# NOTES TO INFORMATION THEORY

From October 2020 to January 2021

Rastko Vuković
Economic Institute Banja Luka (on hold)
1/27/2021

Rastko Vuković:
Notes to Information Theory
(Economic Institute Banja Luka, on hold, 2021)
https://archive.org/details/notes-to-information-theory

Rough translation from the Serbian:

Растко Вуковић:
Прилози теорији информације
(Економски институт Бања Лука, на чекању, 2021)
https://archive.org/details/prilozi-teoriji-informacije

# Preface

These contributions are sporadic, collected so as not to be lost. I believe that it has happened to everyone to have a thought, a budget or an item that got lost somewhere and suddenly became very important, and not to be easily repeated. This also happens with rejected theories.

My best contributions are always unfinished. I am in a constant phase of checking and correcting ideas that seem interesting to me, and there is no point in taking such diggers more seriously and promoting and discussing them on forums. People ask institutions for mutual confirmation, for the sake of advancement in business or simply because they are trained to trust authorities more than themselves, so that is more of a reason to avoid the public.

However, the real reason for the "privacy" of these works is my private experience of them. I could not compete with thousands of brilliant researchers who would spread these "my topics" in all directions, I could no longer consider the discoveries my own and I might be left without inspiration. Working in scientific teams today is and very frustrating (although participants are often unaware of it), which is why they become workers doomed to small steps and eventually to the hope that quantity will grow into quality. They become people with a chronic lack of ideas and, with all the excess of what they have learned, with a constant feeling of insufficient information. I find it difficult to recognize myself in that environment.

The way these articles are written explains the scattering of their points, frequent returns to the same topics and other repetitions, permanent incompleteness. But it each makes part-by-part shifts in the theoretical discovery of the notion of information, not to mention a world that is hypothetically assumed to be made up of it.

Author, November 2020.

## Content

# 1. Violation of Similarities

Rastko Vuković[1]

In the "information universe" I analyze, every (physical) phenomenon consists only of some information, and every information of some uncertainty. The indefiniteness as the essence of the information can be inferred from the unreliability of reading experiments, but also from its definition as a quantity of options. The smallest measure of information is a pure uncertainty, but then again the part of such becomes some certainty.

Less uncertainty means more certainty. If there are the least information, then they are the least free (particles), they are the uncertainties which are the packages of non-free certainties. In this way we find the law of conservation information. Another way to "prove" the preservation would come from faith in what we observe through experiments. Information that could by itself increase and decrease without a reason, arise and disappear from nothing to nothing, would make the measuring useless. On the other hand, the law of conservation goes with finite divisibility, because only infinity can be its rightful part (that is the definition of infinity).

The final divisibility causes layering, different properties of different sizes. We find the vertical multiplicity of the world already in the first comparisons of the processes of quantum and larger bodies, in the transitions from micro to macro physics, but they are also in line with the principled uncertainty. For example, the volume of a body increases with a cube of height, but the surface with a square, so even the geometry itself becomes an accomplice of information theory.

A similar example is one task from the percentage account that is given to elementary school students. It is said that a shirt cost 100 dollars and went up by 10 percent, it was unsold for a while and then it fell in price by 10 percent. Its price is no longer equal to the initial one!

Namely, 10% of \$ 100 is \$ 10, so the increased price of a shirt is \$ 110. But 10% of that increased amount will be \$ 11 and after deduction the new price becomes \$ 99.

I point to proportional increases and decreases which are not seen as similarities in geometry, or in homothety, but in the way of linear mapping characteristic of "quantum evolution" (linear operators). We have known since Heisenberg (1927) that the non commutativity of these operators means the dependence of quantum processes and then tells us about the uncertainty relations: by more accurately measuring the position of a particle, we know its momentum less accurately and vice versa.

These mappings are not that quantum-mechanical, but they are so analogous to them that they confirm the mentioned informational nature of the physical universe to which we belong. In other words, the idea of proportional increase or decrease of quantities, transition from the micro to the macro world of physics, is also information and, consequently, it is subject to analogous laws of physical reality. We have them in various forms.

[1] Gimnazija Banja Luka, math prof.

An ant can carry up to 50 extra its own weights, a man barely two, and an elephant a tenth. We have said that with a proportional increase in the area of the physical body, it increases with the square of the length, the volume with the cube, and we further notice that the specific strength of (linear) muscles, heat radiation and energy consumption decrease. The laws of large numbers also concern small ones differently than large ones, and the influence on probability distorts the information as well.

If we define the steps of linear operators of increasing the size and decreasing the relative power (in relation to the mass of the body), we will establish their non-commutativity and appropriate "uncertainty relations" like those in the mentioned "shirt problem". This phenomenon exists wherever we have some perception of information.

The "mechanism" of the British theoretical physicist Higgs (Peter Higgs, born in 1929) acted in the very high energies of the early universe, according to which elementary particles gained mass. This is due to a very large particle, the Higgs boson, which decays soon after its formation, which is why only extremely strong accelerators can register it, so the first experimental confirmation arrived from CERN 2010-2011. year, Fermilab and their Large Hadron Collider (LHC), and then from ATLAS and CMS (Compact Muon Solenoid) independently in 2012.

Larger masses know physical effects unknown to smaller ones. Such is the separation of electrical and weak nuclear power. The weak force acts only at distances smaller than the atomic nucleus, while the electromagnetic force can propagate over long distances (by the light of stars), weakening "only" with the square of the distance. Between the two protons, the weak force is about 10 million times weaker than the electromagnetic one. However, one of the main discoveries of the 20th century was that these two forces are different faces of one, one higher, fundamental electroweak force.

This was achieved during the 1960s by three physicists, Sheldon (Sheldon Glashow, born 1953, American), Salam (Abdus Salam, 1926-1996, Pakistani) and Weinberg (Steven Weinberg, born 1933, American). They independently discovered that the gauge-invariant theory of weak force derives from the electromagnetic one with the help of four massless "messengers" or particle carriers, two electrically charged and two neutral.

The short range of weak force indicates that it is carried by massive particles, which means that the basic symmetry of the theory is hidden or "broken" by something that gives mass to particles exchanged in weak interactions, but not to photons responsible for electromagnetism. The presumed mechanism involves additional interaction with an otherwise invisible Higgs field that permeates all space.

The existence of force carriers, neutral particles Z and charged particles W, was experimentally verified in 1983 at CERN. The masses of the particles were in accordance with their predicted values. The law of large numbers of probability theory is even more undoubted, but I have already written about it before, and here only can be mentioned it in support of the above thesis on the mapping of worlds of quantities in the way of information.

# 2. Dual Vectors

Quantum states and processes

October 18, 2020

A short semi-popular story about vectors and operators, with a review of their representations in quantum mechanics from the angle of (my) information theory.

## Увод

We first imagine vectors as oriented longer. They are equal when they are parallel, of the same length and the same direction, and we consider the opposite sides of the parallelogram to be a typical one and the same vector. The sum of the adjacent sides of a parallelogram spanned by two vectors is one (larger) diagonal, and their difference is other (smaller) diagonal of the parallelogram. We multiply the vector by as many times as we extend it.

Abstracting, forces and velocities are vectors. The addition of the force vector is recognized in the pull of the buried column in two directions, with the resulting force equal to the diagonal of the corresponding parallelogram. An example is the addition of the velocity vector in the case of a river flowing through a valley at one speed and a boat cutting its course at its own speed, with the resulting speed of the boat relative to the shore corresponding to the diagonal of the imagined parallelogram. With the exception of formalization, this applies approximately also because this vector analogy ceases at velocities close to light, or in strong gravitational fields.

A vector space is a set of $X$, elements that we call vectors, with a addition operation so that $(X,+)$ has the structure of an Abel (commutative) group. We denote the neutral element of that group by 0 and call it the zero vector. In addition, there is a set $\Phi$, whose elements we call scalars, and which with addition and multiplication operations has a field structure $(\Phi,+,\cdot)$, with neutral elements in relation to these two operations 0 and 1.

In addition, the multiplication of the vector by a scalar is defined, which accompanies each vector $x \in X$ and the scalar $\lambda \in \Phi$ with the vector $\lambda x \in X$, so that the axioms apply:

$$1.\ \alpha(\beta u) = (\alpha\beta)u;\ 2.\ \alpha(u+v) = \alpha u + \alpha v;\ 3.\ (\alpha+\beta)u = \alpha u + \beta u;\ 4.\ 1\cdot u = u;$$

for all vectors $u, v \in X$ and all scalars $\alpha, \beta \in \Phi$. We also denote this vector space by $X(\Phi)$.

## Theory and Practice

Logic can derive absolutely accurate proofs as opposed to an experiment, which follows from fundamental measurement errors. From the point of view of (my) information theory, we supplement this explanation with the following. When something is more informative, it is more the news and comes to us with greater uncertainty, besides the world is built only by information. Consistently, the fabric of our world is unpredictability, uniqueness and complexity, and we communicate (living beings, inanimate, particles) because we do not have everything. Subjects can never have everything, because such would not belong to the world of information.

By the way, notice that communication is an exchange of information, and according to what has been said, it is an interaction. There is no transmission of information without action and there is no action without transmission of information; that connection is so great that action and information are equivalent phenomena.

So, the basis of the world of information is unpredictability. It is such that less information has more predictability. Where there is less multiplicity, there is more multiplication, with the reduction of complexity like snowflake crystals, the simplicity of mathematical axioms stands out. Simply put, by reducing the concrete it grows abstract, and vice versa, that is because simplicity is the absence of complexity, and repeatability is the absence of uniqueness.

In that theory, more information means less certainty but more action. Therefore, the more important the action is to the experiment, the less certainty it has. The more corporeal he is, the fewer theorems. Experimental proof is a method of contradiction in action; by concretizing the resulting limitation of the otherwise ubiquitous and all-time truth to space and time.

It is an IT explanation of the "secret connection" between theory and practice. Theoretical assumptions are repeated in various practical things beyond recognition, just as much as it is difficult for us to discover the same models. Abstract truths are therefore subtle, because they are as energyless as they are all-times (always), that is, they are as impulsive as they are all-space.

## Quantum Mechanics

The dualism of states and processes of quantum mechanics is a recently discovered example of such complexity, it is said to be a symbiosis of the unitary spaces of algebra and the physics of the microworld. A vector space supplied with a scalar (inner) product is called a unitary space, and the representations of its vectors are quantum states. The properties of these vectors are also possessed by unitary operators, which are otherwise processes, ie evolutions of quantum states, so quantum processes and states have the same form.

The word unitary here means singular, normalized to a unit. In the case of vectors (states), standardization per unit allows us to treat the components of the vector as distributions of probabilities (independent) outcomes of a random event. Such would be, for example, equal probabilities of the outcomes when throwing a fair dice, one sixth each with the unit sum of all six. In the case of an unfair dice, the probabilities of falling six numbers would be different, but again with a total of one.

The scalars of quantum mechanical vectors are complex numbers, $\Phi = \mathbb{C}$. We know that for every complex number $\lambda = a + ib \in \Phi$ there exists a conjugate complex number $\lambda^* = a - ib \in \Phi$ such that $\lambda \cdot \lambda^* = \lambda^* \cdot \lambda = a^2 + b^2 = |\lambda|^2$ is a real number. These products of conjugate complex components of vectors define their norms and probabilities in the usual sense. Thus, for the vector $x = (\lambda_1, \lambda_2, \dots, \lambda_n)$, defined by the $n$ -torque of complex components, the square of the norm will be $x \cdot x^* = |\lambda_1|^2 +$

$|\lambda_2|^2 + \dots + |\lambda_n|^2 = |x|^2$, and for the unitary operator $A = (\alpha\_ij)_{n\times n}$ the product conjugated to it is $A \cdot A^*$, and it is a unit operator whose norm is unit, $|A| = 1$. Otherwise, when this multiplication is implied, then we do not write a dot between the factors.

Along with the interpretation of the quantum vector, the probability distribution is accompanied by the understanding of the quantum state as a superposition, ie the ability of the quantum system to be in more possibilities at the same time, until measurement. By the act of measuring, ie interacting with the apparatus, the uncertainty is transformed into one of the possible certainties when the potential information becomes actual. The amount of uncertainty before the realization of a random event is equal to the information after.

In the case of operators (processes), the standardization per unit means that the original and the copied vector will be the same norms, and due to the standardization of the vectors themselves, the reversibility of these operators arises. They "remember" so that for each there is a unique inverse operator whose action the copied vector returns into the original.

It is assumed that unitary operators are linear, that a twice-applied operator makes twice as much change, which is why they formally satisfy the assumptions of the vectors. These operators are thus dual to the vectors they act on and vice versa, the vectors are dual to the operators acting on them. They form of a pair mutually dual vector spaces and consistently, we say that the space of quantum states is dual to the space of its processes.

Dualism will map the uncertainty of a particle over time to the equivalent of the uncertainty of the distribution of particles in space. Namely, in order to determine the arrangement of particles, we move with restrictions due to which we do not have accurate knowledge of their current positions. In particular, the limitation of the speed of light is a consequence of the mentioned dualism of operator and vector!

The electric field acts to move the charged particles, and this in the dual interpretation becomes the claim that the charged particles move the electric field. Also, the transformation of an electron into an electron from which it seems to us that particles are picky in relation to their processes, has a dual interpretation that processes choose states to replicate.

## Composition

Each unitary operator can be represented as the product of two operators, $A = BC$, where one of the factors can be given in advance. If the given operator is unitary, then the second factor is also unitary. This follows from $AA^* = (BC)(BC)^* = (BC)(C^*B^*) = B(CC^*)B^* = BB^* = I$. Namely, when both factors are unitary operators then $BB^* = CC^* = I$, where $I$ is the unit operator. This gives us the idea to understand each quantum process as a composition of two processes, which may or may not be observable.

An example of an observable such is the decay of particles, the spontaneous process of converting one unstable subatomic particle into several others, whereby the resulting particles must be less massive than the original, and the total mass of the system preserved. A particle is unstable if there is at least one allowed final state in which it can collapse, and then it will often have several ways of decay, each with its own probability. The resulting particles themselves may be unstable and susceptible to further decomposition.

Particle physicists are persistently discovering new seemingly elementary particles, which they make using special machines like the LHC (Large Hadron Collider). Many of them decompose into other particles in a small fraction of a second (trillions of trillions of parts and less). This decay is already considered the fate of most elementary particles.

# 3. Potential Information

October 19, 2020

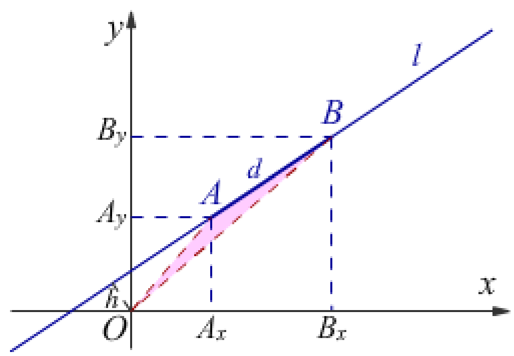
In (my) information theory, physical information is equivalent to action, and both information and action are equivalent to the area, which is demonstrated here in unusual and otherwise familiar ways.

## Introduction

The basic hypothesis of the information theory that I consider in the text says that space, time and matter consist only of information, and that the essence of information is uncertainty. However, nature everywhere tries to avoid its summary so much that we have two fundamentally equal minimalisms, information and actions, and hence the equivalence of those two concepts.

## Commutator

I also use the opportunity to demonstrate the method of commutators that I developed for the needs of "information theory" (still widely unknown). You can treat the calculations that follow as independent examples, but we will eventually combine them into one story.




**Example 1**. In Cartesian rectangular coordinate system OXY we observe the line l, on which are two points $A(A_x,A_y)$ and $B(B_x,B_y)$, as in the figure on the right. The double area of triangle *OAB* is

$$2\Pi(OAB) = [A,B] = A_x B_y - B_x A_y \ .$$

We prove that this formula is correct, for example, by analytic geometry and expressions for the area of a triangle by a determinant:

$$2\Pi(OAB) = \begin{vmatrix} 0 & 0 & 1 \\ A_x & A_y & 1 \\ B_x & B_y & 1 \end{vmatrix} = \begin{vmatrix} A_x & A_y \\ B_x & B_y \end{vmatrix} = (A_x B_y - B_x A_y) = [A,B].$$

**Example 2**. In the same figure, the distance of a given line $l$ from the origin $O$ is $h$, and let the distance between the given points be $d = \overline{AB}$. Then the surface is of the same triangle $\Pi = hd/2$. We further note that by sliding points *A* and *B* along the line *l* so that the distance between them remains constant values of *d*, and by translating longer *AB* along the line *l* the area of the triangle *OAB* remains of the constant value Π.

Let us move this image into the physical space in which the material point *T* moves along the given line *l*, freely, without the action of external forces. Point *T* for equal times exceeds equal distances *d*, which means that the area of triangle *OAB* is always the same *Π*. This applies to a given line *AB* and an arbitrary fixed point *O*.

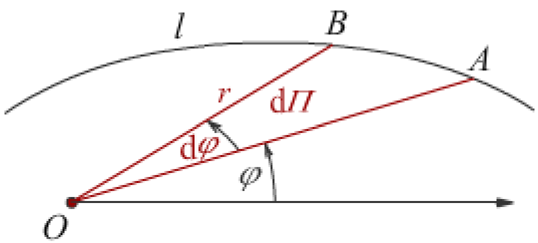
More precisely, the inertial motion is determined by the constant value of the mentioned commutator [A, B]. From theoretical physics we know that bodies generally move according to the principle of least action, and from the said information theory we additionally learn that they move trying to communicate as little as possible. Hence the conclusion that both the action of point *T* and its exchange of information are proportional to the commutator.

When the value of the commutator is zero, the given line contains the origin and we can say that the point *T* moves vertically (orthogonally) towards or from *O*. It does not communicate with the space of the force field at all, but is directed straight to the smaller potential of the field.

## Gravity

Let us now consider this in the gravitational field with respect to Kepler's second law: the radius vector from the Sun to the planet erases equal surfaces at equal times.

**Example 3**.
In the picture on the left, the Sun is at the origin of the polar coordinate system $Or\varphi$ which is the focus of the ellipse $l$, and the planet orbits the line of the ellipse moving from point $A$ to point $B$. Let's say that points $A$ and $B$ are so close that the angle $d\varphi = < AOB$ infinitesimal. Area of $OAB$, where arc $AB$ in the




infinitesimal can be considered as a straight, is $d\Pi = \frac{1}{2}\overline{OA}\cdot\overline{OB}\cdot\sin d\varphi \;= \frac{1}{2}r^2 d\varphi$, because the both pulls, $OA$ and $OB$, are $r$. Radial and vertical velocities are:

$$v_r = dr/dt,\;\; v_o = \frac{rd\varphi}{dt} = r\omega,$$

so for the swept area we can write:

$$d\Pi = \tfrac{1}{2}r^2 d\varphi = \tfrac{1}{2}r^2\omega dt,$$

and for the angular moment of the planet in orbit:

$$L = mrv_o = mr^2\omega.$$

Therefore, the transition speed is proportional to the angular momentum and is equal to $L/2m$.

Newton's laws say that the rate of change of the angular momentum is equal to the torque of the forces acting on the system. We assume that gravity is the only force acting on a planet orbiting the Sun, that it acts by pulling from the planet toward the center of the Sun, and that it therefore has zero torque. There is no lever around that central point, so the angular momentum of the planet around that point is constant.

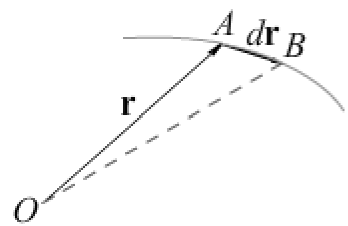
It follows that the speed of sweeping with a radius (vector) is also constant, and that is Kepler's second law, that a radius sweeps equal surfaces at equal times. These are, of course, well-known things, which is why I state to emphasize that the overwritten surfaces are proportional to the action, and this is the emission of information, that is, the communication that the planet has with the gravitational field.

My main thesis, I repeat, is that movements in physics generally occur according to the principle of minimum information, and that this is precisely the well-known principle of least action.

## Generalization

The following figure[2] on the right shows the triangle $OAB$ spanned by the vectors $\mathbf{r} = \overrightarrow{OA}$ and $d\mathbf{r} = \overrightarrow{AB}$ formed by the infinitesimal displacement $d\mathbf{r}$ of the body acting on the (unknown) force from point $O$. Note that this movement $A \to B$ does not have to be under the action of a gravitational force (with center at point $O$), but it must be a constant force, attractive or repulsive, from $O$.

**Example 4**.




The area of the triangle $OAB$ is half the intensity of the vector product of the vector that spanned it, so $d\Pi = \frac{1}{2}\mathbf{r} \times \mathrm{d}\mathbf{r}$. Its derivative in time gives $d\dot{\Pi} = \frac{1}{2}\mathbf{r}\times d\dot{\mathbf{r}}$, and the second derivative $d\ddot{\Pi} = \frac{1}{2}(\dot{\mathbf{r}} \times \dot{\mathbf{r}} + \mathbf{r} \times \ddot{\mathbf{r}})$. The first addition to the right in parentheses is zero, because the vector product of parallel vectors is zero. In the second addition, the vector $\ddot{\mathbf{r}}$ is the acceleration of the body that is proportional to the force, so it has the same direction as r. That's why the second item is zero too. Thus, $\ddot{\Pi} = 0$, and hence $\dot{\Pi} = \text{constant}$. If the force is gravitational, this proves the Kepler's second law. However, the force in that figure also may be different.

The radius (vector) from the source of force $O$ to the body $A$ on which the force acts, at equal times overwrites the same surfaces and with some other types of forces. This will apply equally to the gravitational, electromagnetic, or all attractive or repulsive forces of such a point rise.

The essence of information is uncertainty, ie coincidence, in such a way that a higher probability of a random event means a higher certainty of its occurrence. In short, more probable events happen more often, hence the principle of minimalism of information, that less informative events happen more often. There are examples all around us[3].

We see the principle of minimalism of information in easier coding than decoding, easier spreading of lies than truths on social networks (lies hide information), in forces that always have a direction towards less information. This is a novelty which, compared to the otherwise known principle of the least action of physics, tells us that information and action are equivalent phenomena.

## Epilogue

Erasure by radius vector from the source (constant force) to the moving material-point in equal times gets equal areas, regardless of whether the force in the source is zero, gravitational or some third.

If we compare this with the principle of least action, and this with the principle of least information, we will conclude that the change in surface area is proportional to the action, ie information. Physical bodies try to move so that they do not change the action, that is, the information.

The body moves in the considered fields of force so that its total energy (kinetic and potential) remains constant, so it is proportional to the change in surface area. As the change of the surface is proportional

[2] from the book [1]

[3] see book [2]

to the energy, the area itself is proportional to the product of energy and time, ie action, and then information.

# 4. Information Energy

26. октобар 2020.

The importance of energy for information theory is discussed. The emphasis is on the law of conservation and estimating the direction of attraction depending on the type of energy or information.

## Introduction

The change of energy over time is a physical action which is therefore also in transition. In the short appendix [3] is an unusual and brief observation of the law of conservation of action, and in the book [4] you will find examples of the principle of minimalism of information. In brief, the nature is built only by information, but nature would keep it to a minimum. Due to the two principles of minimalism, action and information are equivalent, so both are changing. The "news" uttered a second time is no longer news, because the essence of information is uncertainty. Information is a measure of the amount of uncertainty.

## Hamiltonian

There are rigorous evaluation of Hamiltonians in many places, e.g. [4], so here I give only a brief explanation. The total energy of the physical system is

$$H(p,q) = T(p,q) + V(q).$$

It is the sum of kinetic $T = T(p,q)$ and potential energy $V = V(q)$, where $p$ and $q$ are the momentum and position of the system, respectively. This $H = H(p,q)$ is Hamiltonian[4]. Velocity is the change of position over time, $v = \frac{dq}{dt} = \dot{q}$, and the momentum is $p = mv$, where $m$ is the mass of a given system and, as we know, the kinetic energy is:

$$T = \frac{1}{2}mv^2 = \frac{1}{2}m\dot{q}^2 = \frac{p^2}{2m}.$$

When the system is in the field of a force, we can measure that force by changing the momentum over time, $F = \frac{dp}{dt} = \dot{p}$, work is the effect of a force on the path, $W = Fq$, and potential energy $V = -W$. Hence it derives position and momentum by time using the sum of total energy by momentum and position:

$$\dot{q} = \frac{\partial H}{\partial p} \text{ и } \dot{p} = -\frac{\partial H}{\partial q}.$$

These are the famous Hamiltonian equations of classical mechanics. We get them in analog form in other areas of physics and they always tell us that there is no change of position without a change of

[4] William Rowan Hamilton (1805-1865), Irish mathematician.

momentum and vice versa, that without a change of position there is no change of momentum, nor does any of that happen without a change of total energy.

Note that the Hamiltonian equations do not contain time, from which follows the law of conservation of energy, and then the law of conservation of action (energy in constant time intervals). Furthermore, due to the equivalence of action and information and the dependence of this on probability, it is possible to report the laws of conservation of both information and probability. The last mentioned law may be a novelty for many, but it should not be the previous one either.

## Final Divisibility

The law of information conservation can be reached in other ways as well. Intuitively, believing in what we get from the measuring apparatus by experiment, we accept that information does not arise or disappear from nothing and that it can be transmitted unchanged. Admittedly, what we read as a result of measurement should be deciphered, but let's add that information, similar to energy, can change its content but not its quantity.

The former paradox of "devouring information" of the universe by "black holes" has been resolved by observing that a body falling towards the horizon of a black hole event flows more slowly and that its radial lengths become shorter and shorter tending to complete disappearance. It is so gravitationally attracted that from the point of view of the relative observer it is never inside. The information of the decaying body only spreads over the surface of the sphere of the event horizon, which is closer to us, and thus remains external forever.

At the other end of the magnitude are quantum evolutions (processes). We present them to unitary operators. Otherwise, such are limited surjections (functions over the whole codomain) that preserve the inner product, and especially in quantum mechanics, the unitary is called the linear operator $U$ whose inverse $U^{-1}$ is adjoined to it $U^*$. We write this:

$$U^{-1} = U^*,\ UU^* = U^*U = I.$$

We can consider them as generalizations of complex numbers whose modulus (absolute value) is one. The unitary operator preserves the "lengths" and "angles" between the vectors (which it maps) and can be considered a type of rotation operator in the abstract vector space. Like Hermitian operators, the eigenvectors of unitarians are orthogonal. However, its own values are not necessarily observable (real), physically measurable quantities.

In addition to the well-known mentioned, I emphasize that from the same properties of unitary operators follows the property of reversibility of quantum evolution, and hence the law of conservation of information. Quantum processes "remember" well, and that is the kind of symmetry from which the law of conservation of information follows, now according to Noether's theorem[5] (where there is symmetry, there is also the law of conservation).

[5] see [1], 1.14 Emmy Noether

What is less known in the mentioned is the conclusion that information is always finally divisible[6]. That's why we have the smallest effects, in quanta. For example, therefore, the proofs of theorems are necessarily in discrete steps, as are legal regulations.

Information is a quantity (of uncertainty), so it makes sense to talk about less of it all the way to the bottom, to some of its atomized values. When we reduce this multitude of random events (such as throwing coins, dice, etc.) to less and less and break it down into elementary experiments, the question of continuation remains. In terms of information, the smallest experiment is in some of its optimum, because the opposite of certainty is uncertainty and vice versa, by reducing uncertainty, certainty is created.

## Types of Energy

We know about chemical, thermal, nuclear and many other types of energy. The diversity of information is multiplied by them, the ways in which energies pass into each other, and then the durations of such changes. For now, what is important to us are only two types of energy: kinetic and potential. Analogously, these are two types of information: active and passive.

Potential energy, it is said, has an object because of its position in relation to other objects, stresses in itself, its charge in the force field, or other factors. Its spontaneous action after the release of the object indicates to us the general aspiration of nature. Therefore, states with zero potential can be taken as those outside voltage, aspirations caused by potential energy, and the potential determined by a negative quantity. Thus deficient states of potential are attractive.

For example, bodies of masses $M$ and $m$ are attracted to each other by the gravitational force $F = GMm/r^2$ when their centers are at a distance $r$, where $G \approx 6{,}67 \times 10^{-11}$ N m2 kg-2 is gravitational constant. Gravitational potential

$$V(\mathbf{r}) = \frac{W}{m} = \int_{\infty}^{r} \mathbf{F} \cdot d\mathbf{r} = -\frac{GM}{r}$$

at a distance $r$ from the center of the mass $M$ can be defined as the work $V = V(\mathbf{r})$ obtained by moving a unit of mass from infinity to a given point. In other words, the higher (absolute value) of the potential energy is more attractive.

In the appendix [3], the constant total energy of the planet (in motion around the Sun) is considered, from which the conclusion about its constant communication with gravity is drawn. The minimalism of such information was proved in the book [4], based on the evidence that geodesic lines of motion in a gravitational field are also trajectories that satisfy the principle of least action. Here, we further emphasize this with the thesis that physical systems tend to have more negative potentials, ie states of absolutely greater potential.

[6] because only infinity can be its own proper part

Due to the law of information conservation, the planet remains on its (elliptical) orbit around the Sun, and the direction of the force[7] is turned towards (absolutely) greater potential. The potential of other forces can be defined analogously and the general conclusion can be reported that passive information is more attractive.

An example of passive information is a lie, untruth. Namely, by transforming (table of values) the relations of the algebra of logic, correct values into incorrect ones and vice versa, tautologies would be transformed into contradictions, from which the conclusion that “the world of truth” is equivalent to “the world of lies”. A lie is a hidden truth. That is why lies are easier to spread on social networks, that is why we read fiction more easily than geometric proofs, because nature prefers passive information.

## Epilogue

This is a detail of my stories about information not far from energy and that is why there is no reference to logarithms, or consideration of encoding and decoding, the former easier than the latter. The information also concerns the difference between the inanimate and the living world, because the one (the second mentioned) who has it in excess has a greater ability to choose. Due to the importance of decision-making, it is the future topic of game theory, and due to options and Everett's “many worlds” of quantum mechanics.

[7] the classical notion of “force” needs to be redefined

# 5. Parallel Realities

Independence from local observers

October 18, 2020.

Objective coincidence leads us to the strange world of the multiverse, but also to an easier understanding of the absurdities of known physics.

## Introduction

The assumption of the objectivity of chance applies. Uncertainty is such a real phenomenon that an event *C* can be perceived differently from observers *A* and *B*. The term “perceived” is freely changed to “measured” or “interacts with”. Moreover, all cases of “perception” are in fact communications, or one-way transmission of messages, because we believe that space, time and matter consist only of information, especially a defined amount of messages, and that the essence of information is uncertainty.

## Dimensions

If a space is of dimension $n = 0,1,2, \ldots$, then the finite set of such spaces is of dimension $n$. When with it we can divide some other, larger space into subspaces, each of dimension $n$, let's say isolate the so-called interior from the outside, other space, superspace, dimensions is $n + 1$; and if we can't, the superspace is larger dimension than $n + 1$. This is an inductive topological definition of dimension.

