

**CONTINUUM-MECHANICAL  
MODELS FOR  
QUANTUM MECHANICS**

**By David Delphenich**

**Neo-classical Press  
Spring Valley, OH  
2017**



# TABLE OF CONTENTS

	Page
<b>Introduction</b> .....	1
§ 1. The roots of quantum physics	1
§ 2. The rise of wave mechanics	5
§ 3. Electron models	7
§ 4. Continuum-mechanical models for wave mechanics	15
§ 5. Chapter summary	16
References	18
<b>Chapter I – Basic continuum mechanics</b> .....	22
§ 1. Extended objects and their deformations.	22
§ 2. Strain.	31
§ 3. Stress.	34
§ 4. Constitutive laws.	36
§ 5. Balance principles.	38
§ 6. Relativistic continuum mechanics	46
§ 7. The principle of virtual work	63
§ 8. The point-particle approximation	64
§ 9. Frame strain	67
§ 10. Cosserat media	75
References	77
<b>Chapter II. – Mechanics of waves</b> .....	79
§ 1. The state of a wave medium	80
§ 2. The kinematical state of a wave	88
§ 3. The dynamical state of a wave	96
§ 4. Derivation of wave equations	99
References	100
<b>Chapter III. – Variational field theory</b> .....	103
§ 1. Variations of fields on space-time regions	103
§ 2. The stationary action principle	105
§ 3. Symmetries of the action functional	110
§ 4. Examples of symmetries	113
§ 5. Gauge invariance	120
References	123
<b>Chapter IV – Non-relativistic, spinless particles</b> .....	124
§ 1. The time-varying Schrödinger equation.	124
§ 2. The Madelung-Takabayasi form of the time-varying Schrödinger equation.	129
§ 3. Planck’s constant as a density	139
§ 4. The Madelung-Takabayasi equations in density form	144

	Page
§ 5. The stationary Schrödinger equation.	147
§ 6. The Madelung-Takabayasi form of the stationary Schrödinger equation.	149
§ 7. Coupling to an external electromagnetic field.	151
§ 8. The introduction of vorticity.	162
References	166
<b>Chapter V – The nature of the quantum terms</b> .....	167
§ 1. The geometrical origins of the quantum stress.	167
§ 2. The Riemannian geometry of instantaneous dynamical spaces.....	170
§ 3. Clifford-Weitzenböck geometry.....	172
§ 4. The Clifford-Weitzenböck geometry of dynamical space.....	179
§ 5. The possible role of the quantum vacuum.	181
§ 6. The relationship between the Madelung transformation and the WKB method.....	187
§ 7. The optical interpretation of quantum mechanics	189
References	191
<b>Chapter VI – Relativistic, spinless particles</b> .....	193
§ 1. The massless, complex scalar wave equation	193
§ 2. The time-varying Klein-Gordon equation.	202
§ 3. The Madelung-Takabayasi form of the time-varying Klein-Gordon equation.	209
§ 4. The stationary Klein-Gordon equation.	218
§ 5. The Madelung-Takabayasi form of the stationary Klein-Gordon equation.	221
§ 6. Coupling to an external electromagnetic field.	224
§ 7. The introduction of vorticity.	235
§ 8. The quantum strain tensor	237
References	238
<b>Chapter VII – Non-relativistic, spinning particles</b> .....	239
§ 1. Electron spin	239
§ 2. The Pauli algebra	241
§ 3. Rotational mechanics with Pauli matrices	252
§ 4. The Pauli equation	258
§ 5. The Bohm, Schiller, and Tiomno form of the Pauli equation	267
§ 6. The method of bilinear covariants	277
§ 7. A more geometric conversion of the Pauli equation	280
References	285
<b>Chapter VIII – Relativistic rotational mechanics</b> .....	287
§ 1. Basic definitions and equations	287
§ 2. Thomas precession	294
§ 3. Fermi-Walker transport	297
§ 4. The Frenkel electron	300

	Page
§ 5. The Bargmann-Michel-Telegdi equations	307
§ 6. The extension $SU(2)$ to $SL(2; \mathbb{C})$	310
<b>Chapter IX – Relativistic, spinning particles.....</b>	<b>316</b>
§ 1. The relativistic Pauli equation	316
§ 2. The Clifford algebra of Minkowski space	324
§ 3. The Dirac equation	329
§ 4. Other forms of the Dirac equation	343
§ 5. Discussion	347
References	348
<b>Chapter X. – The continuum-mechanical form of the Dirac equation.....</b>	<b>350</b>
§ 1. Bilinear covariants of the Dirac wave function	351
§ 2. Yvon form of the Dirac equation	361
§ 3. Takabayasi form of the Dirac equation	366
§ 4. The Weysenhoff fluid	377
§ 5. Relativistic Cosserat media	382
References	383
<b>Epilogue.....</b>	<b>385</b>
§ 1. The introduction of nonlinearity	385
§ 2. The photon wave function	389
§ 3. The electromagnetic vacuum state and constitutive laws	390
References	393
<b>Appendix A – Tensors and differential forms on vector spaces.....</b>	<b>395</b>
§ 1. Tensor algebra	395
§ 2. Exterior algebra	402
§ 3. Exterior derivative	412
§ 4. Divergence operator	414
§ 5. Integration of differential forms	415
§ 6. Lie derivative	419
<b>Appendix B. - Differentiable manifolds.....</b>	<b>424</b>
§ 1. Topological spaces	424
§ 2. Differentiable manifolds	432
§ 3. Differential forms on manifolds	440
§ 4. Lie derivatives on manifolds	442
References	443
<b>Appendix C. – Lie groups, Lie algebras, and their representations.....</b>	<b>445</b>
§ 1. Basic concepts regarding Lie groups	445
§ 2. Fields as representations	448
§ 3. Representations of Lie algebras	460

	Page
References	468
<b>Appendix D. – The mechanics of points and rigid bodies.....</b>	<b>469</b>
§ 1. Point mechanics	469
§ 2. The theory of moments	475
§ 3. Rotation of a rigid body about a fixed point in space	477
<b>Appendix E. – Multiplication tables for the Clifford algebra of Minkowski space.....</b>	<b>485</b>

---

# INTRODUCTION

The focus of this book is a survey of the various attempts that were made since the emergence of wave mechanics to explain the otherwise-enigmatic nature of the quantum wave function by converting the quantum wave equation into a set of differential equations that look intriguingly similar to corresponding equations of continuum mechanics. Historically, that has always been referred to as the “hydrodynamical” interpretation of wave mechanics, but the author is of the opinion that the actual character of a continuous medium (i.e., fluid, anisotropic fluid, elastic solid, plastic, etc.) is something that is rooted in the mechanical constitutive law that associates dynamical states with virtual displacements of kinematical ones, so it would be premature to settle on any particular state of matter at this stage of our knowledge. Furthermore, when one is dealing with macroscopic matter, one finds that the constitutive character of the matter is generally explained by some sort of interaction model between its constituents, such as elastic collisions in gases or atomic bonds in crystal lattices. Since the electric and magnetic field strengths in the vicinity of elementary charges are quite intense, it seems unlikely that one is actually dealing with something as loosely-coupled as a liquid or gas at the level of elementary matter. Hence, the author prefers to speak of “continuum-mechanical” models, instead of “hydrodynamical” ones, and leave open the problem of establishing the deeper nature of the constitutive laws.

**1. The roots of quantum physics.** – In order to properly address modern quantum physics, one must first accept that its present state still represents a semi-empirical system of specialized models and algorithms that do not generally follow from some fundamental system of differential equations. Indeed, most modern quantum physicists have long since given up dreaming of such a system, and now tend to dismiss the search for such a thing as merely a “classical” problem, while quantum physics simply starts with the largely-algorithmic nature of the theory behind the data and hopes that with enough progress in the advancement of experimental physics, perhaps some insight into a better foundation might eventually emerge.

That is why it is essential to accept that the roots of quantum physics (and perhaps all physics) are found in the experimental phenomena that implied the need to adapt the existing theories. Hence, in its early days, progress in quantum physics was mostly a case of experiments leading theories. Nowadays, one sees more examples in which a theory suggests the possible existence of some hitherto-unobserved phenomenon that experimentalists might (or might not) look for. Sometimes, such as the case of the Higgs particle, the phenomenon is eventually observed, while in many other cases, such as magnetic monopoles, wormholes, and tachyons, experiments have yet to prove that they actually exist. However, that does not always deter theoreticians, as one previously saw in the fact that the continuing lack of any experimental confirmation of the existence of gravitational waves did not discourage many physicists from simply accepting that the limits of experimental physics still fell short of the necessary requirements. Fortunately, their faith was eventually vindicated.

*a. The breakdown of Maxwell's theory.* – Historically, one can make a case for saying that classical physics had to start adapting to the world of quantum phenomena because an increasing body of experiments was contradicting the wisdom of Maxwell's theory of electromagnetism. Of course, one might clarify that statement by saying that it was mostly contradicting the *linear* form of Maxwell's equations by moving into a realm in which the electric and magnetic field strengths were large enough that perhaps linearity was becoming an untenable assumption. That might be based in the quantum phenomenon of vacuum polarization, which we shall discuss in due course.

The earliest experimental contradictions to Maxwell's theory mostly related to the discrete structure of bound-state spectra of electromagnetic waves. For example, atomic spectroscopy was amassing an increasing volume of data regarding the spectra of electromagnetic waves that were emitted or absorbed by atoms, although the model for the atom itself had to evolve into something that would account for that fact.

A first definitive step was taken by Max Planck in 1900 [1] when he addressed the theory of black-body radiation. The classical theory of Rayleigh and Jeans was predicting a curve for the intensity of radiation from a black-body versus frequency that agreed for low frequencies, but diverged for high temperatures (later referred to as an “ultraviolet catastrophe”), while Wien had posed a curve with the opposite properties (i.e., an “infrared catastrophe”). The innovation that Planck introduced into the theory that corrected the theoretical curve was the assumption that the spectrum of electromagnetic waves in a black body was discrete, not continuous, and that the separation of the energy levels would be proportional to the separation of frequencies by way of a constant  $h$  that came to be known as Planck's constant. It is important to emphasize that the discreteness is due to the fact that one is again dealing with bound states.

*b. Wave/particle duality.* – Before one gets to the wave/particle duality of matter waves, one must first recall the wave/particle duality of light. The wave theory of light goes back to Christiaan Huygens [2] in 1690, although he imagined that it would propagate in a more mechanical way that would be analogous to the propagation of sound in elastic media, such as compressible gases. Sadly, his theory was immediately dwarfed by the popularity of Newton's corpuscular theory of light and optics [3] in 1704, which was mostly based in the prominence of the one historical figure over the other in the eyes of the scientific mainstream. Huygens's wave theory of light did not re-emerge until the work of Thomas Young and Augustin Fresnel in the early Nineteenth Century. To some extent, the difference between regarding light as a wave or localized point-particle amounts to the difference between wave optics and geometrical optics, respectively, and one must really regard geometrical optics as an approximation to wave optics.

However, two experimental phenomena that were pointing to a similar duality in the context of elementary matter were the photo-electric effect and Compton scattering. The latter effect was first discovered experimentally by Heinrich Hertz in 1887 and expanded upon by Philipp Lenard in 1902, and its theory was due to Albert Einstein in 1905 [4] (<sup>1</sup>), which was the main basis for his Nobel Prize in physics much later in 1921, while the

---

<sup>1</sup>) He referred to 1905 as his *annus mirabilis* (miraculous year), since he also published his first paper on special relativity, his paper on the equivalence of mass and energy, and a fundamental paper on the



former effect was studied by Arthur Compton [5] in 1923. Both effects related to the interaction of *photons* (a term that appeared some time later, and was coined by the chemist Gilbert N. Lewis) with atomic electrons. In the photo-electric effect, the interaction of a photon with an atomic electron caused the ionization of the electron – i.e., its liberation from a bound state – while in the Compton effect, the collision was not sufficient to cause ionization, but simply resulted in a transfer of momentum from the photon to the electron, with a subsequent loss of momentum by the photon, which was observed as a decrease in its wave number that is called the *Compton shift*.

Somewhat later, the experimental work by Sir George Thomson at the University of Aberdeen on the diffraction of electrons by thin metal films, and Clinton Davisson and Lester Germer at Western Electric on the diffraction of electrons by crystal lattices, which led to Thomson and Davisson being conferred the 1937 Nobel Prize in physics, added to the growing suspicion that elementary matter had a wave-like character, in addition to its point-like nature.

*c. The elementarity of matter.* – The idea that macroscopic matter is reducible to some ultimate, irreducible, “atomic” building blocks goes back to ancient Greece and the philosophers Leucippus (5<sup>th</sup> Century BCE) and Democritus (c. 460 – c. 370 BCE). Interestingly, that theory was eventually contradicted by the more influential voice of Aristotle, who felt that nature abhorred a vacuum, because in those materialistic days, everything had to be made of matter, so the non-existence of matter in a vacuum was tantamount to the existence of non-existence, which seemed to imply a logical ouroboros, at least to his way of thinking. He then concluded that atoms could not exist, because the space in between them would have to represent a vacuum, which was impossible to him. Sadly, that position became Holy Writ up until the Renaissance, along with Ptolemy’s picture of the solar system, and Euclid’s foundations for geometry. It was only much later in the mid-Seventeenth Century when Evangelista Torricelli exhibited the existence of a vacuum in the space at the top of an inverted glass tube of mercury that was closed at that end (i.e., a barometer) that the process of returning to the atomic hypothesis could begin.

Mostly, it was the study of chemistry that first led up to the acceptance of the atomic hypothesis as a reasonable basis for the nature of chemical reactions. A “periodic table” of the elements even emerged, which was due to Dmitri Ivanovich Mendele’ev in 1869, although an explanation for why it looked that way seemed quite puzzling. Nonetheless, one already had the elementarity of atoms, while molecules were composite bound states of atoms.

The first step towards a reduction of atoms came from observations that some atoms emitted *rays* of various kinds. For instance, the cathodes in vacuum tubes, which were heated by filaments and subjected to a negative electric potential (relative to the positive plate) emitted “cathode rays.” Henri Becquerel did experiments with more energetic rays that were emitted by radioactive salts around 1896, which had much in common with cathode rays. At the same time, Wilhelm Conrad Röntgen was doing experiments with rays that could make the bones of one’s hand visible when it was placed between a source

---

theory of fluctuations in Brownian motion, and was subsequently awarded a Ph. D. in physics by the University of Zürich.

of Roentgen rays (or “X-rays”) and a phosphor screen <sup>(2)</sup>. An increasing number of elements were found to be “radioactive,” such as uranium and radium, which was discovered by the Curies in 1898. The terminology of alpha, beta, and gamma rays emerged from that era.

Eventually, these rays were identified with elementary particles or composite states of them. Cathode rays were basically low-energy electrons, while Becquerel rays were high-energy ones, and beta rays were electrons generically. Alpha rays were composite states of two positively-charged protons and two neutral neutrons, which could also be regarded as a doubly-ionized helium atom. X-rays and gamma rays were seen to be simply two different energy levels for photons. Hence, the inventory of truly elementary particles was simply the electron, proton, neutron, and photon, although Wolfgang Pauli first proposed the existence of the neutrino theoretically in 1930.

As the experimental basis for the existence of elementary particles advanced, so did the models for the atoms themselves. At first, the fact that atoms were known to contain an equal number of positive and negative charges led to a “raisin-cake” model that put both kinds of charges together into one volume, where they were held together by Coulomb attraction. It was the experiments of Ernest Rutherford in 1909, along with Hans Geiger and Ernest Marsden, on the scattering of alpha particles by gold atoms that led to his 1911 revision of this picture into a positively-charged nucleus surrounded by electrons.

A second attempt at modeling the atom as a nucleus surrounded by electrons was made by Niels Bohr in 1913. He reasoned by analogy with the orbits of planets about stars and came up with a “planetary” electron model. However, since the centripetal acceleration of orbiting electrons would bring about the emission of electromagnetic waves according to Maxwell’s theory of electromagnetic radiation, with a concomitant loss of energy from the electron, one would expect that such electrons could produce only a continuous spectrum of emitted waves as they ran out of energy and crashed into the nucleus, like the degradation of a satellite orbit due to drag from the atmosphere. That would be inconsistent with the experimental facts of quantum physics for two reasons: Atomic spectra are not generally continuous, but discrete, and there seems to be a non-zero “ground state” for atomic electrons. Bohr simply got around those facts by introducing the discreteness in the model as an otherwise-unexplained axiom, and nonetheless came up with a formula for the energy levels of atomic electrons that actually explained quite a bit of spectroscopy. At any rate, one should note that the Bohr model pointed to a fundamental flaw in Maxwell’s theory of radiation, since apparently accelerated charges did not always emit electromagnetic waves, at least at the atomic level.

---

<sup>(2)</sup> In fact, Röntgen eventually lost several fingers from the malignant tumors that resulted from this now-ill-advised entertainment.

**2. The rise of wave mechanics.** – It was Louis de Broglie who first suggested in his doctoral dissertation [6] in 1924 <sup>(1)</sup> that massive matter could behave like waves, and that the angular frequency  $\omega$  of the wave that is associated with a mass  $m$  of kinetic energy  $E$  and linear momentum  $p$  would be equal to  $E/\hbar$  ( $\hbar = h / 2\pi$ ), while the wave number  $k$  would be equal to  $p/\hbar$ . That was consistent with the previously-established relationships that Einstein and Compton had obtained for photons.

By 1926, Erwin Schrödinger [7] attempted to build upon de Broglie’s concept of matter waves by devising a wave equation that would be obeyed by the wave function that would represent a matter wave. His basic intent was to use a pseudo-Hamilton-Jacobi approach to the equation that would make quantum wave mechanics relate to classical point mechanics in the same way that wave optics relates to geometrical optics. In fact, he himself admitted that there was a certain amount of trial and error associated with devising his equation, which now takes the form:

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \Psi + U \Psi, \quad (2.1)$$

in which  $\Psi$  is the complex-valued wave function for the matter of total mass  $m$ , and  $U$  represents its potential energy in the presence of a conservative external force.

That brings us to the main topic of this book: the problem of the correct physical interpretation for the wave function  $\Psi(t, x^i)$ . Typically, that mostly centers on the interpretation of the modulus-squared  $\|\Psi\|^2 = \Psi\Psi^*$  of the wave function, which will then be a real function of space and time. The one thing that all interpretations have in common is that they regard that real function as some sort of density.

Schrödinger himself originally thought in terms of charged particles, such as electrons, so he imagined that  $\|\Psi\|^2$  would describe the electric charge density. He eventually abandoned that idea, although it resurfaced some time later in the Pauli-Weisskopf interpretation of the Klein-Gordon equation, which we shall mention shortly.

The interpretation that seemed to catch on the most definitively was the one that was proposed by the Copenhagen School of quantum theory, which included Bohr, Max Born, Werner Heisenberg, and others, at the Fifth Solvay Conference on physics in 1927, was the statistical interpretation, which regarded  $\|\Psi\|^2$  as a probability density function. More specifically, it described the probability of finding a *point-like* mass  $m$  in a differential element of volume; i.e., its integral over a finite volume would give the probability of finding the point mass in that volume. Schrödinger and de Broglie, as well as Einstein, were originally reluctant to accept that interpretation. Einstein had famously said that “God does not throw dice,” to which Bohr retorted “Who are you to tell God what to do?”

---

<sup>(1)</sup> It is only a recurring urban legend that de Broglie’s thesis was a mere four pages long, which was probably the result of Nazi propaganda during the rise of that party in the 1930’s, which included various attempts to assert the supremacy of “Aryan” science. Indeed, to this day, there is a small, but vocal clique of physicists who will insist that de Broglie did not deserve a Ph. D. in physics for that thesis, much less the Nobel Prize. The author has never cleaved to that belief, since he has read, and even translated, enough of de Broglie’s work to know that he was one of the more radical and innovative thinkers of his era whose understanding of basic physics was beyond reproach.

In the same year that the Copenhagen School was establishing the statistical interpretation, Erwin Madelung [8] suggested an alternative interpretation that made  $n = \|\Psi\|^2$  a number density for an extended mass, so  $mn$  would represent its mass density. Furthermore, by essentially introducing polar coordinates into the field space  $\mathbb{C}$ , so  $\Psi$  would be expressed in the form  $R e^{i\theta}$ , with  $n = R^2$  and  $\theta = S/\hbar$ , which means that  $S$  plays a role that is similar to Jacobi's action function, one found that the complex Schrodinger equation (2.1) would give rise to a pair of real equations that took the form of a continuity equation for the mass density (or number density, for that matter) and another equation that looked like a Hamilton-Jacobi equation that included a potential energy term that seemed to embody all of the quantum content, and thus took on the name of "quantum potential." That paper became the seed from which the present work evolved eventually.

A closely-related problem to that of the interpretation of the wave function and wave equation is that of showing how quantum wave mechanics relates to classical mechanics. The *correspondence principle* of theoretical science, in general, says that any new theory must account for the successes of the previously-accepted theory, and ideally as something that remains as some fundamental parameter approaches a "classical limit." For instance, relativistic physics must coincide with non-relativistic physics in the limit as  $c$  becomes infinite. In the case of quantum mechanics, the classical limit is defined by letting  $\hbar$  go to zero. One also has *Ehrenfest's theorem* that quantum mechanics should give classical mechanics when one takes the means of the quantum expressions while using the probability density function that is defined by  $\|\Psi\|^2$ .

A recurring theme in this book will be the fact that there is a subtlety that was overlooked by the early founders of quantum mechanics that since the support of the wave function  $\Psi$  is not generally a single world line in space-time, but a world-tube, there is something hasty about calling point mechanics (i.e., a single world-line) the classical limit when it would be more methodical to first look at the continuum-mechanical picture that is defined by the world-tube. In effect, one first ignores the phase structure on the world-tube that is defined by the function  $\theta(t, x^i)$  by replacing it with the energy-momentum *density* 1-form  $p = \hbar d\theta$ , which also gives the flow covelocity 1-form  $v = p / m$  and the flow velocity vector field  $\mathbf{v}$  that is dual to  $v$  under the relevant space-time metric. One also has the aforementioned mass density  $\rho = m \|\Psi\|^2$ . In order to then go on point mechanics, one uses the well-established method of moments relative to  $\rho$ . That is, one gets the zeroth-order moment  $m$  by integrating  $\rho$  over all space (i.e., the support of  $\Psi$ ). In non-relativistic physics, one can then define the first moment in the form of the world-line of the center-of-mass (i.e., the mean position). The second moment of  $\rho$  is then the moment of inertia of the mass distribution, and so on.

Something that then becomes clearer when one interpolates this intermediate step of going from wave mechanics to continuum mechanics and then to point mechanics is that the classical observables that conventional quantum mechanics arrives at (i.e., eigenvalues of quantum operators) will always represent *total* quantities that are more appropriate to point mechanics, while the quantities that are appropriate to continuum mechanics are always *densities*. In fact, unless one is integrating a scalar or pseudo-scalar quantity (such as mass, charge, or energy density) over space, there is something less than rigorous about the common practice of integrating the components of vector and

tensor fields to get total quantities, such as total momentum, angular momentum, and the like. We shall have more to say about that topic as it becomes relevant.

One of the first extensions of the Schrödinger equation was to give it a relativistic formulation, which produced the Klein-Gordon equation, due to the efforts of Oskar Klein [9] and Walter Gordon [10] in 1926. There were two early objections to the resulting equation, namely, that its second-order character as a partial differential equation led to solutions with negative kinetic energy, which seemed physically absurd in the years before the experimental discovery of the positron and antimatter, more generally. Furthermore, the conserved current that was associated with the phase invariance of the action functional for the Klein-Gordon has a temporal component that is not positive-definite, and cannot be identified with a probability density function then. As a result, little serious consideration was given to the Klein-Gordon equation until Pauli and Weisskopf resurrected it in 1934 [11], along with Schrödinger's interpretation of  $e \|\Psi\|^2$  as an electric charge density.

One can also see that the non-positive-definite character of the temporal component of the aforementioned conserved current can be interpreted as a *reductio ad absurdum* of the statistical interpretation of  $\Psi$ . That was implicit in the Pauli-Weisskopf paper, since they pointed out that electric charge densities do not have to be positive-definite, by any means.

**3. Electron models.** – Since there were so few elementary particles that were known in the early days, the models for the particles themselves usually focused on the electron. (That was probably because it was easier to configure experiments to study the electron in the years before nuclear reactors.) Most theoreticians tended to assume that electrons were simply point-like objects with a certain mass and charge.

*a. Abraham-Lorentz-Poincaré model.* – However, the ones that were pursuing extended-matter models for the electron included Max Abraham [12], Hendrik Antoon Lorentz [13], and Henri Poincaré [14]. Interestingly, the work of Lorentz and Poincaré on the modeling of the electron was really more directed towards what would now be called a relativistic theory of the motion of matter, more generally, and thus overlapped strongly with the work of Einstein [15] on the electrodynamics of moving bodies.

A recurring theme in the Abraham-Lorentz-Poincaré models <sup>(1)</sup> was that in the eyes of Coulomb's law of electrostatic forces, an extended distribution of negative charge would have to be unstable in the absence of any other forces, since it would tend to expand indefinitely under its mutual electrostatic repulsion. Furthermore, part of the rationale for choosing an extended distribution was that if one wished to attribute all of the rest mass of an electron to the total energy that was contained in its electrostatic field then as one approached a point-like distribution, that total energy would diverge, not converge to the measured mass.

---

<sup>(1)</sup> For more modern discussions of the issues associated with classical electron models, see the books by Rohrlich [16], MacGregor [17], and Yaghjian [18], along with the survey article by Pearle [19].

*b. Discovery of electron spin.* – One of the key advances in quantum physics was the discovery that the electron had a magnetic dipole moment, which seemed to exist in two distinct states that were thought of as the “up” and “down” state. This was the conclusion of the *Stern-Gerlach* experiment [20] in 1922, which passed a beam of electrons through an inhomogeneous magnetic field and let them impinge upon a screen or emulsion. The result was that there were two distinct components to the pattern that was produced by the beam; i.e., a splitting of the beam. The subtlety in this result is that if one thinks of the electron as a small bar magnet then one would expect that the electrons would all interact with the magnetic field in the same way and produce a single component to the final pattern. The fact that it produced two is more reminiscent of the way that birefringent optical media will split an incoming light ray into two rays according to the state of the polarization that they are in.

The most plausible explanation for the existence of a magnetic dipole moment for the electron was that it was due to an *intrinsic angular momentum* – or *spin* – for the electron, which, unlike the conventional classical kind of angular momentum, existed in two distinct states, and not a continuum of intermediate states between them. Although this hypothesis is usually attributed to George Uhlenbeck and Samuel Goudsmit in 1925 [21], they themselves point out that it was proposed previously by Arthur Compton in 1921 [22], and had been suggested by Ralph Kronig in unpublished notes that Pauli had criticized for the fact that if the electron were a spinning charged sphere then the equatorial velocity that would be consistent with measured magnetic moment and the “classical electron radius” would be greater than that of light. Enrico Fermi and Franco Raseti [23] also commented on the magnetic field of the electron and the fact that it would define a different radius for the electron than the one that Abraham, Lorentz, and Poincaré derived from purely electrostatic considerations. Of course, despite the lofty place of the latter three savants in the history of science and mathematics, their classical model of the electron must nonetheless be regarded as incomplete for its lack of a magnetic moment, although that was not their oversight, but simply a limitation of history itself.

*c. The Pauli equation.* – Wolfgang Pauli was the first to attempt to extend the (non-relativistic) Schrödinger equation to include the possibility of particles with spin in his 1927 paper [24]. In effect, the ordinary Zeeman effect represented the splitting of spectral lines for atomic electrons in an external magnetic field that coupled to their *orbital* angular momentum, while the anomalous Zeeman effect represented a further splitting that was harder to explain until one attributed *intrinsic* angular momentum to the electrons, as well. An important advance for quantum wave mechanics was Pauli’s decision to model the splitting of that spin into an up and down state as something that came from the two-to-one map of three-dimensional Euclidian rotations, when represented by  $2 \times 2$  complex unitary matrices with unity determinants [i.e., elements of  $SU(2)$ ], to the same rotations when they are represented by  $3 \times 3$  real orthogonal matrices with unity determinants [i.e., elements of  $SO(3)$ ]. The resulting quantum wave functions  $\Psi$  took their values in  $\mathbb{C}^2$ , which carries the “defining” representation of  $SU(2)$ , and were referred to as *Pauli spinors*. The resulting extension of the Schrödinger equation, which came to be called the *Pauli equation*, not only extended the field space of  $\Psi$  from  $\mathbb{C}$  to

$\mathbb{C}^2$ , but coupled an external magnetic field to the magnetic moment of the charged, spinning particle by way of a set of three  $2 \times 2$  complex Hermitian matrices  $\{\sigma_i, i = 1, 2, 3\}$  with trace zero that allowed one to represent the spin of the wave function as a linear combination of those matrices, which then took on the role of a basis for a linear space, or really, the Lie algebra  $\mathfrak{su}(2)$ , which one gets by turning the Hermitian matrices into anti-Hermitian ones by the use of the imaginary  $i$ .

Heisenberg also commented on the application of the Pauli equation to the problem of the anomalous Zeeman effect in a 1926 paper with Pascual Jordan [25].

*d. The need to make the Pauli equation Lorentz-invariant.* – One could say that the Schrödinger equation was doubly-incomplete, since it was not Lorentz-invariant, and it did not account for the electron spin. Hence, although the Klein-Gordon equation accounted for the latter, but not the former, the opposite statement could be made for the Pauli equation. The challenge to theoretical physics at that point in history was then to find a single equation (or system of equations) that would accomplish both objectives.

Although there is now such a thing as the *relativistic Pauli equation* <sup>(1)</sup>, in the early days it was not clear how to develop such a thing, so it was passed over initially. Indeed, it is easier to see how to formulate a relativistic Pauli equation once one has first developed the Dirac equation. Basically, one simply extends the action of  $SU(2)$  on  $\mathbb{C}^2$  to an action of  $SL(2; \mathbb{C})$  on  $\mathbb{C}^2$ , and the two-to-one map of  $SU(2)$  onto  $SO(3)$  to a two-to-one map of  $SL(2; \mathbb{C})$  onto the identity component of the Lorentz group; viz.,  $SO^+(3, 1)$ .

*e. The Thomas-Frenkel relativistic, spinning, classical electron.* – In 1927 (a year before Dirac's seminal paper on the quantum theory of the electron), the English physicist Llewelyn Thomas devised a relativistic theory of the classical spinning electron [27]. It was still a point-like electron, so it did not attempt to explain how a point could rotate, but it did give a deep and compelling reason for the discrepancy of a factor of 2 in the gyromagnetic ratio that was pointed out by Uhlenbeck and Goudsmit: Thomas showed that the factor of 2 could be explained by the transition from the rotation group to the Lorentz group; i.e., it was an artifact of the relativistic formulation that did not exist in the non-relativistic formulation. As we shall see, it really comes down to the difference between the commutation laws for Lorentz boosts and the corresponding commutation laws for translations.

In 1929 (a year after Dirac's paper was published), the Russian Joseph Frenkel published [28] a paper in which he attempted to improve upon some of the limitations of Thomas's theory. Once again, his model still assumed a point-like electron, but it did introduce the relativistic formulation of electromagnetism in terms of what would now be called exterior differential forms that is mainly due to Minkowski.

*f. The Dirac electron.* – Due to the lack of interest in formulating a relativistic Pauli equation, the way that events played out historically was that in 1928, Paul Dirac

---

<sup>(1)</sup> One might confer the author's discussion of that equation in [26] and the references that are cited therein.

formulated a relativistic theory of the electron [29] that was explicitly intended to be Lorentz-invariant and account for electron spin. As was the custom in early quantum theory, he proceeded in a somewhat heuristic way by starting with the second-order Klein-Gordon equation and looking for a way of taking the “square root” of the Klein-Gordon operator, which one can think of as  $\square + m_0^2$ , where  $\square = \eta^{\mu\nu} \partial_\mu \partial_\nu$  is the d’Alembertian operator [with the sign convention  $(+, -, -, -)$  for  $\eta^{\mu\nu}$ ],  $m_0$  is the rest mass of the particle (which is assumed to be point-like) that is described by the wave function, and one must use “natural” units, which make  $\hbar = c = 1$ . If one does not use natural units, which is often advisable when one is discussing the propagation of waves and does not want to lose sight of some fundamental properties of the waves, then  $m_0$  would be replaced with the *Compton wave number*, namely,  $k_0 = m_0 c / \hbar$ .

Actually, it is not rigorously correct to say that the Dirac operator  $\not{\partial} + ik_0$  is the square root of  $\square + k_0^2$ , in the sense that  $(\not{\partial} + ik_0)^2 = \square + k_0^2$ , but rather, one must have:

$$(\not{\partial} + ik_0)(\not{\partial} - ik_0) = \not{\partial}^2 + k_0^2 = \square + k_0^2, \quad (3.1)$$

which is then equivalent to:

$$\not{\partial}^2 = \square. \quad (3.2)$$

Hence, one is basically looking for a square root of the d’Alembertian operator.

If one assumes, naively, that  $\not{\partial} = \gamma^\mu \partial_\mu$ , where the coefficients  $\gamma^\mu$  belong to some as-yet-unspecified algebra then (3.2) will become:

$$\frac{1}{2}(\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu) \partial_\mu \partial_\nu = \eta^{\mu\nu} \partial_\mu \partial_\nu,$$

which can be satisfied iff:

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu}, \quad (3.3)$$

which can also be expressed in the form:

$$(\gamma^0)^2 = 1, \quad (\gamma^i)^2 = -1, \quad i = 1, 2, 3, \quad \gamma^\mu \gamma^\nu = -\gamma^\nu \gamma^\mu, \text{ when } \mu \neq \nu.$$

However, (3.3) is precisely how one defines the *Clifford algebra*  $\mathcal{C}(4, \eta)$  of Minkowski space  $\mathfrak{M}^4 = (\mathbb{R}^4, \eta^{\mu\nu})$ . It is a 16-dimensional real algebra that is generated by the four linearly-independent vectors  $\gamma^\mu$  in  $\mathbb{R}^4$ ; that is, a basis for the vector space that the algebra is defined over is defined by 1, the  $\gamma^\mu$ , and all linearly-independent products  $\gamma^\mu \gamma^\nu$ ,  $\gamma^\lambda \gamma^\mu \gamma^\nu$ ,  $\gamma^\lambda \gamma^\lambda \gamma^\mu \gamma^\nu = \gamma^0 \gamma^1 \gamma^2 \gamma^3 = \gamma^5$ . Dirac chose to represent  $\mathcal{C}(4, \eta)$  by  $4 \times 4$  complex matrices, not all of which were Hermitian. Hence, since the matrix algebra of all  $4 \times 4$  complex matrices has a complex dimension of 16, and thus, a real dimension of 32, while  $\mathcal{C}(4, \eta)$  has a real dimension of 16, one sees that the Dirac matrices do not give a faithful representation of the one algebra in the other, but only in a linear subspace of the



latter. Other representations of the  $\gamma$  matrices evolved in time, and the most-used ones were due to Hermann Weyl and Ettore Majorana.

Right from the beginning, the Dirac equation touched off a firestorm of objections, most of which were based in the fact that its introduction of such an abstract mathematical structure as a Clifford algebra gave it a much-less-than-intuitive character, so the physical interpretation of the Dirac wave function, which took its values in  $\mathbb{C}^4$ , the equation itself, and how one was to model basic physical processes using that equation were the topic of spirited debates.

Various researchers attempted to give a tensorial form for the Dirac equation, which generally involved the use of *bilinear covariants* – i.e., bilinear expressions in the field  $\Psi$  and its Hermitian conjugate  $\Psi^\dagger$ , or rather its Dirac conjugate  $\bar{\Psi} = \Psi^\dagger \gamma^5$  – and the  $\gamma$  matrices. Darwin [30] proposed a widely-cited component form of the Dirac equation that did not introduce the  $\gamma$  matrices explicitly, but only their components, in effect.

One of the widely-debated issues regarding the Dirac equation was the question of true field space for the wave function. The only significance of  $\mathbb{C}^4$  was the fact that it carried a representation of  $\mathcal{C}(4, \eta)$ , although not a faithful one. Other quantum theoreticians, such as Alexandru Proca [31], and later Sir Arthur Stanley Eddington [32], pointed out that one could also say that since  $\mathcal{C}(4, \eta)$  carried a representation of itself (by either left or right translation), there was nothing to prevent one from using  $\mathcal{C}(4, \eta)$  as the field space. For Cornelius Lanczos [33], the appropriate field space was the eight-real-dimensional algebra of complex quaternions, whose unit vectors define a Lie group that is isomorphic to  $SL(2; \mathbb{C})$ .

Finally, there was the fact that despite having reduced the order of wave operator from two to one, the Dirac equation still admitted negative-energy states, and their physical interpretation was still just as puzzling. Here, one must realize that although the Dirac *operator* is a first-order, linear differential operator, nonetheless, the Dirac *equation* is not a first-order partial differential equation for a complex-valued wave function, but a *system* of four such entities. Hence, one has not really reduced the order of the Klein-Gordon equation from two to one, since one can always replace a single  $n^{\text{th}}$ -order differential equation (ordinary or partial) with a system of first-order equations by introducing auxiliary variables that represent the higher derivatives of the function that one is solving for.

*g. The discovery of the positron.* – After a few years, the debate regarding the Dirac equation cooled considerably, since the existence of the negative-kinetic-energy states was seen to be the most physically damning aspect of that system. Mostly, particle physics in that era was limited by its small inventory of known elementary particles, and all attempts to make the Dirac equation account for protons as the positively-charged counterparts to electrons were clearly doomed, since the mass of a proton is about 1860 times the mass of an electron, from the outset. According to Miller [34], Pauli had become so disenchanted with quantum electrodynamics that for a while he had drifted into the underground café culture of Europe, and was even toying with the idea of directing a movie! However, Hermann Weyl wrote the first edition of his classic book

[35] on the theory of groups and quantum mechanics in that period of time, in which he had the admirable objectivity to suggest that the case for the Dirac equation was not closed, since one could not anticipate what sort of advances that experimental physics would bring.

Indeed, it was the discovery of the positron by Carl David Anderson in 1932 [36] <sup>(1)</sup> that essentially got everyone back to work, because the positron made a much better candidate for the negative-energy counterpart of the electron than the proton. It was also discovered that the collision of an electron and positron (which later came to be called *Bhabha scattering*) in the presence of some other mass, such as an atomic nucleus, could generally produce two gamma photons, while annihilating the original incoming massive particles. Conversely, the reverse process of two colliding gamma photons could produce an electron-positron pair. Indeed, if the photon energies are sufficient, one can also produce particle-anti-particle pairs of higher masses, such as muons and anti-muons, or even strongly-interacting particles.

These processes of pair creation and annihilation prompted Dirac to come up with his *hole theory* of electrons and positrons. One imagined that the negative-energy states, which represented positrons, also represented “holes” in the quantum electromagnetic vacuum, or “Dirac Sea,” so annihilating a positron would be the same thing as filling in a hole. The vacuum state was defined to be the state in which all of the holes were filled, although that seemed to suggest that the vacuum state would have be infinitely massive, infinitely charged, and infinitely energetic. This was the first point at which quantum electrodynamics started to exhibit unphysical infinities, in addition to the infinite self-energy of the point-like electron that one found in classical physics already.

One would also have a somewhat-intermediate state between the creation and annihilation of free particle-anti-particle pairs that one called *vacuum polarization*, and which amounted to the creation and annihilation of *virtual pairs*, which would be essentially unstable bound states of particles and their anti-particles. This phenomenon permeates almost all of the explanations of quantum electrodynamics, so one can say with some security that the most fundamental problem in that branch of physics is to completely account for the structure of the quantum electromagnetic vacuum state (or probably *state space*). One of the consequences of vacuum polarization that has yet to be observed directly in experiments is *photon-photon* scattering, in which the vacuum polarization that is produced at the moment of collision would produce a nonlinear contribution to the scattered photons. However, a closely-related process called *Delbrück scattering*, in which a photon is scattered by the intense electrostatic field of an atomic nucleus as a result of vacuum polarization, has been observed experimentally, so one tends to assume that the limitations to the observation of photon-photon scattering are mainly technological in character.

*h. The emergence of QED.* – During the 1930’s, the increasing volume of theoretical and experimental work that was being done with the Dirac equation, as well as the discovery of new particles, was bringing about an increasing need for a unified theory of

---

<sup>(1)</sup> Apparently, the positron had been observed previously, first by Dmitri Skobeltsyn in 1929, who used a cloud chamber, and in that same year by a Caltech graduate student Chung-Yao Chao, although the latter observations were not pursued further. At any rate, it was Anderson who was awarded the 1936 Nobel Prize for the discovery.

such phenomena; i.e., a quantum theory of electrodynamics (QED) that would replace Maxwell's classical theory.

However, due to the complex nature of quantum electromagnetic vacuum state and the mathematics of the Dirac equation, it was increasingly believed that (at least, for the time being) it was unlikely that anyone could pose a convincing set of partial differential equations for some appropriate tensor or spinor fields that could truly be called the "fundamental equations of QED" in the same way that Maxwell's equations were fundamental to classical electrodynamics or Einstein's equations were fundamental to gravitation. The reasons for that were two-fold:

1. One might regard the presence of vacuum polarization as essentially a nonlinear contribution to the electromagnetic constitutive law that became significant at high-enough field strengths. Hence, in order to pose that system of equations, one would need to have a more detailed picture of that constitutive law, which was not (and is still not) available.

2. Even if one could correctly pose that system, it would undoubtedly represent a coupled system of nonlinear partial differential equations. Consequently, merely proving the existence of solutions to such elementary problems as boundary-value problems in the static approximation or dynamical solutions to the Cauchy problem would probably prove to be mathematically daunting, much less the problem of finding any useful closed-form solutions that might be the basis for explaining the structure of elementary particles and their interactions.

For the most part, it was the work of Werner Heisenberg, Wolfgang Pauli, and Pascual Jordan that proved to be most definitive in laying the foundations for QED. First, Heisenberg introduced the "exchange-particle" concept as a substitute for a detailed picture of the interaction of elementary particles. This amounted to eliminating any discussion of the "forces of interaction," in lieu of simply the exchange of particles that would mediate the interaction; for instance, the electromagnetic interaction would be mediated, not by some extension of Coulomb's law and the Biot-Savart law, but by the exchange of a photon. (Nowadays, the exchange particles are referred to as "gauge" particles.)

Closely-related to the exchange-particle concept was the passage to the *scattering approximation* as a way of linearizing the dynamics of time evolution. Basically, this amounts to replacing the Cauchy problem of describing the time evolution of incoming particle states (i.e., wave functions) from a finite initial time  $t_0$  to outgoing particle states at a finite final time  $t_1$  with the problem of describing the evolution of incoming states at  $t_0 = -\infty$  to outgoing states at  $t_1 = +\infty$ , which should turn the nonlinear time evolution operator  $\Phi(t_0, t_1)$  into a linear scattering operator  $S$ . This is equivalent to enclosing all of the nonlinear complexity into a "black box" that takes the form of the time interval during which the interaction of the particle states takes place, whose length is assumed to be infinitesimal in comparison to the time interval of the linear scattering process that involves only free particle states.

