy which would be equivalent to the energy changes occurring as a consequence of the acceleration of the second body, despite the fact that this must be considered necessary if energy is to be conserved, as I previously explained.

Those problems can be avoided, however, when real negative energy states are allowed for matter, because in a general relativistic context the variations in the gravitational field can effectively balance the changes occurring in the stress-energy of the two interacting matter components and given that the gravitational interaction is responsible for all energy exchange between opposite energy bodies, then no energy variations remain uncompensated. I think that this is a clear indication that the tentative solution of the problem of vacuum decay (the collapse of matter to ever more negative energy states) through the contradictory proposal of a gravitationally repulsive mat-

ter that would have positive energy (from all viewpoints) is misguided and ineffective. Thus, if an observer is allowed to attribute a positive energy to matter of his own kind, regardless of which matter he is made of, it should be clear that once this choice is made the energy sign of the matter which from the viewpoint of this same observer provides a negative contribution to the stress-energy tensor of matter must be assumed negative. In any case I must mention again that from a cosmological viewpoint the growth of negative energy matter overdensities occurring from an initially homogeneous distribution of such matter will always be compensated by an opposite growth of underdensities in the surrounding environment. But given that from my viewpoint those two kinds of inhomogeneities provide opposite contributions to the total stress-energy tensor of matter experienced by a positive energy observer, then it follows that there is an additional constraint regarding the conservation of energy contributed by negative energy matter and this is a further confirmation of the viability of the proposed equations.

Returning to the criteria imposed by the principles enunciated in the preceding section, we can readily assess that the condition set by principle 6 (according to which only density variations over and below the average cosmic density of negative action matter have an effect on positive energy matter) is also reflected in the equations proposed above. Indeed, the modified measure of negative stress-energy provided by the irregular stress-energy tensor $\gamma^{-+}\tilde{\mathbf{T}}^{-+}$ which naturally enters the gravitational field equation associated with a positive energy observer actually allows to fulfill the requirement set by principle 6 given that it provides a measure of stress-energy from which is subtracted the average stress-energy of negative energy matter. This compliance of the proposed gravitational field equations may perhaps appear to be of secondary concern given the negligibility of the average density of positive energy matter (and presumably also of negative energy matter) in comparison with the densities encountered under most circumstances when we are dealing with astronomical objects of interest like stars or even galaxies. But, if it was not for the modified measure of negative stress-energy provided by the second term of equation (2.15), or the corresponding term from equation (2.17), serious problems would occur.

In a preceding section (dealing with the concept of negative mass) I mentioned in effect that if a body with a given mass sign was to interact with all matter of both positive and negative mass that is present on the cosmological scale then the classical phenomenon of inertia itself could not even exist (be-

cause the inertial forces resulting from acceleration relative to positive and negative mass matter would cancel out). However, a Newtonian model is all about inertia, so that if inertial forces were made impossible by the presence of negative energy matter, then reduction of the relativistic equations to a Newtonian gravitation theory with gravitationally repulsive negative mass densities would effectively be impossible, even as an approximation. I believe that ignorance of the requirement to impose a suitable, modified measure of negative stress-energy for the generalized gravitational field equations is in fact the ultimate source of the difficulties which according to certain authors are encountered in trying to obtain an appropriate Newtonian limit from traditional bi-metric theories. This is in addition to the fact that, without the appropriate measure of negative stress-energy, complex hypotheses (of the kind which are often found in the literature) would have to be assumed concerning the expected variation in time of the ratio of the cosmic densities of positive and negative energy matter in order to try to maintain the agreement of the proposed models with astronomical observations regarding the rate of expansion of ordinary positive energy matter, which are already well matched by traditional cosmological models when no negative energy matter is assumed to be present.

Finally, the fact that two maximum contributions of opposite signs to the energy density of the vacuum are now explicitly present in the most general form of each of the gravitational field equations means that both positive and negative contributions to the energy of the vacuum itself (ignoring voids) are allowed to contribute to the gravitational field experienced by positive or negative energy matter on the cosmological scale, as required by principle 7. From this alternative viewpoint what allows one to effectively ignore most of the effects that the vacuum would have on the gravitational field experienced by positive or negative energy matter is merely the fact that those opposite energy contributions nearly cancel each other out. I may also mention that the condition set by principle 8 (that the equivalence principle be valid not merely locally, but really for one unique particle with a given energy sign) is implicitly contained in the structure of the equations at the most basic level, because they describe gravitational fields which are dependent not merely on the location, but also on the sign of energy of the particles submitted to them. On the other hand, principles 9 and 10, which identify requirements that have to do with the elementary properties of matter particles (namely the absence of independent energy contributions for bound systems and the impossibility under ordinary circumstances of a reversal of action on a con-

tinuous particle world-line), are not explicitly contained in the gravitational field equations proposed here, but if we assume the validity of those equations then experimental facts make those constraints unavoidable.

2.16 Summary

To conclude this chapter, I would like to provide a summary of all the results which were obtained concerning the problem of negative energy in the context of the improved understanding of the issue of time directionality which underlies those developments. The decisive results are the following.

1. There is no valid observational argument against the existence of negative energy matter and what is required by the facts is merely that negative energy matter does not interfere under most circumstances with processes involving positive energy matter.
2. The introduction of antiparticles does not constitute a complete and acceptable solution to the problem of negative energy states.
3. There exists a fundamental degree of freedom associated with the direction of propagation in time of elementary particles.
4. The sign of energy is purely conventional given that it cannot be defined independently from the direction of propagation in time of the particle carrying this energy which is itself a matter of coordinative definition.
5. The only significant measure of energy sign from a gravitational viewpoint is that provided by the sign of action obtained by multiplying the sign of energy by the sign of time intervals.
6. The sign of action is also a matter of convention dependent on the choices made regarding what should be the sign of energy of those particles which are considered to propagate forward in time.
7. The sign of action cannot be asserted other than as a relative property between different particles sharing the same convention regarding what shall be the direction of propagation in time of a given particle with an arbitrarily chosen sign of energy relative to this direction of time.

8. As a consequence of a certain condition of continuity of particle world-lines a particle with a given conventionally defined sign of charge relative to a given direction of propagation in time cannot be assumed to also occur as a particle carrying an opposite charge in the opposite direction of time.
9. Any anomalous response of a conventionally defined negative action particle to the gravitational field of a conventionally defined positive action body must be shared by a positive action particle in the gravitational field of a negative action body.
10. As a matter of consistency a negative action or negative mass body must be assumed to have both negative gravitational mass and negative inertial mass, not because of some perceived requirement from the equivalence principle, but because mass as one single physical property cannot be attributed mutually exclusive or contradictory attributes.
11. Contrarily to what is usually assumed the hypothesis that inertial mass reverses along with gravitational mass does not give rise to absolutely defined attractive or repulsive gravitational fields.
12. It is the incorrect assessment of the response of a negative mass body to any applied force, made on the basis of current assumptions regarding the effect of a reversal of inertial mass, that is responsible for allowing an absolute character of attractiveness or repulsiveness to be associated with a given sign of mass.
13. It is inappropriate to assume that inertial mass remains positive for a negative mass body not only because this assumption would not give rise to the kind of ordinary response to forces that is usually assumed of such a mass, but also because even if the response was appropriate a body with such properties would irreconcilably violate the equivalence principle as a consequence of the fact that the same inertial mass would respond differently to a given gravitational field depending on the sign of the associated gravitational mass.
14. The direction of the equivalent gravitational field experienced by an accelerating body must be considered to be dependent on the sign of mass of the body and therefore to be equal rather than opposite the acceleration for a negative mass body, so that the inertial force on a

negative mass body is left invariant despite the fact that its inertial mass must be assumed to be negative.

15. A generalized formulation of Newton's second law involving a dynamic equilibrium between applied forces and the inertial force associated with the equivalent gravitational field, instead of an equilibrium between forces and acceleration, allows to predict that $\mathbf{F} = -m\mathbf{a}$ when the mass m is negative, so that the acceleration of a negative mass body takes place in the direction of the applied force, as is the case for a positive mass body.
16. When the mass experiencing a gravitational field is considered positive definite, while it is the direction of the gravitational field attributable to a given local matter distribution which itself varies under exchange of positive and negative energy observers, the same outcome as would occur when the equivalent gravitational field is reversed along with the mass of the body experiencing an unchanged local gravitational field must be observed.
17. Not only is it allowed that the principle of relativity, which motivates the equivalence principle, be preserved by the proposed alternative concept of negative mass, but in fact it is this very principle that requires such a concept of negative mass according to which only the difference or the identity between the signs of mass of two bodies has a physical significance.
18. Only when local inhomogeneities in the matter distribution are not superposed for positive and negative energy matter can there be an effect of acceleration or rotation relative to those matter concentrations.
19. Given the unavoidable similarity of the large scale distributions of positive and negative energy matter, the phenomenon of inertia as an effect of acceleration relative to the large scale matter distribution can only occur if a body with a given mass sign gravitationally interacts solely with the large scale distribution of matter having the same sign of mass as its own, because otherwise the effects attributable to positive and negative mass matter cancel out.
20. The generalization of the equivalence principle made necessary by the existence of negative mass matter implies the physical nature of the

gravitational field as resulting from particle interactions despite the fact that it still allows a geometrical treatment of gravitation, because the metric properties of space and time are now themselves relatively defined properties which arise as a consequence of an equilibrium of local and inertial gravitational forces which depend on the sign of energy of the bodies experiencing them.

21. The equivalence principle must be generalized in such a way that it applies not merely locally but only for a single elementary particle with one mass or energy sign at once for which there would never be a difference between acceleration and a gravitational field.
22. A void in a uniform positive energy matter distribution is not equivalent in general to the case of a hollow sphere of finite size and positive energy bodies present inside such a void would actually experience a repulsive gravitational force as a consequence of the absence of gravitational attraction from the matter that is missing.
23. Birkhoff's theorem does not contradict the preceding conclusion because it is valid only in a universe that is spherically symmetric around any point and for a homogeneous and isotropic matter distribution this condition is met only in the absence of a local void in a uniform matter distribution.
24. It is inappropriate to assume that when we are considering a spherical region of the universe the rest of the universe surrounding that region can be considered as a hollow sphere simply on the basis of the fact that according to the cosmological principle matter is distributed uniformly in all directions.
25. In the presence of a spherical void in an otherwise uniform matter distribution spherical symmetry exists only at the center of the void so that the presence of the void would necessarily alter the equilibrium of gravitational forces anywhere else inside (and to some extent also outside) the void in such a way as to produce a force that would be the opposite of that which we would attribute to the presence of an equivalent additional quantity of matter with the same energy sign in place of the void.

26. The mistake involved in the traditional interpretation of Birkhoff's theorem consists in considering that the surrounding matter which could influence the particles located inside a chosen spherical region in a homogeneous and isotropic universe is spherically distributed around the center of the spherical region considered, instead of recognizing that the center of mass in a universe without boundary is always located at the position of the observer experiencing the effects of the matter distribution.
27. The gravitational repulsion that would be exerted on a positive energy body as a consequence of the presence of a void in a uniform positive energy matter distribution is actually the consequence of uncompensated gravitational attraction by matter with the same energy sign as that of the body.
28. Negative energy states are phenomenologically equivalent to an absence of positive energy from the vacuum, because removing positive energy from a vacuum with near zero energy is like decreasing energy into negative territory.
29. The expected gravitational repulsion exerted on a positive energy body by negative energy matter would occur as a consequence of the fact that the absence of positive energy from a region of the vacuum that is equivalent to the presence of negative energy matter would result in an uncompensated gravitational attraction from the surrounding positive energy vacuum pulling positive energy matter away from the region where the energy is missing.
30. A void in a uniform positive energy matter distribution remains physically distinct from a local absence of positive vacuum energy, even if in both of those cases the effects are equivalent to the presence of an excess of matter of negative energy sign, simply because an absence of matter (with positive energy sign) is necessarily different from the presence of matter (with negative energy sign).
31. Voids in the negative energy portion of the vacuum are equivalent to the presence of positive energy matter and along with voids in a uniform negative energy matter distribution would produce an equivalent gravitational repulsion on negative energy matter.

32. A void in a uniform negative energy matter distribution remains clearly distinct from a void in the uniform distribution of negative vacuum energy.
33. A description of matter of a given energy sign as voids in a filled distribution of matter of opposite energy sign would involve a violation of the requirement of a relational definition of the sign of energy because it would allow a forbidden absolute distinction between positive and negative energy matter given that there can only be one filled matter distribution.
34. There must be a certain compensation between the usually considered positive contributions to vacuum energy and the usually ignored negative contributions by the same particles, so that the natural value of the cosmological constant which we can expect to observe should actually be as small as the symmetry under exchange of positive and negative energy matter is perfect.
35. It can no longer be assumed that there is a clear distinction between matter and vacuum given that matter is merely a manifestation of missing vacuum energy.
36. When we are considering a void in the positive energy portion of the vacuum we must ultimately always be referring to the absence of a specific elementary particle with specific non-gravitational charges so that a real particle's charges are actually a manifestation of opposite charges missing from the vacuum.
37. There can be no equivalent gravitational repulsion on positive energy matter from the presence of the void of cosmic proportion in the positive energy portion of the vacuum that is equivalent to a uniform distribution of negative energy matter.
38. The rate of universal expansion of matter with a given sign of energy is not influenced by the presence of matter with an opposite energy sign.
39. The voids in the negative energy portion of the vacuum which are equivalent to the presence of positive energy matter would interact with themselves even if the missing negative energy was uniformly distributed throughout all of space and despite the fact that a similar

distribution of missing positive vacuum energy would have no effect on positive energy matter.

40. The effect on a particle with a given energy sign of the presence of a void in a uniform matter distribution of opposite energy sign would be a gravitational attraction directed toward the void.
41. There can be no direct interactions of any kind between positive and negative energy particles, because no definite energy sign can be attributed to the fields of interaction between opposite energy particles and therefore negative energy matter must be dark.
42. Despite the absence of any direct interaction between opposite energy particles there exists an indirect gravitational repulsion between opposite energy bodies as a consequence of the equivalence between the presence of negative energy matter and a void in the positive energy portion of the vacuum which for positive energy bodies gives rise to an uncompensated gravitational attraction directed away from this void.
43. The absence of direct interaction between positive and negative energy particles does not mean that positive energy matter does not experience the gravitational effects of the negative energy portion of the vacuum, because as a particular manifestation of negative vacuum energy, positive energy matter cannot be expected not to interact with the negative energy portion of the vacuum.
44. Opposite action particles with opposite charges cannot annihilate one another under normal conditions, because there are no direct interactions between particles with opposite action signs, which means that they cannot come into contact with one another except under conditions of extremely high opposite energies or very short spatial scale under which the indirect gravitational interaction they do experience is no longer negligible.
45. The creation of pairs of opposite action particles out of the vacuum is prevented from occurring under ordinary circumstances as a consequence of the weakness of the indirect gravitational interaction between such particles which requires very high (positive and negative) energy particles to be created while such creation processes are not favored

from a thermodynamic viewpoint given that they require the concentration of very large amounts of energy from the vacuum into the two created particles.

46. Despite the fact that processes of pair creation involving opposite action particles can only occur as short-lived fluctuations on a time scale characteristic of quantum gravitational phenomena, their effects could nevertheless become permanent during the first instants of the big bang when the expansion of space takes place at a sufficiently high rate and this may effectively explain the presence of matter in our universe.
47. A negative energy matter particle cannot decay to 'lower', more negative energies by emitting positive energy radiation particles, because the positive energy radiation particles could not even have been into contact with the decaying negative energy particle.
48. A positive energy matter particle cannot turn into a negative energy particle by emitting positive energy radiation particles, given that there can be no direct interaction between the now negative energy matter particle and the positive energy radiation it would have released.
49. No interaction vertex involving particles with mixed action signs needs to be taken into account in determining the transition probabilities of quantum processes.
50. For negative energy matter the objectively defined low direction on the energy scale which is associated with the thermodynamically favored degradation of energy is that toward the zero energy level of the vacuum as is the case for positive energy matter, because such lower negative energies are associated with states of matter characterized by a larger number of microscopic degrees of freedom and a higher entropy.
51. A negative energy particle does not have a natural tendency to decay in the direction on the energy scale which is associated with more negative energies for the same reason that positive energy particles do not naturally tend to reach higher positive energies.
52. Based on the preceding results we can expect that there is no interference on the part of negative energy particles into the processes usually

described by quantum field theory except at the energy level associated with quantum gravitational phenomena.

53. The energy that is gained or lost by a positive energy body as a consequence of its indirect gravitational interaction with a negative energy body is compensated by a variation in the negative gravitational potential energy associated with the variation of positive vacuum energy that is equivalent to the variation of energy of the negative energy body (a conclusion which is valid in the context where the negative gravitational potential energy associated with the interaction of this positive vacuum energy with the rest of the matter and energy in the universe can be as large in magnitude as the positive vacuum energy itself).
54. Given that the negative energy of a field of interaction between positive energy particles in a bound system cannot be directly and independently measured it cannot be assumed to contribute independently to the inertial mass of the entangled system as a whole.
55. The perpetual motion argument against gravitational repulsion only rules out the possibility of an anomalous gravitational interaction between ordinary matter and ordinary antimatter, because while positive mass matter can gain potential energy by being raised in the gravitational field of a positive mass planet by negative mass matter, this negative mass matter would lose potential energy in the process, which means that no work can be produced in such a way.
56. Negative energy matter could not be used to provide the conditions necessary to make a traversable wormhole given that it cannot be made to remain near the singularity of a positive mass black hole or even simply be brought inside such a black hole.
57. The fact that negative energy matter cannot cross the event horizon of a black hole means that it cannot reduce the mass of the black hole and the area of its event horizon, so that the existence of negative energy matter would not allow to produce a diminution of the entropy associated with those objects.
58. Negative energy is not equivalent to negative heat for a positive energy system given that from the viewpoint of positive energy matter kinetic energy is exchanged with negative energy particles as if it was a

positive definite quantity and therefore the existence of indirect gravitational interactions between opposite energy systems does not allow the transformation of useless forms of thermal energy into more useful forms in a way that could have given rise to a reduction of entropy.

59. The observer dependence of the gravitational field which must be assumed in the context of a bi-metric general relativistic theory implies that observers with opposite energy signs experience the metric properties of space and time associated with a given local matter configuration in a different way.
60. There are two distinct categories of contributions to the total stress-energy entering the generalized gravitational field equations that determines the metric properties of space experienced by positive energy matter, the first is provided by the conventional stress-energy tensor and is positive definite for all densities of positive energy matter while the other is provided by the irregular stress-energy tensor and can be either positive or negative depending on the value of energy density of negative energy matter relative to its average density.
61. The natural value of positive and negative contributions to vacuum energy density is provided by the Planck scale and when positive energy is missing from the vacuum as a consequence of the presence of negative energy matter the energy of the vacuum is reduced from this maximum positive value.
62. In the proposed generalized gravitational field equations, the value of vacuum energy density observed by positive energy observers and associated with the cosmological constant is determined solely by the conversion factor associated with the map of the metric properties associated with negative energy matter as they are experienced by positive energy observers.
63. The values of vacuum energy density which are observed in the absence of matter by positive and negative energy observers could vary with position and time given that they arise as a consequence of applying the variable metric conversion factors which provide the map of the metric properties associated with matter of a given energy sign as they are experienced by observers of opposite energy sign.

64. Given that the cosmological constant is very small on the spatial scale at which it is observed it follows that there must be a near perfect level of symmetry between the metric properties associated with positive energy matter and those associated with negative energy matter as they are experienced by positive energy observers.
65. Even when the cosmological constant is not null the gravitational field equations describing the motion of negative energy matter are symmetric with the equations describing the motion of positive energy matter.

Chapter 3

Discrete Symmetries

3.1 The problem of discrete symmetries

In this chapter I would like to explain how a more consistent and adequate formulation of the discrete P , T , and C symmetry operations can be obtained that integrates the insights gained while studying the problem of negative energy and that offers a better understanding of why and how such symmetries can under certain circumstances appear to be violated. Discrete symmetry operations are usually assumed to be relevant only in the context of quantum field theory, but in fact they can also be examined from a semi-classical standpoint. Their level of application is actually right at the interface between the classical world of gravitation theory and that of quantum theory and it should not come as a surprise therefore that some of the results which I have obtained allow progress to be achieved concerning the problem of identifying the origin of the degrees of freedom associated with black hole entropy, which effectively arises in a semi-classical context. In order to do so it was necessary to introduce an additional category of discrete symmetry operations that relates positive and negative action matter particles in a way that is similar in many respects with that by which the charge conjugation symmetry operation relates ordinary matter and antimatter.

I had long ago realized that it would be necessary to revise our conception of space and time reversals, because the current formulation of those symmetry operations is based on unreasonable assumptions regarding the significance of time reversal and its relationship with the sign of energy and that of non-gravitational charges. It is indeed presently believed that the

charge conjugation or C symmetry operation is not a discrete space or time symmetry operation, but simply an additional symmetry having to do with charge as an independent concept. But I came to suspect that the relationships which are known to exist between this charge reversal operation and the discrete P and T symmetry operations associated with space and time reversals are an indication that C should be conceived and explicitly defined as a particular instance of discrete spacetime symmetry operation. What constitutes the underlying basis of those considerations is the acknowledgement that the observed sign of physical quantities (including charge) are dependent on their direction of propagation in time. From that viewpoint it would seem indeed that both the T and the C symmetry operations should be assumed to involve some form of time reversal and this is reason enough to suspect that they may also both give rise to a reversal of charge.

The problem, however, does not really have to do with the current conception of the charge reversal operation as such. What is truly inappropriate is the simple kinematic representation of time reversal as involving a backward motion of all particles and their angular momenta, which I believe is too rudimentary to characterize a reversal of the fundamental time direction degree of freedom. I also think that if T is to be assumed as effectively reversing time then it should leave momentum unchanged (despite common expectations) as this is a quantity that should rather be reversed independently, along with the direction of space intervals. In this context if some reversal of momentum may still be of relevance to T it would clearly have to arise as a consequence of the fact that it is effectively equivalent to the effects we should expect from an appropriate reversal of time when we insist on measuring physical quantities against the perceived rather than the actual direction of the flow of time. In any case it must be understood that what we observe from our classical historical perspective is not representative of the true evolution that takes place when we are dealing with the propagation of elementary particles. The subtleties of what is going on at the microscopic level are not directly apparent from the superficial viewpoint associated with a global representation of events ‘after the fact’ that provides a static picture of the spacetime paths followed by elementary particles. Therefore, it is not appropriate to define a reversal of the fundamental (non-thermodynamic) time direction degree of freedom based merely on narrative aspects of phenomena which are all directly discernible at this superficial level of description. Better formulations of the discrete spacetime symmetry operations are required which would reflect the actual and sometimes ignored variations or absence

of variation of physical parameters associated with each of those reversals of the fundamental space and time direction degrees of freedom.

3.2 The constraint of relational description

To begin this discussion, I must first of all mention that once again the most significant constraint which we need to consider and against which our understanding of the discrete symmetry operations must be developed is that of the necessary relational definition of physical quantities and their changes. Those quantities are here the directions of space and time intervals, the directions of momentum and angular momentum and the signs of energy and non-gravitational charges. The main point I want to emphasize is that there can be no meaning in considering a change of any one of those quantities (to its opposite value) that does not occur relatively to some remaining unchanged parameter of the same kind. Breaking that rule is to be considered logically impossible simply because if it was allowed it would mean that we can define an absolute (metaphysical) direction or polarity (in the general sense), which in effect would not be related to any reference point of a physical nature in our universe. What I am suggesting is that the profound reason why a certain level of lopsidedness, such as the observed breaking of P symmetry by the weak interaction, can exist is that such asymmetries merely occur when one or two physical parameters are reversed *relative* to a fixed background of unchanged directional parameters of a similar kind. In other words, what makes these violations of discrete symmetry possible is simply the fact that application of a reversal operation to a single parameter leaves some other properties unchanged which allows the asymmetry to occur as a real feature characterized by a measurable change relative to a distinct physical quantity. In the case of P symmetry, the reversal of space intervals involved occurs relative to the direction of time intervals which remain unchanged by such an operation and therefore it should be expected that violations of P can be observed given that the reversal of physical parameters associated with this operation can be measured against the unchanged properties.

But those asymmetries cannot imply the existence of an absolute lopsidedness or directionality at the most fundamental level for the universe as a whole, because they can be compensated by an appropriate reversal of the unchanged parameters relative to which the original transformation took

place. This is what explains that despite the violation of P symmetry by the weak interaction it remains impossible to provide an absolute definition of left and right, because indeed reversing the sign of charges allows to regain invariance. Thus, contrarily to what is sometimes assumed, the preferred handedness unveiled by the weak interaction is not more profound than that we observe in certain complex structures. As long as invariance under a more general discrete symmetry operation like CP is observed to hold it is impossible to communicate the significance of right and left without knowing which of two C -related particles is to be considered as having positive electric charge. But if it is impossible to distinguish an absolute (non-relational) difference between positive and negative charges themselves, as I previously suggested, then only observers which are effectively sharing the same universe and which are allowed to directly compare physical quantities, could differentiate between left and right. This is a very general feature which I think would always be observed to apply given that it is actually required by the condition of relational definition which is relevant to any change of direction or polarity (such as a reversal of the sign of charges). The directions of space and time which are singled out by any process which appears to violate a discrete symmetry can have meaning only in relation to other aspects of reality which must be identifiable from within the universe in which those processes take place. If in one particular instance it was to be found that no combination of discrete symmetry operations allowed invariance to be regained, then it would mean that there exist physical properties which can refer to elements of reality not shared only by observers within our universe. In other words, if directional asymmetries not occurring merely in relation to unchanged quantities (not defined as relative properties) were allowed, it would effectively be impossible to describe the polarities so revealed by referring only to measurable properties of physical reality.

The problem which there would be if such violations of discrete symmetry were possible is that completeness and self-determination are the defining characteristics of the universe concept, in the sense that the universe is precisely that ensemble of physical elements which are all causally related to one another and to nothing else. Thus, if we were to find that the description of our universe can refer to absolute and immaterial notions of direction not defined merely as relationships between elements of reality which must be part of that universe, then the only logically valid conclusion would have to be that there exists a *causally related* reality outside what we consider to be the universe (this has nothing to do with the concept of the multiverse

whose elements are not to be assumed as causally related to one another) relative to which the otherwise metaphysical polarities could be properly defined. As a consequence, there is definitely no way our universe could be considered lopsided if it is effectively the whole universe and I believe that the fact that it can be shown that the existence of such an irreducible asymmetry would imply that some physical quantities may not be conserved for the universe as a whole is a confirmation of the validity of this conclusion. It must be understood, however, that the identified requisite does not mean that symmetry could never be preserved following a reversal of one single parameter, like space direction alone, which can be defined in a relational way, but simply that such invariance is not absolutely required to apply under all circumstances.

Given those considerations, we can be totally confident that there is no such thing as an absolute direction of space or time intervals, because indeed this would imply a violation of the principle of relativity (as understood in its most general form which predates relativity theory) and the validity of this criterion is necessary for the consistency of any model concerning physical reality. Even without going into elaborate mathematical arguments, such as those entering the CPT theorem, it is therefore possible to appreciate that the only problem there could be in relation to the observation of an asymmetry under a properly defined discrete symmetry operation would have to involve a violation of invariance under a combined operation that reverses all parameters and leaves absolutely none unchanged. I will later explain why an appropriately defined *PTC* transformation must be considered as one instance of such a symmetry operation that reverses all parameters and leaves nothing unchanged (by actually reversing all space- and time-related parameters twice) and which we are thus justified to categorize as inviolable. But I believe that the fact that it would be impossible to provide a mathematical framework for quantum field theory that would satisfy the requirements set by special relativity if the equations of the theory are not invariant under *PTC* (which constitute the substance of the argument behind the traditional CPT theorem) confirms that relativistic imperatives (all measures of space and time intervals are relative) are the true constraints which impose invariance under the most general, combined, discrete symmetry operation.

The fact that this simple but most unavoidable requirement has never been considered as a means to restrict allowed violations of discrete symmetry illustrates the fact that our treatment of space and time reversals is incomplete and inadequate due to multiple misconceptions which do not

concern only the aspect discussed here. The often met remarks to the effect that there is no a priori reason why the universe could not be asymmetric in a fundamental way and that it is only the above mentioned mathematical requirements arising from the CPT theorem that motivate the conclusion that some overall symmetry must nevertheless be obeyed under all circumstances are therefore inappropriate and misleading. But it should also not come as a surprise that the discrete symmetry operations, when performed independently from one another, may not produce invariance. What justified the unexpectedness of the violations of P and CP symmetries when they were first observed is actually the intuitive belief that absolute directionality should not be allowed, while, as I just explained, this is rather the argument that would apply to a more general symmetry operation like PTC whose required conservation, ironically, is usually not believed to be intuitively explainable. The truth is that for an imbalance under reflection to exist all that is required is that the world must be unbalanced with respect to something. This conclusion is the outcome of the most unequivocal interpretation of the requirement of relational definition of physical quantities, which itself constitutes the one rule we can be most confident need to apply to the physical world we experience. In fact, the argument against the possibility of a violation of symmetry under a combined reversal of all space- and time-related parameters is probably the strongest kind of argument which can be proposed from a theoretical viewpoint.

Regarding time reversal in particular and the question of what it would mean to assume that the whole universe is running backward in time and whether there can be any objective meaning to such a reversal operation I think that given the preceding discussion we would have to recognize that such a reversal could effectively be physically significant if it is defined as a reversal that leaves other parameters, such as the direction of space intervals unchanged. But this means that such a time reversal operation cannot consist in a mere reversal of the motions and rotations of objects taking place in a reverse chronological order. A reversal of time that would be relationally defined would have to be meaningful both globally and locally as it would allow a distinction between a physical system with unchanged time direction and one with reversed time. This difference could be determined by directly comparing the physical properties of one of the systems with those of the other, if the two systems are part of the same universe. But a difference could also be identified as occurring for the whole universe in relation to the unchanged direction of space intervals. In any case the above discussed

constraint would require that such a relative backward in time evolution be clearly identifiable from the physical properties of the particles involved, precisely because it is only under such conditions that the change of direction in time could be objectively determined by comparing it with that of the unchanged parameters. But given that those differences would then effectively be determined in relation to the value of parameters which are themselves reversible, it follows that no absolutely characterized notion of asymmetry would be involved.