For example, a point is dimension 0. With two points at the ends of the interval we can isolate a given line interval from the outside, so the line is dimension 1. With a finite number of points we cannot close the surface area, which means that the surface is dimension 2 or larger. However, with a closed line we separate its part, its interior from the exterior, so the dimension of the surface is exactly 2. With a closed surface (sphere) we isolate the interior from the exterior of space and therefore the physical space is of dimension 3.

The special theory of relativity considers systems in inertial motion with constant velocity $v$. The proper, self-observer is the one who rests in the system $S$, and the relative is the one who sees such a person in motion as $S'$. Both systems are inertial (they do not feel acceleration). Their geometry is Minkowski space-time with the main coordinate planes: the abscissas that coincide and are parallel to the direction of motion, their own $x$-axes and relative $x'$-axes, and the ordinates of the time axes $x_4 = ict$, ie $x'_4 = ict'$. The imaginary unit is also $i^2 = -1$, the speed of light in vacuum is $c \approx 300\ 000$ km/s, and $t$ (or $t'$) is the proper (relative) time of the system.

From the point of view of a relative observer, one's own moves along the abscissa with the speed $v = x/t$, and the relative with respect to the own moves with the speed $v' = -v$. These two movements are equal. I repeat the well-known things of this kinematics so that there would be no confusion in the continuation at the point of the story. Proportionate to the so-called Lorentz coefficient

$$\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

the relative abscissa shortens ( $\Delta x' = \Delta x / \gamma$ ),and stretches relative time ( $\Delta t' = \Delta t \cdot \gamma$ ).

In a three-dimensional space, the Minkowski space-time coordinates are $x_1 = x$, $x_2 = y$ and $x_3 = z$, with notations from Cartesian rectangular coordinate system $Oxyz$. When the system moves only on the abscissa, then there are no changes (shortening, contraction) of the other two coordinates, so observation in the mentioned plane is often enough to understand the special theory of relativity.

We know that the notion of simultaneity is not the same for two observers, proper (own) and relative. At the moment “now” what is “here” will be some 3D system of spatial coordinates, but it will not be the same for these two observations. Nevertheless, 3D space is able to divide Minkowski's space-time into two 3D parts, its past and future, which according to the inductive definition of dimension means that each of the observers belongs to some space-time of dimension 4.

The problem with the deterministic conception of the theory of relativity arises when we have one fixed observer and two or more of proper in movements in different directions. As relative time axes (from the point of view of the fixed one) must be represented by leaning towards the direction of movement, the higher the speed, the case of at least two directions of movement (two proper) cannot fit all the necessary space-time in 4D.

In physics, the fact that Minkowski's geometry does not exist in the case of several uniform inertial motions is overlooked, or tacitly ignored. However, three time dimensions, i.e. 6D space-time would be a sufficient framework for all such 3D present.

We come to this unusual conclusion in the general theory of relativity. Due to the impossibility of defining any “now” in the general gravitational field for any observer, which would be 3D space and separate the space-time of the field into the past and the future, it has at least 6 dimensions. The simplest case of a gravitational field is centrally symmetrical, as the Moon, Earth and Sun have approximately, and each other can be obtained by a final union of such, so then each is of dimension 6 or more.

Imagine that at the points of a centrally symmetric gravitational field there are some small (infinitesimal) systems of Minkowski coordinates, with abscissas directed towards the center. The relative observer is fixed somewhere in the distance. He observes the shortening of the abscissa and the inclination of a certain time axis towards the abscissa, which is greater the stronger the field. 6D space-time is just enough to accommodate all of them.

We get a similar conclusion by referring to (my) information theory. We use the uncertainty and definition of an event with four coordinates, that something happened in a given place at a given moment. If the event was perceived differently by different observers and happened for one, then it did not happen for the other. This means that with the help of 4 dimensions of space-time (three spatial and one temporal) it is not possible to separate the reality, say into the past and the future. Therefore, the reality is at least six-dimensional[8].

## Effects

The basic framework for scales of physics quantities are Heisenberg uncertainty relations:

$$\Delta p \cdot \Delta x \geq h/4\pi, \quad \Delta E \cdot \Delta t \geq h/4\pi,$$

where $\Delta p$ and $\Delta x$ are the uncertainties of momentum and position (along the same axis) of the particle, and $\Delta E$ and $\Delta t$ are the uncertainties of its energy and time. The magnitude of the Planck constant, $h \approx 6{,}626 \times 10^{-34}$ J ⋅ s, determines the range of these relations in the world of magnitude of physics. Behaviors of action (products of momentum and position, or energy and time), ie. information, are different in the micro and macro worlds because of them, but also because of some other laws that are not the topic here.

Because there is no clearly defined position in the physical micro-world, Fourier's approximations will apply. We know how the Fourier series divides the periodic function into the sum of sines and cosines, presenting it more and more accurately by taking more summands, but also that the generalization of the Fourier method does the same with other functions instead of trigonometric ones. We add here that these theorems are consistent with the relations of uncertainty and universality of information.

In previous works, I pointed out that the possibility of approximating a function with fragments of various functions indicates the non-existence of “shape” in the micro world, at least not as clear as we imagine it in the macro world.

Similar follows from a recent appendix [6]. The quantum process, the unitary operator, can be broken down in countless ways into factors, always into the products of unitary operators, even when one of them is set in advance (previously determined). The conclusion is now obvious, I hope.

This means that the famous “double slit” experiment can indeed be explained in the way Everett (1957) did with “many worlds” of quantum mechanics. The particle-wave in front of two slits can be realized in any of the two passages, in one it is in our reality and in the other it is in parallel. However, in parallel reality, it can do the same, so in our reality, these two particles can appear, which will then interfere after two slits.

It is clear that the thought experiment “Wigner's friend” is also in line with this theory. Observer *A* observes another observer *B* performing a quantum measurement on physical system *C*, and the two then make different statements about what they saw. The two observations of system *C* may be different.

An appropriate interpretation of the also famous thought experiment “Schrödinger's cat” could have the following sequel. Event *C* is a cat in a box, and with that box in a larger box is Observer *A*. Outside both is

[8] Slightly different proof is in the book [8].

Observer *B*. After an accidental event in a smaller box *C*, the cat remained alive or became dead, and this is seen by *A*. However, this is not must see *B* in the same way.

Namely, if quantum event *C* were always perceived by different observers as one and the same event, then that event would not have uncertainty. We could deceive (overcome) Heisenberg's relations of uncertainty, say by taking the mean value of several observers and, reducing the uncertainty to a classical measurement error, obtain an estimate that could be used indefinitely. But that would then be in contradiction with the algebra of operators that are not otherwise commutative, and from which this uncertainty follows..

It is understood that in this “information theory” experiments such as [7], the local independence of the observer, may prove successful.

## Epilogue

We have seen the detail of the success of information theories in explaining the “strange” phenomena of quantum mechanics and connecting its smallest method with the rest of the known world of physics. Also, that the price of using the new method is to expand the reality of physics.

## 6. Page Evaluation

One application of information perception

30. окобар 2020.

### Introduction

Intelligence $a$ for solving a given problem is defined as a quantity proportional to the amount of options, that is the freedom $s$, which the subject can perceive with the problem and inversely proportional to the constraints $b$. Hence the freedom of the given, $k$-th ($k = 1,2,\dots,n$) temptation $s_k = a_k \cdot b_k$, and the total freedom $S = s_1 + \dots + s_n$. It is the information of perception (book [8]), that is, the vitality of the subject, which we then transfer to inanimate matter.

### Chess tournament

Let's watch a chess competition on several boards, between two teams of players with a rating (evaluation). Let the first team on the boards (first, second,..., $n$-th) have players with descending ratings $a_1$, $a_2$, ..., $a_n$ and the second team has players with ratings $b_1$, $b_2$, ..., $b_n$ which can and do not have to be a series of declining quantities.

If the second n-tuple is also decreasing, then the sum of the products, $S = a_1b_1 + \dots + a_nb_n$, perception information, or tournament vitality, is maximal. That sum of $S$ will be smaller if the second $n$-tuple is not decreasing, so if the other team does not place a stronger player on the stronger board and a weaker player on the weaker one. It is a perception information theorem, which here speaks to the strength of the tournament.

It can happen that the second team is very bad (compared to the first), as much so that at least one victory would be good for them. Then it is better for them to put the best player on the worst board, knowing that they will be the worst player of the first team there, but that lowers the value of the tournament.

### Scientific rating

The situation is similar with the ratings for which scientific journals and authors are fighting. Each author has some references of his own, the literature he lists at the end of his contribution. Authors and (collectively) references have their own rating, and then the journal itself has a rating. A publication in a journal with a higher rating gives the author a higher rating as well as his appearance in the reference of another author. If you are in the reference of an author with a higher rating, you get more rating points. In short, these are well-known rules that now remind us of the rating of a chess tournament..

Here, if authors with a decreasing rating of $a_1$, $a_2$, ..., $a_n$ were published in a given journal, and they had references in their articles with also a decreasing rating, in the order of $b_1$, $b_2$, ..., $b_n$, then the “information of perception” is maximal $S = a_1b_1 + \dots + a_nb_n$ for the journal. However, when better authors refer to worse, or worse to better, it reduces the journal's score.

It is natural then that the magazine will try to be selective in the fight for a better position on the top list of the "most respected", and that then goes in favor of the so-called "brotherly" system where he "pushes his own". Note that the editor of the magazine does not have to be biased or bribed by authors or clans, nor it is nepotism, but simple professionalism. The editor is a successful character in the fight for the success of the company he represents.

I note that the information of perception only takes over the rating scoring, and that in turn has shortcomings. The given system is "mediocre" on some level. Sooner or later, an outsider appears, such as George Bull[9] almost two centuries ago, the author of the algebra of logic [10], who is ahead of his contemporaries, incomprehensible and unaccompanied by them. The rating system values such as "bad", and he has a value above many others.

## Epilogue

This is a short story about the application of the PageRrank algorithm [11], which made Google search famous. I avoided the well-known descriptions of stronger scoring of more frequently opened web pages, or opened page by more valuable pages, considering it familiar. The description of the chess tournament, or the scientific weight of modern magazines, is also not the central theme of this story – as much as it is information of perception.

[9] George Boole (1815-1864), English mathematician.

# 7. Inner Product

About information perception

November 3, 2020.

I answer a frequently asked question: why did you define *perception information* as [8] the sum of the products of corresponding pairs of opposite quantities. Here, of course, is only part of the answer.

## Abilities and limitations

The idea of information perception arose in part by observing the intelligence of a species. On an individual $\omega_k$ element of perception ($k = 1,2, \dots, n$) from the set $\Omega$ of all elements available to it (species or individuals), the intelligence $I_k = I(\omega_k)$ would be proportional to the number of options related to the senses of a given individual, ie some ability of its perception, or freedom $S_k$ and inversely proportional to the external obstacles $H_k$, the limitations there called the hierarchy. Hence $I_k = S_k/H_k$, and hence $S_k = I_k \cdot H_k$. This is a property of one, $\omega_k$ element of perception.

The total perception of all individual elements $\omega_k \in \Omega$ would be $S = S_1 + S_2 + \dots + S_n$. The number $n = 1,2,3, \dots$ of all elements of observation may be very large, but it is always finite. The final divisibility of perception, discretion (from the bottom), or quantization of perception, follows from the law of conservation of physical information[10].

Information is the amount of various options such as the outcomes of tossing coins, dice and the like. Because it is quantity, we can shred (chop) information, but because it is discreet we have to get to the smallest portions. These minimum amounts are some optimums, here’s why.

By reducing uncertainty, certainty grows, and vice versa, by losing certainty uncertainty is gained. Looking at it that way, beneath the ultimate uncertainty are new certainties. This is not a contradiction, because uncertainty is the essence of the choice of information, but on the contrary it is in line with the quantization of information. The parts, the smallest information, do not have the independence to such an extent that we can consider them unreal in the classical sense of reality.

However, in addition to “reality”, the mentioned “unreal” parts can also affect physical phenomena. Aware of it or not, quantum mechanics got to know both types of these and encompassed that multiplicity by using complex numbers. It has been postulated that only phenomena are observable (physically measurable) when the complex expressions associated with them are real. In that direction, we reduce the information of perception from living to non-living beings. Vitality and uncertainty thus become shapes of the same form.

## Unitary space

In the unitary space, it is possible to define a linear mapping, by the so-called unitary operators, so that perception information remains constant. These are spaces that are called Hilbert's in physics, and

[10] Infinitely divisible sets can be their real subsets, and for such we cannot introduce the law of conservation.

whose representation is quantum mechanics. These two reasons alone are sufficient for the great importance of unitary spaces in information theory.

A vector space $\mathbb{X}$ over the field $\mathbb{C}$ of complex numbers is given, with a scalar product, the so-called inner product of the vectors

$$\langle x|y\rangle = x_1 y_1^* + x_2 y_2^* + \dots + x_n y_n^*.$$

In this case, $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_n)$ are arbitrary vectors from $\mathbb{X}$ with coefficients with complex numbers, $x_k, y_k \in \mathbb{C}$, and $y_k^* \in \mathbb{C}$ is conjugated $y_k$. If we have a unitary space, then for each of the vectors and each of the scalars ($\forall x, y, z \in \mathbb{X}$ и $\forall \lambda \in \mathbb{C}$) is:

1) $\langle x|y\rangle = \langle y|x\rangle^*$;

2) $\langle \lambda x|y\rangle = \lambda\langle x|y\rangle$;

3) $\langle x + y|z\rangle = \langle x|z\rangle + \langle y|z\rangle$;

4) $\langle x|x\rangle \geq 0$, where the equal sign is valid for $x = 0$.

The intensity of the vector is defined by $\|x\| = \sqrt{\langle x|x\rangle}$, and the distance between the vectors (points) $x$ and $y$ with:

$$d(x, y) = \|x - y\| = \sqrt{\langle x - y|x - y\rangle}.$$

It is easy to check that this is a function of the metric, because it is:

$$d(x, y) = 0 \Leftrightarrow x = y,\ d(x, y) = d(y, x),\ d(x, y) + d(y, z) \leq d(x, z).$$

The last, the inequality of the triangle, is a consequence of the Cauchy-Schwartz inequality which claims:

$$|\langle x|y\rangle| \leq \|x\| \cdot \|y\|,$$

where the equal sign is valid when $x$ and $y$ are linearly dependent. A complete space with a metric defined in this way is Hilbert space.

Unitary space does not have to be of finite dimensions and, as in Euclid, can use the concept of orthogonality and have an orthonormal system of vectors. In the case of a finite-dimensional space, the existence of an orthonormal base can be proved.

The unitary operator $U$ is a bounded linear operator in Hilbert space for which the equations $U^*U = UU^* = I$ apply, where $U^*$ is an adjunct (transposed and conjugate) $U$, and $I$ is an identical operator. Therefore, domain $U$ is dense in Hilbert (unitary) space and it preserves the internal (scalar) product of the vector, $\langle Ux|Uy\rangle = \langle x|y\rangle$.

## Epilogue

The purpose of this short story is to be used in discussions about (my) theory of information, that is, information of perception. I am tired of repeating the basic concepts of that theory too often, which are shown to be too far from official science, so this “introduction” to those stories is welcome.

# 8. Central Movement

About information perception

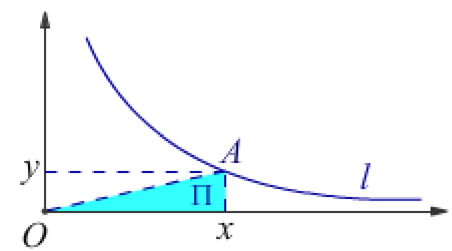
November 8, 2020.

I enclose a few more new but similar to previous interpretations of information by surface, that is, ways of establishing the equivalence of information, surface and physical action.

## Introduction

This is again a slightly different example of “perception information” which, I hope, avoids boring repetitions but leaves sufficiently recognizable similarities with the previous ones.

In the picture on the right, in Descartes' rectangular coordinate system $Oxy$, a curved line $l$ (hyperbola) and a point $A(x,y)$ are given. The area $\Pi$ of the triangle $AOx$ is constant, it does not depend on the choice of the point $A \in l$ as long as $xy =$ const. Let's look further at what this has to do with the information of perception, that is, with freedom or vitality.




Imagine a subject in a situation $\omega \in \Omega$, an element of a set of similar situations $\Omega$. Let it be a problem whose weight, the notation $y$, depends on the possibility of solving $x$. It is clear that $x$ and $y$ are certain numbers and such that they are in the function $x \to y$ and that the equality $xy = s$ applies to them. This is a hyperbola from the picture, which means $s =$ const. This $s$ defines the freedom of the subject in a given situation.

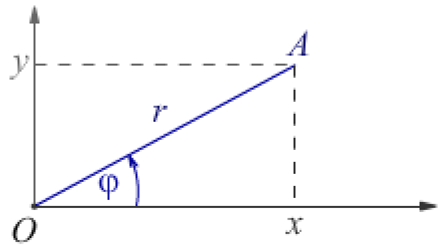
In general, when there are several elements $\omega_1, \omega_2, \dots, \omega_n \in \Omega$ that bring the same subject into similar situations, then we define the total freedom $S = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$. It is information of perception and it is again of dimension two, but no more in the previous way (shown in the picture) than in the next one.

Let us represent the given sequences by the vectors $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and $\mathbf{y} = (y_1, y_2, \dots, y_n)$ of a vector space of dimension $n = 1,2,3,\dots$, and consider them as oriented longer with a common beginning, starting from $O$. They then span the parallelogram in one plane which in turn, in its own way is equivalent to the information of perception.

In general, if the information is two-dimensional and if all space, time and matter consist only of information, then any natural phenomenon can be analogously decomposed on the surface.

## Lagrange equations

In the polar (plane) coordinates $Or\varphi$, in the picture on the left, the point $A(r,\varphi)$ is given. In that picture




we also see the position of that point, then we write $A(x, y)$, in relation to the corresponding Cartesian rectangular coordinate system $Oxy$. From the right triangle $AOx$ we can easily recognize the equations of coordinate transformation:

$$x = r\cos\varphi, \quad y = r\sin\varphi.$$

Then we define the unit vector

$$\mathbf{r}_o = r(\cos\varphi, \sin\varphi),$$

norm $\|\mathbf{r}_o\| = 1$, whose derivative in time is

$$\dot{\mathbf{r}}_o = \dot{r}(\cos\varphi, \sin\varphi) + r\dot{\varphi}(-\sin\varphi, \cos\varphi),$$

where the unit vectors are mutually perpendicular, $\boldsymbol{\varphi}_o = (-\sin\varphi, \cos\varphi) \perp \mathbf{r}_o$. The derivative of the path in time is speed, this intensity

$$v = \|\dot{\mathbf{r}}_o\| = \sqrt{\dot{r}^2 + r^2\dot{\varphi}^2}.$$

Therefore, the expression for the kinetic energy $T = \frac{1}{2}mv^2$ in polar coordinates is

$$T = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\varphi}^2).$$

The force is $\mathbf{F} = f(r)\mathbf{r}_o$ with the center at the origin of the coordinate system. It is given by an arbitrary function of the force $f(r)$ depending only on the distance $r$ of point $A$ from the center $O$ and the direction $OA$. Potential energy can be understood as the work of force on the road, ie as a scalar product of the vector of force and road. Hence, infinitesimally:

$$dU = -\mathbf{F}\cdot d\mathbf{r} = -f(r)\mathbf{r}_o \cdot d\mathbf{r} = -f(r)dr,$$

$$U = -\int f(r)\,dr,$$

so that the Lagrange function $L = T - U$ took shape

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\varphi}^2) + \int f(r)\,dr.$$

Hence the Lagrange's equations [14]:

$$m\ddot{r} - mr\dot{\varphi}^2 = f(r),$$

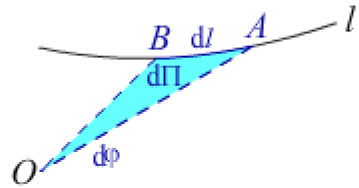
$$\frac{d}{dt}(mr^2\dot{\varphi}) = 0.$$

The second of these equations by integration gives equality

$$r^2\dot{\varphi} = 2\dot{\Pi},$$

from which we recognize the integral of the sector velocity, $S = \dot{\Pi}$. If the radius vector of a given point, $\mathbf{r} = \overrightarrow{OA}$, overwrites equal surfaces at equal times, analogous to Kepler's second law, then the derivative of the surface is constant over time $\dot{\Pi} = \text{const}$). The angular velocity is inversely proportional to the square of the distance of the material point from the origin ($\dot{\varphi} = \mathrm{C}/r^2$).

From the point of view of "information theory" (unofficial), the picture on the right shows the "perception" of the movement A → B of a material point. The point moves along the line l, in the figure for the infinitesimal dl, assuming that the cause of the movement is some central force from O.




The total motion information $dl$ is proportional to the area $d\Pi$, and this, ie both (information and area), are proportional to the action. This is in accordance with the well-known principle of least action, which is the basis for the use of Lagrangian in physics.

## Binet formula

Let's rewrite Lagrange's equations in the form:

$$\ddot{r} = r\dot{\varphi}^2 - f(r)/m, \quad r^2\dot{\varphi} = 2S,$$

then substitute the second ($\dot{\varphi} = 2S/r^2$) to the first. We get

$$\ddot{r} - \frac{4S^2}{r^3} = \frac{1}{m} f(r).$$

The solution of this equation is of the form $r(t)$, the distance from the center of force as a function of time. To find the trajectory, $r(\varphi)$— functions of the distance from the angle, we use the transformation:

$$\ddot{r} = \frac{d\dot{r}}{dt} = \frac{d}{d\varphi}\left(\frac{dr}{d\varphi}\frac{d\varphi}{dt}\right)\frac{d\varphi}{dt} = \frac{d}{d\varphi}\left(\frac{2S}{r^2}\frac{dr}{d\varphi}\right)\frac{2S}{r^2} = -\frac{4S^2}{r^2}\frac{d^2}{d\varphi^2}\left(\frac{1}{r}\right)$$

and write the previous one:

$$-\frac{4S^2}{r^2}\frac{d^2}{d\varphi^2}\left(\frac{1}{r}\right) - \frac{4S^2}{r^3} = \frac{f(r)}{m},$$

$$\frac{d^2}{d\varphi^2}\left(\frac{1}{r}\right) + \frac{1}{r} = -\frac{r^2 f(r)}{4mS^2}.$$

Binet's[11] formula, known differential equation of trajectory, is obtained [13].

## Conics

The intersection of the cone and the plane is an ellipse, parabola or hyperbola. If the plane is on all sides of the cone, the section is an ellipse, if it is parallel to some generatrix, we have a parabola, and if the intersection plane builds an even sharper angle with the axis of the cone, we get a hyperbola. The orbits of bodies in the solar system are often ellipses, but some comets travel in a parabola or hyperbola. The sun is the focus of such trajectories, so we have the typical central motion task for Binet’s formula.

The general equation of conic sections (conic) in polar coordinates with the focus at the origin is

[11] Philippe Binet (1786-1856), French mathematician.

$$r = \frac{ep}{1 \pm e\cos\varphi},$$

where the directrixes of the conic are $x = \pm p$, with a positive real number $p$ and with also a positive real eccentricity $e$. When $0 \le e < 1$ the conic is an ellipse, when $e = 1$ the conic is a parabola, and when $e > 1$ the conic is a hyperbola..

To search for the function of the central force, $f(r)$, due to which the bodies move along the conics, we calculate the second derivative of the reciprocal radius of the point on the conic and include it in Binet formula:

$$\frac{1}{r} = \frac{1 + e\cos\varphi}{ep},$$

$$\frac{d}{d\varphi}\frac{1}{r} = -\frac{\sin\varphi}{p},$$

$$\frac{d^2}{d\varphi^2}\frac{1}{r} = \frac{d}{d\varphi}\left(-\frac{\sin\varphi}{p}\right) = -\frac{\cos\varphi}{p},$$

$$\left(-\frac{1}{r} + \frac{1}{ep}\right) + \frac{1}{r} = -\frac{r^2 f(r)}{4mS^2},$$

because from the conic equation we have $-\cos\frac{\varphi}{p} = -\frac{1}{r} + \frac{1}{ep}$, so the central force is

$$f(r) = -\frac{4mS^2}{ep}\frac{1}{r^2}.$$

It is attractive ($m > 0$) and decreases with the square of the distance from the center. In the case of such a force, as we know, the central motion describes the trajectory of an ellipse, parabola, or hyperbola.

## Epilogue

It is well known in theoretical physics that bodies move in trajectories subjected to the principle of least action. This principle also speaks of constant action, and as demonstrated here, of constant area. The point of this story, however, is on the equivalence of that surface with action and information, and then even further on the principled minimalism of information.

## 9. Energy Leakage

November 16, 2020.

From the occasional conversations with colleagues about the hypotheses I deal with in my free time, there are interesting discussions of two issues that I will try to elaborate popularly here. The first question was what the eventual "leakage of energy" from gravity has to do with information theory, and the second would be why I bother so much with tensors and gravity. The latter comes from a man who does not believe in the general theory of relativity, and the first from a colleague who is skeptical of (my) information theory. I wrote it down because the answers might be interesting to many.

**Question 1.** What does the leakage of energy from gravity have to do with information theory? (It will be clarified later what is meant by "energy leakage" here.)

**Answer:** Various. For example, energy could "leak" into additional dimensions of time, and without these dimensions there is no uncertainty or information. Also, the alleged leakage is the bridge for the action of the past on the present, perhaps visible in the movement of the perihelion of Mercury in the direction of its movement, or the gravity of dark matter, but irrelevant with that, in the ability of the space to "remember". Namely, if the past did not somehow act physically – it would not be the information, that is, action – but the information is everything that the universe consists of.

**Question 2.** Why do you bother so much with tensors and gravity?

**Answer:** If you have concentration, I will try to put together that "puzzle" with a more or less popular story, without a tensor account.

First, photons move at the speed of light. That is why they do not have a mass of rest, and time stands still for them. They are the prisoners of 3D spaces, their own (proper), say, length, width and height. The electromagnetic force, of which they are carriers, is also closed in some three dimensions, always in its own present, and because we see with the help of light and we see only "now". Gravitons, on the other hand, penetrate the layers of time.

Second, in the section "8. Central motion" you will find proof that motion on conics (ellipses, parabolas, hyperbolas) caused by any central force (not only Coulomb's, or gravitational) means that this force decreases with the square of the distance. The planets move around the Sun in ellipses, but Mercury's ellipse (closest to the Sun and in the strongest gravity) slowly rotates in the direction of Mercury's orbit, which means that this orbit is no longer an ellipse. This means that it should be re-examined whether the strength of the strong gravitational force decreases also with the square of the distance.

Then, in an earlier work, it was proved that bosons (photons, gravitons, ...) that carry a field of (arbitrary central) force, move at the speed of light if and only if that force decreases with the square of the distance. Therefore, from the other, it follows that gravitons could (perhaps) move at a speed (at least a little) less than light, which would mean that time does not stand still, that they "see" more than one present, that is, that they can penetrate. through the layers of time. This would further mean that gravitons have a rest mass (even if it was $10^{23}$ times smaller[12] than today's easiest known particle – neutrinos).

This is only as the first confirmation of the above, because as you know, although I follow the news in science a bit, I do not follow a modern course.

The next, and that is the second confirmation of the above, is in the tensor account. The equations of gravity (general theories of relativity) are easily extended to 6D space-time. Additional dimensions are only temporal (three temporal to three spatial, see [1]) and from which we can take any four (three spatial and one temporal) and again obtain the correct Einstein equations. That's weird, isn't it. That's why it's attractive.

In other words, gravity penetrates both time and space, except that time dimensions are calibrated with “$ict$” (an imaginary unit multiplied by the path that light travels in a unit of time). The square of the “time path” is a huge real number, due to the square of the speed of light, so the penetration of gravity through the layers of time is significantly weaker – from the human point viewing the size.

Finally, it would be possible to add to the above proofs of ”energy leakage“ from the gravitational field one more, which is believed to have been first uttered publicly by Landau[13]. Telling it to Einstein, he allegedly replied as he had been knowing that, and that he ”sacrificed the law of conservation of energy in order to preserve the causality“ of order in space. Information theory, of course, does not adhere to causality, but still, energy leakage from the gravitational field cannot be avoided. It now takes on a deeper meaning: that with the help of gravitons we can peek into the parallel dimensions of the universe.

[12] Clifford Will at the University of Florida
[13] Lev Landau (1908-1968), Soviet physicist and Nobel laureate.

# 10. Central Movement II

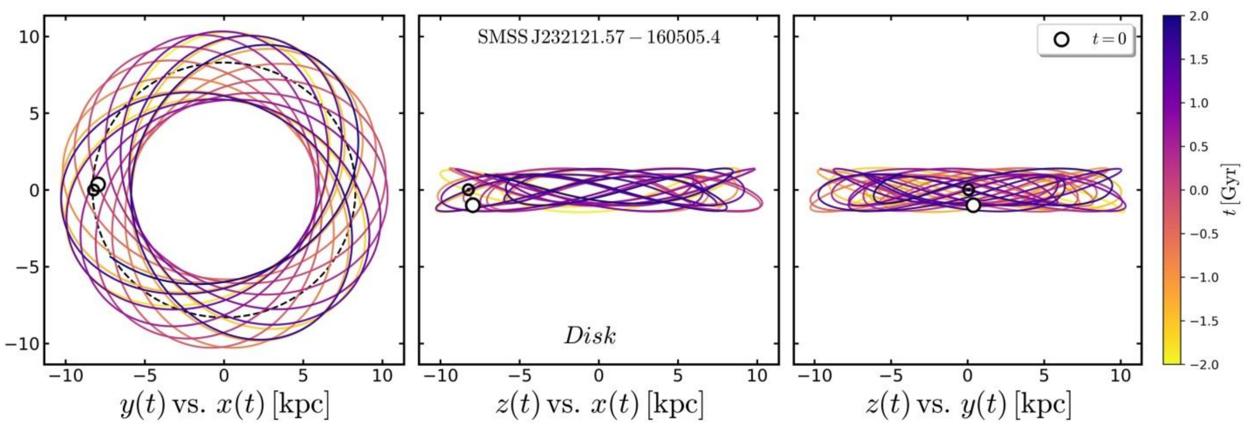
About trajectories due to central force

November 19, 2020.

I prove, the central force that decreases with the square of the distance move the charge along the conics. This is the inverse application proved in the previous application[14], that the conic as a charge trajectory is created by a central force that decreases with the square of the distance.

## Introduction

Are these unexpected orbits, partly from the workbook [14], proof that the gravitational force does not decrease with the square of the distance? That question was asked to me later in the conversation[15]. It doesn't have to be, I replied. In the introduction, I will try to paraphrase the reasons, followed by one piece of evidence.




Conic trajectories (eg ellipses) prove that the central force decreases with the square of the distance (app 8), but only if there are no other forces nearby. These oscillating orbits of metal-poor galactic stars are in the company of other bodies that also act on them gravitationally and which introduce trajectory disturbances. However, there is something “strange” in their movement that gives us additional information.

Mercury is the planet closest to the Sun and has the most visible deviation from the elliptical orbit, such that its perihelion (the major axis of Mercury's ellipse) moves slightly in the direction of orbiting, with each revolution. This was one of the four “great proofs” of Einstein's general theory of relativity, which successfully confirmed[16] (predicted) this shift.

