The Substantiation of Free Will

Abstract

1. First, the difference between reality and description is determined. Based on this, it can be shown that the physical causality – in the following referred to as "causality from below" – is *incomplete*.

2. This is a necessary condition for assuming causality in more complex layers of reality governed by nonphysical laws. This type of causality – in the following referred to as "causality from above" – is explained by an example and then generally justified.

3. The explanation applies also to the human neural network. From this follows that the mental layer is the *causal layer* of the network.

4. In contrast to the laws of physics, mental laws are changeable. Since the mental processes are the causal processes, also these changes must be attributed to the mental activity.

5. Therefore, to a voluntary decision the following applies:

a) It is not a *physical* but a *mental* process.

b) The decision-making process can change the laws that applied before it started. However, if only by this process itself is decided what will happen, the decision cannot be determined beforehand.

So it is free.

1. The Difference Between Reality and Description

In our universe, the following seems to apply:

Everything that exists **consists of elementary objects that interact with each other. How these objects behave *is completely regulated by physical laws*. Thus, the entire future development follows from so-called "initial conditions" – the totality of the attributes of all objects at any point in time – and physical laws.**

In this picture that is so convincingly presented to us by science, there seems to be no room for anything other than physics. No matter how complex the aggregates are into which the elementary physical objects are assembled, no matter what fantastic creations evolution produces – *ultimately* everything remains physics. There is just no room for anything else.

This fact can be specified as follows:

In this so-called *reductionist* view of reality just presented, causality always remains "below", i.e. in the *elementary layer* of reality. All other, more complex layers have lost their independence. Descriptions that refer to these layers – such as neural or psychological descriptions of human actions – are just simplified, approximately valid summaries of processes that are actually of physical nature.

The consequences of these hypotheses are rather strange, if not to say bizarre. If we assume, for example, that we made an assertion B ***because*** it is logically correct, then that would be a self-deception: It would mean postulating a causality at the level of mental processes, so to speak a causality from "above" – which, however, is inadmissible after what has just been said; B would then be "causally overdetermined". If this "causality from above" could actually claim an independent existence – *in addition* to the "causality from below" – then it would have to be possible to decide *against* the physical causality.

There would be only one possibility that B could actually correspond to logic: that evolution had adapted the physical processes in our brain to the requirements of reality to such an extent that we behave and think logically to a sufficient degree for our survival. But I emphasize again: the conviction that we made the assertion B *because* it is logical would be a delusion, a ruse of evolution to reinforce our adapted behavior through a pleasant feeling. And, incidentally, we would never be able to determine whether something like "logic" exists at all, since *understanding* something would also be a mental process that does not exist *as such*. Insights would not be insights, thoughts would not be thoughts, mind would have disappeared, *we ourselves* would have evaporated in the fog of self-delusions ...

So it is a completely absurd picture that follows from the reductionist view, and I believe that it is only so widespread because no reductionist has ever fully considered the consequences of his or her convictions. (If there still were one, however, he or she would have long since fallen silent and would therefore be untraceable.)

I want to briefly touch on the two most popular attempts to "defuse" the problem.

The first objection is, that – because of quantum mechanical uncertainty – in nature itself an "objective indeterminacy" exists, so that it cannot be said that "the future follows from initial conditions and laws". However, it can be said that "the future depends *exclusively* on initial conditions and laws" – save that these laws are no longer deterministic. The following conclusions then remain valid.

However, the most common objection to reductionism is, that in most cases a complete reduction has not been achieved and will probably never be possible. I consider this objection inadequate: whether there *is* a reduction cannot be decided by whether *we* are able to carry it out – the picture of reality sketched above, which is the basis of the incredible success of natural science, is not questioned by the restrictions which *our* means and abilities are subject to, and this applies also to the conclusions drawn from this picture.

Therefore, in order to avoid these strange inferences, it is necessary to put the picture itself into question. So we ask: *Is the hypothesis A true?*

A: ***Everything which happens follows from physical laws and initial conditions.***

Let us start with a thought experiment:

We consider the following scenario: a large number of any material objects in empty space that are moving randomly relative to each other, but in such a way that they remain gravitationally bound to one another.