One of the advantages of linearizing the time-evolution operator to the scattering operator is that one can justify applying the methods of Fourier analysis. Indeed, almost

all of modern QED takes place in “momentum space,” which is Fourier transform of configuration space. The Fourier transform of the scattering operator then takes the form of an integral operator whose kernel, with suitable causality constraints, is thought of as a “propagator.” The construction of that propagator is usually approached perturbatively with the use of (Feynman) diagrams that represent the successive corrections to the elementary scattering picture.

However, even the picture that emerged in 1929 from Heisenberg and Pauli [37], which came to be called “second quantization,” was still somewhat debatable. For one thing, it was still plagued by spurious, unphysical infinities, which were mostly traceable to the structure of the Dirac Sea as the electromagnetic vacuum. Heisenberg proposed a process of “subtracting infinities,” which later came to be called “regularization and renormalization,” although even at that point in time Pauli was unconvinced that such a process could be justified in a mathematically-rigorous way <sup>(1)</sup>. One cannot help but notice that at that point in history, almost all of the subsequent published research that pertained to the theory of quantum electrodynamics was concerned with making the subtraction of infinities an acceptable process, and increasingly less discussion was devoted to the physical interpretation of the theory.

By the time that the various regularization and renormalizations schemes become established in the mainstream, the fact that they were basically “kludges” (i.e., error-correcting algorithms) was lost on all but the most distinguished members of the QED community. Everyone else seemed to be content to use them on the grounds that there was more work to be done in the name of particle physics than just the development of its fundamental theory. As the emphasis shifted to the development of more powerful particle accelerators that could reach higher and higher center-of-mass collision energies, the majority of particle physicists were happy to have even kludges that worked.

Something that gradually emerged from the theory of renormalization was the concept of “effective field theories,” which amounted to starting with the classical field theory as a zeroth-order approximation (more precisely, *zero-loop*, or *tree* approximation), and then looking at the inclusion of an increasing number of renormalization loops in the Feynman diagrams for the process as the successive quantum corrections to the classical field theory. This process is actually quite analogous to the way that one corrects the geometrical optics approximation to wave optics by adding diffraction terms that represent the processes that geometrical optics left out, and also relates to the mathematical theory of *asymptotic expansions*.

A particularly useful one-loop effective theory of quantum electrodynamics is the one that was devised by Heisenberg and his student Hans Euler in 1936 [38]. It basically corrected the Dirac wave function of a photon in the presence of an external electromagnetic field for the polarization of the photon into a bound state of an electron and a positron as a result of the external field being sufficiently strong and led to a corresponding effective theory for photon-photon scattering by Euler [39].

---

<sup>(1)</sup> Apparently, that rift led to an eventual estrangement of Heisenberg from Pauli. Heisenberg came to regard Pauli’s objections as “overblown” and did not even attend Pauli’s funeral.

**4. Continuum-mechanical models for wave mechanics.** – To return to the continuum-mechanical interpretation of quantum wave mechanics, we note that almost every successive advance in mainstream quantum theory, with its statistical interpretation of the wave equation, eventually led to corresponding advances in the continuum-mechanical interpretation. To begin with, after the seminal paper [8] by Madelung, which only addressed the hydrodynamical form of the stationary Schrödinger equation, that program was expanded upon and applied to the time-varying Schrödinger equation, as well.

One of the early proponents of the hydrodynamical interpretation amongst the founders of quantum theory was Louis de Broglie, who also added his own innovations in the form of the “theory of the double solution” and “pilot wave theory” (e.g., [40]). Several of his students, such as Jean-Pierre Vigier [41] and Francis Halbwachs [42], as well as the aforementioned Proca, continued his work along those lines, although eventually de Broglie grew temporarily disenchanted with the hydrodynamical picture and begrudgingly accepted the statistical interpretation. However, he later came back to the former interpretation in the context of what he called “the hidden thermodynamics of the isolated particle [43],” although nowadays one would say that the “hidden thermostat” that the isolated particle interacted with would be called the “quantum vacuum state,” and in fact de Broglie himself mentioned that terminology in a footnote.

Some of the most definitive advances to the hydrodynamical interpretation came from Takehiko Takabayasi at Nagoya University. He not only enlarged the scope of Madelung’s original work to the time-varying case, and looked at various applications of the theory to more traditional topics in quantum mechanics [44a], but also applied the program to the Klein-Gordon equation [44b]. In particular, he derived a “quantum stress tensor” that was related to the quantum potential function that had been defined by Madelung. However, when he addressed the cases of wave functions with spin, namely, the Pauli equation [44c] and the Dirac equation [44d], he switched from the transformation by polar coordinates on the field space (viz.,  $\mathbb{C}$ ) to the use of bilinear covariants as a way of deriving the continuum-mechanical quantities that were contained in the wave function.

When one defines the energy-momentum density 1-form to be  $\hbar d\theta$  that will automatically imply that the resulting motion is (dynamically) irrotational, since the exterior derivative of any exact 1-form will vanish. Although Takabayasi discussed the possible inclusion of vorticial flows in [44a,b], it was M. Schönberg who expanded upon that possibility in a series of papers [45]. That had the intriguing consequence that the terms in the quantum stress tensor could be related to more established expressions in the theory of turbulence, such as momentum transfer.

One of the better known proponents of the hydrodynamical interpretation was David Bohm, who saw it as a way of introducing “hidden variables” [46]. In a paper with Vigier [47], Bohm also discussed the possibility that the hydrodynamical model (or “Madelung fluid”) that one obtains from Schrödinger’s equations is essentially an approximation to a more complex motion, which takes the form of “sub-quantum” fluctuations about the quantum ground state. Again, this sounds like another form of the quantum vacuum, and also overlaps with the discussion in Schönberg [45]. In a paper with Schiller and Tiomno [48], Bohm discussed the conversion of the Pauli equation into a vorticial hydrodynamical form that differed from Takabayasi’s approach by the fact that

it represented a transformation of the field space, not the application of bilinear covariants.

One can also pose the inverse problem to the Madelung transformation, namely, how does one start with continuum-mechanical equations and transform them into a quantum wave equation. Lajos Jànossy addressed that problem in several papers [49].

Some of the more modern work regarding the hydrodynamical interpretation of quantum mechanics was done by Robert Carroll [50] and Iwo Bialynicki-Birula [51], and Takabayasi published a paper in 1983 [52] that discussed some of the advances that had been made since the early days.

This author has investigated some of the aspects of the continuum-mechanical models, in his own right. He first observed [53] that the quantum potential function could be related to the scalar curvature of the space-time metric that is obtained from the Minkowski-Lorentz metric by rescaling it using the density  $n$  or  $\rho$ . He also expanded upon the fact that the basic transformation in the Madelung theory is amounts to the introduction of polar coordinates in the field space in [54]. That led to further work along those lines that included obtaining a quantum strain tensor that would couple to the quantum stress tensor by way of a mechanical constitutive law in [55]. In the process, it was becoming clearer that the usage of the terms “curvature” and “torsion” in continuum mechanics (e.g., the bending and twisting of beams) was closer to their usage in the Frenet-Serret equations than in the Riemann-Cartan approach to differential geometry, in which the terms have more to do with the integrability of parallel translation. That led to the discussion in [56], in which a broader picture of the application of differential geometry to continuum mechanics emerged that sounded more like the geometry of “teleparallelism” in spirit. Further, it was observed that the structure of the conservation laws that are associated with the Dirac electron and the Weyssenhoff fluid (which is an approximation to the latter) was that of a “relativistic Cosserat medium,” and that was discussed in [57].

**5. Chapter summary.** – Although it would be helpful if the reader had some prior exposure to the basic geometrical and topological methods of mathematical physics (cf., e.g., Frenkel [58]), little use of topological methods will be made, since the basic objective is to summarize the various models that were constructed before, and except for some discussion of vortices, topology was not generally an issue. (The author is not trying to trivialize the role of topology in the subject, but only to defer that discussion to a later monograph. Indeed, the role of topological defects as the sources of fields is quite fundamental.)

However, there are times that the calculus of exterior differential forms is simply a more concise way of doing practical calculations, so the basic notions of that topic will be reviewed in Appendix A, while some basic notions regarding differentiable manifolds will be reviewed in Appendix B. Similarly, the discussion of conserved currents in the context of variational field theory is easier to navigate when one has some basic concepts from the theory of Lie groups, Lie algebras, and representations in hand, so the relevant notions are summarized in Appendix C. Finally, before embarking upon continuum mechanics, it helps to review some of the corresponding notions concerned with the

motion of points and rigid bodies, which is presented in Appendix C. Appendix E contains some multiplication tables for the Clifford algebra of Minkowski space.

As for the main body of text, we shall give the following summary:

*Chapter I.* – Since it is no longer the case that the theoretical physics mainstream regards continuum mechanics as something fundamental to everyone’s education, the first chapter attempts to briefly review the basic notions of that discipline. Once again, it is desirable that the reader is not encountering them for the very first time.

*Chapter II.* – The author is of the opinion that there was something hasty about the founding of quantum theory in its early days, and in particular that the very concept of wave mechanics needs to be developed in a more general context that treats a general wave as an object that has a kinematical state and a dynamical state that are coupled by a constitutive law, along with an integrability equation for the kinematical state and a balance principle for the dynamical one. Hence, this chapter is a first attempt at laying such foundations.

*Chapter III.* – Much of the discussion of the individual models for quantum waves is based in the methods of variational field theory, which is summarized to the extent that it will be used in the rest of the book.

*Chapter IV.* – In this chapter, we first address the Madelung-Takabayasi picture in the context of non-relativistic, spinless waves, which are then described by the Schrödinger equation.

*Chapter V.* – In the process of developing the continuum-mechanical model of Chapter IV, some “quantum terms” appeared, such as the quantum potential. In this chapter, the author’s thoughts on the subject of their intrinsic nature are presented.

*Chapter VI.* – The extension of the Madelung-Takabayasi methodology to relativistic, spinless waves is addressed in this chapter; i.e., the Madelung-Takabayasi transformation of the Klein-Gordon equation, and the corresponding alterations to the discussion of the quantum terms are made.

*Chapter VII.* – In this chapter, the introduction of Pauli spinors for the description of non-relativistic, spinning waves is investigated. The methods of Bohm, Tiomno, and Schiller are discussed in more detail than those of Takabayasi, which are mentioned briefly, and the author’s own view of the process is presented.

*Chapter VIII.* – Before discussing the quantum-mechanical formulation of relativistic, spinning matter, we first discuss some of the key ideas of relativistic rotational mechanics in its more “classical” context.

*Chapter IX.* – The Dirac equation for relativistic, spinning waves is reviewed, along with some of the alternative ways of formulating it that had been proposed.. The author

also discusses the relativistic Pauli equation, and makes some suggestions on how to remain consistent with the introduction of “generalized spherical coordinates” into the field spaces as the basis for the transformation of the wave equation.

*Chapter X.* – In this chapter, the continuum-mechanical formulation of the Dirac equation is presented in the manner of Takabayasi [44d]. The Weyssenhoff fluid is discussed as a simplification of the Dirac electron, and the facts that both the Dirac electron and the Weyssenhoff fluid represent relativistic Cosserat media are also presented and examined.

*Epilogue.* – Finally, a few suggestions on directions for further growth in the field of continuum-mechanical models for quantum mechanics are proposed.

### References

1. M. Planck:
  - a. “Über eine Verbesserung der Wienschen Spektralgleichung,” *Verhandlungen der Deutschen Physikalischen Gesellschaft* **2** (1900), 202–204; Translated in D. ter Haar, “On an Improvement of Wien's Equation for the Spectrum” in *The Old Quantum Theory*, Pergamon Press, 1967; pp. 79–81.
  - b. “Zur Theorie des Gesetzes der Energieverteilung im Normalspektrum,” *Verhandlungen der Deutschen Physikalischen Gesellschaft* **2** (1900), 237. Translated in D. ter Haar, *The Old Quantum Theory*, Pergamon Press, 1967, pp. 82.
2. C. Huygens, *Traité de la lumière*, Chez Pierre van der A a, Leiden, 1690.
3. I. Newton, *Opticks*, 1704.
4. A. Einstein, “Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt,” *Ann. Phys. (Berlin)* (6) **17** (1905), 132–148. Translated in D. ter Haar, *The Old Quantum Theory*, Pergamon Press, 1967, pp. 91.
5. A. Compton, “A Quantum Theory of the Scattering of X-Rays by Light Elements,” *Phys. Rev.* **21** (5) (1923), 483–502.
6. L. de Broglie, “Recherches sur la théorie des quanta,” *Ann. de Phys. (10)* **3** (1925), 22–128; English translation by A. F. Kracklauer, 2004, published online.
7. E. Schrödinger, “Quantisierung als Eigenwertproblem,” *Ann. Phys. (Berlin)* (4) **384** (1926), 273–376. English translation in E. Schrödinger, *Collected Papers on Wave Mechanics*, AMS Chelsea Publishing, Providence, Rhode Island, 1982.
- 8\*. E. Madelung, “Quantentheorie in hydrodynamischer Form,” *Z. Phys.* (3–4) **40** (1927), 322–326.
9. O. Klein, “Quantentheorie und fünf-dimensionale Relativitätstheorie,” *Z. Phys.* **37**, 895 (1926).
- 10\*. W. Gordon, “Der Comptoneffekt nach der Schrödingerschen Theorie,” *Zeit. f. Phys.* **40** (1926), 117–133.
11. W. Pauli and V. Weisskopf, “Über die Quantisierung der skalaren relativistischen Wellengleichung,” *Helv. Phys. Acta* **7** (1934), 709–31; also published in translation in [33].



- 12\* . M. Abraham, “Prinzipien der Dynamik des Elektrons,” *Ann. Phys. (Berlin)* **315** (1903), 105-179.
13. H. A. Lorentz, *Theory of Electrons*, Dover, NY, 1952.
14. H. Poincaré, “Sur la dynamique de l’électron,” *C.R. Acad. Sci.* **140** (1905), 1504; “Sur la dynamique de l’électron,” *Rend. Circ. Mat. Palermo* **21** 18 (1906), 129; Translation, with commentary, by A. A. Logunov in “On the articles by Henri Poincaré ‘On the dynamics of the electron’.” Publishing Department of the Joint Institute for Nuclear Research. Dubna. 1995.
15. A. Einstein, “Zur Elektrodynamik bewegter Körper,” *Ann. Phys. (Berlin)* (10) **322** (1905), 891–921. Translated in *The Principle of Relativity*, Dover, Mineola, NY, 1952.
16. F. Rohrlich, *Classical Charged Particles*, Addison-Wesley, Reading, MA, 1965.
17. M. H. MacGregor, *The Enigmatic Electron*, 2<sup>nd</sup> ed., El Mac Books, Santa Cruz, CA, 2013.
18. A. D. Yaghjian, *Relativistic Dynamics of a Charged Sphere*, Springer-Verlag, Berlin, 1992.
19. P. Pearle, “The classical electron,” in *Electromagnetism – Paths to Research*, D. Teplitz (ed.), Plenum, NY, 1982.
20. W. Gerlach and O. Stern, “Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld,” *Zeit. Phys.* **9** (1922), 349–352.
21. G. Uhlenbeck and S. Goudsmit, “Spinning electrons and the structure of spectra,” *Nature* **117** (1926), 264-265.
22. A. Compton, “The magnetic electron,” *J. Franklin Inst.* **192** (1921), 145-155.
- 23\* . F. Raseti and E. Fermi, “Sopra l’elettrone rotante,” (8) **3** (1926), 226-235.
- 24\* . W. Pauli, “Zur Quantenmechanik des magnetischen Elektrons,” *Zeit. f. Phys.* **43** (1927), 601-623.
- 25\* . W. Heisenberg and P. Jordan, “Zur Quantenmechanik des magnetischen Elektrons,” *Zeit. f. Phys.* **43** (1927), 601-623.
26. D. H. Delphenich, “The relativistic Pauli equation,” arXiv:1207.5752.
27. L. H. Thomas, “The kinematics of an electron with an axis,” *Phil. Mag.* (7) **13** (1927), 1-22.
- 28\* . J. Frenkel, “Die Elektrodynamik des rotierenden Elektrons,” *Zeit. Phys.* **37** (1926), 243-262.
29. P. A. M. Dirac, “The quantum theory of the electron,” *Proc. Roy. Soc. (London) A* **117** (1928), 610-624; Part II: *ibid.* **118** (1928), 351-361.
30. C. G. Darwin, “The wave equations of the electron,” *Proc. Roy. Soc. London A* **118** (1928), 654-680.
- 31\* . A. Proca, “Sur l’équation de Dirac,” *J. Phys. et le Rad.* **1** (1930), 235-248.
32. A. S. Eddington, *Fundamental Theory*, Cambridge Univ. Press, Cambridge, 1953.
- 33\* . C. Lanczos:
  - a. “Die tensoranalytischen Beziehungen der Diracschen Gleichung,” *Zeit. f. Phys.* **57** (1927), 447-473.
  - b. “Zur kovarianten Formulierung der Diracschen Gleichung,” *Zeit. f. Physik* **57** (1929), 474-483.
  - c. “Die Erhaltungssätze in der feldmässigen Darstellung der Diracschen Theorie,” *Zeit. f. Phys.* **57** (1929), 484-493.

34. A. I. Miller, *Early Quantum Electrodynamics: a source book*, Cambridge University Press, Cambridge, 1994.
35. H. Weyl, *Gruppentheorie und der Quantenmechanik*, 1928. English translation: *Theory of Groups and Quantum Mechanics*, Dover, Mineola, NY, 1950.
36. Carl D. Anderson, "The Positive Electron," *Phys. Rev.* (6) **43** (1933), 491–494.
- 37\*. W. Heisenberg and W. Pauli:
  - a. "Zur Quantendynamik der Wellenfelder," *Zeit. Phys.* **56** (1929), 1-61.
  - b. "Zur Quantentheorie der Wellenfelder, II" *Zeit. Phys.* **59** (1929), 168-190.
- 38\*. W. Heisenberg and H. Euler, "Folgerungen aus der Diracschen Theorie des Positrons," *Zeit. f. Phys.*, **98** (1936), 714-732.
- 39\*. H. Euler and B. Köckel, "Über die Streuung von Licht an Licht nach der Diracschen Theorie," *Naturwiss.* **23** (1935), 246-247; H. Euler, "Über die Streuung von Licht an Licht nach der Diracschen Theorie," *Ann. Phys. (Leipzig)* **26** (1936), 398-448.
40. L. de Broglie, *Nonlinear Wave Mechanics*, Elsevier, Amsterdam, 1960.
- 41\*. J. P. Vigièr, *Structure des micro-objets dans l'interprétation causale de la théorie des quanta*, Gauthier-Villars, Paris, 1956.
- 42\*. F. Halbwachs, *La théorie relativiste des fluides a spin*, Gauthier-Villars, Paris, 1960.
- 43\*. L. de Broglie, *La thermodynamique de la particule isolée*, Gauthier-Villars, Paris, 1964.
44. T. Takabayasi:
  - a. "On the Formulation of Quantum Mechanics Associated with Classical Pictures," *Prog. Theor. Phys.*, **8** (1952), 143-182.
  - b. "Remarks on the Formulation of Quantum Mechanics Associated with Classical Pictures," *Prog. Theor. Phys.*, **9** (1953), 187-222.
  - c. "The vector representation of spinning particle in the quantum theory, I," *Prog. Theor. Phys.* **14** (1955), 283-302.
  - d. "Relativistic Hydrodynamics of Dirac Matter (Part 1. General Theory)," *Prog. Theor. Phys. suppl.* **4** (1957), 1-80.
45. M. Schönberg:
  - a. "A non-linear Generalization of the Schrödinger and Dirac Equations," *Nuov. Cim.* **11** (1954), 674-682.
  - b. "On the Hydrodynamical Model of the Quantum Mechanics," *Nuov. Cim.* **12** (1954), 103-133.
  - c. "Simple Solution of the Generalized Schrödinger Equations," *Nuov. Cim.* **12** (1954), 300-303.
  - d. "Vortex Motions of the Madelung Fluid," *Nuov. Cim.* (1955), 543-580.
46. D. Bohm, "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables I" *Phys. Rev.* **85** (1952), 166-179; "...II," *ibid.*, 180-193.
47. D. Bohm and J.-P. Vigièr, "Model of the Causal Interpretation of Quantum Theory in Terms of a Fluid with Irregular Fluctuations," **96** (1954), 208-216.
48. D. Bohm, R. Schiller, J. Tiomno, "A causal interpretation of the Pauli equation (A)," *Supp. Nuovo Cimento* **1** (1955), 48-66, (B) 66-91.
49. L. Janossy:
  - a. "Zum hydrodynamischen Modell der Quantenmechanik," *Zeit. Phys.* **169** (1962), 79-89.

- b. “The physical interpretation of wave mechanics, I,” *Found. Phys.* **3** (1973), 185-202.
  - 50. R. Carroll, “On the quantum potential,” arXiv:quant-ph/0403156v3.
  - 51. I. Bialnycki-Birula, Z. Bialnycki-Birula, and C. Sliwa, “Motion of vortex lines in quantum mechanics,” *Phys. Rev. A* **61** 032110 (2000).
  - 52. T. Takabayasi, “Hydrodynamical formulation of quantum mechanics and Aharonov-Bohm effect,” *Prog. Theor. Phys.* **69** (1983), 1323-1344.
  - 53. D. H. Delphenich, “The Geometric Origin of the Madelung Potential,” arXiv:gr-qc/0211065.
  - 54. D. H. Delphenich, “Generalized Madelung transformations for quantum wave equations I: generalized spherical coordinates for field spaces,” arXiv:0802.0305.
  - 55. D. H. Delphenich, “A strain tensor that couples to the Madelung stress tensor,” arXiv:1303.3582.
  - 56. D. H. Delphenich, “The use of the teleparallelism connection in continuum mechanics,” *Mathematics and Mechanics of Solids* **21** (2016), 1260-1275; first published online Dec. 26, 2014.
  - 57. D. H. Delphenich, “Mechanics of Cosserat media: II. relativistic theory,” arXiv:1510.01243.
  - 58. T. Frenkel, *The Geometry of Physics: an introduction*, Cambridge University Press, Cambridge, 1997.
-

## CHAPTER I

# BASIC CONTINUUM MECHANICS

The better part of the problem of formulating the continuum-mechanical models for quantum wave equations is trying to gain a better intuition for the very nature of the state of matter that one is dealing with in the first place. That problem amounts to gaining a deeper understanding of the mechanical constitutive law that associates a dynamical state of energy-momentum and stress with a kinematical state of deformation. As we will see, the crucial notion that one must introduce into the kinematical state of the extended matter distribution that one is addressing is the idea that in addition to the “metric strain” that is traditionally considered in the name of continuum mechanics, one must also consider “frame strain,” which can be present even when there is no change in the distances between points of the object.

Before going on to that more unconventional topic, we shall first summarize the basic ideas that one deals with in conventional continuum mechanics. We feel that such a summary is unavoidable nowadays, because although there was a time that continuum mechanics was the foundation for all theoretical physics, between the ultimate failure of the mechanical ether theory of electromagnetic waves and the discreteness of matter at the atomic level, the physics mainstream seemed to abandon their previous belief that continuum mechanics was truly fundamental to the study of matter. It is our steadfast belief that this judgment was premature.

**§ 1. Extended objects and their deformations.** – When one is dealing with point-like matter moving in space, the kinematical state of a point can be characterized by its position, velocity, acceleration, and any higher-order time derivatives that might be necessary <sup>(1)</sup>. Indeed, since the basic (ordinary) differential equations of motion ultimately prove to be second-order, it is usually sufficient to consider only the first-order derivatives; i.e., the velocity. Furthermore, one of the lessons of rotational mechanics is the idea that having the velocity equal to precisely the time derivative of position is not the only possibility, since the angular velocity of the reference frame might also contribute to the total velocity, in addition to the time derivatives of position. This situation can be regarded as an elementary example of the difference between integrable and non-integrable kinematical states, which we will discuss in due course.

Hence, we shall use a modern mathematical formalism that is best adapted to the modeling of kinematical states in a manner that makes the issue of integrability obvious, namely, the formalism of “jet manifolds.” This formalism has the advantage that it is also closely-adapted to some problems of differential equations and the calculus of variations. However, we shall stay closer to the local coordinate expressions than most purely mathematical treatments [1], simply because that is where one finds most of the

---

<sup>(1)</sup> Some relevant concepts from the motion of points and rigid bodies are summarized in Appendix D, for the sake of completeness.

standard literature of continuum mechanics [2-7], although some of what we shall discuss was established some time ago by Gallisot [8].

*a. Jet manifolds.* – One of the big reasons for introducing the methods of jets is that they allow one to replace the consideration of infinite-dimensional state spaces with the consideration of finite-dimensional ones. Basically, one starts with something that could rapidly lead into functional-analytic complications, namely, the idea that a *configuration* of an extended object  $\mathcal{O}_p$  (which we assume to be an open subset of a parameter space  $\mathbb{R}^p$  for some dimension  $p$ ) in a space, which we assume to be  $\mathbb{R}^n$  for simplicity, is a  $C^k$  (viz.,  $k$  times continuously differentiable) embedding  $x : \mathcal{O}_p \rightarrow \mathbb{R}^n$ ,  $u^a \mapsto x^i(u^a)$ . That is, the map is one-to-one and a homeomorphism onto its image  $x(\mathcal{O}_p)$ , as well. Thus, the image will not intersect itself, and it will have the same dimension as the prototype object  $\mathcal{O}_p$  in  $\mathbb{R}^p$ . Some examples of this situation are when  $\mathcal{O}_p$  is a solid, open ball, disc, or cylinder, and the configurations  $x$  are smooth deformations of the more symmetric, geometric prototypes.

The example of a point in space is an elementary example of a configuration in which  $k = 0$  and the embedding of that point in space is its inclusion. If one wishes to describe the time evolution of a point then one must go to the next dimension ( $k = 1$ ) and let  $\mathcal{O}_1$  be an interval  $(t_0, t_1)$  on the time line  $\mathbb{R}$ . The position will then take the form  $x^i(t)$  for  $t \in (t_0, t_1)$  and its successive derivatives will become the velocity, acceleration, etc. <sup>(1)</sup>, with respect to an inertial frame. However, in statics, it is also important to consider the case in which  $\mathbb{R}$  does not represent time, but simply a curve parameter  $s$ . The first derivatives of  $x^i(s)$  would still define (up to a non-zero scalar constant) the tangent line to the curve at each point, even though one could not really call them “velocities” when nothing is moving. Similarly, the second derivatives would describe the curvature (in the Frenet-Serret sense of the word) of the curve more than its acceleration.

When one goes on to  $k = 2$ , one can again consider  $\mathbb{R}^2$  to be the parameter space for the motion of a curve segment (e.g., a vibrating string) or the embedding of a surface, and similarly,  $k = 3$  can describe the motion of a surface (e.g., a vibrating membrane) or the embedding of a solid object. Ultimately, we shall have no use for any  $k$  that is greater than four, and in that case,  $\mathcal{O}_p$  will usually take the form of a “cylindrical” object; i.e.,  $\mathcal{O}_4 = (t_0, t_1) \times \mathcal{O}_3$ , where  $\mathcal{O}_3$  is a three-dimensional object. Indeed, time evolution, for us, will generally involve objects of the form  $(t_0, t_1) \times \mathcal{O}_{k-1}$ , where  $k = 0, 1, 2, 3$ . The limitation of cylindrical objects as prototypes for moving extended matter is that one must rule out “topology-changing processes,” which are actually quite common in nature (the bifurcation of branches on plants, the formation of soap bubbles, smoke rings, and

---

<sup>(1)</sup> For the record, the next two time derivatives are referred to as “jerk” and “yank.”

the nucleation of bubbles in boiling liquids), although they properly deserve special treatment, due to the topological overhead that must be introduced <sup>(1)</sup>.

The *kinematical state* of the configuration  $x : \mathcal{O}_p \rightarrow \mathbb{R}^n$  is then defined by the set of real numbers  $(u^a, x^i(u), x_{,a}^i(u), \dots, x_{,a_1, \dots, a_k}^i(u))$ , and its *order* will then be  $k$ . We shall mostly be concerned with first-order states of the form  $(u^a, x^i(u), x_{,a}^i(u))$ .

One can back up a step from this definition and define the  $k$ -jet of a  $C^k$  map  $x : \mathcal{O}_p \rightarrow \mathbb{R}^n$  at a point  $u \in \mathcal{O}_p$ , to be the set of real numbers  $j_u^k x = (u^a, x^i, x_{,a}^i, \dots, x_{,a_1, \dots, a_k}^i)$ , where the arrays  $x^i, x_{,a}^i, \dots, x_{,a_1, \dots, a_k}^i$  are equal to the values of the successive derivatives of  $x$  at  $u$ , although they themselves are not functions of  $u$ . One can also define the  $k$ -jet  $j_u^k x$  to be the equivalence class of all  $C^k$  maps  $x : \mathcal{O}_p \rightarrow \mathbb{R}^n$  that are defined in some neighborhood of a point  $u \in \mathcal{O}_p$  and have the same values for their first  $k$  derivatives at that point. Notice that since a jet is a purely local object, it will be irrelevant whether the map in question is an embedding or an immersion, since that distinction is a global consideration that will not typically affect the local derivatives.

One can see that the definition of a tangent vector at a point  $x$  in a differentiable manifold  $M$  as an equivalence class of differentiable curves through  $x$  that have the same first derivative at  $x$  makes a tangent vector essentially a 1-jet of a  $C^1$  map of  $\mathbb{R}$  to  $M$  at some chosen point (e.g., 0) on  $\mathbb{R}$ .

The set of all  $k$ -jets of  $C^k$  maps from  $\mathcal{O}_p$  to  $\mathbb{R}^n$  is then a differentiable manifold  $J^k(\mathcal{O}_p, \mathbb{R}^n)$  whose local coordinate charts look like  $(u^a, x^i, x_{,a}^i, \dots, x_{,a_1, \dots, a_k}^i)$ , so it will have dimension  $p + n + pn + \dots, p^k n$ . In most cases, we will be dealing with  $J^1(\mathcal{O}_p, \mathbb{R}^n)$ , which has coordinate charts of the form  $(u^a, x^i, x_{,a}^i)$ . The coordinate transformations for the higher-order derivatives get rapidly complicated, so we specify them only for 1-jets: Let  $(u^a(u), x^i(u), x_{,a}^i(u))$  and  $(\bar{u}^a(u), \bar{x}^i(u), \bar{x}_{,a}^i(u))$  be two sets of coordinates for  $j_u^1 x$ ; one must be careful to note that:

$$\bar{x}_{,a}^i = \frac{\partial \bar{x}^i}{\partial \bar{u}^a}, \quad (1.1)$$

by definition.

Hence, we already have coordinate transformations:

---

<sup>(1)</sup> Many mathematicians will insist that the appropriate topological techniques for the discussion of topology-changing processes come from the study of “cobordism,” which is a generalized homology that is specialized to the needs of differential topology. This author has long been of the opinion that nothing in everyday nature actually goes beyond the scope of the more computable homologies (e.g., simplicial, singular) so radically that the generalization to cobordism becomes unavoidable, especially if the canonical example of a cobordism is the “trouser manifold,” which can easily be triangulated by more elementary building blocks. The introduction of gratuitous generality for its own sake runs counter to the spirit of Occam’s razor, which is fundamental to the scientific method.

$$\bar{u}^a = \bar{u}^a(u^b), \quad \bar{x}^i = \bar{x}^i(x^j). \quad (1.2)$$

Therefore, we must also account for the transformation of  $x_a^i$  to  $\bar{x}_a^i$ . Since we are dealing with only first derivatives, it will be a linear transformation by way of the differentials of the last two transformations:

$$\bar{x}_a^i = \frac{\partial u^b}{\partial \bar{u}^a} \frac{\partial \bar{x}^i}{\partial x^j} x_b^j. \quad (1.3)$$

If we wished to go on to 2-jets then we would have to differentiate the product on the right-hand side with respect to  $\bar{u}^b$ , which would produce an inhomogeneous transformation law, namely:

$$\bar{x}_{ab}^i = \frac{\partial u^c}{\partial \bar{u}^a} \frac{\partial u^d}{\partial \bar{u}^b} \frac{\partial \bar{x}^i}{\partial x^j} \left( x_{cd}^j + \frac{\partial x^j}{\partial \bar{x}^m} \frac{\partial \bar{x}^m}{\partial x^k \partial x^l} x_c^k x_d^l \right). \quad (1.4)$$

This differs from a tensor transformation law by the appearance of the quadratic term in the first derivatives. One should note, however, that in the case of one parameter, for which the partial derivatives will become ordinary derivatives, if one treats the three-index expression in the second term on the right-hand side as if it represented the components:

$$\Gamma_{kl}^j = \frac{\partial x^j}{\partial \bar{x}^m} \frac{\partial \bar{x}^m}{\partial x^k \partial x^l} \quad (1.5)$$

of a linear connection then the vanishing of the expression in parentheses would amount to the geodesic equation for the symmetric part of that connection, namely:

$$0 = \frac{d^2 x^j}{dt^2} + \Gamma_{(kl)}^j \dot{x}^k \dot{x}^l, \quad \Gamma_{kl}^j \equiv \frac{1}{2} (\Gamma_{kl}^j + \Gamma_{lk}^j). \quad (1.6)$$

Hence, one can already see how the study of jets overlaps with the study of connections.

The manifold  $J^1(\mathcal{O}_p, \mathbb{R}^n)$  admits three canonical projections:

1. The *source* projection:  $\alpha: J^1(\mathcal{O}_p, \mathbb{R}^n) \rightarrow \mathcal{O}_p, \quad j_u^1 x \mapsto u.$
2. The *target* projection:  $\beta: J^1(\mathcal{O}_p, \mathbb{R}^n) \rightarrow \mathbb{R}^n, \quad j_u^1 x \mapsto x.$
3. The *contact* projection:  $p_0^1: J^1(\mathcal{O}_p, \mathbb{R}^n) \rightarrow \mathcal{O}_p \times \mathbb{R}^n, \quad j_u^1 x \mapsto (u, x).$

Sometimes, the manifold  $\mathcal{O}_p \times \mathbb{R}^n$  is denoted by  $J^0(\mathcal{O}_p, \mathbb{R}^n)$ , for completeness.

Consistent with the last projection, one can think of the matrix  $x_a^i$  as defining a *contact element* at  $(u, x)$ , which is a  $p$ -dimensional linear subspace of the tangent space  $T_x\mathbb{R}^n$ , as long as the maps  $x$  in question are always immersions (so  $x_a^i$  always has rank  $p$ ). Moreover, one can then regard the  $p$  columns  $\{x_1^i, \dots, x_p^i\}$  of the rank- $p$  matrix  $x_a^i$  as also defining a  $p$ -frame in  $T_x\mathbb{R}^n$ , which will then span that linear subspace.

It is important to see that the differential map to  $x : \mathcal{O}_p \rightarrow \mathbb{R}^n$  at a point  $u \in \mathcal{O}_p$  is a linear map  $dx|_u : T_u\mathcal{O}_p \rightarrow T_x\mathbb{R}^n$ , so it is really a “two-point” function, and if one wishes to express it in terms of the natural coframe  $du^a$  for  $T_u\mathcal{O}_p$  and the natural frame  $\partial / \partial x^i$  in  $T_x\mathbb{R}^n$  then one can write:

$$dx|_u = x_a^i(u) du^a \otimes \partial / \partial x^i, \quad (1.7)$$

which is well-defined as a tensor field only on  $\mathcal{O}_p \times \mathbb{R}^n$ .

The kinematical state of the extended object that is described by  $x$  then becomes a *section*  $s : \mathcal{O}_p \rightarrow J^k(\mathcal{O}_p, \mathbb{R}^n)$ ,  $u \mapsto (u^a, x^i(u), x_{,a}^i(u), \dots, x_{,a_1, \dots, a_k}^i(u))$ , so each  $s(u)$  belongs to the set  $J_u^k(\mathcal{O}_p, \mathbb{R}^n)$  of all  $k$ -jets that project to  $u$  under the source projection (i.e., the *fiber* of that projection over  $u$ ). Moreover, as defined, it also represents a section of a very special type, namely the  *$k$ -jet prolongation* of a  $C^k$  map  $x$ , which gets denoted by:

$$j^k x(u) = (u^a, x^i(u), x_{,a}^i(u), \dots, x_{,a_1, \dots, a_k}^i(u)). \quad (1.8)$$

In the more general case, the coordinate functions  $x_{,a_1 \dots a_l}^i(u)$  of the value of a section  $s$  at  $u$  do not have to take the form of mixed partial derivatives of  $x^i$  with respect to  $u^a$ . When that is true for a section, i.e.:

$$s = j^k x, \quad (1.9)$$

for some  $C^k$  function  $x : \mathcal{O}_p \rightarrow \mathbb{R}^n$ , one calls the section (i.e., kinematical state) *integrable*. Not all sections of the source projection will be integrable, since for one thing, the lower indices of  $x_{,a_1 \dots a_l}^i(u)$  do not have to always be completely symmetric, as they must be for mixed partial derivatives (with appropriate regularity assumptions about  $x$ ). Even in the case of  $x_a^i(u)$ , if one forms the set of  $n$  1-forms on  $\mathcal{O}_p$ :

$$\xi^i = x_a^i(u) du^a \quad (1.10)$$

then one must have the vanishing of:

$$d \wedge \xi^i = \frac{1}{2} (x_{,a,b}^i - x_{,b,a}^i) du^a \wedge du^b \quad (1.11)$$



for each  $i$  as a necessary condition for the integrability of  $x_a^i(u)$ ; i.e., for one to have:

$$\xi^i = dx^i \quad (1.12)$$

for each  $i$ . (As long as  $\mathcal{O}_p$  is simply-connected, that condition will also be sufficient.)

The condition (1.11) then says that one must have:

$$x_{a,b}^i = x_{b,a}^i. \quad (1.13)$$

More generally, one defines the *Spencer operator* [1] on sections of the source projection, which one denotes by  $D : J^k(\mathcal{O}_p, \mathbb{R}^n) \rightarrow J^{k-1}(\mathcal{O}_p, \mathbb{R}^n)$ , although that is really an abbreviation for saying that it takes sections of the source projection of one jet manifold to sections of the source projection of the other one. If  $s(u) = \mapsto (u^a, x^i(u), x_a^i(u), \dots, x_{a_1 \dots a_k}^i(u))$  is a section of the former projection then:

$$Ds(u) = (u^a, Dx^i(u), Dx_a^i(u), \dots, Dx_{a_1 \dots a_{k-1}}^i(u)), \quad (1.14)$$

in which:

$$Dx_{a_1 \dots a_l}^i(u) = x_{a_1 \dots a_l, a_{l+1}}^i(u) - x_{a_1 \dots a_{l+1}}^i(u). \quad (1.15)$$

Hence,  $s$  is integrable iff:

$$Ds = 0. \quad (1.16)$$

The case of non-integrable kinematical states becomes fundamental to the study of rotational mechanics when one looks at components with respect to non-inertial (i.e., anholonomic) frame fields, such as ones that are fixed in rotating bodies. For instance, the velocity components will then take the form:

$$v^i = \frac{dx^i}{dt} + \omega_j^i x^j, \quad (1.17)$$

which is not generally of the form  $d\bar{x}^i / dt$  for some set of functions  $\bar{x}^i(t)$ .

*b. Finite deformations.* – One advantage of the formalism that we have defined above is that one can talk about a deformation of a  $k$ -dimensional object  $x : \mathcal{O}_k \rightarrow \mathbb{R}^n$  as a (cylindrical)  $k+1$ -dimensional object  $x : \mathcal{O}_{k+1} \rightarrow \mathbb{R}^n$ , with  $\mathcal{O}_{k+1} = (s_0, s_1) \times \mathcal{O}_k$ . Thus, we are really referring to a differentiable one-parameter family of kinematical states that generalizes the motion of a point along a curve when  $k = 0$ . One can also think of  $x$  as a “differentiable homotopy” from the *initial state*  $x_0 : \mathcal{O}_k \rightarrow \mathbb{R}^n$ , with  $x_0(u) = x(0, u)$  to the *final state*  $x_1 : \mathcal{O}_k \rightarrow \mathbb{R}^n$ , with  $x_1(u) = x(1, u)$ , although we shall not go into that further at

this point. (We do not refer to  $x_0$  as the *natural state* since that term makes restricting assumptions on the internal stress distribution, which we have not defined yet.)

For many purposes, it is also useful to have some way of describing the transformation that takes  $x_0(u)$  to  $x_1(u)$ . Since we are assuming that both maps  $x_0$  and  $x_1$  are embeddings,  $x_0$  will be invertible on the points of its image  $x_0(\mathcal{O}_k)$ . If one then composes  $x_0^{-1} : x_0(\mathcal{O}_k) \rightarrow \mathcal{O}_k$  with  $x_1 : \mathcal{O}_k \rightarrow \mathbb{R}^n$  then the composed map  $y = x_1 \cdot x_0^{-1} : x_0(\mathcal{O}_k) \rightarrow x_1(\mathcal{O}_k)$  will be a diffeomorphism of the two images. If the points of  $x_0(\mathcal{O}_k)$  are described by coordinates  $x^i$  (suitably-restricted), and the points of  $x_1(\mathcal{O}_k)$  are described by coordinates  $y^i$  (suitably-restricted) then that transformation can be described locally by a set of equations of the form:

$$y^i = y^i(x^j). \quad (1.18)$$

This is the usual starting point for conventional continuum mechanics. One then describes the finite deformation that is defined by (1.18) by its *displacement vector field*, whose components with respect to the natural frame field for the canonical coordinate system on  $\mathbb{R}^n$  will be:

$$u^i(x) = y^i(x) - x^i. \quad (1.19)$$

Although one can define a vector field  $\mathbf{u}(x)$  on  $x_0(\mathcal{O}_k)$  by way of:

$$\mathbf{u}(x) = u^i(x) \frac{\partial}{\partial x^i} = \mathbf{y}(x) - \mathbf{x}(x), \quad (1.20)$$

with analogous definitions for  $\mathbf{x}(x)$  and  $\mathbf{y}(x)$ , nonetheless,  $\mathbf{u}(x)$  is not a true vector field, any more than  $\mathbf{x}(x)$  and  $\mathbf{y}(x)$ , which take the form of “position vector fields” or “radius vector fields,” since their components are actually *coordinate* functions. Therefore, when one changes to a different coordinate system, unless that transformation is linear, the components of  $\mathbf{u}(x)$  will not transform linearly by the differential of the coordinate transformation. Nonetheless, we introduce the concept of displacement vector field for the sake of completeness, since its use is widespread in continuum mechanics; if one wishes to avoid it, one must deal with the diffeomorphism  $y$  directly.