[14] see 8. Central Movement
[15] see 9. Energy Leakage
[16] the phenomenon has been known before

In (my) information theory, the same movement of the perihelion will also serve as proof that "space remembers" and that its past has a gravitational effect on our present. Namely, the front part of Mercury's orbit from its previous rotation is closer in time to the present and it acts a little stronger than the back part.

Distance through time ($t$) is measured by the path ($x_4 = ict$) that light passes in a given time, and this becomes a very large value by squaring (due to the high speed of light, $c \approx 300\,000$ km/s), so the penetration of gravitational force is very small but in the case of strong forces it is still visible. It can be shown that the shear explained in this way, calculated (by the tensor calculus of the general theory of relativity), corresponds very precisely to the observed one.

Further, the oscillation up and down around the elliptical orbit of a star traveling (billions of years) through the galaxy, with a massive black hole in one of the foci, can also be explained by the "gravitational memory of space."

The stars and other substances of the galaxy orbit around the center in (approximately) one plane, but in the classical way it is difficult to understand the oscillations of the body around that plane, because in the lateral (perpendicular to the plane) attractions there seems to be a lack of substance for the exact account. However, the deficit of "lateral" attraction, which returns the star to the plane of the ellipse, corresponds to the deficit of the mass of the galaxy, without which the matter of the galaxy would explode and the galaxy would disintegrate. In the case of both deficits, the hypothesis of "dark matter" helps, although only the latter was its initiator.

Information theory has an explanation of dark matter with the help of the "space that remembers" and those memories that gravitate to the present. Although the explanation seems to be independent of the decrease in gravitational force with the square of the distance, I will supplement the proof of central motion with the reverse implication, because I have received numerous questions in which such doubt is possible at all, but also because I could refer to this proof in some next discussion.

## Differential equations

In the mentioned previous article on central motion, it was proved that the central force decreases with the square of the distance – if the charge moves along conices (ellipse, parabola, hyperbola). The proof in the other direction of the implication, that due to the central force decreasing with the square of the distance the charge would have to move along conical trajectories, I will demonstrate in a few steps. First, if $r^2 f(r) = \text{const}$, then the Binet's equation, from which the first implication is derived, can be written in the form

$$y'' + y = C, \tag{1}$$

where is the constant $C = -r^2 f(r)/4mS^2$, and $y(x) = 1/r$, so $r = r(x)$ is the distance of the material point (charge) from the center of the force. These are the polar coordinates with the force at the origin. Binet's formula thus becomes a linear differential equation of the second order

$$y'' + ay' + by = C \tag{2}$$

with constant coefficients $a, b, C$.

When $C = 0$ it is a homogeneous linear differential equation of the second order with constant coefficients, we say corresponding to the previous one

$$y'' + ay' + by = 0. \tag{3}$$

The following principle of superposition applies to such. If both $y_1(x)$ and $y_2(x)$ are solutions of the homogeneous equations, then its solution is also $y(x) = C_1 y_1(x) + C_2 y_2(x)$, where $C_1$ and $C_2$ are arbitrary constants. This is easily proved by direct inclusion in the equation (see [16]).

The general solution of the homogeneous differential equation (3) follows from the basic solution represented by the exponential function, $y(x) = \exp(\lambda x)$, by successive application of superpositions with arbitrary constants, $C_1$, $C_2$ and $\lambda$. This is also easily checked by direct shift.

By substitution we get $(\lambda^2 + a\lambda + b)\exp(\lambda x) = 0$, then

$$\lambda^2 + a\lambda + b = 0. \tag{4}$$

This quadratic equation is the characteristic equation of homogeneous (3).

If the complex function $y(x) = u(x) + iv(x)$ of a real argument is a solution of homogeneous equation (3), then the real functions $u(x)$ and $v(x)$ are solutions of that equation also. That claim is also easy to verify. It further generalizes the solutions so that they are also real functions

$$y_1(x) = e^{\alpha x}\cos\beta x \ , \quad y_2(x) = e^{\alpha x}\sin\beta x,$$

the solutions of equation (3). As it is $y_1 / y_2 = \operatorname{tg}\beta x$, and $\beta \ne 0$ this quotient is not constant, the functions $y_1$ and $y_2$ are also linearly independent. We say that they form a fundamental system for solving equation (3). The general solution of that equation is as follows

$$y(x) = e^{\alpha x}(c_1\cos\beta x + c_2\sin\beta x). \tag{5}$$

**Example 1**. Homogeneous differential equation $y'' + y = 0$ has the characteristic equation $\lambda^2 + 1 = 0$ whose roots are $\lambda_{12} = \pm i$. To them correspond the solutions of the given differential equation $y_1 = \cos x$ and $y_2 = \sin x$, so the general solution is

$$y = C_1\cos x + C_2\sin x. \tag{6}$$

Note that (6) is the solution of the homogeneous part of the Binet's equation (1).

**Example 2**. Let $y_1$ and $y_2$ are solutions of the equation (2), the inhomogeneous linear differential equations of the second order with constant coefficients, at some interval. Then there is their difference $y_1 - y_2$ solution of the corresponding homogeneous equation (3) on the same interval. Namely, from

$$y_k'' + ay_k' + by_k = c \quad \text{за} \quad k \in \{1,2\},$$

by subtraction we get $(y_1 - y_2)'' + a(y_1 - y_2)' + b(y_1 - y_2) = 0$.

**Example 3**. If $\bar{y}$ is the solution of the non-homogeneous equation (2), and $\bar{\bar{y}}$ solution of the corresponding homogeneous (3), on some interval, then the sum $\bar{y} + \bar{\bar{y}}$ is the solution of the non-homogeneous equation (2) on the given interval. Namely, from

$$\overline{y''} + a\overline{y'} + b\bar{y} = c \quad \text{и} \quad \overline{\overline{y''}} + a\overline{\overline{y'}} + b\bar{\bar{y}} = 0$$

by adding we get $(\bar{y} + \bar{\bar{y}})'' + a(\bar{y} + \bar{\bar{y}})' + b(\bar{y} + \bar{\bar{y}}) = 0$.

These examples are well-known views of differential equations. We also know that the general solution of inhomogeneous equation (2) on some interval has form

$$y(x) = \bar{y}(x) + \bar{\bar{y}}(x), \tag{7}$$

where $\bar{y}(x)$ is arbitrary solution of the inhomogeneous equation (2), and $\bar{\bar{y}}(x)$ is some particular solution of the corresponding homogeneous equation (3) of that interval. The most well-known general way to obtain this particular solution is the Lagrange method of parameter variation, but in special cases such as ours, special methods are faster.

**Example 4**. Binet's differential equation (1) is inhomogeneous and obviously has a constant function $\bar{y}(x) = C$ for the particular solution. We have found (example 1) that (6) is a general solution of the corresponding homogeneous equation, which we will now denote $\bar{\bar{y}}(x)$. Therefore (example 3) the general solution is inhomogeneous

$$y(x) = C_1 \cos x + C_2 \sin x + C. \tag{8}$$

It is the (general) solution of Binet's differential equation (1).

We transform function (8) in steps:

$$y(x) = \sqrt{C_1^2 + C_2^2}\left(\frac{C_1}{\sqrt{C_1^2 + C_2^2}}\cos x + \frac{C_2}{\sqrt{C_1^2 + C_2^2}}\sin x\right) + C,$$

$$y(x) = \sqrt{C_1^2 + C_2^2}(\sin\gamma \cdot \cos x + \cos\gamma \cdot \sin x) + C,$$

$$y(x) = C_3 \sin(\gamma + x) + C, \tag{9}$$

where are the constants $C_3 = \sqrt{C_1^2 + C_2^2}$, $\sin\gamma = C_1 / C_3$ and $\cos\gamma = C_2 / C_3$.

As (9) represents the solution of Bennet's equation (1), where the function $y = 1/r$ is the reciprocal distance of the material point from the origin $O$ of the polar coordinate system $Or\varphi$ and the source of the force, and the argument is $x = \varphi$, we can write:

$$r = \frac{1}{C_3 \sin(\gamma + \varphi) + C} = \frac{ep}{1 \pm e\cos(\varphi + \phi)}. \tag{10}$$

This is the general conic equation (ellipses, parablolas, hoperballs) in polar coordinates $Or\varphi$. The eccentricity $e = \pm C_3 / C$ and the number that determines the directrix $p = 1/Ce$ are positive real constants, and $\gamma$ or $\phi$ are phase shifts, where the angle $\phi$ represents the slope of the conic axis towards the abscissa (direction $\varphi = 0$).

In other words, if the central force decreases with the square of the distance, then the conic is the trajectory along which the charge moves. This force can be as attractive as the solar gravity that moves the planets, or Coulomb electro force, but it can also be repulsive. This is the second direction of inclusion of the previous article on central motion where the assertion is proved: if the trajectory of the charge is conic then the central force decreases with the square of the distance.

# 11. Force and Information

Annex to information theory

November 22, 2020.

The time of a photon stands, it moves at the speed of light and defines the Coulomb force which decreases with the square of the distance. When the time of a particle flows, it sticks to the time dimensions and has mass, and the force it defines as a gauge boson does not decrease with the square of the distance.

## Introduction

Part of my earlier research is retold in the book "Information Stories" [1]. It is the primary reference here, which is a bit incorrect because the methods here are from a broader framework, but again, it is okay to read because it can be difficult to accept what I am writing about. You will easily understand that photons move at the speed of light[17] and that their time stands still, I hope, but it will be a little harder with two-dimensional information. It is even more difficult to absorb that space, time and matter consist only of information, of which the essence is uncertainty, and the action is equivalent.

Whatever, the relative time of a photon does not flow, so the motion of a photon from our point of view is through discrete (space-time) events of (our) perception. In ever newer layers of time, a spatial change occurs from the point of view of the observer and the new-old photon moves at the speed of light. It is at any moment in a new next position whose movement defines the speed of light. Each new one is the most probable outcome of the previous one, and then the least informative.

I recall, the glass is on the table because such a state of its is about the most probable in the given conditions. Like moving a glass by hand, new conditions are created by force and changes hide the causes. In other words, force changes probabilities, and changes in probability define force[18]. That is, all changes in motion or state, not just photons, are the causes and consequences of some forces.

The direction and intensity of the force are measures of the distribution of uncertainty (information) so that a greater change in uncertainty corresponds to a greater force. Thus says the principle of minimalism of information, that advantages in realization have greater probability and less information, and hence the law of inertia.

Thus, the photon moves in relation to the subject as its concrete information. The uncertainty of light is more direct in the oscillation in two planes, electro and magnetic, and the relativity in the differences of frequencies[19] and wavelengths of the same source, which we know from the Doppler effect[20].

## Coulomb force

---

[17] speed of light in vacuum c≈300 000 km/s

[18] see [1], 2.19 Classical force – figurative term

[19] see [1], 3.22 Light

[20] see [1], 2.16 Doppler effect

Each electron around itself emits virtual concentric spheres of virtual photons in continuous waves. They become real photons when they interact (communicate) with another electro-magnetic charge. But as the surface area of the sphere ($4r^2\pi$, radius $r$) increases, their amplitudes decrease proportionally, not the wavelengths, similar to water waves that propagate in concentric circles from the place where the stone falls on the water surface, with smaller amplitudes and unchanged wavelengths.

However, by decreasing the amplitude of the virtual sphere, the probability of its interaction and transfer of information from the starting electron to some other possibly present charge decreases. At the same time, the amount of transmitted momentum or energy does not change, we said, due to the constant wavelength. The rest of the process of transmitting information by a virtual photon can be considered in the manner of Feynman diagrams[21], now with a probability distribution.

Let $P_n$ be the probability that the virtual sphere in the n-th step (wavelength) interacts with some other charge. According to the previous one, $P_n = \chi/n^2$, where $\chi$ is a constant that we can determine since we have a probability distribution. Namely:

$$1 = P_1 + P_2 + P_3 + \cdots = \chi\left(\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots\right) = \chi \cdot \frac{\pi^2}{6}, \tag{1}$$

and from there $\chi = 6/\pi^2$.

In the following, it would be possible to join the energies of the transition of electrons from one shell of an atom to another with these probabilities, but let it remain for another occasion. For now, let's just notice that the virtual sphere is easier to tie to energy than the classic way of Fanman diagrams.

When we define the length by the path that light (photon) travels in a unit of time, because light is timeless and the observer's own (proper) time always flows at the same speed, the speed of light will not depend on the speed of its source. This is the situation in the geometry of relativity when the space-time interval is zero length, $ds = 0$. The probability of the interaction of a virtual sphere that is decreasing with the square of the distance will give the step of decreasing the force of interaction (here Coulomb, $kq_1q_2/r^2$). The change of state by interaction is equivalent to a force, and in the case of photons it is electromagnetic.

We know that photons are just one of a kind of gauge bosons, particle-waves that transmit fundamental interactions. If such time does not stand still and if it flows in the direction of the observer[22], its perceived speed must be less than the speed of light! This is intuitively clear from the fact that the particle encroaches on the present of the observer, which moves away in relation to the given moment, and at the same time moves less than the photon.

In other words, a particle which time flows with us moves at a lower speed than light, and then the probabilities of its interactions, analogous to (1), would decrease faster. The force whose field it defines

[21] see [1], 2.7 Feynman diagrams
[22] on opposite flaw later

would not decrease with the square of the distance. From the previous subheadings (8th and 10th about central motion) it can be seen that the charge trajectories would not be conical then.

## Mass

I mentioned earlier that mass is not a micro concept, and here I will repeat the reason for such an attitude derived from the properties of spin[23] (internal impulse). There are two types of complete elementary particles in physics, bosons and fermions. Among the first are the carriers of fundamental forces, such as photons and gravitons, and among the others are the particles that these forces act on. Bosons have an integer spin, photons $\pm 1$ and gravitons $\pm 2$, while fermions have a half-integer spin, say one of $\pm\frac{1}{2}$ such as electrons, protons or neutrons.

In the case of interaction, due to the law of conservation of the total spin, the photon of the spin $+1$ can leave the electron of spin $+\frac{1}{2}$ and leave it with spin $-\frac{1}{2}$ to switch to another electron spin $-\frac{1}{2}$ and translated it into a spin electron $+\frac{1}{2}$. Shorter written, +½ - 1 → -½. The opposite interaction of electrons and photons would also be possible: -½ + 1 → +½. However, an analogous interaction of graviton and electron would not be possible, because by adding any of the numbers $\pm 2$ with one of the two numbers $\pm\frac{1}{2}$, neither of the numbers $\pm\frac{1}{2}$ or $\pm 2$ is obtained.

Therefore, gravity is a property of a multitude of particles like a horizontal water wave formed as a separate entity from the vertical motion of its molecules. The mass of a particle is a similar property derived from its proper (own) time existence. From that stretching of the particles through the layers of time and the principled minimalism of the information, hence by duration, the particles gained inertia.

The greater the mass of the body, the greater the part of the body belonging to different present and the greater the restraint of the whole due to the influence (tendency not to act) of each of them. Resistance to action is proportional to the lack of total mass in a given present, or what is the same, is proportional to the excess of its presence in other presents (parallel realities). We derive this from the principle of minimalism of information.

The same change occurs as a relative observation of mass, or total body energy, in order:

$$m = m_0 \cdot \gamma, \quad E = E_0 \cdot \gamma, \tag{2}$$

during inertial uniformly rectilinear motion with velocity $v$, with coefficient

$$\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}, \tag{3}$$

or in the gravitational field of a body of mass $M$ at a distance $r$ from the center of force

[23] see [1], 3.27 Graviton

$$\gamma = \frac{1}{\sqrt{1-\frac{2GM}{rc^2}}}, \tag{4}$$

with the gravitational constant $G \approx 6{,}67 \times 10^{-11}$ m$^3$ kg$^{-1}$ s$^{-2}$. The first of the coefficients (3) is Lorentz from the special theory of relativity. It is better known than the equivalent other (4) defined here on the basis of the Schwarzschild metric. Otherwise, the connection between its proper mass and energy, $E_0 = m_0 \cdot c^2$, it is the same as between the relative ones, $E = m \cdot c^2$, so the formulas (2) are aligned.

The coefficients $\gamma$ are valid for the relative deceleration of time $t = t_0 \cdot \gamma$ of a body moving with speed $v$, or in another case in a gravitational field. Also, they are valid for shortening units of length $r = r/\gamma$ in the direction of movement, ie in the direction of force. Derivations of all these relations are well known, as well as the calculation of the gravitational force of centrally symmetric fields not stronger than the Sun together with (4). So I skip that part.

However, the novelty is that the same coefficient can be obtained from the assumption of the existence of parallel realities[24], ie that they are the cause first of the inertia of mass, and then of gravity[25] itself. Of particular interest is the connection between Shannon's channel noise and the force of gravitational attraction[26], which further connects the principle of minimalism of information with gravitational attraction.

[24] see [1], 3.24 Many worlds
[25] see [1], 3.23 Gravity
[26] see [1], 3.26 Channel noise

# 12. Negative Information

A discussion of minimalism

November 26, 2020.

This is another treatment of the idea of linking information with the algebra of logic, with physical action, potential energy and the principle of information minimalism..

## Introduction

Space, time and matter, as well as all the events of the universe, consist only of information. This is the working assumption of (my) information theory. Hence, and then from the impossibility of proving the inaccuracy of physical phenomena, it follows that we must consider information to be true. The kinds of truths are, therefore, "lies". This is a story about the opposite of the truth, which is not untrue in the classical sense, and in the search for such in physical action.

A lie is covert information[27]. The reflection of truth in its dual image, which we call untruth, is masked, that is, a more complex or higher level of truth. A lie is a place where information can escape into better coded forms, and it does due to its principled minimalism, but by no means it becomes something so inaccurate that it cannot exist.

Decoding lies is demonstrated by logical puzzles. For example, a passenger reaches an intersection for places *A* and *B* where he encounters one of two brothers who he knows one always telling the truth and the other always lying. The passenger does not know which of the two it is and can ask only one binary question to which strangers will answer with "yes" or "no" so that he finds out which of the roads leads to place *A*. What is that question?

The solution is for the traveler to show one of the unknown paths and ask, "Would your brother say this one leads to place A"? If the answer is affirmative, he will continue in the other way, and if he is negative, he will continue with what has just been shown.

That the "world of lies" is some equivalent to the "world of truth" in the amount of truth, can be proved by mapping the correct values into incorrect ones ($\top \to \bot$) and vice versa in the tables of logical operations (eg negations, disjunctions and conjunctions). Then the tautology (statement that is correct for all values of variables) is mapped into a contradiction (statement that is incorrect for all values of variables) and vice versa. And just delving into the "world of lies" to get the truth has long been a well-known maneuver in mathematics that we call proving by the method of contradiction.

However, misinformation is also a type of "negative information". We know that by misinforming, the truth is covered up, which means that it belongs to the "world of lies". But misinformation dilutes the truth, which indicates to us that "untruth" does not have to be something extremely negative, and only

[27] see [1], 3.7 Dualism Lies

that, the opposite of the positive, can be a state of lack and even excess of "truth". In the following, it becomes a question of form.

Making logic processor switches of a classic computer, so-called electrical circuits, gates or shutter, whatever you call them, does not require positive and negative voltage to represent the logic variables true (⊤) and false (⊥). The states "there is current" (digit, or letter 1) and "there is no current" (digit 0), or "higher voltage" and "lower voltage", and even something else are enough. This tells us that formal logic can be set up so that both values, true and false, are from different scales of intensity without losing or gaining in validity.

## Potential energy

In general, we say that potential energy ($U$) is the energy that is stored or is guarded in an object or substance and is based on position, arrangement or state. The two main types of potential energy are gravitational and elastic.

The object has a gravitational potential energy in a vertical position due to the attractive force of gravity. The amount of this energy ($U = mgh$) increases with the mass of the object ($m$), the gravitational acceleration ($g$) and the height ($h$) of the object. The elastic potential energy ($U = kx^2/2$) increases with the square of the length ($x$) of stretching or compression of the spring, trampoline or bungee cable (with different constants $k$).

The force on an object is conservative if the function $\mathbf{F}(\mathbf{r})$ is only of position $\mathbf{r}$. In the one-dimensional case, energy is required to operate the force on the path from $r_1$ to $r_2$

$$U(r) = -\int_{r_1}^{r_2} F(r)dr, \tag{1}$$

which can be taken as accumulated potential energy. As a result, any integration constant is added, which means that the magnitude of the zero potential energy can be chosen arbitrarily.

For example, for gravitational force, $F(r) = -GMm/r^2$, we find potential energy:

$$U(r) = -\int_{r}^{\infty} \left(-\frac{GMm}{r^2}\right) dr = -\frac{GMm}{r}, \tag{2}$$

where $G$ is the gravitational constant, $M$ is the mass of the planet, $m$ is the mass of the body, and $r$ is the distance from which the body is dragged (against the attractive force) to infinity.

Falling from infinity to the height $r$, the potential energy of a given body becomes kinetic $T = mv^2/2$, from which, by equalizing $T = U$, we find the initial velocity $v$ required for the body to leave the planet's surface. At small changes in height $h = r_2 - r_1$ the gravitational force ($F = mg$) is approximately constant and the potential energy is $U = mgh$, where $g$ is the gravitational acceleration.

These are well-known examples to which we now add an IT meaning. By transmitting energy, information is also transmitted – which is higher with more transmitted energy and longer transmission duration. Such physical information is equivalent to action (product of energy and time), so if we define potential energy by a negative value (in a positive flow of time) we have negative information.

From the definition of potential energy (1), the connection between information and force immediately emerges: the lack of information is attractive.

## Principle of information

It is easier to encode than to decode. Free media spread lies more and faster than the truth. Equality generates conflicts, because equally likely outcomes are more informative. Let's look at some more examples of the principled minimalism of information.

In free information transmission networks, concentrators, a small number of nodes with many links and numerous nodes with few links are formed spontaneously, because this reduces equally probable situations. That is why a small number of very rich people spontaneously appear on the free market (capital flow) compared to a large number of much less rich people, there are many less well-known people among public figures, such are the Internet clusters.

Dictatorship requires equality of the masses. The citizens of the Roman Republic aspired to this before the appearance of Caesar, equal before God for the establishment of the Inquisition, the people of the French Revolution for the Emperor Napoleon, the proletarians for the lifelong presidents of socialism. Conversely, it is also true that hierarchies grow more easily in conditions of equality, so more and more regulations, administration and lawyers are needed in the system of rights of principled equality.

The lack of information that could be the cause of gravitational pull can be understood in a variety of ways. For example, we know from the general theory of relativity that a stronger gravitational field slows down the passage of time, which we can now interpret as a relative lack of events, a lack of random outcomes, and a lack of information. The slower flow of time is attractive, therefore, because it is defined by the amount of (random) events, and that is the same one that determines the information.

The second cause of gravity (which is easily added to the first) is interpreted using Everett's idea of "many worlds" of quantum mechanics[28]. In short, the assumption would be that the relative "lack of time" (slowed down time flow) occurs due to the departure into the "parallel reality" of the part of the observed object within the gravitational field. The account shows that such a lack of information in principle could be the one just mentioned.

Third examples would go even further into the beginnings of (my) information theory[29], say, the "generalization of entropy". I would not repeat myself here further. It is enough that I have pointed out the connection between the concepts of negative information, the algebra of logic, potential energy and the principles of information, I hope.

[28] see [1], 3.24 Many worlds
[29] see [1], 2.24 Generalization of entropy

## 13. Increasing Time

November 29, 2020.

---------------------------------------------------------

A discourse of the need and consequences of the existence of time in information theory.

---------------------------------------------------------

In 1961, the German-American physicist Landauer[30] discussed the energy cost of communication. He assumed that information decreases when entropy increases and, given the spontaneous growth of entropy, he considered information to be a wasteful process. He was only a step away from noticing that physical information is equivalent to action (a product of energy and time).

When we store the essence of information in a molecule in a compartment, its deletion will be the removal of the molecule or memory. This move increases the entropy of the gas and the loss of heat to the environment. In isothermal process, the product of temperature and entropy change is the work you have to pay as an energy bill because you deleted the information.

Consistent with Landauer, as the law of conservation of energy applies, so will the law of conservation of information, which leads us to confirm in physics the well-known view that quantum evolution is reversible. As we know, quantum processes are representations of regular linear operators (for which there are inverse ones). This further means that we cannot extract information from a parallel reality[31] without depriving it of energy, or in other words, that its energy does not leak[32] to us.

Because indestructible information lasts and is remembered. When we take a closer look it is a paradox of the law of conservation of information. Due to the storage of information in the present and its accumulation in the past, we could have its overall increase, and thus a contradiction. But the information is less and less in the future by an amount equal to that which reaches from the past to the present[33]. Simply because the entropy (substance) grows spontaneously[34], the information is less and less.

Due to the law of conservation, information is quantized[35] and is in physical action. A quantum of energy always carries some information, just as every information produces some action. As repeated "news" is no longer real news, elementary information is proportional to both the exchanged energy and the elapsed time, that is, it is equivalent to action. This is a particularly interesting topic.

The quantum of action (Planck's constant) of the order of magnitude is a product of energy and time, Heisenberg's uncertainties. Consistently, when there is no penetration along the time axis, then there

[30] Rolf Landauer, 1927-1999
[31] see [1], 2.13 Space and time
[32] https://www.academia.edu/44517838/Energy_Leakage
[33] Note that the flow of information from the past must be slowed down, muffled.
[34] see [1], 3.25 Space memory
[35] By definition, infinitely divisible quantities are those that can be its own, proper part.

remains an immeasurably small moment as a hidden place of relatively large energies of, say, some of Everett's (1957) “many worlds” of quantum mechanics. If the change of time tends to zero then energy tends to infinity and the transition to an imaginary parallel reality in many ways becomes impossible. The multiverse could hide from us in such infinitesimal ways along the time axis.

Particles that move at the speed of light (photons) do not flow time. Such particles cannot know more than three dimensions, and we always observe them in the motion of one plane (electrical and magnetic polarizations). On the contrary, particles that do not move at the speed of light last, which means that they penetrate the layers of time and are present in parts in several “parallel realities”. They actually “get stuck” for layers of time and we say they have inertia.

Deeper, the inertia is a consequence of the principle of minimalism of communication. The world consists only of information kept by the law of conservation against nature that would have as little as possible of it. This principled stinginess becomes an “effort” when going into (new) uncertainties, and they again are the essence of information. That is why we consider the change of uncertainty as a “force”. I'll explain it one more time.

Until the light reaches from the end to the end of (our) body, some time passes for which the body does not exist in the same present. Random outcomes of processes at different times stretch the body unpredictably, creating new situations for random outcomes, and the change in the resulting probabilities is corrected only by forces. These forces arise from the principle of stinginess of communications, or the principle of least action, that is, the spontaneous growth of entropy.

Therefore, unlike light, gauge bosons (particles that carry the force field) that have a mass of rest will pass the “time barrier”, and the price of that kind of freedom is their inertia. Also, the central force they represent does not decrease with the square of the distance, nor do its charges move along the conics[36]. If gravity is that kind of force, then it can act from the past to the present.

Because space remembers it grows. Space, in addition to matter, also collects current information. It thus increases also when we talk about the same 3D space of a 6D space-time. It is never exactly the same for us participants of a 4D space-time (limited by the range of photons), and for all possible “parallel observers”, because space never makes the same memories. However, we have no reason to doubt in the conservation laws of physics in parts of these differences.

[36] https://www.academia.edu/44533846/CENTRAL_MOVEMENT_II

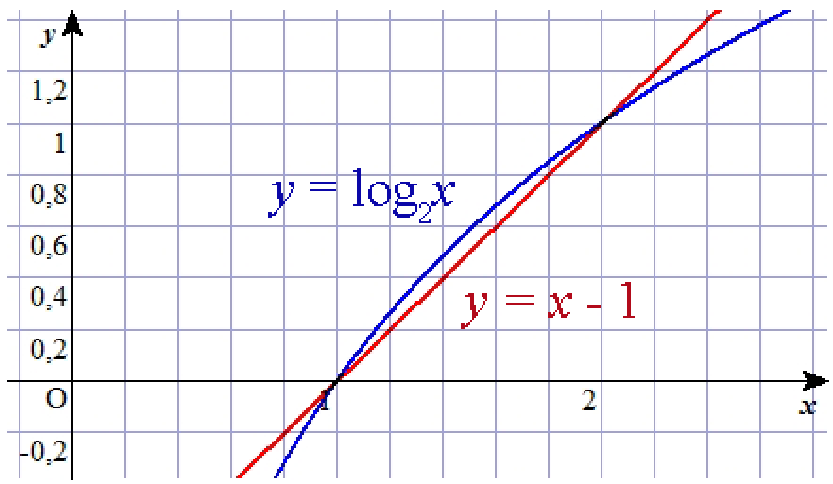
# 14. Uncertainty

December 5, 2020.

Discussion of relative energy and time in the context of probability and information. The first, which increases with vitality, the information of a given particle and decreases with probability, and the second (time), which disappears with increasing probability.

## Introduction

At the interval of probabilities, in the image enlarged by one, logarithmic function $y = \log_b x$ is approximately equal to the straight line $y = (x-1)/(b-1)$, as seen in the pictures for the case $b = 2$.




The difference between these two functions $f(x) = \log_b x - (x-1)/(b-1)$ reaches an extreme value at the point that null the derivative $f'(x_0) = \frac{\log_b e}{x_0} - \frac{1}{b-1} = 0$, when $x_0 = (b-1)\log_b e$. In the case of a given binary logarithm, the maximum is $f(x_0) = 0{,}0860713$ approximately, for $x_0 = \log_2 e = 1{,}4427$. From approximate equality $\log_2 x = x - 1$, for $x \in (1,2) \subset \mathbb{R}$, substituting $x = 1 - q$ we also get approximately $\log_2(1-q) = -q$, that is

$$q = -\log_2(1-q). \tag{1}$$

Number $q \in (0,1) \subset \mathbb{R}$ can represent the probability that some event $\omega$ will not happen, so $p = 1 - q$ is the probability that the same event will happen.