Let us assume that we were able to grasp the initial conditions – the totality of the attributes of all objects of the system – with absolute precision and transfer them to a description. So we ignore that we cannot measure with infinite accuracy, or that we are not even able to write or store the value of a single attribute with infinite accuracy. We also assume that our law of gravitation is correct and that we are able to perform all the necessary calculations.

Now we compare the situation in the *really existing system* with the situation in the *description system*.

Under the above conditions, in the *existing system* exactly what we expect will undoubtedly happen: every object will behave precisely as gravitation dictates. Thus, here, hypothesis A seems to be confirmed.

And in the *description system*? Well, here, at first *nothing at all* happens. Although we have inserted the infinitely precise values of all attributes into our equations, so that they actually represent the objects and their development in time *perfectly,* still the equations do not behave like the objects themselves: While – starting from the point in time that we have chosen to measure their attributes

– the *actually existing objects* move on *by themselves* and, in this way, carry out the gravitationally determined dynamics of the system, the *equations* obviously do *not* do that – they simply remain unchanged as we have noted them.

This is actually completely obvious. Nevertheless, I was a little more explicit than necessary because here we have come across an extremely important issue, which, however, so far almost completely escaped both philosophical reflection and scientific research – presumably precisely *because* of its ostensible obviousness. It reads as follows:

Proposition:

There is a fundamental difference between a really existing system and its representation: the really existing system is active, but the representation is not active.

Let us return to our thought experiment. We have stated: In the *existing system*, every object will behave exactly as gravitation dictates. Does this actually confirm hypothesis A?

The answer is: *No, it does not!* Actually, we have *added* something to the really existing system that is not contained in A: *activity.*

The fact that reality is *active* means: at any point at any time exactly what has to happen happens *by itself*. It means that reality does not need a law or an algorithm, because it simply processes all individual cases at the same time.

Obviously, however, *activity* is precisely that which cannot be transferred from the reality to its representation. It can be said that the *type* of activity of the system, its *specific structure*, must be contained in our equations of the gravitational field, but the *activity itself* is missing.

Let us note: Because of its *activity*, reality advances *by itself* from the present to the future. But the description system refuses to do us this favor. In order to obtain information about the future of the system in our description, we therefore need a *mathematical procedure* that substitutes the missing activity.

Do we have such a procedure? First of all, it is clear that for a "large number" of objects that move randomly, our equations cannot be solved. In fact, we have only one way to obtain knowledge about the further development of the system: Since we know the gravitational field, we can calculate for each object where it *would have moved* after a certain time interval *in this field* – and here, the subjunctive is necessary because of course it does *not* move in *this* field: indeed not only the object we are looking at is moving but also all other objects, and this means that also the field itself is constantly changing. But in order to be able to calculate anything at all, for small time intervals we have to assume the field as *static*. We then do the same kind of calculation for all bodies. Then we repeat this procedure for the next time interval etc.

The crucial point is that from start to finish we depend on *approximations*, and that we also do not know to what extent our calculations deviate from reality. At the latest after the next branching point – that is a point in the development of a system at which an arbitrarily small difference in the initial conditions can lead to completely different states of the whole system – our prediction becomes pure luck.

With this we have shown that hypothesis A is false. Since there is no procedure which enables us to conclude the future from the present, A cannot be maintained.

Proposition:

There are systems whose future development does not follow from physical laws and initial conditions.

But isn't reality itself constantly showing us that the future follows from the present? Not at all. What we see is just that the future "follows" the present. It is only this suggestive picture of reality conveyed by physics that leads us to believe that everything "follows from" initial conditions and laws. However, the expression "follows from" is a logical conjunction that can only relate to a description. To apply it to reality means to replace the "follows" that we observe with the "follows from" that we postulate; But we have to *justify* this act of substitution, and so we are forced to replace our "follows from" by a series of logical steps. Thus we inevitably end up with a mathematical procedure, and finally again with the fact that no such procedure exists – even if we imagine we were freed from all restrictions of measuring and calculating.

