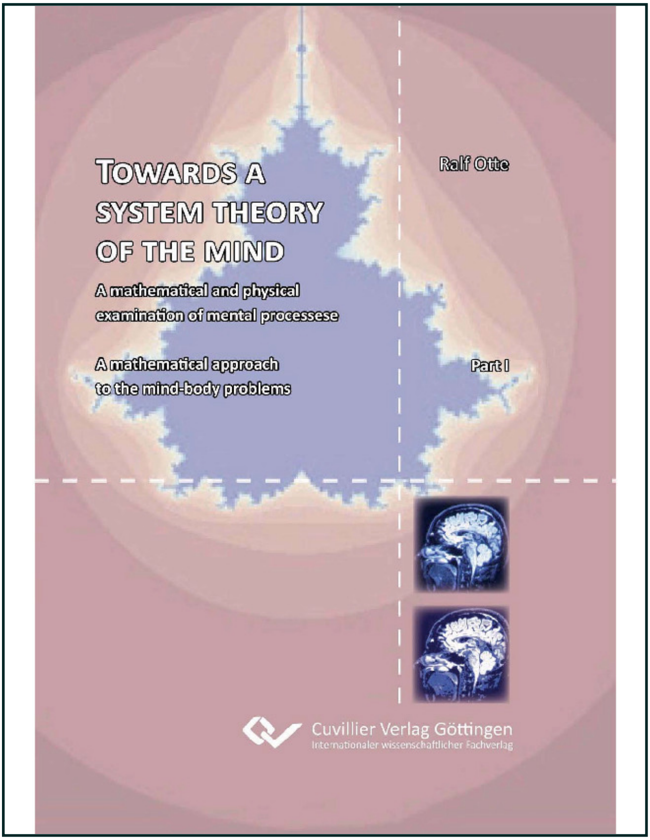
Ralf Otte (Autor)

Towards a system theory of the mind

A mathematical and physical examination of mental processese A mathematical approach to the mind-body problems



https://cuvillier.de/de/shop/publications/40

1. Preliminary considerations and an introduction to the topic

A lively discussion among philosophers concerning the nature of mental states has been ongoing for centuries. In all its infinite detail, this discussion is barely accessible to a layperson. Nevertheless, a basic pattern can be identified in the fundamental philosophical assumptions woven through this discussion which has, from time to time, put its stamp on entire epochs. Going back as far as ancient times, it was assumed that a substantial material difference between actual and mental phenomena existed, a genuine dualism between physical and mental substances. Later philosophical views weakened this fundamental dualism, reducing it to a dualism of properties of the same substance. Around the middle of the last century, many philosophers favoured a monist explanation, one which posited that mental and physical states were one and the same. Monism was later adapted further; an essential difference between physical and mental phenomena was once more acknowledged, and many philosophers and researchers today take a perspective termed functionalism as the best approach towards describing the mind-body problem. All these approaches have continued to develop; today, a variety of different philosophical approaches towards and explanations of the nature of mental states are advanced.

The problem at hand is also no longer the exclusive domain of philosophers; for many years now, scientists have also been attempting to find out how mind states, or the internal mental states of a person, can emerge from the complicated physical activity patterns of the brain. The focus is on discovering nothing less than how consciousness in all its many facets comes into existence. The question as to whether and *how* mental states can, in turn, causally affect physical states is also of paramount importance. Some researchers deny that such effects are possible at all (for reasons related to energy) as a matter of principle. Others acknowledge that such a cause-effect relationship may be possible, but ask how it could function. These are fundamental questions for philosophers, biologists and psychologists working on the "new science of mind" (Eric Kandel). This book introduces formal descriptions of how mental and physical states can be modelled *mathematically* in order to support this interdisciplinary discussion. This should enable the development of models that can be used to gain a better understanding of the interaction between mental and physical phenomena.