Let us look for the mathematical expectation (mean value) of the number of attempts to occur $\omega$. This can happen already in the first attempt with probability $p$, exactly in the second with probability $qp$, in the third $q^2p$, ..., and in the n-th with probability $q^{n-1}p$. The expectation $\mu = \mu(\omega)$ of the required number of repetitions of the experiment for $\omega$ to occur for the first time is (the derivative is by $q$):

$$\mu(\omega) = p + 2qp + 3q^2p + \dots + nq^{n-1}p + \dots$$

$$= p(1 + 2q + 3q^2 + \dots + nq^{n-1} + \dots$$

$$= p(q + q^2 + q^3 + \cdots + q^n + \cdots)'$$

$$= p\left(\frac{q}{1-q}\right)' = p \cdot \frac{1}{(1-q)^2} = \frac{p}{p^2},$$

$$\mu = \frac{1}{p}. \tag{2}$$

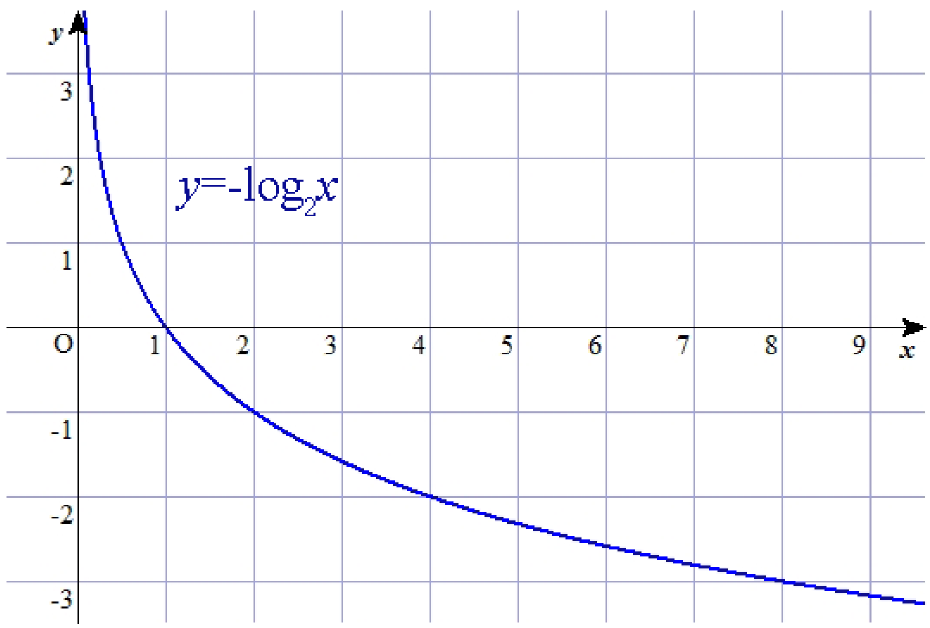
For example, when we toss a fair coin, the probability of the heads falling is $p = ½$ and it falls on average in the second toss, $\mu = 2$. When we roll a dice with six equal chances, the probability of one of them is one-sixth, $p = ⅙$, so we expect the desired outcome in approximately the sixth roll, $\mu = 6$.

## Information

In 1928, Hartley[37] defined information with

$$H(\omega) = -\log_2 p, \tag{3}$$

where $p = p(\omega)$ is the probability of a random event $\omega$. From the shape of the function, in the following figure, it can be seen that the information of the impossible event is infinite ($H \to \infty$ if $p \to 0$), and that it quickly falls to zero when the event becomes known ($H \to 0$ if $p \to 1$). We mostly still stick to these definitions.




In 1948, Shannon[38] used Hartleys information to calculate the mathematical expectation of the distribution of $n \in \mathbb{N}$ probabilities. When disjoint events $\omega_1, \omega_2, \dots, \omega_n$ form a complete set of outcomes (exactly one must occur), with probabilities $p_k = p(\omega_k)$, where $k \in \{1, 2, \dots, n\}$, then their Hartley information are $H_k = -\log_2 p_k$, and mathematical expectation

[37] Ralph Hartley (1888-1970), American researcher.
[38] Claude Shannon )1916-2001), American mathematician.

$$S = p_1 H_1 + p_2 H_2 + \cdots + p_n H_n. \tag{4}$$

If there is no difference in the chances, ie in the case of equal probabilities, $p_1 = p_2 = \cdots = p_n$, each of these amounts $p = \frac{1}{n}$, so Shannon's information becomes $S_0 = -n \cdot p \log_2 p = -\log_2 p$, then

$$S_0 = \log_2 n. \tag{5}$$

This is the mean value (mathematical expectation) of information $n$ equally probable events.

Note that (4) has the form

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n, \tag{6}$$

hence the scalar product of two vectors $\vec{a} = (a_1, a_2, \dots, a_n)$ and $\vec{b} = (b_1, b_2, \dots, b_n)$, which can also be written $\vec{a} \cdot \vec{b} = ab \cos\theta$. Here are $a = |\vec{a}|$ and $b = |\vec{b}|$ the intensities of the given vectors, and the angle between them is $\theta = \angle(\vec{a}, \vec{b})$.

The scalar product[39] (6) is maximal when $\cos\theta = 1$, for $\theta = 0$, ie when the larger coefficient $a_k$ is multiplied by the larger $b_k$ and vice versa, the smaller of the first vector with the smaller of the second. The same product becomes minimal when the higher coefficient of the first vector is multiplied by the smaller of the second and the smaller with the higher, that is, when these vectors tends to be mutually perpendicular.

In the case of one of these extremes, we can arrange the coefficients of the vector into monotone series and renumber the indices so that the first series is non-increasing, $a_1 \ge a_2 \ge \cdots \ge a_n$, so if the second is also non-increasing product (6) it will be maximal[40], and if the second is non-decreasing , $b_1 \le b_2 \le \cdots \le b_n$, that product will be minimal[41]. In each of the two extremes, when we imagine that the differences between the coefficients are smaller, until its final case, the equality $a_1 = \cdots = a_n$ and $b_1 = \cdots = b_n$, which is then both maximum and minimum.

For example, Shannon's information (4), which has the property of a minimal scalar product (6) because it multiplies smaller coefficients with higher ones, reaches its maximum in the case of equally probable outcomes of a given distribution. Result (5) is the highest possible value of Shannon's information[42], given how it is obtained from the probability distribution.

A similar example is Heisenberg's relations[43] with scalar products of particle position and momentum uncertainty

$$(\Delta s)^2 = \Delta x_1 \cdot \Delta p_1 + \Delta x_2 \cdot \Delta p_2 + \Delta x_3 \cdot \Delta p_3 - \Delta t \cdot \Delta E, \tag{7}$$

---

[39] see [8], 1.4 Scalar product
[40] see [8], Theorem 1.5.4.
[41] see [17], Theorem 1.2.7.
[42] Note how much simpler this proof is than usual by differentiation.
[43] see [8], formula (3.16)

which is of the order of Planck constant $h = 6{,}626 \times 10^{-34}$ m$^2$ kg / s. Substituting $x_4 = ict$ and $p_4 = iE/c$, where $c = 300\,000$ m/s is speed of light, and for imaginary unit is $i^2 = -1$, we get expression (7) as the square of the 4D space-time interval.

The interval $\Delta s$ is the length, and its square $\Delta s^2$ is the area. These are two uncertainties that we know from quantum mechanics to represent the lower limit of perception in the microworld. According to the previous explanation, expression (7) is also a minimum, because when we reduce the position uncertainty, the momentum uncertainty increases and vice versa. In particular, the square of indeterminacy of the 4D space-time interval, therefore (7), also represents information.

That (7) indeed can be information is seen from approximate formula (1). With some probability $p$, the occurrence of the event gives the information $-\log_2 p = q$, where $q = 1 - p$ is the probability that the event will not happen. In other words, the information of the occurrence of a given event is (approximately) equal to the vagueness of that event. Consistently further, the information is also the surface of 4D space-time.

## Energy and time

I elaborated all this prompted by a question from a colleague who tried to add axioms to probability theory to which he would formally join energy and time in that part of mathematics. He considered "energy" to be a quantity proportional to the probability of a random event that it represents, so he got everything "somehow upside down", as he writes. What's wrong with that – he asked me.

I have forwarded to him several of my earlier works on a similar subject, shorter and unfinished, together with these descriptions, and, it seems for the time being, he is satisfied with the answer. In that sense, the previously opened topic is the beginning for the next debate.

Space, time and matter are the very information whose essence is uncertainty. The more information the system has, the greater the power of dialing, the more "liveliness", so more energy goes with more information and then less probability. The main mistake of the mentioned colleague is, I suppose, to identify higher energy (always) with a higher probability, which further resulted in "everything upside down".

It is confusing, if formula (1) is also confusing, that with more probability $p = 1 - q$ we get less information, but that is why we have less uncertainty $q = -\log_2 p$ of a given event. If we use such an immediate probability of an event to determine its relative time, we agree with the principle of information minimalism.

Namely, it follows from the definition of information (3) that a more informative event is less probable, and as nature prefers the realization of more probable events, it prefers less informative ones. It inhibits, obstruct and slows down the manifestation and changes of energy, so from the generality of the principle of the highest probability we can say that the principle of the least information is the cause of the inertia of the body. Moreover, that the cause is body weight, and here's how.

Any particle (photon) that would move at the speed of light would be trapped in one present. Time itself would not pass for it, and the impact of the information principle would be minimal. From the point of view of an observer that particle also would be in 3D, in the way that its information is in one plane (electro-magnetic) with the third time dimension of the observer himself.

However, if the particle does not move at the speed of light, it itself penetrates into other time dimensions, its time flows and for that sake it gets stuck there, the principle of information is which hampered it. We perceive this as her inertia, but such has a mass of rest too. Therefore, the greater the certainty of a particle, the greater the probability of its occurrence, means its less presence at other times (outside the observer's), its less information and less energy than mass particles; it also means potentially longer duration, because it has no choice but to stay in the time of the observer.

Consistently, energy increases with expectation (2). The number $\mu = \frac{1}{p}$ also represents the frequency that is otherwise higher for higher energy particles. In the very expression, expectation tells us how many cycles it took to re-realize a given event, of which the observer sees only realizations. In short, $\mu$ is the amount of something we don't see but happens "under the hood."

## 15. Optimum Perception

December 7, 2020.

I demonstrate the formulas of information perception and interpretation. The information and power of choosing a living being is greater than that of an inanimate being. The former may clash greater abilities with greater challenges, as opposed to the passivity of inanimate matter. The principle of the least information certainly applies, but living beings tend to give in to equality and homogenization, and others go towards uniqueness and stratification.

### Perception information

In (this) information theory, we define the *intelligence* ($u$) of biological species as a quantity scaled to *freedom* ($w$) and inversely ratable to *constraints* ($v$). These new[44] terms are seemingly close to the old ones, but they are more operative and actually more precise of them (primarily because of $u = w : v$).

Freedom ($w = u \cdot v$) is then a quantity of options that a living being can have, and the limitations are natural, artificial or personal and in general those that the individual masters with his perceptions, by which we mean interactions of senses, instincts, impulses, mind and everything which can make decisions and choices. We understand the concept of "perception" in such a way that what we are talking about here can be transferred to dead objects.

Perception events are elements of a set ($\omega_1, \omega_2, \dots, \omega_n \in \Omega$) determined by a given subject, by which ($k = 1,2, \dots, n$) we distinguish her individual intelligences $u_k = u(\omega_k)$, constraints $v_k = v(\omega_k)$ and freedoms $w_k = u_k \cdot v_k$. Thus, for example, many animals see but do not perceive the same colors, and those that are not perceived do not directly influence the subject decisions. A photon interacts with an electron, but not with another particle. Set $\Omega$ selectivity expresses an important property of information, not to communicate everything with everyone.

[44] from [8]

In other words, for each being there is a special set of events $\Omega$, the range of perception by which we define a series (of $n$) of abilities $\vec{u} = (u_1, u_2, \dots, u_n)$ and a series of corresponding constraints $\vec{v} = (v_1, v_2, \dots, v_n)$. The first of the sequences defines the intelligence[45] of the individual, and the second the hierarchy of the environment. The scalar product of these two series is $w = \vec{u} \cdot \vec{v}$, where

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n \tag{1}$$

we call total freedom or *information perception* of a given individual. As individual freedoms $w_k$ represent some information (quantities of possibilities), they are additive quantities with the law of conservation[46], so the total freedom is their sum

$$w = w_1 + w_2 + \dots + w_n. \tag{2}$$

We can also denote this sum by $W = W(\Omega)$, but the notation $w = w(x)$ leaves the possibility of evaluating the observed subjects with the variable $x$.

Nevertheless, a living being is distinguished from an inanimate one by a greater power of choice, a greater number of options, and it is also shown by a greater information of perception. The principle of least action applies to all inanimate matter studied by physics, to which we now add the principle of least action.

The above interpretation of perception information is slightly different from the previous one[47], and we will see some more similar interpretations below. Such a variety of applications comes from the abstractness of consideration and the breadth of the very notion of information. We do not want to diminish the meaning of modern computers by limiting them, for example, to solving only algebraic equations, so why would we narrow the scope of information theory.

## Simple examples

Let us check several properties of scalar products of arrays (vectors) that are less common in algebra. My experiences from conversations with colleagues are that the views I present here should be explained slowly, step by step, because I have been in situations to be told “yes, if that were true”, for something that seemed familiar or trivial to me.

**Example 1**. For $n = 2$, consider the inequality

$$(u_1 - u_2)(v_1 - v_2) \ge 0, \tag{3}$$

which is true if the strings are of the same monotonicity, $u_1 \ge u_2$ and $v_1 \ge v_2$, or $u_1 \le u_2$ and $v_1 \le v_2$. The product of real numbers of the same sign is non-negative. Hence we calculate easily

$$u_1 v_1 + u_2 v_2 \ge u_1 v_2 + u_2 v_1, \tag{4}$$

---

45 the theory is new and I have to invent expressions
46 see [18]
47 appendix “14. Uncertainty”

which means that by multiplying “larger with larger and smaller with smaller” members of a given sequence, we get more information of perception than by multiplying “larger with smaller and smaller with larger”. This becomes important in living beings when we realize that intelligence is plastic, that it is usable for various phenomena. ♦

**Example 2**. In the case of longer arrays like $\vec{u} = (1,2,3)$ and $\vec{v} = (4,5,6)$ relation (4) becomes:

$$1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 \geq 1 \cdot 6 + 2 \cdot 5 + 3 \cdot 4,$$

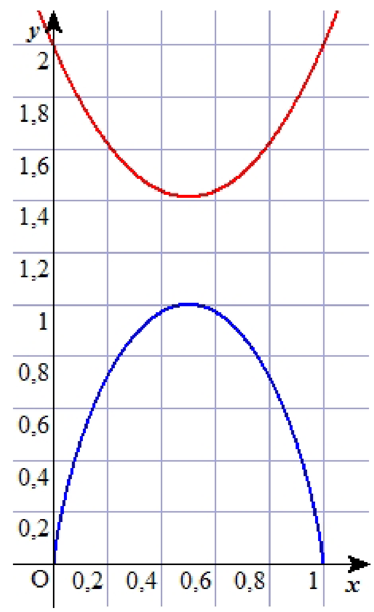
that is $32 \geq 28$, which is also true. ♦

From the previous two examples, we surmise that the confrontation of larger abilities with greater challenges and smaller ones with smaller ones means maximum information of perception. Living beings have this ability, to say stubbornness, to swim upstream and ignore easier paths. Conversely, weak engagement against larger difficulties and larger against smaller ones requires minimal information of perception, characteristic of inanimate matter. Surrendering to fate like a log through the water, a living being imitates the passivity of inanimate matter.

**Example 3**. The outcome of a random probability event $x \in (0,1)$ has Hartley's information $H(x) = -\log_2 x$ bit. The probability that the outcome will not happen is equal $1 - x$, and Shannon's information (mathematical expectation) of such an event is $S(x) = xH(x) + (1 - x)H(1 - x)$, that is

$$S(x) = -x \log_2 x - (1 - x) \log_2(1 - x).$$

The derivative of this function equated to zero gives the stationary point, $S'(x_0) = 0$, and hence the maximum ordinate $S(x_0) = 1$ bit with abscissa $x_0 = \frac{1}{2}$. ♦




Note in the lower graph to the right (blue) that Shannon's information, the function $y = S(x)$, mimics the minimal form of perception information, with its highest value in the case of equally likely outcomes. Otherwise, it is a form of inanimate matter and the principle of least action from which all trajectories known today in physics follow.

Top graph to the right, $y(x) = x \cdot 2^x + (1 - x) \cdot 2^{1-x}$ red, where both factors are in the sums of the growing function, in $y\left(\frac{1}{2}\right) = \sqrt{2} \approx 1.41$ has a minimum.

As incredible events are more informative and vice versa, by spontaneously more frequent realization of more probable outcomes the nature tries to achieve the less informative ones. This is a general phenomenon which we call the principle of information, and from which it follows from the third example that inanimate matter does not like equality, and that information is equivalent to action.

Let us further denote the left and right sides of inequality (4) by $M$ and $m$, so $M \geq m$, and change the coefficients of the second vector to equalization, say to the mean $v_0 = \frac{v_1 + v_2}{2}$, that is $\frac{4+5+6}{3} = 5$, the first and second examples respectively. We find:

$$M \geq u_1 v_0 + u_2 v_0 = \frac{M+m}{2} \geq m,$$

$$32 \geq 1 \cdot 5 + 2 \cdot 5 + 3 \cdot 5 = 30 \geq 28.$$

We would get similar with the changes of the first vector, and then with both. The conclusions are about the uniform engagement of capabilities by constraints. In the case of a living being, equalizing personal efforts by tasks would mean reducing information perception, and in the case of an inanimate being – increasing it.

In accordance with the principle of information, living beings tend to give in to equality and homogenization, and non-living beings go towards unity and stratification. The discovery of these unusual spontaneous developments is worth additional attention, so I leave other news of information theory for later.

## Generralisation

The previous views can be generalized. First on n-tuples of real numbers, then on complex ones, and then on vector and metric spaces of algebra and functional analysis in general.

**Theorem 1**. If $u_1 \geq u_2 \geq \dots \geq u_n$ and $v_1 \geq v_2 \geq \dots \geq v_n$, then

$$u_1 v_1 + u_2 v_2 + \dots + u_n v_n \geq u_1 v_{k_1} + u_2 v_{k_2} + \dots + u_n v_{k_n} \tag{5}$$

where $(k_1, k_2, \dots, k_n)$ is arbitrary permutation of the n-tuple $(1,2, \dots, n)$.

*Proof*. Let us go in order along the summands of the right-hand side of inequality (5) until we find the first pair of factors $u_i v_{k_i}$ of unequal indices ($i \neq k_i$). Then there is a sumand with the opposite pair of indexes $u_{k_i} v_i$ which can exchange second factors with the first summand (the first factors remain the same). Considering (4) and the change of only those two summands, we will get a larger total on the right side. Continuing the same procedure, with each replacement, the amount on the right side increases until we get the expression on the left side. ♦

The mentioned exchanges of factors, I repeat, in the theory of information make sense because of the plasticity of intelligence, its ability to transfer to various problems. It can be said that this is a basic property of intelligence as opposed to the obstacles it solves, made up mainly of unchanging circumstances.

**Theorem 2**. If $u_1 \geq u_2 \geq \dots \geq u_n$ and $v_1 \leq v_2 \leq \dots \leq v_n$, then

$$u_1 v_n + u_2 v_{n-1} + \dots + u_1 v_1 \leq u_1 v_{k_1} + u_2 v_{k_2} + \dots + u_n v_{k_n} \tag{6}$$

where $(k_1, k_2, \dots, k_n)$ is arbitrary permutation of the n-tuple $(1,2, \dots, n)$. ♦

The proof of the second theorem is analogous to the first. The proof of the following is also simple.

**Theorem 3**. When in expressions (5) and (6) we replace second factors with their mean values, $v_0 = \frac{v_1+v_2 \dots + v_n}{n}$, they become

$$M \ge u_1 v_0 + u_2 v_0 + \dots + u_n v_0 \ge m, \tag{7}$$

where $M$ and $m$ are in the order the left side of inequalities (5) and (6), ie the maximum and minimum values of the scalar products of the vector that can be obtained by permutations of the coefficients $\vec{u}$ and $\vec{v}$.

*Proof*. Multiplying we get in order: $u_1 v_0 + u_2 v_0 + \dots + u_n v_0 =$

$$\frac{1}{n}(u_1 v_1 + u_1 v_2 + \dots + u_1 v_n +$$

$$+u_2 v_1 + u_2 v_2 + \dots + u_2 v_n +$$

$$\dots$$

$$+u_n v_1 + u_n v_2 + \dots + u_n v_n),$$

and that is the sum that contains the maximum $M$ and the minimum $m$ with another $n-2$ mixed products. Each of these $n$ sums is not greater than $M$ and is not less than $m$, so the sum of all divided by $n$ is in the same interval of numbers, which means that (7) is true. ♦

In the following, we recall Schwartz's inequality[48], otherwise better known than the previous theorems, which is why I do not prove it here. It claims that for every pair of $n$-tuples (vectors) it holds

$$|u_1 v_1^* + \dots + u_n v_n^*|^2 \le (|u_1|^2 + \dots + |u_n|^2)(|v_1|^2 + \dots + |v_n|^2), \tag{8}$$

where the coefficients can also be complex numbers ($u_k, v_k \in \mathbb{C}$). On the right side of the inequality are the squares of the norms of the given vectors:

$$\|\vec{u}\|^2 = |u_1|^2 + \dots + |u_n|^2, \ \ \|\vec{v}\|^2 = |v_1|^2 + \dots + |v_n|^2, \tag{9}$$

so, if $\|\vec{u}\| = \|\vec{v}\| = 1$, which is always in quantum mechanics, Schwartz's inequality (8) becomes

$$|u_1 v_1^* + \dots + u_n v_n^*|^2 \le 1. \tag{10}$$

I have already given interpretations in information theory[49].

Generalizations to complex numbers of expressions of information perception (1) have shown well, and among them the inequality (10) which tells us that the norm of a scalar product of quantum states could

---

48 see [2], 2.2.2 Schwartz inequality
49 see [1], 3.28 Authority

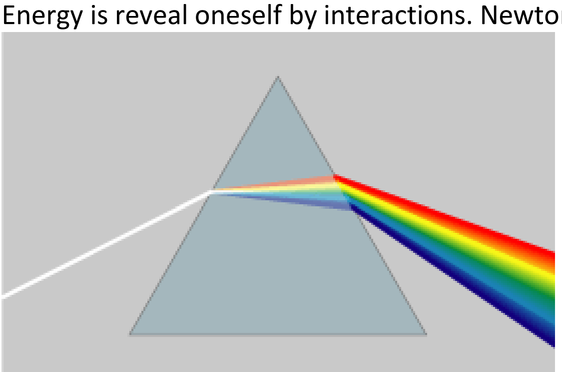
again be some probability[50]. Consistently, states with a higher such probability would be more prone to quantum coupling (entanglement).

## Speculations

In the new theory, any explanation is a speculation. So I have no illusions that I can convince anyone of anything with this, especially because everything else is more and more suspicious, but I can try to keep things going[51] by making the disputes unprovable.

From the application (1) as information of perception to the interpretation of inequality (10) much needs has to be clarified and verified. Does greater information of perception really represent greater "liveliness" and why would such a product of quantum states (vectors) express the probability not only in the formal sense, but also in essence, that because of such a greater one there is a greater chance of coupling given quantum states? This is followed by the question of sequences that determine the probabilities of measurement (Born rule) with the information of perception.

Energy is reveal oneself by interactions. Newton's prism refracts more blue waves of light which get stuck more with more energy. Longer electromagnetic waves reach more through the universe, because with less energy they get less attached to the environment.



Compton effect[52] is the scattering of a photon from an atom in which the photon loses part of its energy, and its wavelength increases with increasing deflection angle. It is otherwise valid for the proof of the corpuscular nature of light, and I also use it for the proof of motion in accordance with the principle of probability, and accordingly information. Namely, some force (collision) is needed to change the state of probability – in relation to the given photon. From the point of view of another subject (experimenter), the photon also turns into larger wavelengths, that is more smeared states of lower probability density.

Heisenberg's uncertainty of the position of a particle is reduced at the expense of momentum, time for the sake of energy, so that the total information is preserved – if we consider the information to be equivalent to the action and assume that it is a measure of uncertainty.

Game theory deals with decision making and it brings it closer to information theory. Energy as a factor of the amount of uncertainty will take a step more like the unexpectedness of the move that contributes to the victory of the player. Additionally, perception information that expresses the degree of confrontation between "strong with strong and weak with weak" is a measure of liveliness in reciprocity games (eye for an eye and kind for a darling) otherwise one of the most successful strategies in winning games organized by mathematicians on computers to test the theory.

[50] The squares of the norms of the coefficients of quantum states (vectors) are observable probabilities.
[51] This theory offers continuations of known, say [19].
[52] https://en.wikipedia.org/wiki/Compton_scattering

We know that the monopolist sees his interest in the absence of competition in profit, but that duopoly and oligopoly are better for the good of society. There is another application of information perception in understanding the market economy of the French mathematician Carnot[53] (Antoine Cournot, 1801-1877) and now in a far simpler way, if the basic thesis were acceptable. Moreover, then, through the same idea of greater vitality of facing equally strong and the advantages of political competition also become clearer.

By accepting that a living being has more information than the inanimate matter it consists of, we accept that it has more options. This is the connection of the mentioned vitality and energy. From the expectation that looking through a telescope or microscope, into cosmic phenomena or in a petri dish, to recognize life by a movement that deviates from the principle of least action of physics and that it seems to us that a living being has additional choices and more information, because of the conservation law, we expect that greater information carries greater obligations of choosing.

As in this story all space, time and matter consist only of information, so that what at least does not work and does not exist, and the essence of information is uncertainty, then change and progress become imperatives. The universe is constantly changing, so a particle that oscillates locally, recurring periodically, is never in the same broader context. At the same time, its future is probably less informative (along with the increase in entropy) and the information of the substance decreases at the expense of the growing space.

The information theory, as I present, can be tested on applications from extremely large quantities of physics to extremely small ones, but also in an additional understanding of fear. The living being, with its excess information, has a certain “comfort zone” above which is the fear of the unknown, and below which is the fear as kind of anxiety. The upper is the fear of freedom, and the lower is the lack of freedom. At the same time, the principle of minimalism of information quietly but persistently pulls the “comfort zone” down.

All the surrounding substance is already filled with information and it is not easy to get rid of it, but the living beings its freedoms also surrender to the organization, with which the hierarchy also becomes a “living being”. The state's aspiration for equality speaks about the information of the perception of the state structure closer to the maximum $M$, than the minimum $m$, ie about its vitality, and the aspiration to ensure the safety of citizens (to protect them from uncertainty) and to exaggerate with restrictions (reduction of freedoms), speaks of its subordination the principle of (minimalism) information as well.

In persons in particular, the propensity for obedience, the need for authority, order and efficiency, are expressions of the same principle of information, often disguised but essentially equal to much more intense forms of “worship of death”.

[53] https://www.academia.edu/39880324/Frequency

## 16. Decomposition of Information

December 11, 2020.

I explain why the deeper "essence" of information can be found in countless causes in such a way that we can equally assume that that essence does not actually exist.

### Elementary information

By its definition, information is a quantity of options and that is why we can say that it has more or less, but the law of conservation applies to it and that is why it is discrete – because only infinitely divisible sets can be their real parts. Information is transmitted in the smallest portions because it arrives in quanta of energy, and travels in quanta because each performs some action. I wrote about that earlier, as well as about the next paradox of uncertainty.

The world consists only of information whose essence is uncertainty. However, less uncertainty means more certainty, so in the case of the smallest parts, if we continued to talk about even less information, we would talk about more. More certain events are less informative, and more informative are more uncertain. These absurdities are the subject of the annex.

We know that Hartley's (1928) information is the logarithm of the number of equally probable outcomes of random events ($n = 2,3,4, \ldots$), that is, the logarithm of the probability ($p = 1/n$) of one of such

$$H = -\log_b p. \tag{1}$$

The base of the logarithm determines the unit of information. It is a bit when the base is $b = 2$, a nat when the base is $e = 2{,}718 \ldots$ or decit if it is the number 10. Of this, the probability is more important to us than as the real number $p \in (0,1) \subset \mathbb{R}$ that can always be represented as the product of several such numbers, from the same interval.

If $p = p_1 p_2$, where $p_1, p_2 \in (0,1)$, so $p/p_2 \le 1$ follows from $p_1 \le 1$, and hence $p_2 \ge p$. In the same way we derive $p_1 \ge p$ from $p_2 \le 1$. Hence the view that probability factors, when all represent some probabilities, all represent more probable events than the given; they are strictly more probable if the given event is not certain ($p \ne 1$). This is in line with the previous explanation.

No matter which base, $b > 0$ and $b \ne 1$, of the logarithm to take, if the given probability is expressed by the product of other probabilities, $p = p_1 p_2 \ldots p_n$, then Hartley's information (1) becomes the sum of information, $H = H_1 + H_2 + \cdots + H_n$ with $H_k = -\log_b p_k$ for $k \in \{1,2, \ldots, n\}$. Compliance with the above description is complete.

Due to the discrete nature of information, both mathematical theorems and their flows of proves, as well as legal regulations or administrative decisions, are always some separate steps.

### Expected value

Shannon's (1948) information is a mathematical expectation, or the mean value of information (1) of the probability distribution

$$S = -p_1 \log_b p_1 - p_2 \log_b p_2 - \cdots - p_n \log_b p_n, \tag{2}$$

where $p_1, p_2, \dots, p_n \in (0,1)$ and $p_1 + p_2 + \cdots + p_n = 1$.

Knowing[54] that each individual outcome ($k = 1,2, \dots, n$) has probabilities of events $p_k$ and non-events $q_k = 1 - p_k$ such that the event information is approximately equal to the probability of non-events, $q_k = -\log_b p_k$, average value (2 ) information of a given distribution the probability is approximately equal to the information of perception

$$W = p_1 q_1 + p_2 q_2 + \cdots + p_n q_n. \tag{3}$$

From $x - 1 > \log_b x$, for $x < 1$, substituting $x = 1 - q$ we get $q < -\log_b(1 - q)$, and hence $W < S$. Let me remind[55] you, the physical information (for which the law of conservation applies) is greater than Shannon's ($S < L$).

Scalar product of vectors $\mathbf{x} = (x_1, x_2, \dots, x_n)$ normed

$$\|\mathbf{x}\| = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} \tag{4}$$

is slightly larger than (3) because

$$\left(\sum_{k=1}^{n} p_k\right)^2 = \sum_{k=1}^{n} p_k^2 + 2\sum_{i \neq j} p_i p_j \geq \sum_{k=1}^{n} p_k^2 = \|\boldsymbol{p}\|^2.$$

So, in the vector space of the norm $\|\mathbf{p}\| = \sqrt{|p_1|^2 + |p_2|^2 + \cdots + |p_n|^2}$ expression $W$ as perception information will grow towards Shannon's information $S$ and be closer to physical information $L$. I will return to this in a future article.

## Complex information

If we allow the probability to be a complex number $z = e^w \in \mathbb{C}$, then $w = \ln z$ (lat. Logaritmus naturalis). In the polar form the complex number is $z = re^{i\varphi}$, with $r > 0$ and $\varphi$ real numbers, and with such Hartley's information (1) becomes

$$w = \ln r + i(\varphi + 2k\pi), \tag{5}$$

where $k \in \mathbb{Z}$ is an arbitrary integer. Probability is then a periodic function.

Probability waves are a proven phenomenon in quantum mechanics, so information theory should not ignore them either. Especially because norms (4) and the condition $\|\mathbf{x}\| = 1$ define Born probabilities.

[54] 14. Uncertainty
[55] see [18], Theorem 2.3.6.

However, quantum mechanics has some secrets[56] that will now help us to understand even more deeply the elementary information with which this story began.

Quantum states are representations of vectors and quantum processes are representations of unitary operators. This means that in the linear algebra of quantum mechanics (3) it becomes:

$$w = p_1 q_1^* + \dots + p_n q_n^* = (p_1, \dots, p_n) \begin{pmatrix} q_1^* \\ \vdots \\ q_n^* \end{pmatrix} = \langle p | q \rangle, \tag{6}$$

where in the end Dirac's bracket brackets are used for the notation of scalar (inner) product of vectors, and $z^* = x - iy$ is a conjugate complex number $z = x + iy$.