Note that although every diffeomorphism onto  $y: x_0(\mathcal{O}_k) \rightarrow \mathbb{R}^n$  defines a displacement vector field on  $x_0(\mathcal{O}_k)$ , the converse is not true. For instance, the negative of the radius vector field (viz.,  $-\mathbf{r}(x) = -x^i \partial_i$ ) would take every point  $x$  to the origin, which does not describe a diffeomorphism.

*c. Infinitesimal deformations.* – The first thing that one defines after the displacement vector field is the *displacement gradient*, which is the differential:

$$d\mathbf{u} = u^i_{,j} dx^j \otimes \frac{\partial}{\partial y^i}. \quad (1.21)$$

Note that from (1.20):

$$d\mathbf{u} = d\mathbf{y} - I \quad (u^i_{,j} = y^i_{,j} - \delta^i_j), \quad (1.22)$$

so the essential information in  $d\mathbf{u}$  is already contained in  $d\mathbf{y}$ . Once again, the displacement gradient is well-defined only as a tensor field on  $\mathbb{R}^n \times \mathbb{R}^n$ .

Notice that  $d\mathbf{u}|_x = 0$  for all  $x \in x_0(\mathcal{O}_p)$  iff the diffeomorphism is a constant (i.e., rigid) translation.

Up to this point, we could just as well define the initial state to be the subset  $x_0(\mathcal{O}_p)$  of  $\mathbb{R}^n$ , and the final state to be the subset  $x_1(\mathcal{O}_p) = y(x_0(\mathcal{O}_p))$ , so the deformation could be defined by the diffeomorphism onto  $y : x_0(\mathcal{O}_p) \rightarrow \mathbb{R}^n$ . The space of  $k^{\text{th}}$ -order kinematical states would then be the jet manifold  $J^k(x_0(\mathcal{O}_p), \mathbb{R}^n)$ . However, we have found it more convenient to use the definitions that we made above, so we simply show how the two relate, as we have just done.

If one assumes that  $\mathbb{R}^n$  has a scalar product defined on it (whether Euclidian or Lorentzian) then one can think of lowering the upper index of  $u^i_{,j}$  to produce the components of the doubly-covariant second-rank (pseudo) tensor on  $x_0(\mathcal{O}_p) \times \mathbb{R}^n$ :

$$du = u_{i,j} dx^j \otimes dy^i. \quad (1.23)$$

If one uses the same coordinate system on  $\mathbb{R}^n$  for both  $x$  and  $y$  then one can polarize this under the permutation of indices:

$$du = e + \theta, \quad (1.24)$$

in which:

$$e = \frac{1}{2} e_{ij} dx^i dx^j, \quad e_{ij} \equiv u_{i,j} + u_{j,i}, \quad (1.25)$$

$$\theta = \frac{1}{2} \theta_{ij} dx^i \wedge dx^j, \quad \theta_{ij} \equiv u_{i,j} - u_{j,i}. \quad (1.26)$$

The former symmetric tensor  $e$  is referred to as the *infinitesimal strain* that is associated with the deformation, while the latter 2-form  $\theta$  is the *infinitesimal rotation*, which can also be expressed as the exterior derivative of the displacement 1-form:

$$\theta = d \wedge u. \quad (1.27)$$

One can further decompose  $e$  into a traceless part and a trace part, where the trace refers to the matrix  $e^i_j$ :

$$e^i_j = e^{\circ i}_j + \frac{1}{n} e \delta^i_j, \quad (1.28)$$

by setting:

$${}^0 e^i_j = e^i_j - \frac{1}{n} e \delta^i_j, \quad (1.29)$$

but that definition of  ${}^0 e^i_j$  is not unique, since one can add any other traceless matrix to  ${}^0 e^i_j$  without changing its trace.

The trace:

$$e = e^k_k = \frac{\partial u^k}{\partial x^k} = \text{div } \mathbf{u} \quad (1.30)$$

is called the *infinitesimal dilatation* of the deformation.

*d. Rate of deformation.* – When one is dealing with dynamics, and not statics, one can single out the  $u_0$  parameter as representing time and regard the set of curves  $x(u^0) = x(u^0, u_0^1, \dots, u_0^p)$  that one obtains fixing the remaining parameters as a *congruence of curves* that represents the evolution of each point  $(u_0^1, \dots, u_0^p)$  of the initial spatial object. That means, in particular, that one can be dealing with matter in any state (e.g., fluid or solid), even though the concept of congruences of curves is usually introduced in the context of relativistic hydrodynamics, nowadays.

We can then define the velocity vector field of the congruence of curves that is defined by  $x$  by:

$$\mathbf{v}(t, u_0^1, \dots, u_0^p) = \frac{\partial x(t, u_0^1, \dots, u_0^p)}{\partial t} = v^i(t, u_0^1, \dots, u_0^p) \frac{\partial}{\partial x^i}. \quad (1.31)$$

Hence, for each  $(u_0^1, \dots, u_0^p)$ , the vector field  $\mathbf{v}(t, u_0^1, \dots, u_0^p)$  will be tangent to the curve of the congruence that goes through  $(u_0^1, \dots, u_0^p)$ .

One can then think of the vector field  $\mathbf{v}$  as the time derivative of the displacement  $\mathbf{u}$ , and the rate of displacement  $d\mathbf{v} = v^j_{,i} dx^i \otimes \partial_j$  (or rather,  $dv = v_{i,j} dx^i \otimes dx^j$ ) can be polarized in the same way as  $du$ :

$$dv = \dot{e} + \omega, \quad (1.32)$$

in which:

$$\dot{e} = \frac{1}{2} \dot{e}_{ij} dx^i dx^j, \quad \dot{e}_{ij} \equiv v_{i,j} + v_{j,i} = \frac{\partial e_{ij}}{\partial t}, \quad (1.33)$$

$$\omega = \frac{1}{2} \omega_{ij} dx^i \wedge dx^j, \quad \omega_{ij} \equiv v_{i,j} - v_{j,i} = \frac{\partial \theta_{ij}}{\partial t}. \quad (1.34)$$

Analogously to the situation with  $u$  in place of  $v$ , one can regard  $\omega_{ij}$  as components of the 2-form that amounts to the exterior derivative of the covelocity 1-form:

$$\omega = d\mathcal{V}. \quad (1.35)$$

These tensors are referred to as the *rate of strain* and *kinematical vorticity*, respectively.

One can also split  $\dot{e}_j^i$  into a traceless part and a trace part:

$$\dot{e}_j^i = \Theta_j^i + \frac{1}{n} \dot{e} \delta_j^i, \quad \Theta_j^i = \dot{e}_j^i - \frac{1}{n} \dot{e} \delta_j^i. \quad (1.36)$$

The trace:

$$\dot{e} = \dot{e}_k^k = \frac{\partial v^k}{\partial x^k} = \operatorname{div} \mathbf{v} = \frac{\partial}{\partial t} (\operatorname{div} \mathbf{u}) = \frac{\partial e}{\partial t} \quad (1.37)$$

is referred to as the *rate of dilatation*, or *kinematical compressibility*.

**§ 2. Strain.** – Strain is basically one way of measuring how the geometry of an object changes under a deformation of the sort that was defined above. In particular, one looks at how the metric on the object changes; i.e., how the distances between all pairs of points change. That is why we shall first introduce “metric strain” in order to give a traditional picture of the deformation of regions of space, and then discuss the “frame strain,” which will be fundamental in explaining the quantum stress tensor later on.

*a. Finite strain.* – As long as one is dealing with finite deformations, one must always deal with the fact that the initial state and the final state are not generally the same set of points in space, so one must choose whether the definitions relate to things that are defined on one state or the other one. This basically amounts to the difference between the *Lagrange picture* of deformation and the *Euler picture*, respectively. The former is more customary in the case of the deformation of solids, which tend to have better-defined exemplars for  $\mathcal{O}_p$ , while the latter is generally used in fluid mechanics, since it is the channel or tank that encloses the fluid that has the consistent exemplar. However, the concept of “fluid cells,” which follow the flow, is often used, which is a more Lagrangian sort of concept.

In order to see how the metric on an object changes under a deformation, one starts by assuming that the ambient space ( $\mathbb{R}^n$ , for us) has a “background metric”  $g$  defined on it, which can be Euclidian or Lorentzian. One then gives the initial and final states the metrics that are induced by restricting it to tangent vectors to the objects.

In the Cauchy-Green theory of strain, one chooses the Lagrange picture and uses the diffeomorphism (onto)  $y : x_0(\mathcal{O}_p) \rightarrow \mathbb{R}^n$  to “pull back” the metric  $g$  on  $x_1(\mathcal{O}_p)$  by means of the diffeomorphism  $y$  to give a metric:

$$\bar{g} = y^* g \quad (2.1)$$

on  $x_0(\mathcal{O}_p)$ . By definition, if  $\mathbf{v}, \mathbf{w}$  are tangent vectors to the initial state at  $x$  then:

$$\bar{g}_x(\mathbf{v}, \mathbf{w}) = g_{y(x)}(dy|_x(\mathbf{v}), dy|_x(\mathbf{w})). \quad (2.2)$$

The component form of this is:

$$\bar{g}_{ij}(x) = y_i^k(x) y_j^l(x) g_{kl}(y(x)). \quad (2.3)$$

If one introduces the displacement vector field  $\mathbf{u}(x)$  for the diffeomorphism  $y$  (so  $y_j^i = \delta_j^i + u_j^i$ ) then one can express the last equation in the form:

$$\bar{g}_{ij}(x) = [\delta_i^k \delta_j^l + \delta_i^k u_j^l + \delta_k^l u_i^k + u_i^k u_j^l](x) g_{kl}(y(x)). \quad (2.4)$$

The *finite strain tensor* that this defines is then the difference between the deformed, pulled-back metric and the initial one:

$$E = \bar{g} - g = y^* g - g. \quad (2.5)$$

Naively, if one uses (2.4) in (2.5) then the components of the finite strain tensor will take the form:

$$E_{ij} = [\delta_i^k \delta_j^l + \delta_i^k u_j^l + \delta_k^l u_i^k + u_i^k u_j^l](x) g_{kl}(y(x)) - g_{ij}(x).$$

However, one sees that actually one is dealing with the components of the  $g$  at two distinct points of  $\mathbb{R}^n$ , so unless one makes the usual assumption that the natural coordinate system is also orthonormal<sup>(1)</sup>, or at least it makes the components of  $g$  constant in space, one can go no further. With that assumption, however, one can cancel the initial and final background metric, and get:

$$E_{ij} = g_{kl}[\delta_i^k u_j^l + \delta_k^l u_i^k + u_i^k u_j^l] = u_{ij} + u_{ji} + g_{kl} u_i^k u_j^l = e_{ij} + g_{kl} u_i^k u_j^l. \quad (2.6)$$

All of the expressions involved are now defined at a point of the initial state.

Not all second-rank, symmetric, covariant tensor fields on any region of  $\mathbb{R}^n$  can serve as the finite strain tensors of some diffeomorphism of the region. If one considers (2.6) then one will see that part of the problem is rooted in the fact that not all functions  $u_j^i(t, x)$  with values in  $GL(n)$  represent the differential matrices of diffeomorphisms. Hence, one can think of (2.6) as a system of first-order partial differential equations for the functions  $y_i(t, x)$ , so the issue at hand is the integrability of that system. Traditionally (see, e.g., [3]), the necessary condition for the integrability of finite strain was based upon the fact that the diffeomorphism of the region should not change the Riemann curvature of the initial metric  $g$  when it becomes the deformed metric  $\bar{g}$ ; in particular, the Riemann curvature should stay zero. That is essentially the *St.-Venant compatibility condition*; more precisely, its linearization will give St.-Venant's compatibility condition for infinitesimal strain. (See, e.g., Murnaghan [9])

Here, we have a perfect example of how the Riemann-Cartan usage of the terms “torsion” and “curvature” can create confusion in continuum mechanics, since the

---

<sup>(1)</sup> This assumption is not as trivial as it sounds, in general, since it implies that the Riemannian curvature of the background metric must vanish.

Riemann-Cartan usage of the terms has more to do with the integrability of parallel translation than it does with the bending and twisting of things. (The author has discussed this subject at length in [10].)

Since we shall have no further use for the issue of the integrability of strain, we shall not introduce the necessary mathematical overhead at this point. However, we will point out that Pommaret [11] has long been emphasizing a different approach to the integrability of strain that is rooted in Cosserat-related considerations, although it demands even more mathematical overhead than the traditional approach.

*b. Infinitesimal strain.* – In the last equation, we saw the reappearance of the infinitesimal strain tensor  $e$  as a component of the finite strain tensor  $E$ . Basically, the infinitesimal strain represents the first-order contribution of the displacement gradient to the finite strain. From its definition, it is always a symmetric, second-rank covariant tensor, and thus its component matrix will always be diagonalizable. The frame in which  $e_{ij}$  (or rather,  $e^i_j$ ) is diagonal is then called the *principal frame* for the infinitesimal strain tensor, and the diagonal elements  $e_x, e_y, e_z$  are referred to as the *principal strains* in the principal directions. There are three conditions under which the principal strains can overlap:

1. Isotropy:  $e_x = e_y = e_z$ .
2. Uniaxiality: Two principal strains are equal, but not the third.
3. Biaxiality: All three are distinct.

In the general case, one can think of the diagonal elements of  $e_{ij}$  as *infinitesimal elongations*, and the off-diagonal elements as *infinitesimal shearing strains*. When one decomposes  $e^i_j$  into a traceless part and a trace, as above:

$$e^i_j = \overset{\circ}{e}^i_j + \frac{1}{3}e \delta^i_j, \quad \overset{\circ}{e}^i_j = e^i_j - \frac{1}{3}e \delta^i_j, \quad (2.7)$$

the traceless part  $\overset{\circ}{e}^i_j$  will be referred to as the *deviatoric strain tensor*, and will represent an infinitesimal volume-preserving linear transformation, while the trace part will be the *volumetric strain tensor*, whose magnitude  $\frac{1}{3}e$  will take the form of a *mean elongation*.

If the final state  $y(t)$  is a differentiable function of time then so is  $e(t)$ , and differentiating with respect to time will product the infinitesimal rate of strain. However, there is another way of characterizing the infinitesimal rate of strain that has a deep geometric significance. If one regards the deformation as a motion from the initial state to the final state then one can express the final state  $y(t)$  as a differentiable function of time, as well as the finite strain tensor:

$$E(t) = y(t)^* g - g. \quad (2.8)$$

When one differentiates  $E(t)$  at  $t = 0$  one will get the infinitesimal strain tensor in the form of the *Lie derivative* of  $g$  with respect to the velocity vector field  $\mathbf{v} = \partial y / \partial t$ :

$$e = \left. \frac{dE}{dt} \right|_{t=0} = L_v g. \quad (2.9)$$

Those readers with a background in Riemannian geometry will know that it is a theorem of Wilhelm Killing that, for the general metric  $g$ :

$$L_v g_{ij} = \nabla_i v_j + \nabla_j v_i, \quad (2.10)$$

in which  $\nabla_i$  represents the covariant derivative in the direction of  $x^i$  when one uses the Levi-Civita connection of  $g$ . When the components of  $g$  are constant, that covariant derivative will reduce to the partial derivative with respect to  $x^i$ , and we will have:

$$L_v g_{ij} = \partial_i v_j + \partial_j v_i, \quad (2.11)$$

which is identical with the rate of strain.

If one wishes to use the Lie derivative to define the infinitesimal strain then one can apply the same process to the displacement vector field  $\mathbf{u}$  for the deformation:

$$e = L_u g. \quad (2.12)$$

Vector fields for which  $L_u g = 0$  are called *Killing vector fields*. They can then be characterized by the fact that their flows consist of one-parameter families of isometries or that their infinitesimal strain (or rate of strain, in the case of  $v$ ) vanishes. Hence, one sees in that the intimate relationship between strain and the measure of the degree to which a diffeomorphism fails to be an isometry (i.e., rigid motion).

Here, one must be careful, since differential geometry has many examples of diffeomorphisms that are not rigid, but still preserve the distances. However, one must go to diffeomorphisms of submanifolds of a higher-dimensional space in order to find them. For instance, if one bends a flexible, extensible wire without stretching or compressing it then the one-dimensional metric along it (e.g., arc length) will not change, even though the deformation is clearly not rigid. Similarly, developable surfaces, such as cylinders and cones, are all isometric to flat plates, even though the deformation of a flat rectangular plate into a cylinder or a flat angular wedge into a cone is clearly not rigid. Hence, one begins to suspect that the Cauchy-Green definition of strain is too coarse-grained to account for all of the possibilities, since it will pass over any non-rigid deformation that is still an isometry. That is why we will introduce “frame strain” shortly.

Note that in the case of infinitesimal deformations, it is irrelevant whether one is considering the Lagrange picture or the Euler picture, since the distinction between “initial” and “final” state has disappeared.

**§ 3. Stress.** – Stress is to strain what force is to displacement. That is, it is the dynamical state that is associated with the kinematical one. In most cases, it is associated with an infinitesimal kinematical state, such as the infinitesimal strain tensor. That is because in reality the phenomena that are associated with finite deformations are much



more complicated than their infinitesimal approximations, and as a result, the material constants that are obtained from engineering tests almost always tend to describe infinitesimal deformations.

*a. The Cauchy definition of stress.* – The units of stress are basically those of pressure, but one also considers the anisotropic possibility that the pressure on a reference area through a point will depend upon the angular orientation of the area – i.e., the direction of its normal vector. Once again, this is typically assumed to be a linear relationship. That is, if  $\mathbf{n}$  is the unit normal vector to a unit square  $A_x$  through a point  $x(u) \in \mathbb{R}^3$  then the vector  $\mathbf{f}$  that describes the force that acts upon  $A_x$  will be given by a linear map  $\sigma(x): T_x\mathbb{R}^3 \rightarrow T_x\mathbb{R}^3$ ,  $\mathbf{n} \mapsto \mathbf{f} = \sigma(x)(\mathbf{n})$  so if one has defined a frame for  $T_x\mathbb{R}^3$  then one can express this relationship in terms of the components of everything:

$$f^i = \sigma_j^i(x) n^j. \quad (3.1)$$

The  $3 \times 3$  component matrix  $\sigma_j^i(x)$ , or rather its doubly-covariant form  $\varepsilon_{ij}(x)$ , then defines the *Cauchy stress tensor*. It is often assumed to be symmetric, but since that is equivalent to the absence of internal couple-stresses, the question arises whether such couple-stresses have any basis in reality. Since the time of the Cosserat brothers, Eugène and François, the fact that internal couple-stresses are physically realistic has been taken seriously by an increasing number of researchers in theoretical mechanics. We shall return to discussing the Cosserat approach to continuum mechanics as it becomes relevant in what follows.

Traditionally, the Cauchy stress tensor is associated with a geometric object that is called the *Cauchy stress tetrahedron*. Actually, the fact that it is a tetrahedron has more to do with the period of history in which Cauchy did his work than anything else. Basically, to the mathematicians and scientists of that era, geometry always meant projective geometry, and the “reference tetrahedron” plays the same role in the context of projective spaces that a linear frame does in the context of linear spaces. Of course, the matrix of the stress tensor does not have to be invertible, so its columns or rows do not have to define linearly-independent vectors. However, even in the non-invertible case, one can still think of the  $i^{\text{th}}$  row of the matrix as representing the force per unit area that acts in the direction of the  $i^{\text{th}}$  frame vector.

In the case of a symmetric  $\sigma_{ij}$ , one can speak of the *principal frame* for the tensor, which will be the frame for which the matrix  $\sigma_{ij}$  is diagonal, which will consist of eigenvectors of the matrix  $\sigma_j^i$ . Those diagonal elements  $\sigma_x, \sigma_y, \sigma_z$  are then called the *principal stresses* for  $\sigma$ . That then defines three possibilities regarding how the principal stresses can overlap, and they are analogous to the cases for the infinitesimal strain tensor.

In the general case, one can think of the diagonal elements of  $\sigma_{ij}$  as *pressures*, and the off-diagonal elements as *shearing stresses*. When one decomposes  $\sigma_j^i$  into a traceless part and a trace, as above:

$$\sigma_j^i = \overset{\circ}{\sigma}_j^i + \frac{1}{n} \sigma_k^k \delta_j^i, \quad \overset{\circ}{\sigma}_j^i = \sigma_j^i - \frac{1}{n} \sigma_k^k \delta_j^i, \quad (3.2)$$

the traceless part  $\overset{\circ}{\sigma}_j^i$  is referred to as the *deviatoric stress*, while the trace part is a *mean pressure*, whose magnitude is  $\frac{1}{n} \epsilon_k^k$ .

*b. The energy-momentum-stress tensor.* – When  $\mathbb{R}^n$  is  $\mathbb{R}^4$ , and one has a time+space decomposition of it into  $\mathbb{R} \oplus \mathbb{R}^3$ , moreover, the time-time and time-space components of a stress tensor would have different interpretations from the space-space ones, just as the exterior derivative of a four-dimensional covelocity splits into a linear acceleration and an angular velocity, due to the difference in interpretation between time derivatives and spatial derivatives. In particular, one finds that pressures have the same basic units as energy densities and momentum fluxes. For instance, if a fluid has a mass density of  $\rho$  and a flow velocity of  $\mathbf{v}$  then the scalar  $\frac{1}{2} \rho v^2$ , which looks like a kinetic energy density, is usually referred to as the *dynamic pressure*. If the fluid were streaming from the end of a hose and impinging upon a surface, such as a window or car body, then the dynamic pressure would be the pressure that the fluid exerted on the plane of impact.

**§ 4. Mechanical constitutive laws.** – The association of a stress state with a state of strain (whether finite or infinitesimal) is the point at which the empirical considerations must be introduced into the model for the equilibrium state or motion of a deformable extended object.

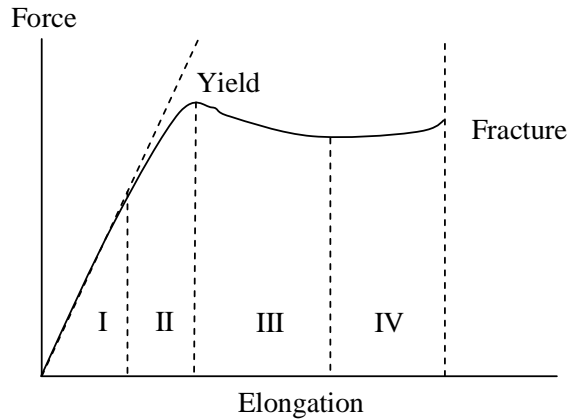


Figure I.1. Typical force vs. elongation graph for a real-world material.

*a. Basic notions.* – Generally, the association is something that is established by engineering test stands, in which a cylindrical sample of the material is machined to have a bottleneck in the middle, so it will not fracture inside the test stand, and threaded on the ends so it can be held firmly.

A programmed sequence of forces is applied to the sample and the resulting elongation is measured. A typical graph of force vs. displacement might take the form that is illustrated in Fig. I.1.

The various regimes are referred to as:

- I. Linear elasticity.
- II. Nonlinear elasticity.
- III. Plastic deformation. (Actually, the illustration would relate to a “softening” material.)
- IV. Work hardening.

The first regime is the one that gets the most attention, since it defines *Hooke’s law* for elastic springs (viz.,  $F = -k \Delta x$ ) and elastic materials, more generally, which is typically assumed to be true for small enough displacements. For three-dimensional materials, and not just springs, the  $F$  would get replaced with an  $\sigma_{ij}$ , which would then be the force per unit area in a specific direction, and the  $\Delta x$  would get replaced by the percentage elongation  $e_{ij} (\Delta l / l)$  in a given direction. A linear relationship between stress and infinitesimal strain would then take the form  $\sigma = C(e)$ , where  $C : S^2(\mathbb{R}^3) \rightarrow S^2(\mathbb{R}^3)$  would be an invertible linear map from the six-dimensional vector space of symmetric, second-rank covariant tensors over  $\mathbb{R}^3$  to itself. It is usually represented in the component form:

$$\sigma_{ij} = C_{ij}^{kl} e_{kl}. \quad (4.1)$$

Various symmetries are commonly attributed to the component array  $C_{ij}^{kl}$  besides the symmetry in  $ij$  and  $kl$ . For instance, if one prefers to regard  $e$  as an element of  $S_2(\mathbb{R}^3)$  (namely, a symmetric, doubly-contravariant second-rank tensor over  $\mathbb{R}^3$ , so the associated stress will be an element of its dual space) then the components  $C_{ijkl}$  of  $C$  will all be covariant, and one will sometimes assume that  $C_{ijkl} = C_{klij}$ , along with the other symmetries.

Further reductions in the possible components  $C_{ijkl}$  follow from assuming various symmetries to the material that it refers to. Ultimately, if the material is linear, isotropic, and homogeneous, then one can express the constitutive law in the form:

$$\sigma_{ij} = G e_{ij}^{\circ} + K e_k^k \delta_{ij}, \quad (4.2)$$

in which  $G$  is a constant called the *shear modulus*, and  $K$  is a constant called the *bulk modulus*.

More generally, a material can be nonlinear, anisotropic, and inhomogeneous. In fact, the very assumption that the relationship between stress and strain is an algebraic one, and not a differential or integral relationship, or even a combination of the three, is not always a good approximation. For a dispersive material, the relationship becomes an integral transformation for which the strain at the neighboring points might affect the

stress at a given point. Some materials have “memory,” which amounts to saying that their constitutive properties can depend upon the time-parameterized program of deformations that they have experienced up to the current state, which amounts to an integral over time.

*b. Examples of constitutive laws.* – One of the recurring themes in our eventual discussion of the continuum-mechanical models for quantum wave equations is that perhaps they are giving us a strongly-worded hint concerning the nature of matter at the atomic-to-subatomic level. Hence, it is good to have some specific examples in mind for the form that mechanical constitutive laws take for the kinds of matter at the macroscopic level that have been examined in many laboratories.

Perhaps the simplest medium is the *inviscid (or perfect) fluid*. Such a material is characterized by the fact that it cannot support any strain. Hence, there is no constitutive law, but just a general form for the stress tensor. Typically:

$$\sigma_{ij} = \pi \delta_{ij}, \quad (4.3)$$

in which  $\pi$  is the pressure. This definition assumes that the fluid is isotropic, as opposed to liquid crystals, which can be anisotropic. Moreover, the pressure does not have to be constant in time or space (even for incompressible fluids), as one can see in the cases of the atmosphere or the oceans.

Usually, the distinction between liquids and gases as examples of fluids is defined by compressibility. That is, a gas is a compressible fluid, while a liquid is incompressible. Of course, perfect incompressibility, like perfect rigidity, is impossible in the eyes of relativistic continuum mechanics, since it would imply an infinite speed for the propagation of waves.

For a *viscous fluid*, the rate of strain couples to the stress tensor, in addition to the bulk pressure:

$$\sigma_{ij} = \pi \delta_{ij} + \eta \dot{\epsilon}_{ij}, \quad (4.4)$$

in which  $\eta$  is then called the *viscosity*.

**§ 5. Balance principles.** – Ultimately, the formulation of differential equations that would determine the equilibrium state of a deformable object in statics or its motion in dynamics is based in certain “first principles” that take the form of balance principles in the case of open systems and conservation laws in the case of closed systems. On the surface of things, they often have a somewhat tautological character, which is why it is only when one can imagine other possibilities that they take on the character of equations, not identities.

We shall now examine some of the balance principles that one encounters most often in continuum mechanics. We shall first discuss them in their non-relativistic form, and then present the modifications that would be necessary if one were to give them a relativistic formulation.

*a. Balance of mass.* – Naively, one expects that mass cannot be created or destroyed but only moved around in space. Of course, that sort of thinking is distinctly non-relativistic and non-quantum, since the relative motion of a body can change its observed mass, and the creation or destruction of mass by pair creation and annihilation at the level of elementary particles is commonplace.

Hence, as long as the nature of problem does not demand the introduction of relativity or quantum theory, one of the oldest conservation laws to be applied to processes, such as chemical reactions, was the notion that the total mass  $M(t)$  of a closed system should stay constant in time. In an open system, the balance principle takes the form of a first-order ordinary differential equation:

$$\frac{dM}{dt} = \sum_{in} \dot{m}_i - \sum_{out} \dot{m}_j, \quad (5.1)$$

where the right-hand side details the various rates at which mass is being added or subtracted from the system. Hence, for a closed system, all of the rates  $\dot{m}_i$  and  $\dot{m}_j$  must vanish. There is also the *steady-state* possibility, in which they do not vanish, but the two summations cancel. (One might imagine water flowing into a sink with its drain open.)

When the mass is distributed continuously over a bounded volume  $V$  with a mass density of  $\rho(t, x^i)$ , the total mass will become the integral of the density over the volume:

$$M(t) = \int_V \rho(t, x^i) dV, \quad (5.2)$$

in which  $dV$  is the differential volume element on space.

Hence, as long as  $V$  does not change in time:

$$\frac{dM}{dt} = \int_V \frac{\partial \rho}{\partial t}(t, x^i) dV. \quad (5.3)$$

The right-hand side of the balance equation (5.1) consists of the resultant of the mass currents (i.e., fluxes) that flow through the boundary surface  $\partial V$ . If one thinks of the mass current that flows through the surface as something that takes the form of the vector field  $\rho \mathbf{v}$ , where  $\mathbf{v}$  is the flow velocity, then that *resultant mass flux*  $\Phi[\partial V]$  will take the form of the surface integral:

$$\Phi[\partial V] = \int_{\partial V} \#(\rho \mathbf{v}), \quad (5.4)$$

in which the *Poincaré dual* of a vector field  $\mathbf{A}$  is the 2-form:

$$\#\mathbf{A} = i_{\mathbf{A}} dV = \frac{1}{2} A^i \varepsilon_{ijk} dx^j \wedge dx^k. \quad (5.5)$$

(In conventional vector calculus, this expression gets denoted by  $\mathbf{A} \cdot dS$ .) For instance, the dual  $\#\mathbf{n}$  of the unit normal vector field  $\mathbf{n}$  on a surface  $\phi(x^i) = \text{const.}$  in  $\mathbb{R}^3$  is the

surface element  $du \wedge dv$  if one uses the adapted coordinates  $(u, v, \phi)$  for the points of the surface, since that will make the components of  $\mathbf{n}$  be  $(0, 0, 1)$ .

One can then think of the 2-form  $\#(\rho \mathbf{v})$  as being the *mass flux density* that is dual to the *mass current density*  $\rho \mathbf{v}$ , which also represents a linear momentum density. Similarly, if  $\sigma$  is an electric charge density then  $\#(\sigma \mathbf{v})$  will be the *electric charge flux density* that is dual to the *electric current density*  $\sigma \mathbf{v}$ .

From Stokes's theorem for exterior forms, which takes the form of Gauss's theorem here, one will have:

$$\Phi[\partial V] = \int_V d \#(\rho \mathbf{v}) = \int_V \# \operatorname{div}(\rho \mathbf{v}) = \int_V \operatorname{div}(\rho \mathbf{v}) dV . \quad (5.6)$$

Hence, if the vector  $\mathbf{v}$  is assumed to point outward on the surface then one must have:

$$\int_V \frac{\partial \rho}{\partial t} dV = - \int_V \operatorname{div}(\rho \mathbf{v}) dV . \quad (5.7)$$

If this must be true for every possible volume  $V$  then one can express this as a partial differential equation for  $\rho$  :

$$\frac{\partial \rho}{\partial t} = - \operatorname{div}(\rho \mathbf{v}). \quad (5.8)$$

This equation then represents the balance of mass or *continuity equation*.

One can expand the right-hand side to give:

$$\frac{\partial \rho}{\partial t} = - \mathbf{v} \rho - \rho \operatorname{div}(\mathbf{v}),$$

which will give:

$$\frac{d \rho}{dt} = - \rho \operatorname{div}(\mathbf{v}), \quad (5.9)$$

with the generic definition:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + v^i \frac{\partial f}{\partial x^i} \quad (5.10)$$

for the derivative of a function  $f$  along the flow of  $\mathbf{v}$ ; i.e., its Lie derivative with respect to  $\mathbf{v}$ ; this derivative is also sometimes called the “material” derivative or “substantial” derivative in the standard literature.

Therefore, from (5.9), the flow of  $\mathbf{v}$  will be *incompressible* [i.e.,  $\operatorname{div}(\mathbf{v}) = 0$ ] iff the mass density is constant along that flow.

One can also express (5.8) in the component form:

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v^i)}{\partial x^i} = 0, \quad (5.11)$$

and if one extends space  $\mathbb{R}^3$  to Galilean space-time  $\mathbb{R} \oplus \mathbb{R}^3$ , so that time  $t$  becomes the coordinate  $x^0$  and one always sets  $v^0 = 1$ , then one can give the last equation the even more concise form:

$$0 = \operatorname{div}(\rho \mathbf{v}) = \frac{\partial(\rho v^\mu)}{\partial x^\mu}. \quad (5.12)$$

That is, it simply describes the vanishing of a four-dimensional divergence of a mass current vector field.

Of course, the four-dimensional picture for mechanics is usually treated relativistically, but one often finds that formulating non-relativistic mechanics in a four-dimensional Galilean space-time can make the transition to relativistic mechanics more natural.

*b. Balance of linear momentum.* – Although it is traditional to discuss the balance of linear momentum by analogy with the balance of mass, actually, the basic step at which one goes from the linear momentum density  $\mathbf{p}(t, x) = \rho(t, x) \mathbf{v}(t, x)$  of a congruence of massive curves to the total linear momentum  $m(t) \mathbf{v}(t)$  of a massive point along a single curve by integrating the component functions  $p^i(t, x)$  over all space at each time point  $t$  has some fundamental limitations.

In the first place, although one can unambiguously define the total mass  $m(t)$  to be the spatial integral of  $\rho(t, x)$  at time  $t$ , the issue of defining the curve  $x(t)$  whose velocity vector field will be  $\mathbf{v}(t, x)$  is not as unambiguous. In non-relativistic mechanics, it is traditional to define the center-of-mass  $x_{\text{cm}}(t)$  of the density function  $\rho(t, x)$  and use that curve as the reference curve, but since the mass density by itself is not a Lorentz-invariant function, one cannot define a Lorentz-invariant center-of-mass, either.

Secondly, the process of integrating component functions with respect to a chosen local frame field is not frame-invariant, unless one restricts oneself to local frame fields that differ from each other by constant transition functions. That is because otherwise one would have to include the component functions of the frame transition function in the integration, and not take them out of the integral like so many constants. For instance, suppose  $\mathbf{p}(t, u) = p^i(t, u) \mathbf{e}_i(t, u)$ , where  $\mathbf{e}_i$  is a local frame field that includes the support of  $\mathbf{p}$ , which is a congruence of curves  $x(t, u)$  that is parameterized by  $u = (u^1, \dots, u^p)$ . Suppose further that one defines the *total linear momentum* at  $t$  to have spatial components:

$$P^i(t) = \int_{\Sigma} p^i(t, u) dV_u \quad (5.13)$$

with respect to  $\mathbf{e}_i$ , and that one makes a transition to another local frame field  $\bar{\mathbf{e}}_i(t, u) = \mathbf{e}_j(t, u) \tilde{L}_i^j(t, u)$ , in which  $\tilde{L}_i^j(t, u)$  is the inverse of the transition function  $L_i^j(t, u)$ , which takes its values in  $GL(3)$ . Here,  $\Sigma$  is the spatial manifold, and  $dV_u$  is its volume element, when it is pulled back to the parameter space.  $\mathbf{p}$  will then have new components  $\bar{p}^i = \tilde{L}_j^i p^j$  with respect to  $\bar{\mathbf{e}}_i$ , and if one defines the new total linear momentum components to be:

$$\bar{P}^i = \int_{\Sigma} \bar{p}^i(t, u) d\bar{V}_u = \int_{\Sigma} L_j^i(t, u) p^j(t, u) |L_j^i| dV_u \quad (5.14)$$

then one will see that it is only if  $L_j^i(t, u)$  is a spatially-constant function with determinant unity that one can, in fact transform the components of  $\mathbf{P}$  linearly as:

$$\bar{P}^i(t) = L_j^i(t) P^j(t). \quad (5.15)$$

Hence, we shall take the position that really the differential formulation of the balance principle is more definitive than the integral formulation. The soul of the differential formulation of the balance principle for any physical observable is that its derivative along the curves of the congruence  $x(t, u)$  – i.e., its Lie derivative with respect to  $\mathbf{v}(t, u)$  – must equal the sum of the external contributions to that density. In the case of the linear momentum density, that would mean the resultant  $\mathbf{f}_{\text{ext}}$  of the external force densities that act upon  $\mathbf{p}$ :

$$\mathbf{L}_v \mathbf{p} = \mathbf{f}_{\text{ext}}. \quad (5.16)$$

Now, when the spatial support  $\Sigma(t)$  of  $\rho$ , and thus the domain of definition of  $x(t, u)$  and  $\mathbf{v}(t, u)$ , has a boundary  $\partial \Sigma(t)$  at each  $t$ , there will be two types of external force densities that act upon the points of  $\Sigma(t)$ :

1. Volume force densities  $\mathbf{f}_V$ , such as gravitation, electrostatic forces, magnetic forces, which act upon the points of  $\Sigma(t)$ .

2. Surface force densities  $\mathbf{f}_S$ , such as surface tension, atmospheric pressure, and some types of applied loads, which act upon the points of  $\partial \Sigma(t)$ .

Now, the surface forces can be converted into volume force densities by using Gauss's law on the vector field  $\mathbf{f}_S$  (i.e., Stokes's law for the  $n - 1$ -form  $\#\mathbf{f}_S$ ), as long as one assumes that the surface force density  $\mathbf{f}_S$  consists of simply the boundary values of a volume force. If  $\mathbf{n}$  is the normal vector field on  $\partial \Sigma(t)$ , so one can express the spatial components of  $\mathbf{f}_S$  with respect to some local frame field  $\mathbf{e}_i(t)$  that includes  $\Sigma(t)$  in the form:

$$f_S^i = -\sigma_j^i n^j \quad (5.17)$$

then that will make:

$$d \wedge \#\mathbf{f}_S = -d(\sigma_j^i n^j) \wedge \#\mathbf{e}_i - (\sigma_j^i n^j) d \wedge \#\mathbf{e}_i. \quad (5.18)$$

The second term will vanish for a natural frame field  $\mathbf{e}_i = \partial_i$ , and since:

$$d_s(\sigma_j^i n^j) = \partial_k(\sigma_j^i n^j) dx^k$$

and

$$dx^k \wedge \#\partial_i = \delta_i^k V_x,$$

that will make:



$$d \wedge \# \mathbf{f}_S = - \partial_i (\sigma_j^i n^j) V_x = - \#(\operatorname{div} \mathbf{f}_S),$$

so:

$$\operatorname{div} \mathbf{f}_S = - \partial_i (\sigma_j^i n^j) = - [(\partial_i \sigma_j^i) n^j + \sigma_j^i (\partial_i n^j)]. \quad (5.19)$$

One best extends  $\mathbf{n}$  into the interior of  $\Sigma$  by assuming that the frame field  $\mathbf{e}_j$  is *adapted* to  $\mathbf{n}$ , so  $\mathbf{n}$  is itself one of the frame members, and the other frame members are tangent to  $\partial\Sigma$ . The components of  $\mathbf{n}$  will then be constants (viz., one of them will be 1, and the others will be 0). One then defines the 1-form on  $\Sigma$ :

$$f_S = - (\partial_i \sigma_j^i) \theta^j, \quad (5.20)$$

where  $\theta^i$  is the reciprocal coframe field to  $\mathbf{e}_i$ . [i.e.,  $\theta^i(\mathbf{e}_j) = \delta_j^i$ .]

It now becomes more convenient to represent  $p$  and  $f_V$  as 1-forms, as well, and the balance equation (5.16) will take the form:

$$\mathbf{L}_v p = f_V + f_S = (f_i - \partial_j \sigma_i^j) \theta^i \quad (5.21)$$

Now, according to a basic property of the Lie derivative [see App. B], if  $p = p_i \theta^i$  then:

$$\mathbf{L}_v p = \mathbf{L}_v p_i \theta^i + p_i \mathbf{L}_v \theta^i.$$

As long as one uses a “convected” frame field, for which  $\mathbf{L}_v \theta^i$  will vanish, that will make:

$$\mathbf{L}_v p = (\mathbf{L}_v p_i) \theta^i = (\mathbf{v} p_i) \theta^i = \frac{dp_i}{dt} \theta^i.$$

The condition for  $\mathbf{L}_v \theta^i$  to vanish reads:

$$0 = \mathbf{L}_v \theta^i = i_v d \wedge \theta^i + d v^i, \quad (5.22)$$

explicitly.

One example of a convected frame field is a natural frame field ( $\theta^i = dx^i$ ) that is adapted to  $\mathbf{v}$ , so  $v^i$  will be constants [e.g., (1, 0, ..., 0)]. Of course, that would imply that  $\mathbf{v}$  would have to be irrotational, in order to have  $d \wedge v = 0$ ; i.e.,  $v$  would have to admit a velocity potential (if only locally).

Ultimately, we arrive at the conventional component form that the balance of momentum takes:

$$\frac{dp_i}{dt} = f_i - \partial_i \sigma_j^i. \quad (5.23)$$

If mass is conserved (so  $d\rho/dt = 0$ ) then that will take the form:

$$\rho \frac{dv_i}{dt} = f_i - \partial_j \sigma_i^j. \quad (5.24)$$

In the case of inviscid fluids, for which  $\sigma_i^j = \pi \delta_i^j$ , this takes the form of *Euler's equation*:

$$\rho \frac{dv_i}{dt} = f_i - \partial_i \pi. \quad (5.25)$$

In the case of elasticity, for which the Lagrangian picture is more commonly used, the Lie derivative becomes the partial derivative with respect to time, and the equations of motion take the form:

$$\rho \frac{\partial v_i}{\partial t} = f_i - \partial_j \sigma_i^j. \quad (5.26)$$

When the problem is one of statics, there will be no acceleration, so the equations will become equations of equilibrium:

$$\partial_j \sigma_i^j = f_i. \quad (5.27)$$

*c. Balance of angular momentum.* – There are two types of angular momentum to consider for a moving, deformable mass distribution  $\rho(t, x)$ , which amount to the external and internal angular momentum densities. It is important for later discussion to note that “internal angular momentum” is quite distinct from the quantum concept of “intrinsic angular momentum,” or “spin,” which is rooted in the weight of the representation of the rotation or Lorentz group in the field space of the quantum wave function. The internal angular momentum is basically an orbital angular momentum density that is associated with the material's response to stress-couples.

