Continuum mechanics and signals in nerves

Jüri Engelbrecht*, Kert Tamm and Tanel Peets

Department of Cybernetics, School of Science, Tallinn University of Technology, Akadeemia tee 21, 12618 Tallinn, Estonia

Received 24 August 2020, accepted 21 October 2020, available online 12 January 2021

© 2021 Authors. This is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/).

Abstract. The review describes how ideas from the fields of physics and mathematics have influenced the studies on signal propagation in nerves, which has classically been related to electrophysiology and chemistry.

Key words: nerve pulses, mechanical waves, temperature, interdisciplinarity.

1. INTRODUCTION

Mechanics is an old discipline of scientific research and there is even a saying that mechanics shaped the world [1]. Indeed, statics and dynamics of material objects are described by Newtonian mechanics over many scales, be it the motion of planets in space, the slow drift of continents, propagation of stress waves in solids or molecular dynamics of nanostructures. Leaving aside quantum mechanics and the theory of relativity, continuum mechanics can describe many processes in solids and fluids [2]. Continuum mechanics is closely related to thermodynamics which governs temperature, energy and entropy. This is all basic physics but many branches of research need more knowledge from other fields, e.g. chemistry, biology, etc. Dealing with living organisms, the level of complexity is increased considerably. This review is concerned with the contemporary state of modelling nerve pulse propagation, demonstrating how the knowledge from continuum mechanics can be applied to building consistent mathematical models for this fascinating phenomenon. Propagation of signals in nerves is a fundamental process needed for understanding cognitive processes and mental phenomena. It involves not only electrical signals (action potentials) but also chemistry, mechanics and thermodynamics, and in this sense can be placed at the interface of physics, physiology and mathematics – a truly interdisciplinary branch of research [3]. Kaufmann [4] has stated that “electrical action potentials are inseparable from the force, displacement, temperature, entropy and other ... variables”. This means that a signal is composed of several interacting components – an electrical signal (action potential) is accompanied by mechanical and thermal components. Below it is shown how the basic laws of physics can be used in modelling of such multicomponent signals in nerves and how that knowledge has enriched the understanding about the process. Attention is paid to possible interactions between the signal components and the modelling of physical effects.

* Corresponding author, je@ioc.ee
Communication across the barriers of various disciplines hinders sometimes the usage of interdisciplinary advantages of research. However, during the last two decades this attitude has been changing. Keener and Sneyd [5] note first that “... many physiologists seek only ... learning the functions of an incredibly complicated array of components; often the complexity of the problem appears to preclude a mathematical description.” Then they conclude that “times are changing and it is rapidly becoming apparent that applied mathematics and physiology have a great deal to offer one another”. Cohen [6] states that “those who understand the calculus, ordinary and partial differential equations and probability theory have a way of seeing and understanding the world, including the biological world, that is unavailable to those who do not”. Hobbie and Roth [7] describe physics for medicine and biology, starting with mechanics and thermodynamics followed by a description of possible electric fields before explaining the physiological processes. Presently it is well understood that the complexity of biological processes needs interdisciplinary approaches to understand the phenomena under consideration. Such an approach is called systems biology [8] or more specifically computational biology [9], biomathematical modelling [10] or just working at the interface of physics and biology [3,11]. Actually, the saying attributed to Galileo Galilei that “the Book of Nature is written in the language of mathematics” has taken on a new value.

In a nutshell the problem is the following [12,13]. An axon, the main carrier of a propagating signal, can be modelled as a tube in a certain environment. Inside the tube is the axoplasmic fluid, called also intracellular fluid, which contains cytoskeletal filaments and has a certain concentration of ions. The wall of the tube has a bilayered lipid structure called biomembrane. This biomembrane is composed of two layers of amphiphilic phospholipids with hydrophilic heads and hydrophobic tails. The lipid bilayer can be asymmetric (uneven distribution of phospholipids between the two sides of the membrane) [14]. The wall permits the passage of ions through the ion channels from inside to the environment and vice versa. This process is needed for the propagation of an electrical signal (action potential) in the axon, which is accompanied by mechanical and thermal effects. Contemporary research has revealed many interesting physiological phenomena of such effects but there is still no consensus about the mechanisms of coupling between the different effects [15].

The well-known models describing the emergence and propagation of the action potential are of the type of partial differential equations (PDEs), i.e. instead of a discrete description (the system of ordinary differential equations – ODEs), the continuum approach is used by Hodgkin [16], Nagumo et al. [17], etc. However, these PDEs can describe also the possible ionic (i.e. molecular) mechanisms needed for the propagation. In continuum physics, such presentations have been well elaborated since the original ideas of Augustin-Louis Cauchy in the 19th century. Contemporary ideas of how the discrete structure of the matter can be described by a continuum which fills completely the space it occupies are presented by Askar [18] and Maugin [19] among others. Continuum mechanics has a good thermodynamical foundation [20] and many practical applications [1,2], including the description of microstructured solids [21]. Within the framework of continuum theories, it is possible to describe the coupling effects of several fields, the effects of dispersion and dissipation, as well as the effects of singularities [2]. It is of great interest to systemize the basic laws of physics within the framework of continuum mechanics for the description of nerve impulses, especially for modelling the coupling of various accompanying effects and the effects of microstructures.