However, the scalar product (6) is equal to $\langle p | \hat{I} | q \rangle$, where $\hat{I}$ is the corresponding unit matrix (operator) which can be represented in countless ways by the product of also unitary operators (matrices), but also some others. Namely, every regular matrix $\hat{A}$ has an inverse matrix $\hat{B} = \hat{A}^{-1}$ so that $\hat{A}\hat{B} = \hat{I}$, and there are countless of them.

An interesting example are Pauli's second-order matrices:

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{7}$$

whose squares are a unit matrix ($\hat{\sigma}^2 = \hat{I}$). Multiplying these matrices by an imaginary unit we obtain three quaternions $\hat{q}_\kappa$ with indices $\kappa \in \{x, y, z\}$, whose product pairs multiplied by an imaginary unit give the initial Pauli matrices.

Therefore, the deeper “essence” of information perception can be found in countless interpretations of such factors, just as we can say that this essence does not actually exist. This is analogous to representing Fourier series[57] in different ways: we can approximate almost every function with a given precision to a given fragment of almost every other function, which is why we can say that micro quantities can be of arbitrary shape, just as there are no shapes at their level.

[56] see 2. Dual Vectors

[57] see [1], 2.20 Fourier Deriving

# 17. Latent Information

December 15, 2020.

I explain Emergence Theory[58] using information theory and discover some more interesting approximations[59].

## Introduction

The latent process, which is currently hidden, is given only in potential form, but will later, eventually, develop and manifest. In philosophy, systems theory, science and art, emergence occurs when the observed entity has properties that its parts do not have on themselves, characteristics or behaviors that occur only when the parts interact with the wider whole[60].

Here I will show that the information theory presented in the book "Physical Information" [18] predicts such a phenomenon, the emergence of latent information. So that this text would not be too repetitive, I would like to add a few, I hope, useful approximations in the future.

## Binomial distribution

In probability theory and statistics, the binomial distribution $\mathcal{B}(n,p)$ is a discrete distribution of probability $p \in (0,1) \subset \mathbb{R}$ and the number of successes in a series of $n = 1,2,3,\dots$ independent experiments. The number of attempts ($n$) is fixed, each attempt is independent (none of them affects the probability of the others), and the probabilities of successful ($p$) and unsuccessful outcome ($q = 1 - p$) are constants.

**Example 1**. When tossing a fair coin, two equally likely outcomes are possible, either from the set $\Omega=$\{tails, heads\} and each with a 50-50 percent chance. If we expect a "tails" to fall in one throw, then the probability of a favorable outcome is $p = \frac{1}{2}$, as well as the probability of an unfavorable $q = \frac{1}{2}$. ♦

**Example 2**. When throwing a fair dice, there are six equally probable outcomes, the numbers from the set $\Omega = \{1,2,3,4,5,6\}$ and each with the chance of 1: 6. If we expect the "3" to fall in the first throw, then the probability of a favorable outcome is $p = \frac{1}{6}$, and the probability of an unfavorable $q = \frac{5}{6}$. ♦

The probability that in $n \in \mathbb{N}$ of the experiment a favorable outcome occurs $k \in \{0,1,2,\dots,n\}$ times is

$$p_k = p^k q^{n-k}. \tag{1}$$

The ways, the combination that in a series of $n$ outcomes $k$ of them are favorable is

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \tag{2}$$

[58] at the suggestion of colleague D. Koščica, prof. Informatics in Gimnazija Banja Luka
[59] besides the 14. Uncertainty
[60] see [20]

where "n!" reads "en-factorial", and for example $4! = 4 \cdot 3 \cdot 2 \cdot 1$. Therefore, the probability that in $n$ experiments in any of (2) the combination occurs exactly $k$ favorable and $n-k$ unfavorable outcomes, each individual probability (1), is

$$P_k = \binom{n}{k} p^k q^{n-k}. \tag{3}$$

It is clear that the sum of all (3) is $\sum_{k=0}^{n} P_k = 1$ and that they form a binomial distribution $\mathcal{B}(n,p)$.

You will find further processing of the binomial, and then other distributions, in the mentioned book[61], and here we deal with only one aspect and one general example.

## Binary information

For $n = 1$ the binomial distribution $\mathcal{B}(1,p)$ consists of one experiment with two possible outcomes – favorable probability $p$ and unfavorable probability $q = 1-p$. Shannon[62] and physical information are then equal and amounts

$$S_1 = J_1 = -p \log_b p - q \log_b q. \tag{4}$$

With the allowed base of logarithms, $b > 0$ and $b \neq 1$, we choose the units of information.

For example, when $b = 4$ by substitute $p = x$ и $q = 1-x$ in (4) we get the function $y_1 = S_1(x)$, ie

$$S_1(x) = -x \log_4 x - (1-x)\log_4(1-x)$$

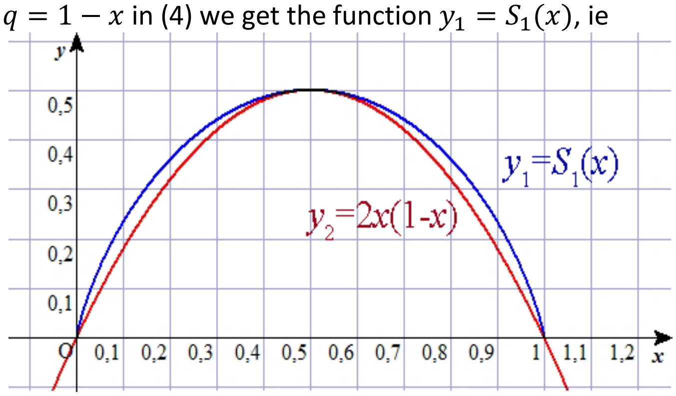
whose graph is upper (blue) in the figure on the right. This information is twice as large as that measured in bits, approximately

$$x = -\log_2(1-x) \text{ и } 1-x = \log_2 x$$

which in (4) gives the lower, red graph

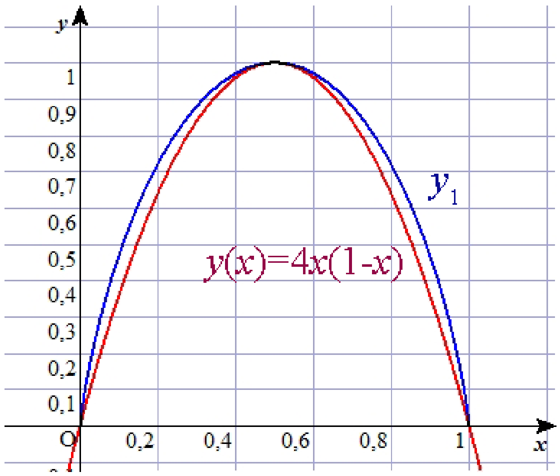
$$y_2 = 2x(1-x).$$

The upper graph is a transcendental function that contains logarithms (bases four), and the lower is a parabola.




[61] see [18]

[62] I also called Shannon's information in the book [18] technical information.




The following example is in the picture on the left. Putting for the base $b = 2$ we get the information (4) in bits and the upper, blue graph of its function $y_1$. Lower, red graph of the parabola

$$y(x) = 2x(1 - x) \cdot 2$$

is its approximation.

In general, for each given logarithm base, $b \in \mathbb{R}_+ \backslash \{1\}$, there is at least one parable, say

$$y(x) = 2x(1 - x) \cdot \log_b 4,$$

which approximates the information (4). The above follows the manner of the previous ones, and I cite them only as the idea that Shannon's information (4) can be roughly replaced with some other "information" (for some future text) besides "physical".

## Trinary information

For $n = 2$ the binomial distribution $\mathcal{B}(2, p)$ consists of two experiments with four possible outcomes – two favorable probabilities $p^2$, favorable and unfavorable probabilities $pq$, unfavorable and favorable probabilities $qp$, two unfavorable probabilities $q^2$. Shannon's (technical) and physical information are then in order:

$$\begin{cases} S_2 = -P_2 \log_b P_2 - P_1 \log_b P_1 - P_0 \log_b P_0 \\ J_2 = -P_2 \log_b p_2 - P_1 \log_b p_1 - P_0 \log_b p_0 \end{cases}. \tag{5}$$

The idea of physical information is not to calculate the logarithms of repeated probabilities and then to take the mean of all combinations of logarithms. The resulting such information is greater than the corresponding Shannon, that is, it is exactly equal to the two binary ones. The physical information of two experiments is equal to the double information of one experiment, $J_2 = 2J_1$.

Namely, $J_2 = -p^2 \log_b p^2 - 2pq \log_b pq - q^2 =$

$$= -2p^2 \log_b p - 2pq \log_b p - 2pq \log_b q - 2q^2 \log_2 q$$

$$= -2p(p + q) \log_b p - 2q(p + q) \log_b q,$$

and hence $J_2 = 2J_1$. So, the law of conservation really applies to this form and it makes sense to call it physical information.

Latent information is the difference between physical and technical, easy to find

$$L_2 = J_2 - S_2 = pq \log_b 4. \tag{6}$$

This number is approximately half the information of one experiment and is not negligible. An example of an experiment is the throwing of a cube, or the superposition of a spin, or qubit (quantum bit) of a quantum particle. It consists of about two such latent information. An example of two experiments would be qutrit (quantum trit), a superposition of three possibilities (mutually orthogonal quantum states) with information of approximately three latent ones (6).

## General case

The binomial distribution $\mathcal{B}(n,p)$ consists of $n = 1,2,3,\ldots$ experiments with $n+1$ possible probability outcomes:

$$p^n, p^{n-1}q,\ \ldots, p^{n-k}q^k, \ldots, pq^{n-1},\ q^n, \tag{7}$$

where probability (1) can occur in (2) combinations. Technical (Shannon) and physical information are then in order:

$$\begin{cases} S_n = -\sum_{k=0}^{n} P_k \log_b P_k \\ J_n = -\sum_{k=0}^{n} P_k \log_b p_k \end{cases}. \tag{8}$$

The physical information of an n experiment is equal to the n-triple information of one experiment, $J_n = nJ_1$. Наиме, $J_n = -\sum_{k=0}^{n}\binom{n}{k} p^{n-k}q^k \log_b p^{n-k}q^k =$

$$= -\sum_{k=0}^{n}\binom{n}{k}(n-k)p^{n-k}q^k \log_b p - \sum_{k=0}^{n}\binom{n}{k}p^{n-k}kq^k \log_b q$$

$$= -p\left[\frac{\partial}{\partial p}\sum_{k=0}^{n}\binom{n}{k}p^{n-k}q^k\right]\log_b p - q\left[\frac{\partial}{\partial q}\sum_{k=0}^{n}\binom{n}{k}p^{n-k}q^k\right]\log_b q$$

$$= -p\left[\frac{\partial}{\partial p}(p+q)^n\right]\log_b p - q\left[\frac{\partial}{\partial q}(p+q)^n\right]\log_b q$$

$$= np(p+q)^{n-1}\log_b p - nq(p+q)^{n-1}\log_b q$$

$$= n(-p\log_b p - q\log_b q),$$

and hence $J_n = nJ_1$. Therefore, the law of conservation really applies to the “physical information” of binomial distribution defined in this way. For other types of distribution, the definition is similar and you can follow that in the mentioned book, and I will end this text with something that is not covered there.

For latent information, the difference between physical and technical, we now find:

$$L_n = J_n - S_n = \sum_{k=0}^{n}\binom{n}{k}p^{n-k}q^k \log_b \binom{n}{k}. \tag{9}$$

These are the numbers in order $L_1 = 0,\ L_2 = 2pq\log_b 2,\ L_3 = 3pq\log_b 3, \ldots,$ ascending series.

# 18. Wavelength

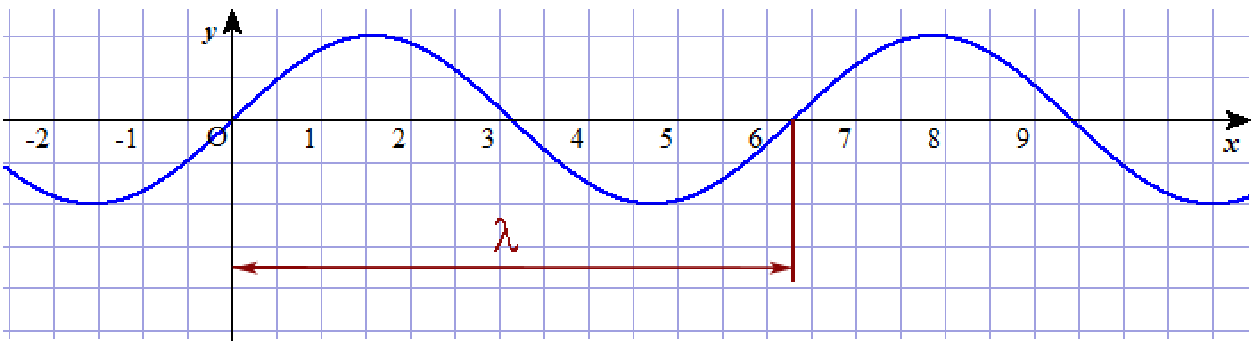
About information and action

December 26, 2020.

The wavelength can be understood as the “smear” of the probability of the position of the particle that represents the wave, so that the lower the probability density, the greater the information. But then the energy of the particle should be looked at a little differently than usual.

## Introduction

The wavelength, the symbol λ (Greek: lambda), is the smallest distance of two points of the same phase of one wave. In the figure it is the length, $\lambda = 2\pi$, of one period of the sinusoid $y = \sin x$. Amplitude means the intensity (existence) of the wave by deviating from the middle (equilibrium) positions, here from the abscissa. An important formal characteristic of the wave is the frequency (lat. Frequentare - to visit, often do), rate or number of repetitions of the period in a unit of time.




The elementary particles of physics, as well as matter in general, have wave nature as well. Electrons can receive energy from the electromagnetic field only in discrete units (quanta or photons) in the amount

$$E = hf, \tag{1}$$

where $E$ is the energy quantum, $h$ is the Planck constant (approximately $6{,}626 \times 10^{-34}$ Js), and $f$ is the frequency. The duration of one repetition, $\tau = 1/f$, multiplied by the corresponding change in energy is the action

$$\tau E = h. \tag{2}$$

The first formula (1) is generated by the product of wavelength and momentum

$$\lambda p = h, \tag{3}$$

and by means of both, (2) and (3), it is easier to understand Heisenberg's relations of uncertainty:

$$\Delta t \cdot \Delta E \ge \frac{1}{2}\hbar, \quad \Delta x \cdot \Delta p \ge \frac{1}{2}\hbar, \tag{4}$$

where $\Delta$ denotes standard deviations, ie uncertainties, and $\hbar = h/2\pi$ is the reduced Planck constant. These are well-known attitudes in physics.

Information theory is a much broader concept that should accurately describe the phenomena of the inanimate micro world. The basic (new, my) ideas of information theory are that time is created by the realization of random events and that greater uncertainty has greater effect (action). This implies the objectivity of uncertainty, that the information is smaller the higher the probability, and that the tendency of nature to realize more probable events is valid as a principle (minimalism) of information, about the tendency to realize less informative.

My explanation of the formation of mass by “sticking” through time is unknown (it looks a bit like the Higgs field). Namely, particles that have mass (in rest) do not move at the speed of light and its time does not stand still, so they “penetrate” through the layers of time and the principle of information applies to them all the more. Those “under the hood” are still doing something[63], which a relative observer from any reality[64] cannot fully observe.

## Light

Electromagnetic waves whose smallest parts are photons move at the speed of light. They therefore do not have a mass of rest or their own (proper) flow of time. Because time stands still, particles that move at the speed of light are captives of some 3D “space”, the specially “present” which consists of two dimensions of information they carry and the time of the observer.

Simply put, particles of the speed of light cannot exist without the reference system (observer) to which their information relates. Otherwise, there would be a certain default wavelength and a corresponding coordinate system, at rest or in motion, of their source.

Because we see with light, our perception of the world is as it is, limited by the range of photons always in some 4D space-time. On the other hand, they are not able to get out of that world, nor to hide something from us. In addition, due to the quantum nature of the action, ie information (1) and (2), the energy of light is in such a simple relation with frequency, as well as the momentum with wavelength, and due to the same, the speed $c = \lambda f$ is invariant.

At the age of 39, Christian Doppler[65] published his most significant work on the effect of increasing and decreasing the frequency of light depending on the relative motion of stars. When the light source (speed $c \approx 300\,000$ km/s) of frequency $f_0$ approaches, by the speed $v$, we observe the frequency[66]

$$f_+ = f_0\sqrt{\frac{1+\beta}{1-\beta}}, \ \ \beta = \frac{v}{c}. \tag{5}$$

[63] 14. Uncertainty
[64] Including "parallel", see [1], 2.13 Space and Time
[65] Christian Doppler (1803-1853), Austrian mathematician and physicist.
[66] Relativistic formula for Doppler Effect.

When the source moves away, the same formula applies with the change of the speed sign, so the mean value (arithmetic mean) of the incoming and outgoing frequencies

$$f = \frac{1}{2}(f_+ + f_-) = \frac{1}{2}\left(\sqrt{\frac{1+\beta}{1-\beta}} + \sqrt{\frac{1-\beta}{1+\beta}}\right) f_0 = \frac{f_0}{\sqrt{1-\beta^2}},$$

that is

$$f = \gamma f_0, \quad \gamma = \frac{1}{\sqrt{1-\beta^2}}, \quad \beta = \frac{v}{c}. \tag{6}$$

It is also a formula for lateral (transverse) change of light frequency.

The product of wavelength and frequency is the (constant) speed of light, so (5) gives

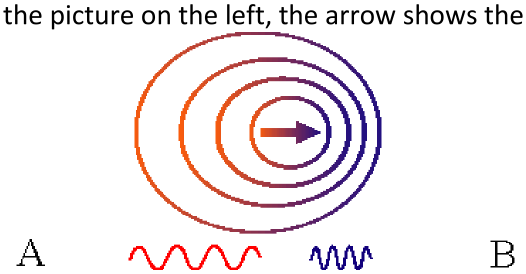
$$\lambda_+ = \lambda_0\sqrt{\frac{1-\beta}{1+\beta}}, \quad \lambda_- = \lambda_0\sqrt{\frac{1+\beta}{1-\beta}}, \tag{7}$$

and the arithmetic mean of these two lengths is again of the previous form

$$\lambda = \gamma\lambda_0, \tag{8}$$

where $\lambda_0$ is the own (proper) wavelength of light whose source is at rest.

In the picture on the left, the arrow shows the direction of movement of the light source, from observer




A to B, and thickening of the wave for the right observer (B) towards which the source is moving, ie thinning for the observer to the left (A) from which the source is moving away. The product of wavelength and frequency, $\lambda_k f_k = c$ for $k \in \{+,-\}$, in both cases is the same speed of light, but this is not the case and for arithmetic means. There is an interesting question. The product of arithmetic means, frequency (6) and wavelength (8), does not give the speed of light.

Also, when we replace relative energy and time $(E, t)$ with the proper $(E_0, t_0)$, by relativistic formulas

$$E = \gamma E_0, \quad t = \gamma t_0, \tag{9}$$

for the appropriate frequencies of light we would get $f = f_0/\gamma$, which would mean slowing down the frequency and, according to (1), reducing the energy. Note that this leads to a similar alleged discrepancy as in the just mentioned "interesting question".

Similar to gravity, which is a macro phenomenon[67], the law of conservation and the final divisibility of action and information have a higher priority in the micro world than relativistic effects. In information theory, space and time should be treated symmetrically, from at least 6D multiverse we extract 4D

[67] 11. Force and Information

space-time by declaring three coordinates spatial ($x_1, x_2, x_3$) and the fourth ($x_4 = ict$) temporal. This idea is justified by such a possibility of forming relativistic equations, both Einstein's macroscopic and Klein-Gordon's microscopic, but it comes from the interpretation of photons mentioned here.

Electromagnetic waves define space-time as we see it. Their source in arrival, due to the relativistic slowing down of time, is in the future of the relative observer until the moment of passing from when in departure it reaches deeper and deeper into the observers past. Shorter relative wavelengths than the proper on arrival and higher on departure do not only mean less blur of photons on arrival and more on departure, but also a higher probability of future positions than past ones, ie they tell us the reason for the movement of our time, more precisely our present towards the future.

Just as the entity we call our present moves from our past to our future because such a movement is more likely, so do particles move in their trajectories, because for them such a movement is most likely. The proof of the first is in the previous paragraph, and of the second is the Compton effect[68].

The glass that is on the table at this moment is because such a state is most likely for those who see it that way – for the glass, the environment and the observer. It will be the same in the next moment due to the inertia of probability, that is, the principle of least action, or the principle of least information, until some force like a hand acts on the glass and move it. Probabilities change by force.

## Mass

Unlike particles which time stand and they move at the speed of light, there are also those that have their own duration. Such penetrate through the layers of time and the principle of information it applies to them all the more, they endure more depreciation, ie inertia. They have their “secret life” which is invisible (to photons) to every single observer. Therefore, they also have a “secret energy” that they do not manifest, do not show to everyone.

If we divide energy into *active* and *passive*[69] depending on the immediate manifestation, that is, on the property that a given observer “sees” it or not, we will say that light has only active energy. The energy that is “more visible” is more certain and carries less information. Hence, the higher the energy of the photon is of the smaller the wavelength, because it is “smeared” in a shorter space and has a lower probability density.

**Example 1**. The wavelength of the photon of blue light is $\lambda = 450$ nm (1 nm = $10^{-9}$ m), and red lights $\lambda = 700$ nm, so is energy ($E = hc/\lambda$) blue lights $4{,}4 \times 10^{-9}$ J, and red 2,8 J.

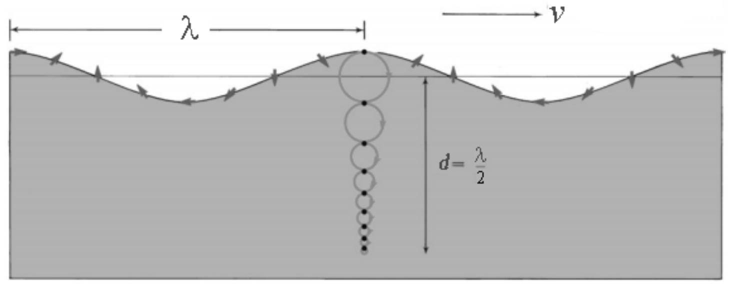
**Example 2**. From the wavelength formula $\lambda = h/p$, where $h = 6{,}626 \times 10^{-34}$ Kg m$^2$/s Planck's constant in SI units, and momentum $p = mv$ particles, for electron masses $m = 9{,}1 \times 10^{-31}$ Kg and speed $v = 10^6$ m/s, we get $\lambda = 7{,}3 \times 10^{-10}$ m, which is approximately the length of one atom.

[68] see [1], 2.6 Compton Effect

[69] I intentionally avoid the terms "kinetic" and "potential".

Note that as the kinetic energy (mass and velocity) increases, the electron momentum increases and the wavelength decreases. In that sense, kinetic energy is mentioned as "active", whose greater manifestation means greater certainty and less information.

**Example 3**. The horizontal motion of waves on the surface of the sea is caused by the circular motion of the water particles below (picture right), and the water particles themselves at the top of the wave and the wave move in the same direction. The formula also applies to the speed of sea waves $v = \lambda/\tau$, where $\tau = 1/f$ is duration of one period. The depth of these waves is about half the wavelength, $d = \lambda/2$, so their kinetic energy grows with the mass of water they capture and the square of the velocity.




With the increase of the mass and speed of the water that participates in the wave on the surface of the sea, the kinetic energy increases, but also the passive, latent power of that water increases, so the wavelength also increases.

## Conclusion

To put it mildly, it is debatable whether a shorter wavelength really means a higher total energy of the system it represents and, therefore, a clearer expression of that system. The thesis that opens here goes with the understanding of the world that is more complex than we usually understand it, and this article is only an introduction to its next elaboration.

# 19. Decomposition of Information II

December 30, 2020.

Information is equivalent to action, not particularly to energy or time. The smoother the energy and momentum of a free particle and the larger the domain, the more continuous and abstract space and time are.

## Infinity

This is the second part of the article of the same name[70], and both are a continuation of the story about one confusing place in information theory, which I have mentioned several times[71], but it always seems with insufficient attention. Here's what it's about.

The main (hypo) thesis of my version of this theory is the universality of information, its ubiquity in physical interactions and their interpretations. Consistent with this, the information should also include abstract mathematical truths, and then we have the problem of accepting infinity. It is difficult to dispute the applicability of the infinitesimal calculus (limits, derivatives and integrals), as well as its accuracy, and it is almost impossible to ignore it in such stories. It is similar with Cantor's[72] method of bijection and the properties of infinite sets (discrete and continuum), the challenge of which is not an easy task.

An infinite set defines[73] the ability to (by quantity) be its true, proper subset. The best known such set is the natural numbers that we see as a real, proper subset of integers, and this proper subset is the proper subset of the rational ($\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$). All three have the same (cardinal) number of elements of a countable infinite set, $\aleph_0$ (read: alef zero). These are the so-called discrete sets and are not the most numerous. Greater infinity than discrete is a continuum such as real numbers ($\mathbb{R}$), or the irrational that are proper subset of real ones.

Opposite the infinities are the information of the physical world which are finite. The law of conservation applies to them[74]. The law of conservation also applies to physical actions, which is why they are finally divisible, quantized, hence the equivalence of physical action and information. Note that this "equation" does not mean "one and the same", nor that it does (not) make sense to talk about parts of the smallest packages of information or action. It means that these parts, if they exist, are not self-sustaining, that is, they cannot manifest themselves as independent entities in physical action (information) or interactions (communications).

---

[70] 16. Decomposition of information
[71] see [1], 3.17 Present
[72] see https://www.storyofmathematics.com/19th_cantor.html
[73] see [21], p. 11. Differences
[74] see [1], 1.14 Ammy Noether

In the previous title (16), I showed[75] that the existence of parts of a quantum is a matter of free assumption. As in Zermelo's[76] set theory, where we can assume that there is a set of greater cardinality than the countably infinite and less than the continuum, but we don't have to. Both theories will be equally correct. Lobachevsky[77] proved something similar when he discovered a non-Euclidean geometry named after him. He proved that the new geometry is as accurate as Euclidean, more precisely that it is incorrect if and only if Euclidean geometry is incorrect.

After the pioneering discoveries, the proofs that the alternative theories are not contradictory, it is still easier to think up sets of independent axioms and build all kinds of "true" constructions on them. Postulates can be so independent that by turning one of them into the opposite statement, we get two alternative and equally correct (incorrect) theories. In that sense, the decomposition of (independent) information into its (non-independent) parts is a correct topic.

The supplement to the previous story follows, and the detailing around these series (the sum of the items) also aims at one of my perhaps next contributions. For now, the point of describing Fourier's analysis is in the (infinite) possibilities of giving geometric shapes to the smallest particles of physics. I consider them as accurate as the alternative that they have no form, each of these theories with some peculiarities.

## Fourier approximation

Representation of a function in Fourier series is a mathematical operation by which the function $f(x)$ decomposes into its "spectral components" by a series of functions $f_n(x)$ more and more equal to the given one for simpler analysis. The given function $f(x)$ should be integrable on an interval of length $L$ which will be the period of each of the members of the Fourier series $f_n(x)$.

Fourier series (for $n = 1,2,3,\ldots$) reads

$$f_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} \left( a_k \cos\frac{2\pi kx}{L} + b_k \sin\frac{2\pi kx}{L} \right), \tag{1}$$

where are the Fourier coefficients[78]:

$$a_k = \frac{2}{L}\int_0^L f(x)\cos\frac{2\pi kx}{L}\,dx, \quad b_k = \frac{2}{L}\int_0^L f(x)\sin\frac{2\pi kx}{L}\,dx. \tag{2}$$

If the given function itself is $L$-periodic, then any interval of length $L$ will suffice. The coefficients $a_0$ and $b_0$ can be reduced to $a_0 = \frac{2}{L}\int_0^L f(x)\,dx$ and $b_0 = 0$, and for the period take the basic period of the sine function, $L = 2\pi$, in order to simplify the expression. Even some of the first members of this development (string $f_n$) are useful, and a particularly common type of approximation in technique.

[75] I quote: "We can find the deeper essence of the information of perception in countless interpretations of such factors, just as we can say that that essence does not really exist."
[76] Ernst Zermelo (1871-1953), German mathematician.
[77] Nikolai Lobachevsky (1792-1856), Russian mathematician.
[78] see [22], 1.3.7 Orthogonality

When the index $n \in \mathbb{N}$ grows indefinitely, then the current item of the Fourier series on a given interval tends to a given function, $f_n(x) \to f(x)$ for $n \to \infty$. But even then, the series may not converge to the exact value of the given function $f(x)$ just for every $x$, say at the points of discontinuity of the initial function. This is again a convenient property of Fourier's development to replace the “bad” function with the “good”.

Using the addition formula for the sine summation, $\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$, putting $c_k^2 = a_k^2 + b_k^2$ и $\beta_k = 2\pi kx/L$ we find:

$$f_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} c_k \left(\frac{a_k}{c_k}\cos\beta_k + \frac{b_k}{c_k}\sin\beta_k\right)$$

$$= \frac{a_0}{2} + \sum_{k=1}^{n} c_k (\sin\alpha_k \cos\beta_k + \cos\alpha_k \sin\beta_k)$$

$$= \tfrac{a_0}{2} + \textstyle\sum_{k=1}^{n} c_k \sin(\alpha_k + \beta_k),$$

where $\sin\alpha_k = \frac{a_k}{c_k}$ and $\cos\alpha_k = \frac{b_k}{c_k}$ is possible substitution when the sum of the squares of the sine and cosine of the same angle is one, which is here. Therefore, the series of the sine gives the Fourier approximation

$$f_n(x) = \tfrac{a_0}{2} + \textstyle\sum_{k=1}^{n} c_k \sin\left(\frac{2\pi kx}{L} + \alpha_k\right), \tag{3}$$

where $\alpha_k$ represent the phase shifts of the basic “angles”. it is also $\lim_{n\to\infty} f_n(x) = f(x)$.

Using the addition formula for the cosine of sums/differences of angles with similar substitutions, an appropriate decomposition of a given function into the Fourier series of cosines is obtained.

Then, using complex numbers and Euler's formula $e^{i\beta} = \cos\beta + i\sin\beta$, that is:

$$\cos\beta = \tfrac{1}{2}\left(e^{i\beta} + e^{-i\beta}\right),\ \ \sin\beta = -\tfrac{i}{2}\left(e^{i\beta} - e^{-i\beta}\right), \tag{4}$$

by substituting to the sine series (3), ie the corresponding cosine series, or directly in (1) we obtain a Fourier approximation using a series of exponential functions. In short

$$f_n(x) = \textstyle\sum_{k=-n}^{n} C_k e^{i2\pi kx/L}, \tag{5}$$

where $C_k$ are the corresponding coefficients. For the sake of shorter writing, there are also added the negative indexes. Further, if we consider $\beta_k$ to be some information (these possibilities have been demonstrated in my previous appendices), then $p_k = e^{-\beta_k}$ are the corresponding[79] probabilities. With complex numbers we are in the domain of quantum mechanics.