The external or *orbital angular momentum density 2-form*  $L = \frac{1}{2} L_{ij} dx^i \wedge dx^j$  of a moving mass distribution that is associated with a linear momentum density 1-form  $p(t, x^j) = p_i dx^i$  is essentially the moment of that momentum with respect to an observation point (e.g., the origin):

$$L(t, x^i) = p \wedge R = \frac{1}{2} (p_i x_j - p_j x_i) dx^i \wedge dx^j, \quad (5.28)$$

in which  $R(t, x^j) = x_i dx^i$  is the position 1-form of the point  $x$  of  $\rho$  at time  $t$  with respect to the observer.

In order to get the conservation law that goes with  $L$ , one first differentiates it with respect to time:

$$\frac{dL}{dt} = \frac{dp}{dt} \wedge R + p \wedge v. \quad (5.29)$$

As long the momentum density  $p$  is of the “convective” type, so  $p = \rho v$ , the last term will disappear, but if  $p$  includes a *transverse momentum density* component, which is not parallel to  $v$ , then it will be non-vanishing. Hence, for the time being, we shall leave it in.

If one applies the balance of linear momentum (5.23) to  $dp_i / dt$  then (5.29) will become:

$$\frac{dL_{ij}}{dt} = M_{ij} - \partial_k \sigma_i^k x_j - \partial_k \sigma_j^k x_i + p_i v_j - p_j v_i,$$

in which we have defined the *moment of external force density 2-form*:

$$M = f \wedge R = \frac{1}{2} (f_i x_j - f_j x_i) dx^i \wedge dx^j. \quad (5.30)$$

An application of the product rule to the second two terms on the right-hand side will give this the form:

$$\frac{dL_{ij}}{dt} = M_{ij} - \partial_k S_{ij}^k + \sigma_{ij} - \sigma_{ji} + p_i v_j - p_j v_i, \quad (5.31)$$

in which we have defined the vector-valued 2-form:

$$S^k = \frac{1}{2} S_{ij}^k dx^i \wedge dx^j, \quad S_{ij}^k = \sigma_i^k x_j - \sigma_j^k x_i, \quad (5.32)$$

which we shall then call the *internal stress-couple density*.

Hence, external angular momentum will be conserved iff one has:

$$M_{ij} = \partial_k S_{ij}^k - \sigma_{ij} + \sigma_{ji} + p_i v_j - p_j v_i. \quad (5.33)$$

In the absence of external force-moments, one can still have:

$$\partial_k S_{ij}^k = \sigma_{ij} - \sigma_{ji} - p_i v_j + p_j v_i. \quad (5.34)$$

Hence, the sources of internal stress-couples are found in the possible asymmetry of the stress tensor and the existence of transverse momentum.

*d. Balance of energy.* – If one writes the balance of linear momentum in the form:

$$L_{\mathbf{v}} p_i = f_i - \partial_j \sigma_j^i \quad (5.35)$$

and contracts both sides with  $v^i$  then since one will have:

$$v^i L_{\mathbf{v}} p_i = v^i \frac{\partial p_i}{\partial t} + v^i v^j \frac{\partial p_i}{\partial x^j} = \frac{\partial (\frac{1}{2} \rho v^2)}{\partial t} + v^j \frac{\partial (\frac{1}{2} \rho v^2)}{\partial x^j} = L_{\mathbf{v}} (\frac{1}{2} \rho v^2),$$

that will make:

$$L_{\mathbf{v}} (\frac{1}{2} \rho v^2) = v^i (f_i - \partial_j \sigma_j^i). \quad (5.36)$$

Now, one can interpret the expression  $\frac{1}{2} \rho v^2$  as either a kinetic energy density or the dynamic pressure of the motion, and since the right-hand side represents a power

exchange density, the last equation says that the power exchange density is equal to the rate of change of  $\frac{1}{2}\rho v^2$  along the flow of  $\mathbf{v}$ .

If  $f_i = -\partial_i U$ , where the potential energy density  $U$  is not a function of time then:

$$v^i (f_i - \partial_j \sigma_i^j) = -v^i \partial_i U - v^i \partial_j \sigma_i^j .$$

If we further assume that a *stress potential vector*  $\psi^j$  exists, so:

$$\sigma_i^j = \partial_i \psi^j , \quad (5.37)$$

then, if  $\psi^j$  is not a function of time, either, we will have:

$$v^i (f_i - \partial_j \sigma_i^j) = -v^i \partial_i (U + \partial_j \psi^j) = -L_{\mathbf{v}}(U + \partial_j \psi^j)$$

and if we define the total energy density to be:

$$\mathcal{E} = \frac{1}{2}\rho v^2 + U + \partial_j \psi^j \quad (5.38)$$

then (5.36) can be put into the form:

$$L_{\mathbf{v}}\mathcal{E} = 0. \quad (5.39)$$

Hence, as long as the external forces and internal stresses are conservative, the total energy density will be constant along the motion of the mass distribution that is described  $\rho$ .

In the case of inviscid fluids, one will have  $\partial_j \sigma_i^j = -\partial_i \pi$ , and if  $\pi$  is also independent of time then (5.38) will be replaced with:

$$\mathcal{E} = \frac{1}{2}\rho v^2 + U + \pi, \quad (5.40)$$

which is often referred to as the *total head* of the fluid. The fact that this is constant along the flow when  $U$  and  $\pi$  are time-independent amounts to *Bernoulli's theorem*. One usually gets to Bernoulli's theorem by starting with Euler's equation (5.25) and assuming that flow is *steady*, which we have implicitly done by assuming that  $U$  and  $\pi$  were time-independent.

**6. Relativistic continuum mechanics.** – One would expect that when one goes from non-relativistic to relativistic continuum mechanics, the mathematics would get more complicated. However, in many respects, it becomes simpler when time gets treated as another dimension. Fortunately, we have been trying to anticipate the transition to relativistic methods all along, so it will not be as necessary to start distinguishing vectors from covectors, which is always an issue if one has treated covariant and contravariant indices as indistinguishable, simply because the use of the Euclidian metric in an

orthonormal frame will not change the values of components when one raises or lowers an index.

One will find that relativistic continuum mechanics is more developed in the context of fluid media than it is in the context of solids. To some extent, that is based in the fact that true “rigidity” cannot exist in relativistic mechanics, so even elastic solids will have some relativistic idiosyncrasies. It is also based in the fact that the main application for relativistic continuum mechanics all along has basically been directed towards developing energy-momentum-stress tensors that would serve as models for the celestial objects that define the sources of gravitational fields in the eyes of Einstein’s equations. However, since the wave equations that we will treat are usually discussed in the context of special relativity, not general relativity, we shall stop short of anything that involves introduction connections, curvature, and Einstein’s equations. (See, however, the book by Vigier [12].) Hence, it will be assumed in all of what follows that the quantum waves propagate in a space where the gravitational force is non-existent.

*a. Minkowski space preliminaries.* – When one goes from non-relativistic motion to relativistic motion, the first thing that changes fundamentally is that the speed of light (*in vacuo*)  $c$  will no longer be infinite, and in fact, it will have to be the same for all observers. As a result of that, the three-dimensional Euclidian metric, whose components are  $\delta_{ij}$  in an orthonormal frame, will have to be extended to the four-dimensional Lorentzian metric:

$$\eta = \eta_{\mu\nu} dx^\mu dx^\nu, \quad (\eta_{\mu\nu} = \text{diag}[+1, -1, -1, -1]). \quad (6.1)$$

We shall denote Minkowski space by  $\mathfrak{M}^4 = (\mathbb{R}^4, \eta)$ . If the components of two tangent vectors to  $\mathfrak{M}^4$  at some point with respect to the natural frame of the coordinates  $x^\mu$  are  $v^\mu$ ,  $w^\mu$ , resp., then their Lorentzian scalar product will be:

$$\eta(\mathbf{v}, \mathbf{w}) = \eta_{\mu\nu} v^\mu w^\nu = v^0 w^0 - v^1 w^1 - v^2 w^2 - v^3 w^3, \quad (6.2)$$

and the Lorentzian square of  $\mathbf{v}$  will be:

$$\|\mathbf{v}\|^2 = (v^0)^2 - (v^1)^2 - (v^2)^2 - (v^3)^2. \quad (6.3)$$

Unlike the Euclidian square, the Lorentzian one does not have to be non-negative. One refers to  $\mathbf{v}$  as *space-like*, *light-like* (or *isotropic*), or *time-like* according to whether  $\|\mathbf{v}\|^2$  is negative, zero, or positive, respectively. One advantage of the present sign convention for the components of  $\eta_{\mu\nu}$  is that the square root of  $\|\mathbf{v}\|^2$  will always be real for time-like vectors. The set of all light-like vectors in any tangent space is referred to as the *light cone* at that point.

A further consequence of the finitude of  $c$  is that one must carefully distinguish between the *proper time* parameter  $\tau$  of a curve  $x(\tau)$  in space-time and the *time coordinate*  $t$  of the points along that curve. It is often convenient to use  $c$  as a units conversion constant in order to define  $x^0 = ct$ , so that one does not have to introduce  $c$  into the Lorentzian metric explicitly. In the latter case, one will have:

$$\eta(\mathbf{v}, \mathbf{w}) = c^2 v^0 w^0 - v^1 w^1 - v^2 w^2 - v^3 w^3. \quad (6.4)$$

Because of the difference between proper time, which presumably describes the time evolution of all processes in one's "rest space" (i.e., everything that is at rest relative to oneself), and the time coordinate, which is more general, there will also be a difference between the spatial components of the four-velocity of the curve  $x(\tau)$  and the three-velocity of the spatial curve  $x(t)$ . The four-velocity that is defined by the proper-time parameter will have components:

$$u^\mu = \frac{dx^\mu}{d\tau} \quad (\mu = 0, \dots, 3), \quad (6.5)$$

while the three-velocity that is defined by the time coordinate will have components:

$$v^i = \frac{dx^i}{dt} \quad (i = 1, 2, 3). \quad (6.6)$$

The way that one gets from  $u^\mu$  to  $v^i$  is by taking advantage of the fact that:

$$u^0 = c \frac{dt}{d\tau}, \quad (6.7)$$

so if one recalls the chain rule for differentiation then:

$$u^i = \frac{dx^i}{d\tau} = \frac{dt}{d\tau} \frac{dx^i}{dt} = \frac{1}{c} u^0 v^i; \quad (6.8)$$

i.e.:

$$v^i = c \frac{u^i}{u^0}. \quad (6.9)$$

Since one can think of  $u^i / u^0$  as the "inhomogeneous coordinates" of a point of  $\mathbb{RP}^3$  that are associated with the "homogeneous coordinates"  $u^\mu$ , that means that the projection from space-time to the rest space has more in common with projective geometry than it does with affine geometry. We mention that only in passing, since we shall not have any cause to refer to that fact, but if one wishes to know more about that approach to special relativity then one can confer the author's discussions of that fact ([13]) and the references that are cited in them.

An identifying characteristic of proper time is that when a time-like space-time curve  $x(\tau)$  has been parameterized by proper-time  $\tau$ , one will always have:

$$\|\mathbf{u}\|^2 = c^2. \quad (6.10)$$

In the rest space of  $\mathbf{u}$  (i.e., all co-moving objects), one will then have  $\mathbf{u} = (c, 0, 0, 0)$ , or  $u^i = 0$ . Hence, the proper-time parameterization is essentially a unit-speed parameterization

when one uses units in which  $c = 1$ . However, light-like curves will not admit proper-time parameterizations, since proper time will always be zero for them.

If we expand (6.10), while using (6.9), then we will get:

$$c^2 = (u^0)^2 - (u^1)^2 - (u^2)^2 - (u^3)^2 = (u^0)^2 \left(1 - \frac{v^2}{c^2}\right),$$

which will make:

$$u^0 = c\gamma, \quad \gamma \equiv \frac{dt}{d\tau} = \left(1 - \frac{v^2}{c^2}\right)^{-1/2}. \quad (6.11)$$

This accounts for the ubiquitous nature of the Fitzgerald-Lorentz factor  $\gamma$  in special relativity.

As a result of (6.10), one must always have:

$$\eta(\mathbf{u}, \boldsymbol{\alpha}) = 0 \quad \left(\boldsymbol{\alpha} \equiv \frac{d\mathbf{u}}{d\tau}\right). \quad (6.12)$$

Hence, the proper velocity will always be orthogonal to the proper acceleration  $\boldsymbol{\alpha}$ .

One can also relate the proper acceleration  $\boldsymbol{\alpha}$  to the non-relativistic acceleration, which is:

$$\mathbf{a} = \frac{d\mathbf{v}}{dt}, \quad (6.13)$$

by differentiating (6.8) with respect to proper time:

$$\alpha^i = \frac{du^i}{d\tau} = \gamma \frac{d(\gamma v^i)}{dt} = \gamma \left( \frac{d\gamma}{dt} v^i + \gamma \frac{dv^i}{dt} \right) = \gamma^2 \left( a^i + \frac{1}{\gamma} \frac{d\gamma}{dt} v^i \right),$$

and if we take into account that:

$$\frac{d\gamma}{dt} = \frac{\gamma^3}{c^2} v_i a^i \quad (6.14)$$

then we can say that:

$$\alpha^i = \gamma^2 \left[ a^i + \left( \frac{\gamma}{c} \right)^2 (v_j a^j) v^i \right], \quad (6.15)$$

So the spatial part of the proper acceleration does not have to be collinear with the non-relativistic acceleration unless the latter is orthogonal to the velocity. Since:

$$v_j a^j = \frac{1}{2} \frac{dv^2}{dt}, \quad (6.16)$$

one can also characterize the latter condition by saying that the motion has constant speed.

The process of splitting the  $\mathbb{R}^4$  of space-time (if only locally, as in general relativity), into a time-plus-space  $\mathbb{R} \oplus \mathbb{R}^3$  is usually intimately linked with choosing a rest space. In particular, if an observer  $x(\tau)$  moves through space-time with a time-like four-velocity  $\mathbf{u}(\tau)$  then one generally defines the *time line*  $\mathbb{R}$  in the decomposition to be the tangent line at each point  $x(\tau)$  that is generated by  $\mathbf{u}(\tau)$  and the *spatial subspace*  $\mathbb{R}^3$  to be the orthogonal complement of the time line.

A time-plus-space splitting of space-time implies a way of partitioning the components of tensors into various combinations of the temporal index 0 and the spatial indices. For instance, a second-rank covariant tensor field  $a$  whose components with respect to a natural coframe are  $a_{\mu\nu}$  can be partitioned into:

$$a = a_{00} dx^0 \otimes dx^0 + a_{0i} dx^0 \otimes dx^i + a_{i0} dx^i \otimes dx^0 + a_{ij} dx^i \otimes dx^j, \quad (6.17)$$

which involves one time-time component  $a_{00}$ , three time-space components  $a_{0i}$ , three space-time components  $a_{i0}$ , and nine space-space components  $a_{ij}$ .

The linear transformations of Minkowski space  $\mathfrak{M}^4$  that preserve the Minkowski scalar product  $\eta$  are called *Lorentz transformations*. Hence, for such a transformation  $L$  one must have:

$$\eta(L(\mathbf{v}), L(\mathbf{w})) = \eta(\mathbf{v}, \mathbf{w}) \quad (6.18)$$

for any two vectors  $\mathbf{v}$  and  $\mathbf{w}$  in  $\mathfrak{M}^4$ .

If one rewrites (6.18) as the matrix equation:

$$\mathbf{v} L^T \eta L \mathbf{w} = \mathbf{v} \eta \mathbf{w}$$

then condition for a  $4 \times 4$  real matrix  $L$  to be a Lorentz transformation is that:

$$L^T \eta L = \eta \quad \text{or} \quad L^{-1} = L^* \equiv \eta L^T \eta. \quad (6.19)$$

Hence, its inverse (which automatically exists) must be equal to its *Lorentz adjoint*  $L^*$ , in analogy with the way that orthogonal matrices have their transposes for their inverses.

If one expresses  $L$  in time+space form then one will see that:

$$\text{if } L = \left[ \begin{array}{c|c} L_0^0 & L_j^0 \\ \hline L_0^i & L_j^i \end{array} \right] \quad \text{then} \quad L^* = \left[ \begin{array}{c|c} L_0^0 & -L_0^i \\ \hline -L_j^0 & L_j^i \end{array} \right].$$

In particular,  $L_0^0$  is not altered, while the spatial submatrix  $L_j^i$  gets transposed. Thus, the spatial submatrices will always be spatial rotations.

However, the Lie group  $O(3, 1)$  of all Lorentz transformations has a dimension of six, while the subgroup  $O(3)$  has a dimension of only three. The remaining three dimensions



are defined by the *pure Lorentz transformations* – or *boosts*. If one thinks of a proper-time hyperboloid:

$$(u^0)^2 - (u^1)^2 - (u^2)^2 - (u^3)^2 = c^2 \tau^2, \quad (6.20)$$

as a one-parameter family of 2-spheres of radius-squared  $(u^1)^2 + (u^2)^2 + (u^3)^2$  then the rotations will preserve that radius while the boosts will change it.

Physically, boosts are the relativistic transformations from one reference frame to another that has a constant relative velocity with respect to the first one. If the relative velocity points in the  $x$ -direction and has a magnitude  $v$  then the corresponding boost transformation will take the form:

$$\bar{t} = \gamma[t + (v/c^2)x], \quad \bar{x} = \gamma[vt + x], \quad (6.21)$$

while the other coordinates  $y, z$  will remain the same.

If one replaces  $t$  with  $x^0 = ct$  and  $\bar{t}$  with  $\bar{x}^0 = c\bar{t}$  and adds a superscript 1 to  $x$  and  $\bar{x}$  then this can be expressed as a matrix equation:

$$\begin{bmatrix} \bar{x}^0 \\ \bar{x}^1 \end{bmatrix} = \begin{bmatrix} \cosh \alpha & \sinh \alpha \\ \sinh \alpha & \cosh \alpha \end{bmatrix} \begin{bmatrix} x^0 \\ x^1 \end{bmatrix}, \quad (6.22)$$

into which, we have introduced the *rapidity parameter*  $\alpha$ , which will make:

$$\cosh \alpha = \gamma, \quad \sinh \alpha = \gamma v / c. \quad (6.23)$$

The subgroup  $SO(3, 1)$  of  $O(3, 1)$  for which all matrices have unity determinant is called the *special Lorentz group*, or sometimes just the *Lorentz group*. Such transformations will preserve the four-dimensional volume  $\mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3$  of a 4-frame  $\{\mathbf{e}_\mu, \mu = 0, \dots, 3\}$ , as well as its orientation. If one identifies  $\mathbf{e}_0$  as the time direction then one can further decompose special Lorentz transformations according to whether they preserve or invert the orientation of the line  $[\mathbf{e}_0]$ . Such transformations are called *orthochronous* in the event that they preserve that orientation. Topologically,  $O(3, 1)$  consists of four connected components depending upon the sign of  $\det L$  and whether  $L$  is orthochronous or not. The identity component shall be denoted by  $SO^+(3, 1)$  and referred to as the *orthochronous Lorentz group*.

*b. Relativistic hydrodynamics.* – The story of relativistic hydrodynamics [14-16] starts with a flow velocity vector field:

$$\mathbf{u}(x) = u^\mu(x) \frac{\partial}{\partial x^\mu} \quad (6.24)$$

that is defined on some region  $\mathcal{R}$  of space-time, and is assumed to be either timelike or lightlike.

Now that we must distinguish proper time  $\tau$ , which is a curve parameter, from  $t$ , the coordinate (which can also serve as a parameter), the business of integrating the system of four ordinary differential equations for the world lines  $x(s)$ :

$$\frac{dx}{ds} = \mathbf{u}(x(s)), \quad \text{i.e.,} \quad \frac{dx^\mu}{ds} = u^\mu(x(s)), \quad (6.25)$$

will become more subtle depending upon whether the parameter  $s$  is equal to  $\tau$  or  $t$ .

For one thing, if the vector field  $\mathbf{u}(x)$  is light-like (as one would expect for a photon gas), there will be no proper time parameter, and one will have to use an affine parameter. Similarly, in time-like cases, one might have to consider the integral curves in both proper time and time coordinate parameterizations.

One thing that simplifies matters is that since the basic configuration manifold  $\mathfrak{M}^4$  includes time already, in effect, all velocity vector fields will be time-varying (i.e., they will be steady only with respect to some reference frames). Hence, one no longer needs to distinguish between path-lines, which are integral curves of time-varying flow velocities, and the streamlines, which pertain to steady flow. Following Lichnerowicz [15], we agree to call all integral curves *streamlines* generically.

The covelocity 1-form:

$$u = u_\mu dx^\mu \quad (6.26)$$

is associated with  $\mathbf{u}$  is associated by way of the Minkowski scalar product:

$$u = i_{\mathbf{u}}\eta \quad (u_\mu = \eta_{\mu\nu} u^\nu). \quad (6.27)$$

Hence, one must be careful to include appropriate signs in its components:

$$u_0 = u^0, = \beta c, \quad u_i = \beta v_i = -u^i = -\beta v^i. \quad (6.28)$$

One then has:

$$u(\mathbf{u}) = \|\mathbf{u}\|^2 = \|u\|^2, \quad (6.29)$$

and for proper-time parameterization:

$$u(\mathbf{u}) = c^2. \quad (6.30)$$

In (6.29), the definition of  $\|\mathbf{u}\|^2$  comes from the Minkowski space structure on the cotangent spaces to  $\mathbb{R}^4$ , which is defined by the inverse matrix  $\eta^{\mu\nu}$  to  $\eta_{\mu\nu}$ .

The rate of deformation  $du$  has more non-zero components now, since one must consider  $v^0$  to be a function of  $x^\mu$ , and not just 1, as it was before:

$$du = \dot{e} + \Omega_k, \quad (6.31)$$

in which:

$$\dot{e} = \frac{1}{2} \dot{e}_{\mu\nu} dx^\mu dx^\nu, \quad \dot{e}_{\mu\nu} \equiv u_{\mu,\nu} + u_{\nu,\mu} \quad (6.32)$$

is the symmetric *rate of strain* tensor field, while:

$$\Omega_k = \frac{1}{2} \Omega_{\mu\nu} dx^\mu \wedge dx^\nu, \quad \Omega_{\mu\nu} \equiv u_{\mu,\nu} - u_{\nu,\mu} \quad (6.33)$$

is the anti-symmetric *kinematical vorticity* 2-form, which satisfies:

$$\Omega_k = d \wedge u. \quad (6.34)$$

If one splits space-time according to  $x^0$  and  $x^i$ ,  $i = 1, 2, 3$  then the two tensor fields will decompose into:

$$\dot{e} = u_{0,0} (dx^0)^2 + \frac{1}{2} (u_{0,i} + u_{i,0}) dx^0 dx^i + \frac{1}{2} (u_{i,j} + u_{j,i}) dx^i dx^j, \quad (6.35)$$

$$\Omega_k = (u_{0,i} - u_{i,0}) dx^0 \wedge dx^i + \frac{1}{2} (u_{i,j} - u_{j,i}) dx^i \wedge dx^j. \quad (6.36)$$

Note that the components  $u_{0,0}$  and  $u_{0,i}$  would vanish in non-relativistic mechanics, for which  $\gamma$  would be a constant, namely, 1. However, if one sets  $u_0 = c \gamma(x^t)$ , as in (6.11), then:

$$u_{0,0} = \frac{\partial u_0}{\partial x^0} = \frac{\partial \gamma}{\partial t} = \frac{\gamma^3}{2c^2} \frac{\partial v^2}{\partial t}, \quad (6.37)$$

$$u_{0,i} = \frac{\partial u_0}{\partial x^i} = c \frac{\partial \gamma}{\partial x^i} = \frac{\gamma^3}{2c} \frac{\partial v^2}{\partial x^i}. \quad (6.38)$$

Hence,  $u_{0,0}$  will vanish when the spatial speed of flow is constant in time, while  $u_{0,i}$  will vanish when it is constant in space.

Since one also has:

$$u_{i,0} = \frac{\partial u_i}{\partial x^0} = \frac{1}{c} \frac{\partial(\gamma v_i)}{\partial t} = \frac{\gamma}{c} \left[ \frac{\partial v_i}{\partial t} + \frac{1}{2} \left( \frac{\gamma}{c} \right)^2 \frac{\partial v^2}{\partial t} \right], \quad (6.39)$$

one can now compute the time-space components in both cases:

$$u_{0,i} + u_{i,0} = \frac{\gamma}{c} \left[ \frac{\partial v_i}{\partial t} + \frac{1}{2} \gamma^2 \left( \frac{1}{c^2} \frac{\partial v^2}{\partial t} v_i + \frac{\partial v^2}{\partial x^i} \right) \right], \quad (6.40)$$

$$u_{0,i} - u_{i,0} = - \frac{\gamma}{c} \left[ \frac{\partial v_i}{\partial t} - \frac{1}{2} \gamma^2 \left( \frac{1}{c^2} \frac{\partial v^2}{\partial t} v_i - \frac{\partial v^2}{\partial x^i} \right) \right]. \quad (6.41)$$

The interpretation of these time-space components is not entirely obvious, beyond the fact that they involve the non-relativistic spatial acceleration covector field  $\partial v_i / \partial t$ , along with corrections to it that originate in the possibility that the speed  $v$  of the non-relativistic velocity vector might vary in both time and space.

One can then express the spatial gradient  $u_{i,j}$  of the relativistic (i.e., proper-time) velocity in terms of the spatial gradient  $v_{i,j}$  of the non-relativistic velocity by the equation:

$$u_{i,j} = \frac{\partial(\gamma v_i)}{\partial x^i} = \frac{\partial \gamma}{\partial x^i} v_i + \gamma \frac{\partial v_i}{\partial x^i} = \gamma \left( v_{i,j} + \frac{\gamma^2}{2c^2} v_i \frac{\partial v^2}{\partial x^j} \right). \quad (6.42)$$

As  $v$  goes to 0,  $u_{i,j}$  will go to  $v_{i,j} + \frac{1}{2c^2} v_i \frac{\partial v^2}{\partial x^j}$ , while  $u_{i,j}$  will become infinite as  $v$  approaches  $c$ . Hence, once again, the possibility that the speed of the velocity vector field will vary in space has changed the nature of the non-relativistic limit of the relativistic expression for the covelocity gradient.

When one symmetrizes  $u_{i,j}$ , the spatial components of the relativistic rate of strain will relate to the non-relativistic components by:

$$u_{i,j} + u_{j,i} = \gamma \dot{e}_{ij} + \frac{\gamma^3}{2c^2} \left( v_i \frac{\partial v^2}{\partial x^j} + v_j \frac{\partial v^2}{\partial x^i} \right), \quad (6.43)$$

while the difference between the spatial components of the relativistic kinematical vorticity and  $\beta$  times the non-relativistic components will be:

$$u_{i,j} - u_{j,i} = \gamma \omega_{ij} + \frac{\gamma^3}{2c^2} \left( v_i \frac{\partial v^2}{\partial x^j} - v_j \frac{\partial v^2}{\partial x^i} \right), \quad (6.44)$$

which can be put into the form:

$$\Omega_k = \gamma \left[ \omega + \frac{1}{2} \left( \frac{\gamma}{c} \right)^2 v \wedge dv^2 \right]. \quad (6.45)$$

The second term in the brackets will vanish when the gradient of the speed-squared is collinear with the covelocity.

Once again, there are two types of  $\dot{e}_{\mu\nu}$ : ones for which the matrix  $\dot{e}_\nu^\mu = \eta^{\mu\kappa} \dot{e}_{\kappa\nu}$  has trace zero and one for which that trace is non-zero. One can then decompose  $\dot{e}_\nu^\mu$  into:

$$\dot{e}_\nu^\mu = \widehat{e}_\nu^\mu + \frac{1}{4} \dot{e}_\kappa^\kappa \delta_\nu^\mu, \quad \widehat{e}_\nu^\mu \equiv \dot{e}_\nu^\mu - \frac{1}{4} \dot{e}_\kappa^\kappa \delta_\nu^\mu, \quad (6.46)$$

in which  $\widehat{e}_\nu^\mu$  is traceless. Once again, this decomposition is not unique, since one can add any traceless matrix to  $\widehat{e}_\nu^\mu$ ,

One finds that:

$$\lambda_k \equiv \dot{e}_\mu^\mu = \frac{\partial u^\mu}{\partial x^\mu} = \frac{\partial \gamma}{\partial t} + \gamma \frac{\partial v^i}{\partial x^i} + \frac{\partial \gamma}{\partial x^i} v^i = \gamma \left[ \frac{\partial v^i}{\partial x^i} + \frac{1}{2} \left( \frac{\gamma}{c} \right)^2 \frac{dv^2}{dt} \right], \quad (6.47)$$

which one then regards as the *relativistic compressibility* of the flow. Its vanishing would not be equivalent to non-relativistic incompressibility, since one would need to have:

$$\frac{\partial v^i}{\partial x^i} = -\frac{1}{2} \left( \frac{\gamma}{c} \right)^2 \frac{dv^2}{dt}. \quad (6.48)$$

Hence, this would vanish iff the speed of the flow were constant along the integral curves of  $\mathbf{v}$ .

One must realize that non-relativistic incompressibility would imply that the speed of propagation of sound waves in the fluid would have to be infinite, so relativistic incompressibility must reflect the fact that there is an upper bound on that speed of propagation that is imposed by  $c$ .

When one introduces a mass density  $\rho$  into the relativistic context, one must recognize that it is not a Lorentz-invariant object, but only one-fourth of a Lorentz-invariant object, namely, the energy-momentum density 1-form:

$$p = \rho_0 u. \quad (6.49)$$

In this expression, we have distinguished the expression that  $\rho$  takes in any rest frame, namely, the *rest mass density*  $\rho_0$ . Such a Lorentzian frame will have the property that  $u^i = 0$  (so  $u = c dt$  and  $p = \rho_0 c dt$ ). Hence, we can say that:

$$p = (\rho_0 \gamma c, \rho_0 \gamma v_i) = (\rho c, \rho v_i), \quad (6.50)$$

in which:

$$\rho = \rho_0 \gamma \quad (6.51)$$

will then become the *relative mass density* for the Lorentzian frame whose relative velocity is  $v^i$ .

One can also characterize the temporal component of  $p$  as  $1/c$  times the *energy density*  $E$  of the motion:

$$c p_0 = E = \rho c^2 = \gamma \rho_0 c^2 = \gamma E_0, \quad (6.52)$$

which will then make:

$$E_0 = \rho_0 c^2 \quad (6.53)$$

the *rest energy density* of that motion.

One then has:

$$\| p \|^2 = \rho_0^2 \| u \|^2 = \rho_0^2 c^2. \quad (6.54)$$

Since light waves are associated with an energy density and a momentum density, but a vanishing rest mass density, one cannot use the definition (6.49) for the energy-momentum density 1-form, but must start with  $p = (E/c, p_i)$  as the definition, which will have the property that:

$$\| p \|^2 = 0. \quad (6.55)$$

Later, we shall discuss the way that such an energy-momentum density is associated with a kinematical state, in the form of the “frequency-wave number” 1-form, rather than the covelocity.

The exterior derivative of  $p$  is called the *dynamical vorticity*  $\Omega_d$ , and since:

$$\Omega_d = d\wedge p = d\rho_0 \wedge u + \rho_0 d\wedge u = d\rho_0 \wedge u + \rho_0 \Omega_k,$$

the dynamical vorticity will be  $\rho_0$  times the kinematical vorticity iff  $d\rho_0$  is collinear with  $u$ .

If  $\mathbf{p} = \rho_0 \mathbf{u}$  is the energy-momentum density vector field then since:

$$\operatorname{div} \mathbf{p} = \frac{1}{c} \frac{\partial(\rho_0 u^0)}{\partial t} + \frac{\partial(\rho_0 u^i)}{\partial x^i} = \frac{\partial(\rho_0 \beta)}{\partial t} + \frac{\partial(\rho_0 \beta v^i)}{\partial x^i} = \frac{\partial \rho}{\partial t} + \frac{\partial(\rho v^i)}{\partial x^i},$$

one will see that this differs from the non-relativistic expression (5.11) for conservation of mass by only a correction to the mass density. Hence:

$$\operatorname{div} \mathbf{p} = 0 \tag{6.56}$$

can serve as the relativistic definition of conservation of mass.

$\mathbf{p}$  also defines the relativistic dynamical compressibility by its four-divergence. Since:

$$\operatorname{div} \mathbf{p} = d\rho_0(\mathbf{u}) + \rho_0 \operatorname{div} \mathbf{u} = \frac{d\rho_0}{d\tau} + \rho_0 \operatorname{div} \mathbf{u}$$

the dynamical compressibility will be equal to  $\rho_0$  times the kinematical compressibility iff  $\rho_0$  is constant along the flow of  $\mathbf{u}$ .

Of course, since mass is not a Lorentz-invariant concept, one usually expands the conservation of linear momentum by embedding the energy-momentum density into the energy-momentum-stress tensor  $T_v^\mu$ , whose doubly-covariant form  $T_{\mu\nu}$  is generally assumed to be symmetric in relativistic hydrodynamics. The components of both decompose as:

$$T_v^\mu = \left[ \begin{array}{c|c} E & cp_j \\ \hline cp^i & \sigma_j^i \end{array} \right], \quad T_{\mu\nu} = \left[ \begin{array}{c|c} E & cp_i \\ \hline cp_i & \sigma_{ij} \end{array} \right], \tag{6.57}$$

in which  $\sigma_j^i$  is the (relative) stress tensor, and one must note the sign discrepancies:

$$p^i = -p_i, \quad \sigma_{ij} = -\sigma_j^i. \tag{6.58}$$

Hence:

$$T_{\nu}^{\mu} = \left[ \begin{array}{c|c} E & cp_i \\ \hline -cp_i & -\sigma_{ij} \end{array} \right]. \quad (6.59)$$

Hence, even though  $T_{\mu\nu}$  is assumed to be symmetric,  $T_{\nu}^{\mu}$  will generally be asymmetric. One should notice that all components have the same basic units of  $M / (LT^2)$ , since pressure has units of force per unit area.

One then sees that the trace of  $T_{\nu}^{\mu}$  will take the form:

$$T_{\mu}^{\mu} = E + \sigma_i^i. \quad (6.60)$$

It will then represent the sum of the energy density and the three times the mean pressure.

Since  $T_{\nu}^{\mu}$  is generally asymmetric symmetric, the eigenvalues of  $T_{\nu}^{\mu}$  will not necessarily be real. One can also pose a generalized eigenvalue problem for  $T_{\nu}^{\mu}$  in the form:

$$T_{\mu\nu} X^{\nu} = \lambda \eta_{\mu\nu} X^{\nu} = \lambda X_{\mu}. \quad (6.61)$$

If one assumes that the eigenvalues are all real then there will be a *principal frame* in which  $T_{\nu}^{\mu}$  is diagonal. A common restriction on  $T_{\nu}^{\mu}$  is that it must have at least one time-like eigenvector. One then calls such a  $T_{\nu}^{\mu}$  *normal* [15].

The divergence of  $T_{\nu}^{\mu}$  has temporal and spatial components that are equal to:

$$\partial_{\mu} T_0^{\mu} = \frac{1}{c} \frac{\partial E}{\partial t} + c \frac{\partial p^i}{\partial x^i} = c \left( \frac{\partial \rho}{\partial t} + \frac{\partial p^i}{\partial x^i} \right) = c \frac{d\rho}{dt}, \quad \partial_{\mu} T_j^{\mu} = \frac{\partial p_j}{\partial t} + \frac{\partial \sigma_j^i}{\partial x^i}. \quad (6.62)$$

Hence, its vanishing:

$$\partial_{\mu} T_{\nu}^{\mu} = 0 \quad (6.63)$$

will yield two sets of conservation laws:

$$\frac{d\rho}{dt} = 0, \quad \frac{\partial p^i}{\partial t} = -\partial_i \sigma_j^i. \quad (6.64)$$

From the first equation in (6.62), one can also write the first conservation law as the conservation of energy in the form:

$$\frac{\partial E}{\partial t} + c^2 \frac{\partial p^i}{\partial x^i} = 0. \quad (6.65)$$

In the case of external force densities  $f_\nu$  that act upon the fluid (e.g., electromagnetic forces that act upon a charged fluid), instead of conservation laws, one will get balance principles. Namely:

$$\partial_\mu T_\nu^\mu = f_\nu, \quad (6.66)$$

which yields:

$$\frac{d\rho}{dt} = \frac{1}{c} f_0, \quad \frac{\partial p^i}{\partial t} = f_j - \partial_i \sigma_j^i. \quad (6.67)$$

Hence,  $f_0 / c$  must represent the source of the net flux of mass through any region of space-time. However, it is more conventional to think of the temporal component of  $f$  as representing a power transfer density, which would be consistent with the elementary association of  $f$  with  $dp / dt$ . One can then make the first equation in (6.67) consistent with that definition by multiplying both sides by  $c^2$ :

$$\frac{dE}{dt} = f_0 c. \quad (6.68)$$

The general form of  $T_\nu^\mu$  for the purposes of relativistic hydrodynamics (see Lichnerowicz [15] or Halbwachs [17]) is:

$$T_\nu^\mu = p_\nu u^\mu + \theta_\nu^\mu. \quad (6.69)$$

The first term is essentially a kinetic energy term, while the second one represents the contribution of internal stresses that are due to the interaction of the constituent molecules with each other, along with the external forces. When there is no transverse momentum present:

$$p_\nu u^\mu = \rho u_\nu u^\mu. \quad (6.70)$$

Otherwise, the kinetic term can be asymmetric in its doubly covariant form.

The most common definitions that we shall use in what follows are:

1. Pure matter (i.e., dust cloud):  $\theta_\nu^\mu = 0$ .

In such a case, there are no mutual interactions, and therefore, no internal stresses. The divergence of  $T_\nu^\mu = u^\mu p_\nu$  will take the form:

$$\partial_\mu T_\nu^\mu = (\partial_\mu u^\mu) p_\nu + u^\mu \partial_\mu p_\nu = \lambda_k p_\nu + \frac{dp_\nu}{d\tau}, \quad (6.71)$$

which differs from the proper-time derivative of the energy-momentum density by a contribution from the kinematical compressibility. Hence, for a kinematically incompressible fluid, they will be equal.



2. Perfect (i.e., inviscid) fluid:  $\theta_v^\mu = \pi \delta_v^\mu$ .

Since  $\partial_\mu \theta_v^\mu = \partial_v \pi$ , one sees that the relativistic Euler equation (i.e.,  $\partial_\mu T_v^\mu = f_v$ ) will take the form:

$$\frac{dp_v}{d\tau} = f_v - \partial_v \pi + \lambda_k p_v. \quad (6.72)$$

Note, however, that the time derivative has become the Lie derivative, which is consistent with the fact that one usually treats hydrodynamics in the Eulerian picture, which means following the evolution of the hydrodynamical state variables along the streamlines.

3. Charged, perfect fluid:  $\theta_v^\mu = (\text{perfect fluid}) + \tau_v^\mu$ , where:

$$\tau_v^\mu = F_\kappa^\mu F_v^\kappa - F^2 \delta_v^\mu \quad (6.73)$$

is the *Faraday tensor* for the electromagnetic field strength 2-form  $F$  that represents the external field. Its divergence will be the Lorentz force that acts upon the charge and electric current that  $\mathbf{J} = \sigma \mathbf{u}$  represents when  $\sigma$  is the charge density:

$$\partial_v \tau_\mu^v = F_{\mu\nu} J^\nu, \quad (6.74)$$

which will get added from the right-hand side of (6.72);  $\partial_v \tau_\mu^v$  might also substitute for  $f_\mu$  if there are no other external forces that act upon the fluid.

Although one can also introduce viscosity by making  $\theta_{\mu\nu}$  proportional to the rate of strain  $\dot{e}_{\mu\nu}$ , we will not be dealing with viscous fluids in what follows, except to compare the form of the “quantum stress tensor” that will emerge later to some of the standard forms that we discussed here.

It should be noted that the state of the relativistic fluid is determined by six variables, namely  $(\rho, \pi, u^\mu)$ . There are four equations in (6.72), and if one makes a definite statement about compressibility then there will be five, which will still leave the system underdetermined by one equation. The most common way of completing the set is to add an *equation of state*, which might take the form:

$$\rho = \rho(\pi). \quad (6.75)$$

A common choice is based upon the ideal gas relation:

$$\pi V = nRT \quad \text{or} \quad \pi \propto \rho = \frac{nM}{V}, \quad (6.76)$$

in which  $V$  is the volume of the gas,  $n$  is the number of moles present ( $M$  = molecular weight of gas molecule),  $R$  is a universal constant, and  $T$  is the absolute temperature.

When a fluid has been given an equation of state, one refers to it as *barotropic*.

*c. Balance of relativistic angular momentum.* – Since the main physical application of the theory of relativistic angular momentum for deformable continuous media is to the Dirac electron, we shall defer a discussion of the conservation of relativistic angular momentum to a later chapter in which that topic will become necessary.

*d. Relativistic solid mechanics.* – If one thinks of the solid state of matter as something that includes elastic, plastic, and rigid matter then one will see why the relativistic theory of solid mechanics is less developed than relativistic fluid mechanics. For one thing, the very notion of the deformation of a solid object already leads to certain relativistic subtleties, such as the idea that a solid sphere will appear to deform into an oblate spheroid as its velocity relative to an observer increases, despite the fact that it will continue to appear to be spherical in its rest space; i.e., to comoving observers.

Hence, we shall only attempt to discuss some of the issues that are associated with the transition from non-relativistic to relativistic solid mechanics, and point to relevant references in which they might have been discussed.

1. *Causality.* – Due to the universal speed limit of  $c$ , not all deformations will be physically realizable. For instance, if one looks at a proper-time-parameterized family of diffeomorphisms  $\phi_\tau: \mathcal{R}_0 \rightarrow \mathcal{M}^4$ ,  $x \mapsto \phi_\tau(x)$ , ( $0 \leq \tau$ ) of an initial spatial region  $\mathcal{R}_0$  then the curves  $\phi_\tau(x_0)$  that are defined by the evolution of each point  $x_0 \in \mathcal{R}_0$  cannot have velocity vectors that become space-like for any  $\tau$ ; if the region  $\mathcal{R}_0$  is massive then they cannot become light-like either. Hence, one must distinguish causal deformations from acausal ones.

Rigidity is to the solid state what incompressibility is to the fluid state in the eyes of relativity. In both cases, the issue is the infinite speed of propagation of waves, which cannot exceed  $c$  in reality. Hence, it has been suggested by Brota and Fernandes [18] that what Born [19] was calling a “rigid” body should have been called an *undeformable* one.

2. *Polar decomposition.* – The fact that an element of  $GL^+(3)$  – viz., an invertible real  $3 \times 3$  matrix with positive determinant – can be expressed uniquely as a product of a non-zero multiple of the identity matrix, a volume-preserving strain, and an oriented Euclidian rotation is due to the fact that topologically the manifold of the Lie group  $GL^+(3)$  is the product of a six-dimensional vector space and the group manifold of  $SO(3)$ . Furthermore, the algorithm by which one usually obtains that decomposition is basically the Gram-Schmidt orthonormalization process. At the level of infinitesimal transformations, one finds that the corresponding decomposition of the vector space that underlies the Lie algebra  $\mathfrak{gl}(3)$  into  $\mathfrak{so}(3) \oplus \mathbb{R}^6$  is nothing but the polarization of a square matrix into the sum of an antisymmetric matrix and a symmetric one by using the matrix transpose operator.