So the future does not always follow from the present. What does this result mean?

The most important consequence is that a *logical free space* is created: If initial conditions and physical laws were sufficient to derive the future, then there would be no room in the set of conditions for the derivation of the future; But since they are *not* sufficient, there is now room for further elements in this set.

Proposition:

Causality from below is incomplete. There is room for causality from above.

2. Non-Physical Causality

Our next step will be to clarify what kind of "further conditions" could exist on which the future development of systems depends – in addition to initial conditions and physical laws. Is it any other kind of data? Or other kinds of laws? To determine this, we change the scene.

We consider a simple glass vessel. When we hit it, it vibrates and makes a sound. What does this tone depend on? What determines its height and character? The answer is: *the shape of the vessel.* It gives rise to a mathematical law that enables us to predict the vibration pattern of the glass. So here we don't have to go into the physical objects – the glass molecules – nor the physical interaction – the electromagnetism – in order to predict the sound. The only physical information needed is the speed of the sound propagation in the glass.

The law that now allows us to predict the future of the system is therefore *not a physical law.* It belongs to another kind of laws which I shall call ***Laws of Form*** or ***Laws of Structure***.

Let us compare our two scenarios, that of the gravitating bodies and that of the vibrating vessel:

In the gravitation scenario, the initial conditions are given as ***local parameters***, as attributes of the individual bodies. Their values are inserted into the ***physical law*** – the law of gravity. Although everything that happens fully conforms to this law, it is still impossible to predict the further development. The future of the system ***does not follow*** from its present.

In the glass scenario, it is not the attributes of the glass molecules that are inserted into the law, but the dimensions of the glass, i.e. ***global parameters***. The law is not a physical law, but a ***Law of Structure***. The further development can be derived from the global parameters and the law. The future of the system ***does follow*** from its present.

The sound that we hear is largely independent of the way we produce it. However, this does not apply to the first moment: initially, there is a transient process that depends on how we strike the vessel. Only after this process it does always vibrate in the same state. This state to which the glass ultimately adapts – the vibrational pattern into which it develops and which it then maintains – is called ***attractor***.

Above, we asked ourselves what types of data and laws could there be in addition to physical initial conditions and laws. The simple example of the vibrating vessel gave us an answer:

1. new data in the form of *global parameters.*
2. new laws in the form of *Laws of Structure* that are based on the global parameters.

Since these new data and laws can be used to predict the future of the system, they are in fact elements of the "set of conditions for deriving the future" mentioned above.

However, most important for our considerations is undoubtedly the following:

The local parameters – such as the positions and velocities of the glass molecules – initially depend on where, with what and how hard we hit the vessel. So at first they can be quite different. Regardless of this difference, the state of the vessel always evolves towards the same vibrational pattern – the attractor.

In the case of a glass vessel, there is only one possible vibration pattern that always develops, regardless of how the vessel is struck. The future movements of the components of the vessel – the glass molecules – are therefore determined by this pattern.

Causality works from the whole to the individual, from the vessel to its components, and not the other way round.

Proposition:

A form of "causality from above" occurs when in a system *attractors* exist, i.e. states which the system will *inevitably* evolve into, if it is "close enough" to the attractor state.

(A necessary condition that it is actually "causality from above" is that the physical causality in the respective system – the "causality from below" – is *incomplete,* just as we have demonstrated in the gravitation scenario. However, since the glass vessel was only intended to demonstrate what our argument is about, we do not need to worry about whether this condition is met here.)

Now we have made all necessary preparations to move on to our final and decisive scenario:

3. The Human Neural Network

Subject of our investigation is the following question:

What kind of causality does the neural network obey?