In the context where absolute lopsidedness is to be considered impossible it follows that it is of primordial importance to identify all the physical properties which can be related to one another and which could be affected by transformations of the kind that involve a reversal of space and time directions at the fundamental level. Indeed, if we are to be able to determine whether there remain quantities not reversed when a certain discrete symmetry operation is performed, we certainly have to be able to determine which quantities are effectively affected by the operation involved. It is my belief that a portion of the violations of discrete symmetries which are usually assumed to be observed are actually a consequence of the fact that the effect of the considered reversals on some quantities are ignored, while invariance would actually be inferred if all quantities dependent on the parameters which are assumed to be reversed were appropriately transformed. I already mentioned the fact that there are indications to the effect that we may, in particular, expect the sign of charges to be dependent on the sign of time intervals experienced by the particles carrying them. Yet the traditional definition of the time reversal operation T does not involve any reversal of charges (from whatever viewpoint) and thus we could observe violations of such a T symmetry that would occur simply because we do not appropriately reverse the sign of charges when we try to verify invariance under a reversal of time (from a certain viewpoint). We must therefore first take care of identifying all unaccounted dependencies which may confuse our assessment of symmetry violations before we can truly appreciate under which conditions they are effectively allowed to occur.

3.3 The concept of bidirectional time

Concerning the problem of discrete symmetries another essential aspect must be recognized in addition to that regarding the necessity of a relational def-

inition of all such symmetry operations. Awareness of what it involves is of the highest importance for a proper resolution of all matters associated with time directionality and given that this is the central problem with which this report is concerned it is crucial to grasp the significance and the implications of the notions involved. Basically, what must be understood is that a distinction is to be made between the traditional concept of time direction associated with changes occurring at a statistically significant level where the notion of entropy is meaningful and a concept of time direction associated with the existence of a fundamental time direction degree of freedom independent from the constraints related to entropy variation. The traditional concept of time direction related to statistically significant changes and the growth of entropy gives rise to what I call the unidirectional or thermodynamic time viewpoint, while the alternative concept of time direction related to the existence of a fundamental time direction degree of freedom independent from statistical constraints gives rise to what I call the time-symmetric or bidirectional time viewpoint.

Associated with this alternative concept of time direction is a different notion of time reversal not limited by the constraints imposed on our description of physical processes by the second law of thermodynamics. Indeed, the traditional notion of time reversal associated with the thermodynamic time viewpoint merely consists in assuming a reversal of the motion of all particles involved in a process, so as to give rise to the same events as observed in the original process, but in the reverse order. However, those events would still be described from the same unique and immutable forward direction of time associated with entropy growth. This is a consequence of the fact that the unidirectional time viewpoint involves considering that there can only be one direction in time at once for the propagation of all particles, indiscriminately, which effectively amounts to ignore the existence of a fundamental time direction degree of freedom. From that viewpoint if time was reversed all particles would have to propagate backward, not relative to some fundamental time direction parameter, but in comparison with the direction of motion which they were all *observed* to have originally. Thus the time reverse of a process would simply be the equivalent process for which the same observations are made, but in the reverse order. The bidirectional or time-symmetric viewpoint on the other hand is at once less restrictive and more distinctive in that it actually recognizes the existence of a fundamental time direction degree of freedom, distinct from the observed direction of motion of particles apparent to an observer constrained by the law of entropy increase.

This time direction parameter must be allowed to vary from one particle to another, even between those of an otherwise identical nature which are involved in the same process at the same time.

Now of course I have already discussed the significance of the existence of a fundamental time direction degree of freedom as being that property which allows to explain the distinction that exists between a particle and its antiparticle, despite the fact that from an observational viewpoint both objects appear to be ordinary particles travelling forward in time, but which merely happen to carry opposite non-gravitational charges. However, I had previously made clear that in fact the sign of charge is *not* affected by any time reversal operation which may relate a particle with its antiparticle and therefore if it is nevertheless observed as being reversed it can only mean that the direction of time relative to which we measure the charge is itself reversed in comparison with the true direction in which the particle is propagating in time, which introduces a relative change in the sign of charge that would not occur for an observer measuring the same property while following the true direction of propagation in time of the particle¹. It is merely the fact that a backward in time observation is indeed impossible that justifies assuming a reversal of charges for a particle propagating toward the past. Indeed, measuring apparatuses always record changes as they occur in the future direction of time due to the fact that the processes involved in the amplification of the signal which gives rise to a measurement can only take place in this direction of time in a universe where a thermodynamic arrow of time governs the evolution of processes involving a large number of independently moving particles. This constraint is therefore what justifies the use of a unidirectional viewpoint relative to which all physical properties are given as they would appear relative to the conventional future direction of time, even when the true direction of time in which the processes involved occur is the past direction. Non-gravitational charges, therefore, effectively remain unchanged from the bidirectional viewpoint when the fundamental time direction degree of freedom is reversed, but this is the very reason why they appear to be reversed from the unidirectional time viewpoint.

¹I will henceforth use the term ‘propagation’ in place of ‘motion’ to designate the true direction in which a particle is traversing space and time intervals, as occurs from a bidirectional time viewpoint. This allows to explicitly refer to those aspects associated with the fundamental time direction degree of freedom which are ignored from the viewpoint of unidirectional time relative to which all changes refer to a particle’s observed (semi-classical) trajectory.

A rule thus emerges which is that, for any particle propagating in the past direction of time, a time direction-dependent physical property of that particle which would be positive when considered from the bidirectional time viewpoint (relative to the true direction of propagation of that quantity in time) would appear as negative from the unidirectional time viewpoint. But this reversal of observed quantities from their true value is not restricted to charge or energy, which I had already identified as properties dependent on the direction of propagation in time, but would actually have to apply to the direction of space intervals associated with the motion of particles (which are always given in relation to time intervals) and thus also to momentum (even if the time intervals entering the traditional definition of momentum were assumed positive definite as a consequence of adopting a unidirectional time viewpoint). Thus, if momentum was assumed to be left unchanged by a properly defined reversal of time (on the basis of the fact that from a fundamental viewpoint the associated direction of space intervals is an independent parameter to be reversed by an independent symmetry operation, as I will later explain), it would nevertheless appear to be reversed in comparison with its actual value, from the unidirectional time viewpoint. But given that the direction of momentum is not fixed for a given type of particle propagating in a given direction of time (it also changes when the direction of propagation of the particle in *space* is reversed) it cannot be taken as a clear indicator of the direction of propagation in time of a particle. That, however, is not the case with charge which from the bidirectional time viewpoint remains unchanged even as a particle reverses its direction of propagation in time (while also reversing its energy sign) and this is why it is possible, from the unidirectional time viewpoint, to identify the true (even if merely conventional) direction of propagation in time of a particle based on the observed value of its non-gravitational charges (in relation to those of an otherwise identical particle).

What is important to understand is this interdependence of space and time intervals even as they would be separately and independently transformed by their respective discrete symmetry operations. Thus, when we reverse the direction of the motion of a particle in space we reverse the sign of the space intervals associated with this motion not merely relative to the space axes, but also relative to time intervals (same time interval, opposite space interval). The sign of space intervals associated with the propagation of a particle submitted to a reversal of space directions would be reversed not merely from what it previously was (or relative to the space intervals

associated with the motion of a particle not subject to the reversal), but also relative to the direction of time intervals in which the particle is still propagating. A particle which was propagating to the right relative to the future direction of time will now be propagating to the left relative to the same future direction of time, which was not affected by the reversal of space directions (this is illustrated in figure 3.1 where I consider the effects of the various discrete symmetry operations as they will be defined below). In other words, the particle is not just propagating left, it is propagating left forward in time, because indeed we are always concerned with the properties of processes involving particles propagating in space and time and not just with the properties of space or time themselves. What matters therefore is not just the direction of space intervals associated with some arbitrarily fixed spatial coordinate system, but the direction of space intervals for a particle propagating in a given direction of time, as asserted from a fundamental bidirectional viewpoint. Similarly, when time is assumed to be reversed it must be considered that the time intervals are reversed relative to the unchanged direction of space intervals in which a particle submitted to the reversal is propagating, so that the same positive space intervals are now travelled in the opposite direction of time. This does not mean that a reversal of both space and time cannot have clear meaning, however, because as I will explain later, even in such a case there would still remain unchanged physical properties relative to which the transformation could be characterized.

This relationship between space and time intervals is what gives a true physical meaning to the notion of time reversal when it is to be considered as a symmetry operation clearly distinct from space reversal and which should therefore leave momentum unaffected (from the bidirectional viewpoint at least). In fact, it is what allows the very notion of a fundamental degree of freedom associated with direction in time to have a definite meaning, because it allows to distinguish (as a theoretical possibility) the process by which a particle is going through a given spacetime trajectory forward in time from the similar process by which an identical particle would be going through the exact same spacetime trajectory, only now backward in time. Such a distinction is crucial given that if we were to ignore it then from a unidirectional viewpoint in time there would be no meaning to assume that it may be possible for a trajectory to be traversed backward in time, given that from such a viewpoint we always observe particles as if they were necessarily going forward in time. But given that charge can be assumed to be left unchanged by a reversal of time (from the bidirectional viewpoint) we

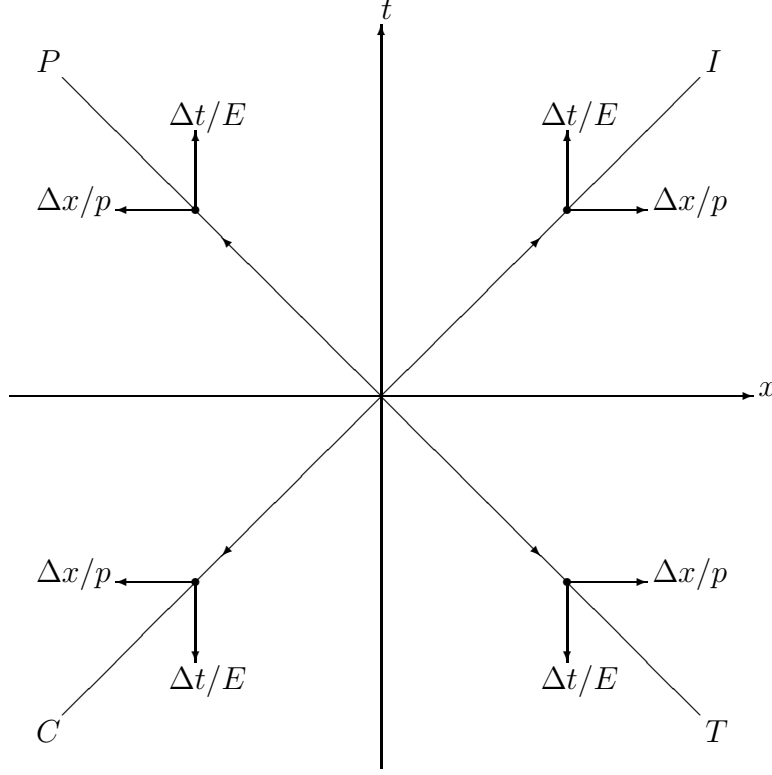


Figure 3.1: Variation of physical parameters under the proposed alternative definition of P , T , and C as described from the bidirectional time viewpoint. In this figure and the other related figures, I represents the original state and the diagonal lines correspond to particle trajectories. The space and time intervals Δx and Δt are indicated by vectors whose lengths correspond to the magnitude of the intervals and whose directions indicate the sign of the intervals relative to the space and time coordinates. The direction of the vectors associated with the energy E of particles corresponds with the sign of energy relative to the direction of the time coordinate.

are effectively allowed to differentiate between those two situations from an observational viewpoint, even in the context where all particle trajectories are necessarily followed as if they were occurring in the ‘normal’ chronological order (forward in time) associated with the growth of entropy, regardless of the true direction of propagation in time of the particles. It is therefore the relation between space intervals and time intervals that allows to distinguish backward in time propagation from forward in time propagation and the fact that the observed value of the sign of charge is dependent on that distinction simply confirms that it is appropriate to consider the existence of such a directionality parameter for the time dimension at the fundamental, elementary particle level.

It must be clear, however, that the coordinate systems for space and time still have a physical significance, because you may reverse the direction of the space intervals travelled by particles in the forward direction of time as well as the associated momenta while keeping the positions of the particles in space unchanged (not reversed as they would under a conventional space reversal operation). Indeed, as a comparison of figures 3.1 and 3.2 allows to reveal, it is only from the bidirectional time viewpoint that the sign of space and time *intervals* corresponding to the directions of propagation of particles always change in association with the sign of *positions* on the space and time coordinate axes, while from the unidirectional time viewpoint that need not be the case. Under such conditions quantities like angular momentum, which depend on both the position in space and the direction of space intervals in time, may not always be left invariant as they would when a complete space reversal operation is performed. This would effectively occur for processes submitted to a reversal of time when they are described from the unidirectional viewpoint in which time is maintained positive even for backward in time propagating particles and all time direction-dependent quantities like the direction of space intervals and the momentum of a particle consequently appear to be reversed, while the positions are left unchanged (which implies that spin would appear to be reversed). In this context it seems that space intervals, as properties defined in relation to the direction of propagation in time, can effectively be reversed in two different ways. They may be reversed because space directions are reversed (which also reverses positions) or they may be reversed because the direction in which they are assumed to be traversed in time is reversed (which leaves positions unchanged). This distinction is what allows the traditional concept of time reversal as affecting the directions of momentum and angular momentum to still be relevant,

even in the context of the existence of a fundamental time degree of freedom, when those directions should in fact be left invariant (from a bidirectional viewpoint) by a properly defined time reversal operation.

Another point must be emphasized regarding the kind of time reversal operation which can be developed in the above described context. Indeed, if we no longer consider appropriate the picture of time reversal as consisting in a simple reversal of the observed motion of each and every particle then it must also be recognized that a properly defined time reversal operation could never give rise to a reversal of the thermodynamic arrow of time for the physical systems involved. In fact, I think that we should already suspect that there is something wrong with the often met suggestion that a reversal of the motion of every particle in a region of space would give rise to entropy decreasing evolution (in the absence of any external perturbation). For such a proposal to be valid it would have to be shown that the origin of the observed time-asymmetry of thermodynamic processes in our universe is to be found in a very precise adjustment of the motion of every single particle in the universe at the present time which would occur in just such a way as to allow a state of minimum entropy to be reached as time unfolds in the past right back to the big bang state. However, given the inherently random nature of quantum processes and the extreme sensitivity to initial conditions (here the ‘final’ conditions giving rise to a given past evolution) which are known to exist even in a classical context, this hypothesis appears highly implausible. But if in addition we admit the existence of a fundamental time direction degree of freedom distinct from the observed motion of particles then we clearly have to reject the possibility that a reversal of time may produce anti-thermodynamic behavior, because time-reversed propagation is in fact already taking place in processes for which there is no apparent change to the direction of the thermodynamic arrow of time. This means that the direction of propagation in time of particles (the sign of time intervals associated with a bidirectional viewpoint) is not necessarily that relative to which entropy increases despite the fact that it may appear unnatural that evolution could proceed in a direction of time other than that in which we do observe time to be ‘flowing’ (as a thermodynamic necessity). The thermodynamic arrow of time and the notion of time directionality occurring from a bidirectional viewpoint are two completely independent concepts.

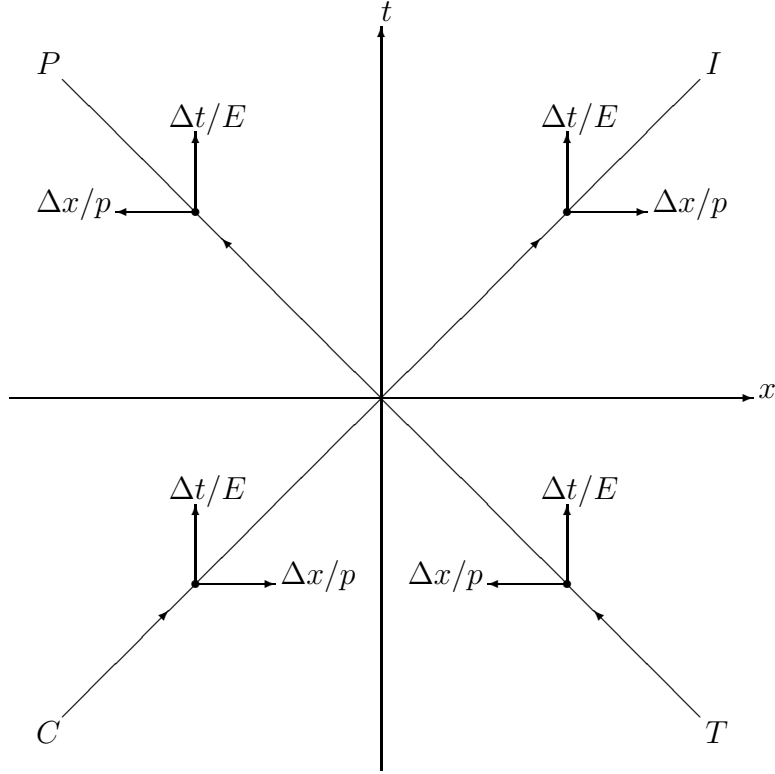


Figure 3.2: Variation of physical parameters under the proposed alternative definition of P , T , and C as apparent from the unidirectional time viewpoint. We can see that from this viewpoint the only difference between the original process and the T -reversed process is that the space intervals are traversed in the opposite direction, just as would be expected according to the traditional definition of backward in time motion. The case of the C -reversed process is also quite in line with traditional expectations given that such a process should not be different from the original process except for a reversal of the sign of charges (which is not illustrated here) which would in fact also occur for the T -reversed process despite traditional expectations.

3.4 Alternative definition of C , P , and T

One last remark is necessary before I can provide a full description of exactly how the fundamental physical properties of matter should be considered to vary under an alternative set of discrete symmetry operations formulated so as to allow the above discussed requirements to be satisfied. I previously hinted at the fact that the direction of momentum should be considered as independent from the direction of time at least from the most consistent viewpoint which is provided by a bidirectional perspective on time. I believe in effect that momentum, as the property conjugate to physical space, should only be considered to reverse along with space and not along with time, just as energy being the physical property conjugate to time should necessarily reverse when time reverses and only then. There is, however, an additional motivation for requiring this kind of joint variation of all space-related properties or time-related properties (independently) besides the fact that consistency may require that it be imposed when what we seek to assert is precisely the dependence of various parameters under reversal operations which are defined after the quantities they are assumed to reverse. This perhaps more unavoidable justification for the joint variation of conjugate quantities is to be found in the requirement that the considered symmetry operations should not change the sign of action of the physical systems on which they operate.

It is my understanding of the true physical significance of a reversal of the sign of action that allows me to recognize the necessity to define the discrete symmetry operations in such a way that momentum would necessarily reverse as a consequence of a reversal of space coordinates while energy would necessarily reverse as a consequence of a reversal of the time coordinate. Indeed, in the context where a reversal of space coordinates would necessarily give rise to a reversal of space intervals, while a reversal of the time coordinate would necessarily give rise to a reversal of time intervals, if the sign of action itself is to remain invariant then it means that a reversal of space must also involve a reversal of momentum and a reversal of time must also involve a reversal of energy. In fact, we always implicitly assume that the P , T , and C reversal operations do not relate physical processes in which the particles involved would have opposite action signs or energies (as measured from the forward direction of time). But the implications this should have for the dependence (under conventional discrete symmetry operations) of the signs of momentum and energy on those of space and time intervals is not always

recognized. I believe that this lack of clarity is responsible for a good part of the misunderstanding regarding what parameters should really be affected by any symmetry operation involving a reversal of time. In tables 3.1, 3.2, 3.3, and 3.4 I will therefore provide an explicit account of the dependence of the signs of momentum and energy, along with those of space and time intervals, under all relevant discrete symmetry operations. It will be apparent from this account that clear distinctions exist between the traditional and the redefined time reversal and charge conjugation symmetry operations. Yet, given that the original definitions actually need to be replaced and cannot even be considered meaningful anymore, I think that it will not be necessary to relabel those operations and associate them with new symbols or letters, so that I will continue to use the T and C notation when referring to those redefined discrete symmetry operations.

In the following tables and in the corresponding diagrams (figure 3.1 corresponds to table 3.3 and the bidirectional viewpoint, while figure 3.2 corresponds to table 3.4 and the unidirectional viewpoint) the position along the space and time axes are denoted x and t (I am assuming a one-dimensional space for simplicity) while the space and time intervals corresponding to the motion, or the propagation of the particles involved in the processes which are transformed by the symmetry operations are denoted Δx and Δt respectively. The energy of the particles involved in the same processes is denoted E and can effectively vary in sign, while the momentum of those particles along the x axis is simply denoted p . The sign of non-gravitational charges (which allows to distinguish between the state of a particle and that of its antimatter counterpart), even though it should be understood not to be reversed by any of the discrete symmetry operations (including C) from the bidirectional time viewpoint (which provides the most accurate description of the transformations involved), is nevertheless defined as q and can effectively appear to be reversed from the unidirectional viewpoint. The sign of angular momentum related to the motion of the particles involved in the processes transformed by the P , T , and C operations, as well as the spin direction of elementary particles, which again should be understood not to be affected by those operations from a bidirectional time viewpoint are together denoted by the letter s , while the associated parameter of handedness (the direction of spin along the axis associated with the momentum of a particle) is here denoted h and should be expected to vary, even from a bidirectional time viewpoint.

From a semi-classical viewpoint, the displayed tables giving the varia-

Trad.	t	Δt	E	x	Δx	p	q	s	h
I	t	$?$	E	x	$?$	p	q	s	h
P	t	$?$	E	$-x$	$?$	$-p$	q	s	$-h$
T	$-t$	$?$	E	x	$?$	$-p$	q	$-s$	h
C	t	$?$	E	x	$?$	p	$-q$	s	h

Table 3.1: Variation of the physical parameters associated with a process transformed by the discrete P , T , and C symmetry operations as they are traditionally defined. The absence of explicit assumption concerning the Δt and Δx parameters (specifically) can be noted. The variation of the direction of angular momentum s as well as that of the handedness h can be derived from those of the other fundamental parameters, but the outcomes are nevertheless indicated in the tables because they may differ from what is traditionally expected. The identity operation I which corresponds to an absence of reversal is shown for reference purpose.

Impl.	t	Δt	E	x	Δx	p	q	s	h
I	t	Δt	E	x	Δx	p	q	s	h
P	t	Δt	E	$-x$	$-\Delta x$	$-p$	q	s	$-h$
T	$-t$	Δt	E	x	$-\Delta x$	$-p$	q	$-s$	h
C	t	Δt	E	x	Δx	p	$-q$	s	h

Table 3.2: Implicitly assumed variation of physical parameters under the discrete P , T , and C symmetry operations as they are traditionally defined. The parameters whose transformation is only implicitly assumed are the space and time intervals Δx and Δt associated with the propagation of the particles involved in the processes transformed by the various discrete symmetry operations. The absence of reversal of Δt when time is assumed to be reversed can be noted.

Bidir.	t	Δt	E	x	Δx	p	q	s	h
I	t	Δt	E	x	Δx	p	q	s	h
P	t	Δt	E	$-x$	$-\Delta x$	$-p$	q	s	$-h$
T	$-t$	$-\Delta t$	$-E$	x	Δx	p	q	s	h
C	$-t$	$-\Delta t$	$-E$	$-x$	$-\Delta x$	$-p$	q	s	$-h$

Table 3.3: Variation of physical parameters under the redefined discrete P , T , and C symmetry operations as described from the bidirectional time viewpoint. The necessary reversal of Δt with E as well as that of Δx with p can be noted, as also the necessary reversal of t with Δt and that of x with Δx . This is the variation of physical parameters which would be produced by the most appropriately defined discrete symmetry operations that can be formulated in a semi-classical context. Here all reversals of physical quantities are seen to occur twice or to not occur at all, as required for explicit invariance under a joint PTC operation.

Unidir.	t	Δt	E	x	Δx	p	q	s	h
I	t	Δt	E	x	Δx	p	q	s	h
P	t	Δt	E	$-x$	$-\Delta x$	$-p$	q	s	$-h$
T	$-t$	Δt	E	x	$-\Delta x$	$-p$	$-q$	$-s$	h
C	$-t$	Δt	E	$-x$	Δx	p	$-q$	$-s$	$-h$

Table 3.4: Variation of physical parameters under the redefined discrete P , T , and C symmetry operations as described from the unidirectional time viewpoint. Again all quantities are reversed either twice or never by a combination of all operations, which guarantees explicit invariance under PTC . The equivalent reversal of charge q by both T and C as well as the apparent absence of any variation of Δt and E and the absence of joint variation of x and Δx when t is reversed can be noted.

tions of the time-related and space-related physical parameters under the traditional or redefined discrete symmetry operations, along with the assumptions which are made concerning the variation of the sign of charge, provide the most precise definitions that can be achieved of the operations involved. Using those definitions one can rebuild the quantum operators which are needed to transform the state vectors or the propagators corresponding to specific quantum states or processes. It must be clear that quantum field theory itself does not dictate how the discrete symmetry operations should be defined and it is merely the assumptions used while formulating the related operators (to achieve transformations that match our expectations regarding which parameters should be affected by a given operation) that provide the necessary constraints on which depend their precise mathematical formulation. What I bring to the table, therefore, is an improved knowledge of the constraints that must apply to those transformations, based on a reexamination of the meaning of space and time reversals as they would occur in a semi-classical context. It is important to recognize indeed that despite the apparent freedom, the discrete symmetry operations cannot be arbitrarily defined, but must be the outcome of the most unavoidable consistency requirements (formulated in an empirically motivated context) which I believe are those I have identified in the above discussion. The fact that greater simplicity has been achieved while redefining those symmetry operations is only a further confirmation of the appropriateness of the alternative viewpoint that emerged from the preceding analysis. Indeed, the pattern of variations of physical parameters which is illustrated in figure 3.1 is strikingly simple in comparison with that we would have according to the traditional definition of the discrete symmetry operations and this simplification was actually one of the objectives I sought to achieve while redefining them. Let me then describe what the elegance of this proposal really embodies.

Looking at the tables in which the outcomes of the various discrete symmetry operations are displayed one thing we may first remark is that the parity operation P remains as it was originally defined, even in the context of the proposed alternative formulation of those transformations and this regardless of whether we use the bidirectional or the unidirectional time viewpoint. Of course the reversal of space intervals associated with the propagation of particles (which from my viewpoint must occur as a result of the reversal of space coordinates) is now more explicitly stated, but otherwise the traditional definition of space reversal remains unchanged. There is one good reason for that, which is that the revision I am operating regards the

concept of time direction essentially and the P operation is unique for being the only one that does not involve any time reversal, regardless of the approach favored. This is what explains that this operation was properly defined already, in the form it originally was, despite the failure of the traditional viewpoint in general. What P expresses indeed is a reversal of space coordinates that produces a reversal of positions, space intervals and naturally also momentum (as a requirement of action sign invariance) while it leaves unchanged (now as a matter of definition) the position in time, the time intervals and the sign of energy. No reversal of charge is to be observed in this case (particles are not replaced by antiparticles), from any perspective, because there is no time reversal involved from a bidirectional viewpoint and thus no change to be associated with the adoption of a unidirectional time viewpoint. There is no reversal of angular momentum either (because both momentum and position are together reversed), which is appropriate given that if angular momentum or spin were reversed a forbidden reversal of action would occur from the bidirectional viewpoint (because spin has the dimension of an action) that would not be associated merely with the shift to a unidirectional time viewpoint. But again this is in perfect agreement with traditional expectations regarding the effects of P . Handedness is to be assumed reversed by such a reversal of space, however, because momentum is reversed while spin is left invariant from all viewpoints.

It should be noted that the explicit mention of a reversal of space intervals Δx under a symmetry operation like P does not mean that a reversal of space intervals must be assumed to occur in addition to that produced by the reversal of space coordinates. In other words, if the space intervals are effectively reversed it is merely as a consequence of the reversal of space coordinates, as otherwise there would be no real change in the direction of space intervals, that is, no change relative to the new coordinates. We may in fact consider it more appropriate to assume that it is the intervals themselves which are reversed along with the position of particles while the coordinates remain unchanged, which would still be equivalent to reversing the coordinates themselves. If I choose to explicitly mention a reversal of space intervals, along with the assumed reversal of positions, it is because there may be situations where the intervals would be reversed independently from the positions on the coordinate axes and we must be able to distinguish between the two situations. What the explicit statement of a reversal of Δx should be understood to imply, therefore, is that there must occur a reversal of the sign of space intervals traversed by the particles involved in the re-

versed processes in comparison with the sign of space intervals experienced by particles involved in processes which would not be submitted to the reversal. Those space intervals, therefore, are those which are traversed during unchanged time intervals and which we may ordinarily associate with the directions of the momenta of the particles involved. Indeed, the reversal of space intervals associated with the motion of particles is usually assumed to be implied by the reversal of momentum itself, but given that I will later suggest that momentum can be reversed without space intervals being equally reversed (when action is to be considered reversed) then it becomes necessary to explicitly define the variation of space intervals under P and to recognize that momentum direction is an independent quantity whose specification is not sufficient to determine the sign of space intervals spanned during a given time interval (except if action sign is in effect required to be invariant).

It must be recognized therefore that the reversal of Δx is not *merely* a reflection of the reversal of space coordinates, but is also a manifestation of the physical changes that occur when a particle reverses its direction of propagation in space while retaining its direction of propagation in time and those changes would be significant even if the position in space was to itself remain unchanged. Likewise, what the specific statement about the reversal of momentum p under space reversal P is intended to mean is that the direction of momentum is now the opposite of what it was, not merely relative to the new coordinates, but also relative to the directions of the momenta of particles which would not be subject to the symmetry operation. I may add that the same remarks would apply to time intervals Δt and the sign of energy, because if the reversal of those physical parameters under the T and C operations (from a bidirectional viewpoint) can be understood to occur as a consequence of the reversal of the time coordinate, it is clear that it also arises in relation to the time intervals experienced by particles which would be left unaffected by the reversal.

3.5 The time reversal operation

Despite a concordance of the rules for deriving the variation of physical parameters under any one of the redefined discrete symmetry operations there are important differences between the case of time reversal T or charge conjugation C and that of space reversal P and this is reflected in the fact that those two symmetry operations would produce results which are unexpected

from a traditional viewpoint. In the case of T it must be required in effect that the physical time intervals Δt associated with the propagation of elementary particles and the energy E be together reversed when the time coordinate is reversed (if action is to remain positive when it already is), while it is traditionally assumed (even if only implicitly) that both energy signs and *bidirectional* time intervals are in fact unchanged by T despite the reversal of the time coordinate. Also, it must now be assumed that there is no *a priori* reversal of the space intervals Δx and momentum p when time is reversed (which is allowed when those parameters are recognized as independent from the time-related parameters Δt and E). This is required despite the fact that traditionally momentum is assumed to be dependent on time intervals (I will explain below how this apparent contradiction is to be resolved). In fact, the traditional assumption that p would be reversed by T , while the position x on the space axis would remain unchanged, would be problematic if in this context we did not again implicitly assume an independent reversal of physical space intervals Δx by presuming an invariance of the sign of action.