When one goes from three dimensions to four, the corresponding orthonormalization process will make  $GL^+(4)$  diffeomorphic to  $\mathbb{R}^{10} \times SO(4)$ . However, the relativistically

interesting subgroup of  $GL^+(4)$  is  $SO(3, 1)$ , not  $SO(4)$ , so the question becomes that of whether one can decompose a matrix in  $GL^+(4)$  into the product of a non-zero multiple of the identity matrix, a volume preserving *Lorentz strain*, and a *Lorentz transformation*. The problem with finite transformations is that Gram-Schmidt will break down whenever one encounters a light-like frame vector in the algorithm, since normalization would then involve division by zero.

One finds that the process is more elementary to describe in the context of the Lie algebras. Basically, one uses the Lorentz adjoint operator  $*$ :  $\mathfrak{gl}(4) \rightarrow \mathfrak{gl}(4)$  that was defined above, and which takes the matrix  $M$  to:

$$M^* = \eta M^T \eta, \quad (6.77)$$

in which  $\eta = \text{diag}[+1, -1, -1, -1]$  is the matrix of the Minkowski scalar product in an orthonormal frame.

If one looks at a differentiable curve  $M(s)$  through the identity matrix  $I = M(0)$  and differentiates it at the identity then the condition  $MM^* = M^*M = I$  will give:

$$m + m^* = 0 \quad m = \left. \frac{dM(s)}{ds} \right|_{s=0}, \quad m^* = \left. \frac{dM^*(s)}{ds} \right|_{s=0} = (m)^*, \quad (6.78)$$

when  $m$  belongs to  $\mathfrak{so}(3, 1)$ . Hence, if one polarizes a general  $m \in \mathfrak{gl}(4)$  by using  $*$  then one will decompose it into a sum:

$$m = l + e, \quad l = \frac{1}{2}(m - m^*), \quad e = \frac{1}{2}(m + m^*), \quad (6.79)$$

in which  $l$  will become an infinitesimal Lorentz transformation, and  $e$  will become what we shall call an *infinitesimal Lorentz strain*.

Although the trace of  $l$  will vanish, since the diagonal elements of real, antisymmetric matrices will always vanish, the same cannot be said of  $e$ . Hence, one can further decompose  $e$  into a traceless part and a trace part in the usual way:

$$e = \overset{\circ}{e} + \frac{1}{4} \text{Tr}(e) I, \quad \overset{\circ}{e} = e - \frac{1}{4} \text{Tr}(e) I. \quad (6.80)$$

One can obtain a corresponding decomposition of a matrix  $M$  in  $GL^+(4)$  into a product of a determinant, a finite Lorentz strain, and a Lorentz transformation in a neighborhood of  $I$  by exponentiation, although details become rapidly involved, due to the non-Abelian nature of the group.

3. *Relativistic constitutive laws.* – As one can imagine, the way that the basic material properties are affected by relativistic considerations can also get rapidly involved. For instance, the basic Hooke law association of stress with strain will carry with it the unbounded velocity of oscillation of simple harmonic oscillators in the rest space of the oscillator. To an observer, if one is dealing with an undamped mass-spring system with a Hookean spring, so the natural frequency and maximum velocity will be:

$$\omega_h = \sqrt{\frac{k}{m}}, \quad v_{\max} = A \omega_h \quad (A = \text{amplitude})$$

as the velocity of oscillation increases, so will the apparent mass, which will reduce the natural frequency. Moreover, the apparent amplitude will be shortened with increasing relative velocity. Of course, to be rigorous, one should really just re-pose the problem of the relativistic harmonic oscillator, but the point of the example was only to show that the simple linear constitutive law would be somewhat naïve in the relativistic context.

One of the earliest attempts to introduce a relativistic constitutive law was made in 1959 by Synge [20], who assumed that one could still deal with a Hookean law if one went to the coupling of rate of stress to rate of strain. A few years later, in 1963, Rayner [21] discussed the Hookean coupling of stress to strain. In 1970, Oldroyd [22] introduced a “tensor of physical constants” for a material, which included, elasticity, viscosity, and visco-elasticity within its scope. Carter and Quintana’s 1973 general-relativistic formulation of perfectly-elastic bodies [23] asserted that Rayner’s constitutive law would be sufficient for low-pressure applications, such as the interaction of gravitational waves with planetary bodies, such as the Earth, but for high-pressure applications, such as deformations and vibrations in the crusts of neutron stars, one will need to correct that law.

More recently (2004), Hermann, et al [24] examined the constitutive properties of relativistic spin fluids, such as the Dirac electron and the Weysenhoff fluid, which is a simplification of the latter. We shall discuss these examples in our last chapter on “hydrodynamical” models for the Dirac equation.

4. *The relativistic propagation of waves in elastic media.* – We have already mentioned the ultimate bound of  $c$  on the speed of propagation of waves. In many of the discussions of relativistic constitutive laws (in particular, Synge [20] and Rayner [21]), the test of whether the law would be useful in a relativistic context was to look at the propagation of shock waves and see how restricting the maximum speed would constrain the constitutive law itself. Synge looked at the Cauchy problem for Einstein’s equations, when the energy-momentum-stress tensor was suitably constrained, while Rayner used Hadamard’s approach to the propagation of wave that involved disturbances being defined across initial discontinuity hypersurfaces. Israel [25] also examined how conditions in the constitutive properties might affect the propagation of shock waves.

5. *Variational formulation.* – The earliest known (special) relativistic treatment of continuum mechanics was by Herglotz in 1911 [26]. In that article, he derived the relativistic equations of motion/equilibrium from an action principle that was defined by a “kinetic potential.”

In 1964, Schöpf [27] gave a general-relativistic treatment of conservative systems. For such systems, one had an elastic potential, but no heat current.

Since we will devote an entire chapter to variational methods, we shall content ourselves for the moment with that brief historical synopsis.

**§ 7. The principle of virtual work.** – Before we get to the principle of least action in the context of the variational field theory, we shall first discuss a first principle of mechanics that is more general in scope, since it also applies to forces that do not admit potential functions and motion with non-holonomic constraints. We are referring to the principle of virtual work.

*a. Virtual displacements.* – If we return to the definition of a finite deformation of an initial state  $x_0$  of a  $k$ -dimensional object in space into a final state  $x_1$  as a  $k+1$ -object then the definition of an infinitesimal deformation will become not only straightforward, but consistent with a good way of defining “variations” for the purposes of the calculus of variations. Namely, if the “time” dimension is defined by  $u^0$  then one can define a congruence of differentiable curves  $x(u^0) = x(u^0, u_0^1, \dots, u_0^k)$  from each point  $x(0, u_0^1, \dots, u_0^k)$  of  $x_0$  to its corresponding point  $x(1, u_1^1, \dots, u_1^k)$  of  $x_1$ . The partial derivative  $\partial x^i / \partial u^0$  will then define a field of velocity vectors on the image of  $x$ , and we define the *virtual displacement* (AKA: *variation*) of the initial state  $x_0$  to be the restriction of that vector field to the image of  $x_0$ :

$$\delta x^i = \left. \frac{\partial x^i(u^0, \dots, u^k)}{\partial u^0} \right|_{u^0=0}. \quad (7.1)$$

Hence, this is simply the restriction of the velocity vector field that was defined in (1.31) to the initial configuration ( $s = 0$ ).

The sense in which these displacements are “virtual” is that since vector fields represent the infinitesimal generators of one-parameter families of finite displacements, until one actually integrates them into such families, in a sense, they are only latent. That is, they only tell one about the other deformations that in exist in a “sufficiently-small neighborhood” of the initial configuration of an object.

The concept of displacement can include rotations, as well as translations. More generally, one has a Lie group  $G$  acting upon a region  $\mathcal{R}_0$  of space, so a one-parameter family of displacements  $x(u^0, u_0^1, \dots, u_0^k)$  can also be generated by starting with a differentiable curve  $g(u^0)$  in  $G$  and letting it act upon some initial configuration  $x_0(u_0^1, \dots, u_0^k)$ . For instance, if  $G$  is represented by  $n \times n$  matrices that act upon the coordinates  $x_0^i$  of the initial configuration then one might define:

$$x^i(u^0, u_0^1, \dots, u_0^k) = g_j^i(u^0) x_0^j(u_0^1, \dots, u_0^k). \quad (7.2)$$

In order to turn such an action into a virtual displacement, one differentiates  $g(u^0)$  at  $u^0 = 0$  to get a tangent vector  $\omega$  to  $G$  at  $g(0)$ , and obtains a vector field on  $x_0(u_0^1, \dots, u_0^k)$ :

$$\delta x^i(u_0^1, \dots, u_0^k) = \omega_j^i x_0^j(u_0^1, \dots, u_0^k), \quad (7.3)$$

which is the restriction of the fundamental vector field that is associated with  $\omega$  to the initial configuration.

*b. Virtual work.* – The dual object to a displacement is a force in the case of translations and a torque (or moment) in the case of rotations. As long as one deals with infinitesimal displacements, one can use the canonical bilinear pairing of a Lie algebra  $\mathfrak{g}$  with its dual vector space  $\mathfrak{g}^*$ , namely:

$$\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \rightarrow \mathbb{R}, (\alpha, \mathbf{a}) \mapsto \langle \alpha, \mathbf{a} \rangle,$$

in which:

$$\langle \alpha, \mathbf{a} \rangle \equiv \alpha(\mathbf{a}) = \alpha_i a^i, \quad (7.4)$$

to define a bilinear pairing of forces  $F$  with virtual translations  $\delta x$  and torques  $M$  with virtual rotations  $\delta\theta$  that gives the *virtual work* that is done by the force (torque, resp.) on the virtual displacement:

$$\delta W = F(\delta x) = F_i \delta x^i, \quad \delta W = M(\delta\theta) = M_i \delta\theta^i. \quad (7.5)$$

As long as elements of both  $\mathfrak{g}$  and  $\mathfrak{g}^*$  are represented by  $n \times n$  matrices, one can also represent this bilinear pairing by using the *Cartan-Killing* form on  $\mathfrak{g}$ :

$$\langle \mathbf{a}, \mathbf{b} \rangle = \text{Tr } \mathbf{a} \mathbf{b} = a^i b_i^j. \quad (7.6)$$

(When  $\mathbf{a}$  and  $\mathbf{b}$  are not represented by matrices, one replaces them with their adjoint representations.)

As mentioned in Appendix C, for semi-simple Lie algebras, such as  $SO(3)$  and  $SO(3, 1)$ , this bilinear functional will define a scalar product on  $\mathfrak{g}$ ; i.e., it will be symmetric and non-degenerate. For  $SO(3)$ , it will be minus the Euclidian scalar product, while for  $SO(3, 1)$ , it will have the signature type of  $(-1, -1, -1, +1, +1, +1)$ .

We shall return to the study of virtual work in the context of the calculus of variations, since that is really the proper place for such considerations. For now, we simply state *d'Alembert's principle* that when a system is in an equilibrium state, the virtual work that is done by any “allowable” virtual displacement must vanish. One can also use this as the basis for its time evolution if one includes “kinetic forces,” which essentially amount to contributions of the form  $ma$ .

**8. The point-particle approximation.** – Since the statistical interpretation of quantum wave functions (or rather, their moduli-squared) comes down to the probability of finding a *point-like* particle within the differential volume element that is located at that point, for the sake of completeness, it is necessary to discuss the way that one goes from the motion of extended, deformable matter to the motion of point-like matter.

The first step is that of replacing the mass density function  $\rho(t, x)$  with an equivalent set of moments by using the *moment theorem* for a distribution. Although that theorem is discussed most frequently in the context of probability density functions, nonetheless, the same theorem is just as applicable to densities of mass, charge, spin, etc. It can be

expressed in  $n$ -dimensional Euclidian space in full generality, although that involves introducing the formalism of multi-indices in order to make the notation manageable, so we shall show how it works in the case of a one-dimensional real vector space, since the essential details are already contained in that demonstration. We shall also omit the time dimension as something of a distraction at this point, but that is really no loss in generality when one treats time as a dimension, anyway.

Suppose that the density function  $\rho(x)$  admits a Fourier transform:

$$\hat{\rho}(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ikx} \rho(x) dx. \quad (8.1)$$

Expand the exponential in a Taylor series:

$$e^{-ikx} = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} (-ikx)^n = \sum_{n=1}^{\infty} \frac{1}{n!} (-ik)^n x^n. \quad (8.2)$$

Assuming that  $\rho$  is also “analytically convenient” for one to invert the order of summation and integration, one will get:

$$\hat{\rho}(k) = \frac{1}{2\pi} \left[ M_0 + \sum_{n=1}^{\infty} \frac{1}{n!} M_n (-ik)^n \right], \quad (8.3)$$

in which we have introduced the successive *moments* of  $\rho$ :

$$M_n = \int_{-\infty}^{+\infty} \rho(x) x^n dx. \quad (8.4)$$

The moment theorem amounts to the fact that if one is given a complete set of moments then one can reconstruct  $\rho$  by taking the inverse Fourier transform of  $\hat{\rho}(k)$ .

In particular,  $M_0$  is the *total mass* of  $\rho$  (assuming that  $\rho$  is a mass density). For a probability density function,  $M_0$  would always be equal to 1, since that would be the probability that the statement is true in at least some case; e.g., in quantum mechanics, it is the probability that the point-particle exists somewhere in space.

The first moment is essentially a *mean position* of  $\rho$  when  $\rho$  has been normalized to have unit mass. For an arbitrary mass distribution, the *center-of-mass* (COM) becomes:

$$x_{\text{cm}} = \frac{M_1}{M_0}. \quad (8.5)$$

For a probability density function, the second moment is its variance, and the square-root of that is the standard deviation, which (inversely) describes the extent to which the points in the support of  $\rho$  cluster about the mean. For a mass density function, the second moment relates to the moment of inertia.

One basically uses the moments of  $\rho$  to reduce the motion of the deformable, extended matter distribution that it describes to the motion of a point-like mass  $M(t)$  that is always located at the center-of-mass  $x_{cm}(t)$  of  $\rho$  ( $\sim M_1$ ). When one wishes to go on from a point to an extended, but rigid, body one will then associate the COM with an orthonormal frame, such as the principal frame of the moment of inertia that can rotate independently of the motion of the COM.

Note that there is generally a difference between defining the total linear momentum of the mass in the point approximation by the product  $M \mathbf{v}_{cm}$ , where  $\mathbf{v}_{cm} = dx_{cm} / dt$ , and defining it by the integral  $\int \rho v^i dV$  over all space, unless  $\mathbf{v}$  is constant in space to begin with. However, that would put one into the case of rigid motion, not deformable motion. Nonetheless, one does have that:

$$v_{cm}^i(t) = \frac{dx_{cm}}{dt} = \frac{d}{dt} \int_{\Sigma} \frac{\rho}{M} x^i(t, x) dV = \frac{1}{M} \int_{\Sigma} \frac{\partial(\rho x^i)}{\partial t} dV = \frac{1}{M} \int_{\Sigma} \left( \frac{\partial \rho}{\partial t} x^i + \rho \frac{\partial x^i}{\partial t} \right) dV ,$$

which will generally equal:

$$\bar{v}^i = \frac{1}{M} \int_{\Sigma} \rho v^i dV$$

when  $\rho$  is not time-varying. The expression on the right-hand side amounts to the mean velocity over the support of  $\rho$ .

Hence, one should be careful about jumping to the conclusion that the operations on densities and points commute with the operation of taking the moments.

The degree to which the approximation of a deformable, extended mass distribution with a moving massive point or frame is accurate can sometimes be estimated by looking at a breakdown of the distribution of the total energy of the original deformable, extended mass distribution over its (potentially-infinite) degrees of freedom. In particular, one might first look at how the work that is done on the extended object in question shows up in its elementary rigid-body degrees of freedom, namely, its bulk translation and rotation. One can then better say whether rigidity is really a close approximation or not. For instance, if one drops a ball of clay from a given height then the work that is done accelerating it as it falls will be converted almost completely into the work of deformation when it hits the floor and stops, except for a certain small percentage that goes into heat. In such a case, rigidity clearly breaks down at the point of impact.

The sloshing of liquids in moving tanks offers an interesting example of a motion of a deformable, extended mass distribution for which the kinetic energy of motion is mostly concentrated in the rigid-body degrees of freedom, while a much smaller percentage is distributed over such things as turbulence in the volume of the liquid and on its surface.

An interesting problem in which to test the various degrees of approximation to the motion of deformable, extended bodies is what one might call the “water balloon” problem, which amounts to determining the motion of a water balloon in space after it has been given some initial translation and rotational velocities, as well as an initial state of deformation. The successive degrees of approximation start with:



1. Point-mass moving along a curve.
2. Rigid-body (i.e., oriented, orthonormal frame) translating and rotating along a curve (e.g., COM).
3. Pseudo-rigid body (i.e., oriented, linear frame) that translates, rotates, and shears about a curve.

The last case includes the possibility that it can deform by elongation and contraction along the principal axes of the strain. Hence, a sphere can deform into an oblate spheroid or ellipsoid about any given oriented, orthonormal frame that is to serve as the principal frame.

**9. Frame strain.** – As mentioned above, the Cauchy-Green approach to strain considers only the relative deformations of neighboring distances in a material medium. Consequently, although it is true that a transformation of an  $n$ -dimensional Euclidian space that preserves all distances between pairs of points must be an isometry – i.e., a rigid motion or a reflection – nonetheless, a transformation of a lower-dimensional submanifold can be an isometry of the lower-dimensional metric without having to be rigid.

For instance, if one bends a wire of negligible cross-section without stretching or compressing it then that motion will preserve the one-dimensional metric along the wire – viz., the arc length – despite the fact that it does not have to be rigid. Similarly, one can deform an infinitely-thin surface (membrane, plate, shell, etc.) without stretching or compressing any of its regions and come up with an isometry of the surface that is not a rigid motion in its ambient space. For instance, all developable surfaces, such as cylinders and cones, are isometric to a plane, even though there is clearly a non-trivial deformation (i.e., diffeomorphism) that will deform a plate into a cylinder or cone, in the sense that work will be necessary in order to perform the deformation.

The way that continuum mechanics usually addresses that situation is to treat the wire and the surface as being, in fact, three-dimensional objects, such that the approximation is not in their dimensions, but their material properties. Hence, the wire would be a thin solid cylinder whose strains and stresses could be treated as constant across the cross-section, and similarly a membrane would be a cylinder whose height (i.e., thickness) was small enough that the strains and stresses could be treated as constant across the thickness.

*a. Monomorphisms of frame bundles.* – However, one might also approximate some of the dimensions of a  $n$ -object  $\mathcal{O}$  as effectively zero and consider the space  $GL_x\mathcal{O}$  of linear  $n$ -frames at each of its points  $x \in \mathcal{O}$  ( $k =$  reduced dimension of object,  $n =$  dimension of ambient space). One then says that the deformation affects not only the points of  $\mathcal{O}$ , but the frames of  $GL(\mathcal{O})$ , which is the set of all linear  $n$ -frames on all points  $x$ . Since that set projects onto  $\mathcal{O}$  in a natural way:

$$p: GL(\mathcal{O}) \rightarrow \mathcal{O}, (x, \mathbf{e}_i) \mapsto x,$$

and the set  $GL_x\mathcal{O}$  of all linear frames at  $x$  is called the *fiber* of  $x$ , one calls the triple  $(GL(\mathcal{O}), \mathcal{O}, p)$  a *fiber bundle* over  $\mathcal{O}$  <sup>(1)</sup>. In particular, it is the bundle of all linear  $n$ -frames in  $\mathbb{R}^n$  at the points of  $\mathcal{O}$ .

One can think of a deformation of  $\mathcal{O}$  as something that affects  $GL(\mathcal{O})$  directly, namely,  $F: GL(\mathcal{O}) \rightarrow GL(\mathbb{R}^n)$ , and takes a linear frame  $(x, \mathbf{e}_i)$  at  $x \in \mathcal{O}$  to a linear frame  $(y, \bar{\mathbf{e}}_i)$  at  $y \in \mathbb{R}^n$  in an invertible, differentiable way such that its inverse  $F^{-1}$  (which is defined only upon the image of  $\mathcal{O}$ ) is also differentiable; i.e.,  $F$  is a diffeomorphism of  $GL(\mathcal{O})$  into  $GL(\mathbb{R}^n)$ ; we call such a map a *monomorphism* of the two bundles when it also has the property that the image of any fiber  $GL_x(\mathcal{O})$  will be the corresponding fiber  $GL_{F(x)}(\mathbb{R}^n)$ . Since the Lie group  $GL(n)$  acts upon the frames in both  $GL(\mathcal{O})$  and  $GL(\mathbb{R}^n)$ , one can also characterize a monomorphism  $F: GL(\mathcal{O}) \rightarrow GL(\mathbb{R}^n)$  by saying that it commutes with the group action; i.e., for any frame  $\mathbf{e}_i$  at  $x$  and any  $A_j^i$  in  $GL(n)$ , one must have:

$$F(x, \mathbf{e}_j A_j^i) = (F(x), F(\mathbf{e}_j) A_j^i). \quad (9.1)$$

The particular form for  $F(\mathbf{e}_j)$  that we shall use is:

$$F(\mathbf{e}_j) = \mathbf{e}_j h_i^j(x), \quad (9.2)$$

in which  $h: \mathcal{O} \rightarrow GL(n)$ ,  $x \mapsto h_i^j(x)$  is a sufficiently-differentiable function that then takes the form of a transition function that acts upon frames.

If one then defines the projection of the map  $F$  onto the base manifolds, namely,  $f: \mathcal{O} \rightarrow \mathbb{R}^n$ ,  $x = p(GL_x(\mathcal{O})) \mapsto \bar{p} \cdot F(x)$ , which is the composition of  $F$  with  $\bar{p}: GL(\mathbb{R}^n) \rightarrow \mathbb{R}^n$  then the fact that  $F$  is a monomorphism will imply that  $f$  is a diffeomorphism onto.

One refers to  $F$  as *vertical* when  $f$  is the inclusion of  $\mathcal{O}$  in  $\mathbb{R}^n$  (i.e.,  $x \mapsto x$ ). Such a deformation will take a frame at  $x$  to another frame at  $x$  for every  $x$ . Otherwise, when  $f$  is not the inclusion,  $df|_x: T_x\mathcal{O} \rightarrow T_{f(x)}\mathbb{R}^n$  will be an invertible linear map for every  $x \in \mathcal{O}$ ,

---

<sup>(1)</sup> Of course, there are some topological conditions that must be added to make the definition of a fiber bundle complete, but we will not need them in what follows. One might, however, confer a good book on geometrical and topological methods in mathematical physics such as Frenkel [28].

and will then define an invertible, bi-differentiable map  $\bar{f}_x : GL_x \mathcal{O} \rightarrow GL_{f(x)} \mathbb{R}^n$ ,  $(x, \mathbf{e}_i) \mapsto (f(x), df_x(\mathbf{e}_i))$  that we call the *lift* of  $f$ .

Generally, the lift of  $f$  will not agree with  $F$  except that both  $\bar{f}$  and  $F$  will project to the same map. If one defines  $F_v = \bar{f}^{-1} \cdot F$  then  $F_v$  will be vertical, and one will have:

$$F = \bar{f} \cdot F_v.$$

That is, any monomorphism of  $GL(\mathcal{O})$  into  $GL(\mathbb{R}^n)$  can be expressed as the product of two unique monomorphisms, one of which is vertical, and the other of which is the lift of its projection. We shall say that a monomorphism  $F$  is *integrable* iff it coincides with the lift of its projection:

$$F = \bar{f}; \tag{9.3}$$

equivalently, its vertical part  $F_v$  will be the identity transformation at every point of  $\mathcal{O}$ . We illustrate the scenario that we have been discussing as Fig. I.2:

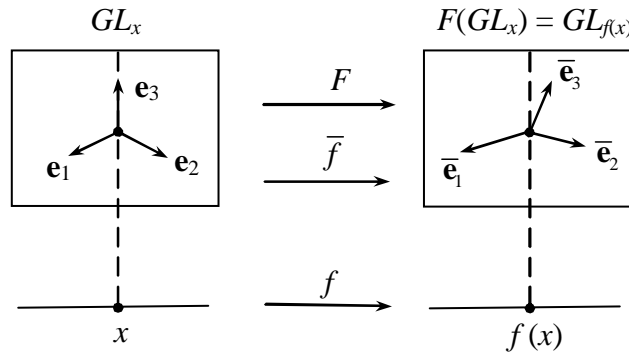


Figure I.2. A monomorphism of a frame bundle.

Although the manifold  $\mathcal{O}$  cannot always be covered with a single coordinate chart, any point  $x$  of it will admit at least one local coordinate chart  $(U, x^j)$  in  $\mathbb{R}^n$  that is *adapted* to  $\mathcal{O}$  (so the last  $n - k$  coordinates will be constant for all points of  $\mathcal{O}$ ), and for which  $GL(\mathcal{O})$  will be topologically equivalent to  $U \times GL(n)$  when one restricts  $U$  to points of  $\mathcal{O}$ . In order to get coordinates for  $GL(n)$ , one defines the natural frame field that goes with  $(U, x^j)$  – viz.,  $\mathbf{e}_i = \partial_i$  – and that will allow one to associate every linear  $n$ -frame  $\mathbf{e}_i$  at each  $x$  with the invertible matrix  $A_j^i$  that makes:

$$\mathbf{e}_i = \partial_j A_i^j. \tag{9.4}$$

Hence, the coordinate charts of  $GL(\mathcal{O})$  will look like  $(x^1, \dots, x^k, A_j^i)$ .

When one has such a chart  $(U, x^j)$  about  $(x, \mathbf{e}_i)$  and another  $(V, y^j)$  about  $(y, \bar{\mathbf{e}}_i)$ , one can express a monomorphism  $F$  as two systems of equations:

$$y^i = y^i(x^j), \quad \bar{\mathbf{e}}_i = \bar{\mathbf{e}}_i(x^k, \mathbf{e}_i) = \mathbf{e}_j \bar{h}_i^j(x). \quad (9.5)$$

The projected map  $f$  is described by simply the first set of equations in (9.5). Its lift  $\bar{f}$  will then be defined by the matrix of  $df$ , namely:

$$\left[ df|_x \right]_j^i = \frac{\partial y^i}{\partial x^j}, \quad (9.6)$$

whose inverse will be  $\partial x^i / \partial y^j$ .

The vertical part  $F_v$  of  $F$  will then be defined by the equations:

$$y^i = x^i, \quad \bar{\mathbf{e}}_i = \mathbf{e}_j \bar{h}_i^j \frac{\partial x^i}{\partial y^j} = \mathbf{e}_j h_i^j, \quad (9.7)$$

in which we have defined the transition function:

$$h_i^j = \bar{h}_i^j \frac{\partial x^i}{\partial y^j}. \quad (9.8)$$

We now see that, in a sense, the scope of the Cauchy-Green picture of deformation and strain is confined to the “horizontal” part of the deformation of  $GL(\mathcal{O})$  that is defined by the monomorphism  $F$ ; viz., the lift  $\bar{f}$  of its projection  $f$ . However, since the dimension of  $GL(\mathcal{O})$  is greater than that of  $\mathcal{O}$  by  $n^2$ , there will be many more deformations of  $GL(\mathcal{O})$  than there are of  $\mathcal{O}$ . In order to show how deformations like  $F$  can be used to enlarge the scope of the definition of strain, we must define strain for them in a way that extends the definition of strain for the deformations of  $\mathcal{O}$ .

*b. Definition of frame strain.* – The definition of the *finite frame strain matrix* will be simply the matrix  $h_i^j$  that represents the vertical part of a monomorphism of the bundle  $GL(\mathcal{O}) \rightarrow \mathcal{O}$  of linear frames on a material object  $\mathcal{O}$  into the bundle  $GL(\mathbb{R}^n) \rightarrow \mathbb{R}^n$  of linear frames on the space  $\mathbb{R}^n$  in which the object is embedded.

The corresponding *infinitesimal frame strain matrix*  $\omega_i^j$  is defined by the sufficiently-differentiable function  $\omega: \mathcal{O} \rightarrow \mathfrak{gl}(n)$  that takes every point  $x$  in the object to the matrix  $\omega_i^j(x)$  that represents the infinitesimal generator of the matrix  $h_i^j(x)$ ; i.e.:

$$\omega_i^j(x) = \exp [ h_i^j(x) ]. \quad (9.9)$$

The matrix  $\omega_i^j(x)$  acts upon the frame field  $\mathbf{e}_i(x)$  to produce minus its differential:

$$d\mathbf{e}_i = -\mathbf{e}_j \otimes \omega_i^j, \quad (9.10)$$

so upon differentiating  $\mathbf{e}_j h_i^j$  one will get:

$$d(\mathbf{e}_j h_i^j) = d\mathbf{e}_j h_i^j + \mathbf{e}_j \otimes dh_i^j = -\mathbf{e}_j \otimes (\omega_i^j - dh_k^j \tilde{h}_i^k) h_i^j.$$

If this vanishes then we must have:

$$\omega_i^j = dh_k^j \tilde{h}_i^k = \tilde{h}_k^j dh_i^k. \quad (9.11)$$

In order to see the last equality, note that when one replaces  $h_i^j$  with its inverse and *vice versa* in the expression  $dh_k^j \tilde{h}_i^k$ , the effect should be to produce minus  $\omega_i^j$ , but one has:

$$d\tilde{h}_k^j h_i^k = -\tilde{h}_k^j dh_i^k, \quad (9.12)$$

which follows upon differentiating the identity  $\tilde{h}_k^j h_i^k = \delta_i^j$ .

Now, let us introduce a metric on  $\mathbb{R}^n$ , which one can restrict to  $k$ -dimensional objects in it, such as  $\mathcal{O}$ . For non-relativistic continuum mechanics, that metric would be the Euclidian one  $\delta$ , while for relativistic problems, it would be the Minkowski (or Lorentzian) metric  $\eta$ .

The introduction of a metric into the tangent spaces to  $\mathbb{R}^n$  also allows one to define a special class of  $n$ -frames in the form of orthonormal ones. Actually, any  $n$ -frame can be *defined* to be orthonormal, and the metric that makes it so will follow from that definition. For instance, if  $\{\mathbf{e}_i, i = 1, 2, 3\}$  is a 3-frame in  $\mathbb{R}^3$  then one can define a Euclidian metric  $\delta$  on  $\mathbb{R}^n$  by demanding that:

$$\delta(\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij}. \quad (9.13)$$

If one wishes to find the scalar product  $\delta(\mathbf{v}, \mathbf{w})$  for any two vectors in  $\mathbb{R}^3$  then one must first express them in terms of the given frame as  $\mathbf{v} = v^i \mathbf{e}_i$ ,  $\mathbf{w} = w^j \mathbf{e}_j$  and then use the bilinearity of the scalar product:

$$\delta(\mathbf{v}, \mathbf{w}) = v^i w^j \delta(\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij} v^i w^j. \quad (9.14)$$

One can do something analogous for a linear frame in  $\mathbb{R}^4$  and the Minkowski scalar product.

Since orthogonal transformations must take orthonormal frames to other orthonormal frames, any linear frame will define an orbit in the space of linear frame under the action of a chosen orthogonal group. One can think of that orbit space as defining the manifold of all metrics  $\mathcal{M}(p, n-p)$  of the chosen signature type. Hence, diffeomorphisms of  $\mathcal{M}(p, n-p)$  will represent deformations of metrics.

The essential point is that since a linear frame defines a metric, a deformation of the linear frame will either be orthogonal for that metric or it will deform the metric in the process, which will represent strain in the usual sense. If the deformation in question is an invertible linear map of the frame then polar decomposition:

$$h_i^j(x) = R_k^j(x) E_i^k(x) \quad (9.15)$$

will give the part  $E_i^k(x)$  of the deformation that actually changes the metric as a finite strain matrix, along with the orthogonal part  $R_k^j(x)$  that does not.

However, it is a serious oversight to think that the orthogonal part of the polar decomposition will play no role in the geometry of deformation merely because it does not alter the metric at each point individually. Indeed, the only way that a rotation at each point will represent a rigid motion (i.e., no deformation) is when the *same* <sup>(1)</sup> rotation acts at all points of the object. Hence, one must expect that the differential  $dR_i^j$  must a crucial role.

*c. Frenet frames.* – A simple example of that situation is given by a skew curve  $x(s)$  in  $\mathbb{R}^3$  for which non-zero first and second derivatives exist, and they are not collinear; one also assumes that the parameterization is by arc length, so the velocity  $\mathbf{v}(s)$  will have unit speed for all  $s$ . One can then define a *Frenet frame*  $\{\mathbf{e}_i(s), i = 1, 2, 3\}$  for the curve by setting  $\mathbf{e}_1(s)$  equal to  $\mathbf{v}(s)$ , and  $\mathbf{e}_2(s)$  equal to the normalized acceleration vector field, while  $\mathbf{e}_3(s)$  is a unit vector that is perpendicular to the plane of  $\mathbf{e}_1(s)$  and  $\mathbf{e}_2(s)$  (which is called the *osculating plane*) and completes the 3-frame in a right-hand oriented way. One refers to the three vectors of the Frenet frame as the *tangent*, *normal*, and *binormal*, resp., of the curve at each point.

By definition, all frame vectors are unit vectors, and the third one is orthogonal to both of the first two. The fact that the first two are orthogonal to each other follows from the condition that  $\mathbf{v}(s)$  must be a unit vector:

$$\langle \mathbf{v}(s), \mathbf{v}(s) \rangle = 1, \quad \text{so} \quad 0 = \frac{d}{ds} \langle \mathbf{v}(s), \mathbf{v}(s) \rangle = 2 \langle \mathbf{v}(s), \frac{d\mathbf{v}}{ds} \rangle.$$

---

<sup>(1)</sup> Of course, the definition of “same” also depends upon topological subtleties that relate to the difference between a general fiber bundle and one whose total space is the product of the base manifold with fiber – viz., *trivial* fiber bundles. However, we shall not need that level of generality in what follows.

Since the frame  $\mathbf{e}_i(s)$  spans the tangent spaces at each  $x(s)$ , which can be identified with  $\mathbb{R}^3$ , any vector in  $\mathbb{R}^3$  can be expressed in terms of the frame  $\mathbf{e}_i(s)$ . In particular, the vectors  $d\mathbf{e}_i / ds$  can be expressed in terms of  $\mathbf{e}_i(s)$  in the form of the *Frenet-Serret equations*:

$$\frac{d\mathbf{e}_i}{ds} = \mathbf{e}_j \omega_i^j, \quad \omega_i^j = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix}. \quad (9.16)$$

The parameter  $\kappa$  relates to an infinitesimal rotation in the osculating plane about the binormal, and is called the *curvature* of the curve. The parameter  $\tau$  relates to an infinitesimal rotation about the tangent vector, and is called the *torsion* of the curve. When the curvature vanishes, the curve must be a straight line, which will make the binormal undefined; when the torsion vanishes, the curve must always lie in the osculating plane.

Now, suppose that a curve  $x(s)$  is deformed into a curve  $y(s)$  with no change in the arc length, so points of the two curves that have the same value of  $s$  will correspond under the deformation. Let this deformation be expressed by the system of equations:

$$y^i = y^i(x^j), \quad (9.17)$$

which are, of course, defined only for the points of  $x(s)$ .

The new velocity vector at  $y(s)$  will be:

$$\frac{dy^i}{ds} = \frac{\partial y^i}{\partial x^j} \frac{dx^j}{ds}, \quad (9.18)$$

and since the arc length does not change, the matrix  $\partial y^i / \partial x^j$  must represent a rotation of the tangent vector, in addition to its translation from  $x(s)$  to  $y(s)$ .

The new acceleration vector will be:

$$\frac{d^2 y^i}{ds^2} = \left( \frac{d}{ds} \frac{\partial y^i}{\partial x^j} \right) \frac{dx^j}{ds} + \frac{\partial y^i}{\partial x^j} \frac{d^2 x^j}{ds^2} = \frac{\partial^2 y^i}{\partial x^j \partial x^k} \frac{dx^j}{ds} \frac{dx^k}{ds} + \frac{\partial y^i}{\partial x^j} \frac{d^2 x^j}{ds^2}. \quad (9.19)$$

However, since the velocity is a unit vector, the acceleration will still be orthogonal to that velocity. Hence, when one normalizes that acceleration, and completes the orthonormal triad that represents the new Frenet frame  $\bar{\mathbf{e}}_i(s)$ , one will see that the net effect of the deformation, apart from the translation of points of the curve has been a one-parameter sequence of rotations of the initial frame:

$$\bar{\mathbf{e}}_i(s) = \mathbf{e}_j(s) R_i^j(s). \quad (9.20)$$

One can then express the new Frenet-Serret equations as:

$$\frac{d\bar{\mathbf{e}}_i}{ds} = \bar{\mathbf{e}}_j \bar{\omega}_i^j = \mathbf{e}_j R_k^j \bar{\omega}_i^k.$$

But since:

$$\frac{d\bar{\mathbf{e}}_i}{ds} = \frac{d\mathbf{e}_j}{ds} R_i^j + \mathbf{e}_j \frac{dR_i^j}{ds},$$

one must have:

$$\frac{d\mathbf{e}_i}{ds} = \mathbf{e}_j \left( R_k^j \bar{\omega}_i^k \tilde{R}_i^l - \frac{dR_k^j}{ds} \tilde{R}_i^k \right),$$

which will imply that:

$$\omega_i^j = \tilde{R}_k^j \bar{\omega}_i^k R_i^l - \frac{d\tilde{R}_k^j}{ds} R_i^k,$$

or:

$$\bar{\omega}_i^j = R_k^j \omega_i^k \tilde{R}_i^l + \tilde{R}_k^j \frac{dR_i^k}{ds}. \quad (9.21)$$

One can then think of this last equation as describing the deformation of the curvature and torsion of the curve by way of the action of the curve  $R(s)$  in  $SO(3)$  on  $SO(\mathcal{O})$ , where  $\mathcal{O}$  is the curve  $x(s)$ , in this case. Note that if  $R(s)$  is rigid (i.e., constant) then the transformation from  $\omega_i^j$  to  $\bar{\omega}_i^j$  will be an overall change of reference frame for the components. Furthermore, if the initial curve is straight ( $\omega_i^j = 0$ ) then the deformed  $\bar{\omega}_i^j$  will be equal to:

$$\bar{\omega}_j^i = R_k^i \frac{d\tilde{R}_j^k}{ds}. \quad (9.22)$$

Actually, the method that we used for obtaining (9.21) did not depend essentially upon our choice of the Frenet frame for the orthonormal frame field along  $x(s)$ . If we had used any other orthonormal frame field for  $\mathbf{e}_i(s)$  then the first difference in the *differential equations for the moving frame* – namely, (9.16) – would be in the definition of the matrix  $\omega_i^j$ , which would not need to have zeroes in outermost off-diagonal entries, but only along the diagonal; it would still have to be antisymmetric, though.

Due to the linearity of the system (9.16), one can solve the initial-value problem for it by matrix exponentiation:

$$\mathbf{e}_i(s) = \mathbf{e}_i(0) \exp \int_0^s \omega_i^j(\sigma) d\sigma. \quad (9.23)$$

One can define the deformation of the frames by simply  $R_j^i(s)$ , *a priori*, and one will eventually obtain (9.21) for a measure of the deformation of the frames along  $x(s)$  as a result of  $R(s)$ .

*d. Higher-dimensional Frenet frames.* – There is also nothing special about the choice of one parameter objects, since one can extend to higher-dimensional objects, such



as surfaces, by extending the differential equations for the moving frame for a linear system of partial differential equations for the frame field on the object, and if the parameters are  $u^a$ ,  $a = 1, \dots, k$  then the equations of the frame field  $\mathbf{e}_i(u^a)$  will be:

$$\frac{\partial \mathbf{e}_i}{\partial u^a} = \mathbf{e}_j \omega_{ia}^j, \quad \omega_{ia}^j = \begin{bmatrix} 0 & \kappa_a & -\lambda_a \\ -\kappa_a & 0 & \tau_a \\ \lambda_a & -\tau_a & 0 \end{bmatrix}, \quad a = 1, \dots, k. \quad (9.24)$$

However, integrating this system will no longer be as simple as (9.23).

If one defines the deformation of the frame field  $\mathbf{e}_i(u^a)$  by way of  $R(u^a)$ ,  $a = 1, \dots, k$  then the analogue of (9.21) will be:

$$\bar{\omega}_{ia}^j = R_k^j \omega_{ia}^k \tilde{R}_i^l + \tilde{R}_k^j \frac{\partial R_i^k}{\partial u^a}. \quad (9.25)$$

The author has discussed some of these issues in more detail in his article [10]. However, in the present context, it will not be necessary to say quite so much, and we will only make occasional comments that relate to this aspect of the kinematics of deformation. It should be pointed out that Kelvin and Tait [29] essentially looked at the deformation of frames as a way of describing the bending and twisting of curves and surfaces in the years before Riemannian geometry began attracting all of the attention.

**10. Cosserat media.** – In the last section, we discussed the deformation of frames on objects without mentioning the associated displacement of the points to which they were associated. Eventually, one needs to include some accounting for the deformation of those points, as well.

One might notice that system of equations (9.5) could just as well define an affine transformation, depending upon the nature of the functions involved. That really comes down to the fact that the manifold that underlies the Lie group  $A(n)$  – i.e., the  $n$ -dimensional affine group – is diffeomorphic to the (trivial) bundle  $GL(\mathbb{R}^n)$  of linear frames on  $\mathbb{R}^n$ . The diffeomorphism is defined by choosing a linear frame  $(x, \mathbf{e}_i)$  at a point  $x$  in  $\mathbb{R}^n$  and applying an affine transformation  $(\mathbf{a}, A_j^i)$  to  $(x, \mathbf{e}_i)$  as follows:

$$(x, \mathbf{e}_i) (\mathbf{a}, A_j^i) = (x + \mathbf{a}, \mathbf{e}_i A_j^i) = (y, \bar{\mathbf{e}}_j),$$

so:

$$y = x + \mathbf{a}, \quad \bar{\mathbf{e}}_j = \mathbf{e}_i A_j^i. \quad (10.1)$$

Since every element of  $A(n)$  has a unique inverse, this pair of equations will be uniquely invertible, and there will be a one-to-one correspondence between all  $(y, \bar{\mathbf{e}}_j)$  in  $GL(\mathbb{R}^n)$

and the  $(\mathbf{a}, A_j^i)$  in  $A(n)$ , which will be differentiable with a differentiable inverse, moreover.