In the network, there are three levels of increasing complexity:

1. the physical level
2. the neural level
3. the mental level

In relation to this classification, our question is:

Of which kind of processes does it depend what happens in the net? Of physical, neural or mental processes? Which level is the causal level? – Or, to put it another way: Which level is dominant?

First to the physical level. Let us assume we had complete knowledge of the values of the attributes of all physical objects in the network and could thus set up the system of equations that represents the state of the network and its further development. (Of course this idea is completely absurd, but in the form of a thought experiment it is permissible – *in principle*, this system of equations must exist.)

But now we are again confronted with the problem that already prevented the calculation of the development of the system in the gravitation scenario: An enormous number of processes are

running at the same time, and each of them is directly networked with several others. In order to be able to calculate any process, we have to assume at least for a small time interval that its immediate environment is constant – i.e. we have to isolate it for a short time. Then we can do the same for all other processes, and after that we repeat the whole procedure for the next time interval etc.

As with the gravitation scenario, we are therefore dependent on approximations that can deviate considerably from reality already after a short time. It is not possible to predict how the network will develop. The claim "What happens in the network follows from initial conditions and physical laws" is wrong.

And here, too, the following applies again: Reality does what we are not able to do: due to its *activity*, it executes the enormous number of processes at the same time, so that we get the impression that everything "follows from" initial conditions and physical laws.

Proposition:

In the neural network, the physical causality is incomplete. There is room for causality from above.

Let us now consider the *neural level*. It consists of many billions of neurons. Each neuron is directly connected to hundreds or even thousands of other neurons, and *all* neurons are linked to one another via a few intermediate steps.

The neural activity is regulated by a law that follows from the neural input-output mechanism.[1] This law can be understood as the *law of interaction* of the neurons. (It also serves as basis for computer simulations.)

Also at the neural level, it initially seems completely natural to us that what will happen in the network follows from the initial conditions of the neurons and their law of interaction. And again we have to recognize that we succumbed to the same deception, in that we have not differentiated between reality and description or confused them:

Since the neural interaction law is a summary of physical circumstances, the argument with which we have just refuted the claim that everything follows from initial conditions and physical laws remains valid. Thus for the neural level the following applies: The high degree of networking of the neurons – the permanent feedback that results from it – precludes the existence of a mathematical method for calculating the further development.

Proposition:

Also the description by neural initial conditions and the neural interaction law leaves room for causality from above.

This brings us at last to the most complex level, the *level of the mind*. We make the following assumptions:

1. Every kind of mental activity (thoughts, chains of associations, sequences of images, etc.) is a sequence of neural activation-patterns.
2. Sequences of neural activation-patterns can be representations of facts.[2]

Let us look at the neural patterns. How do they become representations?

---

1 The expression "input-output mechanism" means the following: The dendrites of each neuron are stimulated or inhibited by other neurons via synapses. The electrical excitation caused in this way is passed on to the cell body and added up there. When a certain limit is exceeded, it is released to the axon and distributed to its branches, so that ultimately it influences other neurons via synaptic connections.

2 Here, "facts" must be understood in the widest-possible sense.

Let us imagine a neural network in which there are no representations yet. An object perceived for the first time will cause a certain pattern in this network, starting from the primary visual cortex. The neural connections that are active are strengthened because of this very activity. The same is the case with each repetition. This gradually creates a stable connection between the object and a specific neural pattern (or rather an ensemble of specific neural patterns).

In addition, the following applies: Although the neural patterns are initially caused by external stimuli, after a sufficient number of repetitions they are also produced by the neural network independently of these stimuli. This means:

Neural patterns that are connected to objects in the manner just described are* attractors *of the network.

Previously we have stated:

Under the condition that the causality from below is incomplete, from the existence of attractors follows that the respective system – provided it is "close enough" to the attractor state[3] or in this state itself – is governed by causality from above.

However, according to our first premise, a mental process consists not only of neural patterns, but also of the transitions between these patterns. But to this transitions the same applies as to the patterns themselves: First, they are determined by the sequence in which the causative objects appear. If this sequence is repeated, the corresponding neural activity is reinforced, and this has the consequence that the patterns occur again in the same sequence even if they are generated by the network itself. In the same way, also the spatial relationships of the objects are transferred to the patterns.