What must be recognized therefore is that from a consistent bidirectional viewpoint, when the time coordinate is reversed it must be assumed that the time intervals of propagating particles (associated with the fundamental time-direction degree of freedom) along with their energies (as defined relative to the true direction of propagation in time) are reversed, while momentum and space intervals are left unchanged, just like a reversal of space coordinates is assumed to imply a reversal of the space intervals and momenta, but no change to energy sign and no reversal of time intervals. This independence of space- and time-related physical parameters (from one another) is a requirement of the constraint of relational definition of those quantities which imposes that something remains unchanged when T or P is applied and those invariant properties are in fact the spatial directions themselves (when the direction of time is reversed) or the direction of time itself (when space directions are reversed).

Now, if we appropriately assume that the spatial positions, the space intervals, and the momenta remain unchanged under a properly defined time reversal operation it follows that the spin and the handedness must also remain invariant. Those relationships may appear unnatural (spin is usually considered to be reversed under a reversal of time), but from a bidirectional time viewpoint they are perfectly acceptable and in the context where we want to define time reversal as effectively affecting time-related parameters

in a specific way they actually constitute unavoidable requirements. What's more, the discussed invariance is derived from the bidirectional time viewpoint according to which the values of physical properties are such as they would appear to an observer following the direction of propagation in time of the particles involved in the processes submitted to this reversal. But from a unidirectional time viewpoint (of the kind that is required from a practical perspective) the only quantities which would appear to be left unchanged when time is reversed would actually be the time intervals Δt and the energies E , because they would be submitted to twice the same reversal, once as time-related quantities and once as a consequence of the additional reversal occurring when we are forcing a forward in time perspective. This is what justifies the validity of the assumption that energy would not appear to be reversed from the forward in time viewpoint in which it is observed and it means that if energy was not effectively reversed from the time-symmetric viewpoint, then from the unidirectional viewpoint it would actually appear to be reversed by T , which is certainly not desirable.

On the other hand, the physical space intervals and the momenta associated with the propagation of particles do need to be reversed (once) when time is reversed if we insist on describing the motion of particles as it appears to take place from the conventional forward in time viewpoint and this despite the fact that only the physical time intervals experienced by the particles should effectively be reversed by T . Indeed, given that the direction of space intervals is defined in relation to the direction of time intervals, if time intervals are followed in the wrong direction, then space intervals are also traversed in the wrong direction, so that the observed directions of the motion of particles are opposite the true directions of their motion, which means that those directions are effectively reversed under a properly defined T operation when the outcome of this operation is considered from a unidirectional time viewpoint (this is made apparent when we reverse the direction of the arrows associated with the time reversed states in figure 3.1 to produce those in figure 3.2). Thus, when the direction of time is reversed, but the time intervals in which the particles effectively propagate are kept unchanged as a consequence of practical limitations imposed by the thermodynamic nature of the observation process, the associated space intervals actually appear to be reversed (they are the negative of those really experienced by the particles) even though the spatial positions remain unchanged. This is true again despite the fact that at the most fundamental level of description, which is that of bidirectional time, the direction of space intervals is to be consid-

ered unchanged by a reversal of time. As a consequence, we obtain results which comply with the traditional definition of time reversal according to which momentum (and implicitly also space intervals) should effectively be reversed by T along with angular momentum or spin, because given that momentum is here reversed independently from the position parameter x it follows that angular momentum would also appear to be reversed.

From the unidirectional viewpoint it may effectively seem like the traditional conception of time reversal as involving a reversal of motion which simply allows the particles to follow a trajectory backward could be valid. We must recognize, however, that just as there is no reason to assume that momentum is affected by a reversal of time from a bidirectional viewpoint (which explains that it is reversed from a unidirectional viewpoint), there is also no reason to assume that the sign of charge, as distinct from that of energy (the gravitational charge), would be affected from this same viewpoint when T is applied, because charge is not constrained to reverse by the requirement of action sign invariance when the direction of propagation in time reverses. This may also appear to comply with traditional expectations, but in fact (as I previously remarked) it rather constitutes the one aspect which introduces a radical departure from what is normally assumed concerning time reversal. Indeed, it means that the same reversal that does apply to momentum from the unidirectional time viewpoint would have to apply to non-gravitational charges as well, because if the direction of propagation in time of the charges is effectively reversed as required, then the fact that time is followed in the same forward direction relative to which the charges were originally propagating means that the charges would now appear to be reversed. We must therefore consider a reversal of charges to be associated with a reversal of time, as a result of the fact that this physical property is not experienced along the true direction of time in which it is propagated. This is a very important result which is definitely not expected from a traditional viewpoint given that it asserts that a quantity which was previously assumed to be unaffected by a reversal of time (namely the sign of charge) would actually appear to be reversed under such a transformation and if the preceding argument is valid then this conclusion would have to be considered unavoidable.

Thus it seems that considering a reversal of time without assuming a consequent reversal of charge is incorrect and may give rise to violations of symmetry which are a simple artifact of the inappropriateness of traditional assumptions concerning which quantities are reversed along with the time

coordinates, from the unidirectional viewpoint. To be meaningful, the experiments which seek to verify invariance under T would effectively have to assume a reversal of momentum and spin retracing a process backward, but combined with a reversal of charge (a permutation of particle and antiparticle). In other words, to test the invariance of physical laws under time reversal we would have to use antimatter, which may explain why a violation of T symmetry is so difficult to observe despite the fact that violations of the combined CP symmetry were effectively observed (which implies that T should also be violated given that CPT is inviolable). It appears that we are simply not using the right kind of matter to probe for T violation. It is not the invariance of a process relative to the thermodynamic arrow of time which must be probed, but invariance under a reversal of the true directions of propagation in time of elementary particles. I believe that the improved consistency of the interpretation suggested here from both an observational and a theoretical viewpoint confirms that the traditional definition of time reversal as involving nothing more than a reversal of the directions of motion and rotation of particles can no longer be considered appropriate.

It may also be noted that from a unidirectional viewpoint the reversal of charge and the reversal of spin under a properly defined time reversal operation are now the only aspects that differentiate this T operation from the P operation, apart from the respective reversals of the time and space coordinates themselves. But given that spin can also vary independently from the direction of propagation in time of a particle this means that the only unmistakable distinction between the time-reverse of a given state and the space-reverse of the same state is effectively the sign of charge, which again emphasizes the importance of recognizing the dependence of this parameter on the direction of time. In such a context it seems possible that the violations of T which may have been observed despite all the previously mentioned experimental difficulties could actually be violations of P symmetry, or violations of combined symmetries under which charge is left invariant by being reversed twice, because indeed those experiments do not compare matter and antimatter processes. Yet it might be considered that, despite what is commonly believed, violations of time reversal symmetry had already been observed, even before the violations of traditional T symmetry were reported, because, as I will explain below, the C operation also involves some time reversal and violations of charge conjugation symmetry do occur. In any case it is clear that a violation of the time reversal symmetry operation T as it was here redefined would not provide us with an absolute direction of time

at a fundamental level, but merely with a preferred direction of time relative to some arbitrarily chosen direction in space, or relative to some arbitrarily chosen sign of charge.

Another particularity of the alternative definition of time reversal proposed here is that it implies that it would now be electric fields which would reverse under application of the T operation instead of magnetic fields, because electric fields depend only on the sign of charge of the source particles and charge must be assumed to reverse under time reversal. Magnetic fields on the other hand would now remain unchanged under time reversal, because from the unidirectional viewpoint the direction of motion of the source particles would reverse, as is currently understood, but charge would also reverse, despite what is currently assumed, so that currents (which are the source of magnetic fields) would remain unchanged as a consequence of being submitted to this additional reversal. We must therefore assume that there is effectively a relative change between the direction of an electric field and that of a magnetic field under a properly defined time reversal operation, only it is not attributable to a variation of the magnetic field, but rather to a variation of the electric field. The failure to recognize the dependence of the sign of charge on the direction of propagation in time of elementary particles therefore gives rise to an incorrect appraisal of the response of electromagnetic fields to a reversal of time.

A more consistent definition of the operation of time reversal on the other hand allows to avoid the troubling conclusion that certain phenomena involving electromagnetic fields would actually constitute a challenge to the necessary relational definition of discrete symmetry operations. Indeed, violations of time symmetry could arise for example in the case where neutrons would be observed to have an electric dipole moment and as such could effect a movement of precession around the direction of an external electric field, because this movement would appear to vary depending on the direction of time, but independently from the direction of the field and the sign of the electric dipole. However, while the direction of the dipole is not affected by the reversal of a neutron's spin angular momentum occurring as a consequence of the reversal of time, according to my proposal it would nevertheless be reversed together with it, because it depends on the sign of the constituent particles' electrical charges, which we must now also assume to be reversed as a consequence of applying the T operation. It is not possible in this context to assume that a reversal of time would allow a change in the precession motion of the neutron (associated with the direction of the neutron's spin) to

occur *independently* from the direction of its electrical dipole in the presence of an invariant external electric field, because in fact both the spin and the dipole must be assumed to be reversed by T , along with the external electric field. In other words, it is no longer possible to assume that while we should observe the precession motion to occur in reverse upon reversing time, the same dipole would nevertheless be interacting with the same electric field, as would happen if applying T effectively reversed spin, but left the direction of the dipole and the external electric field unchanged. When the appropriate time reversal symmetry operation is considered, only *relative* differences can occur between the direction associated with the precession motion and the direction of the dipole.

Still concerning the T operation, it must be clear that it is not possible to assume that what the traditional definition of this transformation involves is a reversal of the time coordinate that reverses physical time intervals and leaves energy unchanged, combined with a reversal of momentum that leaves both space coordinates and physical space intervals unchanged, even if that would appear to correspond with the explicit definition of T as it is usually conceived. Such a definition of time reversal would be inapplicable simply because it would reverse the sign of action of the physical systems involved and this is certainly not desirable knowing that negative action matter (propagating positive energies backward in time) would be an entirely different kind of matter from a gravitational viewpoint and therefore certainly cannot be involved in those processes which we currently assume to be the time-reverse of processes involving positive action matter. This has nothing to do with the fact that a unidirectional viewpoint is used traditionally. It is a different problem that would be unique to the T operation despite the fact that I am here assuming that C also involves some time reversal, because charge conjugation is simply not assumed to involve any space or time reversal traditionally and as such cannot be mistaken to involve action sign reversal. From the viewpoint of unidirectional time we can therefore only assume that the space intervals are reversed by T , along with the momenta, and that the time intervals, along with the energies, are left unchanged by the same operation despite the reversal of the time coordinate. In other words, an appropriate (action sign preserving) time reversal operation needs to reverse both momentum and space intervals together (from a unidirectional viewpoint) or leave them unchanged together (from the time-symmetric viewpoint) and those constraints must be explicitly stated in the definition of the symmetry operation. This again illustrates how important it is to identify the variability

of all physical parameters under any discrete symmetry operation, in particular for what regards the sign of charge and that of energy in relation to the direction of propagation in time, as otherwise we may misinterpret ordinary phenomena for potentially forbidden, symmetry violating occurrences.

3.6 The charge conjugation operation

I think that in the context of the preceding analysis it becomes clear that the common assumption that time reversal amounts to simple motion (including rotation) reversal is what prevents a proper understanding of the nature of the charge conjugation symmetry operation. The problem is that if we ignore the dependence of the observed sign of charges on the true direction of propagation in time of the particles carrying them then this direction of propagation becomes impossible to assert, which explains that the existence of such a degree of freedom has traditionally been ignored altogether. Thus I believe that the mistake we do when we consider time reversal as it is traditionally defined (even if we can now recognize that this error is not *only* a consequence of using a unidirectional viewpoint) is that we do not consider an evolution according to which the direction of propagation in time of particles is really reversed, but instead consider processes for which a series of events occur forward in time, merely in the reverse order to that in which they would otherwise be observed to occur. But given that non-gravitational charges are not affected by a reversal of the direction of propagation in time of the particles carrying them (which is distinct from the observed direction of their motion) we have a means to determine the direction of propagation in time of particles which therefore becomes a meaningful, well defined concept which must be taken into consideration². It would therefore be incorrect to argue that only thermodynamic phenomena allow to distinguish a direction of time (even in the absence of violations of T symmetry), because from a unidirectional time viewpoint the sign of charge is dependent on the direction of time. It is thus simply the fact that the sign of charge itself cannot be characterized in an absolute manner that prevents a direction of time from

²This conclusion is also justified by the fact that if an observer was ‘following’ the actual direction of propagation in time of an antiparticle then this antiparticle would appear to have the same charge as its particle counterpart, but then it would be all the other particles in the universe which would appear to have a reversed charge, which is certainly a significant change.

being singled out as objectively distinct, in the way thermodynamic processes may appear to allow.

Now, what makes the acknowledgement of the existence of a relationship between direction of time and sign of charge unavoidable is the recognized validity of the interpretation of antiparticles as particles propagating backward in time, which allows to identify reversal of time as the very cause of the apparent reversal of charge occurring from the unidirectional time viewpoint. I believe indeed that despite what is often suggested, the interpretation of antiparticles as particles propagating in the opposite direction of time is not merely a helpful analogy with no real significance. Given the absence of a rational motive for rejecting the existence of a fundamental time direction degree of freedom equivalent to the space direction degree of freedom and given the simplification made possible by the discussed interpretation of antimatter in a relativistic context, I think that we must recognize that there definitely exists a relationship between the direction of time and the sign of charge. But it must also be clear that despite what is sometimes proposed there is no equivalence between a reversal of *space* directions and a reversal of the sign of charge (which could imply that antiparticles are merely the enantiomorphic equivalent of their corresponding particles), even if there does occur situations when reversing the space coordinates may appear to counteract asymmetries associated with the sign of charge, because the relationship between space direction and sign of charge is in fact always a consequence of the existence of a relationship between the direction of space intervals and that of time intervals. In any case, if the relationship between time reversal and charge reversal which is suggested by the above mentioned interpretation is considered valid then it would mean that the charge conjugation symmetry operation must actually be understood as itself involving some time reversal.

What I propose therefore is that we should recognize that the charge conjugation symmetry operation C must actually be conceived as a combined space and time reversal operation that leaves the sign of non-gravitational charges invariant relative to the direction of time in which particles would be propagating following such a reversal. Thus C must be understood to reverse the time parameter t (associated with the ‘position’ in time), along with the physical time intervals Δt associated with the propagation of particles, and the energy sign E of those particles (which is reversed as a requirement of action sign invariance). But it must also reverse the space position parameter x , the physical space intervals Δx associated with the

propagation of particles, and the momentum p of those particles (which is also reversed as a requirement of action sign invariance). Here again we must recognize that the charge q is actually left unchanged, along with the spin of elementary particles, from a fundamental viewpoint, even by this reversal operation we call charge conjugation. Yet it still makes sense to consider C as a reversal of charge given that, from the viewpoint of unidirectional time, non-gravitational charges would effectively appear to be one of the few physical properties of elementary particles which would effectively be reversed by this symmetry operation, while the space and time intervals, along with the energies and the momenta would appear to remain unchanged.

This must happen for the same reasons that justified assuming that momentum and space intervals are reversed by T from a unidirectional time perspective, even though they are left invariant by this symmetry operation from the bidirectional viewpoint. Indeed upon applying C we are in a situation where all intervals and their conjugate properties are reversed from a fundamental time-symmetric viewpoint, which means that to satisfy the needs of a unidirectional perspective we must reverse the time-related parameters Δt and E again, but given the relationships that exist between the physical time intervals and the space intervals this means that the space-related parameters Δx and p must also be reversed a second time, just as they were shown to be reversed (once) by T from this unidirectional viewpoint. It is therefore the fact that from the unidirectional viewpoint we use the wrong direction of time which requires that the physical time intervals and the energies be reversed from what they really are (what they have become as a result of applying the operation in the first place), but given that following time in the wrong direction also implies that the space intervals are followed in the wrong direction (the relational aspect) then it effectively means that the space intervals must also be reversed from what they really are (what they have become), along with the momenta. As a result, there appears to be no change to space and time intervals from applying C , even though it is here defined as a space and time reversal operation. Yet, as charge is not a spacetime related physical property, because it is associated with interactions distinct from gravitation (unlike energy or momentum which can be conceived as the charges determining the metric properties of spacetime), it should be considered that it effectively remains unchanged from the fundamental bidirectional viewpoint under a space and time reversal operation such as the properly defined C , which means that it would appear to be reversed, as we would normally expect, from the unidirectional time viewpoint

(because time is then followed in the wrong direction).

There is a slight difference, however, between the outcome of a properly defined C operation and the expected outcome of a traditionally defined charge conjugation operation, because the reversal of the space and time position parameters x and t themselves (which now occurs from both the bidirectional and the unidirectional time viewpoint), even if it is without any effect on the sign of the space and time intervals associated with the propagation of particles from a unidirectional viewpoint (given that those intervals must then be reversed a second time), actually implies that angular momentum would appear to be reversed by C (because momentum is effectively unchanged while the position in space is reversed). Thus, despite common expectations, a C -reversed process would also appear to involve reversed angular momentum or spin, because relative to positive definite time the space intervals would be experienced as positive, while spatial positions are still negative. Contrarily to what is sometimes suggested, therefore, the behavior of spin under charge conjugation is not a mere matter of convention and its reversal (apparent from a unidirectional time perspective) must be considered an unavoidable outcome of applying this symmetry operation. The reversal of spin under C is certainly unexpected according to the traditional approach, but from my perspective it appears natural, given that C involves a reversal of time. It must be clear though that this reversal of spin is only apparent and does not occur at the most fundamental level of description, in accordance with the requirement that an action sign preserving symmetry operation like C should not reverse the sign of action associated with angular momentum. This is to be required even if in general the sign of spin is not uniquely tied to the sign of action associated with energy and momentum, because the only way spin can reverse is when either the position in space or the momentum are independently reversed and an action sign preserving reversal symmetry that reverses momentum would necessarily also reverse spatial position given that it must reverse space intervals (which is not required from the unidirectional time viewpoint relative to which momentum can be made to vary independently from the sign of space position, even when action is to remain positive).

We are now therefore in the situation where we must recognize that, from a certain viewpoint, charges are reversed by a properly defined time reversal operation T , while spin angular momenta are reversed by a properly defined charge reversal operation C , despite what had traditionally appeared to be required from such discrete symmetry operations. Another distinc-

tion of the proposed approach is that handedness is now also reversed by C from whatever viewpoint, because either momentum is reversed and spin is invariant (as from the bidirectional viewpoint), or momentum is invariant and spin is reversed (as from the unidirectional viewpoint), so that there is always a relative change between the direction of spin and that of momentum. The outcome of the proposed charge reversal operation C as it was redefined would therefore differ from that of a properly defined T operation mainly through the fact that unlike C , T would reverse the momentum and space intervals (from a unidirectional viewpoint), but would not reverse the handedness of particles, just as we would also expect traditionally. Thus both the P operation and the redefined C operation would alone and from any viewpoint reverse the handedness. In this context the fact that under certain circumstances, such as when the weak interaction is involved, particles of a given handedness seem to be naturally related to antiparticles with opposite handedness could be understood to follow from the fact that the handedness is reversed by a properly defined charge conjugation operation (which still relates particles to antiparticles), so that if there can be invariance under such a symmetry operation then reversing both charge and handedness should not be expected to produce any change. This is an important result which confirms that the suggestion, usually made on the basis of purely phenomenological considerations, that charge conjugation should perhaps involve a reversal of handedness, was in fact justified from a theoretical viewpoint.

3.7 Invariance under combined reversals

I think that I have appropriately justified the inevitability of the above discussed conclusions regarding which parameters should be expected to reverse under the various discrete symmetry operations (in particular when I discussed the requirement of action sign invariance and the constraint of relational definition of the reversal operations), but I must nevertheless mention how remarkable it is that the described variations of physical parameters under the redefined P , T , and C operations happen to be just such that they *explicitly* require invariance to occur under a combined PTC operation. This happens because all the parameters which are independently reversed by any of the symmetry operations are effectively reversed twice when the operations are combined and this regardless of whether we are considering a

unidirectional or a bidirectional time viewpoint (a look at tables 3.3 and 3.4 allows to quickly confirm this fact). Either a parameter such as Δt is reversed twice or either it is not reversed a single time by a properly defined PTC and this effectively guarantees that there is invariance under a combination of the three discrete symmetry operations, because anything that may be reversed is reversed again and only once. In fact, as I will explain below what we actually need is twice a reversal of *all* fundamental space- and time-related parameters (that is both the time-related parameters t , Δt and E , and the space-related parameters x , Δx and p) under a properly defined PTC and this effectively occurs when the appropriate bidirectional time viewpoint is considered. Charge and spin on the other hand need not reverse at all from such a viewpoint under a PTC operation as they necessarily transform independently from the action sign preserving discrete symmetry operations and only reverse as a consequence of adopting a unidirectional viewpoint and in such a case they effectively reverse twice as required. This is in contrast with the traditional definition of the discrete symmetry operations (described in table 3.2) according to which some parameters like the space and time coordinates, the charge, and the spin can be reversed a single time only by the combined PTC operation.

We can understand, however, why it is that this combined symmetry operation should be expected to produce invariance even as it is traditionally defined (as required by the CPT theorem). This is possible simply because, according to the traditional conception, while charge would be reversed only once (by C), spin would also be reversed only once (by T), but as can be shown, there is a kind of equivalence, at least for fermions, between a reversal of the polarization state associated with spin and a reversal of charge and this is why even under its traditional definition the combined PTC symmetry operation would have to leave physical states invariant (although it would seem to alter the direction of space and time coordinates, which could turn out to be physically significant under particular circumstances). It is also interesting to observe that in the context of my revised definitions of the discrete symmetry operations any two operations applied together is *explicitly* equivalent to the remaining operation, so that applying PT , for example, is totally equivalent to applying C , which again demonstrates that charge conjugation must really be conceived as a space and time reversal operation and that time reversal must involve a reversal of charge from a certain viewpoint. What those relationships really show is that the discrete symmetry operations as they are now defined are all necessary and together sufficient

to provide a complete account of the possible transformations involving a reversal of any of the fundamental properties of matter aside from the sign of action.

In this regard I must also mention that it is not possible to assume that applying either P or T alone but twice should necessarily produce invariance (in the sense that it would leave any system with no discernible change that could be related to unchanged physical parameters) despite the fact that it would appear to leave all parameters unchanged, because such a combined transformation may not leave the quantum phase associated with fermions unchanged given that it would only be equivalent to a rotation in space by 2π radiant (as a single space reversal introduces a π radiant rotation and a single time reversal introduces an equivalent additional π radiant rotation in space) and only twice such a complete rotation would necessarily produce invariance in the presence of fermions. Of course applying P or T alone twice would already be more likely to produce invariance than applying P alone or P combined with T only once, because at least some of the effects of applying P or T once would effectively be neutralized by a second application of the same operation, but the point is that in such a case invariance would not *necessarily* follow. The case of C is different, however, given that this operation involves a reversal of both space and time parameters all at once, which produces an equivalent 2π radiant rotation with only one application (therefore allowing the changes involved to be related to the incomplete transformation of fermion wave functions), so that applying C twice reverses *all* parameters twice and introduces twice a 2π rotation that must leave even the quantum phase of fermions invariant. The C operation as I redefined it is thus unique, because it is the only one of the three relationally distinct discrete symmetry operations that reverses both space- and time-related parameters together and from its alternative definition it can be seen that applying C is effectively and explicitly equivalent to applying a combined PT operation. In this context applying PTC could be considered equivalent to applying PT twice, which clearly shows that the PTC operation involves a reversal of all parameters twice and is also equivalent to two complete rotations, which can only produce invariance.

In fact, any one of the three basic discrete symmetry operations can be considered as equivalent to a combination of the other two, so that T for example would here be equivalent to CP and P would be equivalent to the combined CT . Therefore, applying T twice would be equivalent to applying CP twice, which would amount to reverse both space- and time-related pa-

rameters twice (which considered alone would have to produce invariance) and then also reverse space-related parameters twice (the order of application of the discrete symmetry operations in a combined operation has no importance and only the number of times a parameter is reversed is significant). But such a combined operation would not leave fermion wavefunctions invariant for the same reason that applying P alone twice should not be expected to necessarily leave things invariant. It remains, however, that the fact that some combinations of basic discrete symmetry operations which are not required to necessarily produce invariance can effectively involve twice a reversal of some specific physical parameters, allows one to expect that an invariance which was lost when one of those fundamental operations was applied alone can sometimes be regained by application of such combined operations. This should indeed be expected to occur given that, as I mentioned above, reversing one physical parameter twice, even if it is not guaranteed to leave all processes invariant, still allows the possibility of neutralizing some asymmetries which would occur as a consequence of the reversal of this single parameter.

What must be retained here is that there may be a difference between applying a symmetry operation twice and applying the *outcome* of this operation only once (which would effect no change), even if in certain cases, as when the operation considered is the C symmetry operation, we would necessarily observe no change when the same operation is applied twice. This particularity of the C operation is merely a consequence of the fact that it reverses more individual parameters all at once so that applying it in combination with itself effectively allows to leave no parameter unchanged relative to which an asymmetry could be properly defined. It must be understood, however, that despite their equivalence with combinations of distinct operations, the three basic operations defined above are all essential to a description of the allowed discrete transformations of physical parameters and none is more fundamental than any other. Indeed, two operations are distinct from a relational viewpoint, when one of them reverses one category of parameter, say space, relative to the other category, say time, while the other reverses another category of parameter, say time, relative to the previous one, say space, and each one of those operations is relationally distinct from yet another one that reverses both categories of parameters together and which constitutes the necessary complement to the other two operations.

3.8 The significance of classical equations

We can now return to the problem of understanding how it is possible for the momentum p to be left unchanged by a properly defined time reversal operation T which from the most fundamental viewpoint must be assumed to reverse time intervals dt , but to leave space intervals dx unchanged. A problem would effectively appear to arise from the fact that according to the classical equation that defines the momentum of a particle with mass m we should have $p = m dx/dt$, which would clearly imply that if dt is reversed or negative while dx is invariant or positive then p should be negative, which is contrary to my proposal that both space intervals and momentum are unaffected by a reversal of time. But I would like to suggest that this contradiction is only apparent and a result of the fact that the classical equation for momentum is actually valid only from a unidirectional viewpoint, because it was originally introduced under the implicit assumption that physical properties are always measured in the conventional forward direction of time. Indeed, what the classical equation is telling us is merely that from the unidirectional viewpoint of an observer always following events in the unique direction of time associated with entropy increase and providing an account of physical quantities like momentum and space intervals in relation to that unique direction of time, relative to which time intervals dt are effectively positive definite, independently from the true direction of propagation in time of the particles involved, some quantities like dx which we might assume not to be reversed by T are effectively observed to be reversed while dt itself is kept unchanged. Thus if we use the viewpoint relative to which we are allowed to assume that the above equation is valid then dt would actually remain positive definite despite the reversal of time, while dx would have to be assumed reversed (for reasons I have already explained), which means that momentum would also be reversed according to this action sign preserving classical equation, which agrees with the definitions I provided for the unidirectional viewpoint and which is certainly appropriate given that particles submitted to such a time reversal operation must have invariant momentum in the apparent (but false) direction of their motion, which is satisfied when momentum has the same negative sign as space intervals.

There is no contradiction here, despite the fact that we must assume that the true signs of conjugate physical parameters such as the space intervals and the momenta are together invariant under a reversal of time from the alternative time-symmetric viewpoint (according to which the sign of time

intervals is itself reversed), because in such a case the classical equation no longer applies, simply because as a traditional formula it *never* really applied to such situations. The classical relation between momentum and the space and time intervals was deduced on the basis of the validity of a thermodynamic viewpoint of time and therefore does not apply in a context where time intervals are allowed to change sign. The classical equations are logical deductions dependent on a certain viewpoint of time which must be considered inappropriate at the most fundamental level of description. In other words, it is not the validity of the classical equations in a limited context which implies that the assumptions made from a time-symmetric viewpoint (concerning the sign of physical quantities) are contrary to experimental evidence, but really the limited value of the classical equations which imply that the assumptions associated with a unidirectional viewpoint are not generally valid. We must recognize that the assumptions used in the more appropriate time-symmetric context regarding the variations of space- and time-related quantities under a reversal of time are not just theoretically well motivated, but that under the right interpretation they are fully supported by observations, while the variations deduced from a unidirectional time viewpoint are explainable merely in the context where they are assumed to derive from the more fundamental bidirectional description.

It must be clear that in this context we would also be unjustified to make use of the classical formula for angular momentum \mathbf{L} , to which the spin of elementary particles is related, to decide what would happen to spin from a fundamental viewpoint under a reversal of time effected by a properly defined T or C operation. Indeed, the classical formula defines the angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ in terms of the position vector \mathbf{r} and the momentum $\mathbf{p} = m(dx/dt)\mathbf{i}$ and if we assume a reversal of time intervals dt to follow from both a T and a C reversal operation then according to this equation it would seem that \mathbf{L} should reverse under both types of time reversal, because either dt reverses alone (as under a properly defined T) or it reverses along with \mathbf{r} and dx (as under a properly defined C). But, as I already mentioned, and for reasons I have previously discussed, it would be incorrect to assume that angular momentum reverses under either T or C from the bidirectional time viewpoint relative to which dt effectively reverses. Yet there is no problem here, because the classical formula is only right when we consider things from the unidirectional viewpoint according to which dt is positive definite, but under such conditions either dx and p reverse together with unchanged \mathbf{r} (as occurs when T is applied), or else dx and p are unchanged and \mathbf{r} is reversed

(as occurs when C is applied and only space positions are reversed) so that in both cases spin angular momentum should effectively reverse. Again it must be emphasized that the incompatibility of the classical equation for angular momentum with the definition of time reversal as it occurs from a fundamental bidirectional viewpoint must not be considered to imply that the proposed fundamental definition is inapplicable, because all that it means is that the equation itself is of limited scope, having been developed in the context of a unidirectional perception of the evolution of physical systems, when it had not yet even been realized that there exists a fundamental degree of freedom associated with the direction of propagation in time.

3.9 Reversal of action

The clarification of the situation which was achieved in the preceding sections regarding the interdependence of fundamental physical properties as they vary under application of any of the three essential discrete symmetry operations has allowed to established that that none of the traditionally considered discrete symmetry operations engenders a reversal of the sign of action. This is of course a consequence of the fact that regardless of the viewpoint we adopt, those symmetry operations always reverse the sign of energy in combination with the sign of time intervals associated with the propagation of particles, just as they always reverse the direction of momentum in combination with the direction of space intervals. Thus the T operation in particular, despite the ambiguity of its traditional definition, cannot be assumed to reverse the action, because while it reverses the time position parameter and leaves the sign of energy unchanged from the unidirectional time viewpoint, it is also implicitly assumed to preserve the sign of time intervals associated with the propagation of elementary particles. The role of inverting the sign of action must therefore be attributed to some symmetry operations distinct from all of those which are usually considered.