The action of  $A(n)$  on  $GL(\mathbb{R}^n)$  above differs crucially from the action of  $GL(n)$  on  $GL(\mathbb{R}^n)$  as a structure group, because the latter action is only vertical; i.e., it will project to the identity transformation on  $\mathbb{R}^n$  under the canonical projection  $GL(\mathbb{R}^n) \rightarrow \mathbb{R}^n$  that associates a linear frame at  $x$  with  $x$ . However, as we see from (10.1), the action of  $A(n)$  on  $GL(\mathbb{R}^n)$  also moves the points in  $\mathbb{R}^n$  that the frames are associated with. That is why one must be careful about committing to the methods of principal fiber bundles (such as the various frame bundles) when one needs to include the displacement of points in space along with the displacement of the frames at those points.

When one has defined an orthogonal structure on  $\mathbb{R}^n$ , one can restrict the aforementioned action to the subgroup of rigid motions or the Poincaré group, as the case may be. The corresponding diffeomorphism will then be with the bundle of orthonormal (Lorentzian, resp.) frames over  $\mathbb{R}^n$ . For three-dimensional Euclidian space  $E^3$ , the diffeomorphism is between  $ISO(3)$  – viz., the Lie group of rigid motions – and  $SO(E^3)$ , which is the bundle of orthonormal 3-frames on  $E^3$ . For four-dimensional Minkowski space  $\mathfrak{M}^4$ , the diffeomorphism is between  $ISO(3, 1)$  – viz., the Poincaré group – and  $L(\mathfrak{M}^4)$ , which is the bundle of Lorentzian frames on  $\mathfrak{M}^4$ .

The thought of expanding the scope of non-relativistic continuum mechanics from regions of space to the bundles of orthonormal frames on the regions goes back to the work of the brothers Eugène and François Cosserat, and was set down in their *magnum opus* that was entitled *La théorie de les corps déformables* [30] in 1909. For them, the precedent for such a theory already existed in some work of Woldemar Voigt [31] on the elasticity of crystals, as well as some of the mechanical ether models for the propagation of electromagnetic waves. The essential innovation from the phenomenological standpoint was the existence of internal couple stresses that would act upon the medium, so one could define a *Cosserat medium* to be the bundle  $SO(\mathcal{R})$  of orthonormal frames on a region  $\mathcal{R}$  of space that is occupied by matter that is subject to internal couple stresses. One symptom of the existence of such stresses is the asymmetry of the stress tensor, which ultimately factors in the equations for the balance of angular momentum as an effective torque or moment. One will also generally have a momentum 1-form that is not collinear with the covelocity 1-form; i.e., there will be a “transverse momentum” that is added to the collinear kind.

Since the basic starting point for the derivation of the Cosserat equations for the equilibrium or time evolution of a Cosserat medium is in the theory of action functionals that are invariant under the action of the group of rigid motion, we shall defer a more detailed discussion of that topic to later chapters as it becomes appropriate. For now, we mention that when one goes on to quantum wave equations that involve matter with spin, one will find that the continuum-mechanical equivalent of the Dirac electron is

essentially a relativistic Cosserat medium, although the stress tensor of the non-relativistic spinning Pauli electron is symmetric.

Some more recent references on Cosserat theory will be given in later chapters as they become relevant.

### References (\*)

1. D. J. Saunders, *Geometry of Jet Bundles*, Cambridge University Press, Cambridge, 1989.
2. C. Truesdell and R. Toupin, "The Classical Field Theories," in *Handbuch der Physik* III/1, ed. S. Flügge, Springer, Berlin, (1960), pp. 226-793.
3. A. E. H. Love, *A Treatise on the Mathematical Theory of Elasticity*, 4<sup>th</sup> ed., Cambridge University Press, Cambridge, 1927; reprinted by Dover, NY, 1944,
4. L. D. Landau and E. M. Lifschitz, *Theory of Elasticity*, Pergamon, London, 1959.
5. Sir H. Lamb, *Hydrodynamics*, 6<sup>th</sup> ed., Cambridge Univ. Press, Cambridge, 1932; First ed. 1879, reprinted by Dover, Mineola, NY, 1945.
6. Batchelor, D., *Introduction to Fluid Mechanics*, Cambridge University Press, Cambridge, 1967.
7. Landau, L., Lifschitz, L., *Fluid Mechanics*, Pergamon Press, London, 1987.
- 8\* . F. Gallisot, "Les formes extérieures et le mécanique des milieux continus," *Ann. Inst. Fourier (Grenoble)* **8** (1958), 291-335.
9. F. W. Murnaghan, *Finite Deformation of an Elastic Solid*, Wiley, New York, (1951).
10. D. H. Delphenich, "The use of the teleparallelism connection in continuum mechanics," *Mathematics and Mechanics of Solids* **21** (2016), 1260-1275; first published online Dec. 26, 2014.
11. J. F. Pommaret, *Lie Pseudogroups in Mechanics*, Gordon and Breach, New York, 1988.
- 12\* . J. P. Vigièr, *Structure des micro-objets dans l'interprétation causale de la théorie des quanta*, Gauthier-Villars, Paris, 1956.
13. D. H. Delphenich, "Projective geometry and special relativity," *Ann. Phys. (Leipzig)* **15** (2006) 216-246.
14. J. L. Synge, "Relativistic Hydrodynamics," *Proc. London Math. Soc.*, **43** (1937), 376-416.
- 15\* . A. Lichnerowicz, *Théorie relativistes de la gravitation et de l'électromagnétisme*, Masson and Co., Paris, 1955.
16. A. Anile, *Relativistic Fluids and Magnetofluids*, Cambridge Univ. Press, 2008.
- 17\* . F. Halbwachs, *La théorie relativiste des fluides à spin*, Gauthier-Villars, Paris, 1960.
18. A. Brota, J. C. Fernandes, "The relativistic elasticity of rigid bodies," arXiv:physics/0307019.
- 19\* . M. Born, "Die Theorie des starren Elektrons in der Kinematik des Relativitätsprinzips," *Ann. Phys. (Leipzig)* **30** (1909), 1-56.

---

(\*) References marked with an asterisk are available in English translation at the author's website: neo-classical-physics.info.

20. J. L. Synge, “A theory of elasticity in general relativity,” *Math. Zeit.* 72 (1959), 52-87.
  21. C. B. Rayner, “Elasticity in general relativity,” *Proc. Roy. Soc. London* **2724** (1963), 44-53.
  22. J. G. Oldroyd, “Equations of state of continuous matter in general relativity,” *Proc. Roy. Soc. London A* **316** (1970), 1-28.
  23. B. Carter and H. Quintana, “Foundations of general relativistic high-pressure elasticity theory,” *Proc. Roy. Soc. London A* **331** (1973), 57-83.
  24. H. Hermann, G. Rückner, *et al*, “Spin axioms in relativistic continuum physics,” *Found.Phys.* 34 (2004) 1005-1021.
  25. W. Israel, “Relativistic theory of shock waves,” *Proc. Roy. London A* **259** (1960), 129-143.
  - 26\*. G. Herglotz, “Über die Mechanik des deformieren Körpers vom Standpunkt der Relativitätstheorie,” *Ann. d. Phys.* (1911), 493-533.
  - 27\*. H.-G. Schöpf, “Allgemeinrelativistische Prinzipien der Kontinuumsmechanik,” *Ann. Phys. (Leipzig)* **12** (1964), 377-395.
  28. T. Frenkel, *The Geometry of Physics: an introduction*, Cambridge University Press, Cambridge, 1997.
  29. W. Thomson (Lord Kelvin) and P. G. Tait, *Treatise on Natural Philosophy*, 2 vols., Cambridge University Press, Cambridge, 1895; available as a facsimile reprint through Elibron Classics, [www.elibron.com](http://www.elibron.com), 2005.
  - 30\*. E. Cosserat and F. Cosserat, *Théorie des corps déformables*, Hermann, Paris, 1909.
  - 31\*. W. Voigt:
    - “Theoretische Studien über die Elasticitätsverhältnisse der Krystalle, I, II, *Abh. d. kgl. Ges. d. Wiss. zu Göttingen*, **34** (1887), 3-52, 53-100.
    - “L’état actuel de nos connaissances sur l’élasticité des cristaux,” report presented at the International Congress of Physics, Paris (1900), t., 277-347.
-

## CHAPTER II

# THE MECHANICS OF WAVES

In retrospect, there seems to have been a certain amount of hastiness in the early attempts to lay the theoretical foundations of quantum physics. Of course, that is to be expected, since the founders were dealing with a class of natural phenomena that lies far beneath the scale of macroscopic laboratory measurements, and those laboratory measurements were producing results that were inconsistent with the intuition that physicists had gained from dealing with more directly observable phenomena. Hence, the early attempts to devise some sort of theoretical framework were predictably semi-empirical and phenomenological. Indeed, that basic approach to quantum theory persists to this day.

Part of the hastiness that was associated with going from the mechanics of waves to the mechanics of points was due to the fact that no one seemed to have taken the time to formulate a general theory of wave motion that would be as fundamental in character as the mechanics of points and deformable, extended matter. That is, no one had started by defining a *wave* as an elementary object that moved in a configuration space, and thus had a kinematical state associated with it. Thus, they also did not take the time to define dynamical states that would be associated with the variations (i.e., virtual displacements) of the kinematical states, but they seemed to simply start with the partial differential equation that would govern the time evolution of a wave function. By Schrödinger's own admission, the process of obtaining his equation involved a certain amount of trial-and-error that mostly used the Hamilton-Jacobi approach to geometrical optics (i.e., the *eikonal*) as its basic model.

In this chapter, we shall attempt to fill in some of the missing steps that would connect the main field theories that involved some form of wave motion – namely, elasticity, electromagnetism, and more recently gravitation – to the Hilbert space formalism of quantum wave mechanics. In particular, since many of the most fundamental experimental phenomena of quantum physics seemed to point to a breakdown of Maxwell's theory of electromagnetism, one would expect that the class of waves that is most relevant to quantum phenomena the that of electromagnetic waves.

The basic flow of ideas amounts to an application of a general theory of physical models that the author has been developing over the years [1], and has previously discussed in the context of models for wave motion in [2]. Basically, a physical model (whether static or dynamic) generally involves the following components:

1. A differentiable manifold of kinematical states whose tangent vectors represent infinitesimal (i.e., virtual) displacements (or variations) of those states.
2. A system of differential equations that define the integrability of kinematical states and their variations.
3. A space of dynamical states that are, in some sense, dual to the kinematical states or their variations.
4. A constitutive map that associates dynamical states with kinematical ones.

5. A system of differential equations that expresses the balance principle or conservation law that governs the dynamical state.

**1. The state of a wave medium.** – Our starting point for a general theory of wave motion must necessarily be a definition of a wave. One finds that this problem is actually more complicated than it sounds. One could make a first attempt by saying that a wave is a “wave-like” solution of a system of field equations, but that obviously just replaces one undefined concept with another.

It is easier to say what the definition of a wave should *not* be than to say what it is. Most treatments of waves in elastic and electromagnetic media, and many treatments of gravitational waves, start by introducing small-amplitude perturbations of fields, which predictably results in linear partial differential equations for the waves. However, one of the most active topics in modern nonlinear physics is that of nonlinear waves, such as one finds in nonlinear optics or large-amplitude elastic waves. Indeed, one of the most common wave phenomena that people observe, namely, the breaking of waves on a beach, is manifestly nonlinear in character. Similarly, despite the ubiquitous use of plane-wave solutions of wave equations, one should probably not start with them as elementary objects, since the Fourier analysis of a more general wave-like solution into a linear superposition of plane waves of varying frequencies and wave numbers is more useful when the operator that takes the Fourier transform of an “incoming” wave to the Fourier transform of an “outgoing” wave is actually a linear operator, which then restricts the wave equation to a linear partial differential equation with constant coefficients. It should also be pointed out that plane waves are actually physically absurd unless one restricts them to a compact region, since otherwise their total energy and momentum will diverge, but restricting them to a region with boundary will introduce higher harmonics due to the cutoff, which would imply that one is no longer dealing with a true plane-wave anymore.

Another wrong turn in the name of basic definitions is to start with any specific wave equation as the basis for all waves, since the proliferation of specialized linear and nonlinear wave equations by now would suggest that any such definition would be too limited in scope to represent a fundamental statement about natural law as it pertains to waves in general.

Many physicists and engineers agree that the essence of wave phenomena is rooted in the “response” of a “wave medium” to a “disturbance.” Hence, we shall attempt to first clarify those terms.

We define a *wave medium* to be a region  $M$  of space-time (where the dimension of space can range from one to three) in which each point is associated with an “elementary oscillator” and those oscillators are coupled by some principle <sup>(1)</sup>. Typically we shall assume only a “cylindrical” topology to  $M$ , namely,  $M = (t_0, t_1) \times \Sigma_n$ , where  $\Sigma_n$  ( $n = 1, 2,$

---

<sup>(1)</sup> If one prefers the language of fiber bundles, a wave medium might be defined to be a “bundle of oscillators” with some coupling law. The definition of “oscillator” then bears upon the choice of manifold for the fiber of that bundle.

3) is a spatial manifold, which will usually be an open subset of  $\mathbb{R}^n$ , such as the set of points at which the wave function is non-zero. (The closure of that set is called the *support* of the wave function, although it is possible for a wave function to vanish at a limit point.)

*a. Elementary oscillators.* – Of course, even the concept of an elementary oscillator can lead to its own complications, since the theory of oscillators is quite vast in its own right [cf., e.g., **3, 4**]. Even if one assumes that an oscillator is most fundamentally characterized by a dynamical system that exhibits periodic motions, some of the issues that one must still address are:

1. The dimension of the space in which oscillation takes place.
2. Linearity vs. nonlinearity.
3. Damping vs. the absence of damping.
4. Forced vs. unforced.

We believe that it is best to simply regard these issues as the basis for a classification scheme that one could apply to oscillators. In most cases, one begins by choosing a type of basic oscillator upon which to base wave motion, just as most discussions of problems in continuum mechanics often start with a choice of mechanical constitutive law.

As far as the issue of dimension is concerned, one understands that it amounts to the problem of characterizing the basic “amplitude” of the oscillatory motion. That is, when the amplitude is characterized by a real number, the space of oscillation is one-dimensional. Although one might say that in the case of complex wave functions that amplitude is a complex number, nevertheless, it is common to regard that number as having a real amplitude and phase factor of the form  $e^{i\theta}$ . Of course, waves in elastic media typically involve oscillators that oscillate in three-dimensional space, while possibly changing direction, as well, and electromagnetic waves include the dimension of time, as well. Since gravitational fields are defined by symmetric, second-rank covariant tensors of globally hyperbolic normal type, one assumes that their basic oscillators would live in a ten-dimensional space (<sup>1</sup>).

By now, the issue of linearity seems to be widely recognized to be one of “small-amplitude” approximations. Hence, one suspects that only a nonlinear theory would be truly definitive. This is especially true when one considers that as far as electromagnetic phenomena are concerned, the quantum domain is characterized by large electric and magnetic field strengths at which non-classical phenomena, such as vacuum polarization, begin to emerge, as they might with a critical-point phase transition. However, one of the defining characteristics of the elementary wave equations of quantum mechanics (e.g., Schrödinger, Pauli, Klein-Gordon, Dirac) is their linearity, although many believe that the transition from quantum mechanics to quantum field theory must include some introduction of nonlinearity into the wave equations, since it is widely believed that the

---

<sup>(1)</sup> In fact, it is not a vector space, either, but a homogeneous space that is only homeomorphic to a vector space, namely, the connected component of  $GL(4) / SO(3, 1)$  that includes  $SO(3, 1)$  as a coset.

interaction of elementary particles will destroy that linearity. Indeed, that is why quantum field theory immediately passes over any discussion of the Cauchy problem for interacting particle fields to the scattering approximation, in which one assumes that the incoming fields are defined at time minus infinity, while the outgoing fields are defined at time plus infinity. That is essentially equivalent to assuming that the time interval during which the interaction occurs is quite short, and linearizes the time-evolution operator for the particle fields to a scattering operator of the linear, integral type [cf., e.g., **5, 6**].

Presumably, an oscillator gets damped by interacting with an absorptive medium, such as a viscous fluid. Although the concept of absorptive medium is certainly relevant to virtually all experimental work, since there is bound to be some damping of local oscillations, one also expects that damping can only be due to unmodeled complexity in the medium. However, there are those who suspect that the nature of the quantum vacuum that all fields must ultimately interact with has something of a thermodynamic character, which was the essence of de Broglie’s theory [7] of the “hidden thermodynamics of isolated particles.” Hence, one should not dismiss the possibility that even isolated particle/waves can still exchange energy with their environment.

A forcing function that acts upon an oscillator can take the form of gravity, the momentum that is carried by incident waves, or more artificial loadings that one might encounter in engineering practice. Such forces tend to be defined on the base manifold of a bundle of oscillators and not the total space. In the next subsection, we shall discuss the way that systems of oscillators can be coupled by forces of interaction that become essentially “internal” forcing functions.

Ultimately, one must define some manifold  $\mathfrak{D}$  that represents the generic state space of the oscillators. In the simplest case, it might be  $\mathbb{R}$ , and its points would represent the amplitude of the oscillation. Another elementary one-dimensional state space might be  $S^1$ , which might also describe the direction of a unit vector in a plane.

*b. Examples of elementary oscillators.* – Some elementary examples of oscillators are:

1. One-dimensional simple harmonic oscillator: For this oscillator, the basic relation takes the form of Hooke’s law:

$$F = -k \Delta x, \quad (1.1)$$

or expressions that are analogous to it.  $F$  represents the force that is associated with a displacement  $\Delta x$  from the equilibrium position  $x_0$  of the oscillator, and  $k$  is an empirical spring constant. If one assumes that  $x_0 = 0$ , by definition, then  $\Delta x$  can be replaced with  $x$ , but it is important to note that the force is not independent of the location of  $x_0$ . Hence, the basic constitutive law for such a dynamical system must be linear, non-dispersive, and homogeneous.

Such a force is conservative and admits a potential function of the form:



$$U(x) = \frac{1}{2} kx^2. \quad (1.2)$$

The differential equation for the forced, linearly-damped ( $F_d = -b \, dx / dt$ ) simple harmonic oscillator is obtained from Newton's second law and takes the form:

$$\frac{d^2x}{dt^2} + 2\omega_n \zeta \frac{dx}{dt} + \omega_n^2 x = a(t), \quad (1.3)$$

in which:

$$\omega_n^2 \equiv \frac{k}{m}, \quad 2\omega_n \zeta \equiv \frac{b}{m}, \quad (1.4)$$

and  $a(t)$  is the forcing function, which has units of acceleration.  $\omega_n$  is referred to as the *natural frequency* of the system, while  $\zeta$  is the *damping constant*.

Initial-value solutions to the equation (1.3) take the exponentially-damped sinusoidal form:

$$x(t) = e^{\alpha t} (A e^{\omega t} + B e^{-i\omega t}), \quad (1.5)$$

in which  $s_{\pm} = \alpha \pm i\omega$  are the roots of the quadratic equation:

$$s^2 - 2\omega_n \zeta s + \omega_n^2 = 0, \quad (1.6)$$

which is called the *characteristic equation* of the ODE. Those roots will then be:

$$s_{\pm} = \omega_n \left( \zeta \pm i\sqrt{1 - \zeta^2} \right), \quad (1.7)$$

which are then essentially damped frequencies. Note that if  $\zeta > 1$  then the roots will both be real, which is the case of overdamping.

Solutions to the problem of forced oscillations are usually best treated with the method of Laplace transforms, and sinusoidal forcing functions tend to exhibit the phenomenon of *resonance*, which makes the amplitude of the oscillation a maximum when the frequency of the forcing function equals the natural frequency of oscillation.

2. The forced, undamped, one-dimensional, anharmonic oscillator. – For this oscillator, one extends the constitutive law (1.1) by the addition of a cubic term, which is based upon the assumption that the force  $F(x)$  is an odd function of  $x$ :

$$F = -kx + bx^3. \quad (1.8)$$

Such a force is still conservative and its associated potential function takes the form:

$$U(x) = \frac{1}{2} kx^2 - \frac{1}{4} bx^4, \quad (1.9)$$

which is then an even function of  $x$ .

The differential equation of such an oscillator takes the form:

$$\frac{d^2x}{dt^2} + \omega_n^2 x - \frac{b}{m} x^3 = a(t), \quad (1.10)$$

which is nonlinear then.

The solutions to the initial-value problem for such an equation are more involved than simple sinusoids and generally involve elliptic functions.

3. The physical pendulum. – A physical pendulum of length  $l$  and mass  $m$  has a restoring force that is given by the torque that acts upon it due to gravity, which is assumed to be constant:

$$\tau = -mgl \sin \theta, \quad (1.11)$$

in which  $\theta$  is the angle that the pendulum makes with the  $z$ -axis.

Although this constitutive law is nonlinear, for small enough angles ( $\theta \ll 1$  rad.),  $\sin \theta$  can be approximated by  $\theta$ , which will then yield a simple harmonic oscillator. The differential equation then takes the form:

$$\frac{d^2\theta}{dt^2} + \omega_n^2 \sin \theta = 0. \quad (1.12)$$

in which the natural frequency is now:

$$\omega_n = \sqrt{\frac{g}{l}}, \quad (1.13)$$

if one assumes that the mass  $m$  is concentrated at the end of the pendulum, so the moment of inertia about the pivot is  $m l^2$ ; more generally,  $l$  must be replaced with the radius of gyration of the pendulum about its pivot.

Once again, the solutions to the initial-value problem for this nonlinear ODE are generally given by elliptic functions.

*c. Coupling of oscillators.* – The type of coupling between neighboring oscillators is a property of the medium more than a property of the oscillators themselves. The two main issues that affect the type of coupling are:

1. Discreteness vs. continuity of the medium
2. Dispersion vs. the absence of dispersion.
3. Causality.

Either discreteness or continuity can be regarded as an approximation to the other one. For instance, although one might think of a compressible gas as composed of discrete molecules and an elastic crystal lattice as being composed of discrete atomic ions, nonetheless, when it comes to the propagation of mechanical waves in those media, it is often more convenient to resort to a continuum model. Conversely, when one goes looking for numerical solutions to the continuum wave equations in real-world situations

(e.g., fluids, antennas), it is usually unavoidable that one must resort to discrete models for the purposes of computer software design. Hence, it is important to understand the dual relationship between both approximations.

The coupling of oscillators at different space-time points is typically assumed to be at the infinitesimal level; i.e., a given oscillator is coupled to only the other oscillators that are “infinitesimally close” to it. That sort of picture, which involves partial differential equations, can often be obtained by starting with a system of ordinary differential equations for a spatial-distributed, discrete system of finitely-separated oscillators that are coupled in some tangible way (e.g., connecting springs to the nearest neighbors) and passing to the limit as the number of lattice points becomes infinite and their separations become infinitesimal.

It is heuristically illuminating to see how the systems of ODE’s for the oscillation of an indefinite sequence of masses  $m(i)$  at points  $x(i)$  along the real line that are coupled by springs with spring constants  $k(i-1, i)$  and  $k(i, i+1)$  to the nearest neighbors becomes the one-dimensional, linear wave equation when one passes to the continuum limit, in which  $m(i)$  becomes a linear mass density  $\rho(x)$ ,  $k(i, j)$  becomes a function of  $x$ , and the sum of the forces that act on each mass becomes proportional to the second partial derivative of the amplitude of oscillation with respect to  $x$ . We illustrate this arrangement in Fig. II.1:

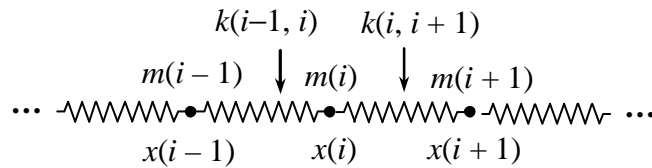


Figure II.1. An indefinite sequence of masses coupled by linear springs.

If one assumes Hooke’s law for the springs then the equations of motion for each mass point will be:

$$m(i) \frac{d^2 x(i)}{dt^2} = -k(i-1, i) x(i-1) + [k(i-1, i) + k(i, i+1)] x(i) - k(i, i+1) x(i+1). \quad (1.14)$$

If one regards the basic state of the system of masses as being defined by the indefinite vector  $X = (\dots, x(i-1), x(i), x(i+1), \dots)$  then one can think of (1.14) as being a very-large-dimensional system of second-order linear ODE’s for  $X$  of the form:

$$M\ddot{X} = -KX, \quad (1.15)$$

in which  $M = \text{diag}[\dots, m(i-1), m(i), m(i+1), \dots]$  and  $K$  is a “band” matrix. That is, its elements can be non-zero only on the diagonal and its nearest super-diagonal and sub-diagonal, which is typical of nearest-neighbor coupling for finite lattices.

The right-hand side of (1.14) takes the form of a second-order finite difference expression. Hence, in the limit as the spacing between masses becomes zero and the number of masses becomes infinite, it will become a second partial derivative with

respect to  $x$  times  $dx$ . However, since  $\partial x / \partial x$  is always unity, we need to replace  $x(i)$  with a coordinate  $x_0(i) + \xi(i)$ , where  $x_0(i)$  is the equilibrium position of  $x(i)$ , which does not change in time, and  $\xi(i)$  is the amplitude of the oscillation of  $x(i)$ , which does change in time. Ultimately, the continuum limit of (1.14) is:

$$\rho(x) \frac{\partial^2 \xi}{\partial t^2} = k(x) \frac{\partial^2 \xi}{\partial x^2} \quad [\xi = \xi(t, x)], \quad (1.16)$$

which is the linear, one-dimensional, wave equation when the speed of propagation of (longitudinal) waves is:

$$c(x) = \sqrt{\frac{k(x)}{\rho(x)}}. \quad (1.17)$$

If the state of an oscillator at one point in time depends upon its states at some finitely-separated previous time points then there will be *temporal dispersion*, while if the state of an oscillator at one point in space depends upon its state at some finitely-neighboring points of space then there will be *spatial dispersion*. In other words, dispersion is a form of “memory” in the medium. It has the effect of making the constitutive map into an integral operator, instead of an algebraic one. In the discrete lattice model, spatial dispersion would correspond to each mass being coupled to more than just its nearest neighbors.

Causality in the coupling of oscillators not only includes the idea that the state of an oscillator at a given space-time point cannot be influenced by the states of the oscillators at any future time points, but also that the finiteness of the speeds at which waves propagate implies that there will also be space-time points in the past and present that cannot affect the state of a given oscillator. The boundaries of causality ultimately come down to the dispersion law for the waves, which we shall discuss in due course below.

*c. State of a wave medium.* – A *state* of a wave medium  $M$  is a differentiable association of a state of an elementary oscillator with each point of  $M$ . For us, if  $\mathfrak{D}$  is the basic state space for the elementary oscillators then it will be sufficient to define that association by a differentiable function  $\psi: M \rightarrow \mathfrak{D}$  or its graph  $\psi: M \rightarrow M \times \mathfrak{D}$ ,  $x \mapsto (x, \psi(x))$  <sup>(1)</sup>. In the event that the space  $\mathfrak{D}$  is a vector space, one can think of the “zero section” of the projection  $M \times \mathfrak{D} \rightarrow M$ ,  $(x, \psi) \mapsto x$ , which makes  $\psi(x) = 0$  for every  $x$ , as essentially the “quiescent” state.

To make contact with the conventional terminology regarding waves, we too shall refer to  $\psi$  as a *wave function*. That is, the purpose of a wave function is to define the

---

<sup>(1)</sup> If one chooses to replace  $M \times \mathfrak{D}$  with the total space  $B$  of a fiber bundle  $B \rightarrow M$  then the state would be a global section of its projection – if one exists – or perhaps a local one, more generally. Conceivably, the lack of a Euclidian topology on  $M$  might relate to topological defects in the medium, such as lattice defects, vortices, and even the sources of waves themselves. However, we shall try to avoid that discussion in what follows, since we are trying to stay closer to a survey of the existing literature of continuum-mechanical models for quantum wave mechanics, which did not go into those matters.

state of a wave medium  $M$  to be the association of an oscillator amplitude with each point of  $M$ .

*d. Disturbances in the state of a wave medium.* – We shall think of a *disturbance* as something that alters the state of a wave medium. Hence, some of the issues that pertain to them would be:

1. Impulsive vs. continuous.
2. Finite vs. infinitesimal.

An *impulsive* disturbance is generally represented by a jump discontinuity in the state at some time point, such as when one plucks a tensed, elastic string or drops a stone into a pond. It has the effect of defining the initial values for the time evolution of the state on some discontinuity surface. Hence, one can think of the subsequent dynamics of the motion as being “driven by the initial conditions.” This approach to defining waves was developed in great detail by Hadamard [8] and others that followed him (e.g., [9]).

A *continuous* disturbance is a higher-dimensional analogue of a forcing function for an oscillator; the signals that are broadcast from antennas can take that form, at least for some finite length of time. Of course, one must keep in mind that the elementary oscillators are already being forced by the motion of the neighboring oscillators via the coupling law, so one might distinguish between *internal* and *external* forcing functions on the state of the medium, accordingly.

We shall think of continuous disturbances as possibly being composed of discrete pulse trains, rather than invent a separate category for recurring impulsive disturbances. Of course, there is a fundamental distinction to be made by pulse trains that have a low enough frequency that successive disturbances might damp out completely in between pulses, and ones of high enough frequency that the previous disturbances will still be affecting the state of the medium when the next one comes about.

The difference between finite and infinitesimal disturbances of the state of the medium brings us back to the difference between finite and infinitesimal deformations, in effect. We shall eventually address this in the context of variational field theory, but for now, we point out that the finite disturbances take the form of differentiable, one-parameter families  $\psi_s$  of states  $\psi_s : (a, b) \times M \rightarrow \mathfrak{D}$ ,  $(s, x) \mapsto \psi_s(x)$ ; in this case, differentiability means that the curve  $\psi_s(x)$  in  $\mathfrak{D}$  that one obtains by fixing each  $x$  is a differentiable curve. Typically, the parameter interval  $(a, b) \subset \mathbb{R}$  will include 0, so the state  $\psi_0(x)$  will represent an unperturbed state.

An *infinitesimal disturbance*  $\delta\psi$  is then obtained by differentiating  $\psi_s$  with respect to  $s$  – say, at 0:

$$\delta\psi(x) = \left. \frac{d\psi_s(x)}{ds} \right|_{s=0}. \quad (1.18)$$

It then associates a tangent vector  $\delta\psi(x)$  to  $\mathfrak{D}$  to each point  $\psi_0(x)$  in  $\mathfrak{D}$ .

Typically, most linear wave equations are the result of applying infinitesimal disturbances to the state of a wave medium.

*e. Response to a disturbance.* – The *response* of a wave medium to a disturbance in its state is simply a time-varying state of that medium that is first excited from its quiescent state when the disturbance becomes active. Thus, the wave function  $\psi: M \rightarrow \mathfrak{D}$ ,  $(t, x) \mapsto \psi(t, x)$  can serve to describe either the state of the wave medium or its response to a disturbance, such as an impulsive disturbance at  $t = 0$  that basically defines the Cauchy data for its subsequent time evolution; hence, the response is the solution to the wave equation for that Cauchy problem.

**2. The kinematical state of a wave.** – We can now define a *wave* to be the response of a wave medium to a disturbance in its state, and assume that it takes the form of a wave function  $\psi$ , so we now have some hope of defining the kinematical state of a wave. Basically, we shall take the position that if a wave in a medium  $M$  is described by a  $C^k$  function  $\psi: M \rightarrow \mathfrak{D}$ ,  $(t, x) \mapsto \psi(t, x)$  then the  $k^{\text{th}}$ -order *kinematical state* of the wave is a section of the source projection  $J^k(M, \mathfrak{D}) \rightarrow M$  of the manifold of  $k$ -jets of wave functions. Hence, for each  $(t, x) \in M$ , it will take the form:

$$s(t, x) = (t, x^i, \psi(t, x), \psi_t(t, x), \psi_i(t, x), \dots, \psi_{i_1 \dots i_k}(t, x)). \quad (2.1)$$

The kinematical state will be integrable iff it is the  $k$ -jet prolongation of  $\psi$ :

$$s = j^k \psi. \quad (2.2)$$

Since the wave equations that we shall be considering in this book are all second-order partial differential equations in the wave function (except for the Dirac equation, which is a first-order *system*), we shall usually be dealing with only first-order kinematical states. Hence, the kinematical state will take the form:

$$s(t, x) = (t, x^i, \psi(t, x), \psi_i(t, x), \psi_i(t, x)) \quad (i = 1, \dots, n), \quad (2.3)$$

and it will be integrable iff:

$$\psi_t = \frac{\partial \psi}{\partial t}, \quad \psi_i = \frac{\partial \psi}{\partial x^i} \quad (2.4)$$

for all  $(t, x)$  in  $M$ .

When one is dealing with wave motions, it is usually more convenient to regard the coordinates of a point  $x$  as  $x^\mu$ , with  $x^0 = t$  <sup>(1)</sup>, and the components of vectors and covectors as also having temporal component that is given the index 0. Hence, the coordinates of  $J^1(M, \mathfrak{D})$  are  $(x^\mu, \psi, \psi_\mu)$ , so  $s$  can be given the form:

---

<sup>(1)</sup> The reason that we are setting  $x^0 = t$ , and not  $ct$ , is that the use of  $c$  as a units conversion constant would imply that we were dealing with a specific type of wave medium (viz., linear, isotropic, homogeneous, non-dispersive, electromagnetic) in order for  $c$  to be meaningful in the context of  $\phi$ .

$$s(t, x) = (t, x^\mu, \psi(t, x), \psi_\mu(t, x)), \quad (2.5)$$

and it will be integrable iff:

$$\psi_\mu = \frac{\partial \psi}{\partial x^\mu}. \quad (2.6)$$

*a. The frequency-wave number 1-form.* – So far, there is nothing to say that  $\psi$  is any different from any other field on  $M$  that takes its values in the field space  $\mathfrak{D}$ . In order to narrow it down to something more “wave-like,” we shall assume that  $\mathfrak{D}$  is a vector space with scalars in a field  $\mathbb{K} = \mathbb{R}, \mathbb{C}$ , and that  $\psi$  takes the form of a product:

$$\psi(t, x) = A(t, x) \Theta(t, x), \quad (2.7)$$

in which  $A : M \rightarrow \mathfrak{D}$  is an *amplitude function*, and  $\Theta : M \rightarrow \mathbb{K}$  is a *phase function*. Often,  $A$  is time-invariant, so it will define the shape of the wave envelope, and  $\mathbb{K} = \mathbb{C}$  with:

$$\Theta = e^{i\theta}, \quad (2.8)$$

where  $\theta$  is a real function on  $M$ ; its level surfaces are called *isophases*. Hence, one thinks of the elementary oscillator that lives in  $\Theta$  at  $(t, x)$  as being something whose oscillation relates to the phase function  $\Theta$ .

Some particular forms that (2.7) can take are:

1. Standing waves:  $A = A(x), \quad \Theta = \Theta(t).$
2. Traveling waves:  $A = A(x), \quad \Theta = \Theta(x^i - v^i t).$
3. Geometrical optics:  $A = A(x), \quad \theta = \theta(t, x),$

although one also imposes an approximation in this last case that relates to the derivatives of  $\theta$ , namely, that the absolute values of the partial derivatives of  $\theta$  are much larger than those of  $A$ ; one calls that the *high-frequency* or *small wave length* approximation.

4. Plane waves:  $A = \text{const.}, \quad \theta = k_\mu x^\mu = \omega t - k_i x^i \quad (k_\mu = \text{const.})$
5. Cylindrical waves:  $A = \text{const.} / r, \quad \theta = \omega t - k_r r \quad (\omega, k_r = \text{const.})$

In this case, the other coordinates of space are  $x^2 = \vartheta, x^3 = z$ , where  $\vartheta$  is either the azimuth or right ascension.

6. Spherical waves:  $A = \text{const.} / r^2, \quad \theta = \omega t - k_r r \quad (\omega, k_r = \text{const.})$

In this case, the other coordinates of space are  $x^2 = \vartheta, x^3 = \zeta$ , where  $\zeta$  is the codeclination.

With the special form (2.7), the kinematical state of  $\psi$  can be replaced by:

$$s(t, x) = (t, x^i, A(t, x), \theta(t, x), A_\mu(t, x), k_\mu(t, x)) \quad (\mu = 0, 1, \dots, n), \quad (2.9)$$

and it will be integrable iff:

$$A_\mu = \frac{\partial A}{\partial x^\mu}, \quad k_\mu = \frac{\partial \theta}{\partial x^\mu} \quad (2.10)$$

for all  $(t, x)$  in  $M$ .

One can construct a 1-form  $\alpha$  with values in  $\mathfrak{D}$  out of the component functions  $A_\mu$ , and a single 1-form  $k$  (with values in  $\mathbb{R}$ ):

$$\alpha = A_\mu dx^\mu, \quad k = k_\mu dx^\mu, \quad (2.11)$$

so the integrability of the kinematical state would be equivalent to the exactness of these 1-forms:

$$\alpha = dA, \quad k = d\theta. \quad (2.12)$$

The 1-form:

$$k = \omega dt - k_i dx^i \quad (2.13)$$

shall be referred to as the *frequency-wave number* 1-form, since its components are  $\omega$ , which represents the frequency of the local oscillator, and  $k_i$ , which represents the wave number of the wave at that point. When  $k$  is integrable (i.e., exact), one will then have:

$$\omega = \frac{\partial \theta}{\partial t}, \quad k_i = \frac{\partial \theta}{\partial x^i}. \quad (2.14)$$

Of course, unless the function  $\theta$  is truly periodic in  $t$  and each  $x^i$ , the interpretation of these expressions as a frequency and three wave numbers, respectively, is somewhat debatable. Fortunately for the people who are content to work with linear wave equations, that is not generally an issue, since to them everything comes back to the behavior of plane waves, for which the temporal and spatial periodicity is a basic property.

If one divides the frequency  $\omega$  by the wave numbers  $k_i$  then one will get the components of a spatial covector with the units of velocity, namely, the *phase velocity* of the wave:

$$v_p^i = \frac{\omega}{k_i}. \quad (2.15)$$

However, if one considers the process of going from the four dimensions of space-time to the three dimensions of space then one will see that there is something more geometrically natural about making the inverse definition, which will give components with the units of indices of refraction:

$$n_i = \frac{k_i}{\omega}. \quad (2.16)$$



This will also produce a spatial covector, not a spatial vector.

What makes this more geometrically natural is the fact that the map that takes  $(\omega, k_i)$  to  $(n_i)$  is precisely the map that takes the homogeneous coordinates of a point in  $\mathbb{RP}^{3*}$  to its inhomogeneous coordinates in one of the four Plücker coordinate systems. Since this is not the only time that relativistic physics makes this transformation (it is also how one gets from four-velocities to three-velocities), one begins to suspect that there is more to the idea that “space” should really mean “projective space” in physics. The author has done considerable research along those lines, but we shall not go further into the details here, but only refer to [10] and the references therein.

*b. Dispersion laws.* – Since waves are always assumed to propagate with finite speeds (even though the speed might vary in time, position, direction, frequency-wave number, and polarization), it is never true that every 1-form  $k$  is physically admissible. Rather, they are always confined to some hypersurface in each cotangent space  $T_x^*M$ , and more to the point, the hypersurfaces are usually algebraic in character. The algebraic function  $\mathcal{D}(k)$  is generally homogeneous of even degree  $d$ , and more to the point, a homogeneous polynomial of degree  $d$ , although the equation itself:

$$\mathcal{D}(k) = k_0^d, \quad (2.17)$$

which is called the *dispersion law* for the class of waves in question, can be homogeneous or inhomogeneous.

Often, one finds the equation (2.17) solved for  $k_0 = \omega$

$$\omega = \alpha(k_i). \quad (2.18)$$

One can think of a dispersion law as something that relates the coupling of neighboring oscillators, since it converts a local frequency for the oscillator at a point into a wave number for the propagation of a wave in each direction, and vice versa. Indeed, in many linear cases, the dispersion law is simply the *symbol* of the second-order linear differential operator that defines the wave equation that results from the coupling of oscillators in space-time. When that operator has constant coefficients, one can think of that symbol as being the Fourier transform of the operator, which can also be obtained by replacing all partial derivatives as follows:

$$\frac{\partial^n}{\partial t^n} \rightarrow (i \omega)^n, \quad \frac{\partial^n}{\partial (x^i)^n} \rightarrow (i k_i)^n.$$

Note that this also leads to half of the usual canonical quantization rules for quantum mechanics. One can then think of the dispersion law as something like the frequency-wave number space analogue of the space-time coupling.

The common way of obtaining a dispersion law in theory is to start with a set of field equations, choose a certain class of “wave-like” solutions (such as plane waves), and see what algebraic relationship ensues when one does the differentiations of the wave

function and combines them in the algebraic manner that the field equations specify. Of course, it is only in the linear case that it would be sufficient to examine the dispersion law for plane waves, since even though wave functions can (by hypothesis) always be analyzed into their Fourier components, unless the field equations are linear in their derivatives, the dispersion law that one would get for a different class of solutions would not have to be the same as the one for plane waves.

When one has defined a dispersion law, one can also define a second kind of wave velocity in terms of the derivatives of  $\mathcal{D}$  with respect to  $k_\mu$  that one calls the *group velocity* [11]:

$$v_g^i = - \frac{\mathcal{D}^i}{\mathcal{D}^0} = - \frac{\partial \mathcal{D} / \partial k_i}{\partial \mathcal{D} / \partial \omega}. \quad (2.19)$$

When  $\mathcal{D}$  has been solved for  $\omega$  in the form (2.18), that will give:

$$v_g^i = \frac{\partial \omega}{\partial k_i}. \quad (2.20)$$

Some examples of dispersion laws are:

1. Light cones:

$$\mathcal{D}(k) = \omega^2 - c^2 \delta^{ij} k_i k_j = 0. \quad (2.21)$$

This leads to the most elementary “linear” dispersion law:

$$\omega = \pm c \kappa \quad (\kappa = \sqrt{k_1^2 + k_2^2 + k_3^2}). \quad (2.22)$$

One sees that  $c$  is basically the norm of the phase velocity in this case, which is still (2.15):

$$c = \pm \frac{\omega}{\kappa} = \pm \sqrt{(v_p^1)^2 + (v_p^2)^2 + (v_p^3)^2}. \quad (2.23)$$

This makes it clear that the only thing that the sign refers to is the direction of propagation of the waves (which is the line through  $k_i$ ), and we thus omit the sign unless it is essential.

The fact that there is only one  $c$  points to the isotropy of the medium, and if it is also a constant then the medium would be homogeneous, as well. Of course, the optics and acoustics of refraction would become trivial in that case, although most commonly the spatial change in  $c$  is assumed to be a jump discontinuity across a boundary in the context of refraction.