The present review provides a fresh look into mathematical modelling of signal propagation in nerve fibres, demonstrating how the concepts of continuum mechanics can help not only the understanding of this fascinating phenomenon but also grasping the complexity of the process. In Section 2 the main principles of continuum mechanics are presented with attention to wave motion and wave equation. The importance of the Fourier’s law and the Joule’s law is stressed and the basic wave equations for mechanical and electrical waves are presented. The notion of internal variables is introduced and several coupling mechanisms of various fields known in mechanics are briefly described. Section 3 is a brief overview of mathematical solutions and effects described by differential equations that could be used further for modelling of biological processes. Section 4 is the central part of this review, which demonstrates how the ideas briefly described from the viewpoint of physics and mathematics in Sections 2 and 3 are used for modelling signal propagation in nerves. Section 5 presents the conclusions, demonstrating the results of interdisciplinarity.
2. BASIC IDEAS FROM PHYSICS FOR DYNAMICS AND COUPLING

Dealing with dynamical processes in time and space means dealing with wave motions. In continuum mechanics it is assumed that the matter is continuously distributed and deformable while a wave is defined as a state moving into another state [20]. The analytical basis of continuum mechanics is well elaborated, see for example [22]. The cornerstones of the analysis are the conservation laws: the conservation of mass, the balance of momentum, the balance of moment of momentum (the balance of angular momentum), the conservation of energy (the first principle of thermodynamics), as well as the entropy inequality [22,23]. For describing the wave motion, the balance of momentum or Newton’s second law is the fundamental starting point: the acceleration $a$ is proportional to the magnitude of the force $F$ acting on the body. This means that in the 1D setting

$$F = ma, \tag{1}$$

where $m$ is the mass of the body. The usage of the 1D setting is justified because a nerve can be taken as a long 1D object (more on the structure in Section 4). The next step is to propose constitutive equations that are stress-strain relations and then the balance of momentum can be expressed in terms of displacements as wave equations. In the simplest 1D case with a linear stress-strain relation, the outcome is the classical wave equation

$$(\lambda + 2\mu)u_{xx} - \rho_0 u_{tt} = 0. \tag{2}$$

Here $u$ is the displacement, $\lambda$ and $\mu$ are the Lamé coefficients reflecting elasticity, $\rho_0$ is the density reflecting inertia and the indices $x$ and $t$ denote the differentiation with respect to space and time, respectively. The wave propagates with the finite velocity $c (c^2 = (\lambda + 2\mu)/\rho_0)$. The notions of elasticity, inertia and finite velocity are characteristics of the wave motion – see a more detailed description by Engelbrecht [24].

The conservation laws and constitutive equations are complemented by axioms of the constitutive theory [23]. Of the nine axioms, the axioms of causality, determinism, equipresence, and admissibility should be underlined. They mean respectively: the motion and temperature are self-evident and observable; every variable is determined by its history; all the response functionals must be considered as being dependent on the same list of constitutive variables; constitutive equations must be consistent with the balance laws and the entropy inequality. In order to reflect better the reality, the classical wave equation must be modified but the requirements for the conservation laws, constitutive equations and axioms must always be satisfied.

Bearing in mind the physical processes in nerves, besides Newton’s second law, two other laws of physics are of importance that relate heat to the electric current and temperature. First, the Joule’s law states that the amount of heat $q$ is proportional to the square of the electric current $I$:

$$q = I^2 R t, \tag{3}$$

where $R$ is the resistance and $t$ is the time.

Second, according to the Fourier’s law the heat flux $q$ is proportional to the negative thermal gradient, the proportionality constant $k$ is the thermal conductivity:

$$q = -k \Theta_x, \tag{4}$$

where $\Theta$ is the temperature and the index $x$ denotes the differentiation, as indicated above.

Johnston and Wu [25] mention also the laws governing the ion motion: the Fick’s law which governs the diffusion of particles caused by concentration differences and the Ohm’s law governing the drift of ions due to potential differences. We shall not elaborate here further on the usage of these laws.

The laws of physics are universal and must be satisfied for all the processes in matter. The important notions are electric current, temperature and heat.

Besides the wave equation (2) describing mechanical waves, the transmission of the voltage and current is described by telegraph equations. First proposed by Oliver Heaviside in 1876 for the transmission of electromagnetic waves in a wire, these equations also form the basis for the propagation of an electric signal
The telegraph equations for an axon are written in the form of two first-order differential equations:

\[ \pi a^2 C_a \frac{\partial v}{\partial t} + \frac{\partial i_a}{\partial x} + 2\pi a I = 0, \]  
\[ L \frac{\partial i_a}{\partial t} + \frac{\partial v}{\partial x} + \frac{R}{\pi a^2} i_a = 0, \]

where \( v \) is the potential difference across the biomembrane, \( i_a \) refers to the axon current per unit length, \( I \) is the ion current density while the coefficients are: \( C_a \) – the axon self-capacitance per unit area per unit length, \( L \) – the axon specific self-inductance, \( R \) – the axon specific resistance. The spatial coordinate is denoted by \( x \), time by \( t \) and \( a \) indicates the axon radius. It is possible to rewrite the system (5), (6) in the form of a second-order equation:

\[ \frac{\partial^2 v}{\partial x^2} - \frac{L C_a}{\pi a^2} \frac{\partial^2 v}{\partial t^2} = R C_a \frac{\partial v}{\partial t} + \frac{2}{a} R I + \frac{2}{a} L \frac{\partial I}{\partial t}, \]

This is also a wave equation (cf. with Eq. (2)) but with losses described by the terms at the r.h.s.