In the algebra of quantum mechanics, when we work with "good" functions (without discontinuities) such as mainly those with which we describe natural phenomena, it is shown that fragments of any can be taken to form a Fourier approximation, such as (5), an arbitrary other function. This tells us that the tiny parts of the trajectory of each particle have an arbitrary pre-given shape, which also means that the particles do not have a shape. Each of these possibilities is correct in its own way.

## Action

Information is a quantity (a measure of uncertainty) and because it is a discrete phenomenon, there are its smallest (positive) quantities. On the other hand Planck's constant, $h = 6{,}626 \times 10^{-34}$ m$^2$kg/s approximately, is a quantum of action. That is one of the reasons that information and action travel together.

The smallest amounts of information are the bearers of the purest possible uncertainty. But to have less of the uncertainty means to have more certainty[80], so it makes sense to talk about the "structure of the quantum", if not about the self-sustainability of its parts. In this way, we justify the fact that even the smallest (positive) number can be broken down into (various) factors in countless ways. In addition, it means that it is possible to observe abstract "parts" of information – which are then no longer forms of "pure" uncertainty.

Today, the more well-known factors of action are energy and time, or momentum and position. That energy is not the equivalent of information[81], unlike action, we see in the example of photons (light). There is a long spectrum of frequencies $f = 1/\tau$ of electromagnetic radiation (photons) and with such a free particle-wave there are no necessary restrictions of values. The frequencies are reciprocal with the duration $\tau = 1/f$ of one period, the wavelength $\lambda$, with the movement of the wave, so the speed of the wave is $c = \lambda f$, which is about $c = 300\ 000$ km/s for light in vacuum.

Different frequencies of light determine its different energies $E = hf$. Considering energy itself as the equivalent of information, we would collide with the view that information is finally divisible, and then with the view that there is no point in observing energy without the notion of time. The latter, that there is no change in energy without a change in time or a change in momentum without a change in position, is quite clear from the characteristics of Hamiltonian[82].

The ultimate range of energy possibilities could result[83] in time (duration) and space (length) being a continuum, and the information itself still being a discrete. This conclusion arises from the assumption,

[79] Hartley information$H = -\ln p$
[80] see [1], 3.20 Dichotomy
[81] I emphasize, because commentators often mislay that
[82] see [1], 3.9 Hamiltonian
[83] and whether it is here is not the word

ie the knowledge that the information could have (non-independent and more certain) parts, and the previous consideration should be added to that (16).

## Epilogue

Discussions about the deeper “essence” of information in countless causes, otherwise equally correct as the understanding that these essences do not actually exist, I believe, have just been scratched by this and my previous article of the same name.

# 20. Wavelength II

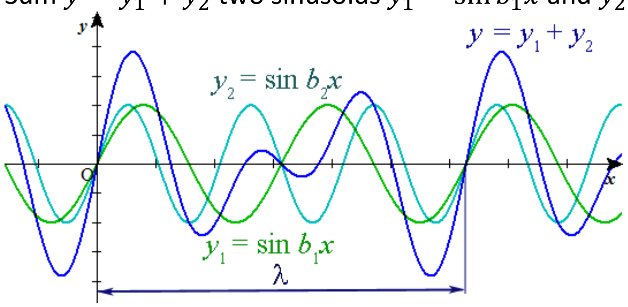
January 5, 2021.

The wavelength is the smallest part of the wave that is constantly repeated. It is a smeared particle of a wave with the size of which its probability density decreases. Greater particle-wave elongation corresponds to greater uncertainty and greater information. We observe this information formally geometrically with additional physical interpretations.

## Sinusoid

A typical wave is a graph of a sine function $y_0(x) = \sin b_0 x$. After increasing the argument $x$ (angles in radians) for its period $\lambda_0$ the whole angle $b_0 x$ is increased by $2\pi$ which is the basic period of the sine. From the equation $b_0(x+\lambda_0) = b_0 x + 2\pi$ follows $\lambda_0 = 2\pi/b_0$. The basic period $\lambda_0$ with respect to the variable $x$ is called the wavelength of a given sinusoid. The higher the wavelength, the smaller the coefficient $b$ (wave number) and the elongated curve, stretched along the abscissa, we say "smeared".

Sum $y = y_1 + y_2$ two sinusoids $y_1 = \sin b_1 x$ and $y_2 = \sin b_2 x$ is on the graph left. The calculation gives:




$$y = \sin b_1 x + \sin b_2 x$$
$$= 2\sin\frac{b_1+b_2}{2}x\cos\frac{b_1-b_2}{2}x,$$

and hence

$$y = a\sin bx.$$

This $a(x) = 2\cos\frac{b_1-b_2}{2}x$ defines the amplitude, the deviation of the point of the graph from the abscissa, and the wave number $b = \frac{b_1+b_2}{2}$ the repetition density of a part of the wave, only one of its aspects and not the wavelength $\lambda$ in the figure, which is there greater than $2\pi/b$.

Note that the amplitudes (deviations from the central axis) are also wave functions $a = a(x)$, except when $b_1 = b_2$, ie when the wavelengths of the given sinusoids are equal. In this special case, when the amplitude is invariable, $a =$ constant, it is twice the amplitude of the each initial sinusoids, and their wavelengths are equal to the resulting wave function.

The previous views are more widely known before, and the following are some novelties. We cannot consider the amplitude only as "wave strength" in the usual sense, because its doubling (in the given example) does not increase the wavelength, which if it represents "smearing" of the wave particle and thus the information, it should have the same "strength". I remind you, with the growth of information of a system or element, its uncertainty and vitality should grow.

In order to fit into the understanding that larger amplitudes of sea waves have higher water power, that larger amplitudes of sound waves determine stronger sounds, larger amplitudes of ground trembling, stronger earthquakes, etc., we can understand amplitudes as prescribed by Born law[84] in quantum mechanics. Transferred to the macro world, the amplitude can still be interpreted as observability

[84] see [22], 1.1.6 Born rule

probabilities, with a slight adjustment. More often ongoings leave some bigger mark, has the more interactions, it is more intensely present in the actuality.

In other cases, $b_1 \neq b_2$, of different wave numbers, and thus of different corresponding wavelengths of the initial sinusoids, $\lambda_k = 2\pi/b_k$ for $k \in \{1,2\}$, the resulting wave function $y = y(x)$ may have a larger wavelength from the sum of the input two $\lambda > \lambda_1 + \lambda_2$). For example, if $b_1 = 1$ and $b_2 = \frac{2}{3}$, then $\lambda_1 = 2\pi$ and $\lambda_2 = 3\pi$, so $\lambda = 6\pi$, which is a number greater than $\lambda_1 + \lambda_2 = 5\pi$.

This possibility, that the total uncertainty of the combined waves is greater or less than the simple sum of the uncertainties of the additions ($\lambda_k$), tells us something more about latent information[85]. Not all the information of our own body is in one present (reality), because it takes some time for light to reach from head to toe – due to the limited speed of light ($c = 300\,000$ km/s approximately), and light defines the "present" of the observer because it does not have its own (proper) flow of time.

## Uncertainty relations

It is impossible to prepare states of a quantum system in which the momentum (with measurement error $\Delta p$) and position (with error $\Delta x$) would be arbitrarily localized at a given direction ($x$-axis). It is possible to organize an approximate measurement of the position and momentum of a quantum particle under the condition of Heisenberg uncertainty relations[86] (product of uncertainty, $\Delta p \cdot \Delta x$, at least of the order of Planck's constant, $h = 6{,}626 \cdot 10^{-34}$ m$^2$ kg / s), so that by increasing the accuracy of momentum measurement, we lose the accuracy of position measurement and vice versa, only with greater momentum uncertainty can we get less position uncertainty.

It can be shown that these limitations are a consequence of a wider impossibility, non-commutativity of the process, that by changing the order of actions we do not always get the same results. It does matter if we firstly lit up the car's signal then turn at the intersection, or we turn first and signal after. Some processes are not commutative and some like multiplication of numbers are. Appropriate uncertainty relations do not apply to commutatives, while non-commutative ones do.

Because quantum mechanics is a representation of Hilbert's abstract algebra, such that vectors represent quantum states and unitary operators represent quantum processes, the noncommutativity of operators (as they are mostly) generalizes Heisenberg's relations into the "uncertainty principle." New evaluation[87] of these generalizations and relations are constant topics, and I have written about many of them[88]. I will not repeat that proves, but I have to mention part of the explanation.

We write relativistic equations (celestial and quantum mechanics) in 4D space-time coordinates so that we denote the abscissa by $x_1 = x$, the ordinate by $x_2 = y$, the applicate by $x_3 = z$, and time by the

[85] 17. Latent Information
[86] see [8], 3.3 Quantum Mechanics, Figure 3.4: Heisenberg's microscope
[87] see [23]
[88] see [22], 1.4.4 Uncertainty principle

imaginary length $x_4 = ict$ which light would pass during $t$. In such a coordinate system, the momentums are $p_1 = p_x$, $p_2 = p_y$, $p_3 = p_z$ with the fourth coordinate energy $p_4 = iE/c$. So are:

$$\hat{p}_n = -i\frac{h}{2\pi}\frac{\partial}{\partial x_n}, \quad \hat{E} = i\frac{h}{2\pi}\frac{\partial}{\partial t}, \tag{1}$$

three momentum operators ($n = 1,2,3$) and an energy operator, while the coordinate operators are simple multiplications by these coordinates.

The vectors on which the operators (1) act are wave functions of the form

$$\psi_k = A[\cos(\omega t - kx) - i\sin(\omega t - kx)] = Ae^{-i(\omega t - kx)} = Ae^{-i\omega t}e^{ikx}, \tag{2}$$

where the probability density $\left|\psi_k(x,t)\right|^2 = A^2$ is uniform and independent of time. The corresponding uncertainty relations ($k = 1,2,3$) are:

$$\Delta p_k \Delta x_k \ge \frac{h}{4\pi}, \quad \Delta E \Delta t \ge \frac{h}{4\pi}. \tag{3}$$

They can be obtained from the non-commutativity of the operators, by calculating:

$$(\hat{p}_k\hat{x}_k - \hat{x}_k\hat{p}_k)\psi_k = \frac{ih}{4\pi}\psi_k,$$

and by removing the wave function $\psi_k$, because the obtained equality holds for each (2).

These are well-known attitudes from before, and the following are new. Consider, for example, the second relation (3), which we can write approximately, in the best case of measurement

$$\Delta E \Delta t = h. \tag{4}$$

On the left side of this equation is the product of the uncertainties of energy and time, on the right is the quantum of action.

Since we consider free information to be quantum, and uncertainties in general to be some (related) information, then right in equation (4) is quantum information. Increasing its factor $\Delta E$ decreases $\Delta t$ so that the product remains constant. By drawing one of the factors into the present, into reality, its certainty increases and the certainty of the other decreases to that extent.

We see similarly in the product of the uncertainty of momentum and position

$$\Delta p \Delta x = h, \tag{5}$$

taken along one, any of the spatial axes. By making (by measuring, interacting with measuring devices) the position of the particle more certain, its momentum becomes more uncertain, because it is not possible to have less than a quantum of information ($h$ – quantum of action) freely.

On the other hand, these $\Delta x$, ie $\Delta t$, are the wavelength or period of the particle-wave, so (4) and (5) talk about information in a way that is the point of this text. A wave is a periodic change in time or space, with periods or wavelengths that represent “blurring”, ie the uncertainty of a wave particle, and thus its information.

With this explanation we supplement the understanding of the wave function (2). The amplitude $A$, or the square of the intensity $|\psi|^2$, defines the probability. We can add that the logarithm of probability defines the information, and because

$$k = \frac{2\pi}{\lambda},\ \omega = 2\pi f,\ f = \frac{h}{p},\ E = hf,\ \hbar = \frac{h}{2\pi},$$

we have another form of the same values of the wave function

$$\psi(x,t) = Ae^{i(px-Et)}. \tag{6}$$

From that form (6) it can be seen that there is an action in the exponent, that it is information, first as a logarithm of probability, and then on the basis of interpretations (4) and (5).

## Epilogue

Information is a measure of uncertainty. Physical information should be a type of “amount of options” that basically matches a mathematical and perhaps a technical definition; Harley's (logarithm of the number of equally probable possibilities) and perhaps Shannon's (mathematical expectation of Hartley's). On the other hand, the law of conservation should apply to it, and that is why Hartley's seems quite fine, while Shannon's is only approximate.

Due to the law of conservation, candidates for “physical information” are seemingly both energy and action (product of energy and time), but further selectivity comes from the requirement that this information be a measure of “itality”. Especially when we notice that each wave carries some information, that a longer wavelength can mean a greater uncertainty of the position of that wave, ie a lower probability density of its place and therefore greater information (more uncertainty – more information).

In the case of water waves, sound and the general wave of a substance, a higher wavelength means a larger amount of the affected substance (higher mass and energy according to $E = mc^2$), higher kinetic energy of particles participating in wave transmission. In this sense, it seems correct to seek the agreement of energy with wavelength and information. Compared to action, in the macro-world of physics, when we observe the motion of waves in equal time intervals (which is then always possible), equal energies can mean the same as equal actions.

In the micro world of physics, that is no longer possible. Think only of a photon (particle-wave of light) whose energy is the product of a quantum of action and frequency ($E = hf$), and whose velocity ($c = \lambda f$) is the product of wavelength ($\lambda$) and frequency ($f$), so the energy is inversely proportional to the wavelength ($E = hc/\lambda$), and still proportional to the action ($h$) and otherwise constant speed of light ($c$). This smallest action remains proportional to the wavelength ($h = E\lambda/c$).

That is one of the reasons why “information” is more an action than an energy. Others, perhaps more important, come from “information perception,” a sub-theory of this information theory, but it’s an even longer story.

## 21. About Parallel Reality

January 5, 2021.

The attachment is an excerpt from one of my older conversations with, say, an anonymous colleague. I quoted his questions and my answers from memory, and I hope that he will not resent if he recognizes himself and eventually discovers that I have exaggerated in some places.

------------------------------------------------------------------------

**Question:** Where is this “parallel reality” in the theory of relativity and where is your “information theory”?

**Answer:** In the special theory of relativity, inertial uniform rectilinear motions, the relative time of a physical system (a body observed in motion) flows slower than its own, proper (observed at rest). I hope we agree with that?

**Q**: Yes, of course. What next??

**A**: From the side of the microworld, the vacuum is full of virtual particles that manage to become real and annihilate in an extremely short duration (annihilate, cancel, and return to the virtual state). Due to the slowing down of time, these temporarily realized particles do not exist in the reality of a relative observer during their short life, but only in the proper reality, of their own.

**Q**: An interesting observation. Are you saying that the physical reality for one of the mutually real subjects may be different than for the other?

**A**: That's right, but there's more. The body of one's own observer (proper, subject A resting besides the body) in relation to the relative (subject B in relation to which A moves) is partly in a different reality, inaccessible to the relative observer. The time of the proper (subject A) therefore flows more slowly from the point of view of the relative (subject B) because part of the time is spent on another reality.

**Q**: As if the proper (A) tramples two time streams at once in relation to the relative (B)?

**A**: It's kind of like that. The total proper flow of time for each of them (either A or B) is the same amount, but each of them does not see some part of the flow of the other. Each calibrates its speed using its own environment, so the own values of both observers are the same, but not relative, which, defined in one system are used to measure something in another. Whenever there is a time-slowing effect, we have a similar result, and that story has a broader context[89].

**Q**: Great. Already in the theory of relativity itself, we have an entrance into the theory of “many worlds” of quantum mechanics (Everett, 1957). As far as I understand, this can cover some more recent theories about different realities of the same in relation to mutually real observers. And where is your information theory?

[89] see [1], 2.13 Space and Time

**A**: In the book Space-Time from 2017, see [17], or some of my even earlier text (complicated books are the consequences of a lot of previous work), at the very beginning you will find the thesis (now I do not consider it a hypothesis) that time and the present are created by realizations events. The fewer of these events are realized, from the point of view of the relative observer, the slower the course of time.

**Q**: The relative velocity of the flow of time is defined by the amount of observed random outcomes?

**A**: Yes. Moreover, the (relative) time of a moving system flows slower (than the proper) by exactly as many virtual events as relative observer does not see as real in relation to the proper.

**Q**: The total amount of random outcomes of a physical system is information measurable by the speed of time?

**A**: This is a presumed consequence of information whose essence is randomness (unpredictability, uncertainty). It is a logical measure of such; information is the intensity of the amount of options.

**Q**: Can the theory of "general uncertainty" explain cause-and-effect laws?

**A**: Of course, it's not just my theory of information that explains[90] certainty with uncertainty, but there's also a theory of probability, and seemingly a "string theory"[91] (unverified) that isn't mine at all but looks like her fugitive sister.

**Q**: What do you see as good in string theory and what do you resent about it?

**A**: I would be lying if I said that I took it more seriously and that my opinion could be of good quality. But I have considered some differential equations about "strings" whose solutions could be very different realities. It is as difficult and inspiring a part of her, encouraging to deal with the "many worlds" of quantum mechanics of Everett, and now information theory. However, those equations pull to another side, their conception is deterministic unlike mine.

**Q**: Are there such immeasurable "realities" elsewhere in classical physics?

**A**: Actually yes. It is not possible to measure the speed of light in only one direction[92], and according to the convention, it is assumed that it is the same in all directions and aims. This cannot be proved or disproved, nor is it even a postulate.

**Q**: How do mathematicians view such alternatives?

**A**: Zermelo's set theory (1908) shows that it is possible to declare the existence of infinity greater than the countable and less than the continuum, and that it is also possible to assume that there is no such infinity. Both theories will be equally correct.

---

[90] see [1], 4.1 Concrete and Abstract

[91] String theory, https://www.britannica.com/science/string-theory

[92] The speed of light cannot be measured, https://www.youtube.com/watch?v=pTn6Ewhb27k

Another example is given by Lobachevsky (1826), who built geometry on Euclid's postulates by turning one to the opposite, the one on parallel lines. He then proved that these two geometries, his hyperbolic and flat Euclidean, are equally true.

**Q**: When we have an independent set of axioms, then we can replace each of them with its negation and obtain equally correct or equally incorrect mathematical theories?

**A**: That would be the definition of “independent axioms”, and Lobachevsky and Zermelo paved the way for us to notice and understand them. Similar ideas appear now in physics, about the existence of realities with which we cannot communicate.

**Q**: What would that mean if this “information theory” were so “both true and false”?

**A**: It would mean that with it we have a nice description of reality that would be irrefutable by physical methods, would fit fantastically well into all known and new interpretations, such that we could not confirm or dispute with any real or thought experiment. Hereon we would notice that the rest of our theories, especially the ones we consider tested, are mostly so alternative, not to say phantom.

**Q**: The truth is not just one?

**A**: When we have a physical phenomenon in practice, or in an experiment, and which we cannot explain by some theory, then we consider it a miracle, a paradox, or a challenge to science until further notice. However, we always expect the future scientific presentation to have mathematical logic, its sharpness and manner. Is it then reasonable to hope that mathematics further discovers alternatives, based on opposite but independent axioms, which are equally correct, without having its reflection in the understanding of physics?

## 22. About parallel reality II

January 10, 2020

Addition the "set method" to the explanation of parallel realities.

### Introduction

Based on the relativistic deceleration of time of system $B$ moving inertial and uniformly in a straight line with $x$-axis by velocity $v$ in relation to system $A$, and due to very short realizations of virtual particles of quantum mechanics in system $B$, which cannot be observed in $A$ due to short life, we conclude that the relative observer ($A$) cannot observe all real proper events ($B$). That is the topic of the previous article of the same name[93].

The relative duration $\Delta t$ of the proper time $\Delta t_0$ is $\Delta t = \Delta t_0 \cdot \gamma$, where is

$$\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}, \tag{1}$$

the so-called Lorentz coefficient, and $c = 299\,792\,458$ m/s is the speed of light in vacuum. Due to the longer relative duration of its own (proper) second, from $B$ time $A$ is observed to be slowed down $1/\gamma$ times.

The units of length ($\Delta x = \Delta x_0/\gamma$) are also so many times shorter in the direction of movement, relative ($\Delta x$) in relation to the proper ($\Delta x_0$), but there is no shortening of units of length in the vertical plane (ordinate and applicate) to the direction of motion (parallel to the abscissa), because there is no motion in that plane. The relative energy in the same proportion with the increase of units of time is greater than its own, the proper ($E = E_0 \cdot \gamma$), and therefore (due to $E = mc^2$) the relative mass of the body is equally increased. Calculations of these relations are well known in physics.

### Sets

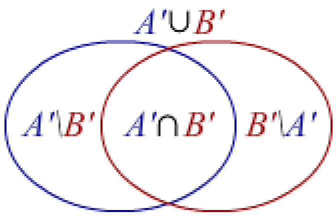
The set of events that the *relative* observer $A$ can perceive from the moving system $B$ is further denoted by $B \to A$, and the information, the amount of these events by $|B \to A|$. It is clear that $B \to B$ should be a set of its own (proper) events, of what the observer or physical system $B$ can notice on itself. Due to the limited speed of light, the larger body exceeds its any own present, so $(B \to B) \subset B$, and of course $(B \to A) \subset (\mathrm{B} \to B)$. Also $|B \to A| < |B \to B| < |B|$.

It follows from the described impossibilities of perception that physical bodies, which can interact (communicate) with each other, do not see exactly the same real phenomena. There is no single reality. This inconceivability defines a new type of "objective coincidence" and with it the additional uncertainty of relative systems, and then their greater information and action.

[93] see https://www.academia.edu/44856563/About_parallel_reality

Greater information of relative systems goes with less entropy. I have written about lower relative entropy on several occasions[94], but now there are new moments. Also, to the earlier interpretation of the increase in relative inertia and body mass by slowing down time, we now add the increase in mass by increasing uncertainty. The novelty is also in the use of the following set equality, otherwise known from Kolmogorov's[95] theory of probability.

When quantity is given by a set and some additive function, such as the Kolmogorov probability (or now information), so that for disjoint sets on the left image, applies:




$$|A' \cup B'| = |A' \backslash B'| + |A' \cap B'| + |B' \backslash A'|,$$

then is $|A'| = |A' \backslash B'| + |A' \cap B'|$ и $|B'| = |A' \cap B'| + |B' \backslash A'|$, so

$$|A' \cup B'| = |A'| + |B'| - |A' \cap B'|.$$

Putting $A' = B \to A$ and $B' = A \to B$, therefore, by replacing the given sets with the above-mentioned relative observations, we obtain equality

$$|(B \to A) \cup (A \to B)| = |B \to A| + |A \to B| - |(B \to A) \cap (A \to B)|. \quad (2)$$

The interpretation is unexpected at first glance. The union of mutual relative observations of two systems gives less than a simple sum. This reduction is all the greater as the relative observations are closer, the more they have in common, when with more of the same their union loses on unpredictability and hence on overall information.

## Outer space

We know that places (galaxies) in space move away the faster they are farther away. With a greater distance, the light from them travels to us longer, so we actually look further and further into their past. These are already two reasons that increase the relative uncertainty. First, the relativistic slowing down of time in proportion to the Lorentz factor (1) and then the aforementioned formalism of sets with equation (2) – can be criticized in an unexpected way.

With the greater distance of galaxy $B$ from our galaxy, us at place $A$, the higher the velocity $v$ and the higher the coefficient $\gamma$, so we should observe a relatively slower flow of time of system $B$ than its own is (proper). I speak with reserve, because there is a possibility that our, the proper time (actually anywhere in the universe) is slower than an arbitrary event fixed in the history of the cosmos, so that the course of time $B$ observed from $A$ would not have to go with a factor (1).

A different uncertainty concerns formula (2). Namely, due to the limited speed of light by which information is transmitted, ie the symmetry of space and time ($x_4 = ict$), just as the outcome of future events is uncertain, so is the unfolding of distant states uncertain. In any case, whatever it is, each of the uncertainties, (1) or (2), takes object $B$ partly into some parallel reality in relation to $A$ in such a way that these two alienations complement each other.

## Epilogue

[94] see [1]

[95] Andrey Kolmogorov (1903-1987), Russian mathematician.

This is just a hint, a note about the “method of sets” by which I sometimes (rarely) speed up or clarify private thoughts in information theory, and which I promised my colleague that he might elaborate on. Some of that I belike bring out later.

# 23. Action and Information

January 12, 2020

You say information is waves, then information is action, um, ... vague[96]. How that?

I will explain in a few, I hope, simple steps. First, Emmy Noether[97] theorem on the law of conservation in physics is derived from Euler-Lagrange's equations of motion. The second is about the smallest pieces of information, and the rest are examples.

## Conservation law

First, if symmetry applies then the law of conservation applies and vice versa (Noether, 1915). To paraphrase: If a system has a continuous symmetry property (so that Euler-Lagrange differential equations can be valid), then there are corresponding quantities whose values are conserved in time[98]. Or, a little more professionally: To every differentiable symmetry generated by local actions there corresponds a conserved current.

In geometry, symmetries are "isometric" (Greek: *isos-metron* = same measures) transformations, those that preserve the distance of points. There are not many of these "immutabilities" (mirror, axial and central symmetry, translation and rotation) and everyone can be reduced to rotations. They are the basis for the use of this theorem in physics.

For example, the immutability of a physical system to spatial displacement (translation) gives the law of conservation of a linear (ordinary, $p$) momentum. Invariance in relation to rotation is given by the law of conservation of rotational (angular) momentum.

Noether's theorem is difficult to prove, but to this day it is widely accepted in physics.

## Discrete sets

Second, if the law of conservation applies then the phenomenon is discrete. They are discrete sets with the smallest but larger than zero parts, of which there can be infinitely many. This seems to contradict the first, due to the requirement of differentiability, but it is not because the "Noether theorem" is broader than the differential equations from which it is derived.

Proof of information discretion is simpler, but it is unknown and at first glance absurd. It follows, for example, from the definition of infinity, here infinite divisibility, that infinite sets are those and only those that can be their proper (real) subsets (the first is all in the second, and the second has some more elements). For example, the set of natural numbers $\mathbb{N}$ = {1,2,3, ...} and the set of integers $\mathbb{Z}$ = {0,+1,-1,+2,-2, ...} are infinite, and the first is real subset of another, $\mathbb{N} \subset \mathbb{Z}$.

[96] A colleague, not from the profession, asked me.
[97] see [1], 1.14 Emmy Noether
[98] see [24]

Here are the absurdities of that theorem. For example, energy is not quantized, because from the known radiation relation $E = hf$ (energy $E$ is the product of Planck's constant $h = 6{,}626 \times 10^{-34}$ $m^2kg/s$ and the frequency of electromagnetic radiation $f = 1/T$, where $T$ is period of one oscillation) does not follow that there are some smallest frequencies $f$, ie the longest period $T$, but only that the action $h = ET$ is quantized.

Therefore, first of all, the question arises where does the "law of conservation of energy" come from? The answer is in the macro-world, where we define "energy of the system" as a quantity during a constant given time interval. The action is invariant, but then the energy is also invariant. If we imagine, on the contrary, a steam engine (the law of conservation of energy is actually the "first law of thermodynamics", later generalized to other fields of physics) that produces energy, but we do not define the duration of that production, then there is no law of conservation of energy.

So, it is not the "law of conservation of energy" that follows from Noether's theorem, but it is Planck's radiation law[99]. Quantum (product of energy and duration) is the smallest portion of a substance that can exist independently and which we call a particle.

Third. I am extensive, but without this step the information (my theories) cannot be understood. Quantum systems (physical structure, from an elementary particle to their largest set of them) are representations of vectors (abstract Hilbert algebra), and their changes, the so-called quantum evolution, or processes, are unitary operators. The point is that these operators are reversible, they remember the originals (the image can be reconstructed from the image), which means that information is not lost in quantum processes. Quantum evolutions are types of symmetry (translation in time or space), so we have the law of conservation, and then discreteness.

Note that in this derivation (proof) one can go directly to discretion and avoid the problem with differential equations and the application of Noether's theorem, but this is not necessary. Namely, if in Planck's formula energy, i.e. time (periods of frequencies), and similarly length are not quantized, then we have no problem with space-time continuity, nor with the validity of Euler-Lagrange equations.

Action is quantized and information is quantized, and further it is only necessary to show that there is "at least some action" that transmits "at least some information" (which is not disputable), from which it follows that actions and information are equivalent. These are sizes that travel together in such a way as to form the so-called isomorphism[100] (in abstract algebra).

## Waves

Finally, why are there waves? The answer is in the hypothesis of Louis de Broglie (1924) and later in Schrödinger's equation (1925), from which it follows that all matter is of a wave nature[101]. Actions are quantized, so all matter is in "particles", and as all its properties are described by the wave equation

[99] see https://www.britannica.com/science/Plancks-radiation-law
[100] Isomorphism in mathematics is a two-sided unambiguous mapping.
[101] see [1], 3.5 Quantum States and Processes

(Schrödinger), it is in the packets that we call "particles-waves". The information is also quantized, so it, any free information, is in those particles-waves.

The velocity of the wave ($v = f\lambda$, where the frequency is $f$, and $\lambda$ is the wavelength) tells us something about the mass ($m$) of the particles. When it is the speed of light ($v = c =$ 299 792 458 m/s), the wave particles do not have a rest mass, their time stands still and all their content is in the time of the observer. In that sense, they do not have additional uncertainties, and their frequency is an indicator of "kind of liveliness". There is no further, because, remember, information is the amount of data and does not deal with their type.

When the speed of the waves is less than the speed of light ($v < c$), they have a mass (of rest), their time flows and their content is not all in the time of the observer. The opposite is also true, when a particle has its own (proper) flow of time, it is partly in parallel reality (in relation to any given observer) and due to the principle of minimalism, it gets inertia, so it has its own mass (proper, in rest). Its speed is not the speed of light, because its relative mass would then be infinite.

De Broglie's wavelength, $\lambda = h/p = h/mv$, associated with the mass $m$ of the particle in relation to the momentum $p$ and Planck's constant, was first confirmed experimentally by Thomson[102], for which he received the Nobel Prize in 1937, and independently the same was done by Davisson–Germer experiment (1923-1927), in both cases on electrons. Hence, $h = \lambda p$, which means that the action ($h$) is proportional to the wavelength and momentum, and in that sense to energy and mass. Therefore, we can say that the information is proportional to the wavelength and momentum.

[102] George Paget Thomson (1892-1975), English physicist.

## 24. Bernoulli's Attraction II

January 14, 2021

This is a continuation of the eponymous text [25] and similar texts on the connection between Bernoulli's law, quantum mechanics and gravity, now with an emphasis on the principled minimalism of information.

### Introduction

That there is a connection between "information of perception" and Bernoulli's law I wrote interpreting entropy (August 2016, [26], pp. 10-11), or space-time (May 18, 2017, [17], pp. 76-77). Shortly afterwards, Bernoulli's attraction was observed especially in quantum mechanics (May 30, 2017, [27]), or more recently in gravity and an attempt to explain "dark matter" (January 2021, [28]).