This means: In the processes that are generated by the network itself, the neural patterns that are in a stable connection with specific objects appear in the same spatial and temporal contexts as the objects themselves. Therefore, *the patterns can be understood as representations of the objects, and the processes as representations of the facts in which the objects appear.*

So, in human neural networks it is not the physical or neural conditions and laws that determine what happens in the network, but *the structure of the network* – the fact which attractors there are and how their sequence is regulated – on which the processes depend that run in the network.

Causality acts from the whole to the individual, from the network on its components, and not the other way round.

We have thus achieved our first goal:

<u>Proposition:</u>

The neural network is regulated by *causality from above*. The mental level is the dominant level. In it lie the *causes* for the processes running in the network.

So the statements we made so far were *actually* conclusions and not just physical processes! Or – to follow up on the formulations used in the criticism of reductionism: Insights are insights, thoughts are thoughts, mind is set in its rights, *we ourselves* are indeed we ourselves …

So far, so good, but that doesn't take us to where we actually want to be. Just because we have moved causality up doesn't mean we are free. We have only replaced physical or neural causality with mental causality. We have thus achieved that our mind is not ruled by physical or neural laws, but *by its own law: the Law of Structure, which the sequence of neural patterns obeys that represent something*.

3 Without the concept of phase space, this "close enough" cannot really be defined. In any case, the neural network is always "close enough" to an attractor state.

But don't we ultimately remain trapped in the scheme of initial conditions and laws from which we wanted to escape? Fortunately, that's not the case. To show this, we need to look at the difference between physical and mental laws.

4. The Difference Between Physical and Mental Laws

Human neural networks differ greatly from one another, even if they have not yet been structured by external stimuli. From this follows immediately that the patterns that represent something are also different in all people, even if the represented facts are identical.

As stated above, initially the order of the patterns is determined by the order in which the objects or circumstances that cause the patterns occur. But as soon as the network itself is able to produce these patterns, the transition rules of the patterns – what we have called the *mental law* – increasingly depend on their use in internal processes. This dependence on external and internal conditions means that the transition rules differ from person to person.

So we have already determined the first difference:

*While physical laws are **generally valid**, mental laws are **individually valid** – they only apply to one singular person.*

Connections between neurons are strengthened when they are active,[4] and weakened when they are inactive. This means that every mental activity alters the structure of the network. But if the structure can change, then obviously also the rules that determine the sequences of the neural patterns can change.

So this is the second difference:

*Physical laws are **immutable**, mental laws are **modifiable**.*

Proposition:

Physical laws are universal and immutable. Mental laws are individual and modifiable.

5. The Substantiation of Freedom

The most obvious implication of the strengthening of active neural connections is that what we always think, feel and do is self-reinforcing. Basically, however, it goes without saying that also the opposite can occur:

We have shown that causality is to be found at the mental level. *Will* and *intention* must be understood as elements of mental causality. Now let us imagine concretely we were faced with an important decision. When we enter the decision-making process, we are initially guided onto certain, well-known paths by the regularities that are valid up to that point – i.e. by our own mental law.

But at any time we are able to leave these paths, for example by simply considering the opposite of what we have assumed up to then, or by taking a path we never tried before; We are able to do so precisely for the reason that the causes for what happens in the network – and thus also for the modifications of the network structure – lie at the mental level.

4 This finding goes back to Donald Hebb, who stated in *The Organization of Behavior* in 1949: "When an axon of cell A is near enough to excite B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased."

In other words:

The law that determines the sequence of neural patterns in our network that represent something, i.e. our own mental law, can be altered *by ourselves*: we ourselves can change the laws of our thinking and acting through our thinking and acting, and we can do it *deliberately.*

This means at the same time:

Although mental processes are governed by their own rules, it is not possible to derive a volitional decision from them: the decision cannot be contained in these rules because they can be changed by the mental process that precedes the decision. While this process is taking place, the laws that it obeys can change – or, more precisely, *it itself* can change the laws that applied before it started.