### 2.1. Modifications of governing equations

Clearly, in reality it means that one should take the dissipative effects and the possible heat generation caused by thermal conductivity into account. In addition, the internal structure of the material (although continuous) may influence the wave motion. Neither can one forget the possible body forces that act on the volume and affect the wave motion [22, 24]. The basic equations and the laws of physics presented above need to be modified in order to describe the complex situation and coupling in nerve fibres, bearing in mind that the whole process of signal propagation has an electro-mechano-thermal character [3, 4, 27]. Based on the general description of an axon (see Section 1), it is evident that the dynamic processes in it should apply the concepts of continuum mechanics which must certainly be consistent with physiological data from experiments. Thus, the biomembrane can be modelled as a tube with a microstructure, the axoplasm is a viscous fluid, the heat generation either from an electrical signal or mechanical processes should obey the Fourier’s law, etc.

First, all the real processes are dissipative and therefore the models of wave processes in axons should involve dissipative terms. Modelling of the main component of a signal – the action potential AP – is based on the telegraph equation (7), but with additional assumptions. According to Hodgkin [16], the inductance is neglected, resulting in the governing equation:

\[ \frac{\partial^2 v}{\partial x^2} = R C_a \frac{\partial v}{\partial t} + \frac{2}{a} R I, \]

with the ion current

\[ I = g_K n^4 (v - v_K) + g_{Na} m^3 h (v - v_{Na}) + g_L (v - v_L) + C_m \frac{\partial v}{\partial t}, \]

where \( C_m \) denotes the membrane capacitance while \( v_K, v_{Na}, v_L \) represent the equilibrium potential of the corresponding ions. The phenomenological (internal) variables \( n, m, h \) govern ‘turning on’ and ‘turning off’ of individual membrane conductance. In Eq. (9) \( g_K, g_{Na}, g_L \) refer to the maximum conductance of the corresponding ion components. This is a reaction-diffusion type equation which includes diffusion and reaction terms and leads to the formation of a travelling wave – the action potential (AP). The simplified model named after FitzHugh and Nagumo (the FHN model) involves only one ion current [17]

\[ \frac{\partial^2 u}{\partial s^2} = \frac{1}{c} \frac{\partial u}{\partial t} - w - \left( u - \frac{u^3}{3} \right), \]
\[
\frac{c}{\rho} \frac{\partial w}{\partial t} + bw = a - u.
\]

Here the original notations are preserved: \( s \) and \( t \) are the space and time coordinates, \( u \) is the membrane voltage and \( w \) is the recovery current. The constants \( c \), \( b \) and \( a \) are positive, satisfying

\[
1 > b > 0, \quad c^2 > b, \quad 1 > a > 1 - \frac{2}{3}b,
\]

while \( h \) is related to the resistance [17]. It is stated [17] that the variables \( u, w \) in Eqs (10), (11) correspond to the pairs of variables \((v,m)\) and \((h,n)\) in Eqs (8), (9), respectively. However, FitzHugh [28] stresses that such modelling describes only qualitative similarity to the Hodgkin model, i.e. Eqs (8), (9). Engelbrecht et al. [29,30] have later used the FHN model in the dimensionless form for several numerical simulations where the main goal was to model the coupling effects.

Second, the waves in the cylindrical biomembrane can be modelled by a wave equation like Eq. (2) but taking into consideration the microstructure of the lipid bilayer, as well as nonlinearities. Heimburg and Jackson [31] have proposed to model longitudinal waves in biomembranes by the equation

\[
u_{tt} = \left[ \left( c_0^2 + pu + qu^2 \right) u_s \right]_x - hu_{xxxx},
\]

where \( u \) denotes the change of density \( -\Delta \rho^A = \rho^A - \rho_0^A \), \( c_0 \) is the sound velocity in the unperturbed biomembrane, \( p \) and \( q \) are the nonlinear coefficients (determined experimentally) related to the change of density which determines the velocity \( c (c = c_0^2 + pu + qu^2 + ... \), and the subscripts \( x, t \) refer to spatial and temporal partial derivatives, respectively. Equation (13) is of the Boussinesq-type [32] and its form shows that the nonlinearity is of the displacement-type (depending on \( u \)). This is obvious from the definition of the variable \( u \) – the change of density in the longitudinal direction. The coefficients \( p < 0 \) and \( q > 0 \) have the dimensions of \( c_0^2/\rho_0^A \), \( \rho_0^2/\rho_0^4 \), respectively and \( h > 0 \) has the dimension of \( m^4/s^2 \) [31].

This model, however, does not describe the microstructure (phospholipids in the biomembrane) properly. Based on the studies of microstructured materials [21], Engelbrecht et al. [33] have proposed an improved model

\[
u_{tt} = \left[ \left( c_0^2 + pu + qu^2 \right) u_s \right]_x - h_1u_{xxxx} + h_2u_{xxxx},
\]

with two dispersive terms at the r.h.s. This equation considers both important properties of the lipid bilayer – elasticity (modelled by \( u_{xxxx} \)) and inertia (modelled by \( u_{xxxx} \)). Here \( h_1 = h \) in Eq. (13) and \( h_2 \) has the dimension of \( m^2 \) [33].