I have come to understand that there is not a unique single operation relating positive and negative action states, but that there are basically four different ways by which action can be reversed, which give rise to four different action sign reversing symmetry operations, whose four different outcomes are each related to phenomenologically distinct states of negative action matter. If any one of those operations is applied independently from the others, it may not necessarily produce invariance. I will collectively denote those op-

erations by the letter M to emphasize the fact that they constitute a different category of reversal transformations which are unlike those already studied. The states produced by those four distinct operations can be transformed into one another by individually applying each of the three action sign preserving symmetry operations P , T , and C and therefore I will denote the various action sign reversing operations by applying the appropriate indices corresponding to the operations which relate the states they generate to the state which is produced by one of those action sign reversing operations chosen arbitrarily as the basic operation, which will itself be denoted M_I . The four discrete symmetry operations so defined are thus the M_I , M_P , M_T , and M_C operations displayed in table 3.5. It must be clear, however, that the choice of which action sign reversing transformation must be associated with the basic operation M_I is completely arbitrary and we could, for example, have defined the operation originally denoted M_C to be the basic operation, which we would instead denote M'_I and we would then obtain the states produced by the other three operations by applying P , T , and C to the state generated by M'_I . That way it would appear that it is the redefined M'_C which would be equivalent to the original M_I , while M'_P would be equivalent to M_T , and of course M'_T would be equivalent to M_P and therefore we see that attribution of the indices is purely a matter of convention. The letter M was chosen to denote action reversal because the operations it represents would effectively alter the gravitational properties of the matter submitted to such reversals and mass (which is usually denoted m) is the property that was traditionally associated with the gravitational interaction.

From the tables it is possible to see that there are two different ways by which a given type of fundamental physical parameter, either space- or time-related, can be reversed in such a way that the sign of action is reversed. We can either assume a reversal of the signs of momenta and energies relative to unchanged space and time intervals or we can assume a reversal of the space and time intervals associated with the propagation of particles that would occur while keeping the signs of momenta and energies invariant. But given that those two different kinds of reversal can be applied differently to space- and time-related parameters (you can apply one kind of reversal to space and the other to time or vice versa as long as you do apply any one type of reversal to each type of parameter), it means that there are four different kinds of operations in all which can reverse the sign of action. From those definitions it is clear that what the M_I , M_P , M_T , and M_C operations really involve is the reversal of an additional degree of freedom relationally distinct from

Bidir.	t	Δt	E	x	Δx	p	q	s	h
$M_I = M'_C$	t	Δt	$-E$	x	Δx	$-p$	q	$-s$	h
$M_P = M'_T$	t	Δt	$-E$	$-x$	$-\Delta x$	p	q	$-s$	$-h$
$M_T = M'_P$	$-t$	$-\Delta t$	E	x	Δx	$-p$	q	$-s$	h
$M_C = M'_I$	$-t$	$-\Delta t$	E	$-x$	$-\Delta x$	p	q	$-s$	$-h$

Table 3.5: Variations of physical parameters under the four relationally distinct action sign reversing symmetry operations as described from the bidirectional time viewpoint. Here I chose the basic action reversal operation M_I to be that which reverses energy E independently from time intervals Δt , and momentum p independently from space intervals Δx . Under an equivalent definition it would be the time intervals Δt and the space intervals Δx which would reversed by the basic action reversal operation M'_I while the energy E and the momentum p would be kept invariant.

those already affected by the P , T , and C operations, because indeed even the state obtained by applying the basic M_I operation effectively involves a reversal of action, which means that all possible states related by application of P , T , and C , including the original state obtained by application of the identity operation I have their counterpart as M -reversed states and under such conditions we can only conclude that we are effectively dealing with a transformation that applies to a distinct property of matter. The illustration of the effects of the various action sign reversing operations depicted in figure 3.3 allows to clearly identify this degree of freedom as the relative orientation of momentum p compared to space intervals Δx or equivalently that of energy E compared to time intervals Δt , which for negative action states is the opposite of what it is for positive action states.

The C , P , and T operations, therefore, do not together operate a reversal of all fundamental physical parameters, because they merely reverse all parameters while leaving the sign of action invariant. The four action sign reversing symmetry operations proposed here are then the additional operations which are required to complete the set of discrete symmetry operations, because they perform the only remaining possible changes that the traditional operations do not produce, by effectively reversing the sign of momentum and energy relative to the direction of space and time intervals. From that viewpoint it appears that even though they are usually ignored

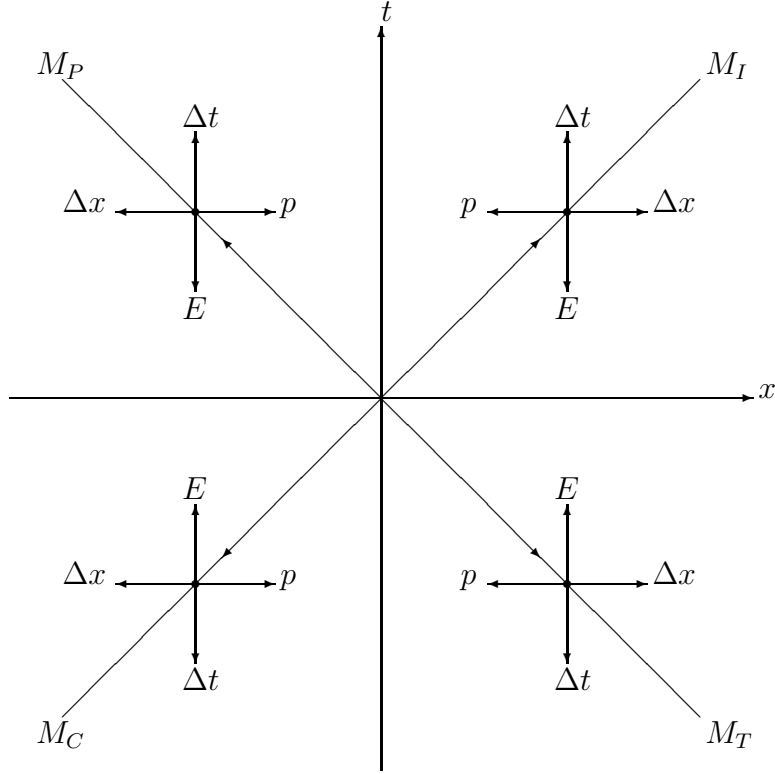


Figure 3.3: Four different outcomes of applying each of the relationally distinct action reversal symmetry operations as described from the bidirectional time viewpoint. Here we notice that the orientation of the vectors which correspond to the signs of space and time intervals is always opposite that corresponding to the signs of momentum and energy, as we should expect to observe when action is effectively negative. If we were to consider a unidirectional time viewpoint we would have to reverse all space and time intervals and all momentum and energy signs for the processes obtained by application of both the M_T and M_C operations, which means that all four operations would give rise to the propagation of negative energies forward in time.

the M_I , M_P , M_T , and M_C operations cannot in fact be avoided. The fact that there are actually four distinct operations that can perform a reversal of action on the other hand simply means that it is not possible to associate a unique state of momentum or energy, or of propagation in either space or time, to negative action matter and that all the different action sign preserving variations of the direction of fundamental physical parameters which can apply to positive action matter would also apply to negative action matter. We can thus effectively expect that there would, for example, be a charge conjugation symmetry operation C applying independently to negative action matter, which would therefore have its own antimatter particles distinct from ordinary antiparticles.

In this context it appears that the distinction that exists between matter and antimatter must be attributed essentially to the true direction of propagation in time of particles, independently from their sign of action. An antiparticle is therefore always just a particle which reversed its energy while changing its direction of propagation in time, which is not very different from the situation of a particle which reverses its momentum by changing its direction of motion in space. Indeed, by reversing its momentum when it changes its direction of propagation in *space* a particle is allowed to keep the sign of its momentum relative to the direction of its motion unchanged, so that its action sign is also unchanged, just like a positron retains the sign of action of the electron with which it annihilates, because the electron reverses its energy when it starts propagating backward in time (which is viewed as the annihilation process forward in time). But a negative action particle would be clearly distinct in this respect as a consequence of the fact that it would have not only negative energy carried forward in time (or positive energy carried backward in time, which is equivalent from a unidirectional time viewpoint when the sign of charge can be ignored), but also negative momentum in the observed direction of its propagation in space (the momentum would point in the direction opposite the observed velocity of the particle), unlike any ordinary matter particle (including antiparticles). It must be clear, however, that according to the proposed definition of action sign reversing symmetry operations which is described in table 3.5, non-gravitational charges are assumed to be unaffected by a reversal of action, just as they were left invariant by the action sign preserving reversal operations. Only the practical necessity of a forward in time viewpoint would for negative action matter also imply that charges appear to be reversed when a process is submitted to an action sign preserving reversal of time.

Another particularity of the operations of action reversal defined above is that spin is deduced to be reversed under all such relationally distinct operations when their effects are considered from the bidirectional time viewpoint. This is certainly just as appropriate as is the invariance of spin observed for all action sign preserving symmetry operations, because as I previously mentioned spin has the dimension of an action and should therefore vary in correspondence with the sign of action associated with momentum and energy from a fundamental viewpoint. The constraint on the variation of the direction of spin is actually the same constraint that requires that either both space- and time-related parameters are such as characterizing a positive action state, or else that they are both such as characterizing a negative action state and that it should not be possible for one single particle to propagate, say, positive momentum in the direction of its motion in space and at the same time propagate negative energy forward in time. This is a simple matter of consistency, because a physical system cannot have at once both the gravitational properties associated with positive action matter and those associated with negative action matter if, as I suggested in the previous chapter, the attractive or repulsive nature of the gravitational interaction between two particles effectively depends on the difference or identity of their action signs. This does not mean, however, that spin cannot vary independently from the sign of action associated with energy and momentum, but merely that while it cannot reverse as a consequence of applying an action sign preserving discrete symmetry operation, it also *must* reverse as a consequence of applying a reversal of action.

It may also be noted that just as is the case for the action sign preserving discrete symmetry operations, some combinations of two of the four operations describing a reversal of action are equivalent to a combination of the other two operations (in the case of the action sign preserving operations one operation, which is that of charge conjugation C , was equivalent to the other two, but in fact this single operation was implicitly combined with the invariant operation I which effected no additional change and thus could be ignored). Here a combination of M_I and M_C or a $M_I M_C$ operation would be equivalent to a combination of M_P and M_T and this is what allows a combination of all action sign reversing symmetry operations (or a $M_I M_P M_T M_C$ operation) to necessarily produce invariance, given that all relevant parameters are effectively reversed twice by such a combined operation. In fact, it turns out that combining any of M_P , M_T , or M_C with M_I produces an operation equivalent to the above defined P , T , or C respectively (while a

combination of M_I with itself produces an operation equivalent with the identity operation I) so that a combination of the other two remaining action sign reversing operations would also be equivalent to those action sign preserving operations. For example, the combined $M_P M_T$ operation is mathematically equivalent to a C operation because it reverses both space- and time-related parameters once and reverses the action twice, which is equivalent to leave action unchanged.

One must understand, however, that even though applying any one action sign reversing operation twice would be equivalent to applying the identity operation I , such a combined operation would not necessarily produce invariance and this for the same reason that applying P or T twice would not necessarily leave everything invariant despite the fact that it would also appear to be equivalent to applying the I operation, which effect no change. This is, again, because applying an operation that does not reverse all physical parameters twice, even if it may appear to return a system to its original state, may still produce a change which can be characterized in a relational way, because some parameters would be reversed relative to other parameters which remain unaffected by the transformation and this may not leave the processes involved invariant. Still regarding the conditions for necessary invariance, it should be clear that simply combining a PTC operation with the basic M_I or any other action sign reversal operation as a way to try to regain invariance which may be lost upon reversing the action (in the way we would apply T to a CP violating process) cannot be expected to produce invariance given that the action sign degree of freedom would then be reversed only once. Thus, a violation of any of the M symmetries would not imply that there must be a violation of PTC symmetry, as we may understand to be independently required on the basis of the fact that invariance under PTC alone must itself be considered unavoidable. The appropriate generalization of the PTC (or really $IPTC$) symmetry must then be recognized to be the $M_I M_P M_T M_C$ symmetry which combines all the relationally distinct action reversal symmetry operations and which must therefore (because there would remain no unchanged physical parameter relative to which a change could be determined) be equivalent to no change at all. Indeed, as indicated in table 3.5, a physical parameter may either not be reversed by any of the action reversal operations or else be reversed by two or all of those symmetry operations, which explicitly guarantees invariance under a combination of the four operations.

Now, in order to avoid confusion, it is important to understand that the

action sign reversal symmetry operations must be considered as operations distinct from one another that apply to an identical state, rather than as an identical operation that applies to different states. In such a context it transpires that the fact that the M_I , M_P , M_T , and M_C operations are related to one another through application of the various action sign preserving symmetry operations merely shows that the states obtained by applying the four action sign reversing operations are themselves related to one another through the same action sign preserving operations that transform unchanged action sign states into one another. Thus, despite the fact that all of the action sign reversing symmetry operations are equivalent to a combination of some arbitrarily chosen action sign reversing operation with one of the four action sign preserving operations, it would not be appropriate to assume that invariance can be obtained by applying all the action sign preserving symmetry operations together as *IPTC* (which must leave everything invariant) along with the combination of a single particular action sign reversing operation with itself, like $M_I M_I M_I M_I$ (which would not necessarily produce invariance).

What must be clear is that no action sign reversing symmetry operation can be identified as *the* action reversal operation and under such circumstances it is not possible to avoid having to consider the many operations as distinct from one another despite the fact that all such operations can be obtained by combining in turn each of the action sign preserving symmetry operations with just one single action reversal operation. In this context it is important to realize that action reversal symmetry can be violated to different degrees when one transform a state of positive energy matter into the different states of negative energy matter which are related to one another by the redefined action sign preserving reversal operations P , T , and C , because each of those states is related to a corresponding state of positive energy matter by a specific action sign reversing symmetry operation and these operations do not necessarily produce invariance when applied separately. Thus the P , T , and C operations can be violated to different degrees by negative energy matter (compared to how they are violated by positive energy matter) when applied independently from one another and this precisely because M_I , M_P , M_T , and M_C can themselves be violated to different degrees in comparison with one another, so that they relate the different asymmetric states of positive energy matter to corresponding states of negative energy matter which can be asymmetric in different ways relative to one another. The only requirement is that the different states of negative energy matter which are

related to the different states of positive energy matter by the various action sign reversal symmetry operations be subject to the same invariance under a combined PTC transformation as are states of positive energy matter, even if P , T , and C are violated to different degrees by negative energy matter in comparison with the violations occurring for positive energy matter. The four action reversal symmetry operations, therefore, simply allow to relate all the positive energy states which are transformed into one another by the action sign preserving symmetry operations to all the negative energy states which are transformed into one another by similar operations. Thus, despite the existence of four distinct action sign reversal symmetry operations, action reversal must really be conceived as transforming one single degree of freedom and this means that I am justified in referring to the action reversal operations collectively as the M symmetry.

In any case it appears that the commonly met remark to the effect that gravitation is invariant under a reversal of time must be nuanced. What I mean is that while it is certainly true that there would be no change to the attractive or repulsive nature of the gravitational interaction if time was locally reversed for some physical system by a time reversal operation such as T , we should certainly expect a reversal of time independent from the sign of energy, such as that produced by an M_T operation, to exert a change on the nature of the interaction of the affected system with the rest of the universe. Indeed, such a transformation would reverse the sign of action and as I previously explained the repulsive or attractive nature of the gravitational force between two bodies depends on the relative value of their action signs (because gravitation is always attractive only for particles with the same sign of action). But even if we consider a reversal of time as produced by an action sign reversing operation like M_T to apply to the whole universe (in which case we would have to use negative energy matter in place of positive energy matter when testing for invariance), the preceding discussion made clear that we should not necessarily expect to observe phenomena which would be entirely identical with those of the original universe, because M_T applied alone could be violated, just as any operation which is not reversing all physical parameters twice. This would also be true of M_P for example, because just as the change in the sign of time intervals produced by an M_T operation can be related to an unchanged sign of energy, so the change in the direction of space intervals produced by an M_P operation can be related to an unchanged direction of momentum.

Yet the fact is that there *could* effectively be invariance under a reversal of

time that does not preserve the sign of action if the operation is applied to all particles in the universe, because in such a case the difference or the identity of the signs of action of the various particles would not be affected and this is the only aspect that would be significant from a gravitational viewpoint. But this invariance would apply only to the extent that there is effectively no violation of the symmetry between the positive and the negative action states of matter which are related through such a reversal operation in our universe. In fact, though, we may have to consider that such a violation is effectively occurring given that, based on the developments introduced in the preceding chapter it seems that the observed small, but non-vanishing positive value of the cosmological constant would have to arise from just such a minute violation of the symmetry under exchange of positive and negative action matter and those are the two kinds of states which are related to one another by a time reversal symmetry operation such as M_T . I may finally add that, as I have explained in a preceding section, simply reversing the direction of motion of the particles involved in a process cannot be considered to consist in a true time reversal operation in any meaningful way, so that assuming that such a transformation would leave processes unaffected, even when gravitation is involved, could not be understood to mean that gravitation is invariant under time reversal.

3.10 The problem of matter-antimatter asymmetry

It has been suggested more than once that the violations of CP symmetry which have been observed in certain experiments and which are believed to imply a violation of time reversal symmetry T could perhaps be the cause of the observed thermodynamic time asymmetry in our universe. It is usually recognized, however, that the weakness of the T violation that is involved would prevent it from being responsible for such an extreme difference between past and future evolution as that which gives rise to the thermodynamic arrow of time. A less common proposal is that it might be the thermodynamic time asymmetry itself which is giving rise to the violation of T symmetry. But this intrusion of macroscopic physics into the affairs of microscopic quantum processes is usually not believed to be a likely possibility, at least by those who do not expect a complete overthrow of conventional

particle physics. In fact, I think that what really justifies this attitude is the recognition that what currently remains unexplained is the macroscopic arrow of time, while any fundamental time asymmetry observable at the elementary particle level could be accommodated by the same rules that make violations of parity possible. Yet if we were to have a satisfactory independent explanation of thermodynamic time asymmetry then it could make sense to try to explain microscopic time asymmetry as a consequence of the existence of this thermodynamic arrow of time. I believe, however, that I have appropriately explained why we do not in effect need to appeal to macroscopic time asymmetry to legitimize the violation of T symmetry, given that it is allowed to occur as a relationally defined asymmetry (it does not need to be defined relative to the direction of thermodynamic time as it is already defined in relation to other fundamental directional parameters).

Now, even if thermodynamic time asymmetry is probably not the cause of violations of T symmetry, the direction of time singled out by T violations can be related to the macroscopic arrow of time and this might allow one to conclude that our universe is characterized by a phenomenologically apparent fundamental lopsidedness. Given that the time reversal symmetry operation can now be understood to involve a transformation of matter into antimatter the question of whether there effectively exists such a preferred direction in time would be equivalent to ask if there really is an absolutely definable asymmetry between matter and antimatter in our universe. However, when I examined this question in the light of the more appropriate conception of antimatter which arose from the developments featuring in the preceding sections I found out that there is after all no absolutely definable lopsidedness if we recognize the validity of a certain hypothesis concerning the continuity of the flow of time along a particle's world-line in spacetime. This hypothesis is that which I had at some point contemplated as potentially offering the required constraint that would prevent transitions in which the direction of propagation in time of a particle reverses without being accompanied by a reversal of the energy of the particle (thereby giving rise to processes of creation and annihilation of pairs of opposite action particles). I mentioned in the discussion of this problem that appeared in the previous chapter that this condition of continuity must effectively prevent certain changes from occurring on a continuous particle world-line, even though all by itself the limitation involved is not restrictive enough to prevent a reversal of energy independent from the direction of propagation in time (a reversal of action). I am now allowed to assert that what such a condition of continuity

requires in effect is merely that there needs to be a continuous flow of the fundamental time direction parameter associated with the sign of physical time intervals along a particle world-line. This restriction becomes relevant in the context where it is recognized that there effectively exists a fundamental time direction degree of freedom distinct from the observed direction of motion of elementary particles.

Compliance with such a continuity requirement would imply that any particle-antiparticle annihilation process, whether it involves particles with the same action sign, or particles of opposite action signs can only occur as the kind of events during which a particle bifurcates in spacetime to start propagating in the opposite direction of time and not as a chance encounter of two opposite-charge particles propagating in the same direction of time. This requirement would then also impose that events cannot occur which would appear to involve a particle turning into its antiparticle by releasing twice its charge without ceasing to exist from the unidirectional time viewpoint, because such processes would imply that the continuous path of a particle in spacetime (the arrow along a particle world-line) could come to an end as a consequence of a particle by chance meeting its backward propagating antiparticle from the future. Yet we have no choice but to effectively assume that antiparticles are backward in time propagating particles (and not particles propagating opposite charges forward in time), because, as I mentioned in the discussion concerning the time-direction degree of freedom appearing in the previous chapter, if we are to view any transformation along a particle world-line as a continuous process then given that the annihilation of a particle with an antiparticle must be allowed to occur with the same probability for all such pairs and cannot only take place for those pairs where the two particles would happen to be those propagating in opposite directions of time, then anti-particles must always be considered to propagate in the direction of time opposite that in which the corresponding particles are propagating. Thus particles defined as propagating backward in time cannot have sometimes positive charge and sometimes negative charge (the same is true for particles propagating forward in time) and in the context where continuity of the flow of time is required to apply on a particle world-line this means that no particle can turn into an antiparticle without effectively reversing its direction of propagation in time at the instant where the transformation event takes place (therefore describing a particle-antiparticle annihilation process forward in time) even if charge could be conserved when a particle would turn into its antiparticle without bifurcating in time (through the emission

of a compensating amount of charge carried by interaction bosons).

What I would like to suggest, therefore, is that we must consider as a necessary rule rather than as a convenient assumption that the arrow associated with the direction of propagation in time of a matter particle (from a *bidirectional* viewpoint) can never reverse along a continuous world-line. This requirement can be formally expressed using the following definition.

Condition of continuity: There must be continuity of the world-lines of elementary fermions under all conditions, so that particles cannot turn into antiparticles (and vice versa) without changing their direction of propagation in time in such a way as to preserve the continuous flow of the fundamental time direction parameter.

If this assumption is valid then from the unidirectional time viewpoint a particle cannot appear to continue propagating forward in time after changing into its antiparticle (just as from the same viewpoint an antiparticle could not have kept propagating backward in time *before* an event at which it transformed into its particle counterpart). Therefore, if a particle continues to propagate forward in time then it must actually retain the sign of its charge, because if it does not then either the condition of continuity would be explicitly violated or we would have to assume that a forward in time propagating particle can sometimes have an opposite sign of charge, which must be considered to be empirically ruled out in the context where the condition of continuity must apply (because annihilation processes cannot only occur for a subset of particle-antiparticle pairs), so that again the same constraint would be implicitly involved in forbidding the occurrence of processes of the kind discussed here.

It must be clear, however, that this limitation is not a requirement of current elementary particle theories, even though no process that violates this rule has ever been observed. In fact, some unconfirmed grand unification theories actually predict the existence of processes which would violate this continuity constraint, but in my opinion this is probably reason enough to doubt their validity given the awkwardness of the kind of evolution they would describe in the context of the best interpretation we have for the nature and the origin of antiparticles. But even if that type of processes was to eventually be observed then the following conclusions may still be valid, only we would have to conclude that the interaction bosons emitted when the charge reversal occurs and which are currently assumed to be elementary are

in fact composed of more elementary particles which would carry the missing matter particle world-lines which appear to be interrupted by the reversal process. It must be clear, however, that under the proposed constraint the charge of a particle could still change on a continuous world-line (as when a blue quark turns into a red quark), because all that is to be required is that a particle does not change into an antiparticle on such a continuous path (the charges cannot reverse), particularly in the case of uninterrupted fermion paths, so that if a particle was propagating forward in time it can still be assumed to propagate in the same direction after the transformation has occurred.

To summarize, if the transformation of a particle into an antiparticle (or vice versa) could occur forward in time on a continuous world-line then there is no way such a hypothetical transformation could be described as an actual change in the particle's properties at the point in time where the transformation occurs, because the phenomenon could only be appropriately described as the encounter of two independent particles approaching the same event from opposite directions in time and nevertheless meeting at a very specific point in space and this is precisely why the requirement of continuity could not be satisfied in such a case. But if we allow for such discontinuity we would then require unlikely coincidences (involving the coordination of distinct forward and backward particle propagation processes) to produce the required meeting of world-lines without any local causality being responsible for this otherwise improbable coordination. Such events would not even be explainable in the way we could explain the chance meeting of two distinct particles at a point in space which would need to occur in the case of a traditional particle-antiparticle annihilation process inappropriately described as the encounter of two opposite-charge particles both propagating forward in time. Now if the kind of processes described above cannot occur then it becomes possible to predict that there should be as many forward in time propagating particles as backward in time propagating particles, so that there should be no fundamental lopsidedness involving the direction of time in our universe. Indeed, if we impose as a condition that there must be continuity of elementary particle world-lines then any forward in time propagating particle world-line present at a given moment must be accompanied by a corresponding backward in time propagating particle world-line, because no forward in time propagating particle can be created without its backward in time propagating counterpart also being created in the process.

But even if highly suitable, this conclusion may at first seem problem-

atic, because there is in effect more forward in time propagating matter than backward in time propagating antimatter in our universe. This observation probably explains why it was never considered that the continuity of particle world-lines may impose that the number of forward in time propagating particles be equal that of backward propagating particles. I would like to suggest, however, that in the context where the existence of negative energy matter is recognized as unavoidable the absence of antimatter in our universe would not rule out the validity of the above discussed conclusion, because we are allowed to assume that the number of backward in time propagating particles could be larger than that of forward propagating particles for negative action matter and if that is effectively the case then the condition of continuity of particle world-lines which requires an equal number of forward and backward in time propagating particles could still be satisfied despite the observed asymmetry between positive action matter and antimatter. The truth would then simply be that the matter-antimatter asymmetry is reversed for negative energy matter (despite the fact that negative energy observers would likely refer to particles propagating forward in time as their own antimatter if those particles are less abundant than backward propagating particles) and that there is actually the same number of otherwise identical positively and negatively charged particles (as observed from the viewpoint of thermodynamic time) when we appropriately take into account the contribution of the unseen negative energy matter. This would in effect be allowed in the context where (as I previously explained) it seems possible for particles of opposite action signs to be permanently created together under the conditions which prevailed during the big bang.

On the basis of the preceding arguments it appears necessary to assume that negative action matter is mostly composed of protons and electrons with charges opposite (from the forward in time viewpoint) that of our most abundant protons and neutrons, a conclusion which is particularly appropriate in the context where any reversal of the sign of action is assumed to leave charge invariant, so that the opposite directions of propagation in time of the most abundant forms of positive and negative action matter should alone determine any difference in the sign of their charges that would be apparent from the unidirectional time viewpoint. What's interesting is that this regained equilibrium between matter and antimatter would have to be observed regardless of the exact nature of the phenomenon which is responsible for the violation of T symmetry that gives rise to the imbalance affecting positive action matter when it is considered independently from negative action mat-

ter. Thus, if the condition of continuity defined above is valid, the number of ordinary matter particles may still be allowed to change independently from that of ordinary antiparticles, but only if there is an opposite variation in the relative number of negative action matter particles over negative action antimatter particles such that the total number of matter particles of *all* action signs remains rigorously equal to the total number of properly defined antimatter particles of all action signs.

What I am suggesting in effect is that this compensation of the observed matter-antimatter asymmetry made possible by the presence of negative action matter is not just a mere possibility, but that the requirement identified above actually implies that there must necessarily be an equal number of forward and backward in time propagating charges when all possible forms of matter are considered together. The direction of entropy growth in our universe would thus correspond to the direction of propagation in time of the most abundant form of positive action matter, but also to the direction of propagation in time of the less abundant form of negative action matter and this contributes to somewhat restore the required symmetry that is lost as a consequence of the existence of a thermodynamic arrow of time. The apparent asymmetry between matter and antimatter would merely be a consequence of the fact that the presence of appropriately conceived negative action matter is not taken into consideration by traditional models. The plausibility of the identified requirement concerning the continuity of particle world-lines is therefore what allows me to conclude that the matter-antimatter asymmetry characterizing our universe cannot be used to identify a fundamental lopsidedness in time (assuming that there is a correspondence of the thermodynamic arrows of time, independent of the sign of time intervals, for positive and negative action matter). But given that I have argued (based on independent motives) against the possibility of absolutely (non-relationally) defined space and time directions occurring at a fundamental level we may consider that the solution of the issue discussed here is a confirmation of the validity of the hypothesis involved.

It must be understood, however, that if the universe is required to be invariant under PTC , as is unavoidable under the above proposed alternative definition of those discrete symmetry operations, then any asymmetry associated with the sign of charge would have to be compensated by asymmetries associated with other physical parameters of matter with the same sign of action, because none of P , T , or C involve a reversal of action. Invariance under PTC cannot in this context be invoked as possibly requiring a compen-

sation of the matter-antimatter asymmetry affecting positive action matter by some asymmetry involving negative action matter, because invariance under PTC is preserved independently from invariance under $M_I M_P M_T M_C$ and therefore only the above described constraint of continuity of particle world-lines effectively requires that there is a compensation between the lopsidedness of positive energy matter and that of negative energy matter. It should be noted as well that if the above conclusion is valid then we should expect that there would not only be an equal number of forward and backward propagating matter particles of any type, but also that there would be an equal number of positive and negative action particles. Given that the sign of action is the significant parameter when the gravitational interaction is concerned, this can be considered appropriate despite the fact that the discussed constraint would imply that there are actually more positive energy particles than there are negative energy particles propagating in any direction of time (because there would then be more positive action particles propagating positive energy forward in time, but also more negative action particles propagating positive energies backward in time). It would then remain to establish if the prediction that there should be an equal number of positive and negative action matter particles in our universe is viable from an observational viewpoint.

One less obvious consequence which would emerge if the above proposed solution of the problem of matter-antimatter asymmetry is valid is that there should then *necessarily* exist conditions in our universe under which positive and negative action states could transform into each other when the appropriate reversal of time is involved. This is indeed a consequence of requiring a continuity of elementary particle world-lines, because given the observed imbalance between the number of positive action particles and antiparticles it must be assumed that some of the bifurcation points in time involve pairs of opposite action particles. This may appear to contradict the previously discussed conclusion that processes of pair creation involving opposite action particles cannot occur as permanent outcomes under ordinary conditions (because such processes are not thermodynamically favored). But the only conclusion we can draw from the above analysis is that it must actually be possible, when the density of matter is very high and space is expanding sufficiently rapidly, for a positive action particle to reverse its direction of propagation in time without reversing its energy sign (from a bidirectional viewpoint) even if it does not immediately annihilate back to the vacuum. It should be clear in this context that the condition of continuity of the flow of

time cannot *alone* be invoked for requiring that a particle reverses its energy sign when it reverses its direction of propagation in time (so that its sign of action may remain invariant), because it is not required that a given particle propagating in the past direction of time always have negative energy like it is required that the same particle propagating in this same direction of time always have the same sign of charge. Indeed, the postulated invariance of the sign of charge under a reversal of the direction of propagation in time is what allows the existence of the time direction degree of freedom to have physical significance when a condition of continuity is imposed, but no such constraint is required for the variation of the sign of energy. The creation and the annihilation of pairs of opposite action particles propagating in opposite directions of time is therefore not independently ruled out by the condition of continuity.