Proposition:

Volitional decisions are causes of actions. Since only by the decision-making process itself is decided what will happen, the decision is not determined beforehand.

So the decision is free.

To the question of why a (sane) person has decided so and not otherwise, there is then only one permissible answer:

Because he/she wanted it that way.

Note:

Of course this does not mean that volitional decisions cannot be analyzed with respect to their neural, chemical, physical, genetic, social etc. causes. It means, however, that these analyses necessarily remain incomplete and never lead to a secure result, because mental phenomena cannot be reduced to other layers of reality. The will remains the final authority.

Postscript:

In reviewing the text, it seemed to me that I followed my goal of presenting the topic as briefly and simply as possible perhaps a little too radically. Therefore I will try to explain the most important points of my argument one more time:

Let us assume we have to describe a system that consists of a large number of physical processes that are linked to one another. Then the equations of the processes are also networked with one another. For an exact description, we would therefore need the values of all parameters of any process at every moment in order to insert them into the equations of all other processes – in other words: it is (except in very simple cases) *impossible* – for reasons of principle, and not just because of the limitations of measurement and calculation – to make accurate predictions about the system that consists of all these processes by using *physical means*.

And with that we would have actually reached the end of our possibilities – *unless* the processes could be understood as elements of a "structure of a higher order", in which further laws apply. These "higher order laws", however, are then *no physical laws*, and with that we have left the field of physics.

If these new laws make it possible to predict the development of the overall system, then the following applies:

1. The development of the overall system ***does not follow from physical laws***.
2. The development of the overall system ***does follow from higher-order laws***.

Of course, everything continues to happen ***in accordance*** with the laws of physics – but these laws now take place within a ***higher-level structure***. (Think of the vibrating glass vessel.)

Thus causality is no longer *below,* which means: in the elementary, physical realm. It has migrated *upwards*, into a realm of higher order, in which ***new, non-physical laws*** apply.

Exactly these conditions can be found in the neural network, and in fact several times:

In a neuron, numerous physical processes take place at the same time. Although the physical approach allows us to understand what is going on in the neuron, still the coupling of the processes prevents any exact calculation of the further development. However, due to the shape and structure of the neuron, these processes are embedded in a system of higher-order, so that they obey a "structural law" – the one that we previously called "neural input-output law".

Now, however, it again applies that this law does not allow us to make any precise predictions about the future development of the many neurons that are coupled to one another. But the neurons themselves are again elements of a higher-order system: the neural network with its imprinted patterns (attractors). So the neurons are also subject to a new law: a structural law of again higher order: the law of the sequence of neural patterns, and that means: ***the law of the mind***. Thus mind is the causal layer; It determines the processes that take place in the network – including those that change this law itself.

Finally, I shall repeat the difference between description and reality:

In order to get from the present to the future in the ***description*** of a system, we need some kind of procedures. These can be mathematical procedures, algorithms or equations, but also methods to combine facts in such a way that conclusions can be drawn. In some cases we are able to do this so well that we can state: *B follows from A*.

In the ***reality***, none of this is necessary. If what has to happen happens in every place at every time, then the future will arise *by itself*, and then all complex objects and structures, including their laws, will develop *by themselves*.

But from the fact that in the reality the execution of elementary processes is sufficient for the creation of the future, it cannot be concluded that the future ***follows from*** elementary processes, because that would presuppose that that, what in the reality happens ***by itself***, can be expressed by ***a series of logical steps***, and that is impossible.

Note:

In this justification of free will, it is not necessary that a "bifurcation" exists in the development of the world. The key point here is that *the future is not contained in the present* – that is, it does not *follow* from the present but merely *arises* from it, and that the reasons for what will then actually happen are of a mental nature.

Note:

In order to recognize objects, artificial neural networks must be trained on large data sets. In numerous repetitions the connection strengths of their neurons are varied until a sufficiently high recognition rate is achieved.