Concerning the heat and temperature generated in an axon, the Fourier’s law leads to the heat equation

\[
T_t = \frac{k}{C\rho} T_{xx},
\]

where \( k \) is the conductivity, \( C \) is the heat capacity, \( \rho \) is the density of the axoplasmic fluid, \( T \) is the temperature, and as described above, the subscripts \( x \) and \( t \) denote the corresponding partial derivatives. This is a parabolic equation, one of the basic equations of mathematical physics. Here we assume that the bilayer is so thin that the main change in temperature occurs in the axoplasmic fluid. Note that while temperature changes might occur both in the lipid bilayer and in the axoplasm through various mechanisms like Joule heating, mechanical deformations or chemical reactions, the medium through which the thermal signal reaches the location of measurement in the experiments recorded in the published literature is the axoplasm [34,35]. Moreover, it has been shown that the temperature within a typical axon cross-section equals much faster (\( \approx 0.6 \mu s \)) [14] than the typical duration of a nerve pulse (which is typically several milliseconds). However, as in experimental setups the measurement location is of some distance away from the axon wall, what can be measured, typically, is the ‘integral’ signal of all the thermal processes combined together. Not all the phenomena of heat generation can be described by Eq. (15) or similar equations. Some experiments [1] have demonstrated slow relaxation, for which it was only suggested that exo- and endothermic chemical
reactions may cause such an effect. We shall give a possible explanation of this phenomenon in Sections 2.2 and 2.3.

One of the well-studied objects in mechanics is the 1D rod which can be compared with the 1D axon. Disregarding the detailed analysis [36] which started with the studies of Gustav Kirchhoff in the mid-19th century, the main question is how the longitudinal deformation of a rod is related to the transverse deformation. For asymptotic models of rods, the Rayleigh–Love correction is used, which assumes that the transverse displacement $w$ along the radial axis $r$ of a rod is related to the longitudinal strain $u$ along the axis $x$ (see, for example, [37] for a detailed description of waves in rods, as well as [38]):

$$w = -\nu ru_x,$$  \hspace{1cm} (16)

where $\nu$ is the Poisson coefficient. Expression (16) along with considering the shear leads to proper dispersion effects of longitudinal waves [38]. Physically it means that the longitudinal and transverse deformations are coupled and expression (16) works in both directions – any longitudinal deformation is accompanied by the transverse deformation and vice versa. This is an important remark and will be elaborated in Section 3.

Finally, one should note that in continuum mechanics the body forces play an important role [22]. They originate from other fields and therefore their presence could model the interactions.

### 2.2. Internal variables

In order to model phenomena hidden in terms of direct measurements, the concept of internal variables is introduced in continuum mechanics. This concept has proved to be useful also in modelling the nerve pulses [39], which is why the ideology of internal variables is briefly presented. The idea is traced back to P. Duhem, P. Bridgman, and J. Kestin (see the overview by Maugin [40]). Leaving aside the thermodynamical considerations, this concept makes a clear distinction between the variables that can be measured, i.e. are observable, and the variables that are hidden, called internal.

Such a concept is widely used in continuum mechanics with regard to modelling of processes in liquid crystals, localization of damage, phase transitions, description of dissipative structures, relaxation oscillations, etc. [40]. Internal variables describe the influence of the underlying hidden processes and in this way compensate for the lack of knowledge of the precise description. The formalism of using the internal (hidden) variables was derived by Maugin [41].

Let us denote the internal variable by $\alpha$ and the observable variable by $\chi$. Then the simplified description [41] involves the governing equations presented below. For the dependent variable $\sigma$ one has in original notations

$$\sigma = \sigma(\chi, \alpha),$$  \hspace{1cm} (17)

which means that the dependent variable is a function of two independent variables – the internal and the observable. This equation must be complemented by the governing equation (evolution equation) for $\alpha$

$$\dot{\alpha} = f(\chi, \alpha) + g(\chi, \alpha)\dot{\chi},$$  \hspace{1cm} (18)

where $f(\ldots)$ and $g(\ldots)$ are certain functions [41] and the dot indicates the time derivative. Quite often, $g(\ldots) = 0$. The formalism for deriving a governing equation like Eq. (18) is based on exploiting the balance of energy (kinetic and potential energies) and the dissipation potential in terms of the observable and internal variables [41,42].

### 2.3. Physical effects

Following the Hodgkin–Huxley paradigm which states that the main carrier of information is the electrical signal (AP), the coupling effects are caused by the AP or the accompanying ionic currents. Thus, the starting point is:
(i) the generation and formation of an electric signal.
Given the structure of an axon (see Section 1), the possible physical mechanisms of interaction are the following:
(ii) electric – solid (biomembrane) interaction resulting in a mechanical response;
(iii) electric – fluid (axoplasm) interaction resulting in a mechanical response;
(iv) electric – solid (biomembrane) interaction resulting in a thermal response;
(v) electric – fluid (axoplasm) interaction resulting in a thermal response.
Several ideas are suggested to model those interactions in order to build a general model of nerve signalling.

First, (i) the generation of an electrical signal. The celebrated Hodgkin–Huxley model for describing the ionic mechanisms responsible for pulse formation [43] involves Eq. (8) with the ion currents (9). These currents depend on the Na and K ions and their 'turning on' and 'turning off'. The conductances depend on the phenomenological variables (gating variables) \( n, m \) and \( h \), for which an empirical description was ingeniously proposed. These variables are actually internal variables [44]. Typically for internal variables, the limit values of those variables lie between zero and unity, i.e. between two levels. Each of them \( \alpha_i \) \( (\alpha_1 = n, \alpha_2 = m, \alpha_3 = h) \) is described by an evolution equation

\[
\dot{\alpha}_i = p_{i1}(1 - \alpha_i) - p_{i2}\alpha_i,
\]

or equivalently by

\[
\dot{\alpha}_i = -\frac{\alpha_i - \alpha_{i0}}{\tau_i},
\]

where

\[
\alpha_{i0} = \frac{p_{i1}}{p_{i1} + p_{i2}}, \quad \tau_i = \frac{1}{p_{i1} + p_{i2}},
\]

where the dot denotes the time derivative, \( \alpha_{i0} \) is the equilibrium value, \( \tau_i \) is the relaxation time, while \( p_{i1} \) and \( p_{i2} \) are coefficients.