It is now possible to understand the significance of the remarks I originally made to the effect that there could be departures from the rule enunciated in the preceding chapter that a particle cannot reverse its direction of propagation in time without also reversing its energy. Indeed, this tenth principle was formulated under the assumption that it may no longer be valid under the very unusual conditions where the energy of matter is sufficiently large that gravitation is no longer negligible at the elementary particle level (so that the indirect gravitational interactions between opposite action particles could allow the forbidden transmutations despite the absence of contact between those particles). This exception to principle 10 would, in the context of the preceding discussion, constitute an actual requirement given that it is needed to restore the symmetry of our universe under a reversal of time. It is therefore possible to independently confirm that the whole explanation for the absence of creation of matter out of the vacuum which is embodied in this tenth principle is fully appropriate even in the context where the constraint it expresses is mostly of a practical nature and does not constitute an absolute requirement that would be valid under absolutely all circumstances. But what is remarkable is that despite the conditions of short duration and extreme energy density under which the creation and the annihilation of pairs of opposite action particles would be expected to have occurred, it seems to be appropriate to assume that the requirement of continuity of particle world-lines would still be fulfilled, given the very possibility that this assumption offers to solve the problem of the asymmetry between matter and antimatter, which occurs in the context where the distinction between those two forms of matter originates from the existence of a fundamental time direction degree

of freedom whose preferred direction could have been related to the direction singled out by the thermodynamic arrow of time.

3.11 Black hole entropy

We are now entering the realm of a more uncertain domain of scientific enquiry where classical gravitation theory reaches the limits imposed by quantum indeterminacy. In order for the following discussion to be meaningful it will first be necessary to recognize that the theoretical justifications and the indirect evidence for the existence of black holes is sufficiently well established that these objects can be considered legitimate subjects of study. The objective I will try to achieve is then simply to show that it is possible to identify the degrees of freedom of matter which give rise to the exact measure of black hole entropy derived from the semi-classical theory of black hole thermodynamics. This explanation will be based on the results achieved in the previous sections while deriving an improved formulation of the discrete symmetry operations, as well as on a better understanding of the implicit assumptions entering the derivation of the semi-classical formula for black hole entropy. More specifically, I will explain that based on certain plausible hypotheses concerning the constraints that should apply on matter particles approaching a spacetime singularity, it is possible to deduce that a finite number of discrete degrees of freedom characterizes the microscopic states of elementary particles which are captured by the gravitational field of a black hole. As a consequence, it becomes possible to actually confirm the existence of an exact relationship between those matter degrees of freedom and the binary measure of information or entropy which according to the semi-classical theory should be distinctive of those situations in which event horizons are effectively present.

I will be working here under the hypothesis (now commonly recognized as appropriate) that the information concerning the matter which produced the gravitational collapse that gave rise to a black hole (or the matter which was later captured by the same object) is not lost, but is rather contained in the detailed microscopic configuration of certain degrees of freedom associated with the event horizon of the object. Ignorance of this microscopic configuration when a black hole is described using the classical macroscopic physical parameters of total mass, angular momentum and charge is what gives rise to gravitational entropy. What is not fully understood presently

is how we can reconcile the fact that matter appears to be characterized by physical parameters that vary in a continuous fashion, while the information contained in the microscopic degrees of freedom on the surface of a black hole must be given in binary units. What is the exact nature of the fundamental degrees of freedom of matter that should be associated with the information encoded in the physical degrees of freedom present on the event horizon of a black hole? Given the limitations imposed by the Bekenstein bound it would appear that this question actually applies to the microscopic configuration of matter contained within *any* region enclosed by a surface through which information about this microscopic state must be obtained.

It therefore seems that the problem of identifying the fundamental degrees of freedom of matter which are associated with the binary measure of entropy encoded on a two-dimensional boundary is not one that concerns only situations in which black holes are present, even though its significance is made more obvious when we are effectively dealing with event horizons. I think that the fact that there is a similar measure of gravitational entropy associated with both event horizons and ordinary surfaces means that we must admit the reality of what would be occurring beyond the limits of any event horizon, despite the fact that the processes involved cannot be subject to direct observation. Thus, regardless of the practical limitations which clearly exist for actually determining the exact state of whatever microscopic degrees of freedom are to be associated with the particular measure of information encoded on the surface of a black hole, this problem should nevertheless be considered a tangible one, even if only because under appropriate conditions information about this microscopic state could be obtained. In fact, I believe that the constraints imposed by quantum theory concerning the conservation of information require that we recognize the reality of the microscopic degrees of freedom which encode all the relevant information about the matter which was captured by the gravitational field of a black hole and whose existence appears to be necessary for the consistency of the semi-classical theory.

Indeed, what the semi-classical theory of black hole thermodynamics implies is that there does exist information about what lies behind event horizons, but that this information is missing from the description of a black hole in terms of its classical macroscopic parameters and therefore we must assume that it could only be obtained through measurements of the microscopic configuration of some physical parameters associated with the surface delimited by the event horizon of the object. The fact that a consistent theory of black hole thermodynamics effectively exists means that we have no rea-

son to expect that when such objects are involved there could be departures from the rules which govern ordinary physical systems with a large number of degrees of freedom, for which it is already recognized that any apparent information loss merely occurs as a practical limitation. In the context where it is understood that, from a physical viewpoint, information must involve a distinction, this assumption is effectively supported by the existence of a relation between the mass of a black hole and its entropy, because microscopic distinctions must be carried by elementary particles and when the number of particles absorbed by a black hole grows its mass necessarily becomes larger. This observation would remain significant even if it was determined that the actual microscopic degrees of freedom which are allowed to vary for matter that fell into a black hole do not consist of mere energy differences. Also, if we recognize that information, as a physical distinction, can be conserved without the knowledge of some such distinction being shared by any specific observer then we are certainly allowed to assume that information persists even when black holes are involved.

Some well-known results appear to confirm that the information about the state of the matter which was captured by the gravitational field of a black hole can effectively be encoded in the detailed configuration of certain degrees of freedom associated with the event horizon of the object. Those conclusions are all dependent, basically, on one assumption, which is that there is a finite maximum level of accuracy applying to our description of spatial distances. This limitation would then also apply to the description of surfaces such as those which are associated with event horizons. Indeed, the still largely uncertain quantum gravitational theories which were used to achieve those results all have as a key characteristic that they involve a discrete description of physical space on the shortest scale. Based on what I have learned concerning this issue I think that I can safely argue that it is this unique particularity of current quantum gravitation theories which allows to explain that they can predict that black hole event horizons are characterized by a finite number of microscopic degrees of freedom which vary as binary parameters and which appear to encode the information about the unknown microscopic state of the matter contained within the objects. Current quantum theories of gravitation would therefore have succeeded in unveiling at least one distinctive aspect of the structure of space and its associated gravitational field in the context where quantum indefiniteness can no longer be ignored.

What was learned, more exactly, is that two events must be considered

indiscernible from the viewpoint of any measurement when they would occur within intervals of space and time smaller than the natural scale of quantum gravitational phenomena. We now understand that trying to describe the state of matter and energy at a level of definition of spatial distances and during time intervals more precise than those provided by the Planck scale would constitute a superfluous characterization of physical reality. Despite the fact that this constraint now appears clearly inescapable it is still often ignored, as when someone is talking about what may have happened at a time shorter than the Planck time after the big bang. Here I will assume that the limitations imposed by quantum indeterminacy, which imply the existence of a smallest meaningful spatial distance, constitute a fact which will gradually become as well established as the existence of elementary particles of matter and on which further insights can therefore be based. In such a context it would appear that if the degrees of freedom on the event horizon of a black hole are to be associated with the state of some particles (perhaps gravitons) crossing this horizon then in no circumstance could two particles effectively be present at the same moment in a unit of surface smaller than that which is associated with the scale of quantum gravitational phenomena. It would therefore be impossible for any physical parameter associated with such a unit of area to be attributed more than one degree of freedom at any particular time.

Thus it seems that it is from discrete elements of structure with a size of the order of the Planck interval that a proper description of the exact configuration of the microscopic degrees of freedom associated with the event horizon of a black hole can be formulated that may also be valid to some extent in the case of ordinary surfaces. What is remarkable is that, given that the parameters involved can only vary in a binary way, it follows that the physical distinctions associated with this configuration effectively provide a binary measure for the entropy or ignored information which characterizes those objects. Indeed, the relevant microscopic degrees of freedom on a surface can only be this or that, or yes or no, rather than assume any value from a continuous spectrum of possibilities as we go from one discrete surface element to the next. It appears that not only must we accept that space is divided in elementary units on the shortest scale, but we must also recognize that the values taken by the physical parameters associated with those discrete elements of surface can only be either one thing or another and nothing in between. Therefore, at the most fundamental level of description it would appear that the physical properties of a surface must be described

using discrete elements of structure corresponding to the smallest physically meaningful measures of area to which are associated only two possible states of some microscopic degree of freedom. In such a context the entropy of a black hole would derive merely from ignorance of the detailed configuration of this microscopic degree of freedom (characterizing elements of surface on its event horizon) which arises as a consequence of the difficulty to obtain experimental data about what is effectively occurring at this level of precision of measurement.

Given the current state of knowledge concerning quantum gravity, it is not possible to determine the exact physical nature of the elementary degrees of freedom present on an event horizon, but it seems natural to assume that if a black hole of sufficient mass was isolated and matter was no longer crossing its event horizon then this information would have to be contained in the microscopic configuration of the particles mediating the gravitational interaction as they are being released or absorbed by this event horizon. In any case it is necessary to distinguish between the degrees of freedom characterizing the states of the particles which were captured by the gravitational field of a black hole and the degrees of freedom on the event horizon of the object which merely reflect the microscopic state of the matter and which may be of a different nature from a physical viewpoint. But despite this ambiguity it must be assumed that there exists a clear relationship between the state associated with the microscopic degrees of freedom encoded on the event horizon of a black hole and that of the matter from which an observer has become separated as a consequence of the presence of this theoretical boundary. What's more, given the small scale of the elementary units of surface in which the information about the state of matter contained inside an event horizon is encoded, it appears that we would be justified to assume that the degrees of freedom of matter which we must identify are those which would apply to a description of matter at the Planck scale.

In any case I think that the existence of such a correspondence between the microscopic degrees of freedom associated with an event horizon and those of the matter it contains should be considered unavoidable even if only because we can never get more information concerning what is located beyond any surface than is obtainable by observing through this surface. But if there effectively exists a limit to the accuracy of measurements that can be effected on a surface (due to the existence of a smallest meaningful spatial distance) then it necessarily follows that there must be a limit to the amount of information that could be obtained through a detailed probing

of the processes effectively occurring on that surface and this limit should naturally be expected to be proportional to the number of discrete surface elements through which the information must flow. It should not come as a surprise, therefore, that the total area of a black hole must effectively provide a measure of the number of ignored elementary units of information which should ultimately be related to the exact microscopic state of the matter which is located past the event horizon of the object. What's more difficult to explain is why this constraint effectively appears to be relevant to what is actually taking place beyond event horizons rather than merely to what we can *tell* about what is going on there. Despite the enduring uncertainty associated with this question I believe that the following discussion will help clarify the nature of the relationship between the microscopic degrees of freedom on a surface and the microscopic state of the matter located within that surface.

Before I undertake the task of explaining why it is that the states of the elementary particles which have been absorbed by a black hole can become so constrained that they are allowed to match the required binary measure of ignored information which is encoded on the event horizon of the object it would be appropriate to first recall what the semi-classical analysis of black hole thermodynamics has revealed. What we know in effect is that for a non-spinning black hole of mass m with an event horizon of area $A_{BH} = 4\pi R_S^2$, where $R_S = 2mG/c^2$ is the Schwarzschild radius of the black hole, the entropy is given by $S_{BH} = \frac{1}{4}A_{BH}/A_P$, where $A_P = l_P^2$ is the Planck area given in terms of the Planck length which is defined as $l_P = \sqrt{\hbar G/c^3}$ and the units are chosen so that Boltzmann's constant k is equal to unity. In general, a black hole would therefore have an entropy that is determined by the value of the area of its event horizon in Planck units of surface divided by a factor of four. Given that entropy is simply a measure of the information that is missing from the description of a black hole in terms of its macroscopic parameters of mass, radius, or area it seems that the amount of information encoded in the unobserved microscopic degrees of freedom characterizing the surface of the object is equal to one fourth its area in natural units. It was pointed out by Gerard 't Hooft, before the previously mentioned results obtained from quantum gravity were derived, that this effectively means that information appears to be encoded on the surface of the black hole in binary units corresponding to an area equal to four Planck areas.

Now, if we are willing to accept that the Planck unit of area may actu-

ally be given as equal to $A_P = 4\pi l_P^2$ (following the traditional formula for the area of a sphere in terms of its radius) when the mass of a black hole approaches the Planck mass (from higher values associated with macroscopic event horizons) then an interesting result can be shown to follow. Indeed, using the above equation for A_{BH} we can deduce that the event horizon of what I would call an elementary black hole, with a mass equal to exactly one Planck mass $m_P = \sqrt{\hbar c/G}$, should have an area that is effectively equal to four such Planck areas. Using the formula for the entropy of a conventional black hole I would thus be allowed to conclude that the detailed configuration of the microscopic degrees of freedom on the surface of a Planck mass black hole must carry one single binary unit of information. I think that the outcome of this simple derivation is extremely significant, because on the basis of the hypothesis that there can be no significance in attributing existence to a particle which would occupy a volume smaller than that which is associated with the most elementary unit of area (as current quantum gravitational theories appear to require) it seems necessary to assume that such an elementary black hole, would be formed of at most one single elementary particle and in such a case we have no choice but to attribute the information encoded in the microscopic degrees of freedom on the surface of the black hole to its matter content.

But if, in the case of an elementary black hole at least, the ignored information encoded on the event horizon of the object must definitely be associated with the single Planck energy particle it contains (which need not necessarily have a large rest mass) then even for a black hole of larger mass it should be possible to associate microscopic information with particle content despite the fact that according to the above equations the entropy of a black hole S_{BH} is not in general proportional to its mass m , but rather to its mass squared (so that entropy rises faster than the matter content). The fact that no simple relationship between entropy and matter content appears to exist in the general case of a macroscopic black hole is simply due to the fact that the gravitational field must itself carry a portion of the entropy when large accumulations of matter are involved. However, in the context where the particles mediating the gravitational field are to ultimately also be understood as being a form of matter we would have no choice but to associate the entire amount of missing information associated with a black hole's event horizon with the 'matter' content of the object, which would then include gravitons. In any case, if an elementary Planck mass black hole, containing

a most elementary particle with an energy of the order of the Planck energy, can be associated with the smallest unit of information then it requires that we recognize that the detailed microscopic configuration of the degrees of freedom associated with the event horizon of any black hole is attributable to states of matter which can therefore only vary in a discrete way.

So, what are exactly those degrees of freedom prevailing for matter trapped by the gravitational field of a black hole? When we ask this question in the context where the information associated with an *elementary* black hole is understood to provide a complete description of the state of a most elementary particle in the conditions where an event horizon is constraining the motion of this particle it appears necessary to assume that its state must be completely definable by one single binary unit of information. It may therefore appear that we should be seeking to identify a unique physical parameter that reverses under a given discrete symmetry operation as being the binary degree of freedom related to the information encoded on the event horizon of our elementary black hole. But if we are to assume that the same fundamental parameters characterize the spacetime-related properties of matter under all conditions then it rather seems that all the truly independent discrete symmetry operations, like the previously defined T , P , and M operations should have their counterpart in the information associated with the state of the particle forming an elementary black hole. Indeed, all of those reversal operations allow to distinguish the sign or the direction of some physically significant property of elementary particles and there is no a priori reason why only a subset of those variable properties should need to be taken into account in the characterization of the discrete degrees of freedom applying at a fundamental level in the presence of an event horizon.

It must be clear that if all of the independent discrete symmetry operations were considered to determine one distinct degree of freedom of a particle confined by the event horizon of an elementary black hole then we would need not one binary unit of information or one bit to be encoded in the microscopic configuration on the surface of the object, but rather three bits. Indeed, with two yes or no questions we can determine the action sign preserving direction of time intervals (reversed by T or not reversed) and the action sign preserving direction of space intervals (reversed by P or not reversed), which already allows to distinguish four states of matter (identity being the state where neither space nor time is reversed). The distinctions which exist between each of those four states as they appear from the bidirectional and the unidirectional time viewpoints are illustrated in figures 3.1 and

3.2. With an additional yes or no question we can then determine the sign of energy or action (reversed by M or not reversed), which doubles the number of states of matter that can be distinguished, so that we can differentiate between the eight possible states of matter related by the discrete symmetry operations defined in a preceding section. The C symmetry operation being a combination of T and P does not provide an additional distinct degree of freedom and therefore need not be considered here (even though we may as well consider only T and C or only P and C to provide the relevant discrete degrees of freedom and then it would be P or T which could be ignored). But three bits is not equal to one bit and so it may seem that there is a problem with associating the ignored information encoded on the surface of a black hole with the degrees of freedom transformed by the discrete symmetry operations, despite the fact that those parameters should effectively characterize the states of elementary particles under all circumstances.

However, I believe that this discrepancy cannot be assumed to rule out the validity of the theoretically unavoidable conclusion that any binary distinction between the states of the matter particles that crossed the event horizon of a black hole must be a reflection of the structure underlying the previously defined discrete symmetry operations which together allow to transform all physically meaningful states of matter that vary in a binary way. I will show that very restrictive constraints actually limit the variability of certain microscopic physical parameters whenever black holes are involved. Those limitations imply that some parameters which may otherwise appear to be independent from one another actually vary together when subjected to various reversal operations. Some microscopic physical parameters are also restricted to a subset of the values they would otherwise be allowed to take. This effectively contributes to reduce the number of binary units of information needed to specify the microscopic states of particles trapped by the gravitational field of a black hole. A further insight will be needed, however, to allow the number of binary units of information required for achieving this complete description of the state of gravitationally collapsed matter to effectively be made compatible with the measure of black hole entropy derived from the semi-classical theory.

In order to clarify the situation regarding what variations are allowed for the various microscopic properties of elementary particles when matter has become confined by the gravitational field of a black hole we may first recall that the three macroscopic physical parameters characterizing a black hole

are its total mass m , its total charge q , and its overall angular momentum j . To those three parameters I would add the momentum p , which is not usually considered to define the macroscopic state of a black hole, but which I believe provides essential information required to identify the parameters which must be taken into account in defining the microscopic state of the particles that form such an object. It must be clear that each of those macroscopic parameters must be allowed to vary not just in magnitude, but also in sign or in direction. The total mass m in particular must be conceived as being either positive or negative depending on the overall sign of energy of the black hole. This is also an aspect that is usually not taken into consideration in the conventional treatment of black hole thermodynamics, but which must be recognized as a necessary assumption in the context where the existence of negative energy matter is theoretically unavoidable.

A different question would be to ask whether the sign of energy or action is a variable parameter for the particles forming a black hole. Given that I have already argued that negative energy matter cannot be absorbed by a positive energy black hole it may seem that only positive energy states need to be taken into account for describing the microscopic configuration of the matter that was captured by the gravitational field of a positive mass black hole. It is important to understand, however, that we cannot assume that all black holes with a given mass sign must at all times be formed only from particles with the same mass sign as that of the object itself, because even if no particle of energy sign opposite that of a given black hole can cross its event horizon from the outside, the possibility that a positive energy black hole may already contain negative energy matter (or vice versa) is very real and must be taken into consideration. Indeed, it is indisputable that a positive mass black hole with a very large radius and a rather low density could potentially form despite the initial presence of some comparatively small amount of negative energy matter inside the surface that is to become its event horizon. Thus, it is not strictly forbidden for a positive energy black hole to contain negative energy matter even though this matter would only be allowed to be present inside the event horizon associated with such a black hole if it was already contained inside the surface that became this event horizon before the gravitational collapse occurred.

Of course, even if a positive energy black hole was to contain negative energy matter this matter would not remain in this situation for very long, because it would rapidly be expelled by a gravitationally repulsive force equivalent in strength to that which is attracting the rest of matter toward the

central singularity, so that the black hole would effectively end up containing exclusively particles having the same sign of energy as its own. Furthermore, even if the sign of energy of the particles contained within any surface was to constitute a relevant degree of freedom (transformed by the M symmetry operation) that could contribute to the measure of microscopic information encoded on this surface, there is an independent motive for assuming that black holes are composed of matter with a sign of energy that is necessarily the same as that associated with their own total mass. Indeed if I want to explain the results of the semi-classical theory of black hole thermodynamics I have no choice but to first assume that the energy sign of every matter particle forming a black hole corresponds to the energy sign of the object itself, because the conventional theory is based on the implicit hypothesis that positive energy black holes exist in a stable state and are not in the process of releasing negative energy matter, which means that they must be formed exclusively of positive energy matter. Of course when we are considering the case of an elementary black hole it is quite unavoidable that the sign of energy of its constituent particle should be constrained by the sign of the generalized mass parameter m given that the sign of energy of this unique particle must alone determine that of the object. What I am assuming here, however, is that even for macroscopic black holes we are allowed to assume that a knowledge of the sign of mass of the object would determine the energy signs of *all* the matter particles whose states are encoded on the event horizon of the object.

It should be clear that under such conditions it cannot be assumed that the energy sign of particles, which is transformed by the action sign reversal symmetry operation M , constitutes the one binary degree of freedom per elementary unit of area which is associated with the measure of black hole entropy provided by the semi-classical theory, because if that was the case then all the microscopic physical parameters of a black hole would be fixed by knowledge of the sign of mass of the object and no information could be missing from the macroscopic description. It would thus follow that entropy would always be minimum, which is certainly not desirable given that the semi-classical theory rather requires entropy to be maximum when matter collapses into a black hole. The constraint imposed by the sign of mass of a black hole on the energy sign of its constituent particles may not be so significant, however, given that even if we ignore any additional degree of freedom which could be associated with the other discrete symmetry operations, a determination of the sign of energy cannot alone be considered to exhaust

the requirements of a complete description of the state of the matter particles forming a black hole, because in principle energy must also be allowed to vary in magnitude.

For now, we may choose to leave aside that difficulty, but then we are still left with having to explain how it may be that the other two independent discrete symmetry operations which should also characterize the states of matter under all conditions provide at most only one single binary unit of information, even though they together transform two degrees of freedom. As I suggested above, those two symmetry operations may be chosen to be the action sign preserving time reversal operation T and the action sign preserving space reversal operation P . You may recall that in the context of the redefinition of the discrete symmetry operations which I proposed in a previous section the T symmetry operation must be assumed to reverse all momenta, as well as all angular momenta and all non-gravitational charges, even if merely from the unidirectional time viewpoint. The P operation on the other hand has absolutely no effect on the direction of angular momentum or the sign of charge, from any viewpoint, but must be considered to reverse the direction of momentum and the handedness of particles (as indicated in table 3.4). Thus, taken together the T and P symmetry operations would affect all the physical parameters defining the microscopic states of particles which add up to produce the total momentum p , angular momentum j , and charge q parameters that characterize the macroscopic properties of a black hole with a given sign of mass.

Yet this does not necessarily mean that all that must be specified to determine all of those macroscopic physical parameters are the signs of the microscopic parameters transformed by the T and P operations (reversed or not reversed for each of the two symmetry operations) for every elementary particle that forms a given black hole. There is, in effect, no a priori reason to assume that the momentum of elementary particles (like their energy) can vary only in sign and it would rather seem that not only must this parameter be allowed to vary in magnitude like energy, but it must in addition be allowed to vary in direction, not in a binary way like energy, but as a continuous two-dimensional angular variable, which would forbid its complete determination through a knowledge of the value that would be taken by one single binary degree of freedom. What's more, even if under ordinary circumstances an action sign preserving reversal of time intervals generated by T would affect the direction of angular momentum (because it would reverse momentum independently from position) it would not affect the handedness

of particles and in the context where we are trying to identify the microscopic configuration associated with the states of elementary particles present at the Planck scale it appears necessary to restrict our account of the spin state of matter particles to handedness. But while an action sign preserving reversal of space intervals obtained by applying P would effectively reverse the handedness (because it would reverse the momentum of particles without also reversing their spin), we have no reason to assume that the spin could not itself reverse independently from momentum, thereby also reversing the handedness. It would then appear necessary to specify the handedness of particles independently from the other degrees of freedom which are reversed by those two symmetry operations. As a consequence, only the sign of charge of a given particle can be assumed to be entirely determined by its dependence on the redefined time reversal symmetry operation T when the effects of such a transformation are considered from the unidirectional time viewpoint, which usually applies in a classical context.

Now, despite the fact that the T operation reverses both momentum and charge, it certainly seems appropriate to assume that as far as those microscopic physical parameters are concerned we are effectively dealing with two distinct degrees of freedom, because momentum can also be independently reversed by the P operation. But even though it may appear obvious that the sign of charge should be independent from the direction of momentum it is reassuring to observe that from a bidirectional time viewpoint this hypothesis is unavoidable given that the variation of the sign of charge would in fact occur as a consequence of a reversal of time intervals obtained while leaving the sign of action invariant, which would effectively reverse the energy but leave invariant the direction of momentum. In any case the outcome of this reflection is that we have to accommodate three independent microscopic degrees of freedom which are the sign of charge, the direction of momentum and the handedness of elementary particles. The variation of any other physical parameter (except for the sign of action) can then be derived from a knowledge of the variation of those three independent parameters. It is also important to mention that despite the fact that we are seeking to determine the degrees of freedom which would apply on a very small scale at which the fundamental interactions would presumably be unified, I am nevertheless assuming that the sign of any non-gravitational charge would remain a parameter distinct from the sign of action (the gravitational charge), because the variation of the sign of charge would here occur merely as a secondary consequence of a reversal of the direction of propagation in time which must still be considered

a significant change clearly distinct from a reversal of action, even under such conditions.

If we recognize the appropriateness of those remarks it would then seem that the situation may be even more problematic than I indicated above, because despite the fact that we are considering as relevant only those parameters which are affected by the redefined discrete symmetry operations, the degree of freedom associated with handedness would provide an independent contribution to the measure of information concerning the microscopic state of the matter that crossed the event horizon of a black hole. This contribution would add to those provided by the degrees of freedom associated with the sign of charge and the momentum of a particle (or the sign of time intervals and the momentum from a bidirectional time viewpoint). It would then seem that we still need at least three binary units of information to completely describe even just the *signs* of all the relevant physical parameters characterizing the microscopic state of matter under such conditions. But, as I will explain, the existence of an independent degree of freedom related to handedness, far from creating a problem is in fact precisely what allows the appropriate measure of entropy to be derived in the presence of an event horizon.

It is while I was trying to visualize what would happen to a negative mass body which would find itself inside some surface that was about to become the event horizon of a positive mass black hole that I realized that both positive and negative mass particles would actually be submitted to very restrictive constraints when experiencing the effects of the gravitational field which exists inside the region delimited by the event horizon of a black hole. Indeed, a negative energy particle which would happen to be located near the center of a positive mass black hole at the time of its formation would immediately be repelled outward by a force as large as that it would experience inside the most powerful of particle accelerators. While it is being ejected outside the event horizon, the negative energy particle would reach an arbitrarily high (negative) energy and its momentum would also become arbitrarily large in the direction opposite the forming central singularity (considered as the point where the density of the dominant form of energy reaches its theoretical limit), regardless of what its initial state of motion was. The nearer to the center of the object the particle would initially be, the larger its final negative energy would be when it would emerge from the event horizon of the positive mass black hole. But given that the force which accelerates the

particle is always directed away from the forming singularity, it follows that the components of the momentum in any other direction would become completely negligible in comparison with the component directed away from the singularity. In fact, if we were to consider only particles literally emerging from a positive mass singularity, we would end up (in the absence of collisions with infalling positive energy matter) obtaining particles reaching the event horizon with a maximum (negative) energy and a momentum invariably directed along the positive normal to the surface of the black hole. In other words, we would always obtain (in the absence of interference) particles in a very specific state of motion.

The process would be even more constraining for a positive mass body in the gravitational field of a positive mass black hole given the rising tidal effect which in this case compresses the object laterally and stretches it vertically as it is accelerated in the direction of the singularity. In such a case we would necessarily end up with a very focused beam of particles whose lateral motions would again be completely negligible. Indeed, the force attracting the particles toward the singularity of the black hole would grow with time from the moment they cross the event horizon, eventually becoming so large that the energy of the particles would become as high as it can be, while the horizontal components of their momenta would become completely negligible in comparison with the vertical component of their momenta oriented toward the central singularity. Any residual lateral motion would simply contribute to increase or decrease the total angular momentum of the black hole whose rotation is shared by all the particles that fell into the singularity (as a consequence of collisions and relativistic frame dragging) and should not therefore contribute to entropy (as a measure of missing information concerning microscopic degrees of freedom). Thus when a positive energy particle reaches the singularity of a positive mass black hole its momentum (in the frame of reference relative to which the object is not rotating) has basically become a unidirectional variable. In fact, space itself must be considered to become analogous to unidirectional time for a positive energy particle that crosses the event horizon of a positive mass black hole, but what I came to understand is that this effectively means that momentum would then become a fixed parameter with a unique direction and a maximum magnitude. As a result, we once again obtain a unique final state of maximum energy and invariant momentum.

The crucial assumption in the present context is that a maximum energy must effectively exist. I believe that this conjecture is appropriate given that

in a quantum gravitational context the existence of a minimum meaningful time interval or spatial distance implies the existence of a similar limit for the magnitude of energy. Indeed, if we are considering the state of a particle that occupies a region of space of the order of that which is associated with the smallest physically meaningful spatial distance then upon reaching an energy of the order of the Planck energy the particle would itself become a black hole. Thus, if the particle was to gain even more energy the area of the event horizon enclosing it would simply grow to encompass a larger region, which we could only associate with the presence of a larger number of elementary particles instead of assigning the original single particle with a larger energy. Hence, there would be no sense in attributing one single elementary particle at the Planck scale with an energy larger than the Planck energy. The situation we encounter here is somewhat similar to that which we have in quantum chromodynamics, where beyond a certain threshold the energy spent at trying to separate two oppositely charged quarks in a meson no longer contributes to increase the distance-dependent attractive force between the two quarks, but merely end up splitting the original particle into two new mesons thereby neutralizing the force that existed between the original two quarks.