In contrast, we started from the following hypothesis: A perceived object, which causes a neural activation pattern, is represented *by this pattern itself*. Therefore, here the relationship between object and representation is not established by varying the connection strengths of the neurons, rather it exists already from the beginning and is only stabilized and specified by *strengthening* the active connections, whereby the neural pattern becomes an *attractor.*

This hypothesis is confirmed most clearly by the so-called "imprinting". (As e.g. in the case of the gray geese of Konrad Lorenz). There are neither "large data sets" nor "numerous repetitions" – the process occurs almost instantaneously.

Furthermore, thereafter *immediate recognition* occurs, despite the inevitable variability of the sensory impression to be recognized. Thanks to the attractor concept, this – otherwise hardly explainable – performance becomes self-evident: as long as the sensory input is within the catchment area of the attractor, it obviously applies: *perceiving = recognizing*, since the newly activated attractor already represents the object, so that further calculations are unnecessary.

Note:

Finally, a comment on the scenario of the gravitating bodies in the first section:

Even a Laplacian demon with infinite resources of space, time and information could not carry out the calculation: In order to *accurately* determine the future of the system, the demon must perform the calculation for infinitely small consecutive time intervals. If the interval boundaries are as close as the *real* numbers, the calculation will not be finished even after an infinitely long time, but if they are less close (like the *rational* numbers, for example), it will happen that an instability is missed that occurs *between* two time points of his calculation.

In fact, even with this argument, we have still not grasped the full extent of the problem: We have assumed that – because we possess complete knowledge of the initial conditions – we know the gravitational field. However, this assumption is wrong for the following reason:

Let us denote the point in time at which we have precise knowledge of the initial conditions – and at which our calculation should begin – by t_0. If we want to calculate for any of the bodies, let's say for body A, where it will move in the first time interval, then we must know all effects from the other bodies which A is exposed to at time t_0.

For example, let's look at body B: we know the position where it is at time t_0. However, the effect originating from B that A is exposed to at time t_0 does ***not*** originate from ***this*** position, but from a position where B was ***before*** – exactly as long before as it took gravity to move ***from there*** and reach body A at time t_0. Therefore, in order to determine the effect of B on A at time t_0, we have to put B on its path *into the past*, and exactly the same applies to all other bodies: they all have to be put into the past – the further, the further they are away from A.

This means: Before we can even ***begin*** to determine the path of A, we first have to determine the paths of all other bodies. But for that it is necessary to also know the effect that A has on the other bodies at time t_0, and therefore we also have to shift A itself on its path into the past, i.e. on the path that is *not known to us*, since we just wanted to calculate it!

The same applies to *every* body: in order to shift it into the past, we must know the paths of all other bodies. However, since we do not know *a single one* of these paths, it is impossible to determine the exact positions where the bodies were before, and therefore it is also impossible to determine the effects which they are exposed to at time t_0.

In other words, we – and by "we" I mean all of us ***and*** Laplace's demon – are not only unable to ***perform*** an **accurate** calculation of the future, we are even unable to ***begin*** with it.

The scenario is not computable. *Reality* is not computable.

[So the formal version of our ontological argument about free will is as follows:

The behavior of all elementary objects is determined exclusively through physical laws.

But if you try to derive the future (or, if objective chance should be factored in: *any* version of the future) in a physical way, you fail because it would require an uncountable number of logical procedures.

In some cases, however, the uncountable set of logical procedures can be replaced by a finite set of statements about a higher, *non-physical* level of reality. The facts which these statements refer to can then be understood as causes (or *reasons*) for the future state.]

Heinz Heinzmann

Vienna, 2021

(The arguments which free will is based on can also be used to prove that robots cannot have feelings and consciousness. Read *Why Robots Cannot Feel*, or *Why free will exists and why robots are not sentient*, where the proof is more elaborated.

The argument against computability presented in the last note also serves as *proof of the impossibility of time reversal*.)