Since:

$$\mathcal{D}^0 = 2\omega, \quad \mathcal{D}^i = -2c^2 \delta^{ij} k_j, \quad (2.24)$$

the group velocity will have components:

$$v_g^i = \pm c \hat{k}^i, \quad (2.25)$$

in which  $\hat{k}^i = \delta^{ij} k_j / \kappa$  is the unit vector in the direction  $k^i$ ; hence,  $v_g$  is also equal to  $c$ . As a result, one will always have:

$$v_p^i v_{gi} = \frac{\omega}{k_i} \frac{ck_i}{\kappa} = c^2. \quad (2.26)$$

One can complete the spatial group velocity  $v_g^i$  to a four-velocity that lies on the dispersion hypersurface (i.e., light cone) by using the dispersion law, which will determine  $v^0$  uniquely (up to sign). If  $v_g$  is the Euclidian norm of  $v_g^i$  then one defines:

$$v^0 = \pm \frac{v_g}{c}. \quad (2.27)$$

One of the key properties of this type of dispersion law is that it has no “rest system.” That is, one cannot find a frame to which a non-zero  $k$  can be referred that will make  $k_i = 0$  for all  $i$ , since that would make  $\omega = 0$ , as well.

2. Massive matter waves:

$$\mathcal{D}(k) = \omega^2 - c^2 \delta^{ij} k_i k_j = \omega_0^2. \quad (2.28)$$

We shall encounter this in the context of the Klein-Gordon equation, but there are also some plasmas that present such a dispersion law for the propagation of electromagnetic waves.

This dispersion law can also be put into the form:

$$\omega = \sqrt{\omega_0^2 + \delta^{ij} c^2 k_i k_j}. \quad (2.29)$$

(We have omitted the sign ambiguity, since it relates to only the direction of propagation.) Whereas all values of  $\omega$  were allowable in the homogeneous case, one now sees that as  $k_i$  ranges from  $-\infty$  to  $+\infty$ , the corresponding values of  $\omega$  will fall between  $\omega_0$  and  $+\infty$ .

This type of dispersion law is distinguished from the previous one by the fact one can now define a rest frame, since  $k_i = 0$  no longer implies that  $\omega = 0$ , but only that:

$$\omega = \omega_0. \quad (2.30)$$

Hence, one can characterize  $\omega_0$  as the frequency of some basic oscillator in the rest system for which the wave will be a standing wave.

Although the phase velocity of these waves has not changed in form from  $\omega / k_i$  (since the definition of  $v_p$  is indifferent to the choice of dispersion law), nonetheless, it does not have to equal  $c$ , anymore,

$$v_p = \frac{\sqrt{\omega_0^2 + c^2 \kappa^2}}{\kappa} = c \sqrt{1 + \left(\frac{v_0}{c}\right)^2} \quad (v_0 \equiv \frac{\omega_0}{\kappa}). \quad (2.31)$$

In fact, as  $v_0$  ranges over all real values,  $v_p$  will range from  $c$  to  $+\infty$ .

Furthermore, the group velocity has changed:

$$v_g^i = \frac{\partial \omega}{\partial k_i} = \frac{c^2}{v_p} \hat{k}^i = c \left(1 + \left(\frac{v_0}{c}\right)^2\right)^{-1/2} \hat{k}^i, \quad (2.32)$$

although (2.26) will still be true. This time,  $v_0$  can range from 0 to  $+\infty$ , but  $v_g$  can range from 0 to only  $c$ . Therefore, one expects that it is the group velocity that is more physically meaningful.

Once again, one can complete  $v_g$  to a four-velocity on the dispersion hypersurface [we change the temporal component to  $u_0$  to avoid confusion with the  $v_0$  in (2.31)]:

$$u_0 = \pm \sqrt{1 + \left(\frac{v_g}{c}\right)^2}. \quad (2.33)$$

3. Bimetric dispersion laws:

$$\mathcal{D}(k) = (g^{\kappa\lambda} k_\kappa k_\lambda) (\bar{g}^{\mu\nu} k_\mu k_\nu) = k^2 \bar{k}^2 = 0, \quad (2.34)$$

where both scalar products  $g$  and  $\bar{g}$  on the cotangent spaces have the same signature type, namely, the Lorentzian one  $\eta^{\mu\nu} = \text{diag}[+1, -c^2, -c^2, -c^2]$ . In the simplest case, they are both diagonal in the same frame, but with different values of  $c$ :

$$\mathcal{D}(k) = (\omega^2 - c^2 \delta^{ij} k_i k_j) (\omega^2 - \bar{c}^2 \delta^{mn} k_m k_n) = 0. \quad (2.35)$$

When  $c = \bar{c}$ , the dispersion law will reduce to the degenerate case  $(\eta^{\mu\nu} k_\mu k_\nu)^2 = 0$ .

One sees that since the polynomial factors into a product of quadratic polynomials, and their product must be zero, the zero locus of  $\mathcal{D}(k)$  will consist of the union of two (possibly intersecting) light cones. Therefore,  $k$  will lie on one light cone or the other one, and possibly both.

Since:

$$\frac{\partial \mathcal{D}}{\partial k_\rho} = 2 [(\eta^{\rho\lambda} k_\lambda)(\bar{\eta}^{\mu\nu} k_\mu k_\nu) + (\eta^{\kappa\lambda} k_\kappa k_\lambda)(\bar{\eta}^{\rho\nu} k_\nu)], \quad (2.36)$$

which we rewrite in the form:

$$\frac{\partial \mathcal{D}}{\partial k_\rho} = 2 (\bar{k}^2 k^\rho + k^2 \bar{k}^\rho), \quad (2.37)$$

with the obvious definitions, the group velocity will take the form:

$$v_g^i = -\frac{\partial \mathcal{D} / \partial k_i}{\partial \mathcal{D} / \partial \omega} = \frac{c^2 \bar{k}^2 k^i + \bar{c}^2 k^2 \bar{k}^i}{\bar{k}^2 \omega + k^2 \bar{\omega}} = \frac{c^2}{v_p \left(1 + \frac{k^2 \bar{\omega}}{\bar{k}^2 \omega}\right)} \hat{k}^i + \frac{\bar{c}^2}{\bar{v}_p \left(1 + \frac{\bar{k}^2 \omega}{k^2 \bar{\omega}}\right)} \hat{\bar{k}}^i. \quad (2.38)$$

In particular, the two extreme cases are when  $k$  lies on the first light cone ( $k^2 = 0$ ), which will make  $v_g = c^2 / v_p$ , and when  $k$  lies on the other one ( $\bar{k}^2 = 0$ ), which will make  $v_g = \bar{c}^2 \bar{v}_p$ .  $v_g$  will become undefined on the intersection of the two light cones, since it will take the form of  $0 / 0$ .

*c. The eikonal equation.* – Whenever one is given a function, such as  $\mathcal{D}$ , on the cotangent bundle  $T^*M$ , one can always define a first-order partial differential equation for the phase function  $\theta$  by evaluating  $\mathcal{D}$  on  $k = d\theta$ :

$$\mathcal{D}(\partial_\mu \theta) = \omega_0^d. \quad (2.39)$$

For instance, in the Lorentzian case, the homogeneous equation:

$$g^{\mu\nu}(x) \frac{\partial \theta}{\partial x^\mu} \frac{\partial \theta}{\partial x^\nu} = 0 \quad (2.40)$$

is referred to as the *eikonal equation*, and it is very important in geometrical optics.

As one can see, that partial differential equation for  $\theta$  is first-order and nonlinear. Finding solutions to it usually involves appealing to Hadamard's method of bicharacteristics, which turns into Hamilton's equations.

*d. Virtual displacements of the kinematical state.* – If  $s : M \rightarrow J^1(M, \mathfrak{D})$  is a kinematical state of a wave function  $\psi : M \rightarrow \mathfrak{D}$  that takes the coordinate form (2.3) then a virtual displacement of that state will be a vector field  $\delta s(t, x)$  on  $s$  that takes the component form:

$$\delta s(t, x) = \delta x^\mu(t, x) \frac{\partial}{\partial x^\mu} + \delta \psi(t, x) \frac{\partial}{\partial \psi} + \delta \psi_\mu(t, x) \frac{\partial}{\partial \psi_\mu}, \quad (2.41)$$

and the virtual displacement will be integrable iff:

$$\delta \psi_\mu = \frac{\partial(\delta \psi)}{\partial x^\mu}. \quad (2.42)$$

If the kinematical state takes the form (2.9) then  $\delta s$  will take the form (we drop the functional dependency of the components, for brevity):

$$\delta s = \delta x^\mu \frac{\partial}{\partial x^\mu} + \delta A \frac{\partial}{\partial A} + \delta \theta \frac{\partial}{\partial \theta} + \delta A_\mu \frac{\partial}{\partial A_\mu} + \delta k_\mu \frac{\partial}{\partial k_\mu}, \quad (2.43)$$

and its integrability will be equivalent to:

$$\delta A_\mu = \frac{\partial(\delta A)}{\partial x^\mu}, \quad \delta k_\mu = \frac{\partial(\delta \theta)}{\partial x^\mu}. \quad (2.44)$$

**3. The dynamical state of a wave.** – The approach that we shall take to defining the dynamical state of a wave is to regard it as dual to the virtual displacement of the kinematical state under the duality that is defined by virtual work.

*a. Fundamental 1-form.* – If  $s : M \rightarrow J^1(M, \mathfrak{D})$  is a kinematical state of a wave  $\psi$ , and  $\delta s$  is a virtual displacement of that state then the dynamical state of  $\psi$  will be a 1-form  $\phi$  on  $J^1(M, \mathfrak{D})$  that we call the *fundamental 1-form* that couples the virtual displacement  $\delta s$  to the virtual work  $\delta \mathcal{W}^o$  that is performed in the process:

$$\delta \mathcal{W}^o = \phi(\delta s). \quad (3.1)$$

When  $s$  is expressed in the coordinate form (2.3),  $\phi$  will have the component form:

$$\phi = P dt + f_i dx^i + \zeta d\psi + \tilde{\varepsilon} d\psi_t + \tilde{\pi}^i d\psi_i. \quad (3.2)$$

All of the component functions must be understood to mean functions on  $J^1(M, \mathfrak{D})$ , such as  $f_i(t, x^i, \psi, \psi_t, \psi_i)$ , for instance. Hence, the mechanical constitutive laws for the waves in the medium  $M$  will already be included in the functional dependency of the components of  $\phi$ .

If the elements of  $\mathfrak{D}$  have “amplitude units,” which we abbreviate by AU, then the components of  $\phi$  will have the following interpretations:

1.  $P$  Power density.
2.  $f_i$  External force density.
3.  $\zeta$  Energy density per AU.
4.  $\tilde{\varepsilon}$  Energy density per time rate of change of AU.
5.  $\tilde{\pi}^i$  Energy density per AU gradient.

Usually, it is more convenient to write  $\phi$  in four-dimensional form without distinguishing spatial components from temporal ones:

$$\phi = f_\mu dx^\mu + \zeta d\psi + \tilde{\pi}^\mu d\psi_\mu. \quad (3.3)$$

The most common way of obtaining a fundamental 1-form is to start with a Lagrangian density  $\mathcal{L}$  on  $J^1(M, \mathfrak{D})$  (which we will discuss in the next chapter) and define  $\phi$  to be:

$$\phi = d\mathcal{L}. \quad (3.4)$$

If  $\mathcal{L} = \mathcal{L}(x^\mu, \psi, \psi_\mu)$  then that will make:

$$f_\mu = \frac{\partial \mathcal{L}}{\partial x^\mu}, \quad \zeta = \frac{\partial \mathcal{L}}{\partial \psi}, \quad \tilde{\pi}^\mu = \frac{\partial \mathcal{L}}{\partial \psi_\mu}. \quad (3.5)$$

When  $s$  is expressed in the coordinate form (2.9),  $\phi$  will have the component form:

$$\phi = f_\mu dx^\mu + \zeta_A dA + \zeta_\theta d\theta + \tilde{\pi}^\mu dk_\mu. \quad (3.6)$$

*b. Examples of fundamental 1-forms for waves.* – The simplest wave equation that one encounters in physics is the massless, linear wave equation for a real wave function  $\psi$ :

$$0 = \square\psi = \eta^{\mu\nu} \psi_{,\mu,\nu}. \quad (3.7)$$

Since this can be associated with the Lagrangian density:

$$\mathcal{L}(\psi_\mu) = -\frac{1}{2} \eta^{\mu\nu} \psi_{,\mu} \psi_{,\nu}, \quad (3.8)$$

the fundamental 1-form that one associates with it will be:

$$\phi = d\mathcal{L} = -(\eta^{\mu\nu} \psi_{,\nu}) d\psi_{,\mu}; \quad (3.9)$$

i.e.:

$$\tilde{p}^\mu = -\eta^{\mu\nu} \psi_{,\nu}, \quad (3.10)$$

which is minus the four-dimensional gradient of  $\psi$ .

When  $\psi$  is complex-valued, the field space  $\mathfrak{D}$  will be two-real-dimensional, so if one wishes to make  $\mathcal{L}$  real-valued then one must regard it as a function of both  $\psi$  and  $\psi^*$  and use the modulus function to turn  $\psi$  and  $\psi^*$  into a real number, namely:

$$\mathcal{L}(\psi_\mu, \psi_\mu^*) = -\eta^{\mu\nu} \psi_{,\mu} \psi_{,\nu}^*, \quad (3.11)$$

which will make:

$$\phi = -(\eta^{\mu\nu} \psi_{,\nu}^*) d\psi_{,\mu} - (\eta^{\mu\nu} \psi_{,\nu}) d\psi_{,\mu}^*; \quad (3.12)$$

i.e.:

$$\tilde{p}^\mu = -\eta^{\mu\nu} \psi_{,\nu}^*, \quad \tilde{p}^{\mu*} = -\eta^{\mu\nu} \psi_{,\nu}. \quad (3.13)$$

If one wishes to go on to the massive, linear (i.e., Klein-Gordon) wave equation for a complex wave function, namely:

$$0 = \square\psi + k_0^2\psi, \quad (3.14)$$

then one must add another term to  $\mathcal{L}$ :

$$\mathcal{L}(\psi, \psi^*, \psi_{,\mu}, \psi_{,\mu}^*) = k_0^2 \psi\psi^* - \eta^{\mu\nu} \psi_{,\mu} \psi_{,\nu}^*. \quad (3.15)$$

One will then get:

$$\phi = k_0^2 \psi^* d\psi + k_0^2 \psi d\psi^* - (\eta^{\mu\nu} \psi_{,\nu}^*) d\psi_{,\mu} - (\eta^{\mu\nu} \psi_{,\nu}) d\psi_{,\mu}^*; \quad (3.16)$$

i.e.:

$$\zeta = k_0^2 \psi^*, \quad \zeta^* = k_0^2 \psi, \quad \tilde{p}^\mu = -\eta^{\mu\nu} \psi_{,\nu}^*, \quad \tilde{p}^{\mu*} = -\eta^{\mu\nu} \psi_{,\nu}. \quad (3.17)$$

In more elaborate field spaces,  $\mathfrak{D}$  must typically include some way of turning its elements into real numbers, such as a norm or scalar product when it is a vector space. We shall deal with those cases as they become relevant in what follows.

*c. Energy-momentum of a wave.* – In the mechanics of points and extended massive objects, where the target variable is the position of a point in space or space-time, one can also consider the kinematical state to be something that involves the velocity or displacement gradient, which we collectively describe by the coordinates  $x_a^\mu$ . The dual object to that kinematical object in the eyes of virtual work is then energy-momentum-stress:

$$\delta\mathcal{W}^\circ = p_\mu^a \delta x_a^\mu = p_\mu d v^\mu + \tau_i^b \delta e_b^i \quad (b = 1, \dots, p). \quad (3.18)$$

We have previously discussed two different ways of associating a spatial velocity vector with a wave, namely, its phase and group velocity, respectively. Indeed, one can then complete the spatial group velocity to a velocity four-vector on the light cone uniquely (up to sign), since if  $v_{gi}$  is given and its Euclidian norm is  $v_g$  then its temporal component can be only  $v_0 = \pm v_g / c$ .

We shall now show that if one uses the group covelocity  $v_{gi}$ , in particular, and associates a spatial momentum density 1-form with it in the conventional way:

$$p_i = \rho v_{gi}, \quad (3.19)$$

in which  $\rho$  is the mass density, then in the case of the most common dispersion law (i.e., the Lorentzian structure), for which  $v_{gi} = c^2 k_i / \omega$  one will get:

$$p_i = \frac{\rho c^2}{\omega} k_i, \quad (3.20)$$

and if one puts  $\rho$  into the form  $mn$ , where  $m$  is the integral of  $\rho$  over all space, and  $n = \rho / m$  is the number density of the matter, then if one also defines:



$$\hbar = \frac{mc^2}{\omega} \quad (\text{i.e., } mc^2 = \hbar\omega), \quad (3.21)$$

one can express  $p_i$  in the form:

$$p_i = \hbar n k_i. \quad (3.22)$$

Except for the inclusion of the number density  $n$ , this is essentially the de Broglie relation for linear momentum [12]. Of course, one must realize that the de Broglie relation couples the linear momentum of a *point mass* to the wave number of a wave, not the linear momentum density of an *extended mass*. Hence, one might consider that the constant  $\hbar$  is to  $\hbar n$  what  $m$  is to  $mn$ ; that is,  $\hbar$  is the integral of a density over all space.

If  $p$  denotes the norm of  $p_i$  then the energy density that will complete  $p_i$  to an energy-momentum density 1-form is (we omit the sign that comes from the square root) then derived from (3.20), (3.21), and the fact that  $\omega = c\kappa$  for this dispersion law:

$$\mathcal{E} = pc = \rho c^2 = \hbar n \omega. \quad (3.23)$$

Hence, we have also recovered the de Broglie relation for energy (density), as well.

Note that in the previous discussion the only actual *assumption* that one must introduce in order to be talking about quantum waves is that the constant  $\hbar$  that we have *defined* will be the same for all matter, and not something that varies from one wave to the next.

**4. Derivation of wave equations.** – We shall use the principle of virtual work as the basis for the derivation of wave equations, namely, that the virtual work that is done by any “allowable” virtual displacement, viz.:

$$\delta W^\circ(\delta\psi) = \int_M \phi(j^i \delta\psi) V, \quad (4.1)$$

must be zero. The way that we are using the term “allowable” is usually defined by the boundary conditions on the virtual displacement.

*a. The principle of virtual work.* – In order to derive a wave equation from the knowledge of  $\phi$ , one must apply it to an integrable virtual displacement  $\delta s = j^1 \delta\psi$  of the kinematical state of  $\psi$  and set the resulting total virtual work equal to zero. If  $\delta s$  has the form (2.41) and is integrable then that will give:

$$0 = \int_M \phi(j^i \delta\psi) V, \quad (4.2)$$

with:

$$\phi(j^1 \delta\psi) = f_\mu \delta x^\mu + \zeta \delta\psi + \tilde{p}^\mu \partial_\mu(\delta\psi), \quad (4.3)$$

and  $V$  is the volume element on  $M$ . [We have suppressed the specific reference to the fact that the function  $\phi(j^1\delta\psi)$  must be pulled down to  $M$  by a section of the source projection of  $J^1(M, \mathfrak{D})$ , for brevity.]

An application of the product rule for differentiation will put this into the form:

$$\phi(j^1\delta\psi) = f_\mu \delta x^\mu + (\zeta - \partial_\mu \tilde{p}^\mu) \delta\psi + \partial_\mu (\tilde{p}^\mu \delta\psi). \quad (4.4)$$

When one substitutes this in (4.2), the result will be:

$$0 = \int_M D^* \phi(\delta\psi) V + \int_M [f_\mu \delta x^\mu + \partial_\mu \tilde{p}^\mu (\delta\psi)] V, \quad (4.5)$$

in which we have defined:

$$D^* \phi = (\zeta - \partial_\mu \tilde{p}^\mu) d\psi, \quad (4.6)$$

or, for complex wave functions:

$$D^* \phi = (\zeta - \partial_\mu \tilde{p}^\mu) d\psi + (\zeta^* - \partial_\mu \tilde{p}^{\mu*}) d\psi^*. \quad (4.7)$$

In a sense, the operator  $D^*$  is “adjoint” to the Spencer operator, as it acts upon vector fields on  $J^1(M, \mathfrak{D})$ , at least with respect to the bilinear pairing that is defined by virtual work.

When the virtual displacement is vertical,  $\delta x^\mu = 0$ , which will make the second integral in (4.5) take the form:

$$\int_M [\partial_\mu \tilde{p}^\mu (\delta\psi)] V = \int_{\partial M} \tilde{p}^\mu (\delta\psi) \# \partial_\mu. \quad (4.8)$$

If (4.5) is to be true for all  $\delta\psi$  that make  $\tilde{p}^\mu (\delta\psi)$  vanish on  $\partial M$  (such as when  $\delta\psi$  itself vanishes on  $\partial M$ ) then one must have:

$$0 = D^* \phi, \quad (4.9)$$

which will give:

$$0 = \zeta - \partial_\mu \tilde{p}^\mu \quad (4.10)$$

in the real case, or:

$$0 = \zeta - \partial_\mu \tilde{p}^\mu, \quad 0 = \zeta^* - \partial_\mu \tilde{p}^{\mu*} \quad (4.11)$$

in the complex case.

*b. Examples.* – If we go back to the examples of fundamental 1-forms that were given in the last section then we will see that either equation (4.10) or equations (4.11) do, in fact, reproduce the wave equations. We summarize this in a table:

Table II.1 Examples of calculations for elementary wave types

Wave type	$\zeta, \zeta^*$	$\tilde{p}^\mu, \tilde{p}^{\mu*}$	$D^* \phi$
Linear, massless, real	0	$\tilde{p}^\mu = -\eta^{\mu\nu} \psi_{,\nu}$	$(\square\psi) d\psi$
Linear, massless, complex	0	$\tilde{p}^\mu = -\eta^{\mu\nu} \psi_{,\nu}^*$ $\tilde{p}^{\mu*} = -\eta^{\mu\nu} \psi_{,\nu}$	$(\square\psi^*) d\psi + (\square\psi) d\psi^*$
Linear, massive, complex	$\zeta = k_0^2 \psi^*$ , $\zeta^* = k_0^2 \psi$	$\tilde{p}^\mu = -\eta^{\mu\nu} \psi_{,\nu}^*$ $\tilde{p}^{\mu*} = -\eta^{\mu\nu} \psi_{,\nu}$	$(\square\psi^* + k_0^2\psi^*) d\psi$ $+ (\square\psi + k_0^2\psi) d\psi^*$

c. *Wave equations in amplitude-phase form.* – So far, we have not said anything much about the form that the wave equations would take when one expresses the kinematical state and dynamical state of the wave in amplitude-phase form. That is because that discussion essentially amounts to the essence of what follows in this book. In particular, the Madelung transformation of the Schrödinger and Klein-Gordon wave functions amounts to introducing polar coordinates in the complex plane, so the amplitude function will be real-valued. Thus, we shall not into the topic further at the moment, since we shall give it considerable attention in the balance of this book.

### References (\*)

1. D. H. Delphenich, "On the general structure of mathematical models for physical systems," *Ann. Phys. (Berlin)* **523** (2011), 965-989.
2. D. H. Delphenich, "On the pre-metric foundations of wave mechanics I: massless waves," *Ann. Phys. (Leipzig)* **18** (2009), 206-230, and [arXiv:0904.1960](https://arxiv.org/abs/0904.1960).
3. A.A. Andronov, E.A. Vitt, S.E. Khaikin, *Theory of Oscillators*, Dover, NY, 1966.
4. P. Hagedorn. *Nonlinear Oscillations*, Clarendon Press, Oxford, 1981.
5. P. D. Lax and R. S. Phillips, *Scattering Theory*, Academic Press, New York, 1967.
6. W. O. Amrein, J. M. Jauch, and K. B. Sinha, *Scattering Theory in Quantum Mechanics*, W. A. Benjamin, Reading, MA, 1977.
- 7\*. L. de Broglie, *La thermodynamique de la particule isolée*, Gauthier-Villars, Paris, 1964.
- 8\*. J. Hadamard, *Leçons sur la propagation des ondes et les équations de l'hydrodynamique*, Chelsea, NY, 1949.
- 9\*. G. Juvet, "Mécanique analytique et mécanique ondulatoire," fasc. 39, *Mem. des Sci. Math.*, Gauthier-Villars, Paris, 1937.
10. D. H. Delphenich, "Projective geometry and special relativity," *Ann. Phys. (Leipzig)* **15** (2006), 216-246.

---

(\*) References that are marked with an asterisk are available in English translation at the author's website [neo-classical-physics.info](http://neo-classical-physics.info).

11. L. Brillouin, *Wave Propagation and Group Velocity*, Academic Press, New York, 1960.
  12. L. de Broglie –
    - a. “Recherches sur la théorie des quanta,” *Ann. de Phys.* (10) **3** (1925), 22-128; English translation by A. F. Kracklauer, 2004, published online.
    - b. *Ondes et mouvements*, Gauthier-Villars, Paris, 1926.
-

## CHAPTER III

### Variational field theory

This chapter is intended to merely summarize the basic notions and formulas from the calculus of variations that will be used in the context of fields that generally take the form of wave functions. Since the basic intent is to present a *calculus* of variations, and not an *analysis* of variations, the discussion will be more local and component-oriented than many modern theorists might prefer. That decision was made in the interests of staying closer to the published physics research that was being surveyed, since much of the justification for the introduction of the modern mathematical generalities is to make the field theory more sensitive to topological issues, which were not generally addressed in the books and papers that are in question here. More to the point, the generalization of a soluble problem might very well be insoluble, which is why science typically advances by successive incremental approximations, and not leaps of faith.

A similar justification will be made for the fact that we will be concerned with a flat background metric for space-time. Since one of the fundamental issues in the continuum-mechanical models for quantum wave mechanics is how the basic density function that one obtains from the quantum wave function changes the geometry of the space of kinematical states into the geometry of the space of dynamical states, it would distract from one's focus on that topic to also consider the way that the geometry is being altered by the presence of gravitating matter. For some theoretical purposes, it is better to regard the formal machinery of general relativity as being most appropriate to the study of physical phenomena in the presence of strong gravitational fields, such as one finds near neutron stars and black holes, rather than a universal background geometry on which to superimpose all physical phenomena, simply in the interests of mathematical generality.

**1. Variations of fields on space-time regions.** – Although it is often useful to think of the calculus of variations as something like “the calculus of infinity variables” or differential calculus on infinite-dimensional differentiable manifolds, for the practical business of deriving field equations, conserved currents, and equations of motion, that viewpoint is usually more heuristically useful as a tool for visualization than actually useful for the derivation of equations. That is because the analytical overhead that is associated with making all of one's statements mathematically rigorous rapidly turns the calculus of variations into the analysis of variations. Hence, although we shall make occasional motivating statements that suggest the infinite-dimensional picture, nonetheless, no attempt at making them analytically rigorous will be made, since the real objective in all of this is simply to obtain systems of differential equations that will tell us more about the nature of quantum wave mechanics.

*a. Finite variations of fields.* – If one imagines that  $\Psi$  is basically a point in the infinite-dimensional vector space  $\Gamma(S, \mathbb{C}^r)$  then one can also imagine that a *finite*

variation of  $\Psi$  is a differentiable curve  $\Psi(s)$  through that point; for convenience, assume that  $\Psi(0) = \Psi$ . Actually, since only the local behavior of the curve at  $\Psi$  will be of any interest to us, the range of the curve parameter  $s$  can be a small, but finite, interval  $(-\varepsilon, +\varepsilon)$  around 0. One can then represent the curve  $\Psi(s)$  as a curve in the graph of  $\Psi$ :  $(x(s), H(s, x))$ , where  $H : (-\varepsilon, +\varepsilon) \times \mathbb{C}^r \rightarrow \mathbb{C}^r$ ,  $(s, x) \mapsto H(s, x)$  is a differentiable map that has the property that:

$$H(0, x) = \Psi(x). \quad (1.1)$$

One can also think of this as a “differentiable homotopy” of the map  $\Psi$  to some unspecified final map from  $S$  to  $\mathbb{C}^r$ , but we shall skip over that fact. (The curious can confer Dedecker [1] or some of the author’s work [2-4].)

*b. Infinitesimal variations of fields.* – Just as one can think of differentiating a differentiable curve through a point in a finite-dimensional differentiable manifold at that point in order to obtain a tangent vector, one can think of differentiating the curve  $\Psi(s)$  at  $\Psi$  to obtain a “tangent vector” at  $\Psi$ . We shall formally write:

$$\delta\Psi \equiv \left. \frac{d\Psi(s)}{ds} \right|_{s=0} \quad (1.2)$$

and refer to  $\delta\Psi$  as a *variation* of  $\Psi$ , which is short for *infinitesimal variation* of  $\Psi$ . However, it is more convenient to regard the “tangent vector at  $\Psi$ ” that we have just derived as actually a *vector field* on the graph of  $\Psi$  in the space  $S \times \mathbb{C}^r$ , whose coordinates are  $(x^i, z^a)$ :

$$\delta\Psi(x) = \delta x^i(x) \frac{\partial}{\partial x^i} + \delta\Psi^a(x) \frac{\partial}{\partial z^a}. \quad (1.3)$$

Hence, the vector field  $\delta\Psi(x)$  is defined only on all points of  $S$  and the corresponding image points of  $\Psi$  in  $\mathbb{C}^r$ . One can then also think of  $\delta\Psi(x)$  as the infinitesimal generator of a one-parameter family of fields  $\Psi(s, x)$  on  $S$  (and its deformations), at least locally (i.e., for a sufficiently-small  $|s|$ ).

A variation, as we have defined it here, is then a generalization of the virtual displacement of the configuration of an extended body that we defined previously.

The variation  $\delta\Psi(x)$  will be called *vertical* iff  $\delta x^i(x) = 0$ ; it will then take the form:

$$\delta\Psi(x) = \delta\Psi^a(x) \frac{\partial}{\partial z^a}. \quad (1.4)$$

In such a case, the finite variation that is generated by the variation will affect only the values of  $\Psi$  in field space.

As mentioned before, the techniques of jet manifolds are directly applicable to the formulation of the calculus of variations, since a Lagrangian density is basically a

differentiable function on a jet manifold. For us, since the wave equations in question will generally be second-order, it will be necessary to consider only 1-jets  $j^1\Psi$  of differentiable maps  $\Psi : S \rightarrow \mathbb{C}^r$ . Thus, the coordinate representation of a point of  $J^1(S, \mathbb{C}^r)$  will then look like  $(x^i, z^a, z^{a*}, z_i^a, z_i^{a*})$ , a section of the source projection  $s : S \rightarrow J^1(S, \mathbb{C}^r)$  will then have values that look like  $(x^i, s^a(x), s^{a*}(x), s_i^a(x), s_i^{a*}(x))$ , and it will be integrable iff there is some differentiable map  $\Psi : S \rightarrow \mathbb{C}^r$  such that:

$$s_i^a(x) = \Psi_{,i}^a(x), \quad s_i^{a*}(x) = \Psi_{,i}^{a*}(x). \quad (1.5)$$

Actually, for our immediate purposes, the real issue is the integrability of a vector field on (the graph of)  $s$ ; i.e., a variation:

$$\delta s(x) = \delta x^i(x) \frac{\partial}{\partial x^i} + \delta s^a(x) \frac{\partial}{\partial z^a} + \delta s^{a*}(x) \frac{\partial}{\partial z^{a*}} + \delta s_i^a(x) \frac{\partial}{\partial z_i^a} + \delta s_i^{a*}(x) \frac{\partial}{\partial z_i^{a*}}. \quad (1.6)$$

$\delta s(x)$  will be called *integrable* iff:

$$\delta s_i^a(x) = \delta s_{,i}^a(x) = \frac{\partial(\delta s^a)}{\partial x^i}(x) \quad (\text{and complex conjugate}). \quad (1.7)$$

Since  $\delta s^a = \delta s_i^a dx^i$  and  $\delta s^{a*} = \delta s_i^{a*} dx^i$  will then become a set of  $2r$  exact 1-forms, it will be necessary that one must also have that they are closed (viz.,  $d \wedge \delta s^a = d \wedge \delta s^{a*} = 0$  for all  $a$ ). Locally, that says:

$$\delta s_{i,j}^a = \delta s_{j,i}^a, \quad \delta s_{i,j}^{a*} = \delta s_{j,i}^{a*}. \quad (1.8)$$

**2. The stationary action principle.** – In finite-dimensional differential calculus, the theory of extrema of differentiable functions is essentially identical with the theory of critical points of those functions; i.e., points at which the differential of the function vanishes. In the calculus of variations, one has an analogous picture, in which the “differentiable function” on the “differentiable manifold” takes the form of a differentiable *functional* on the objects that are being varied by the variations (fields, for us). The “critical points” of that functional are then extremal objects, and the differential equations that define those extremal objects will define the equations of motion or equilibrium for the objects in question.

*a. The action functional.* – The first significant step towards obtaining the field equations is that of replacing the differentiable function on the infinite-dimensional vector space of fields  $\Psi : S \rightarrow \mathbb{C}^r$  whose critical points should determine extremal fields

with a differentiable function  $\mathcal{L}$  on the finite-dimensional space  $J^1(S, \mathbb{C}^r)$ . One calls that function  $\mathcal{L}$  a *Lagrangian density* for the fields  $\Psi$ , and it takes the coordinate form:

$$\mathcal{L} = \mathcal{L}(x^i, z^a, z_i^a).$$

When one composes  $\mathcal{L}$  with the 1-jet prolongation  $j^1\Psi$  of a field  $\Psi$ , one will get a real function  $j^1\Psi^*\mathcal{L}$  on  $S$ :

$$(j^1\Psi^*\mathcal{L})(x) = \mathcal{L}(j^1\Psi^*(x)) = \mathcal{L}(x^i, \Psi^a(x), \Psi_i^a(x)) \quad (2.1)$$

that one calls the *pull-back* of  $\mathcal{L}$  to  $S$  by  $j^1\Psi$ .

In practice, since  $\Psi$  is assumed to take complex values, in order to get a real function for  $\mathcal{L}$ , one will regard  $\mathcal{L}$  as a function of both  $\Psi$  and its complex conjugate  $\Psi^*$ , as well as its first derivatives:

$$\mathcal{L} = \mathcal{L}(x^i, \Psi^a(x), \Psi^{*a}(x), \Psi_i^a(x), \Psi_i^{*a}(x)). \quad (2.2)$$

If:

$$V = dx^1 \wedge \dots \wedge dx^n = \frac{1}{n!} \varepsilon_{i_1 \dots i_n} dx^{i_1} \wedge \dots \wedge dx^{i_n} \quad (2.3)$$

is a volume element on  $S$  (which is assumed to be orientable) then one can integrate the  $n$ -form  $(j^1\Psi^*\mathcal{L})V$  over  $S$  and obtain the *action functional* for the fields  $\Psi$  and  $\Psi^*$ :

$$\mathcal{S}[V, \Psi, \Psi^*] = \int_S \mathcal{L}(j^1\Psi)V = \int_S \mathcal{L}(x^i, \Psi^a(x), \Psi^{*a}(x), \Psi_i^a(x), \Psi_i^{*a}(x))V. \quad (2.4)$$

That construction will then allow one to circumvent the necessity of first defining the infinite-dimensional complex vector space in which  $\Psi$  and  $\Psi^*$  live, as long one can also define the differential of the functional  $\mathcal{S}$  accordingly.

*b. The first variation of the action functional.* – The differential  $df$  of a function  $f: M \rightarrow \mathbb{R}$  on a finite-dimensional differentiable manifold  $M$  will define a linear functional  $df|_x$  on the tangent vectors at each point  $x$  of that manifold. In the present (infinite-dimensional) case, where the points are  $\Psi$  and  $\Psi^*$ , and the function  $f$  is the functional  $\mathcal{S}$ , the tangent vectors will be variations  $\delta\Psi$  and  $\delta\Psi^*$ . The differential of  $\mathcal{S}$  will be then referred to as the *first variation of  $\mathcal{S}$* , and one defines it by (we suppress the reference to  $\Psi$  and  $\Psi^*$  for brevity):

$$\delta\mathcal{S}[\delta V, \delta\Psi, \delta\Psi^*] = \int_S [\delta\mathcal{L}(j^1\Psi)V + \mathcal{L}(j^1\Psi)\delta V], \quad (2.5)$$

in which:



$$\delta\mathcal{L}(j^1\Psi) = \frac{\partial\mathcal{L}}{\partial x^i} \delta x^i + \frac{\partial\mathcal{L}}{\partial\Psi^a} \delta\Psi^a + \frac{\partial\mathcal{L}}{\partial\Psi^{*a}} \delta\Psi^{*a} + \frac{\partial\mathcal{L}}{\partial\Psi^a_{,i}} \delta\Psi^a_{,i} + \frac{\partial\mathcal{L}}{\partial\Psi^{*a}_{,i}} \delta\Psi^{*a}_{,i} \quad (2.6)$$

and:

$$\delta V = d \wedge \# \delta x, \quad \# \delta x = \delta x^i i_{\partial_i} V = (-1)^{i+1} \delta x^i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n. \quad (2.7)$$

In this last expression, the caret signifies that the relevant differential is missing from the exterior product. One can then show that:

$$\delta V = \# \operatorname{div}(\delta x) = \operatorname{div}(\delta x) V. \quad (2.8)$$

As long as  $\delta\mathcal{L}$  is integrable, one can apply the product rule (AKA: integration by parts) and put  $\delta\mathcal{L}$  into the form:

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial x^i} \delta x^i + \frac{\delta\mathcal{L}}{\delta\Psi^a} \delta\Psi^a + \frac{\delta\mathcal{L}}{\delta\Psi^{*a}} \delta\Psi^{*a} + \frac{\partial}{\partial x^i} (\Pi_a^i \delta\Psi^a + \Pi_a^{*i} \delta\Psi^{*a}), \quad (2.9)$$

in which we have defined the *variational derivatives* of  $\mathcal{L}$  with respect to  $\Psi$  and  $\Psi^*$ :

$$\frac{\delta\mathcal{L}}{\delta\Psi^a} = \frac{\partial\mathcal{L}}{\partial\Psi^a} - \frac{\partial}{\partial x^i} \frac{\partial\mathcal{L}}{\partial\Psi^a_{,i}}, \quad \frac{\delta\mathcal{L}}{\delta\Psi^{*a}} = \frac{\partial\mathcal{L}}{\partial\Psi^{*a}} - \frac{\partial}{\partial x^i} \frac{\partial\mathcal{L}}{\partial\Psi^{*a}_{,i}}, \quad (2.10)$$

as well as their *canonical momenta*:

$$\Pi_a^i = \frac{\partial\mathcal{L}}{\partial\Psi^a_{,i}}, \quad \Pi_a^{*i} = \frac{\partial\mathcal{L}}{\partial\Psi^{*a}_{,i}}. \quad (2.11)$$

If we also define the canonical forces:

$$f^a = \frac{\partial\mathcal{L}}{\partial\Psi^a}, \quad f^{*a} = \frac{\partial\mathcal{L}}{\partial\Psi^{*a}} \quad (2.12)$$

then we can express the variational derivatives in the form:

$$\frac{\delta\mathcal{L}}{\delta\Psi^a} = f^a - \partial_i \Pi_a^i, \quad \frac{\delta\mathcal{L}}{\delta\Psi^{*a}} = f^{*a} - \partial_i \Pi_a^{*i}. \quad (2.13)$$

One can similarly apply the product rule to the term  $\mathcal{L}(j^1\Psi) \delta V$ :

$$\mathcal{L}(j^1\Psi) \delta V = \mathcal{L}(j^1\Psi) d \wedge \# \delta x = d \wedge [\mathcal{L}(j^1\Psi) \# \delta x] - d_x \mathcal{L}(j^1\Psi) \wedge \# \delta x, \quad (2.14)$$

in which the notation  $d_x$  means that one differentiates only with respect to the  $x^i$ . However, the last term will then take the form:

$$d_x \mathcal{L}(j^1\Psi) \wedge \# \delta x = \left( \frac{\partial \mathcal{L}}{\partial x^i} \delta x^i \right) V, \quad (2.15)$$

which will cancel the first one in (2.9). Hence, it does not seem to matter whether  $\mathcal{L}$  does or does not depend upon  $x^i$  to begin with, since that dependency will not appear in the final expression for the first variation.

If one assumes that the region  $S$  has a boundary  $\partial S$  then an application of Stokes's theorem for differential  $n$ -forms will put the first variation of  $\mathcal{L}$  into the form:

$$\begin{aligned} & \delta \mathcal{S}[\delta x^i, \delta \Psi, \delta \Psi^*] \\ &= \int_S \left( \frac{\delta \mathcal{L}}{\delta \Psi^a} \delta \Psi^a + \frac{\delta \mathcal{L}}{\delta \Psi^{*a}} \delta \Psi^{*a} \right) V + \int_{\partial S} [\Pi_a^i \delta \Psi^a + \Pi_a^{*i} \delta \Psi^{*a} + \mathcal{L}(j^1\Psi) \delta x^i] \# \partial_i. \end{aligned} \quad (2.16)$$

The rest of what we shall consider in the name of variational field theory will follow from various specializations of this expression.

By definition, an extremal field  $\Psi$  will be one for which  $\delta \mathcal{S}[\delta x^i, \delta \Psi, \delta \Psi^*]$  vanishes for every possible choice  $\delta \Psi$  (and therefore, its conjugate  $\delta \Psi^*$ ). One will then have to contend with the vanishing of the integrand in the first integral, as well as the vanishing of the boundary contribution. The former condition will imply field equations for  $\Psi$  and  $\Psi^*$ , while the latter will involve their boundary conditions.

The idea that the extremal fields should represent the physically-meaningful ones specializes to *Hamilton's principle* or *the stationary action principle*<sup>(1)</sup>. Whether or not that extremal actually represents a minimum then reverts to the methods of the second variation of the action functional, which we shall not go into here.

*c. The Euler-Lagrange equations.* – For the purposes of obtaining field equations from the stationary action principle, one generally makes some restrictions on the allowable variations of  $\Psi$  and  $\Psi^*$  so that they will always make the boundary integral vanish. There are two basic ways that this condition can be satisfied:

1. *The fixed-boundary condition:* Allowable variations  $\delta x^i$ ,  $\delta \Psi^a$ , and  $\delta \Psi^{*a}$  always vanish on  $\partial S$ . Hence, one varies the fields only at their interior points.

$$\delta x^i(x) = \delta \Psi^a(x) = \delta \Psi^{*a}(x) = 0, \quad x \in \partial S. \quad (2.17)$$

---

<sup>(1)</sup> Some authors, especially those of the Nineteenth Century, prefer to specialize the stationary action principle even further in order to state Hamilton's principle by applying it to a more specific class of Lagrangians.