Although the essay begins with a philosophical treatment of the subject, it pursues a systems theory approach to brain phenomena. Specifically, this entails describing the behaviour of the brain just as cybernetic systems theory describes other systems: in terms of input functions (stimuli), (internal) state functions and output functions (responses). It is absolutely clear that only an initial approach can be outlined here, a method for developing stationary and dynamic state functions of the brain. The complexity of even the smallest parts of the brain renders its in-depth mathematical modelling virtually impossible. A formal mathematical approach does, however, have the advantage that it must be accurate enough – that it is, indeed, precise enough –

to cross-check theoretical models against data gathered through experiments. This, in turn, makes it possible to identify contradictory statements quickly. In this way, improvements to models can be made – at least in the ideal case – in a more sophisticated fashion than is possible when complex phenomena are described only verbally. On the other hand, formulating problems mathematically is to risk trivialising them; often, so many simplifications and framework conditions have to be introduced that models may correspond only tenuously with the reality they are supposed to represent. The author's view is that the opportunities outweigh the risks in this specific case.

The following text draws on complex-valued mathematics to describe the state variables of the brain. The author believes that it is not possible to describe brain states correctly only by employing real-valued mathematics. With some simplification, the reasons why this is so can be stated as follows: even a simple synapse in the brain is capable of inductive and capacitive behaviour, since its channels constantly transport charged ions. System theory has demonstrated that the state functions of such systems are best described using complex (real and imaginary) functions. It is, therefore, unthinkable that parts of the brain could be analysed correctly using real-valued mathematics without resorting to precisely the sort of disproportionate oversimplification which needs – as has just been stated – to be avoided as much as possible. For if even the transfer function of a synapse has to be a complex-valued function, it is clear at the very least, that higher-level functional groups in the brain will have to be described with complex-valued mathematics; it begins to become apparent here that hypercomplex mathematics might be required.

The complex nature of the brain which has been established as a result of these initial systems theory considerations has some unexpected consequences. Asking what other phenomena in the natural sciences can solely be described using complex-valued functions inevitably leads us towards problems from the realm of quantum physics. For example, the wave equation of a free electron can only be written using complex numbers. The famous double-slit experiment in quantum mechanics shows that the unexpected interference phenomena of single photons can only be described in terms of the superposition of complex functions. Here, already, we can observe nature itself dictating how its structures should be modelled. We do not have the option of choosing the mathematical tools which are adequate to describe fundamental phenomena at random. Surprisingly enough, it transpires that phenomena which can be described in this manner have "surprising" properties. In this light, we can expect that the brain is also likely to exhibit unexpected properties, ones which are unlikely to be any less startling than the surprising effects found in quantum physics – and this regardless of whether inner brain states are based on physical quantum effects or not.

Particular dynamic effects of the brain could be included among the interesting systems theory properties which might be expected to result directly from the complex-valued mode of description, since complex-valued feedback systems possess much “richer” systems dynamics than real-valued feedback systems. We know from the theory of dynamic systems that particular first-order differential equations can serve to model complicated chaotic behaviour. In the discrete case, differential equations correspond to recurrence equations, or discrete feedback systems. As neurobiology provides numerous indicators for the presence of feedback loops in the brain and we can assume, in a first approximation, that the brain processes information in discrete cycles, it seems appropriate to describe the brain using discrete recurrence equations. However, if the recurrence equations typically used to model feedback in systems theory are extended so that complex-valued difference equations are used, this then results, in even the simplest case, in systems with system dynamics that can be described by complex Mandelbrot and Julia sets. As such, even the simplest complex-valued system models (recurrence equations of the first order with quadratic polynomials) reveal a vast range of possibilities for convergent, divergent and chaotic dynamics. Therefore, in a first approach, they can be used to describe the dynamics of the brain, if the mathematical perturbation parameters of the recurrence equation are interpreted as a stimulus constituting the input the system responds to dynamically. This corresponds very neatly to the theory of chaotic dynamic systems, extended to complex-valued variables. Following this concept, a recurrence equation that allows us to come up with an initial model for the conscious perception of a visual stimulus will be developed, albeit only qualitatively, since the author does not have access to specific experimental data. *Conscious perception* can already be seen as belonging to consciousness, so that the mathematical modelling of the conscious perception of a visual stimulus can present an approach towards achieving a better understanding of the complex phenomenon constituting *consciousness* from the perspective of systems theory in the future.