I would therefore suggest that we assume that the particles that reach a singularity after having been accelerated by its gravitational field must be in a state of maximum energy which we must recognize as the Planck energy. Given that it is not that difficult to visualize what would happen to a positive energy particle which would cross the event horizon of a positive mass black hole it is surprising that it had never been fully realized what the outcome of such a process would mean for any description of the final state of a gravitational collapse. But I believe that it is crucial to recognize, in order to clarify the whole question of black hole entropy, that what happens when matter collapses to form a black hole is that it invariably reaches a state in which every particle has a Planck energy and a correspondingly large momentum characterized by a unique invariant direction which is straight toward the singularity, regardless of the initial motion of the particles at the time when they crossed the event horizon of the black hole.

Here it must be understood that despite the fact that the wavelength of the light emitted by a positive energy particle which is about to be absorbed by a positive mass black hole would be infinitely redshifted (from the viewpoint of a motionless observer far from the event horizon of the object) and would show time as standing still, we are nevertheless allowed to as-

sume that the events occurring after the particle crosses the event horizon of the black hole can be characterized in a physically meaningful way. It is certainly appropriate to consider, in particular, that a particle's momentum will effectively keep increasing in the direction toward the singularity, as I am suggesting would be the case, because time dilation does not mean that the particle itself would become motionless, but merely that the signals it emits are infinitely redshifted by the gravitational field of the black hole (still from the viewpoint of a remotely located, motionless observer). Thus, despite the fact that signals would show the particle as apparently immobilized on the event horizon we must still assume that this particle effectively crosses the event horizon and in a finite time acquires an energy which relative to a motionless outside observer would be arbitrarily high.

Also, the notion that, from the viewpoint of an external observer, a positive energy particle could in fact acquire a negative energy after it crosses the event horizon of a black hole would only apply if we were to consider that the negative gravitational potential energy reduces the energy of the particle itself into negative territory. But in fact this is not an appropriate approach to define the energy of matter (especially in the context where the true properties of negative energy matter are understood to make such a notion implausible) given that as far as this potential energy is concerned we are actually dealing with a distinct contribution to energy which is that of the gravitational field. The truth is that the kinetic energy of the particle itself would keep increasing even if it is compensated by a growing negative contribution to the energy of the gravitational field which it produces.

However, one may perhaps question the conclusion that momentum would have a fixed magnitude for any positive energy particle that reaches the singularity of a positive mass black hole in the context where the rest mass itself may be a variable parameter. It is true in effect that the magnitude of this momentum would depend on the rest mass of the particle which is accelerated in the gravitational field of the black hole given that all masses have the same acceleration and therefore acquire the same velocity. But in the context where we are dealing with kinetic energies which are so large it appears appropriate to assume that the energy associated with the rest mass of the particles which are reaching a singularity after having been absorbed by a black hole would be negligible or null, so that if the total energy of those particles is the Planck energy (the maximum physically meaningful measure of energy) we are allowed to conclude that the final magnitude of momentum would always be what we may call a Planck momentum, understood as the

physical equivalent of the Planck energy which would define the maximum theoretically meaningful value of momentum (associated quantum mechanically with the smallest meaningful measure of space interval). Under such conditions we would have no choice but to recognize that the final magnitude of the momentum of all particles absorbed by a black hole should effectively be an invariant property, just like the direction of this momentum.

Now, it should be clear that if space effectively comes to an end for matter that reaches a black hole singularity, then there is certainly no sense in arguing that momentum, as the conjugate attribute to space, could continue to evolve after the final stages of a gravitational collapse. In such a context there is not much sense in arguing that the energy of particles reaching a singularity could itself be subject to any classically significant changes that would need to be reflected in the configuration of the microscopic degrees of freedom on the event horizon of the black hole, even if time (as the conjugate attribute to energy) does not come to an end at a singularity. Indeed, even if changes of a classically well-defined nature could occur after the formation of a singularity, when new matter collapses on it, the fact that the area of the event horizon of the black hole harboring this singularity would increase as a result of this incoming flow of energy means that any changes to the singularity would then be reflected in a corresponding change involving the event horizon. As a consequence, in the absence of new matter input there would be no motive to consider any change to the state of matter reached during the last stages of collapse into the singularity and therefore this final state can always be considered to be the state that must be reflected in the microscopic configuration of the gravitational field on the event horizon of an isolated, stable-state black hole. Under such circumstances it emerges that the microscopic configuration of the degrees of freedom present on the event horizon of a black hole constitutes a kind of magnification of the state of matter at the time it reached the singularity of the object. It may therefore be appropriate to assume that the radial position of a matter particle at the moment it reaches a black hole singularity is what determines which element on the surface of the object contains the information concerning the state of that specific particle. But if each particle that fell through the event horizon of a black hole effectively reaches the singularity at a unique radial position, as appears to be required by the above argument, we would then have an explanation for why it is possible for the microscopic state of the matter contained in the entire volume of a black hole to be encoded in the two-dimensional boundary delimited by its event horizon.

If we agree on the plausibility of the above conclusions concerning the final state of any particle involved in a gravitational collapse, then we need to recognize that the consequences of the assumption that the sign of mass of a black hole would determine the sign of energy of each of its constituent particles are much more dramatic than one may have expected. Indeed, it now appears that not only must the signs of energy of the particles in the final state of a gravitational collapse be considered to be completely determined by a knowledge of the sign of mass of the object, as I already suggested, but the magnitude of those energies is also to be considered an invariant parameter, which therefore cannot contribute to the entropy of the black hole. What's more, a similar conclusion applies for the momenta of those particles present in the final stages of a gravitational collapse which must be considered to be completely determined not just in magnitude, but also in direction, once the sign of mass of the black hole is known. Therefore, the momenta of the particles which crossed the event horizon of a black hole, like their energies, cannot contribute to the measure of entropy or ignored information which determines the temperature of the object. Only in the presence of negative energy matter would the direction of the momentum of particles be allowed to vary (in sign) inside the region delimited by the event horizon of a positive mass black hole. Indeed, even if only one magnitude of energy can be considered significant for matter located inside the event horizon of a black hole this would not exhaust the number of possibilities regarding the momentum states of particles in the context where the energy could effectively be either positive or negative, because the direction of momentum would then depend on the sign of energy of the particles. But if we are to concentrate on accounting for the microscopic configuration of black holes which have reached a stable state then this possibility can effectively be ignored.

As a matter of fact if the only parameters relevant to a description of the microscopic state of the matter encoded on the event horizon of a black hole were the sign of energy (or action) and the direction of momentum, then given that only one magnitude of energy is allowed for matter that reached a singularity, we would have to conclude that stable-state black holes have minimum entropy, because they only contain matter of one particular energy sign which also happens to determine the direction of all the momenta. Indeed, for a stable-state black hole, the microscopic states of energy and momentum of all the matter particles can be fully determined merely by providing the *sign* of energy of the black hole. Under such conditions there

would be no meaning in trying to associate the energy sign or momentum direction degrees of freedom with some measure of entropy that would be significant from a thermodynamic viewpoint. Yet given that the case of stable-state black holes is more representative of the situation we have in practice when event horizons are effectively present, I would argue that this difficulty does not mean that such black holes cannot be used to derive very general results, but rather that it is necessary to recognize the relevance of additional degrees of freedom also transformed by the discrete symmetry operations. In other words, the microscopic degrees of freedom of matter which are reflected in the detailed configuration of the gravitational field on the event horizon of a stable-state black hole simply cannot be those which are associated with the energies and the momenta of the particles that collapsed to form the object.

I am therefore allowed to conclude that it must be the remaining degrees of freedom, which I previously identified as being the sign of charge and the handedness of matter particles, that would freely vary for particles which have reached the final stages of gravitational collapse and in such a way potentially contribute to the microscopic information content of a black hole. Given that momentum direction itself is fixed it seems that the handedness of particles could effectively allow to determine one binary degree of freedom which would vary upon a reversal of the direction of spin. Indeed, when the direction of momentum is fixed the variable direction of spin relative to this momentum direction is the only parameter that can still vary. But then what about the contribution by the sign of charge of those particles that crossed the event horizon of a black hole? Shouldn't this free parameter also contribute to the measure of entropy derived from the semi-classical theory of black hole thermodynamics? Even if we assume that there is only one type of charge for the elementary particles present at the unification scale, certainly information should be needed to specify whether this charge is positive or negative, given that charge appears to reverse when the direction of time intervals is itself reversed.

I have explained why one binary unit of information would be enough to account for all but one of the fundamental degrees of freedom of any positive energy particle present in the final stages of gravitational collapse, but it would seem that another bit is required in each corresponding elementary unit of area on the event horizon of the associated black hole to determine either the handedness of particles or the sign of charge associated with the direction of propagation in time of those particles, which clearly varies inde-

pendently from that of momentum (which is associated with the direction of propagation in space). We have gone from four bits to two bits per unit of area, but that is still one bit away from the single bit that the semi-classical theory of black hole thermodynamics indicates must be encoded in the detailed configuration of the gravitational field on the surface of an elementary black hole. Given that what I am seeking to allow is a complete determination of all the physical properties of the particles present inside the event horizon of a black hole from a knowledge of the value of all the relevant discrete degrees of freedom it would seem that I have fell short of this objective. I would like to suggest, however, that in fact the problem we seem to have encountered is not real.

The truth is that there is no contradiction between my account of the quantity of information required to completely describe the state of an elementary particle which was captured by the gravitational field of a black hole and the measure of ignored information encoded in the microscopic state of the gravitational field on the event horizon of such an object. To understand what motivates this conclusion we must first acknowledge that the formula for black hole entropy was derived from arguments related merely to the thermodynamic properties of the gravitational field itself and only in the context where the measure of information involved must be used in setting the strictly thermodynamic relationships between quantities like the surface gravitational field and the temperature of the thermal radiation emitted by a black hole. But if it is effectively the case that only one out of two bits concerning the state of matter contained within an elementary black hole is encoded in the detailed configuration of the gravitational field on its surface I think that this is because there is more information encoded in some other physical properties of black holes that do not contribute to the measure of entropy provided by the semi-classical theory of black hole thermodynamics and associated merely with their surface gravitational fields.

Once we have recognized that there must be more information about the state of matter contained within a surface than is provided by the detailed configuration of the gravitational field on that surface what becomes crucial to understand is that there is no reason to assume that the gravitational field should provide information about the microscopic distribution of non-gravitational charge, because that information must actually be contained in the detailed microscopic configuration of the field of interaction associated with this charge. It is surprising in fact that this requirement was never considered before, because when one carefully thinks about this question,

it is hard to arrive at a different conclusion. If information about energy, momentum and angular momentum (as the physical properties of particles which constitute the source of gravitational fields) are to be associated with the microscopic state of the gravitational field then it is also quite unavoidable that information about, say, the electric charge is to be associated with similar microscopic aspects of the electromagnetic field. There is in fact absolutely no reason to assume that the detailed configuration of the electric charges which are the source of the electromagnetic field should be determined from information contained in a different force field, which would here be the gravitational field.

Now, if the information concerning the microscopic distribution of electric charges or electric charge signs inside a given surface (whether or not this surface is that of a black hole) can only be encoded in the detailed configuration of the electric field on the boundary delimited by that surface (even when the total charge inside the surface would be null) rather than in the configuration of the gravitational field on the same boundary, then it means that a theory that would seek to derive a measure of the amount of information necessary to determine the state of the matter contained inside this surface based only on features of the gravitational field present on the surface (which in the case of black holes would be the event horizon) would necessarily fall short of providing the accurate value. Therefore, the results derived from the semi-classical theory of black hole thermodynamics concerning the relationship between the entropy and the area of an event horizon (considered as a gravitational phenomenon) would not rule out the existence of an additional amount of missing information associated with the exact microscopic state of the matter trapped within such a surface.

I believe in effect that the information associated with the sign of charge of every particle forming a black hole and which is transformed by the T symmetry operation (from a unidirectional time viewpoint) cannot contribute to the measure of disorder or entropy associated with the gravitational field of the object and this explains that it need not be taken into consideration when deriving the statistical mechanical properties of black holes associated with the various properties of their event horizons. This is why we were allowed to ignore the existence of this information when elaborating the semi-classical theory of black hole thermodynamics from which the conventional measure of black hole entropy was derived. It thus appears that we really have one additional binary unit of information (distinct from that which must be associated with handedness) concerning every elementary particle in the

final stages of gravitational collapse. This information allows to determine the sign of charge of each and every particle which contributes to fix the total charge q of a black hole. We are then allowed to assume that this is the binary unit of information which is effectively associated with the T symmetry operation (or alternatively the C symmetry operation) defined in a previous section.

It would therefore seem that there is in effect more information associated with the microscopic state of the matter contained in a black hole than is encoded in the detailed configuration of the discrete gravitational field degrees of freedom present on the event horizon of the object. But I have explained why we should not expect this information to contribute to the conventionally derived measure of black hole entropy. Instead, the additional information should be associated with the entropy contained in the interaction field associated with the distribution of non-gravitational charge, which would give rise to its own independent contribution to the temperature of a black hole. In this context it is important to note that there effectively exists an analogue to the Hawking radiation process associated with the gravitational field of black holes and which involves the electromagnetic field. It is a known fact indeed that past a certain magnitude the electrostatic field surrounding a nucleus would induce pair creation processes similar to those associated with the radiation emitted by a black hole and I believe that this phenomenon would allow a similar treatment of the thermodynamic properties which according to the above proposal should be associated with any distribution of non-gravitational charge. Only, in the case of non-gravitational charge we are usually dealing with situations where the total charge is effectively null even when large amounts of positive and negative charges are present inside a surface. Such situations are therefore more analogous to that which is occurring when the measure of gravitational entropy is constrained merely by the Bekenstein bound and both positive and negative energy matter may be present together inside a surface.

Regarding those conclusions, I may add that if the non-gravitational charge did not reverse (from the unidirectional time viewpoint) under application of an action sign preserving time reversal operation T , as it does according to the redefinition of the discrete symmetry operations which I proposed in an earlier portion of this chapter, then we would not be able to attribute only two binary units of information to the degrees of freedom of matter particles reaching a black hole singularity. Indeed, when reversing time is not considered to reverse charge (from the unidirectional time view-

point), the charge conjugation operation C is different from a combined space and time reversal operation PT and we need an extra bit of information to distinguish a PT reversed state from a C reversed state, at least when we take into considerations only quantities which are explicitly reversed by those symmetry operations. Therefore, if T was not defined in the way I suggested, the above proposal for accounting for the degrees of freedom of the matter forming a black hole (in particular an elementary black hole with one Planck mass) would not work. I believe that this is a strong motive for considering that the redefinition of the discrete symmetry operations which I proposed in this report, based on independent motives, is totally appropriate and from a theoretical viewpoint in effect constitute an absolute requirement.

If you have understood the essence of my argument, then there should be no doubt that the only information which is effectively encoded in the microscopic configuration of the degrees of freedom associated with the surface gravitational field on the event horizon of a black hole is that which allows to determine the handedness of every particle it contains using merely one single bit of information for every elementary particle. This conclusion should perhaps have been expected given that current theories of quantum gravitation appear to indicate that at the spatial scale we are dealing with here spin networks become the defining structure. In any case, if we are willing to accept the validity of the arguments on which this deduction is based it would then follow that we now have an explanation not only for the fact that the states of the matter particles trapped by the gravitational field of a black hole vary as discrete variables, but also for why only one such variable (instead of three or four) actually contributes to the measure of missing information which must be taken into account in determining the thermodynamic properties of such an object, therefore allowing the measure of information associated with the matter content of an elementary black hole to match the value of entropy derived from the semi-classical theory of black hole thermodynamics which requires each elementary unit of surface (equal to four Planck areas) to encode one binary unit of information.

Therefore, it is now effectively possible to at least confirm the existence of a definite relationship between the microscopic state of the quantized gravitational field which is associated with the missing information encoded on the surface of a black hole and actual states of the matter it contains. What held the key to a better understanding of the exact nature of the degrees of freedom characteristic of the states of matter submitted to a gravitational

collapse was the recognition that for matter particles reaching a black hole singularity the only relevant variables are the signs of all those physical parameters which are transformed by the previously discussed discrete symmetry operations. It is remarkable that the sign of handedness in particular should be one of the only fundamental parameters of elementary particles (along with the sign of charge) that is not constrained to any specific value by the conditions prevailing in the final stages of collapse into a spacetime singularity and that it must therefore alone contribute to the measure of entropy of a black hole. This is certainly the most significant outcome which has emerged from my reexamination of the question of discrete symmetries in a semi-classical context.

If we now return to the more general case in which the density of matter is not large enough to produce an event horizon and the possibility for positive and negative action matter to be present together inside a surface cannot be ignored, it transpires that this is a situation in which more information would be required to describe the microscopic configuration of matter, because more states of motion are allowed for the particles in the period before such a configuration reaches a stable state. Indeed, even when an event horizon associated with a positive mass black hole is present it is clear that while a positive energy particle would be drawn toward the center of mass of the object, its negative energy counterpart if it was present in the same location would be repelled in the exact opposite direction by a force of similar magnitude (to the extent that the average cosmic density of positive energy matter can be neglected). Thus, in such a case, we would need to take into account at least one additional binary degree of freedom associated with the sign of energy of the matter particles present inside the surface, which would also determine their momentum directions.

But this would actually be the simplest case, as more complex momentum states would occur if the matter was not contained within a surface that constitutes a black hole event horizon, because under such conditions not only would the momentum directions of the particles be allowed to vary, but it seems that their magnitudes could also vary significantly. It is important to understand, however, that the validity of the Bekenstein bound would be preserved even if more information was required to determine the exact microscopic state of matter under those less constraining conditions. This is again because while more information may be required to describe the state of matter when the magnitude of energy and the direction of momentum is

not fixed, this information gain would be offset by a decrease in gravitational entropy (the amount of information required to describe the microscopic state of the gravitational field) that would result from the lower density of the matter distribution associated with lower (nearer to zero) particle energies or a mixed configuration of matter of both energy signs.

Now, it may appear contradictory that under ordinary circumstances, when no event horizon appears to be present and the matter distribution is smoother, it is more difficult to tell the energy sign of the particles present within a surface. How could it be more difficult in effect to determine the microscopic state of the matter when it seems that you can actually see or directly probe more of the content of the surface? But actually the presence or the absence of an event horizon has nothing to do with the fact that it may be more or less difficult to identify the microscopic state of the degrees of freedom of the matter which is contained within a surface, as this difficulty arises merely from the fact that the number of such degrees of freedom effectively grows when the density of matter itself grows. Therefore, it appears that the fact that information is missing from the macroscopic description of a black hole is not consequent to the presence of an event horizon, but is rather attributable to the microscopic nature of the degrees of freedom which encode the information about the state of the matter that is trapped inside the object. What is not known to an observer outside a black hole is not what is inside the black hole, but simply the exact microscopic state of the degrees of freedom associated with the event horizon, from which information about what fell into the black hole could be obtained. Of course actually obtaining this information could only occur at the expense of an even larger increase of entropy in the environment, as would be the case for any system in thermal equilibrium from which we would try to obtain fine-grained information, but in principle the operation could be performed.

The validity of this viewpoint can perhaps only be appreciated when we recognize that the classical gravitational field, as it is described in a general relativistic context, is merely the smooth and predictable statistical average of what is actually a randomly fluctuating quantum field which in its ultimate form would be mediated by the exchange of elementary particles. Indeed, if there are local variations in the gravitational field or in the curvature of spacetime above those described by the classical theory then it is only natural that if some property of the field was to be measured in a very precise location this usually unobserved substructure and the information associated with it would become apparent. It is my belief that the existence of such microscopic

degrees of freedom in the gravitational field on a surface is what allows the information about the state of matter located inside an event horizon to be obtained under proper conditions.

In any case, what's most significant regarding those situations where the entropy associated with the gravitational field is not maximum is that we are necessarily dealing with transitional states which will continue to evolve until the configuration described in the preceding paragraphs is reached. Thus the negative energy matter which may be present inside a positive mass black hole will eventually be expelled from the object, while the positive energy matter will necessarily reach the singularity. By releasing all matter with an energy sign opposite its own, a black hole effectively *increases* its total mass and therefore the area of its event horizon and this effectively means that its entropy grows larger in the process. We are therefore in a situation where a black hole containing less matter (but not less mass) can have a larger entropy. This counter-intuitive outcome is allowed because in those situations where matter contributes to diminish rather than increase the gravitational field on a surface (a general surface, not that associated with the event horizon of a black hole) it also contributes to reduce the portion of entropy attributable to the gravitational field on that surface, which apparently contributes more to the total measure of entropy than the matter itself. It should not come as a surprise therefore that when negative energy matter is released outside the surface of a positive mass black hole the total amount of information required to describe both the microscopic state of the matter particles still contained within its surface and their associated gravitational field grows larger. A negative energy particle inside a positive mass black hole does contribute (positively) to the information content of the object, but at the same time it reduces the amount of information attributable to the gravitational field, which happens to be larger than that attributable to the matter, so that overall the amount of information in the black hole is smaller than it would be without the presence of the negative energy particle. From that viewpoint it certainly appears appropriate that negative energy matter cannot be absorbed by a positive energy black hole, given that this would require entropy to decrease.

The more general situation where only the Bekenstein bound may apply is therefore not incompatible with the results I have derived from a study of stable-state black holes from which all matter with an energy sign opposite that of the object has been expelled. In fact, it seems that there is no real difference between the situation we observe in general when opposite energy

particles are necessarily allowed to be present within a surface and that which arises when we are considering the surface delimited by the event horizon of a black hole. Yet the fact that the presence of negative energy matter within a positive energy black hole would only be temporary and would always give way to a more stable state in which only positive energy matter would remain may suggest that such end states play a role in gravitational physics which is analogous to that which is played by thermal equilibrium states in statistical mechanics. But the real question regarding the Bekenstein bound is how it can be that under the more general conditions in which it applies, the energy and the momentum states of matter particles located within a surface would be allowed to vary in a continuous way, in both magnitude and direction, while the measure of information encoded on the surface must still be provided in binary form.

What my investigations have led me to understand is that in fact this freedom is only apparent. It turns out that even under the more general circumstances discussed here the magnitude of the energies and the directions of the momenta of elementary particles are restricted to binary values. What allows me to draw such a bold conclusion is that I have recognized the consequences of the fact that event horizons are actually always present on the shortest distances where quantum fluctuations in energy continuously give rise to the formation of ephemeral Planck mass black holes. It is clear that the fluctuations in energy occurring at the Planck scale do not all by themselves imply that the energy of particles must be fixed to some maximum value, but the fact that such fluctuations are omnipresent when we reach this scale means that elementary black holes are actually the substance of physical space and time at this level of precision of measurement and if that is the case then it means that matter is always shrouded in the event horizons of those microscopic black holes and therefore we can only conclude that locally it is submitted to the same constraints that would apply in the presence of a macroscopic black hole. Thus, locally, the energies involved would always be of the order of the Planck energy, because the particles trapped within those microscopic black holes would be accelerated to arbitrarily high energies by the gravitational fields present on their surfaces. Indeed, the surface gravitational fields which we may associate with black holes of such small masses would be extremely large, therefore compensating for the short time intervals during which they would effectively be allowed to accelerate the particles which are captured by their gravitational fields. It must be clear, however, that there can still occur variations of energy in smaller units than

the Planck energy on larger scales. The Planck energy must not therefore be conceived as a minimum unit of energy, because to the contrary it constitutes a maximum level of energy which must nevertheless be the only significant measure of energy *magnitude* concerning the state of matter at the fundamental level of precision of space and time intervals set by current quantum gravitational theories.

The case of momentum direction is a little more complex, because we are here dealing with a scale at which quantum indefiniteness in position cannot be ignored. This is reflected in the fact that the same elementary unit of surface would actually correspond to every possible direction normal to the surface of an elementary black hole. But even if it may never be possible to associate a classically well-defined direction to the momentum of a particle submitted to the gravitational field of such a microscopic black hole, it remains that quantum mechanically there would exist a definite (but superposed) state of momentum even for particles in such a situation and this state would still be constrained by the configuration of the local gravitational field. In other words, there would still be a constraint on momentum direction to be fixed by the direction of the gravitational field. Thus I believe that when we are considering the states of particles on the scale of an elementary unit of volume, corresponding to an elementary unit of area (equal to four Planck areas), the momentum direction of a particle may still vary only in a discrete way, even when no macroscopic event horizon is apparent on a larger scale.

Indeed, as a result of quantum indeterminacy, it is not possible to specify the direction of the local momenta any more precisely than there are elementary surface elements associated with the microscopic black hole in which a particle is trapped. If the microscopic black hole has two elementary surface elements (and contains two elementary particles), then any query concerning the orientation of the momentum of one of the particles it contains can only indicate along which of the two portions of surface the momentum is pointing. Thus, regardless of how the elementary units of surface on the microscopic black hole are considered to be oriented relative to the features associated with the macroscopic surface containing those black holes, it is always possible to tell only if the momentum of the particle is oriented along this or that unit of surface and not in which direction precisely. So each elementary unit of area on the event horizon of a local microscopic black hole still contains the same amount of information as would an elementary unit of area associated with a macroscopic black hole which is *not* in a stable state and in which

particles with both energy signs could be present. Again, this is true even if it would be possible to define the orientation of the elementary units of microscopic black hole surface in a very large number of ways, because the momentum direction itself is not determined to any better precision. The orientation of the elementary surface elements of the microscopic black hole could vary in a near continuous way, but given that the momenta of the particles constrained by the event horizon of this black hole are in a state of superposition their directions cannot be identified any more precisely than by specifying if they do point in the arbitrarily defined direction of a particular one of the surface elements, regardless of the exact orientation of those units of area. Thus on a local scale there would be a finite number of possibilities (associated with the finite number of surface elements on the microscopic black hole event horizon) for the momentum direction of a particle, which can therefore be specified using binary units of information.

Now, given that the developments introduced in the earlier portions of the present section appear to indicate that there exists a correspondence between the state of a matter particle reaching a black hole singularity and a given precise elementary unit of surface on the event horizon of the object, then it would seem appropriate to consider that a precise unit of area on a macroscopic surface that is not an event horizon should in general also correspond to a specific matter particle inside that surface. In such a context it would be possible to associate the information which allows to identify the direction of the momentum of a particle contained in a microscopic black hole present inside a macroscopic surface with some precise element (or perhaps with a precise group of elements) on that surface. Thus, if all particles inside some surface can be considered to be locally constrained by a microscopic event horizon, then even in the absence of a macroscopic event horizon we would be allowed to assume that the information about the exact state of those particles must be provided in a binary form corresponding to specific elements on the surface enclosing the volume in which the particles are located. But this effectively occurs only when we assume that event horizons must be present locally at the Planck scale so as to constrain the magnitude of the energies and the direction of the momenta of the particles on such a scale.

Of course under such conditions more binary units of information would have to be encoded in the elementary units of area on the macroscopic surface to specify the exact microscopic state of each of the matter particles it contains, because in addition to specifying the handedness of a particle

we would now need to determine its energy sign and also independently its momentum sign, because the momentum direction is dependent on the sign of energy of the local microscopic black hole which must itself be allowed to vary. Therefore, the amount of ignored information associated with the microscopic state of *matter* inside an ordinary surface would be larger than it would be if this surface was the event horizon of a black hole. In fact, the configurations for which the entropy associated only with the signs of energy and the directions of momentum would be the highest would be those where the microscopic surfaces provided by the local event horizons would be the smallest and would be found in the largest number. But given that this occurs when the total energy density of positive and negative energy matter is the smallest and the distribution of matter is as smooth as it can be, so that no macroscopic black hole is present, it follows that there would then be a compensation between the increase of entropy associated with the energy signs and the momentum directions of matter particles and the decrease of entropy related to the weaker surface gravitational field involved, which would still allow the Bekenstein bound to effectively apply.

If this account of the physical degrees of freedom associated with the microscopic information encoded on a surface is accurate it means that we would not be justified to assume that there is no longer anything physically significant going on *at* the Planck scale, because in fact the state of matter associated with an elementary unit of area on the surface of a macroscopic black hole would effectively also characterize the physical reality which exists when we reach the shortest intervals of space and time. I was able to draw this conclusion only at a relatively late stage of my research program, because for a long period I had assumed without much thinking that the possibility that matter could exist in a negative energy state would imply a cancellation of all quantum fluctuations in energy at the Planck scale, which would not allow the presence of microscopic black holes on such a scale. But in fact all that is truly implied by the possibility that negative energy states can be occupied is that the fluctuations in energy can occur in both positive and negative territory. Thus, not only do fluctuations associated with positive and negative energy states not compensate one another at the smallest physically significant scale of space and time, but it seems that their basic distinction effectively provides one of the only significant degrees of freedom applying to matter on such a scale.

The fact that the proposed description of the constraints applying on

states of matter trapped by the gravitational field of stable-state black holes can be generalized, in the particular manner described above, to situations in which the density of matter is lower and particles of opposite energy signs can be present together within a surface strengthens the argument for the existence of a correspondence between certain properties of black holes and general features of the physical systems described by conventional statistical mechanics (the discussion featuring in the following section will add weight to this conclusion). Indeed, I have already pointed out that the situation we have in the presence of a large black hole containing only matter with one energy sign is analogous, from the viewpoint of gravitational entropy, to a state of thermal equilibrium such as we may describe in a conventional statistical physics context. But if we are justified to assume that the proposed description of the microscopic degrees of freedom characterizing stable-state black holes can be generalized by assuming the existence of states (the microscopic black holes) which are locally similar to those equivalent thermodynamic equilibrium states then the analogy could be carried over to the field of non-equilibrium thermodynamics. This is because in effect the basic assumption of the thermodynamic theory of irreversible processes is that even systems evolving irreversibly are to be conceived as being locally in a state of near thermal equilibrium. What we have then is an ensemble of subsystems in a momentary state of near equilibrium exchanging energy and evolving in such a way that static equilibrium is not required at the level of the system as a whole (which in the current analogy would be any matter-enclosing surface) like it would be in equilibrium thermodynamics.

It is true that in the present case the stability of the configurations occurring on the shortest scale would be limited because microscopic black holes are continuously being created and destroyed, but then the local subsystems in the theory of near-equilibrium thermodynamics are also not in states of perfect equilibrium. What is reflected in this particularity is merely the fact that we are here effectively dealing with statistical laws applying to randomly fluctuating systems for which deviations away from thermal equilibrium continuously occur locally even for a system in a state of overall equilibrium. In fact, the situation we would be dealing with in general would be one where a surface may enclose a configuration where a relatively large number of black holes of various sizes and variable stability (including macroscopic black holes) are present and interact with one another. In this context the states of matter particles would be locally constrained, but in a more or less stable way, as in the local subsystems of the theory of non-

equilibrium thermodynamics, while the system as a whole would be allowed to evolve irreversibly through the merger of smaller mass black holes into ever more massive ones with larger event horizons. One could hardly think of a more perfect analogy between two theories and I believe that this is not a coincidence, but rather a clear indication that the proposed application of the insights derived while studying the problem of discrete symmetries in the context of the existence of negative energy matter allows a better understanding of the problem of black hole entropy as a pure thermodynamic phenomenon. It is clear to me that whatever explanation for the quantization of information would be more accurate than the one provided above would have to derive from a more detailed knowledge of quantum gravitation than is currently available.