Although not expressed in physiological studies, this forms a clear bridge between continuum mechanics and physiology. A simplified FHN model (10) with one ion current (11), as indicated earlier, is also widely used [17]. In some cases other ions should be taken into account, like in the Morris–Lecar model with the ion mechanism based on K and Ca\(^{2+}\) ion currents [45].

Whatever the models are – the HH model or the FHN model – the main features of the AP are taken into consideration:
(i) the existence of a threshold for an input;
(ii) the all-or-none phenomenon for a pulse;
(iii) the existence of an asymmetric localized pulse with an overshoot;
(iv) the existence of a refraction length;
(v) the possible annihilation of counter-propagating pulses.
The iconic shape of the measured AP by Hodgkin and Huxley in 1939 [46] has been verified by numerous later experiments.

A completely different idea for modelling signals in nerves is elaborated by Heimburg and Jackson [31] who consider a nerve pulse an ‘electromechanical soliton’. According to this idea, the main element of the signal is a longitudinal wave in the biomembrane (see Eq. (13)) which generates all the other phenomena (electrical signal, temperature). Although this model can describe the process from different aspects if compared with the HH model [47], it is not clear how the electrical signal is formed from the mechanical one. Heimburg and Jackson assume that the membrane potential is linearly proportional to the density change [31] but this assumption does not explain the measured asymmetric shape of the AP (soliton is strictly symmetric) or the refractive overshoot, which is a normal effect of either the HH or the FHN model. One should stress that the model for the longitudinal deformation of a biomembrane [31] is of great importance in the general theory of cells but its application in nerve pulse dynamics is still under discussion [15].

Second, (ii) the mechanism of interaction between the electrical signal and the deformation of the axon wall (the biomembrane). In principle, this is a problem of electroelasticity. For example, Maugin [19] has
derived a model for elastic ferroelectrics. In the 1D setting the model couples the longitudinal displacement \( u \) with the rotation angle \( \phi \) (dimensionless space-time coordinates \( x, t \) are used):

\[
utt - c_L^2 u_{xx} = -\frac{1}{2} \alpha \left( \phi_x^2 \right)_x, \quad (22)
\]

\[
\phi_{tt} - \phi_{xx} - \Psi \sin \phi = - \alpha \left( u_x \phi_t \right)_x, \quad (23)
\]

where \( c_L \) is the velocity, \( \alpha \) refers to the piezoelectric coefficient and \( \Psi \) denotes the electric susceptibility (original notations are used). Equation (22) is a wave equation with forcing and Eq. (23) is a sine-Gordon equation with forcing. Following this example, which is not related to propagation of nerve impulses, the longitudinal wave in the axon wall is described by the wave equation (14) with forcing, which depends on the amplitude of the AP and/or the ion current(s) – see Section 4.

Another possibility for the coupling mechanism is to use the concept of flexoelectricity \[48\], which has been suggested to model the coupling of electrophysiology and mechanics. Direct flexoelectricity means that the mechanical strain generates the electric polarization in dielectric materials (like liquid crystals). Reverse flexoelectricity signifies that an electric field in the material causes the deformation. In the case of an electric field across the biomembrane, the result is the change in the membrane surface curvature, i.e. the transverse displacement. Such a model is described by Chen et al. \[27\] by using the classical HH model for the AP and the viscoelastic Maxwell model for the membrane. The coupling force applied to the biomembrane depends on the local change of the membrane potential. It is also argued that the electrical-mechanical coupling might be caused either by electrostriction or piezoelectric effects \[49\] but these assumptions have not been validated by experiments.

Third, (iii) the interaction between the electrical signal and the pressure wave in the axoplasm. The experimental evidence of the pressure wave in the axoplasm \[50\] requires a clarification of the possible model of coupling. It is proposed that the dynamics of the axoplasm, which is a viscoelastic fluid, could be described by a wave equation with dissipation and forcing \[29\]. The forcing depends on the amplitude of the AP and/or the ion current(s) – see Section 4. As shown in experiments \[50\], the pressure wave is in phase with voltage changes. The possible faster parts of a generated pressure wave dissipate and do not carry a significant part of energy.

Fourth, (iv) and (v) the interaction between the electrical signal and temperature. Although much discussed, there is still no consensus about the possible mechanisms of heat generation during the passage of the AP. However, it is agreed that there are several mechanisms of interaction to be taken into account. Detailed experiments of Abbott et al. \[34\], Howarth et al. \[51\], Tasaki \[35\] have revealed the role of the Joule heating but the behaviour of the slow decay of thermal profiles has raised a serious discussion. It is suggested \[34\] that such a phenomenon is caused by chemical reactions. Here the ideas of continuum mechanics have proved to be useful: Engelbrecht et al. \[39,52\] have proposed using the concept of internal variables (see Section 2.2) to model the influence of exo- and endothermic reactions. The governing equation is the classical heat equation with forcing, which includes the Joule heating as well as an internal variable related to the endo- or exothermic reactions. The internal variable itself is governed by a usual kinetic equation depending on the ion current (like in the HH model) and needs a relaxation time as a parameter to control the decay. The modifications of such an approach include two internal variables if endo- and exothermic influences can be separated.