On closer inspection, it can also be determined that it is advantageous to extend the model by using hypercomplex algebra (and not only complex-valued mathematics) to describe the “higher states” of the brain, the mental and neural processes of perception and self-awareness. This is a postulate, not an inevitable logical conclusion like the one which led us towards the complex-valued form of description for the transfer behaviour of synapses. Examples will be used to show that mathematical description using quaternions could explain certain properties of the brain optimally. In particular, it is suggested that phenomena relating to self-awareness should be described by hypercomplex numbers, as quaternions allow two imaginary dimensions (i,j) to be multiplied, yielding a further dimension, (k). This form of description would provide an *emergence theory of the mind* to be taken into account. Completely new properties come into being during emergence, and this process can be mathematically modelled effectively by utilising a new (imaginary) dimension. The potential consequences of this for the nature of mental processes have not yet been examined; in particular, the potential to develop hypercomplex functions with no real part (functions based on vector quaternions, for example) pose many questions as to the nature of states which could be described in this way.

It is not necessary to support this latter postulate to recognise that the method for describing neural and mental states using complex values outlined above has consequences not only in systems theory, but also for the recognition of the real nature of the underlying phenomena. While the real-valued part of the function can clearly be reconciled with derivable electric and magnetic potentials, or, in other words, with measurable neural activity patterns, the question remains as to what is *really* meant by saying that states of the mind (mental states) can be described as imaginary. As a matter of principle, this important question needs to be addressed by philosophers. This work addresses the nature of other imaginary phenomena which up to now have as being imaginary.

With respect to quantum physics, it seems clear that these phenomena should be seen as a variety of “pre-physical” phenomena. While this answer may seem surprising and also somewhat unsatisfactory, it no longer perturbs quantum physicists, who have been using complex probability amplitudes in their calculations for decades - without being able to fully resolve the nature of these mathematical constructs. The concept “pre-physical” is introduced here only to emphasize that these quantities exist objectively (that they are physically real), but that they are not (yet) so real that they could be measured directly; all known sensors measure only real-valued physical quantities. Where physical phenomena are described with imaginary functions, they cannot be measured directly, despite the fact that they can realise effects in the physical world. The interference bands on the diffraction grating of a double-slit are the results of such physical effects, resulting from the superposition of components which include imaginary components.

Quantum physics thus teaches us that, or rather *how*, such “pre-physical” phenomena become real-valued as a result of specific “operations”. If these insights are applied analogously to the brain, an approach explaining how imaginary, mental phenomena (of the mind) can exert causal effects on neural (and therefore physical) processes emerges. The mathematical “reason” for this lies in the fact – to simplify somewhat – that simple operations allow imaginary functions to become real-valued: if an imaginary number (bi) is multiplied by itself, or a complex number (a+bi) is multiplied by its complex conjugate (a-bi), the result obtained is purely physical; ultimately, the complex-valued equation i*i=-1 holds.

This raises question as to how this mathematical multiplication of complex state variables is implemented biologically? System theory can help us here. If we assume that the complex functions describing the state variables are the particular complex functions, commonly referred to as Fourier or Laplace transforms, this has a decisive (biological) consequence, thereby underlining the power of this analytic method. Fourier transforms are what is termed “the spectra of time or position functions”. Time signals (of a stimulus) or localised patterns of a response to a stimulus (excitation patterns) in the brain can be represented as time or position functions, or as functions of their respective spectra. The value of spectral representation lies precisely in its uncoupling of the original functions from specific (absolute) times and positions; the spectral representation contains the same information as the original functions, but is independent of their “absolute” positions in time and space. It will be assumed for this reason that biological systems carry out such transformations of their own accord and process input signals in transformed (spectral) form. It is also assumed that brain states are coded spectrally, as that would make an “absolute” localisation of system states in the brain unnecessary.