3.12 Negative temperatures

It is not a widely known fact that while temperatures are usually confined to positive values, it is nevertheless unavoidable that some physical systems be attributed negative temperatures under certain conditions. Those who have considered the issue have recognized in effect that negative measures of temperature must necessarily occur when we are dealing with certain systems characterized by a finite number of energy levels. What happens is that as temperature rises it must in general be assumed that more energy states become available for the constituent particles, so that information or entropy is itself rising. Therefore, entropy must be assumed to be minimum when a system is at zero temperature. But for systems with a finite number of energy levels it turns out that as temperature increases we may reach the point where entropy is maximum and temperature therefore must be considered infinite. This may occur for example in the case of a spin system in a magnetic field where the number of levels of orientation of each nuclei is finite. For such a system the lowest energy configuration is that where all the spins are in the direction of the magnetic field, while the highest energy configuration is that which would occur when all the spins would be oriented in the direction opposite that of the magnetic field. At infinite temperature all spins would be oriented in the most random way, with as many spins oriented in the direction of the magnetic field as there would be in the opposite direction. If we add more energy to a system in such a state we would witness a decrease of its entropy, as more spins would become oriented in the direction opposite

the magnetic field and less information would be required to describe the microscopic state of the system.

Given that temperature merely defines the relationship which exists between energy and entropy, if an increase of energy produces a decrease of entropy then it must necessarily be assumed that the temperature has become negative. But if adding more energy decreases the entropy only slightly when it reaches its maximum point at which the temperature is infinite then it means that the temperature is not ‘minus zero’ but actually ‘minus infinity’. Thus as even more energy is added to the system the entropy would gradually decrease back to a minimum at which point the negative temperature would effectively reach the zero value again. In the case of the spin system this point would be reached when *all* the spins would be oriented in the direction opposite that of the magnetic field and no further change could occur. I may also mention that it was found that when we combine two such systems which happen to have opposite temperatures of equal magnitude the outcome must be a system with infinite temperature. It must be understood that despite common expectation to the effect that temperature is a positive definite quantity, the conclusion that negative temperatures may occur in nature is not just a consequence of adopting some particular definition for what temperature should be or of choosing a particular reference scale for this quantity. Specialists are unequivocal concerning the fact that negative temperatures cannot be avoided in a general context, because they are associated with actual states of any system with a finite number of energy levels.

Now, what I would like to point out is that if the constraints I unveiled in the previous section concerning the microscopic states of matter in the presence of an event horizon are valid, then it would mean that black holes are somewhat similar, from a thermodynamic viewpoint, to those more conventional systems for which negative temperatures are allowed. Indeed, I have explained that in the presence of event horizons the microscopic states of matter can be described using only discrete degrees of freedom, so that a maximum number of microscopic configurations (similar to the energy levels in the conventional theory) must be assumed to exist for black holes of any mass. In fact, given that the number of degrees of freedom of a black hole decreases continuously as they lose mass, it appears that they become increasingly similar to the above described spin systems as they decay through the process of Hawking radiation. This similarity is all the more appropriate given that it would seem that if a positive energy black hole had a positive

value of surface gravitational field, then a negative energy black hole would have a negative value of surface gravitational field and knowing that the surface gravitational field is the quantity which is associated with the temperature of a black hole in the semi-classical theory, I am led to conclude that this temperature itself needs to be allowed to vary not just in magnitude, but also in sign, as would indeed be required on the basis of the fact that a negative mass black hole would radiate particles with an energy sign opposite that of the particles which would be radiated by a positive mass black hole. Thus, if negative energy matter exists, it would seem that some black holes could effectively be attributed negative temperatures which would be made conspicuous by the reversal of their surface gravitational fields.

The correspondence with the above described thermodynamic phenomenon involving spin systems is complete, because as a positive energy black hole evaporates through the emission of thermal radiation and its mass decreases toward zero (in positive territory) its temperature would rise until it becomes infinite when the object reaches the Planck mass at which point if we were to continue to remove energy (by actually adding negative energy) its mass would start to increase into *negative* territory with an initial temperature that would be infinite, but also negative, and which would decrease (toward zero) as the negative mass of the object increases. Of course the dependence of temperature on total energy is reversed in the case of black holes, given that a larger mass black hole would have a lower temperature, but if we consider only the relationships between thermodynamic properties then the analogy is valid. Also, if we were able to combine a positive energy black hole (to which is associated a positive temperature) with a similar negative energy black hole (to which is associated a negative temperature) then what we would obtain is not a zero temperature object, but an object with a larger and possibly infinite temperature (just like when we combine two opposite temperature systems in the conventional theory), because the mass of the resulting configuration would be smaller and such an object would radiate energy at a higher rate. Of course it may not be possible from a practical viewpoint to combine opposite energy black holes so as to cancel their masses, but mathematically the correspondence between the quantities involved is valid and matches the expectations derived from conventional thermodynamics theory.

The fact that the existence of such a beautifully perfect correspondence between the semi-classical theory of black hole thermodynamics and the classical thermodynamics of systems with a finite number of microscopic levels

of energy is allowed to occur under the hypothesis that two signs of mass are relevant for a description of the thermodynamics of black holes constitutes an additional argument for recognizing the legitimacy of this theoretically motivated insight. In fact, I am surprised that the conclusion drawn by specialists concerning the unavoidable character of the concept of negative temperature was never considered to imply that energy itself should be allowed to vary in sign rather than only in magnitude. But as I have always believed that the inherited motivation behind the widespread idea that energy can only be positive originates from the thermodynamic conception of energy as a measure of heat (which is itself a positive definite quantity from a classical viewpoint), I was quite satisfied when I learned that this most thermodynamic concept of all, the temperature, must itself vary in sign. If there is no reason to assume that negative temperatures cannot have a clear significance in physical theory and if it turns out that they must ultimately be associated with the state of objects whose energy is predominantly negative, then we have one less argument to assume that the concept of negative energy itself cannot be given clear meaning.

3.13 Summary

Once again I would like to conclude this chapter by providing a summary of the decisive results which were obtained concerning the problem of discrete symmetries in the context of the alternative approach to fundamental time directionality that was developed in this report. The key results are thus the following.

1. It would violate the requirement of relational definition of physical quantities to consider a reversal of the directions of space and time intervals or those of momentum and angular momentum or of the sign of energy or that of charge that does not occur relatively to some remaining unchanged parameter of the same kind and therefore such changes must be considered impossible.
2. The reversal of space intervals produced by the P symmetry occurs relative to the unchanged direction of time intervals and therefore a violation of this symmetry does not imply that the universe is fundamentally lopsided, because this violation of symmetry can be compensated by an appropriate reversal of time intervals.

3. For an asymmetry under reversal of some physical parameter to exist all that is required is that the relevant properties be asymmetric with respect to something.
4. Only a combination of discrete symmetry operations that reverses all fundamental physical parameters and leaves absolutely nothing unchanged can be categorized as inviolable.
5. The notion that absolute directionality should not be allowed cannot be considered to restrict the violation of the P , T , or C symmetry operations, but merely the violation of the combined PTC operation.
6. A time reversal operation cannot consist merely in a reversal of the motions and rotations of objects taking place in a reverse chronological order, but must allow to establish a distinction between a physical system left unchanged by the operation and one experiencing reversed time intervals.
7. A distinction is to be made between the bidirectional concept of time direction associated with the existence of a fundamental time direction degree of freedom characterizing the propagation of elementary particles and the traditional unidirectional concept of time direction associated with changes occurring at the thermodynamic level where the notion of entropy is a meaningful property.
8. The bidirectional or time-symmetric concept of time direction is less restrictive and more distinctive than the unidirectional concept of time direction, because it recognizes the possibility for elementary particles to propagate backward in time and also allows to differentiate between identical particles effectively propagating in opposite directions of time.
9. It is the impossibility of actually observing processes from a backward in time perspective that justifies the use of a unidirectional time viewpoint relative to which the physical properties attributed to elementary particles are always those which are effectively apparent from the conventional future direction of time, even when the true direction of time in which the particles propagate is the past.
10. Any time direction-dependent physical property of a backward in time propagating elementary particle which would be positive when consid-

ered from the bidirectional time viewpoint would appear to be negative from the unidirectional time viewpoint.

11. Even if momentum is to be left unchanged by a properly defined operation of time reversal it would appear to be reversed along with the space intervals associated with the motion of particles from the unidirectional time viewpoint, because when time intervals are followed in the wrong direction space intervals are also traversed in the wrong direction.
12. The fact that from the bidirectional time viewpoint charge remains unchanged even as a particle reverses its direction of propagation in time allows this physical property to be used as a means to distinguish time-reversed processes independently from the direction of motion of particles which is necessarily observed from a forward in time perspective.
13. When the time intervals associated with the motion of a particle are reversed as a consequence of applying a T operation this change occurs relative to the unchanged direction of space intervals, so that the same positive space intervals are now traversed in the opposite direction of time.
14. A properly defined operation of reversal of the fundamental time direction parameter cannot give rise to a reversal of the thermodynamic arrow of time given that such a T operation has nothing to do with the perceived direction of motion of particles.
15. It must be required that momentum, as the physical property conjugate to space, only reverses when space is reversed, while energy, as the physical property conjugate to time, only reverses when time reverses.
16. If the sign of action is to remain unaffected by properly defined P and T symmetry operations, then momentum must necessarily reverse as a consequence of a reversal of space coordinates while energy must necessarily reverse as a consequence of a reversal of the time coordinate.
17. It is necessary to explicitly define space intervals as being reversed by a P operation even though the direction of space intervals is usually

assumed to be determined by the direction of momentum, because momentum can be reversed without space intervals being equally reversed when the sign of action is reversed and in such a context it must be recognized that momentum direction is an independent quantity whose specification is not sufficient to determine the sign of space intervals.

18. The time intervals associated with the propagation of elementary particles and the sign of energy must be reversed by T even if traditionally it is implicitly assumed that both the energy signs and the bidirectional time intervals are unchanged under time reversal despite the reversal of the time coordinate.
19. The spin of elementary particles must remain invariant under a properly defined time reversal operation described from the viewpoint of bidirectional time, even though this physical property would appear to be reversed from a unidirectional time viewpoint from the perspective of which momentum is reversed while the position of particles remains unchanged.
20. Charge must be considered to be reversed from a unidirectional time viewpoint under a properly defined time reversal operation T despite what is traditionally assumed, which means that to test the invariance of physical laws under time reversal we need to use antimatter.
21. Under an appropriately defined time reversal operation as experienced from a unidirectional time viewpoint it would be electric fields which would reverse while magnetic fields would remain unchanged and not the opposite, because magnetic fields depend on both the direction of currents and the sign of charge of the source.
22. The charge conjugation symmetry operation C must be understood to consist in a combined space and time reversal operation that leaves the sign of charge invariant from the bidirectional time viewpoint, while it appears to reverse the charge and leave the time intervals, the sign of energy, the space intervals and the momentum unchanged from the viewpoint of unidirectional time, as a consequence of the additional reversal to which those quantities are submitted when time is not followed in the right direction.

23. Despite what is traditionally assumed the direction of spin must reverse from a unidirectional time viewpoint under a properly defined charge conjugation operation given that the space coordinates are reversed while the momentum is left invariant by being reversed twice and in such a context it can no longer be assumed that the behavior of spin under application of C is a mere matter of convention.
24. Handedness must be assumed to be reversed by a properly defined C operation from both the bidirectional and the unidirectional time viewpoints because from the former viewpoint momentum is reversed and spin is invariant, while from the latter viewpoint momentum is invariant and spin is reversed, which actually explains why particles of a given handedness often seem to be naturally related to antiparticles with opposite handedness.
25. Invariance under a combined PTC operation is explicitly required in the context of the redefined P , T , and C operations which I proposed, because from both a unidirectional and a bidirectional time viewpoint all the fundamental physical parameters are reversed twice or never when the three operations are combined.
26. The classical equations for momentum and angular momentum as a function of space and time intervals and spatial positions do not need to apply from the viewpoint of bidirectional time, because they were formulated in the context of a unidirectional time perspective according to which time intervals are positive definite and it is the space intervals themselves which are reversed. Therefore, it is not possible to argue that the fact that those equations predict outcomes which differ from those provided by the redefined discrete symmetry operations when time intervals are assumed to be reversed is an indication that the new definitions of P , T , and C are inappropriate, because in this context it is rather the classical equations which are inapplicable.
27. There are four different action sign reversing symmetry operations which can be denoted M_I , M_P , M_T , and M_C and whose four different outcomes are each related to phenomenologically distinct states of negative action matter which can be transformed into one another by individually applying the three action sign preserving symmetry operations P , T , and C .

28. There are two different ways by which space- or time-related parameters can be reversed in such a way that the sign of action is reversed, because it is possible to either reverse the signs of the momenta and the energies while keeping space and time intervals unchanged, or else to reverse the space and time intervals associated with the propagation of particles while keeping the signs of the momenta and the energies invariant, but those two different ways to reverse the action can be applied differently to space- and time-related parameters.
29. A negative action particle would propagate negative energies forward in time or positive energies backward in time and would also have negative momentum in the observed direction of its propagation in space.
30. From the bidirectional viewpoint of time the sign of charge remains unaffected by a reversal of action, while spin must be assumed to be reversed under all action sign reversing operations.
31. Applying any of M_I , M_P , M_T , or M_C alone once or twice is not required to produce invariance, but the $M_I M_P M_T M_C$ operation obtained by combining of all those action sign reversing symmetry operations must necessarily produce invariance given that such an operation reverses all fundamental physical parameters twice.
32. The M_I , M_P , M_T , and M_C operations can be violated to different degrees when applied independently, because the action sign preserving reversal operations P , T , and C which relate the different states of negative energy matter to one another can be violated to different degrees by negative energy matter compared to how they are violated by positive energy matter and it is merely required that the different states of negative energy matter which are related to each other by the action sign preserving symmetry operations are invariant under a combined PTC operation. In such a context the action sign reversing symmetry operations can be conceived as together transforming merely one single additional degree of freedom.
33. There appears to be a violation of the M symmetry relating positive and negative action states in our universe given that based on the generalized gravitational field equations which I have proposed the observed small positive value of the cosmological constant appears to arise from just such a minute violation of symmetry.

34. When a condition of continuity of the direction of the flow of time (associated with the sign of physical time intervals) along a particle world-line is considered to apply, antiparticles must be described as particles which are necessarily propagating backward in time, because the annihilation of a particle with an antiparticle must be allowed to occur with the same probability for all such pairs and cannot only take place for those pairs where the two particles happen to be propagating in opposite directions of time.
35. If the condition of continuity of the flow of time applies, then no particle can turn into an antiparticle without effectively reversing its direction of propagation in time regardless of whether or not it also reverses the sign of its energy and this means that there should be as many forward in time propagating particles as backward in time propagating particles in the universe, which allows to conclude that no fundamental asymmetry under reversal of the direction of time can be related to the thermodynamic arrow of time. This conclusion is not ruled out by observations given that the most abundant form of negative action matter can consist of backward in time propagating particles. In the context where the condition of continuity must apply the compensation of the observed matter-antimatter asymmetry which is made possible by the presence of negative action matter is no longer a mere possibility and there must necessarily be an equal number of particles and antiparticles of all action signs taken together, which in fact also means that there must be as many positive action particles as there are negative action particles of any kind in our universe.
36. If we recognize the necessity for a compensation of the positive action matter-antimatter asymmetry by an opposite asymmetry involving negative action matter and antimatter then it follows that under the conditions existing in the very early universe it must definitely be possible for pairs of opposite action particles to be permanently created out of the vacuum even if this is forbidden under ordinary circumstances, but only if we require the condition of continuity of the flow of time to effectively apply even under such extreme conditions.
37. In the context where the limitations imposed by quantum indeterminacy are assumed to imply the existence of a smallest meaningful unit of area it would be impossible for two particles to be present at the

same moment in such a unit of surface, so that if the state of some elementary particles crossing a surface are to encode the information about the microscopic state of the matter contained inside that surface then it would be impossible for any physical parameter associated with such a unit of area to be attributed more than one degree of freedom at a given time. But given that the measure of entropy or ignored information is known to vary in a binary way it must be assumed that this parameter of particles which provide the microscopic degrees of freedom on a surface can itself only vary in a binary fashion.

38. The microscopic degrees of freedom on a surface must be considered to reflect the microscopic state of the matter that is located within that surface, particularly when this surface is the event horizon of a black hole, even though the degrees of freedom of the matter itself may not be of the same nature as those associated with the surface.
39. An elementary black hole with a mass equal to one Planck mass and an area that is four times that corresponding to a sphere with a radius equal to the Planck length and which we must assume to contain at most one elementary particle should carry exactly one binary unit of microscopic information which means that it is possible and even necessary to associate each unit of microscopic information encoded on an event horizon with the state of a particle it contains which can therefore only vary as a binary parameter.
40. If all distinct degrees of freedom associated with the discrete symmetry operations and only those degrees of freedom needed to be reflected in the microscopic state of a particle confined by the event horizon of an elementary black hole we would need three bits to be encoded on the event horizon of the object.
41. When we restrict our attention to stable-state black holes it must be assumed that the sign of mass of the black hole determines the sign of energy of all the matter particles it contains and therefore the degree of freedom associated with the energy sign of particles, which is transformed by the M symmetry, cannot contribute to the entropy of a macroscopic black hole.
42. It is necessary to specify the handedness of particles independently from the other degrees of freedom which are reversed by the P and

T symmetry operations and therefore this parameter can contribute independently to the entropy of a black hole.

43. The sign of charge of the particles forming a black hole is the only physical parameter that is entirely determined by its dependence on the redefined time reversal symmetry operation T as would be experienced from the unidirectional time viewpoint.
44. It must be assumed that it is merely the momentum direction of the particles forming a black hole which constitutes the degree of freedom that is transformed by the P operation, because while the handedness is reversed by P along with space directions it can also be reversed when the spin reverses and spin does reverse under application of other discrete symmetry operations from a certain perspective.
45. If negative energy particles were present inside a positive mass black hole they would be ejected from the object and in the process would acquire a maximum energy and a momentum which would be invariably directed away from the center of the object in the reference system relative to which the black hole is not rotating.
46. A positive energy particle crossing the event horizon of a positive mass black hole would gain a maximum energy and a momentum invariably directed toward the center of the object in the reference system relative to which the black hole is not rotating, while the lateral components of its momentum would become negligible and would merely contribute to the total angular momentum of the object whose motion of rotation is shared by all particles.
47. The maximum energy that is reached by particles accelerated in the gravitational field of a black hole is the Planck energy associated with the smallest physically meaningful measure of area characterizing an elementary black hole.
48. Given that the energy associated with the rest mass of a particle reaching a spacetime singularity or emerging from one with a mass opposite its own would be negligible in comparison with its kinetic energy then it must be assumed that the magnitude of momentum also constitutes a fixed variable under such circumstances.

49. When no new matter crosses the event horizon of a black hole there is no change to the state reached during the last stages of collapse into the singularity and this final state can always be considered to be that which must be reflected in the microscopic configuration of the gravitational field on the event horizon of a stable-state black hole. In this context the microscopic configuration of the degrees of freedom associated with the event horizon of a black hole may be considered to constitute a kind of magnification of the state of matter that reached the singularity and the radial position of a particle at the time it reaches that singularity would determine which elementary unit of area on the event horizon of the object encodes the information about that particle. If the argument that allows to explain the binary nature of the degrees of freedom of the matter reaching a singularity is considered valid it would therefore help motivate the assumption that all particles must necessarily reach the singularity at a unique radial position which would effectively explain why it is that the detailed three-dimensional configuration of the matter inside a surface can be described using merely the microscopic degrees of freedom present on this two-dimensional boundary.
50. When the sign of energy of all the particles that became trapped by the gravitational field of a black hole is assumed to be determined by the sign of mass of the object, it follows that the sign of the momentum of all those particles is also fixed by the mass of the black hole, so that this microscopic physical parameter cannot contribute to the entropy of such an object.
51. If the sign of energy of the matter particles forming a black hole was not considered to be a fixed parameter, then the sign of the momentum of those particles would also be a free parameter that could contribute to the microscopic information content of the object.
52. Given that the direction of the momentum of all component particles is fixed for a stable-state black hole it follows that the handedness of particles allows to determine one microscopic binary degree of freedom which varies as a function of the direction of spin and which can contribute to the entropy of the object.
53. Given that the traditional formula for black hole entropy was derived

from properties of the gravitational field and given that it would not be appropriate to assume that the gravitational field provides information about the sign of charge of elementary particles then it would also be incorrect to assume that the sign of charge of the particles that form a black hole can contribute to the measure of entropy determined by the semi-classical theory of black hole thermodynamics.

54. Information about the microscopic configuration of the electric charges inside some surface must be provided by microscopic degrees of freedom associated with the electromagnetic field on that surface, which would provide an independent contribution to the temperature of a black hole and the same is true for any other charge and its associated field.
55. The only information which must be encoded in the microscopic configuration of the degrees of freedom associated with the surface gravitational field on the event horizon of a stable-state black hole is that which allows to determine the handedness of every particle contained in the object using one single binary unit of information for every elementary particle, a conclusion which complies with the fact that an elementary black hole with a surface of four Planck areas carries one binary unit of missing information or entropy. This result confirms that the only relevant physical parameters for matter that becomes trapped by the gravitational field of a black hole are those which are transformed by the redefined discrete symmetry operations.
56. Despite the fact that for a general surface, in the absence of event horizon, the direction of momentum as well as the magnitude of energy of the particles inside the surface would be allowed to vary more freely, the limit to entropy imposed by the Bekenstein bound would still apply even if more information would be required to describe the microscopic state of matter contained inside the surface, because the lower density of matter involved would mean that less information would be required to describe the microscopic configuration of the gravitational field.
57. When the entropy associated with the gravitational field attributable to a positive mass black hole is not maximum as a consequence of the fact that some negative energy matter is present within the event horizon of the object along with positive energy matter, the situation is unstable and will continue to evolve until a stable-state black hole

forms. This is unavoidable in the context where the presence of negative energy matter within a positive mass black hole actually contributes to decrease the entropy of the object and it suggests that stable-state black holes play a role in gravitational physics which is analogous to that which is played by thermal equilibrium states in statistical mechanics.

58. Despite appearances, even in the more general case where a macroscopic event horizon is absent the energies and the momenta of elementary particles are still restricted to vary as binary parameters locally, because in fact event horizons are always present on the shortest distances as a consequence of quantum fluctuations in energy which give rise to ephemeral microscopic black holes which constrain the motion of particles present on such a scale.
59. Despite quantum indefiniteness there would still be a constraint on the momentum direction of a particle to be fixed as a binary parameter by the direction of the gravitational field in the presence of a microscopic black hole, because it is not possible to specify the direction of momentum any more precisely than there are elementary surface elements associated with the microscopic black hole and there is only a finite number of surface elements.
60. Each elementary unit of area on the event horizon of a local microscopic black hole contains the same amount of information as would an elementary unit of area associated with a macroscopic black hole in which particles with both energy signs could be present.
61. In the context where there must be a unique correspondence between any matter particle inside some surface and a precise elementary unit of area on that surface it would be possible to associate the binary information encoded on the event horizon of any microscopic black hole located inside a larger surface with a finite number of elements on that surface, so that the ensemble of such elements on the macroscopic surface would provide a binary measure of the information concerning all the matter particles contained inside the surface.
62. The amount of information required to describe the state of matter constrained by the presence of a local microscopic black hole (arising as a consequence of quantum fluctuations) is larger than that required

to describe the state of matter inside a macroscopic black hole, because the sign of energy of the microscopic black hole is not a fixed parameter and given that it does affect the direction of the momentum of the particles submitted to its gravitational field it follows that we must independently determine the energy sign and the momentum sign of each particle in addition to specifying its handedness.

63. Given that it is possible to assume that the description of macroscopic stable-state black holes can be generalized by assuming that locally the states of matter particles are constrained by the presence of microscopic black holes which are also the gravitational equivalent of states of fluctuating thermodynamic equilibrium, it follows that the analogy between the physics of black holes and the classical theory of equilibrium thermodynamics can be carried over to the field of non-equilibrium thermodynamics.
64. In the context where a negative mass black hole must be assumed to have a surface gravitational field opposite that of a similar positive mass black hole, the fact that the temperature of a black hole is proportional to its surface gravitational field implies that a negative mass black hole would have a negative temperature, which is appropriate given that such a black hole would radiate negative energy particles.
65. The fact that a negative temperature can be attributed to a negative mass black hole strengthens the case for an exact correspondence between black hole thermodynamics and the classical theory of thermodynamics according to which negative temperatures are unavoidable when a limited number of energy levels are available for a system as its temperature rises.

Chapter 4

Conclusion

4.1 Main results

I have come a long way since first asking what would happen to a negative mass object dropped in the gravitational field of the Earth. Yet I was able to confirm that my early intuition was right and that consistency dictates that the negative mass would need to ‘fall’ up despite the fact that this goes against current expectations. This is a conclusion for which I have provided ample justification and even if that was all I had been able to establish I would already be very satisfied with the outcome of my undertaking. But several other developments were introduced in this report which are all related to the issue of time directionality as a concept independent from the thermodynamic arrow of time. In fact, the hypothesis of the existence of a fundamental time-direction degree of freedom has become the vital lead that allowed me to better understand many aspects of gravitational phenomena at the semi-classical level of description. Yet despite the fact that the objective of this report was mainly to provide a consistent account of the way by which the concept of negative energy that emerges from those considerations can be integrated into a classical theory of gravitation, I have also made use of those theoretical developments to provide solutions to various specific problems in fundamental theoretical physics.

Using a revised formulation of the gravitational field equations I was able to show, in particular, that the cosmological constant, conceived as a manifestation of the existence of a residual value of vacuum energy density averaged over a very large scale, can be expected to be as small as the

symmetry between the perceived metric properties of space associated with positive energy matter and those associated with negative energy matter as perceived by positive energy observers is itself perfect. In this context it is natural to expect some level of cancellation between all contributions to the resulting value of the cosmological constant above any which may occur due to already known contributions. In fact, given the high level of invariance that is usually observed to apply under most discrete symmetry operations, it seems that the natural value we should expect to measure for the cosmological constant would now effectively be zero rather than the very large quantity which is provided by most current estimates. The fact that the actual value is not perfectly null would then simply mean that the symmetry involved is not absolutely perfect even though it does contribute to reduce the measure of vacuum energy density to a value that is negligible in comparison to what it would otherwise be.

Such a slightly imperfect symmetry is already known to be involved in giving rise to the very existence of matter. Indeed, it is usually recognized that a violation of the symmetry under reversal of the sign of charge (combined with a parity operation in the traditional approach) is likely responsible for the slight imbalance in the relative number of matter particles over anti-matter particles which existed during the big bang and which allowed some of the matter to survive the early period of annihilation. One of the most important results of my analysis of the issue of discrete symmetry operations is that it becomes possible to regain some of the symmetry that is lost as a consequence of this imbalance between matter and antimatter by recognizing that the requirement of continuity of particle world-lines imposes a similar, but compensating imbalance between what we may describe as negative action matter and antimatter particles. This outcome is highly desirable given that it allows to avoid the conclusion that there would exist an absolutely characterized lopsidedness of the universe with respect to the direction of time in the context where the redefined time reversal symmetry operation T effectively involves a transformation of matter particles into antimatter particles.

As I noted when I introduced the alternative interpretation of negative energy matter as consisting of voids in the positive energy portion of the vacuum, the very existence of states of negative *vacuum* energy, like those already predicted by conventional quantum field theory, actually constitutes an additional motive for recognizing the validity of the relational description of negative energy matter. This is because it is only from such a viewpoint

that those negative vacuum energy states can be allowed to be reached without giving rise to catastrophic vacuum decay and the creation of states of ever more negative energy. This is an additional benefit of the approach favored here in the context where the existence of those negative vacuum energy states must be recognized as unavoidable, even from the viewpoint of a traditional interpretation. In this particular sense the prediction of an absence of vacuum decay can be considered as one of the most significant result of the alternative framework for classical gravitation theory which was developed in this report.

What constitutes perhaps the most remarkable outcome surrounding the revised formulation of the discrete symmetry operations, on the other hand, is the derivation of an exact binary measure of entropy for the matter contained within the event horizon of an elementary black hole. This result is particularly noteworthy in that it effectively matches the constraints set by the semi-classical theory of black hole thermodynamics. The possibility that is allowed in the context of the proposed interpretation of negative energy states to generalize the analogy between classical thermodynamics and black hole theory through an application of the classical concept of negative temperature can then be considered to merely confirm the relevance of the concept of negative energy matter for gravitation theory. Those unexpected benefits come in addition to the solutions which were offered in the first chapter of this report to the more traditional problems usually associated with the concept of negative energy matter and which allow to demonstrate the viability of a theory based on such an alternative interpretation of negative energy states. But in my opinion the most interesting outcome of this whole approach will remain the elaboration of a quantitative framework which actually generalizes relativity theory in a way that increases its simplicity rather than adding in complexity over the already elegant gravitational field equations.

4.2 Historical perspective

The significance of the developments introduced in this report can be better appreciated by describing the progress achieved from a historical perspective. If we start with general relativity I think that what the theory allowed us to understand is that all motion, including acceleration, is relative and that the state of motion of an object must be defined in relation to the state of

motion of the rest of the matter in the universe. Thus relativity theory embodied in its structure the requirement of a relational definition of physical properties. But it also failed to integrate the requirement of the relativity of the sign of energy. The common belief which existed since the creation of the general theory of relativity is that energy must be considered positive definite, because otherwise apparently insurmountable problems would arise. Now, what quantum field theory allowed us to understand is that negative energies are unavoidable for properly estimating the probability of all possible transitions involving particles and antiparticles. But the current interpretation of this theory also failed to accommodate the fact that no constraint exists that would justify assuming that those negative energies are only relevant for computational purposes and do not show up as properties of real matter particles distinct from ordinary particles and antiparticles when gravitation comes into play. What I have tried to achieve with this report is to generalize relativity theory to produce a fully relativistic theory compatible with the requirement that the sign of energy should also be a relative property. What motivated those developments was a better understanding of the relationship between the sign of energy of a particle and its direction of propagation in time which again arose from applying the requirement of relational description of physical properties. In such a context it effectively appeared that a concept of negative energy distinct from that which is usually assumed to be relevant in quantum field theory was not only allowed, but was required by a truly consistent classical theory of gravitation.