Several reviews should be brought into attention which summarize the analysis of physical effects: Tasaki \[35\] has collected the earlier results, later reviews are provided by Heimburg and Jackson \[53\]; Andersen et al. \[54\]; Appali et al. \[47\]; Drukarch et al. \[15\]. A general overview of the measured physical effects and possible driving forces following the principle of equipresence is presented by Engelbrecht et al. \[3\].
3. REMARKS ON MATHEMATICS NEEDED FOR MODELLING

The mathematical models describing propagation of signals in nerves are first of all partial differential equations. Much attention is paid to solving them and analysing their solutions [5,55–58]. An important part of these studies is related to stationary solutions, i.e. to ordinary differential equations. In this case, the phase-plane analysis has revealed many aspects of solution types and their behaviour.

As a matter of fact, one deals with nonlinear equations or systems of equations where the leading terms are of the second order but might as well involve the higher-order terms (in modelling dispersion). For solving such equations or systems of equations, numerical methods are widely applied and this trend is growing in biology [59]. Here we present only some remarks from the wide area of analysis needed for building a full model of a signal like a wave ensemble.

Both the original (13) and the improved (14) HJ models fall under the Boussinesq’s paradigm [32]. The analysis of the improved HJ equation (14) has certainly demonstrated that the soliton-type solutions exist for equations like Eq. (13). However, the dispersion effects involving both the elasticity and inertia of the biomembrane (Eq. (14)) lead to the bounded solution and permit regulation of the width of the soliton needed to match the experimental results. In addition, it has been shown that depending on the coefficients, several types of solutions may exist, including trains of solitons, oscillatory solutions, and the soliton doublet [60,61]. It has also been demonstrated that as Eq. (14) is of the Boussinesq-type, the soliton type solutions are not fully elastic and lose energy at collisions due to radiation. These results may be of importance for cell mechanics in general. It is essential to indicate that the nonlinearities in Eqs (13) and (14) are of the displacement-type which is different from the usual models of continuum mechanics of solids where the nonlinearities are of the gradient-type. Certainly, this affects the formation of solitons or soliton trains. One important aspect must be stressed: the formation of a soliton or a solitary train from an arbitrary (not soliton-type) input takes time.

Some small technical details must also be noted. First, in the analysis (Section 4) the derivatives of field variables are used, such as $u_x$ and $u_t$, where $x$ and $t$ are the axial coordinates of an axon and time, respectively. One should note that in physical terms $u_x$ denotes a change along the axis and $u_t$ across the axis. This simple feature will be used further in constructing the coupling forces. Second, the signals in nerves are more or less of pulse-type, i.e. unipolar. The derivative of a unipolar pulse has a bipolar structure. If such derivatives are used to model the coupling forces, it means that they are energetically balanced by definition. This property is suitable for modelling adiabatic processes. In general, such mathematical details are analysed by Peets and Tamm [62].

4. NERVE PULSE MODELLING

Modelling of signals in nerves is at the interface of physics, mathematics and physiology [63]. The excellent experiments about axon physiology such as [12,13] need to be supplemented by the knowledge from other fields of research to build models which describe the formation and propagation of signals in nerves.

Based on the structure of an axon (see Section 1) and the knowledge about the physiological processes, the following basic assumptions are made:

(i) there is strong evidence on electrical transmittance of signals from one neuron to another [64,65], which means that electrical signals are the carriers of information [13] and trigger all the other processes in nerves;
(ii) the axoplasm in a nerve fibre can be modelled as a fluid where a pressure wave is generated due to the electrical signal;
(iii) the axon wall (biomembrane) can deform (stretch, bend) under mechanical impact [31];
(iv) the ion channels in the biomembrane can be opened and closed under the influence of electrical signals or mechanical input [66];
(v) not all the phenomena are described by observable variables but may be described by internal variables [19,39].
These assumptions mean that the Hodgkin–Huxley paradigm is taken as a basis for all the processes. As a result, one has to deal with an ensemble of waves which consists of the following constituents: the action potential AP with the ion current $J$ (or currents $J_i$); the longitudinal wave LW in the axon wall (biomembrane); the transverse wave TW (displacement) of the axon wall, the pressure wave PW in the axoplasm; the temperature change $\Theta$.

A different scheme of signal generation is proposed by Heimburg and Jackson [31] where the mechanical wave LW in the biomembrane has the leading role. The differences between the HH model and the Heimburg–Jackson model are analysed in detail by Appali et al. [47]. Subsequently we shall collect the ideas of modelling based on the assumptions listed above, using also the basic knowledge from physics and continuum mechanics (Section 2).

The trivial idea for constructing a mathematical model for the ensemble of waves is to collect all the single models (equations) describing the processes in a nerve fibre and couple them into a joint system. The crucial question is then how to model these coupling forces which should physically reflect the electromechanical (or mechano-electrical) transduction mechanisms.

The main hypothesis in such modelling is [67]: all mechanical waves in the axoplasm and in the surrounding biomembrane together with the heat production are generated due to changes in electrical signals (the AP and/or the ion currents) that dictate the functional shape of coupling forces. In principle, the reciprocity is also possible. Based on this hypothesis and remarks in Section 3, the functional shapes of coupling forces are proposed in the form of first-order polynomials of gradients or time derivatives of variables [30,67]. The detailed analysis of the equipresence of variables is carried out [30] along with the analysis of experimental results and possible physical mechanisms.

As a result of these assumptions, the main hypothesis and a careful analysis of physical mechanisms, a full mathematical model is proposed where all the components of the ensemble are described by their own governing equations united into a system by coupling forces [3,63]. The model involves the following governing equations and expressions:

(i) the equations for the AP with an amplitude $Z$ and the ion current $J$; at this stage this is the FitzHugh–Nagumo (FHN) model;
(ii) the improved Heimburg–Jackson equation with an amplitude $U$ for the LW in the biomembrane involving two dispersive terms, a dissipation term and a coupling force;
(iii) the wave equation for the PW in the axoplasm with an amplitude $P$ involving the dissipation term and a coupling force;
(iv) the diffusion equation for the temperature $\Theta$ involving a coupling force with an internal variable $Q$ governed by its own kinetic equation;
(v) the transverse displacement $W$ is proportional to $U_x$.