Under these conditions, which are biologically plausible and, for example, can be seen in the hearing apparatus with its spectral decomposition of sound, it can now be concluded that the multiplication of mental states in the brain takes place constantly, since the multiplication of complex spectra is mathematically equivalent to what is generally called a *convolution* of their time and position functions; this can be expressed in simpler terms by saying that position and time structures interpenetrate each other.

If it were possible to prove (anatomically) that two parts of the brain, G1 and G2, were interconnected and interpenetrated each other structurally, in systems theory terms, this would correspond to the multiplication (G1*G2) of the spectra of both of their complex-valued state functions. This, in turn, would inevitably result in the imaginary state variables of G1 and G2 becoming real, and vice-versa. This would be, so to speak, the sought-after place where structural conditions allow real-valued, measurable potentials to emerge from imaginary mental states. Even a separate neuronal assembly with the transfer function Z and an output Y which is neurobiologically linked to its input would correspond, in systems theory terms, to a permanent multiplication of the output state Y by the transfer function Z; this case would also lead to the multiplication of complex-valued functions, causing real-valued components to change (and vice-versa). It is possible to conceive of a third (and exceptionally interesting) case involving neuronal assemblies linked by feedback loops. The author strongly suspects that neuronal assemblies exist in the brain where the system output is not just fed back into the system as an input, but where the complex-valued transfer function Z (the structure, in other words) is fed back upon itself. These systems will be described as self-referential systems in this essay, and their system dynamics can be described mathematically with the equation $Z_{n+1}=A*Z_n*Z_n + B*Z_n + C$, at least at first approximation. Such a self-referential system possesses exceptionally interesting dynamic properties. In particular, certain recurrence equations of this type allow simple self-organisation processes to be modelled. These will be examined more fully in **Part II** of this study.

The present study attempts to develop a formal tool to be used in discussions of the mind-body problem and to elaborate theories and hypotheses and give them specific mathematical expressions. Of course, this particular discourse can only introduce a framework and suggest some provisional approaches towards working within such a framework; given the diversity of structures and processes in the brain, experts in specific areas will have to use whatever mathematical tools required to explain their own observed data. It is already clear that some readily accessible data [Libet, 2005] can be explained well using the suggested description method, and that this approach can make a contribution to the question of the nature of mental processes, mainly because this question has already been asked and answered in the other fields of science which have long been using complex-valued approaches towards describing problems.

Mathematics does not affect the nature of things. The history of science shows rather, that problems which could only be solved by developing "new" mathematical approaches cropped up from time to time. Just as real numbers (e.g. transcendental numbers) which can result from the solution of simple geometry tasks can lead us closer to a better understanding of infinity, imaginary numbers can draw our attention to physical phenomena which take place beyond what we can directly measure but nevertheless influence reality directly. Consciousness or mental phenomena could possess such "pre-physical" properties. Were this to be verified in separate experiments, mental phenomena would regain their status as a separate entity and be viewed as independent phenomena existing "in parallel" with physical phenomena and interact cooperatively with them. The working hypothesis which will be used in this essay equates these imaginary "pre-physical" components of state variables to the *private states* of the first person (with his or her qualia) and takes real-valued components to be objectively measurable neural activity potentials. Whether this simplification is legitimate or whether it leads to inconsistencies, cannot be established at this point. However, one thing is already evident at this stage: we do not need to concern ourselves with the theoretical question of how (objective) spatial patterns of brain activity can lead to (private) internal states; real and imaginary components of complex state variables *always* emerge *simultaneously*, and each is continuously converted into the other. Just as real numbers determine imaginary numbers (and vice-versa) in the field of complex numbers, no linear cause and effect chain connects neural and mental phenomena; both form an "organic whole" in a more profound sense than is suggested by the "two sides of the same coin" metaphor.