These days, the topic was imposed on me again by acquaintances (they don't want me to mention them, they are not from the profession), hoping for additional explanations, primarily fluid dynamics from the point of view of (my) "information theory". I believe that these attempts are interesting, and perhaps correct, and that these contributions have something in common that can be connected by the "information principle". That's why I answer.

### Time slowdown

The relative speed of the time flow is defined by the amount of realized random events. This is a (working) idea that I follow from the period before writing the book "Space-Time" [17] and it is reasonable to expect me to try it again. I do not run away from extreme applications.

The deceleration of the time range is observed on a body moving inertially at a uniform velocity $v$, proportional to the Lorentz coefficient

$$\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}, \tag{1}$$

where approximately $c = 300\ 000$ km/s speed of light in vacuum. This is known from the special theory of relativity. From the general theory, in the Schwarzschild metric, of centrally symmetric gravitational fields, we have a similar deceleration where in the given gamma coefficient we should use the substitution $v^2 = 2GM/r$, where $G = 6{,}674 \times 10^{-11}$ $\text{m}^3\ \text{kg}^{-1}\ \text{s}^{-2}$ gravitational constant, $M$ is the mass of the body that produces gravity, and $r$ is the distance of the body that gravity acts from to the center of the force.

Slightly different examples are given by centripetal and centrifugal force as a pair of action-reaction forces associated with circular motion. When a body moves in a circle of radius $r$ with speed $v$, then it has centripetal acceleration

$$a = \frac{v^2}{r}. \tag{2}$$

It is directed from the center, trying to move the body away by the force $F = ma$, where $m$ is the mass of the given body. The elapsed self-time (proper time) of the body $\Delta t_0$ is relatively observed by the observer at rest

$$\Delta t = \Delta t_0 \cdot \gamma, \tag{3}$$

which means the relative slowing down of the time course, the proper multiplied by $1/\gamma$, ie so many times less amount of relatively observed random events related to a given body. This deficit of time comes from the presence of parts of one's own (proper) events in parallel reality that are not visible to the relative observer, about which I have written several times (see [29]).

## Principled attraction

Due to the principle of minimalism of information, according to which the physical system spontaneously tends to a state of less information, there will be an attractive force that pulls the body from a system with a faster to a system with a slower time course (if some other force doesn't stop it). In other words, the body will spontaneously try to move into a system with a smaller amount of realizations of random events. All gravitational pulls and all centrifugal forces can be reduced to this principle.

We will understand with the help of the following "experiment" that the centrifugal repulsive force can be reduced to gravitational attraction by this principle. Imagine the inverse situation of the rotation of a rigid body, where in a plane we have ever faster rotations of points in space that are getting closer to the center. More generally, imagine the "movement" of points in space whose velocity increases as we approach a given center. Consistent with the mentioned interpretation, then an inverse of the centripetal acceleration will occur, the one that like gravitational pulls the body towards the center.

Consistent with the same interpretation, Bernoulli's attraction arises. The total mechanical energy of a fluid (gases or liquids) exists in two forms: potential and kinetic[103]. The kinetic energy of the fluid is stored in the static pressure $P$ and dynamic pressure $\frac{1}{2}\rho v^2$, where $\rho$ is the density of the fluid (in SI units: kg/m$^3$) and $v$ is the velocity of the fluid (in SI units: m/s). The unit of the SI system for static and dynamic pressure is "pascal". Bernoulli's equation is

$$P + \frac{1}{2}\rho v^2 = \text{constant.} \tag{4}$$

Static pressure ($P$) is that at a given point of fluid, and dynamic ($\frac{1}{2}\rho v^2$) is the kinetic energy per unit volume of fluid particles. The fluid has no dynamic pressure when it is not moving. When there is no change in potential energy along the flow, Bernoulli's equation (4) assumes that the total energy along the flow is constant and expresses the balance between static and dynamic pressure. It expresses the pressure along the current.

[103] Next I follow the text from the textbook of fluid dynamics.

If there are significant changes in height or if the density of the liquid is high, the changes in potential energy cannot be ignored and the addition of $\rho gh$ should be considered. Then Bernoulli's equation reads

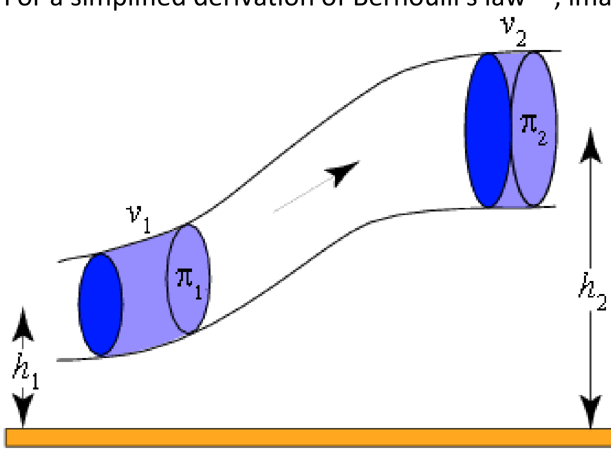
$$P+\frac{1}{2}\rho v^2+\rho gh=\text{constant}, \tag{5}$$

where $g=9{,}8\ \mathrm{m/s^2}$ the gravitational acceleration of the earth weighs approximately, and $h$ is the height (depth) to which the fluid climbs (descends).

## Derivation of the equation

Classically, Bernoulli's equation is derived by integrating Newton's second law along the current with gravitational and compressive forces on the fluid. Since any energy exchange comes from conservative forces, the total energy along the current is constant and is easily replaced between potential and kinetic.

For a simplified derivation of Bernoulli's law[104], imagine a pipe through which an ideal flows of fluid at a constant velocity, pictured left. Let the work $W$ be done by pressing $P$ on the surface π, which produces a shift $\Delta l$ or a change in the volume $\Delta V$. Indices 1 and 2 indicate the start and end position of the fluid in the pipe. The work done by the force of pressure is




$$dW=PdV.$$

The work is at the points of the index:

$$\Delta W_1=P_1\pi_1\Delta l_1=P_1\Delta V$$
$$\Delta W_2=P_2\pi_2\Delta l_2=P_2\Delta V$$

and the difference between these values is

$$\Delta W=\Delta W_1-\Delta W_2=P_1\Delta V-P_2\Delta V.$$

By equating this change with the change in total energy (the sum of kinetic $K$ and potential $U$) we get in order:

$$\Delta W=\Delta K+\Delta U=\left(\frac{1}{2}mv_2^2-\frac{1}{2}mv_1^2\right)+(mgh_2-mgh_1)=P_1\Delta V-P_2\Delta V,$$

where we equate the current and previous expressions. Next is:

$$\frac{\Delta mv_1^2}{2\Delta V}+\frac{\Delta mgh_1}{\Delta V}+P_1=\frac{\Delta mv_2^2}{2\Delta V}+\frac{\Delta mgh_2}{\Delta V}+P_2,$$

so due to the definition of density, $\rho=m/V$, we $\frac{1}{2}\rho v^2+\rho gh+P=$ constant, which is Bernoulli's equation (5).

From these derivations ("classical" and "simplified"), we can easily recognize the previously mentioned "principled attraction", I hope. And that is that.

---

[104] see https://scienceworld.wolfram.com/physics/BernoullisLaw.html

# 25. Multiplicity of Explanations

The truth is not just one?

January 16, 2021

Not every mathematical model is for every situation, but different interpretations of the same exist.

## Introduction

Maybe action (energy over time) and information are not just “equivalents” but are exactly the same, but I do not do their “separation” only because of “excessive caution” (I answer an interesting question asked about the essence of information).

Honestly, after several years, I'm still not quite sure. The main reason for the suspicion is of a very theoretical nature, seemingly unrelated to information stories, I will explain. After Lobachevsky[105] in his book “Geometry” (1823) studied geometry without “V Euclid's postulate” (on parallel lines), later turning it in the opposite meaning, to gradually reach the “hyperbolic geometry”, which we call today , and finally to the proof that his new geometry and Euclidean are both equally correct, or what is the same, that both are equally incorrect. For some of us he opened Pandora's Box of mathematics.

The proof that was unthinkable until Lobachevski, incomprehensible that such a thing could exist at all, gave us a completely new view of the truth, and then of reality. It is a possibility that there are independent and different theories, equally accurate and equally real!

Before I say that such two theories might be “physics of action” and “physics of information,” let us consider a different pair of examples for ease of abstraction later. It will be “free networks” in the way (A) as researched by the American-Hungarian mathematician Barabási[106] (since 2000) with the help of probability, and the same in the way (B) with the help of what we can call “attractions”. Invent the models yourself, and I list one pair (A, B) as a guide. The network consists of “nodes” and “links”, and is free because the links are equal.

**A**. By adding new links (from new or old nodes) the network grows. However, if some nodes have more links than others, then they are more likely to get a new one. Nodes with multiple connections grow spontaneously and the “free network”, precisely because of the equality of connections, becomes a network of unequal nodes. There are a proportionally smaller number of intersections (concentrators) with an increasing number of roads compared to a very large number of intersections with little roads. Such is the situation with the free market (equality in the flow of money, goods and services) in which a smaller number of the very rich versus a large number of the very poor will spontaneously stand out (in proportion to the number of participants).

**B**. Instead of equal connections and the probability of nodes, let's assume that with the growing number of connections of the node, its (some imaginary) attractive force grows, which pulls new connections to

[105] Nikolai Lobachevsky (1792-1856), Russian mathematician.
[106] https://barabasi.com/publications/22/ten-most-cited

itself. It is clear that the intensity of the force can be calibrated so that we get exactly the same results with the previous explanation (method A) in all possible examples and applications of free networks. For example, in the free market, we will say that it is more worthwhile to open a new connection (make a new deal) with an experienced, proven and rich company, than with some unknown newcomer there. And that's what's happening in the business world.

The dilemma gets complicated, doesn't it? It is not possible to say in advance that model A is better than model B in the sense of "correct interpretation of reality", and Lobachevski's work deepens such doubts. It is clear that not every interpretation is good, moreover, we will almost certainly not find the right one by blindly search, but there is still the possibility that there are more equally correct interpretations of the world.

This is the situation because of which I say that information and action are "equivalents", and not "one and the same". This is a continuation of the conversation mentioned in the note "23. Action and information".

## Vector spaces

Vectors are $n$-tuples of numbers, such as $\mathbf{a} = (a_1, a_2, \dots, a_n)$ and $\mathbf{b} = (b_1, b_2, \dots, b_n)$, where the natural number $n = 1,2,3, \dots$ is arbitrary but fixed, and whose corresponding components, at the same positions, are added to vectors of the same type, $\mathbf{c} = \mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) = (c_1, c_2, \dots, c_n)$. The multiplication of numbers and vectors is given by the equation $\lambda\mathbf{c} = (\lambda c_1, \lambda c_2, \dots, \lambda c_n)$ when the multiplier $\lambda$ is called a scalar. The coordinates of the points of the $n$-dimension system are vectors, oriented lines.

In algebra, it is common to define a vector space with a few simple axioms[107], but this can also be done using their consequences. From the property of the exact number of components of the vector of a given vector space and the independence of individual components in the vectors addition law, it turns out that the sum of $m$ vectors, $\mathbf{x}_k = (\xi_{k1}, \xi_{k2}, \dots, \xi_{kn})$ for $k = 1,2, \dots, m$, in general form the vector

$$\mathbf{y} = \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \dots + \lambda_m \mathbf{x}_m \tag{1}$$

where $\lambda_k$ are arbitrary scalars. As $\mathbf{y} = (\eta_1, \eta_2, \dots, \eta_n)$ it is

$$(\eta_1, \eta_2, \dots, \eta_n) = (\lambda_1 \xi_{11} + \lambda_2 \xi_{21} + \dots + \lambda_m \xi_{m1}) +$$

$$+(\lambda_1 \xi_{12} + \lambda_2 \xi_{22} + \dots + \lambda_m \xi_{m2}) + \dots + (\lambda_1 \xi_{1m} + \lambda_2 \xi_{2m} + \dots + \lambda_m \xi_{mm}),$$

and hence the system of $n \times m$ linear equations

$$\begin{cases} \lambda_1 \xi_{11} + \lambda_2 \xi_{21} + \dots + \lambda_m \xi_{m1} = \eta_1 \\ \lambda_1 \xi_{12} + \lambda_2 \xi_{22} + \dots + \lambda_m \xi_{m2} = \eta_2 \\ \dots \\ \lambda_1 \xi_{1n} + \lambda_2 \xi_{2n} + \dots + \lambda_m \xi_{mn} = \eta_n \end{cases} \tag{2}$$

[107] see [22], Definition 1.1.9 (Vector space), p. 17

by unknown $\lambda_k$.

We know that for $n > m$, when there are too many equations in relation to the number of unknowns and the system may be in contradiction, that there does not have to be a series of scalars $\lambda_k$ that would be solution (2), while for $n = m$ that sequence exists and is unique – when the equations are linearly independent. When $n < m$, the system (2) has several (countless) solutions. Consistently, we say that the vectors $\mathbf{x}_k$ are linearly independent if the vector equation

$$\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \cdots + \lambda_m\mathbf{x}_m = 0, \tag{3}$$

has only a trivial solution $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0$. The set of $n$ linearly independent vectors is called the base of an $n$-dimensional vector space.

Comparing (2) with (3), a system of ordinary linear equations with vector's, we find that any two bases of the same vector space (same dimensions $n$) have the same number of vectors. With systems of linear equations, such as (2), we transform one base into another. Then there is always an inverse transformation, a linear system like (2) allows the vectors $\mathbf{x}_k$ of one base to be computed using the vector $\mathbf{y}_j$ of the other base. We say that a given system is regular, or invertible, that the transformation remembers vectors, when we can get originals from copies.

We can reformulate the linear system (2) into a matrix one

$$\hat{A}\vec{x} = \vec{y}, \tag{4}$$

where the matrix $\hat{A} = (a_{ij})$ is quadratic of type $n \times n$, and the vectors it transforms are $\vec{x} = (\xi_k)$ and $\vec{y} = (\xi_k)$, both with $n$ components. The corresponding system of equations is then

$$\begin{cases} a_{11}\xi_1 + \cdots + a_{1n}\xi_n = \eta_1 \\ \qquad\qquad \cdots \\ a_{n1}\xi_1 + \cdots + a_{nn}\xi_n = \eta_n \end{cases} \tag{5}$$

and it can transform base to base if it is regular. If the matrix $\hat{A}$ is regular, then there exists an inverse matrix $\hat{A}^{-1}$ such that $\hat{A}^{-1}\,\hat{A} = \hat{A}\hat{A}^{-1} = \hat{I}$, where $\hat{I}$, is unit matrix, which has ones on the main diagonal and all other coefficients are zero. Multiplying (4) by the inverse matrix on the left, we get

$$\vec{x} = \hat{A}^{-1}\vec{y}, \tag{6}$$

hence the inverse transformation of the vector.

These are known things of linear algebra, and I quote them just to remind you that square matrices of order $n$, in place of the vector in expression (3), do not have to be more than $n$. Therefore, regular matrices also form a (new) vector space with the same number of base vectors as the vectors that transform these matrices. We call it the dual vector space of the vectors on which they act.

We can also reformulate the matrix system (4) into vector transformations by linear operators, and then we can (almost always) reduce these operator equations to matrix ones. The resulting matrices are then called matrix representations of the starting operators, and what is important to us, the both are types of vector spaces. Moreover, they are vector spaces of the same number base vectors and they are isomorphic in that sense: there are mutually unambiguous mappings into each other.

Matrix (4), among other things, can be multiplied by itself and form an equation

$$\beta_0 \hat{I} + \beta_1 \hat{A} + \beta_2 \hat{A}^2 + \dots + \beta_m \hat{A}^m = 0, \tag{7}$$

which for some $m = 0,1,2,\dots,n$ and numbers $\beta_k$ which are not all equal to zero must be correct, because all powers of a square matrix are again square matrices of the same order, so sooner or later, a non-trivial solution appears analogously (3 ). Therefore, the set of exponents of a regular square matrix also constitutes a vector space.

Similarly to (7), polynomials of the same degree $n$ form a vector space as matrices of order $n+1$, i.e. isomorphic to any vector space with $n+1$ base vectors. Indeed, a polynomial

$$f_n(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_n x^n \tag{8}$$

we can represent only by its coefficients, $n+1$-tuples $(\beta_0, \beta_1, \beta_2, \dots, \beta_n)$, because two polynomials are identically equal when all their corresponding coefficients are equal. We define the addition of polynomials in the usual way (we add the coefficients of the same exponent) and we also have the correct usual axiomatic of the vector.

These examples are instructive in themselves. We can declare very different mathematical entities as vectors and obtain the same theorems of vector spaces. When we do not go deeper into the "essence" of these entities, we will not notice their mutual differences, but we will not have any obstacles with their so impoverished logic. As in the nature of the universe of information is the non-communication of everything with everything, the uniqueness and multiplicity of phenomena, so finding logical such "unfinished" entities (separated from the "essence") is possible in their world.

## Quantum interpretations

Quantum mechanics is a representation of (Hilbert's) vector spaces. A vector space is a quantum system (a group of particles), a vector is a quantum state (of a given system), and an operator is a quantum evolution (state change process). The "numbers" are complex numbers, and the product of the corresponding pair of conjugate complex coefficients ($p_k = \xi_k^* \xi_k = |\xi_k|^2$) represents the amplitude of the $k$ –th coordinate of the given vector, $\mathbf{x} = (\xi_k)$, i.e. the probability of observable (measurable physical quantities) represented by that coordinate.

In other words, we choose the coordinate axes so that they are observable, and that the projections of quantum states on them define the probability of finding the state in the measurement. That is why

$$|\mathbf{x}|^2 = |\xi_1|^2 + |\xi_2|^2 + \dots + |\xi_n|^2 = 1. \tag{9}$$

Only vectors of unit norms and, therefore, only unit (unitary) operators are interpreted. The summands in expression (9) represent the probabilities of individual outcomes, so each quantum vector represents some distribution of observables probabilities, which we call superposition. When a measurement occurs, the superposition of the state then collapses into some of the possible outcomes.

Operators on vectors, now quantum processes on quantum states, are also vectors. They form a dual vector space with the quantum system on which they act. In that sense, the formal laws of process and state are equivalent. There is an isomorphism between the phenomena that these vectors represent.

In addition to the dualism between quantum particles and the quantum evolutions that can occur with them, the two most famous representations of quantum mechanics are matrix (Heisenberg, Born, Jordan) and wave (Louis de Broglie, Schrödinger), both discovered around 1925. The first comes from the already mentioned vector space of matrices, and the second from the vector space which consists of solutions of the Schrödinger equation. Namely, it is a wave equation, a differential equation whose sum of solutions is also its solution, and because of the structure itself and because otherwise the derivative of the sum is equal to the sum of the derivatives – its solutions form a vector space.

Two quantum mechanics were a precursor to the discovery of many later. What they have in common are interpretations of universal attitudes of abstract vector spaces and an extraordinary coincidence of theoretical prediction with experimental findings, hitherto unseen accuracy in physics.

## Living world

The diversity of flora and fauna on Earth is another confirmation of the multiplicity of universal information. Not every model of interpretation is suitable for many situations, but for any given situation, different theoretical models are possible by which it can (very, but not extremely, absolutely) be accurately explained. The only questions are whether we can and will succeed in finding those models, and above all the situations.

# 26. Gravitation Multiplicity

About good theories of gravity

January 19, 2021

There are alternative theories of gravity and although seemingly opposed, some of them could prove to be equal with minor fixes and additions. This is an article that was created together with the previous one[108].

## Force

The concept of classical force is in crisis[109] in today's physics, but I believe that it can be improved and useful in the future. Newton introduced force into physics with his theory of gravitation[110]. Somewhat mystical, Newton accepted the assumption of some unlimited gravitational action between celestial bodies. Without knowledge of electrical phenomena and Faraday's cage, for example, the alleged force could penetrate all obstacles instantly, without even noticing them.

He was an excellent mathematician and knew how to calculate from Kepler's laws that the attractive gravitational force ($F$) decreases with the square of the distance ($r$) between the centers of the bodies (mass $M$ and $m$), according to the formula

$$F = G\frac{Mm}{r^2}, \tag{1}$$

where approximately $G = 6{,}67 \times 10^{-11}$ N m$^2$ kg$^{-2}$ is the universal gravitational constant. It is further known that Einstein (1905-15) demolished (repaired) such a notion in a way that I will now try to retell with as few repetitions as possible.

Actions in physics do not go faster than light in vacuum, $c = 300\,000$ km/s approximately, and that is the main problem of "instantaneous gravity transmission". Additionally, Lorentz transformations (Einstein's special theories of relativity) apply to inertial systems and I take them with caution[111].

Imagine an accelerating rocket, with its own (proper) observer inside and relative on the ground. If the proper finds a constant acceleration of the rocket, the relative notices its decrease, and vice versa, if the acceleration of the rocket for the relative observer is constant, the proper will measure the increase – so as not to exceed the speed of light. At velocity $v$, the relative mass increases in proportion to the coefficient

$$\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}, \tag{2}$$

[108] see 25. Multiplicity of Explanations

[109] see [2], 2.19 Classical Force

[110] Isaac Newton: Philosophiae naturalis principia mathematica, 1687.

[111] see [17], 1.1.8 Force

and with it the same slows down the time and shortens lengths in the direction of movement. A constant relative acceleration requires an increase in the power consumption of the rocket propulsion.

The relativistic differences of forces are visible from the basic definitions. The first type of force is equal to the product of mass and acceleration ($F = ma$), and the second to the change in momentum over time ($F = dp/dt$). As the momentum is the product of mass and velocity ($p = mv$) and the change in velocity over time is acceleration ($a = dv/dt$), in order for the two types of forces to be equal, it is sufficient that the mass does not change with time ($dm/dt = 0$). In the motion of the rocket, Newton's physics, unlike Einstein's, does not predict such a change in mass.

When differences in definitions and perceptions of forces are accepted, their improved concept could be useful. Even if it is approximately correct, i.e. moderately inaccurate like other areas of physics, it would have its value.

For example, cosmology tells us that galaxies (on average) are moving away from each other like points on a balloon we are inflating, and are moving away faster and faster. What astronomers actually see is the light from those galaxies that took billions of years to date in the state of those galaxies from the past on the beginning of the journey. Holding that the force of mass gives acceleration ($F = ma$) and that the force $F$ on the path $dr$ gives work, or energy $dE = Fdr$, we conclude that the observed energies of galaxies grow, that is, that our galaxy from their point of view is relatively higher energy[112].

The problem of Mercury's deviation from the elliptical orbit around the Sun, the movement of the perihelion in the direction of rotation, is not so much a problem of the general definition of force as of understanding gravity itself.

## Space curvature

How can we know that a line is really “straight”, can it be said that it consists of “shortest paths” between points and is there any invariant by which a flat Euclidean space would be clearly distinguished from a non-Euclidean one? This is how Gauss[113] thought about geometries when he discovered his famous method for defining “curved” surfaces (Theorema Egregium, 1827). He also performed measurements on local hills in Germany, finding his curvature too small, if at all.

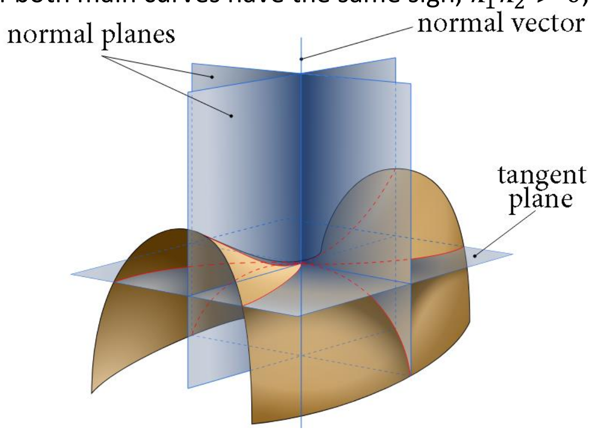
At an arbitrary point on a given surface, he placed a normal (vertical) vector on a tangent plane, at right angles to the surface. He called the planes containing that vector *normal planes*, and the intersection of a normal plane and a surface is a curved line called a *normal intersection*. For example, the curvature of that intersection $\kappa$, the so-called *normal curvature*, is the reciprocal radius of the circle given by its part at the base of the normal. A sphere of radius $r$ everywhere has a curve of normal $\kappa = 1/r$.

For most points on many surfaces, different normal parts have different curves. Their maximum and minimum values are the *main curvatures*, here $\kappa_1$ and $\kappa_2$, and the *Gaussian curvature* is the product of the two main curvatures $K = \kappa_1\kappa_2$. A sphere of radius $r$ everywhere has a Gaussian curve $K = 1/r^2$. In

[112] I state, because as a consequence, I have get it from the theory of information.
[113] Carl Friedrich Gauss (1777-1855), German mathematician.

the figure on the left[114], the saddle surface has the main curvatures of opposite directions, so its Gaussian curve is a negative number.

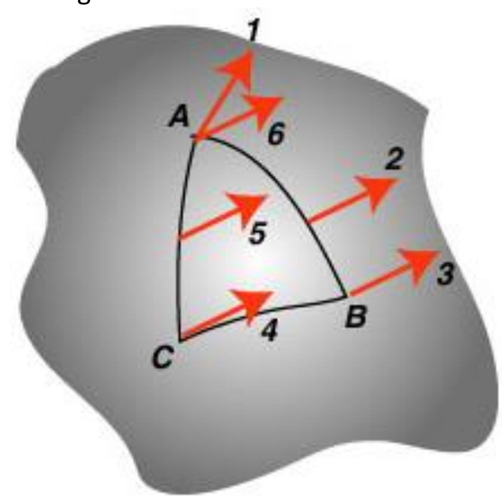
If both main curves have the same sign, $\kappa_1\kappa_2 > 0$, then the Gaussian curve is positive and it is said that the surface has an elliptical point. If the main curves have different signs, $\kappa_1\kappa_2 < 0$, then the Gaussian curve is negative and the surface is said to have a hyperbolic or saddle point. If one of the principal curves is zero, $\kappa_1\kappa_2 = 0$, the Gaussian curve is zero and the surface is said to have a parabolic point.




When a curved surface develops on any other surface, the measure of curvature at each point remains unchanged. In that sense the Gaussian curvature is the internal invariant of a surface.

For example, a sheet of paper cannot be bent into a sphere without creasing, that is, the surface of the planet Earth cannot be projected on a flat map by isometry (preserving the distance between points). The (lateral) surface of the cylinder has zero Gaussian curvature, because one main curvature comes from a circle of finite radius, but the other is a straight line – the derivative of the cylinder which is a circle of infinite radius, and $1/r \to 0$ when $r \to \infty$.

Gauss's student Riemann[115] continued that work. The figure on the right shows one curved surface and the normal vector (red) that is translated, moves in parallel along the curvilinear triangle ABC (black line) on the surface. The vector slides through positions 1-2-3-4-5-6, from A over vertices B and C back to A. However, when it returns to the starting position (vertices A) it does not match its initial value[116].




Riemann calculated the differences of the finite vectors, 1 and 6, and with the obtained values $R^{\rho}_{\sigma\mu\nu}$ defined the local curve of the infinitesimal place (triangle ABC), today the so-called *Riman tensor*. He established special interesting symmetries and algebraic forms which were further discovered by Christoffel[117], Bianchi[118], Ricci[119] and others.

In differential geometry, the Ricci curvature tensor $R_{ij}$ is a set of quantities obtained by choosing the Riemann or pseudo-Riemann metric on the so-called manifolds. This tensor can be considered a

[114] Saddle surface, https://en.wikipedia.org/wiki/Gaussian_curvature
[115] Bernhard Riemann (1826-1866), German mathematician.
[116] see [1], 2.13 Space and Time
[117] Elwin Bruno Christoffel (1829-1900), German mathematician and physicist.
[118] Luigi Bianchi (1856-1928), Italian mathematician.
[119] Gregorio Ricci-Curbastro (1856-1925), Italian mathematician.

measure of the degree to which the geometry of a given metric tensor differs locally from the geometry of ordinary Euclidean space or pseudo-Euclidean space.

Einstein continued this work to obtain his general theory of relativity (1915). In the internal invariance of curved spaces we recognize the property of energy[120] (law of conservation), so the mass that creates the gravitational field and, like he who saw them in inertia too, to accept its famous equations

$$G_{\mu\nu} = kT_{\mu\nu} \tag{3}$$

where on the left side of the equation is the space geometry tensor $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$. This is the "Einstein curvature" of space, the difference between Ricci's tensor $R_{\mu\nu}$ and half of the scalar curvature $R$ multiplied by the metric tensor $g_{\mu\nu}$ . On the right-hand side of equation (3), the stress-energy (energy–momentum) tensor $E_{\mu\nu}$ is multiplied by the Einstein gravitational constant

$$k = \frac{8\pi G}{c^4} \approx 2{,}077 \times 10^{-43}\ \mathrm{N}^{-1} \tag{4}$$

which aligns physical units.

The following year, Schwarzschild (1916) solved[121] equation (3) for centrally symmetric gravitational fields not stronger than the solar one, and showed the coincidence of the solution with Newton's theory with very high accuracy. The new results included a previous anomaly in Mercury's motion and additionally predicted the deflection of light toward gravity. By adding $\Lambda g_{\mu\nu}$ to the left side of equation (3), and later revoking it, Einstein included the accelerated expansion of the universe in the equations.

## Action

Einstein's general equations can be derived[122] from the principle of least action previously known in theoretical physics. The action is a product of the change of energy and the elapsed time (momentum and length), and almost all[123] the movements known today in the physics of the solution are Euler-Lagrange equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x}, \tag{5}$$

where $t$ is the time, $\dot{x}$ is the derivative of the path $x$ by time, and $L = E_k - E_p$ is the difference between the kinetic and potential energy we call Lagrangian. Equations (5) express the condition that Lagrangian be such that it's integral over time, the action

$$S = \int_{t_1}^{t_2} L\, dt\ , \tag{6}$$

[120] property of information
[121] see [17], 1.2.10 Schwarzschild solution
[122] see [4], 2.5 Einstein's general equations
[123] in fact all, because the chaotic movement is reduced to the same thing

be minimal.

Derivations of equations (5) using the principle of least action can be found in various textbooks and discussions of theoretical physics, and I have dealt with this as well[124]. What was particularly interesting to me was the fact that they were first obtained in the 1750s by classical and limited definitions of kinetic and potential energy and the methods of the time, and yet proved to be accurate with the metrics of general relativity. Einstein and his contemporaries were unaware of this possibility, or could not be sure of the success of such a method.

In the same, classical way of working, without Riemann geometries and without Einstein's theory of relativity, the geodesic lines of relativistic gravity are derived from Euler-Lagrange's equations. This does not stipulate that the principles of relativity are not needed for such physics, but only that they are not necessary. And that is the point of this story[125].

## Other alternatives

If we accept the thesis that action and information are equivalent, then Euler-Lagrange equations speak of physical movement in the paths of least interaction and least communication. Physical bodies, particles and waves spontaneously find paths along which they will have as little information emission as possible. It is the “principle of information”, the gentle but omnipresent “force” of the universe.