Note that here the variables are described by capital letters to denote dimensionless quantities used in the numerical calculations. The full description of this system with the analysis of coupling forces is presented by Engelbrecht et al. [3,39]. Some relevant remarks concerning the system:

(i) the AP is described by the FHN model which takes only one ion current into account (similar to the sodium current [17]); a possibility to distinguish between the voltage-sensitive and mechano-sensitive ion channels [66] is accounted for by different parameters governing the ion current;
(ii) the model at this step can be taken as a proof of concept and the possible next step in modifying the analysis is to use the HH model instead of the FHN model; it is also possible to use an experimentally measured profile of an AP for triggering the other components;
(iii) the fundamental novel hypothesis for constructing the joint model is related to proposing the coupling forces between the single components of the full signal; these forces depend on the changes of coupled signals i.e. on their derivatives, not on their amplitudes [67];
(iv) the properties of the biomembrane (elasticity, inertia, microstructure) are accounted for in calculating the LW and the TW [33];
(v) there are several physical mechanisms to model the temperature changes, including the Joule heating and the endo- and exothermic processes [39,52];
(vi) it is possible to adjust the parameters of the system for the temperature [68,69];
(vii) although such a model is a robust one, it describes qualitatively correctly the profiles of experimentally measured components of the signal [3,39];
(viii) it might be useful to distinguish between primary and secondary elements in the wave ensemble: the primary components (AP, LW, PW) are characterized by velocity while the secondary components (TW, Θ) have no characteristic velocity [62].

5. FINAL REMARKS

The mathematical model described above is not the only one proposed for describing signal propagation in nerves. There are several recent attempts to build a joint model which stresses one or another mechanism of interaction. In the following brief analysis, we leave aside the studies on possible ionic mechanisms as it is not the scope of this overview. The ‘soliton’ model of Heimburg and Jackson [31] has already been mentioned. A solitonic model is also used for the description of the AP in plant cells [70]. Actually, this model employs the approach from continuum theory by using the Lagrangian and Euler–Lagrange equations resulting in the kink-type solitonic solutions for describing potassium, chloride, and calcium fields. El Hady and Machta [71] have elaborated a mechanism of electro-mechanical coupling based on the assumption that the potential energy is stored in the biomembrane and the kinetic energy in the axoplasmic fluid. As a result, mechanical surface waves (relating to the surface of the fibre) are generated. The AP was taken without calculation as a Gaussian profile and the temperature changes were calculated by summing up the influence from the LW and the TW. Another coupled model of electrical and mechanical signals based on the spring-dampers (dashpots) system is proposed by Jérusalem et al. [72]. Chen et al. [27] have modelled the electrical and mechanical (the transverse displacement) effects by applying the idea of flexoelectricity. Rvachev [73] has proposed that the axoplasmic pressure pulse PW triggers all the processes. In this model, the PW activates first the Na$^+$ channels which generate the local HH voltage spike and then the outcome is the radial displacement of the lipid bilayer. The crucial problem in the modelling of the propagation of signals in nerves is related to the function of biomembranes. The studies of coupled electrical, chemical and mechanical signals in lipid monolayers [74,75] and bilayers [76] have explained the coupling mechanisms, the importance of nonlinearity and the existence of acoustic pulses at the interface. In this way, “a physical basis to describe acoustics in living and nonliving soft systems is provided” [76]. From the viewpoint of wave mechanics, the description of nonlinear fractional waves in phospholipid monolayers [74] might be a fruitful approach to grasp the properties of such layers. This is again an excellent example of how the knowledge from continuum mechanics could lead to a better understanding of the behaviour of lipid membranes which could be described as a transition between the hyperbolic (wave-type) and parabolic (diffusive-type) behaviours [77]. As indicated by Kappler et al. [74], this is really a new avenue of studies worth developing in more detail, for example modelling dissipative and dispersive effects in the biomembrane as well as the influence of the environment (extracellular fluid) on processes in a nerve fibre. Attention should also be drawn to an interesting analysis of possible capillary waves in cylindrical elastic tubes (like axons) [78].

Moreover, the analysis of the improved Heimburg–Jackson equation (Eq. (14)) demonstrates that such periodic waves in cylindrical biomembranes may exist under the assumptions made for deriving this model [60].

Given the variety of models and the effects described by them, Holland et al. [79] suggest using distinct models as “a comprehensive ‘mosaic’ framework”. However, the physical process of signalling in nerves involves all the effects, combining them into a whole. Consequently, it should also be possible to construct a joint mathematical model as suggested by Andersen et al. [54]. In the modelling described in Section 4, the path was chosen from the behaviour of structures to dynamics, trying to grasp the leading effects. In principle, the possible reciprocity is modelled by additional parameters which might be of importance in reality, especially in pathological cases.