Brain states have real-valued and imaginary parts which continuously interact with each other, but we must abandon the ideas of independent mental phenomena which affect physical ones, and separate physical phenomena which are the root cause of mental phenomena. As an "organic whole", both brain phenomena are converted into each other continuously; each phenomenon interpenetrates and is the prerequisite for the other; even simple multiplication of complex functions is sufficient to establish the *different nature, but also the unity* of both phenomena. This could help us understand why the interactions observed in all specialist disciplines focusing on the brain appear to be so extraordinarily complicated.

This essay is not intended to be a philosophical document, but a contribution to a discussion taking place among scientists and engineers. It begins, nevertheless, with a short (historical) overview of the philosophy of mind. The reason for this theoretical excursion is that theoreticians in the past have already thought intensively about how the mind and the body interact; the arguments and possible inconsistencies they have discovered form an excellent research framework for all empirical disciplines.

What *are* the issues which are at stake here? *Metzinger* points out that all the discussion in recent years can be summarized in a philosophical trilemma which probably represents the best way of describing the philosophical mind-body problem [Metzinger, 2007]. Whenever two statements in this trilemma agree, they always contradict the third.

The trilemma 1. Mental phenomena are non-physical phenomena. 2. Mental phenomena have causal effects in the area of physical phenomena. 3. The physical world is causally closed.

Table 1: The body-mind problem trilemma (Metzinger, 2007, vol. 2, p. 14)

For example, if statements 1 and 2 in Table 1 apply, this combination is at variance with statement 3; if non-physical phenomena could have causal effects on physical phenomena, then the realm of physical phenomena would obviously no longer be causally closed. This third point in the trilemma is of paramount importance; whatever the effects of mental phenomena on the physical basis (of the mind) might be, a "mental intervention" cannot be permitted to disturb the laws of cause and effect in physics.

For our purposes, the philosophical discussion can be reduced to a small number of points. If we believe that mental states exist as independent phenomena – and here we will refer quite deliberately to mental or conscious states – then where in the brain should these mental states intervene in physics, in physical states? Where, for example, should they interact with the neural networks of the brain? This is a crucial point because there seems to be a particular difference in the nature of mental and physical phenomena, one which was already highlighted by *Descartes*. Physical states are always spatial states; mental states, in contrast, seem to occupy no space. Since mental states cannot be localised in the brain, we cannot imagine where our thoughts, feelings and wishes reside if they don't map the activity patterns of neuronal assemblies.

If there is no physical location for thoughts and emotions, where should the "causal handover point" be located, the specific site of interaction between physical and mental states? This question arises independently of the requirement that physics be causally closed (Statement 3 in Table 1). Even if such a place could be determined, we cannot really imagine how mental non-physical phenomena should lead to physical effects. In contrast to this, it is widely accepted that physical phenomena can affect mental states causally. Indeed, it has also been demonstrated experimentally that physical phenomena, such as neural activity patterns in the brain, can certainly lead to mental states. Modern brain research can show that simple stimulation of specific areas of the brains of experimental subjects can evoke mental states such as pain or the perception of colours. Correlations between physical and mental phenomena have repeatedly been shown in experiments, but correlations can have different reasons and causal directions. Some philosophers and scientists still believe that mental phenomena can have causal effects on physical ones; others deny, as a matter of principle, that such effects can exist.

However, if it could be shown that mental phenomena are also ultimately physical phenomena (albeit "pre-physical"), just as some effects of quantum physics are "pre-physical" phenomena, nevertheless able to impact on "observable" physics under certain conditions, then this would resolve the trilemma.