Wherever we have physical trajectories, the principle of information has been their godfather. However, it is possible to build a system with redundant information. Just as water evaporation (upwards), geysers, volcanoes occur on earth, although gravity (downwards) is more or less equally present in all its places, which occur despite the principle of least action, so life is possible, the creation of beings with accumulating information that allows them excess freedoms. This excess is manifested in a larger number of choices and in choosing more incredible options. Life defies obstacles contrary to the principled minimalism of action (or information) that we would like to indulge in like a log down the water.

We come to the next alternative theory of gravity directly from the principled minimalism of information. Namely, as the more probable event is less informative, the more probable is more common. That would be, for example, an obvious and difficult to refute, but mostly invisible “principle of probability”, which I have also been writing about for a long time.

The probability in particle-wave trajectories of physics is found, for example, in their wavelengths. It has long been known that wave amplitudes indicate the probabilities of observables[126], and this is widely used in quantum physics, but it is less known that wavelengths also indicate a probability density. That the wavelength of a particle-wave represents blurring and in that sense the indeterminacy of its position is used in the explanation of Heisenberg's relations of uncertainty, but mostly this is the end of such uses.

[124] see [4], 2.4 Euler-Lagrange equation
[125] see 23. Action and Information
[126] see [17], 1.2.2 Born information

However, the Compton Effect[127], in addition to proving the particle nature of light waves, is also proof that light travels the most probable trajectories from its point of view and from the point of view that longer wavelengths of light mean lower probability densities of its positions. I've written a lot about it before, on various occasions, so now I wouldn't repeat myself.

## Epilogue

We have seen at least three official theories of gravity (Newton's, Einstein's and Euler-Lagrange's) and at least two more unofficial alternatives (informatics and probabilistic). It is difficult to put them all in the same basket, but again it is possible that this will be inevitable in the further development of physics.

This is part of a story about different postulates, perhaps opposite but independent, with which we can formulate seemingly different but equally accurate notions of reality. Let's not forget that it takes a lot of work, intelligence and scientific heritage to discover such theories, and that this is a consequence of the principle of information too, that the nature is made of information and, therefore, of truths or actions, but to act as if she do not want them.

[127] see [17], 1.1.9 Compton Effect

# 27. Gravity of Chance

Moving by probability trajectories

January 23, 2021

Why don't you just say that celestial bodies move in their orbits because they seem them most likely, but you complicate things with your "information theory"? – The question is interesting to me and I single out parts of the answers. – Because the devil hides in the details, and besides, the alleged information theory significantly simplifies the theory of gravity.

## Entropy

Assume that it is accepted that coincidence exists, that bodies, particles or waves move in straight lines, because their deflections are less probable. They will go where to them are more likely, and with the "more likely" are going less information and more entropy. The following are complications.

Entropy is a thermodynamic term that is not easily transferred to celestial mechanics. According to the laws of gases and statistics, the product of pressure and volume is proportional to temperature ($PV = kT$), and temperature and entropy change is directly proportional to the change of heat, thermal energy ($T\Delta S = \Delta Q$). Hence, the change in entropy is directly proportional to the change in heat and inversely proportional to pressure and volume $\Delta S = k\Delta Q/PV$). In a constant volume ($V$), entropy ($S$) increases with heat ($Q$) and decreases with pressure ($P$).

In the unit volumes of the interior of the star (planet), at greater depths, the pressure of the upper layers is higher, so if we assume that the heat increases proportionally with the pressure, we will have a constant quotient, unchanged entropy. But then there will be no spontaneous (entropy) heat transfer from the warmer to the neighboring colder substance (the second law of thermodynamics) and there is no (appropriate) heat flow from the interior of the star to the surface. As dubious as the latter conclusion is, so suspicious is the assumption about the ratio of heat and pressure. I note that this calls into question official physics.

If the heat from inside the star (spontaneously) spreads to the surface, then there is more entropy in the upper layers. Inside the star, gravity decreases as it approaches the center, which is known (due to less active mass), and the force is constantly pulling down. In other words, the direction of entropy growth has the opposite direction of gravity increase! So much for the interior of the star (planet).

It's even harder to talk about classical entropy outside the star, in a vacuum, but let's try. I said, the change in the entropy of a gas is proportional to the change in its heat (thermal energy) and inversely proportional to the temperature (Clausius, mid-19th century). On the other hand, heat and temperature are caused by oscillations of gas molecules (Boltzmann, 1877), so when warmer gas is in contact with colder vessel walls, larger oscillations are transferred to smaller, warmer gas to colder wall molecules.

As the heat shifts from less to higher entropy of the environment, the oscillations of the molecules weaken. The denominator (temperature) decreases faster than the numerator (heat) to increase the quotient (entropy). It is a spontaneous process of increasing entropy, the transfer of heat from a higher body to a lower temperature environment in the second law of thermodynamics.

Note only that this “spontaneous growth of entropy” is in fact still a phenomenon without deeper explanation, and that further explanation comes with “spontaneous reduction of information”, which is again a consequence of “principled minimalism of information”, and which is a consequence of “principled maximalism of probability”; and let’s continue.

We can imagine a similar phenomenon by simply slowing down time. The oscillations subside, heat and temperature decrease, the numerator is slower than the denominator and the entropy increases. If relative observers saw this in the case of inertial rectilinear motion and within the gravitational field, special and general theories of relativity, they would rightly say that the entropy of these slowed-down systems is higher. However, then we would have a problem with the question why the body does not spontaneously pass from a state of rest to a state of uniform rectilinear motion, that is, why the planet does not simply turn from its elliptical orbit to the sun.

On the contrary, if the relative entropy is smaller, both of the system in inertial motion and of the state outside the elliptical orbits of celestial bodies, then the body will spontaneously remain in a state of rest (or uniform rectilinear motion) or in its geodesic, until it is affected by another body or force. In other words, the lower relative entropy is in accordance with the law of inertia.

Note that this is in line with the previous conclusion (inside the star) about the decrease in entropy in the direction in which gravity increases. This is also in line with Boltzmann's explanation of entropy using the most probable arrangement of molecules. We know from combinatorics that the most probable case is uniformity, when the objects of arrangement do not accumulate in certain positions.

Namely, if the layouts are uniform (by maximum entropies) in their own (proper) system, in a relative (moving) system where the lengths are shortened only in the direction of movement – the layouts are no longer so uniform. Also, in the weightless state of a spaceship orbiting the planet inertial, the molecules are evenly distributed while in the room on the ground the lower ones would be denser.

Thus, force changes probabilities and entropy. But don't lie to the devil; if the relative entropy is smaller and the relative information is smaller, then we need additional explanations.

## Spacetime

After all, the concept of classical entropy may need to be left to the substance itself and for space to devise a different one, so then the previous story is not very important to us. This idea comes from the division of elementary particles of physics into bosons and fermions, the first of which are the carriers of the field of forces (gauge bosons) and the second of which are those on which these forces act.

The same bosons can be in the same quantum system while fermions cannot. I will explain this property on the example of flipping two coins, or flipping one coin twice. Possible outcomes are “heads” and “tails”, so each result is one of: HH, HT, TH, TT. If all four are equally probable, then by repeating the throws many times we will notice that, say, the result of HH is about a quarter of the total throws. This means that the results of HT and TH appeared each in a quarter of all cases, ie that the outcomes of the toss behave like fermions – in the first and second toss the "same" coin is actually a different entity.

If the order in which the "heads" falls after the "tails" does not differ from the reverse order of the experiment (two throws), then the outcome has three equal possibilities: two "heads", two "tails" and a mixed case. Then the "heads-heads" appears in approximately a third of all experiments and the coins behave like bosons. It is unnatural to distinguish these two coins. In 100 experiments, in the first case (fermion) for HH the expectation is 25 outcomes, and in the second case (boson) the expectation is 33 outcomes.

A slightly more complex example is the random arrangement of nine balls in nine equal boxes (they do not have to form a square). Each ball with equal probability can be in any of the boxes and there is room for all the balls in any single box.

If we assign names to the balls, in such a way that they are different individuals, then arranging one ball in the box has $9! = 9 \cdot 8 \cdot 7 \cdot ... \cdot 2 \cdot 1 = 362\,880$ permutations. That number of layouts is bigger than any other, when there are at least two balls in one of the boxes. It is therefore typical for the expansion of gas molecules in a room, because it is the most probable and the molecules are fermions, so the logarithm of this number (factorial $9!$) is a measure of Boltzmann entropy.

However, if the balls are nameless and like a boson, then the way of arranging one ball in each box has only one possibility (combination) and is no longer representative of Boltzmann entropy. Then we have to pay attention to the positions (boxes) more than the bosons themselves (balls). Let us further imagine an arbitrary number of these boxes.

By changing the sizes of the boxes, while maintaining equality with respect to the balls, we create scenes to explain the metrics of space (generalized Pythagorean theorems). We have seen that the Gaussian curve[128] is an invariant of space. Preservation of the value of such a defined "curve" in (tensor) coordinate transformations is like the law of conservation of energy in physical processes. That is why metrics are more important than choosing a coordinate system and, on the other hand, it defines the gravitational field as a physical phenomenon that cannot simply disappear.

If smaller and smaller boxes began to appear on one side, a ball that would move "straight" jumping from one to the other would turn to that side. This is reminiscent of reducing units of length to that side and forming a (non-zero) Gaussian curvature. If the boxes were arranged spatially, in three dimensions (length, width and height), and there was a reduction only in the direction toward a fixed point, we would have a scene that is even more reminiscent of the gravitational field.

The space-time theory of relativity is 4D, three dimensions are spatial and one is temporal. We imagine time as a continuation of space so that a moving particle is present in each of them and is static there. It is a model of a deterministic world that does not fit well with the "universe of uncertainty". In the world of information, the essence of which is uncertainty, we cannot be sure that the particle was in every 3D

[128] Gravitation Multiplicity, https://www.academia.edu/44936839/Gravitation_Multiplicity

"layer of space" of the imagined "stringed" 4D space-time, so we cannot place or non-place it there for sure.

The 6D space-time model I proved earlier[129] helps here, along with the symmetries of space and time itself. The first means that there are as many temporal dimensions as there are spatial ones, and the second means that we can take any four of those six to define a reality like ours. If $x$ is one of the "normal" lengths ($x_1$, $x_2$ or $x_3$) of the space itself, then $x_4 = ict$ is the corresponding "duration", where the imaginary unit is also $i^2 = -1$, and $c = 300\ 000$ km/s is (approximately here) the speed of light in a vacuum.

The particle both "is and is not" present in the layers of space, if it is such in our reality. In order not to deny the uncertainty with the model, it is not enough to add just one more dimension of time to one course of events, because the particles that are not present (in our reality) would then necessarily be present in a single other place of "parallel reality". To avoid "determinism in absentia", in this consideration, two more dimensions of time are also needed, which makes this method consistent with the previous ones.

Further, the "penetration of the particle" through the layers of time, i.e. space-time, can be compared (tested) with the capacity of the classical channel, which I did earlier[130]. It is surprising how the results of this method coincide with the known theories of gravity. Another surprise is how much the "time deficit" treatment is also in line with these theories[131]. But I have already written about that and there is no need to repeat myself here.

## Epilogue

The introduction of "objective coincidence" into reality, whatever that means, will not go easy and will have its problems in interpreting the physical world. I believe it will lead to the correction of many theories as well as to hasty and unnecessary rejection.

For example, classical theories of the macro-world (cosmology, gravity, and dynamics of continuous media) could be underestimated in some stochastic physics, although the law of large numbers of probabilities would clearly stand on their side. Also, there will be opposite cases, that causal phenomena are violently probabilized, even though, for example, the theory of deterministic chaos is on their side.

It is in our nature to prefer dogmas to truth.

If it turns out to be correct, the probabilistic model of physics will include "many worlds" of quantum mechanics with the explanation that the particle-wave can interfere with itself (in the experiment double-slot), by going into some parallel reality and hence appearing in our reality as double.

[129] Space and time, https://www.academia.edu/40603049/Space_and_time
[130] Channel noise, https://www.academia.edu/43753255/Channel_noise
[131] Many worlds, https://www.academia.edu/43647244/Many_worlds

Everett wrote about it (1957), which is why he was ridiculed and excluded from the scientific community. He did not even dare to think further in a similar way, for example about the tunnel effect[132], which could make the uncertainties of physics especially interesting for future researchers.

[132] A phenomenon in which an atomic particle penetrates a potential obstacle despite its own lower energy.

## 28. Tunnel Effect

Talks about quantum tunneling and parallel realities

January 25, 2021

The questions arose from different conversations and I was free to select and summarize them, and to supplement my answers. The interlocutors mostly do not know each other and I guess they want to be anonymous or I do not know their exact identity.

------------------------------------------------ *** ------------------------------------------------

**Question**: What does the tunnel effect (quantum tunneling) have to do with your information theory? I read the attachment[133] but it idea only lists that sentence at the end.

**Answer**: The tunnel effect (or tunneling) was first observed experimentally by Robert Wood[134] in 1897, observing the motion of electrons in the emission field, but he failed to interpret it. This is a phenomenon when a micro particle of less energy passes through a barrier of more energy. In macro physics, that is impossible, you don't go through a closed door, but quantum physics predicts that.

As we know from quantum mechanics, all particles have a wave nature, and their wave functions are solutions of the Schrödinger equation[135]. The amplitudes of these waves (functions) define the probabilities observable, ie the probability that a particle can be measured (appear) at a given place. With this knowledge, when we calculate the probability of its occurrence in a place where it could not be in classical physics (according to classical mechanics, a particle can be found in space only where its potential energy is less than the total), it turns out that it also can be “in an impossible place”.

For the so-called barrier permeability coefficient of the particles behind the (potential) barrier, we get a number that is exponentially dependent on mass, as well as the probability of observation, of the mass depends the observable (physically measurable quantities) information. This is further an informatics explanation and considers it still speculative. I would like you to make an effort and find a possible mistake, before I run into it and archive it somewhere (maybe even publish it).

So this new explanation goes like this. Unlike electromagnetic radiation (photons) for which time stands still, everything that has mass moves slower than light in vacuum ($c = 300\ 000$ km/s) and therefore has its own course of events, which means that it penetrates the layers of time independently of the observer. Aside from the dynamics, that having mass means having inertia due to the principle of minimalism of information, rather than focusing on the kinematics of “walking” of mass bodies or their parts through “parallel realities”.

[133] Gravity of Chance, https://www.academia.edu/44960994/Gravity_of_Chance

[134] Robert Williams Wood (1868-1955), American physicist.

[135] Schrödinger equation, https://en.wikipedia.org/wiki/Schr%C3%B6dinger_equation

Everything that can and does happen somewhere in the "many worlds" of quantum mechanics Everett's[136] from 1957, and from there it can also come back to "our reality". In this way, it can bypass any barrier, only if such in the "many worlds" do not exist in at least one of them available to the particle, and if our world is accidentally available to it from such a "world".

**Q**: Then which explanation is correct, the one with probabilities or this?

**A**: Both, they pull each other. Both probabilities, Born's law[137] and Heisenberg's relations of uncertainty, as well as parallel realities, are consequences of the same principled objective coincidence. Both explanations are parts of information theory, a universe in which information is the basis and whose essence is uncertainty.

***

**Question**: So the narrower the tunnel the slower the flow of information[138]?

**Answer**: The point is in "informatics interpretation". In order for the "tunnel effect" to occur, an obstacle (potential barrier) of a particle that has mass and only such – is bypassed by going through pseudo-realities. The assumption is that the penetration does not happen only in "our reality" (say by pure coincidence), but also because of avoiding obstacles through "parallel realities". This is a speculative part of that theory, according to which not only are such "strange realities" possible, but they can also be used to explain some "impossible phenomena".

**Q**: I didn't know that math was a speculative science?

**A**: The information theory that I am developing in its small part is mathematics, and it is also hardly physics. By the way, I do not consider mathematics a science (experiments are not valid for the proof of her views, but that is also a matter of definition).

Otherwise, what they have in common, in mathematics and natural sciences, is that the methods of work in teaching are very different from the methods of work in research. That is why we have such a large number (throughout history and today) of fantastically good professors, lecturers, and desperately unsuccessful researchers, and vice versa. What more, it is hard to find a successful researcher who, no matter how vain, will say for himself that he was a good teacher.

For example, Einstein was not a professor (aside from the fact that Princeton University paid him to call him his own). Gauss, the greatest German mathematician of all time (or with two or three more) whom they called the "Prince of Mathematics", said that he was a bad teacher, and he was, and did not like to hold classes, although they brag about him created better German mathematicians (Riemann, Dedekind) than all the otherwise great lecturers of his time.

[136] Hugh Everett III (1930-1982), American physicist.
[137] Born rule, https://en.wikipedia.org/wiki/Born_rule
[138] A colleague, a retired microbiologist, asks and jokes a bit.

Newton tried to give lectures at Cambridge and it was a disaster (in the end he admitted it himself and gave up). Euler, whom Russia paid until his death (he is the first man in history to receive a pension), but it is not known (me) that he ever gave a lecture. .. That enumeration is a very long story.

Well, the difference in methods comes from the great diversity of goals. For example, in the “good teaching” of mathematics, you first teach the axioms and settings of the scene, and only then do you go to the content, the consequences. Exploring, you first wander in the dark for a long time and when you get to the axioms and the “stage”, for you the job is mostly done. When you have to solve problems at school level (even if it was at the highest international competitions), there is always the awareness that the task is well set and that the solution is at hand and it is just waiting for the skills of the competitors to work.

On the contrary, when you research, you don't really know what the goal is, and if something comes to your mind, you always face the possibility that what you are looking for does not exist (which is the most common case) and that you will spend your whole life banging your head against the wall. What you think you are looking for may be possible to find (rarely) but it is too difficult for you (in those rare cases often).

And in the best (and rarest) case when you might discover something bigger, then you have to hide it because your colleagues would make fun of you (you're not sure also), in a better case they would steal you (few discovered something and then were known for it), or it will impose on you such filtering and race with institutions (they are vain as political organizations, and you threaten their conceit, magnificence) that you have to give up. Many “great” discoveries were made once so that nothing like it ever appeared again by the same author (I also have a long list).

In short, as a teacher you act from the point of view of authority, and in the research phase you are in the position of renegade, speculator and quasi-scientist. The larger the theory you find, the more speculative the period before it and the slower the recognition.

***

**Question**: This explanation of the tunnel effect of quantum mechanics by bypassing the barrier through parallel realities seems interesting enough that, regardless of the possible (in) truth, it obliges physics researchers to more seriously review, elaborate and verify experiments. Why don't you deal with it?

**Answer**: Well, I started that topic. It never occurred to anyone before me to connect such things. However, not everything “inexplicable” will be explained by “parallel realities” and forcing the idea will be trivialized.

**Q**: Yes, understandable. What doubts do you mean in particular?

**A**: For example, I check for “disappearing” (so called evanescent) fields and waves.

Everything that moves at the speed of light does not have its own time (time does not flow). The three dimensions of the photon are the information of that particle (2D) which is alternatively in the planes of electro-magnetic induction, plus the relative time of the observer. The photon itself does not exist in time, has no time course, does not penetrate the layers of time like particles that have mass. The latter necessarily move slower than light and have their own (proper) flow of time which (due to the principle of minimalism of information, aside from the Higgs mechanism[139] which I see as a confirmation of that theory) produces their inertia.

Mass particles can go into “parallel reality” on their own and through it “bypass” the obstacle and appear where they physically “cannot be” (tunnel effect). But the greater their mass, the less chance they have of such a circumvention.

The solutions of Schrödinger's equation fit very well into this story. However, there are also the mentioned phenomena of the evanescent effect (said “disappearance” or “passing”, depending on the translation) of the wave of light, which for now is not associated with quantum physics, but it could be in some of its development. The logical explanation could be that the “passage” of light is a phenomenon caused by the mass of the observer himself.

**Q**: Explain?

**A**: You are big and light takes time to cross the path from your feet to your head, which means you are not all your own. Joke aside, but that is why you are not in only one reality, instead you are always in some “parallel realities” with some of your tiny parts. There, that light has some (at least a little) different life and that is that almost infinitesimal deviation we call evanescent effect.

**Q**: And what's the problem with publishing it like that?

**A**: I have to somehow see (theoretically of course) how in a body of greater mass, or relative bodies observed in motion or greater mass (in the environment of strong gravity), this “disappearing effect” behaves, not only in light.

**Q**: So you have that Hawking[140] radiation (1973) around the “black holes” that made him famous?

**A**: Yes, bravo, I have that in mind. Hawking explained this phenomenon brilliantly with the help of virtual particles, but now a unifying explanation is needed. By the way, I’m close to rounding out that story.

***

**Question**: The more findings we get about dark matter, the better the (hypo) thesis about the space that remembers looks to us[141]. Can you briefly describe to us how to come up with this idea and wheather it could have anything to do with quantum tunneling?

[139] Higgs mechanism, https://en.wikipedia.org/wiki/Higgs_mechanism
[140] Hawking radiation, https://en.wikipedia.org/wiki/Hawking_radiation
[141] Ask me from a group who think I am a representative of a group.

**Answer**: The basic (hypo) thesis is about the universe of information whose essence of the future is uncertainty and the law of conservation of the past. First of all, let us note that the latter can be proved by the former!

In short, information is (always) a quantity of unpredictability, its smallest volume (however many) are pure uncertainties, and less than uncertainty is certainty. That is why the information is quantized, atomized, or say has the smallest portions. Because they are discrete phenomena (like natural numbers), they are finally divisible and the law of conservation applies to them.

This can be demonstrate and vice versa. Starting from quantum mechanics, where we know that action (not energy itself, but the product of energy and time) is quantized, and information travels with interactions (physical action), it follows that it is also "in packets". Additionally, starting from quantum mechanics whose states are vector representations (unit intensity, norms), and operator representation processes which are also unit norms (this to map unit vectors into unit vectors). But they are therefore invertible; it is possible to get the original from the copy. In other words, quantum processes remember and the law of conservation applies to information. Then, the information is quantized, finally divisible, because only infinite sets can be their real subsets and do not maintain quantities.

We need this to understand that there is free least information as elementary particles. We know from the previous that they travel through space and cannot be smaller, but otherwise we also know that they do not become bigger, and that they leave their "biography" somewhere along the way. That is why "space remembers" because there are elementary particles that "do not remember" and last.

**Q**: In what kind uncertainties are the "elementary information"?

**A**: A useful question. They are in the uncertainty of their environment. Just as it is uncertain "what will be tomorrow", it is also uncertain "what will be there". In the wider environment, each of these "particles of the same" (which of course is not entirely true, at least due to Heisenberg's relations of uncertainty) – is always something different.

**Q**: How is that memory "stored"?

**A**: The space is getting bigger. The increase in space could also be apparent, as a melting of the substance, and this as a spontaneous increase in the entropy of the substance, i.e. a decrease in the total information of the substance. With information coming from the past, it makes sense to assume that the total information of the universe (substance plus space) remains constant.

**Q**: Well, let's say that explains dark energy, and where are the gravity and dark matter?

**A**: We should first notice that light moves at the speed of light and that is why it (photon, electromagnetic wave) has no its own (proper) time. Everything that moves at the speed of light belongs only to the present of the observer and from there it gets one dimension (temporal), plus two dimensions (spatial) which it has as information. In the case of photons, for example, these are the electro and magnetic planes of oscillation. We do not see the past, nor parallel realities, because we look with photons.

Particles that have mass have it because they have their own duration (non-zero flow of time) and according to the principle of minimalism, information “gets stuck” in time through which they pass, from where their inertia comes from. In order for space to have memory, it must somehow act from the past to the present (which we perceive with light) and a good candidate for that is gravity.

**Q**: So the gravity does not have to be what acts from the past to the present. And why do you think it is?

**A**: The calculation shows, firstly, that due to the elliptical (conical in general) orbit of the planets around the Sun, the attractive gravitational force decreases with the square of the distance. However, Mercury and bodies near strong gravitational fields do not follow the trajectory of a conic (ellipse) and gravity then does not decrease with the square of the distance.

Secondly, we know that if the force does not decrease with the square of the distance, then its field does not expand at the speed of light. That is why gravitational waves near large masses travel slower than light and, due to the principle of information, have mass. It can be calculated that this mass is fantastically small (perhaps never immeasurable) but it exists, which means that it penetrates through the layers of time.

That is why the past has a gravitational effect on the present, but to say again, only if there used to be a mass in a given place.

**Q**: All right, it's worth reconsidering. And where is the quantum tunneling?

**A**: That part is in the hint. The mass particle passes through the layers of time, either from the past to the present or from one “parallel reality” to another. Let me remind you, in (my) information theory, there are three dimensions of time for the three dimensions of space. These “three plus three” dimensions are so symmetrical that you can take any of the four of them and declare three “spatial” and one “temporal”.

The idea is that each time dimension gives space a special duration, however, I'm still not sure if quantum tunneling is just that bypassing obstacles through parallel realities in order to do that.

**Q**: Why is there no macro body tunneling?

**A**: Because of the law of large numbers of probability theory.

## Epilogue

Do you know that paper books are no longer in fashion – a friend tells me and adds – everything is being digitized today, what is left behind will be buried in the past, and those recently recorded analog texts will soon become rarities or will be unusable. I will keep that in mind – I told him.


Gimnazija Banja Luka Library, January 27, 2021.

## Literature

[1] Rastko Vuković: *INFORMATION STORIES* – of physics, biology, low. Economic Institute Banja Luka, 2020 (on hold). https://archive.org/details/stories_202008

[2] Rastko Vuković: *ACTION of INFORMATION* – energy, time and communication; Economics Institute Banja Luka, 2020. (https://www.scribd.com/document/463637180/Action-of-Information )

[3] Rastko Vuković: *Information Potential*, October 2020. https://www.academia.edu/44373395/Information_Potential

[4] Rastko Vuković: *MINIMALISM of INFORMATION* – physical information and applications; Economic Institute Banja Luka, 2020. https://www.scribd.com/document/421999486/Minimalism-of-Information

[5] Muneo HORI, Lalith WIJERATHNE, Rizwan RIAZ and Tsuyoshi ICHIMURA: *Rigorous derivation of Hamiltonian from Lagrangian for solic continuum*; Journal fo JSCE, Vol 6, 1-11, 2018. (https://www.jstage.jst.go.jp/article/journalofjsce/6/1/6_1/_pdf )

[6] Rastko Vuković: *Dual Vectors* – quantum states and processes, October 18, 2020; (https://www.academia.edu/44320266/Dual_Vectors_quantum_states_and_processes )

[7] Massimiliano Proietti, Alexander Pickston, Francesco Graffitti, Peter Barrow, Dmytro Kundys, Cyril Branciard, Martin Ringbauer, and Alessandro Fedrizzi: *Experimental test of local observer-independence*, ArXiv:1902.05080v2 [quant-ph] 4 Nov 2019; (https://arxiv.org/pdf/1902.05080.pdf )

[8] Rastko Vuković: *INFORMATION of PERCEPTION* – freedom, democracy and physics; Economic Institute Banja Luka, 2016. https://www.scribd.com/document/360939276/Information-of-Perception

[9] Cen Wu, Qingquan Tang: *Reward meritocracy or nepotism*: The case of independent financial advisors appointed by Chinese listed companies; China Journal of Accounting Research journal, 2019; www.elsevier.com/locate/cjar

[10] Boole, George (2003) [1854]: *An Investigation of the Laws of Thought*; Prometheus Books.

[11] Sergey Brin and Lawrence Page (1998): *The Anatomy of a Large-Scale Hypertextual Web Search Engine*; Computer Science Department, Stanford University, Stanford, CA 94305, USA. (http://infolab.stanford.edu/pub/papers/google.pdf )

[12] Unitary space. Encyclopedia of Mathematics. https://encyclopediaofmath.org/wiki/Unitary_space

[13] Hanno Essén: Space Mechanics; Department of Mechanics Royal Institute of Technology S-100 44 Stockholm,Sweden; 1998,October.

[14] Dr Božidar S. Milić: *ZBIRKA ZADATAKA IZ TEORIJSKE FIZIKE* – I deo – mehanika sistema i neprekidnih sredina; Univerzitet Beograd; BIGZ, 1971.

[15] Orbits of ancient stars prompt rethink on Milky Way evolution, NOVEMBER 16, 2020. https://www.scienceinpublic.com.au/media-releases/kinematics

[16] Momir V. Ćelić: *MATEMATIKA II* – Obične diferencijalne jednačine, Višestruki integrali, vektorska analiza, Kompleksna analiza, Furijeova analiza; Abakus sistemi, Banja Luka, 1997.

[17] Rastko Vuković: *SPACE-TIME*, principles physics of chances; Economic Institute Banja Luka, 2017. https://www.scribd.com/document/349336253/Principles

[18] Rastko Vuković: *PHYSICAL INFORMATION* – in social phenomena and mathematics; Economic Institute Banja Luka, 2019. https://www.scribd.com/document/408301124/Physical-Information

[19] Michael A. Nielsen & Isaac L. Chuang: *Quantum Computation and Quantum Information*. Cambridge University Press, 10th Anniversary Edition, 2010. http://mmrc.amss.cas.cn/tlb/201702/W020170224608149940643.pdf

[20] Emergence – How Stupid Things Become Smart Together https://www.youtube.com/watch?v=16W7c0mb-rE

[21] Rastko Vuković: *MULTIPLICITIES* – Choices for animate and inanimate; Economic Institute Banja Luka, 2018. https://archive.org/details/Multiplicity_Rastko/

[22] Rastko Vuković: *QUANTUM MECHANICS* – the representation of continuum (of metrical and vector spaces); Economic Institute Banja Luka, 2018 (on hold). https://www.scribd.com/document/377641503/Quantum-Mechanics

[23] Robertson-Schrödinger formulation of Ozawa's uncertainty principle. Catarina Bastos et al 2015 J. Phys.: Conf. Ser. 626 012050. https://iopscience.iop.org/article/10.1088/1742-6596/626/1/012050/pdf

[24] Thompson, W.J. (1994). Angular Momentum: an illustrated guide to rotational symmetries for physical systems. 1. Wiley. p. 5. ISBN 0-471-55264-X.

[24] [3] Thompson, W.J. (1994). Angular Momentum: an illustrated guide to rotational symmetries for physical systems. 1. Wiley. p. 5. ISBN 0-471-55264-X.

[25] Rastko Vuković: *Bernoulli's attraction* (September 13, 2019) https://www.academia.edu/40349717/Bernoullis_attraction

[26] Растко Вуковић: *ЕНТРОПИЈА* – информација термодинамике; Архимед Бања Лука, август 2016. https://archive.org/details/Informacija2/

[27] I.A. Pshenichnyuk: *Pair interactions of heavy vortices in quantum fluids*; Skolkovo Institute of Science and Technology Novaya St., 100, Skolkovo 143025, Russian Federation (Dated: May 30, 2017). https://www.researchgate.net/figure/Bernoulli-force-responsible-for-the-repulsion-and-attraction-of-heavy-vortices-plotted-as_fig4_317229899

[28] Ofer Moshe, Gravity, January 2021. https://www.academia.edu/s/116ac78a12

[29] Rastko Vuković: *About parallel reality II* (January 10, 2020).
https://www.academia.edu/44876245/About_parallel_reality_II

[30] Thomas Cuff, Frederick Community College: *The STM* (Scanning Tunneling Microscope) – The forgotten contribution of Robert Francis Earhart to the discovery of quantum tunneling. February 2016