The summary of the interdisciplinary experience from basic physics and mathematics to modelling of signals in nerves is the following:
(i) the signals in nerves constitute an ensemble of waves including electrical, mechanical and thermal parts along and across the axis of a nerve;

(ii) the governing equations for the elements of the ensemble stem from the laws of physics including the conservation laws of continuum mechanics;

(iii) the conservation laws of continuum mechanics form a consistent system which must be preserved in modifications by satisfying the axioms of the constitutive theory;

(iv) in dynamical processes every variation of fields acts as a stimulus (coupling force) to other fields;

(v) in the first approximation, the coupling forces are determined by first-order polynomials of gradients (space derivatives) or time derivatives of variables; gradients mean changes along the axis, time derivatives across the axis;

(vi) unipolar pulses have bi-polar derivatives and if considered as structural parts of coupling forces, are energetically stable;

(vii) the hidden (in terms of direct measurements) processes can be described by internal variables which need additional physical parameters to be determined (equilibrium level and relaxation time);

(viii) the mechanical waves in the biomembrane (the LW and the TW) are reciprocally coupled by the Rayleigh–Love correction \( \omega \) and \( u_0 \), meaning that the longitudinal deformation is always coupled with a transverse displacement and vice versa;

(ix) the axiom of equipresence must be followed (the elastic and inertial properties of the phospholipids must be taken into account simultaneously).

Note that the issue (iv) is a reformulation of the du Bois-Reymond law: “the variation of current density at any given time acts as a stimulus to muscle or motor nerve” (cited after Hall, [80]).

The focus of the modelling described above is on the coupling of different physical effects. For that purpose the AP is modelled by a simple FHN model which, however, may be changed to more exact models in order to grasp many ionic mechanisms responsible for the emergence of an AP. The \textit{in silico} experiments permit covering a large area of possible physical parameters in order to find suitable sets verified by experiments in various nerves \textit{in vivo} and \textit{in vitro}. This could provide also a route for estimating the values of many unknown parameters such as the relaxation times of endo- and exothermic reactions. The physical background of biological processes [4] requires even more attention and more experiments [81]. Theoretical studies on energy redistribution between the elements of the wave ensemble might also cast more light on interaction processes.

Surely, 	extit{modelling at the interface of physiology, physics and mathematics} provides more insight into the fascinating problem of nerve impulse propagation. In this context the US National Research Council Report [82] has indicated the importance of computational biology for biological discoveries. Mathematical modelling does not only provide a coherent framework for interpreting data but also opens many new avenues of knowledge, such as uncovering new phenomena to explore or identifying key factors or components of a system, which can be used to generate quantitative predictions or to enable to formalize intuitive predictions, etc. The modelling described in this review also follows these ideas. The fundamental process of nerve impulse propagation is crucial for better understanding conscience and the function of the human brain, for which also interdisciplinary approach is needed [83]. Enrico Fermi is accredited with the saying, “There are two ways of doing calculations in theoretical physics. One way, and this is the way I prefer, is to have a clear physical picture of the process that you are calculating. The other way is to have a precise and self consistent mathematical formalism” [84]. As living organisms might be a bit too complex for a “precise and self consistent mathematical formalism” with all the nonlinearities and interactions between scales, structures, and dynamical processes, then formulating a model based on the best “physical picture” appears to be our path to enlightenment.
ACKNOWLEDGEMENTS

This research was supported by the European Union through the European Regional Development Fund (Estonian Programme TK 124) and by the Estonian Research Council (project IUT 33-24). The publication costs of this article were covered by the Estonian Academy of Sciences. The authors would like to thank the reviewers for their extremely valuable and constructive remarks.

Jüri Engelbrecht (PhD, DSc) is currently Consultant and Professor Emeritus at Tallinn University of Technology. Previously, he was Head of the Centre for Nonlinear Studies (CENS) at the Department of Cybernetics at Tallinn University of Technology and held the position of President of the Estonian Academy of Sciences (1994–2004). His research interests focus on the field of nonlinear wave mechanics, complexity theory and biophysics. J. Engelbrecht is the author of several monographs on wave mechanics and has published a number of articles on complex systems. He has been active in EuroMech, IUTAM, ALLEA, ICSU and is a member of many scientific societies and academies. J. Engelbrecht is the recipient of several honours and awards from Estonia and other countries.

Kert Tamm (PhD) is currently Researcher at Tallinn University of Technology, previously held the position of Researcher at the Centre for Nonlinear Studies (CENS) at the Department of Cybernetics at Tallinn University of Technology. His research interests include wave mechanics in micro-structured solids, computational mechanics and biophysics. K. Tamm is engaged in the teaching of mechanics at the graduate level. He is a member of EuroMech.

Tanel Peets (PhD) is currently Researcher at Tallinn University of Technology, previously held the position of Researcher at the Centre for Nonlinear Studies (CENS) at the Department of Cybernetics at Tallinn University of Technology. His main research interests lie in wave mechanics, mathematical modelling and biophysics. T. Peets teaches physics at the undergraduate level and mathematical modelling at the graduate level. He is a member of EuroMech and a member of the Executive Committee of the Nordic Association of Computational Mechanics.

REFERENCES


58. Nelson, P. C., Radosavljevic, M., and Bromberg, S.
Pideva keskkonna mehaanika ja signaalid närviüududes

Jüri Engelbrecht, Kert Tamm ja Tanel Peets

Eksperimendid on tuvastanud, et elektrilistele signaalidele närviüududes kaasnevad ka mehaanikalised efektid ja temperatuuri muutus. Ülevaade kajastab füüsikaseaduste, pideva keskkonna mehaanika jäävusseaduste ja matemaatilise füüsika põhivõrrandite kasutamist närviüududes leviva mitmekomponendilise signaali modelleerimisel. Interdistsiplinaarsus võimaldab füüsika, matemaatika ja elektrofüüsioloogia põhimõistet ning ideed ühtseks tervikuks siduda.