Once it had been shown that the difficulties usually associated with negative energy matter can be solved without rejecting the physical relevance of the whole concept of negative energy, there appeared to no longer be any rational motive for rejecting the possibility that negative energies can propagate forward in time and give rise to gravitational phenomena distinct from those involving exclusively positive energy matter. It thus became clearly inappropriate to attribute a preferred status to positive energy matter and this in turn meant that we are no longer justified to assume (as some authors did) that even as it becomes integrated with general relativity the concept of gravitationally repulsive matter cannot involve negative energy, but must merely give rise to the notion of a variable metric devoid of any theoretical justification. Indeed, it has been clearly emphasized in this report that it is only when the concept of negative energy is well integrated to classical gravitation theory, by considering the equivalence between the presence of negative energy matter and an absence of positive energy from the vac-

uum that a consistent theory (for which all measures of energy are relative) emerges which agrees with all experimental and observational constraints. The original approach which was developed in the preceding chapters is thus unique in that it effectively allows to account for the very existence of the phenomenon of inertia, despite the fact that both positive and negative mass matter must be present on the largest scale. It alone also enables the success of the standard model of cosmology at predicting the rate of expansion of positive energy matter to be reproduced in a bi-metric theory. Furthermore, it allows to explain the stability of negative energy matter against collapse to ‘lower’ (more negative) energy states. Quite significantly it also predicts an arbitrarily small value for the average energy density of vacuum fluctuations, which is no longer in severe conflict with the upper limit on the value of the cosmological constant derived from astronomical observations. This approach complies as well with the requirement that there can be no absolutely characterized direction of time in our universe, such as would arise from the observed matter-antimatter asymmetry. Finally it provides a description of the microscopic state of gravitationally collapsed matter which is compatible with the quantitative constraints set by the semi-classical theory of black hole thermodynamics.

It must be clear that the concept of negative energy already existed before the developments I proposed in order to make it a consistent notion were introduced. But negative energy was always defined in an absolute or non-relational manner which I have shown leads to serious inconsistencies, in particular because it would give rise to violations of the principle of inertia. Indeed, the idea that energy could be negative in an absolutely defined way and should therefore gravitationally repel all matter regardless of its energy sign, as if this repulsion was a distinctive property of negative energy matter itself, was here shown to give rise to undesirable effects, even aside from the plain logical inconsistency it would involve. The alternative interpretation of negative energy states which I proposed has allowed to avoid those problems while also making unnecessary the hypothesis that only positive energy matter can exist, because it explains why matter in such a state is unobservable from the viewpoint of observers made of positive energy matter. It has also become possible to predict that negative energy matter cannot be found in large concentrations in regions of the universe occupied by positive energy stars and galaxies as a consequence of the mutual gravitational repulsion which must exist between particles of opposite energy signs and because negative energy bodies gravitationally attract each other. Thus it

was actually explained why negative energy matter has remained mostly out of reach of astronomical observations, so that this property no longer needs to be merely postulated.

It is the fact that the notion of time-directionality remained so poorly understood, even after the progress which was achieved in this area by quantum field theory, that explains that the whole problem of appropriately defining the discrete symmetry operations was itself in such a mess from the viewpoint of clarity and consistency when I began studying the subject. It is indeed the stubbornness to consider time from a traditional viewpoint where only one direction is allowed for this degree of freedom that explains that the time reversal operation was never appropriately described. This was allowed to occur despite the clues arising from the discovery of antimatter and the successful description of antiparticles as particles propagating backward in time. The common sense feeling inherited from our experience of thermodynamic time is so strong that it is still commonly believed that antiparticles are merely identical particles which happen to have opposite charges rather than being the same particle propagating backward in time, as seems to be required from a mathematical viewpoint. This is what explains that time reversal was never considered to involve a reversal of charge as I suggested must actually be required. But once this was recognized the possibility opened up to explain other facts. It is in effect by using this insight that I was able to propose an explanation for the binary nature of the degrees of freedom associated with the state of matter constrained by the gravitational field of an elementary black hole. In such a context it can no longer be argued that the notion of backward in time propagation is merely an expedient for facilitating the calculations of probability amplitudes. Our notion of time direction has been irretrievably altered and there is no going back.

Acknowledgements

I would like to thank the government of Québec and its taxpayers which through the generous social programs they offer have allowed me to benefit from a steady source of income during the years in which I was working on the present project without any support from academia or the industry. They have not only allowed me to benefit from the conditions necessary to achieve the depth of knowledge required to conduct this research, but they have also saved my life at times when studying physics was the only activity that still gave enough meaning to my existence that it actually felt endurable.

Bibliography

- [1] L. H. Ford and T. A. Roman. Averaged energy conditions and quantum inequalities. *Phys. Rev. D*, 51:4277, 1995.
- [2] L. H. Ford and T. A. Roman. Restrictions on negative energy density in flat spacetime. *Phys. Rev. D*, 55:2082, 1997.
- [3] M. J. Pfenning and L. H. Ford. Quantum inequalities on the energy density in static robertson-walker spacetimes. *Phys. Rev. D*, 55:4813, 1997.
- [4] M. J. Pfenning and L. H. Ford. Scalar field quantum inequalities in static spacetimes. *Phys. Rev. D*, 57:3489, 1998.
- [5] D. R. Brill and S. Deser. Variational methods and positive energy in general relativity. *Ann. Phys. (New York)*, 50:548, 1968.
- [6] D. R. Brill and S. Deser. Positive definiteness of gravitational field energy. *Phys. Rev. Letters*, 20:75, 1968.
- [7] S. Deser. Timelike character of gravitational field energy-momentum. *Nuovo Cimento*, 55B:593, 1968.
- [8] D. Brill, S. Deser, and L. Faddeev. Sign of gravitational energy. *Phys. Lett.*, 26A:538, 1968.
- [9] R. M. Schoen and S.-T. Yau. On the proof of the positive mass conjecture in general relativity. *Comm. Math. Phys. (Germany)*, 65:45, 1979.
- [10] R. M. Schoen and S.-T. Yau. Proof of the positive-action conjecture in quantum relativity. *Phys. Rev. Lett.*, 42:547, 1979.

- [11] R. M. Schoen and S.-T. Yau. Positivity of the total mass of a general space-time. *Phys. Rev. Lett.*, 43:1457, 1979.
- [12] R. M. Schoen and S.-T. Yau. The energy and the linear momentum of space-times in general relativity. *Comm. Math. Phys.*, 79:47, 1981.
- [13] R. M. Schoen and S.-T. Yau. Proof of the positive mass theorem. ii. *Comm. Math. Phys.*, 79:231, 1981.
- [14] M. M. Nieto and T. Goldman. The arguments against ‘antigravity’ and the gravitational acceleration of antimatter. *Physics Reports*, 205:5, 1991.
- [15] H. Reichenbach. *The Philosophy of Space & Time*, page 256. Dover Publications, 1957.
- [16] G. D. Birkhoff. *Relativity and Modern Physics*, page 253. Harvard University Press, 1923.
- [17] S. Hossenfelder. A bi-metric theory with exchange symmetry. *Phys. Rev. D*, 78:044015, 2008.

Index

- axioms
 - negative energy, 141–147, 164, 165, 167, 168
 - negative energy matter, 141–147, 164, 165, 167, 168
 - negative mass, 141–147, 164
- Bekenstein bound, 236, 237, 258, 260, 262, 265, 281
 - surface information and entropy, *see* black hole
- black hole, 137, 138, 236
 - angular momentum, 236, 244, 247, 250, 279
 - charge, 236, 244, 247, 257, 258
 - conservation of information, 237
 - electromagnetic field, 256–258, 281
 - entropy, 138, 140n, 180, 236, 237, 239, 241–244, 246, 249–251, 253–259, 262, 265, 267, 277, 278, 280, 281, 286
 - event horizon, 138, 177, 236, 245, 249–255, 260–265, 267, 279, 281, 282
 - event horizon degrees of freedom, 236–243, 245, 246, 253, 254, 256–259, 261, 277–279
 - gravitational collapse, 138, 245, 246, 249, 251, 253–255, 257, 259, 279
 - Hawking radiation, 258, 269, 270, 283
 - information, 236–244, 247, 249, 250, 253, 255–257, 259–262, 264–266, 277–282
 - macroscopic parameters, 236, 237, 247, 261
 - mass, 236, 237, 241, 242, 244–247, 253, 254, 262, 270, 278, 280, 283
 - mass reduction, 138, 141, 177, 262
 - matter absorption, 138, 141
 - matter degrees of freedom, 180, 236, 237, 240–247, 249, 250, 253–261, 265, 266, 269, 277–282, 286, 289
 - merger, 267
 - momentum, 244, 247
 - negative energy matter, 177, 245, 246, 249, 252, 253, 258, 260, 262, 265, 266, 269, 270, 279, 281–283
 - particle charges, 255–259, 280, 281
 - particle energies, 253, 254, 261–265, 279–282
 - particle handedness, 255, 259, 265, 280–282
 - particle momenta, 249–255, 260, 262–265, 279–282
 - redshift, 251

Schwarzschild radius, 241
 semi-classical theory, *see* thermodynamics, black hole
 spacetime singularity, 137, 138, 177, 236, 245, 249–254, 258, 259, 262, 279
 stable state, 245, 253, 254, 260, 262, 264, 266, 267, 278–281, 283
 state magnification, 253, 265, 279, 282
 surface area, 138, 177, 238, 240, 241, 250, 253, 257, 259, 262, 281
 surface gravitational field, 140n, 256–259, 263, 269, 281, 283
 temperature, 140n, 253, 256, 258, 269, 270, 281, 283
 thermodynamics, 140n, 236, 237, 241, 244–246, 254–257, 259, 267, 269, 270, 280, 283, 286, 287
 time dilation, 251
 Boltzmann’s constant, 241
 constraint of relational definition, 12
 absolute direction, 182, 184, 205, 272
 absolute lopsidedness, 182, 183, 186, 228, 231, 233, 271, 285, 287
 absolute space, 65, 69
 acceleration, 62, 286
 center of mass of the universe, 84–86, 172
 completeness, 183
 directional asymmetry, 182–185, 215, 271
 discrete symmetry operation, 182–186, 271, 273
 fundamental lopsidedness, *see* absolute lopsidedness, constraint of relational definition
 gravitational force, 33, 45–47, 49, 51, 56, 103, 104, 107, 147, 148, 164, 170, 171
 imbalance, 184
 interaction field energy, 106
 lower energies, 121–124, 176, 285
 polar asymmetry, 182, 183, 271
 principle of relativity, 65, 68, 76, 78, 171, 184
 relativity of acceleration, 65, 67, 69–71, 74–76, 78
 reversal of energy, 182, 271
 reversal of momentum, 182, 196, 224, 271
 self-determination, 183
 sign of charge, 182, 196, 206, 208, 271
 sign of energy, 91–95, 103, 141, 146, 147, 162, 164, 174, 196, 224, 226, 271, 286, 287
 sign of mass, 44–47, 49, 51, 55, 56, 59, 67, 78, 141, 146, 147
 space and time directions, 185, 196, 202, 205, 224, 227, 231, 233, 235, 271, 285
 space and time reversals, 182–185, 196, 206, 271
 time direction-dependent property, 40
 universe, 12, 44, 45, 71, 121, 182, 183, 185, 226
 coordinative definition, 30, 47, 169
 Coriolis force, 73

cosmological principle, 68, 84

discrete symmetry

- matter-antimatter asymmetry, 231, 233–235, 277, 285, 287
- violation, 180, 182–184, 186, 204, 205, 207, 214, 224–228, 232, 271, 272, 276, 285
- weak interaction, 182, 211

discrete symmetry operations

- action sign degree of freedom, 219, 222, 224, 225, 243, 248, 276
- alternative formulation, 180, 181, 195, 196, 199, 201, 206, 207, 213, 236, 247, 258, 267, 286
- angular momentum, 181, 182, 192, 196, 199, 202–206, 209, 211–213, 217, 222, 247, 255, 274–276, 279, 280
- antimatter, 180, 196, 204, 205, 228, 273–275
- backward motion, 181
- basic action reversal operation M_I , 218, 223, 224, 275
- charge conjugation C , 180–182, 184, 195, 196, 201, 205, 207–215, 217–220, 223–225, 227, 233, 243, 257, 258, 272, 274–276, 285
- classical equations, 216, 217, 275
- combined operations, 184, 205, 212–215, 223, 224, 272, 275, 276
- conjugate properties, 195, 210, 216, 253, 273
- CPT theorem, 184, 213
- currents, 206, 274
- dependencies, 186, 207
- electric field, 206, 274
- enantiomorphic equivalent, 209
- equivalent operations, 213, 214
- fermion quantum phase, 214
- fermion wavefunction, *see* fermion quantum phase, discrete symmetry operations
- fundamental degrees of freedom, 243, 244, 246–249, 254, 255, 259, 263, 269, 277, 278, 281, 282, 289
- gravitation, 180, 226
- handedness, 182, 196, 199, 202, 211, 247–249, 255, 257, 259, 275, 278–282
- identity operation I , 219, 223, 224
- invariance of the sign of action, 195, 199–201, 204, 207, 209, 211, 212, 218, 220, 222, 248, 273
- invariance of the sign of charge, 118, 274, 276
- joint variation, 195
- kinematic representation, 181, 185, 187, 193, 204, 208, 272
- magnetic field, 206, 274
- momentum, 181, 182, 188, 190, 192, 195, 196, 199–204, 207, 209–211, 216–220, 222, 226, 247–249, 253–255, 260, 264, 265, 273–275, 279, 280, 282
- parity P , *see* space reversal P , discrete symmetry operations
- PTC transformation, 184, 204, 212–214, 224, 225, 233, 272, 275, 276
- quantum field theory, 180, 184, 196
- quantum operators, 196

- reversal of action M , 180, 195, 200, 218–220, 222–226, 232, 233, 243, 245, 246, 248, 273, 275, 276, 278, 279, 284
- reversal of motion, 187, 193, 201, 203, 204, 208, 222, 226, 272, 273
- semi-classical viewpoint, 180
- sign of charge, 180, 182, 186, 188, 190, 196, 199, 204–213, 222, 232–234, 247–249, 255, 257–259, 273, 274, 276, 279, 280, 289
- sign of energy, 180, 188, 195, 196, 199, 201, 202, 204, 207, 209, 218–220, 222, 226, 245, 246, 248, 253, 254, 260, 265, 273–275, 277, 278, 280, 282
- space and time coordinates, 189, 192, 195, 196, 199–205, 207, 209, 211, 213, 218, 274, 275
- space intervals, 181, 182, 184, 185, 188–190, 192, 195, 196, 199–203, 207, 209–211, 216–220, 226, 243, 273–275
- space reversal P , 180, 182, 184, 186, 189, 190, 192, 195, 196, 199–202, 205, 207, 209, 211, 212, 214, 215, 218–220, 223–225, 227, 233, 243, 247, 248, 271–273, 275, 276, 278, 279, 285
- space rotation, 214
- space-related properties, 195, 196, 202, 210, 212, 214–216, 219, 222, 223, 243, 275
- spacetime reversal, 180, 209, 210, 213, 274
- spin, *see* angular momentum, discrete symmetry operations
- time intervals, 182, 184–186, 188–190, 195, 196, 199–203, 207, 209–211, 216–220, 226, 228, 243, 248, 249, 255, 271–276
- time reversal T , 180, 181, 185, 186, 188–190, 192, 193, 195, 196, 201–208, 210–220, 222–228, 232–235, 243, 247, 248, 257, 258, 272–279, 285, 289
- time-related properties, 195, 196, 202, 210, 212, 214–216, 219, 222, 223, 243, 275
- traditional conception, 180, 181, 184, 186, 192, 195, 196, 199, 201, 203, 204, 211–213, 258, 274, 289
- dynamic equilibrium of forces, 52, 53, 56–58
- electrostatic field pair creation, 258
- energy out of nothing problem, 101, 102, 111–119, 124, 134, 135
 - work and useful energy, 136, 177
- equations of state, 94
- equivalent gravitational field, 51–64, 67, 68, 70, 104, 130, 170, 171
- fermion, 213, 230
- Feynman, Richard, 105
- general relativistic theory, 284, 286, 287
 - additional variables, 155
 - alternative notation, 155n
 - alternative proposals, 152, 154
 - assumptions, 149

average stress-energy tensors, 157, 166, 167, 178
 bi-metric theories, 149, 151, 155, 157, 159, 164, 166, 167, 178, 287
 conjugate metrics, 149
 consequences, 149
 conservation of energy, 126, 166, 177
 curvature tensors, 150–152, 165, 261
 distinctive features, 148, 149
 earlier interpretations, 149, 150
 earlier publications, 148, 149
 Einstein tensor, 155, 161
 energy sign convention, 148
 general covariance, 152
 generalized gravitational field equations, 141, 147–169, 178, 276, 284, 286
 geodesics, 150
 gravitational field energy, 126, 127, 166, 177, 252
 irregular stress-energy tensors, 157–159, 162, 166, 167, 178
 justifications, 149
 maps, 155
 mathematical requirement, 155
 mathematical structure, 147–150, 152, 154, 155, 157, 159, 161, 162, 164
 metric conversion factors, 155, 157, 158, 161, 162, 164, 178, 284
 natural vacuum stress-energy tensors, 161, 162, 164, 168, 178, 284
 natural viewpoints, 148
 Newtonian limit or approximation, 167
 observer dependent energy sign, 148, 150–153, 166
 observer dependent gravitational field, 59–61, 79, 147, 148, 150–152, 164, 168, 171, 178
 observer dependent metric properties, 148, 155, 158, 161, 168, 171, 178
 physical requirements, 157, 161
 pull-overs, 155
 redefined energy ground state, 153, 154, 157–159, 162, 166, 167
 stress-energy tensors, 149, 150, 152, 153, 155, 157–159, 161, 162, 165, 166, 178
 vacuum energy terms, 157, 159, 162, 164
 vacuum stress-energy tensors, 161, 162, 168
 variational principle, 159
 global inertial frame of reference, 74–76
 grand unification theories, 230
 gravitational repulsion, 43, 105, 137–139, 141, 152, 166, 167, 173, 226, 245, 249, 284, 287, 288
 antigravity, 104, 134, 136, 177
 antimatter experiment, 134, 136, 177
 from missing positive vacuum energy, 89, 143, 165, 174, 175
 from voids in a matter distribution, 80–87, 143, 165, 172, 173
 uncompensated, 99
 uncompensated gravitational attraction, 83, 87, 89, 96, 108,

109, 126, 141, 143, 144, 165, 172, 173, 175
 weakness, 114, 115, 121, 129, 175
 inertial gravitational force, 51–58, 61, 62, 68, 70, 72, 73, 104, 132, 167, 171
 from identical matter distributions, 68–71, 74, 76, 98, 167, 171
 interaction boson, 119, 121, 142, 230
 interaction vertex
 mixed action signs, 120, 121, 176
 kinetic energy, 252, 279
 local causality, 231
 local inertial frame of reference, 52, 67, 69, 71–73, 75, 76, 79, 84, 87
 Mach’s principle, 84
 Mach, Ernst, 62
 material nature of gravitation, 72, 73, 76
 electromagnetic field analogy, 73
 matter creation, 102, 119, 121
 big bang, 117, 176, 231, 277
 conservation of energy, 115, 116
 cosmic expansion, 117, 176, 234
 favorable conditions, 111, 115, 117, 175, 234, 235, 277
 observational evidence, 102, 111
 permanence, 115, 117, 176, 277
 quantum gravitational scale, 117, 176
 meson, 250
 multiverse, 183
 negative action matter, *see* negative energy matter
 negative energy
 antiparticles, 18, 20, 23, 24, 28, 29, 37, 38, 40, 91, 93, 101, 104, 119, 120, 134, 146, 169, 233, 277, 286
 bound systems, 36, 104, 106, 131–133, 146, 168, 177
 Dirac’s solution, 18, 19, 37, 90–93
 energy conditions, 34, 35
 filled energy continuum, 91–94, 174
 in general relativity, 7, 14, 64
 in quantum field theory, 7, 15, 18, 39, 40, 101, 109, 111, 119, 124, 134, 176, 286, *see* negative densities, vacuum energy
 interaction constraint, 16, 109, 111, 176
 kinetic energy, 129, 139, 140, 177
 motivations, 7, 107, 270
 myths, 141
 negative action, 26, 27, 29, 31, 41, 47, 101, 112, 125, 146, 169, 170, 222, 233, 276, 284, 286, 287
 of attractive force field, 36, 104, 106, 131, 142, 146, 177, 252
 pair annihilation, 101, 102, 111, 114, 115, 117–120, 134, 135, 175, 229, 234, 235
 pair creation, 101, 102, 111, 112, 115, 117, 118, 121, 175, 176, 228, 234, 235, 277
 positive energy theorems, 36
 propagation constraint, 15, 20, 228
 steady state cosmology, 20
 the problem of, 6, 11, 14–21, 34–41, 161, 180

- traditional interpretation, 7, 32, 102, 104, 125, 133, 138, 141, 149, 288
- transition constraint, 19, 37, 111, 120, 175, 176, 228, 234, 235
- versus negative charge, 16
- negative energy matter
 - absence of interactions with, 106–108, 110, 114–116, 119–121, 126, 129, 134, 142–144, 152, 165, 169, 175, 176, 288
 - antimatter, 220, 231–233, 277, 285
 - antimatter experiment with, 134–136, 177
 - colliding opposite energy bodies, 125–129, 166
 - concentrations, 111, 288
 - conservation of energy, 101–103, 116, 119, 121, 125–129, 134, 136, 140, 177
 - conservation of momentum, 103, 112, 125, 129
 - cosmological models, 162, 167
 - dark matter, 105, 107–109, 142, 175
 - discrete symmetries, 220, 222, 225, 244, 267, 275, 276
 - dominant paradigm, 105, 288
 - energy of force fields, 106, 142, 175
 - gravitational potential energy, 127–129, 136, 177
 - heat, 139, 177, 270
 - homogeneous distribution, 96–98, 144, 154, 166, 174
 - implicit assumptions, 105
 - inhomogeneities, 144, 171
 - initial distribution, 122
 - momentum direction, 112, 125, 276
 - observational evidence, 15–17, 47, 101, 105, 109, 111, 169, 288
 - outstanding problems, 101–105, 133, 137–139, 141, 149, 286–288
 - overdensity, 80, 144, 162, 166
 - potential energy, 136, 177
 - radiation, 139, 141
 - rarity, 111
 - ratio of cosmic densities, 167
 - requirement of exchange symmetry, 31, 47, 49, 123, 141, 144, 149, 157, 164, 170, 174, 178, 179, 226, 285, 288
 - sign of charge, 118, 174
 - temperature, 139–141
 - thermal energy, 139, 140
 - transformation into, 113
 - underdensity, 80, 82, 144, 154, 158, 162, 166
 - uniform distribution, *see* homogeneous distribution, negative energy matter
 - universal expansion, 162, 167, 174
 - voids in positive vacuum energy, 88–90, 95–97, 99, 108, 109, 114, 126, 143, 144, 154, 157, 160, 161, 165, 173–175, 178, 285, 287
- negative mass, 41–80, 171, 284
 - absolute gravitational force, 44, 170, 288
 - absolute inertial mass, 130, 131, 133, 146, 177
 - acceleration, 53, 56–58, 61, 62, 74, 104, 130, 131, 146, 170, 171
 - generalized Newton’s second law, 53–58, 62, 67, 104, 130, 171

- gravitational mass, 42, 43, 46, 49–51, 104, 130, 131, 133, 142, 170
- motivations, 6
- negative inertial mass, 32, 33, 42–46, 48, 51, 53, 56, 58, 62, 63, 66, 104, 130, 133, 142, 145, 164, 170
- Newton’s second law, 53
- Newton’s third law, 43
- Newtonian gravitation, 167
- positive inertial mass, 49–51, 58, 59, 170
- principle of inertia, 42, 43, 49, 50, 66–68, 288
- traditional concept, 32, 33, 42–54, 146, 284, 288
- negative temperatures, 140n, 268, 269
 - black hole, 269, 270, 283, 286
 - decrease of entropy, 268, 269
 - energy levels, 268–270, 283
 - infinite temperature, 268–270
 - negative energy, 270, 283
 - spin system in magnetic field, 268–270
- neutron’s electric dipole moment, 206
 - direction of dipole, 206
 - precession movement, 206
- nucleus, 258
- open questions, 11
- particle accelerator, 249
- particle beam, 250
- perpetual motion problem, 104, 134, 136, 177
- Petit, Jean-Pierre, 149n
- physical nature of geometry, 73, 171
- polarization, 213
- principle of equivalence, 33, 64–80, 170, 171
 - Einstein’s elevator experiment, 54, 65, 67, 68, 70, 72, 78
 - entangled system, 132, 133, 177
 - equivalent source, 54–58
 - relativized, 63, 78, 79, 145, 168, 171, 172
 - violation of the, 49–51, 58, 63, 65–67, 71, 76, 80, 104, 130–133, 170
- propagator, 196
- quantum chromodynamics, 250
- quantum gravitation, 160, 176, 236, 238, 240, 267
 - discrete space, 238–240
 - elementary black hole, 241–243, 245, 250, 255, 256, 258, 263, 264, 278, 279
 - elementary unit of area, *see* elementary unit of surface, quantum gravitation
 - elementary unit of surface, 238–241, 255, 259, 264–266, 277, 279, 281, 282
 - energy fluctuations, 266, 282
 - fluctuating gravitational field, 261, 263, 266, 282
 - gravitons, 171, 238, 242, 261
 - microscopic black hole, 263–267, 282, 283
 - momentum direction, 264, 265
 - negative energy, 266
 - Planck area, 241, 259, 264, 278
 - Planck energy, 114, 160, 162, 164, 178, 242, 249–252, 263, 279

- Planck length, 241, 250, 252, 278
- Planck mass, 241, 258, 263, 270, 278
- Planck momentum, 252
- Planck scale, 178, 238–240, 247, 248, 250, 259, 263, 265, 266, 282
- Planck time, 238, 250
- spin networks, 259
- quark, 230, 250
- relativistic frame dragging, 68, 75, 76, 250
- repulsive force field
 - energy sign of, 106
- rest mass, 252, 279
- Rutherford atom model, 121
- second law of thermodynamics
 - coarse-graining, 261
 - constraint on matter creation, 115, 116, 121, 175, 234
 - constraint on process description, 187
 - degrees of freedom, 138, 176
 - energy, 269, 270
 - entropy, 115, 121–123, 139–141, 176, 177, 186, 187, 190, 193, 216, 233, 268, 269, 272
 - fluctuations, 115, 117
 - gravitational entropy, 242, 260–262, 265, 266, 281
 - matter disintegration, 121, 123, 176
 - microstates, 115, 123
 - smoothness of matter distribution, 122, 260, 265
 - spreading of energy, 115, 176
 - stable state, 116
 - static equilibrium, 266
 - temperature, 268, 269
 - thermal equilibrium, 121, 261, 262, 266, 267, 281, 283
 - violation, 138–141, 177, 193
- special relativity, 184
- state vector, 196
- statistical mechanics, 262, 266
 - equilibrium thermodynamics, 266, 270, 281, 283, 286
 - near-equilibrium thermodynamics, 266, 267
 - non-equilibrium thermodynamics, 266, 267, 283
- strong nuclear interaction, 131
- ‘t Hooft, Gerard, 241
- thermodynamic time asymmetry, *see* time irreversibility
- tidal effect, 250
- time direction degree of freedom, 21–34, 169, 187, 190, 192, 193, 228, 235, 273, 284, 289
 - antiparticles, 188, 209, 211, 222, 229–231, 272, 276, 289
 - bidirectional time, 186–190, 192, 193, 195, 196, 199, 201–204, 207, 210–212, 216, 217, 222, 230, 234, 248, 249, 272–276
 - chronological order, 185, 190, 272
 - condition of continuity in time, 31, 112, 113, 118, 169, 228–231, 233–235, 276, 277, 285
 - direction of propagation, 16, 21–25, 28, 30, 120, 146, 169, 180, 181, 185, 187–190, 192, 193, 196, 199, 201, 202, 204, 205,

207–209, 216, 217, 220, 222,
 230–235, 248, 255, 272–277,
 286, 289
 direction of the flow of time, 118,
 229–231, 276, 277
 Feynman’s interpretation, 23, 37,
 289
 grand unified theories, 26
 pair creation and annihilation, 112,
 229–231, 234, 276, 277
 particle world-line, 112, 228–231,
 233–235, 276
 relativity of the sign of charges,
 21–23, 25, 26, 30, 31
 relativity of the sign of energy,
 21–23, 25, 26, 47, 169, 286
 requirement of continuity, *see* con-
 dition of continuity in time,
 time direction degree of free-
 dom
 reversal of action, 112, 113, 146,
 168, 234, 235
 reversal of energy, 113, 120, 146,
 234, 235
 time direction-dependent property,
 188, 192, 272
 time-symmetric viewpoint, *see* bidi-
 rectional time, time direction
 degree of freedom
 time irreversibility
 backward in time propagation, 22,
 190, 272, 289
 irreversible processes, 266, 267
 origin, 15, 193
 thermodynamic arrow of time, 188,
 193, 204, 208, 227, 228, 233,
 235, 272, 273, 277, 284
 thermodynamic time, 186, 187, 231,
 289
 unidirectional time, 181, 186–188,
 190, 192, 193, 196, 199, 202–
 205, 207–212, 216–218, 222,
 230, 232, 247, 250, 257, 258,
 272–275, 279, 289
 time travel paradox, 104, 137, 138
 causality violation, 138, 141
 unidirectional variable, 250
 unification scale, 255
 universal force, 72, 73
 vacuum charge neutrality, 96
 vacuum decay problem, 19, 102, 119–
 125, 166, 176, 285, 287
 vacuum energy
 conservation of energy, 116, 166,
 177
 cosmological constant, 8, 88, 94,
 110, 157, 162, 164, 168, 174,
 178, 179, 226, 276, 284, 287
 equilibrium state, 90, 160
 from quantum fluctuations, 88, 160
 gravitational potential energy, 127–
 129, 136, 139, 166
 ground state, 102, 123, 176
 interactions with matter, 110, 120,
 141, 144, 165, 168, 175
 local absence of absence, 100, 175
 natural maximum densities, 160–
 162, 168
 negative contributions, 8, 88, 94,
 95, 159, 168, 174, 284
 negative densities, 34, 35, 88, 90,
 102, 119, 124, 173, 285
 virtual particles, 132

- virtual processes, 88, 132
- zero point, 123
- voids in a matter distribution, 80–
 - 101, 108, 143, 144, 173, 175
 - Birkhoff’s theorem, 82, 83, 172
 - effects on expansion of space, 80, 82
 - gravitational dynamics, 80, 81, 110, 144
 - inertial mass, 80
 - negative energy matter distribution, 90, 99, 100, 173
 - spherical voids, 82, 172
 - surrounding overdense shell, 86, 96, 97
 - the hollow sphere analogy, 82–86, 172
- voids in negative vacuum energy
 - mutual interactions, 99, 174
 - positive energy matter, 90, 94–97, 99, 110, 143, 144, 157, 160, 161, 173, 174
- wavelength, 251
- wormhole, 137, 138
 - exotic matter, 137, 138
 - instability, 138
 - throat, 137
 - traversable, 137, 138, 177