2. *The transversality condition:* Allowable variations  $\delta\Psi^a$  and  $\delta\Psi^{*a}$  are always transversal to the canonical momenta at the boundary points, while  $\delta x^i$  fixes the boundary:

$$\delta x^i(x) = \Pi_a^i \delta\Psi^a(x) = \Pi_a^{*i} \delta\Psi^{*a}(x) = 0, \quad x \in \partial S. \quad (2.18)$$

With those restrictions, the first variation of  $\mathcal{L}$  will take the reduced form:

$$\delta\mathcal{S}[\delta\Psi, \delta\Psi^*] = \int_S \left( \frac{\delta\mathcal{L}}{\delta\Psi^a} \delta\Psi^a + \frac{\delta\mathcal{L}}{\delta\Psi^{*a}} \delta\Psi^{*a} \right) V. \quad (2.19)$$

This will vanish for every allowable  $\delta\Psi$  and  $\delta\Psi^*$  iff:

$$\frac{\delta\mathcal{L}}{\delta\Psi^a} = 0, \quad \frac{\delta\mathcal{L}}{\delta\Psi^{*a}} = 0 \quad (2.20)$$

for the  $\Psi$  and  $\Psi^*$  in question. These equations, which are each called the *Euler-Lagrange equations*, then take on the character of systems of differential equations for the extremal fields. In the case of point mechanics, for which  $S$  is an interval along the time line, they are ordinary differential equations, but when  $S$  has a dimension that is higher than one, they will be partial differential equations.

A common specialization of these equations is to the *static* case, for which nothing is a function of time, and  $S = \Sigma^3$ . The extremal fields then take on the character of equilibrium states, and the Euler-Lagrange equations are the equations of equilibrium.

One notes that with the introduction of canonical forces and momenta, the Euler-Lagrange equations can be expressed in the “ $F = ma$ ” form:

$$f^a = \partial_i \Pi_a^i, \quad f^{*a} = \partial_i \Pi_a^{*i}. \quad (2.21)$$

When we get into the examples of how this gets applied to wave equations, we shall see that this is actually somewhat illusory, since in many cases, the forces can get split between the two sides of the equation.

*d. Extension of the stationary action principle.* – To continue the calculus of infinity variables analogy, we point out that the first-variation functional  $\delta\mathcal{S}$  can be regarded as an exact 1-form on an infinite-dimensional differentiable manifold, since it is basically the differential of a 0-form, namely, the action functional. However, one can define that 1-form more generally in a manner that does not make it exact from the outset and still be talking about physically-meaningful concepts. The first variation 1-form  $\delta\mathcal{S}$  will then become a sort of infinite-dimensional equivalent of a conservative force 1-form.

The generalization of the first variation that one defines is essentially the virtual work functional that associates an increment of virtual work  $\mathcal{W}[\delta X]$  with the virtual

displacement  $\delta X$ . In place of the Lagrangian density  $\mathcal{L}$ , one starts with the *fundamental 1-form* on  $J^1(M, N) = (x^i, y^a, y_i^a)$ :

$$\phi = P_i dx^i + f_a dy^a + \Pi_i^a dy_i^a. \quad (2.22)$$

When  $\phi$  is exact (say,  $\phi = d\mathcal{L}$ ), the generalized forces  $f_a$  and momenta  $\Pi_i^a$  will agree with the previous ones, while the power densities  $P_i$  will also be expressible as partial derivatives:

$$P_i = \frac{\partial \mathcal{L}}{\partial x^i}, \quad f_a = \frac{\partial \mathcal{L}}{\partial y^a}, \quad \Pi_i^a = \frac{\partial \mathcal{L}}{\partial y_i^a}. \quad (2.23)$$

Of course, the 1-form  $\phi$  can also be defined in the case of non-conservative forces, as well as conservative ones, and even for non-holonomic constraints, unlike the first variation of the action functional.

The virtual work functional that one defines then amounts to:

$$\mathcal{W}[\delta X] = \int_S [\phi(\delta X) \cdot X] V, \quad (2.24)$$

in which  $X : M \rightarrow N$  is a differentiable map, and  $\delta X$  is the variation of its values in  $N$ .

As long as  $\phi$  is vertical ( $P_i = 0$ ) and the variation  $\delta X$  is integrable – i.e.:

$$\delta X = \delta X^a \frac{\partial}{\partial y^a} + \partial_i (\delta X^a) \frac{\partial}{\partial y_i^a}, \quad (2.25)$$

one can still define equations of motion or equilibrium by the vanishing of the functional  $\mathcal{W}[\delta X]$  for all allowable  $\delta X$ , which is basically d'Alembert's principle or the principle of virtual work. We shall not be using that extension for our present purposes, but it represents an important extension in scope of variational methods. The author has published several papers [2-4] on this rarely-mentioned aspect of the calculus of variations, and we refer the curious to those papers for a deeper discussion of it.

**3. Symmetries of the action functional.** – Now let us return to the general expression (2.16) for the first variation of the action functional for general variations of the fields  $\Psi$  and  $\Psi^*$ .

For extremal fields, and regardless of the constraints that were imposed upon the variations in order to obtain the extremal equations, the first variation will take the form:

$$\delta \mathcal{S}[\delta x^i, \delta \Psi, \delta \Psi^*] = \int_{\partial S} [\Pi_a^i \delta \Psi^a + \Pi_a^{*i} \delta \Psi^{*a} + \mathcal{L}(j^1 \Psi) \delta x^i] \# \partial_i. \quad (3.1)$$

If this expression vanishes then another application of Stokes's theorem will give:

$$\int_S d_\wedge \{ [\Pi_a^i \delta \Psi^a + \Pi_a^{*i} \delta \Psi^{*a} + \mathcal{L}(j^1 \Psi) \delta x^i] \# \partial_i \} = 0. \quad (3.2)$$

If this is to be true for every choice of  $S$  then one must have:

$$d_\wedge \{ [\Pi_a^i \delta \Psi^a + \Pi_a^{*i} \delta \Psi^{*a} + \mathcal{L}(j^1 \Psi) \delta x^i] \# \partial_i \} = 0 \quad (3.3)$$

for that particular variation  $\delta \Psi$  and extremal field  $\Psi$ .

Now, since the expression in square brackets in (3.3) is a set of functions, one can also express this last equation in the form:

$$\partial_i [\Pi_a^i \delta \Psi^a + \Pi_a^{*i} \delta \Psi^{*a} + \mathcal{L}(j^1 \Psi) \delta x^i] = 0; \quad (3.4)$$

i.e., the functions in the brackets are the components of a vector field on  $S$ :

$$\mathbf{J} = [\Pi_a^i \delta \Psi^a + \Pi_a^{*i} \delta \Psi^{*a} + \mathcal{L}(j^1 \Psi) \delta x^i] \frac{\partial}{\partial x^i} \quad (3.5)$$

that has vanishing divergence. One can then think of  $\mathbf{J}$  as a conserved current that is associated with the variation  $\delta \Psi$ .

If one defines  $\text{Var}(x^i, \Psi, \Psi^*)$  to be the infinite-dimensional vector space of all variations of the fields  $\Psi$  and  $\Psi^*$  then one will have a linear map  $J : \text{Var}(x^i, \Psi, \Psi^*) \rightarrow \mathfrak{X}(S)$ ,  $(\delta x^i, \delta \Psi, \delta \Psi^*) \mapsto \mathbf{J}(\delta x^i, \delta \Psi, \delta \Psi^*)$ . It is not one-to-one, since, for one thing, there will generally be many variations  $\delta \Psi^a$  and  $\delta \Psi^{*a}$  that are transverse to the canonical momenta  $\Pi_a^i$  and  $\Pi_a^{*i}$ .

In general, the variations  $\delta \Psi$  and  $\delta \Psi^*$  of the fields consist of two components: There is a vertical part –  $\delta_\vee \Psi$  or  $\delta_\vee \Psi^*$ , resp. – which is independent of the points of  $S$  (and is often called the “substantial variation” of  $\Psi$  or  $\Psi^*$ ) and a lift –  $\Psi_* \delta x$  or  $\Psi_*^* \delta x$ , resp. – of the variation  $\delta x$  of the points of  $S$ . Those lifts take the forms:

$$\Psi_* \delta x = d\Psi|_x(\delta x), \quad \Psi_*^* \delta x = d\Psi^*|_x(\delta x), \quad (3.6)$$

or, in components:

$$\Psi_*^a \delta x = \Psi_{*i}^a \delta x^i, \quad \Psi_*^{*a} \delta x = \Psi_{*i}^{*a} \delta x^i. \quad (3.7)$$

The vertical parts  $\delta_\vee \Psi$  or  $\delta_\vee \Psi^*$  of the variations of  $\Psi$  and  $\Psi^*$  are what is left when one subtracts the lifts of  $\delta x$  from the total variations  $\delta \Psi$  and  $\delta \Psi^*$ :

$$\delta_\vee \Psi = \delta \Psi - \Psi_* \delta x, \quad \delta_\vee \Psi^* = \delta \Psi^* - \Psi_*^* \delta x. \quad (3.8)$$

If variations  $\delta \Psi$  and  $\delta \Psi^*$  are replaced with their vertical parts then the expression (3.5) for the current that is associated with  $\delta x + \delta_\vee \Psi + \delta_\vee \Psi^*$  will now take the form:

$$J^i = T_j^i \delta x^j + \Pi_a^i \delta \Psi^a + \Pi_a^{*i} \delta \Psi^{*a}, \quad (3.9)$$

in which:

$$T_j^i = \mathcal{L} \delta_j^i - \Pi_a^i \partial_j \Psi^a - \Pi_a^{*i} \partial_j \Psi^{*a} \quad (3.10)$$

is referred to as the *canonical energy-momentum-stress tensor*; its doubly-covariant form  $T_{ij}$  does not generally have to be symmetric.

*a. Infinitesimal symmetries of the action functional.* – We refer to a general variation  $\delta s$ , as in (1.6), that makes  $\delta \mathcal{S}[\delta s]$  vanish for any extremal field  $s$  as an *infinitesimal symmetry of the action functional*. Such an infinitesimal symmetry is then associated with a conserved current (i.e., a vector field on  $S$  with vanishing divergence) by way of (3.5). If the variation  $\delta s$  is not an infinitesimal symmetry of the action functional (so  $\delta \mathcal{S}[\delta s] \neq 0$ ) then one can still define the vector field  $\mathbf{J}$  naively. The only thing that will change is that it will no longer need to have vanishing divergence.

One should distinguish between an infinitesimal symmetry  $\delta s$  of the action functional and an infinitesimal symmetry  $\delta \mathcal{L}$  of the Lagrangian density itself, for which:

$$0 = \delta \mathcal{L} = \delta x^i \frac{\partial \mathcal{L}}{\partial x^i} + \delta z^a \frac{\partial \mathcal{L}}{\partial z^a} + \delta z^{*a} \frac{\partial \mathcal{L}}{\partial z^{*a}} + \delta z_i^a \frac{\partial \mathcal{L}}{\partial z_i^a} + \delta z_i^{*a} \frac{\partial \mathcal{L}}{\partial z_i^{*a}}. \quad (3.11)$$

Although an infinitesimal symmetry of  $\mathcal{L}$  will always be an infinitesimal symmetry of  $\mathcal{S}$ , from Stokes's theorem, as long as the  $n$ -form  $(\delta \mathcal{L}) V$  is exact [say,  $(\delta \mathcal{L}) V = d \wedge \eta$ ], it will give rise to an  $n-1$ -form  $\eta$ , and if  $\eta$ , in turn, vanishes on  $\partial S$  then  $X$  will be an infinitesimal symmetry of  $\mathcal{S}$  that is not an infinitesimal symmetry of  $\mathcal{L}$ . One can also express this by saying that:

$$\delta \mathcal{L} = \text{div} \#^{-1} \eta. \quad (3.12)$$

The linear map  $\mathbf{J} : \text{Var}(x^i, \Psi, \Psi^*) \rightarrow \mathfrak{X}(S)$ ,  $(\delta x^i, \delta \Psi, \delta \Psi^*) \mapsto \mathbf{J}(\delta x^i, \delta \Psi, \delta \Psi^*)$  then has the property that if  $(\delta x^i, \delta \Psi, \delta \Psi^*)$  is an infinitesimal symmetry of the action functional then  $\mathbf{J}(\delta x^i, \delta \Psi, \delta \Psi^*)$  will have vanishing divergence. Both of the vector spaces  $\text{Var}(x^i, \Psi, \Psi^*)$  and  $\mathfrak{X}(S)$  have Lie algebras defined on them. When one restricts to vector fields on  $S$  with vanishing divergence, one will get the Lie algebra of infinitesimal generators of one-parameter families of volume-preserving diffeomorphisms, and the latter diffeomorphisms define an infinite-dimensional (Banach) Lie group.

*a. Noether's theorem.* – Although the association of conserved currents with infinitesimal symmetries of the action functional that was just described is more general in scope, the original paper by Emmy Noether [5] was confined to the association of conserved currents with infinitesimal symmetries that were due to the action of finite-dimensional Lie groups in the fields, so the variations themselves took the form of

fundamental vector fields for the group action. Since that is also how the physicists usually think of Noether's theorem, we shall discuss that restricted form of it.

As discussed above, there are actually two ways that  $G$  can act upon  $\mathbb{C}^r$ : Independently, and by way of mapping the action on  $S$  to one on  $\mathbb{C}^r$  using the differential of  $\Psi$  or  $\Psi^*$ ; i.e., the lift of the action. Hence, the fundamental vector field  $\tilde{\mathfrak{g}}$  on  $\mathbb{C}^r$  that corresponds to the action of an element  $g \in G$  will decompose into two parts again:

$$\tilde{\mathfrak{g}} = \tilde{\mathfrak{g}}_v + (\Psi_* \mathfrak{g}_x + \Psi^*_* \mathfrak{g}_x), \quad (3.13)$$

in which  $\tilde{\mathfrak{g}}_v$  is the vertical part and  $\Psi_* \mathfrak{g}_x + \Psi^*_* \mathfrak{g}_x$  is the sum of the lifts of the fundamental vector field  $\mathfrak{g}$  of the action of  $g$  on  $S$ .

The action of  $G$  on  $S$  will enter into only the definition of  $\delta x$ , while the representation of  $G$  in  $\mathbb{C}^r$  (more precisely, the representation of  $\mathfrak{g}$ ) will then enter into only the definitions of  $\delta\Psi$  and  $\delta\Psi^*$ .

**4. Examples of symmetries.** – We shall now discuss the specific examples of this situation that we shall be concerned with in all of the wave equations that follow.

*a. Phase invariance.* – One of the simplest symmetries of the action functional (which is also a symmetry of the Lagrangian density) that one can consider for complex wave functions is the possibility that when their values are expressed in polar coordinates, the choice of *global* phase origin will be irrelevant. Hence, one makes the finite replacements:

$$\Psi \mapsto e^{i\alpha} \Psi, \quad \Psi^* \mapsto e^{-i\alpha} \Psi^*,$$

in which  $\alpha$  is an arbitrary real constant.

In order to turn these into infinitesimal transformations – i.e., variations – one lets  $\alpha$  become a differentiable function  $\alpha(s)$  and differentiates the expressions at  $s = 0$ :

$$\delta\Psi = \left. \frac{d}{ds} \right|_{s=0} (e^{i\alpha(s)} \Psi) = i \alpha' \Psi, \quad \delta\Psi^* = \left. \frac{d}{ds} \right|_{s=0} (e^{-i\alpha(s)} \Psi^*) = -i \alpha' \Psi^*, \quad (4.1)$$

in which  $\alpha' = d\alpha/ds$  at  $s = 0$ .

Since this variation does not involve the points of  $S$  (i.e., it is vertical), the conserved current will take the form:

$$J^i = i\alpha(\Pi_a^i \Psi^a - \Pi_a^{*i} \Psi^{*a}), \quad (4.2)$$

in which we have dropped the otherwise irrelevant prime, for brevity.

As long as the divergence of  $J^i$  vanishes the inclusion of the constant  $\alpha$  will be superfluous since the divergence is a linear operator, and  $\alpha$  can be omitted.

*b. Scale invariance.* – Closely related to the concept of phase symmetry is that of *scale symmetry*, which represents a global rescaling of the space-time coordinates by a positive factor  $\lambda$ , combined with a global rescaling of the field space by the reciprocal of that factor:

$$\bar{x}^i = \lambda x^i, \quad \bar{\Psi}(\bar{x}) = \lambda^{-1} \Psi(\lambda x), \quad \bar{\Psi}^*(\bar{x}) = \lambda^{-1} \Psi^*(\lambda x). \quad (4.3)$$

Note that the space-time part of this transformation is not generally defined independently of a choice of coordinate system.

The associated variations will then take the form:

$$\delta x^i = x^i \delta \lambda, \quad \delta \Psi = (x^i \partial_i \Psi - \Psi) \delta \lambda, \quad \delta \Psi^* = (x^i \partial_i \Psi^* - \Psi^*) \delta \lambda. \quad (4.4)$$

If we suppress the explicit mention of the constant  $\delta \lambda$  then the Noether current that corresponds to this will be:

$$J^i = \mathcal{L} x^i - \Pi_a^i \Psi^a - \Pi_a^{*i} \Psi^{*a}. \quad (4.5)$$

If the action functional is indeed scale-invariant then one will have:

$$0 = \partial_i J^i = T_i^i + x^i \frac{\partial \mathcal{L}}{\partial x^i} - \Psi^a \frac{\partial \mathcal{L}}{\partial \Psi^a} - \Psi^{*a} \frac{\partial \mathcal{L}}{\partial \Psi^{*a}}. \quad (4.6)$$

Otherwise, that divergence, and with it, the right-hand side will not vanish.

In particular, for Lagrangian densities that are independent of  $x^i$ ,  $\Psi$ ,  $\Psi^*$ , one will have:

$$0 = \partial_i J^i = T_i^i; \quad (4.7)$$

i.e., the energy-momentum-stress tensor will be traceless. Such a Lagrangian density will then have the form  $\mathcal{L}(\Psi_i^a, \Psi_i^{*a})$ .

*c. Translational invariance.* – Translations are transformations that presumably act upon the points of space-time, and in fact, unless the space-time manifold actually has the structure of an affine space, such as  $\mathbb{R}^n$ , the action of the translation group for some dimension can usually have only a local character. Of course, since we are only going to consider regions  $S$  that belong to  $\mathbb{R}^n$ , that will not be an issue for us.

If the dimension of  $S$  is  $n$  then the action of  $\mathbb{R}^n$  on  $S$  by a translation  $a^i$  is simply to take a point  $x \in S$  that is described by coordinates  $x^i$  to the point  $x^i + a^i$ . Naturally, if one changes the coordinate system for  $x$  then one must change the coordinates for  $a^i$  accordingly. Of course, the translation of  $S$  does not have to still be  $S$  itself, and that will be the case only for the identity translation.

In order to get a fundamental vector field for  $a^i$ , one replaces  $a^i$  with a differentiable curve  $a^i(s)$  and differentiates  $x^i + a^i(s)$  at  $s = 0$ . If we let  $\varepsilon^i$  denote  $da^i / ds$  for  $s = 0$  then the variation of  $x$  will be the vector field  $\delta x(x)$  whose components are:



$$\delta x^i(x) = \varepsilon^i. \quad (4.8)$$

In particular, it is a constant vector field on  $S$ .

Since the translation group typically does not also act upon the field space, one usually sets:

$$\delta \Psi = \delta \Psi^* = 0. \quad (4.9)$$

Hence, the current that is associated with this type of variation will take the form:

$$J^i = T_j^i \varepsilon^j. \quad (4.10)$$

If the action functional is invariant with respect to translations then this current will have vanishing divergence, and since  $\varepsilon^i$  are constants, that would imply that:

$$\partial_i T_j^i = 0. \quad (4.11)$$

Otherwise, if translational symmetry is not present then one must have:

$$\partial_i T_j^i = f_j, \quad (4.12)$$

where  $f_j$  represent the components of the resultant of the external forces that account for the breakdown of the conservation of energy-momentum.

Explicitly, we have:

$$\partial_i T_j^i = \frac{\partial \mathcal{L}}{\partial x^j} - \frac{\partial \mathcal{L}}{\partial \Psi^a} \frac{\partial \Psi^a}{\partial x^j} - \frac{\partial \mathcal{L}}{\partial \Psi_i^a} \frac{\partial \Psi_j^a}{\partial x^j} - \frac{\partial \mathcal{L}}{\partial \Psi^{*a}} \frac{\partial \Psi^{*a}}{\partial x^j} - \frac{\partial \mathcal{L}}{\partial \Psi_i^{*a}} \frac{\partial \Psi_j^{*a}}{\partial x^j}, \quad (4.13)$$

in which we have used the field equations for  $\Psi$ ,  $\Psi^*$ . If the right-hand side vanishes then we must have:

$$\frac{\partial \mathcal{L}}{\partial x^j} = \frac{\partial \mathcal{L}}{\partial \Psi^a} \frac{\partial \Psi^a}{\partial x^j} + \frac{\partial \mathcal{L}}{\partial \Psi_i^a} \frac{\partial \Psi_j^a}{\partial x^j} + \frac{\partial \mathcal{L}}{\partial \Psi^{*a}} \frac{\partial \Psi^{*a}}{\partial x^j} + \frac{\partial \mathcal{L}}{\partial \Psi_i^{*a}} \frac{\partial \Psi_j^{*a}}{\partial x^j}, \quad (4.14)$$

which suggests that  $\mathcal{L}$  must take the form:

$$\mathcal{L} = \mathcal{L}(\Psi, \partial_i \Psi, \Psi^*, \partial_i \Psi^*). \quad (4.15)$$

In particular, it must be independent of  $x$ .

*d. Rotational invariance.* – Infinitesimal rotations are usually assumed to act upon both the points of space-time and the vectors in field space. One can also deal with those transformations by raising the lower index of the matrix that represents them:

$$\omega^{ij} = \delta^{ik} \omega_k^j. \quad (4.16)$$

The resulting matrix  $\omega^{ij}$  will also be antisymmetric in its indices:

$$\omega^{ij} = -\omega^{ji}. \quad (4.17)$$

Hence, we shall set the components of the fundamental vector field on  $S$  that comes from the action of  $\omega \in \mathfrak{so}(3; \mathbb{R})$  on  $\mathbb{R}^3$  by matrix multiplication equal to:

$$\delta x^i = \omega^{ij} x_j. \quad (4.18)$$

Similarly, the action of the group  $SO(3)$  on  $\mathbb{C}^r$  will result in a representation of its Lie algebra  $\mathfrak{D}: \mathfrak{so}(3) \rightarrow \mathfrak{gl}(r, \mathbb{C})$  in the Lie algebra of infinitesimal linear transformations of  $\mathbb{C}^r$ . If one chooses a basis  $\{\varepsilon_i, i = 1, 2, 3\}$  for  $\mathfrak{so}(3)$  and another  $\{E_a, a = 1, \dots, r\}$  for  $\mathfrak{gl}(r, \mathbb{C})$  then one can represent  $\mathfrak{D}$  by an  $r \times 3$  complex matrix  $\mathfrak{D}_i^a$ . However, if one is representing infinitesimal rotations by antisymmetric  $3 \times 3$  matrices then it will often be more convenient to represent the homomorphism  $\mathfrak{D}$  by a three-index array of the form  $\mathfrak{D}_{ij}^a$ . The way to get from elements of  $\mathfrak{so}(3)$ , when expressed as components  $\omega^k$  with respect to the basis  $\varepsilon_i$ , to elements of  $\mathfrak{so}(3)$ , when expressed as antisymmetric matrices  $\omega^{ij}$ , is by way of the “adjoint” map, whose matrices have the components:

$$[\varepsilon_{ij}]_k = \varepsilon_{ijk}, \quad k = 1, 2, 3, \quad (4.19)$$

so:

$$\omega^{ij} = \varepsilon^{ijk} \omega^k \quad (4.20)$$

and

$$\mathfrak{D}_{ij}^a = \varepsilon_{ijk} \mathfrak{D}_k^a. \quad (4.21)$$

As a result, the fundamental vector fields of the action of  $SO(3)$  on  $\mathbb{C}^r$  by way of the representation  $\mathfrak{D}$  will take the form:

$$\delta \Psi^a = \frac{1}{2} \mathfrak{D}_{ij}^a \omega^{ij}, \quad \delta \Psi^{*a} = \frac{1}{2} \mathfrak{D}_{ij}^{*a} \omega^{ij}. \quad (4.22)$$

When one substitutes (4.18) and (4.22) into the expression (3.9) for the Noether current that is associated with  $\omega^{ij}$ , one will get:

$$J^i = \frac{1}{2} (L_{jk}^i + S_{jk}^i) \omega^{jk}, \quad (4.23)$$

in which:

$$L_{jk}^i = T_j^i x_k - T_k^i x_j \quad (4.24)$$

represents the *orbital angular momentum* tensor, and:

$$S_{jk}^i = \Pi_a^i \mathcal{D}_{jk}^a + \Pi_a^{*i} \mathcal{D}_{jk}^{*a}, \quad (4.25)$$

is the *intrinsic angular momentum* – or *spin* – tensor.

As one can see, this kind of angular momentum has less to do with the rotation of things than it does with the way that the rotations get represented in the field space. It is the type of angular momentum that will be associated with quantum wave functions later on.

If the action functional has rotational symmetry (i.e., no external torques are present) then the vector field  $J^i$  will have zero divergence. Since  $\omega^{ij}$  is a constant matrix that would imply the equations of conservation of total angular momentum:

$$\partial_i (L_{jk}^i + S_{jk}^i) = 0, \quad (4.26)$$

which can also be written:

$$\partial_i L_{jk}^i = - \partial_i S_{jk}^i. \quad (4.27)$$

When one goes back to the expression (4.24) for orbital angular momentum, one will see that:

$$\partial_i L_{jk}^i = \partial_i T_j^i x_k - \partial_i T_k^i x_j + T_j^i \delta_{ik} - T_k^i \delta_{ij} = \partial_i T_j^i x_k - \partial_i T_k^i x_j + T_{ij} - T_{ji}. \quad (4.28)$$

Hence, if linear momentum is also conserved, one must have:

$$\partial_i L_{jk}^i = T_{jk} - T_{kj}. \quad (4.29)$$

That implies that if linear momentum is conserved then orbital angular momentum will also be conserved iff the canonical energy-momentum-stress tensor  $T_{ij}$  is symmetric. When one substitutes (4.29) in (4.27) that will give:

$$T_{jk} - T_{kj} = - \partial_i S_{jk}^i. \quad (4.30)$$

Hence, the symmetry of  $T_{jk}$  would also imply the conservation of spin when one has conservation of total angular momentum.

In the event that the action functional is not symmetric under rotations (i.e., external torques are present), one will have

$$\partial_i (L_{jk}^i + S_{jk}^i) = M_{jk}, \quad (4.31)$$

where  $M_{jk}$  represents the external torques (moments).

*e. Lorentz invariance.* – Lorentz transformations can be dealt with in close analogy to three-dimensional rotations since when one raises the lower index on the matrix of an infinitesimal Lorentz transformation:

$$\omega^{\mu\nu} = \eta^{\mu\kappa} \omega_{\kappa}^{\nu}, \quad (4.32)$$

one will again obtain an antisymmetric matrix:

$$\omega^{\mu\nu} = -\omega^{\nu\mu}. \quad (4.33)$$

In fact, this result is true for scalar products of any signature type. Of course, this time the matrix  $\omega^{\mu\nu}$  is  $4 \times 4$ , not  $3 \times 3$ .

The fundamental vector field that is associated with the action of  $\omega \in \mathfrak{so}(3, 1)$  on  $\mathbb{R}^4$  by matrix multiplication will then take the form:

$$\delta x^{\mu} = \omega^{\mu\nu} x_{\nu}. \quad (4.34)$$

This time, we write the matrix of the representation  $\mathfrak{D} : \mathfrak{so}(3, 1) \rightarrow \mathfrak{gl}(r; \mathbb{C})$  as  $\mathfrak{D}_{\mu\nu}^a$ , so the fundamental vector fields of the action of  $\mathfrak{so}(3, 1)$  on the field space by way of  $\mathfrak{D}$  will now take the forms:

$$\delta \Psi^a = \frac{1}{2} \mathfrak{D}_{\mu\nu}^a \omega^{\mu\nu}, \quad \delta \Psi^{*a} = \frac{1}{2} \mathfrak{D}_{\mu\nu}^{*a} \omega^{\mu\nu}. \quad (4.35)$$

One will then get a Noether current:

$$J^{\mu} = \frac{1}{2} (L_{\kappa\lambda}^{\mu} + S_{\kappa\lambda}^{\mu}) \omega^{\kappa\lambda}, \quad (4.36)$$

with:

$$L_{\kappa\lambda}^{\mu} = T_{\kappa}^{\mu} x_{\lambda} - T_{\lambda}^{\mu} x_{\kappa}, \quad S_{\kappa\lambda}^{\mu} = \Pi_a^{\mu} \mathfrak{D}_{\kappa\lambda}^a + \Pi_a^{*\mu} \mathfrak{D}_{\kappa\lambda}^{*a}. \quad (4.37)$$

If the action functional is Lorentz-invariant (so no external torques or sources of boosts are present), the current  $J^{\mu}$  will be conserved, and since  $\omega^{\mu\nu}$  is a constant matrix, one will get:

$$\partial_{\mu} (L_{\kappa\lambda}^{\mu} + S_{\kappa\lambda}^{\mu}) = 0 \quad (4.38)$$

or

$$\partial_{\mu} L_{\kappa\lambda}^{\mu} = -\partial_{\mu} S_{\kappa\lambda}^{\mu}, \quad (4.39)$$

in analogy with (4.26) and (4.27).

Since:

$$\partial_{\mu} L_{\kappa\lambda}^{\mu} = \partial_{\mu} T_{\kappa}^{\mu} x_{\lambda} - \partial_{\mu} T_{\lambda}^{\mu} x_{\kappa} + T_{\kappa\lambda} - T_{\lambda\kappa}, \quad (4.40)$$

if one also has conservation of energy-momentum, one will have:

$$\partial_{\mu} L_{\kappa\lambda}^{\mu} = T_{\kappa\lambda} - T_{\lambda\kappa} = -\partial_{\mu} S_{\kappa\lambda}^{\mu}, \quad (4.41)$$

in analogy with (4.29) and (4.30).

When the action functional is not Lorentz-invariant, one will have:

$$\partial_{\mu} (L_{\kappa\lambda}^{\mu} + S_{\kappa\lambda}^{\mu}) = M_{\kappa\lambda}, \quad (4.42)$$

in which  $M_{\kappa\lambda}$  represents the external torques (moments).

*f. The Belinfante-Rosenfeld theorem.* – If one raises the lower index on  $T_j^i$  to give  $T^{ij}$  then one can polarize  $T^{ij}$  into a symmetric and an antisymmetric part:

$$T^{ij} = T^{(ij)} + T^{[ij]}, \quad (4.43)$$

with:

$$T^{(ij)} = \frac{1}{2}(T^{ij} + T^{ji}), \quad T^{[ij]} = \frac{1}{2}(T^{ij} - T^{ji}). \quad (4.44)$$

If one raises both lower indices on  $S_{jk}^i$  then one can completely antisymmetrize the indices to produce the components of a 3-vector field:

$$\mathfrak{B}^{ijk} = S^{[ijk]} = \frac{1}{3}(S^{ijk} + S^{jki} + S^{kij}) \quad (4.45)$$

that some [6] call the *Belinfante-Rosenfeld* tensor field. Explicitly, one has:

$$\mathfrak{B}^{ijk} = \frac{1}{3}(\Pi_a^i \mathfrak{D}^{a,jk} + \Pi_a^j \mathfrak{D}^{a,ki} + \Pi_a^k \mathfrak{D}^{a,ij}). \quad (4.46)$$

The *Belinfante-Rosenfeld theorem* [7, 8] then says that:

$$T^{[jk]} = \partial_i \mathfrak{B}^{ijk}. \quad (4.47)$$

Actually, they also showed that if one is dealing with one's field theory in the context of general relativity, so the space-time metric  $g_{\mu\nu}$  is not constant, and the Lagrangian density is a function of it, one can also derive the symmetric part of  $T^{\mu\nu}$  from the formula:

$$T^{(\mu\nu)} = \frac{\partial \mathcal{L}}{\partial g_{\mu\nu}}. \quad (4.48)$$

This is very much in the spirit of Sahkarov's [9] concept of general relativity being a theory of "metric elasticity." If one thinks of  $g_{\mu\nu} = \eta_{\mu\nu} + E_{\mu\nu}$ , where  $E_{\mu\nu}$  is a finite strain tensor (which does not have to be infinitesimal as one usually assumes in the case of linearized gravitation), that would make:

$$\frac{\partial \mathcal{L}}{\partial g_{\mu\nu}} = \frac{\partial \mathcal{L}}{\partial E_{\mu\nu}}, \quad (4.49)$$

and the equation:

$$T^{(\mu\nu)} = \frac{\partial \mathcal{L}}{\partial E_{\mu\nu}} \quad (4.50)$$

would represent a mechanical constitutive law for a hyper-elastic medium (i.e., one for which such an  $\mathcal{L}$  exists).

**5. Gauge invariance.** – When one allows the global constant in a phase transition to become a differentiable function of space-time, one enters the realm of local phases, or *gauges*. There are two types of possible gauge transformations of the fields that enter into a Lagrangian density, which are called *gauge transformations of the first kind* and *gauge transformations of the second kind*, accordingly. Gauge transformations of the first kind are something of a generalization of phase transformations that become an issue whenever one couples the particle that is described by the wave function  $\Psi$  to an external electromagnetic field  $F$  by assuming that the particle is charged. Gauge transformations of the second kind pertain to the Lagrangian density of the electromagnetic field  $F$  alone. Hence, when one is dealing with only an electromagnetic field, but no wave functions, in effect, the first kind of transformations will play no role. Hence, we shall start with the gauge transformations of the second kind.

When one represents one of the Maxwell equations in the form:

$$d \wedge F = 0, \quad (5.1)$$

one can solve it (if only locally) with a 1-form  $A$  such that:

$$F = d \wedge A \quad (F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu), \quad (5.2)$$

but not uniquely. That is because any other 1-form  $A + \alpha$ , where  $d \wedge \alpha = 0$  (i.e.,  $\alpha$  is closed) would produce the same  $F$ . Since every closed 1-form is locally exact, one can also say that  $A$  can be replaced with  $A + d\lambda$  for any 0-form  $\lambda$ . The replacement of  $A$  with  $A + d\lambda$  is what one calls the *gauge transformation of the second kind*.

If one then regards the Lagrangian density  $\mathcal{L}(F)$  for the electromagnetic field as something that really takes the form  $\mathcal{L}(A, dA)$ , then a variation of  $\mathcal{L}$  will take the form:

$$\delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial A} \delta A + \frac{\partial \mathcal{L}}{\partial dA} \delta dA = \frac{\partial \mathcal{L}}{\partial A} \delta A + \frac{\partial \mathcal{L}}{\partial dA} d(\delta A). \quad (5.3)$$

One should note that from the second equation in (5.2), one will have:

$$\frac{\partial \mathcal{L}}{\partial A_{\mu,\nu}} = - \frac{\partial \mathcal{L}}{\partial F_{\mu\nu}}. \quad (5.4)$$

Actually, if one wishes to formulate Maxwell's second set of equations:

$$\partial_\mu F^{\mu\nu} = J^\nu \quad (5.5)$$

as the Euler-Lagrange equations for an action functional then one will find that one cannot avoid the introduction of the electromagnetic potential 1-form  $A$ . When one uses a gauge dependent action functional  $\mathcal{L}(A, dA)$  for the electromagnetic field, one can then express (5.5) in the form:

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial A_{\mu,\nu}} = -\partial_\mu \frac{\partial \mathcal{L}}{\partial F_{\mu\nu}} = \frac{\partial \mathcal{L}}{\partial A_\nu}. \quad (5.6)$$

Hence, one can use the Lagrangian density:

$$\mathcal{L}(A, dA) = \mathcal{L}(F) + \mathcal{L}(A) = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + A_\mu J^\mu, \quad (5.7)$$

and get:

$$\partial_\mu \frac{\partial \mathcal{L}(F)}{\partial F_{\mu\nu}} = \partial_\mu F^{\mu\nu}, \quad \frac{\partial \mathcal{L}}{\partial A_\nu} = J^\nu. \quad (5.8)$$

The Lagrangian density  $\mathcal{L}(F)$  is automatically invariant under the gauge transformations of the second kind:

$$F_\lambda \equiv d^\wedge (A + d\lambda) = d^\wedge A = F. \quad (5.9)$$

As for the Lagrangian density  $\mathcal{L}(A)$ , one has:

$$(A_\mu + \partial_\mu \lambda) J^\mu = A_\mu J^\mu + \partial_\mu \lambda J^\mu = A_\mu J^\mu + \partial_\mu (\lambda J^\mu),$$

since one must have the vanishing of  $\text{div } \mathbf{J}$  identically (it follows from the identical vanishing of  $d^\wedge d^\wedge$ ). When one integrates the second term on the final right-hand side over a space-time with no boundary, it will vanish, and that will imply that the invariance of  $\mathcal{L}(A, dA)$  under a gauge transformation of the second kind. One can also use the second equation in (5.8) as the definition of the conserved current that follows from gauge invariance of the second kind, as we shall do in what follows.

The canonical energy-momentum-stress tensor that is associated with  $\mathcal{L}$  in (5.7) is:

$$T_\nu^\mu = -\frac{1}{4} F_{\kappa\lambda} F^{\kappa\lambda} \delta_\nu^\mu + F^{\mu\kappa} F_{\kappa\nu} - J^\mu A_\nu, \quad (5.10)$$

which is often associated with the name Faraday, although it is somewhat dubious whether the methods of tensor analysis were contemporaneous to his era.

One sees that the trace of this tensor is:

$$T_{\mu}^{\mu} = -A_{\mu} J^{\mu}, \quad (5.11)$$

which will then vanish outside of the support of  $\mathbf{J}$ .

The divergence of the tensor  $T_{\nu}^{\mu}$  will be:

$$\partial_{\mu} T_{\nu}^{\mu} = F_{\mu\nu} J^{\mu}, \quad (5.12)$$

which is then the Lorentz force law in its relativistic form. Hence, energy-momentum will be conserved only outside of the support of  $\mathbf{J}$ .

When one couples the electromagnetic field  $F$  to a wave function  $\Psi$  and its complex conjugate  $\Psi^*$ , the gauge transformations of the first kind will play a role. If one replaces  $A$  with  $A + d\lambda$  then the replacement  $\Psi \mapsto e^{i\lambda} \Psi$ ,  $\Psi^* \mapsto e^{-i\lambda} \Psi^*$  will be called a *gauge transformation of the first kind*. Hence, the first kind of transformation enlarges the scope of phase transformations to ones that have local choices of phase origin, not just global ones. Therefore, when  $\lambda$  varies with position in  $S$ , one can no longer replace  $d\Psi$  with  $e^{i\lambda} d\Psi$  and  $d\Psi^*$  with  $e^{-i\lambda} d\Psi^*$ , since one must also differentiate  $\lambda$ , and the replacement will become:

$$d\Psi \mapsto e^{i\lambda} (d\Psi + i d\lambda \Psi), \quad d\Psi^* \mapsto e^{-i\lambda} (d\Psi^* - i d\lambda \Psi^*).$$

If one defines the *covariant differentials*:

$$\nabla\Psi = d\Psi + i A \Psi, \quad \nabla^*\Psi^* = (d\Psi^*) - i A \Psi^* \quad (5.13)$$

then one will see that:

$$\nabla\Psi \mapsto e^{i\lambda} [d\Psi + i (A + d\lambda) \Psi], \quad \nabla^*\Psi^* \mapsto e^{-i\lambda} [d\Psi^* - i (A + d\lambda) \Psi^*].$$

Hence, the gauge transformation of the first kind will induce a gauge transformation of the second kind in the covariant differentials.

The replacement of the ordinary differentials  $d\Psi$ ,  $d\Psi^*$  with covariant differentials  $\nabla\Psi$ ,  $\nabla^*\Psi^*$ , resp., is commonly referred to as *minimal electromagnetic coupling*.

Naturally, one now needs to extend the action functional and Lagrangian density to something that also depends upon  $A$  and  $F$ , as well as  $\nabla\Psi$  and  $\nabla^*\Psi^*$ , if one is account for gauge symmetries of both kinds, namely:

$$\mathcal{L} = \mathcal{L}(x, A, F, \Psi, \Psi^*, \nabla\Psi, \nabla^*\Psi^*). \quad (5.14)$$

Hence,  $\delta\mathcal{L}$  will now include an extra term that comes from  $A$ :



$$\delta\mathcal{L}_A = \frac{\partial\mathcal{L}}{\partial A_\mu} \delta A_\mu + \frac{\partial\mathcal{L}}{\partial F_{\mu\nu}} \delta F_{\mu\nu}, \quad (5.15)$$

while the terms that relate to  $\Psi, \Psi^*$  will take the form:

$$\delta\mathcal{L}_{\Psi, \Psi^*} = \frac{\partial\mathcal{L}}{\partial\Psi^a} \delta\Psi^a + \frac{\partial\mathcal{L}}{\partial\Psi^{*a}} \delta\Psi^{*a} + \frac{\partial\mathcal{L}}{\partial(\nabla_\mu\Psi^a)} \delta(\nabla_\mu\Psi^a) + \frac{\partial\mathcal{L}}{\partial(\nabla_\mu^*\Psi^{*a})} \delta(\nabla_\mu^*\Psi^{*a}). \quad (5.16)$$

The variational derivatives of  $\Psi, \Psi^*$  will then take the covariant forms:

$$\frac{\delta\mathcal{L}}{\delta\Psi^a} = \frac{\partial\mathcal{L}}{\partial\Psi^a} - \partial_\mu \frac{\partial\mathcal{L}}{\partial(\nabla_\mu\Psi^a)}, \quad \frac{\delta\mathcal{L}}{\delta\Psi^{*a}} = \frac{\partial\mathcal{L}}{\partial\Psi^{*a}} - \partial_\mu \frac{\partial\mathcal{L}}{\partial(\nabla_\mu^*\Psi^{*a})}. \quad (5.17)$$

### References (\*)

- 1\*. P. Dedecker, "Calcul de variations, différentielles, et champs géodésiques," *Géométrie différentielles*, Colloq. Int. de la Recherche Sci., Strasbourg (1953), 17-34.
2. D. H. Delphenich, "Integrability and the variational formulation of non-conservative mechanical systems" *Ann. Phys. (Berlin)* **18** (2009), 45-56.
3. D. H. Delphenich, "On the variational formulation of systems with non-holonomic constraints" *Ann. Phys. (Berlin)* **18** (2009), 649-670.
4. D. H. Delphenich, "A generalization of Noether's theorem based on the virtual work functional," (2011) arXiv:1109.0461.
- 5\*. E. Noether, "Invariante Variationsprobleme," *Kgl. Ges. d. Wiss. Nachrichten, Math.-phys. Klasse* (1918), 235-257.
6. F Halbwachs, *La theorie relativiste des fluides a spin*, Gauthier-Villars, Paris, 1960.
7. F. J. Belinfante, "On the current and the density of the electric charge, the energy, the linear momentum and the angular momentum of arbitrary fields," *Physica* **7** (1940) 449-474.
- 8\*. L. Rosenfeld, "Sur le tenseur d'impulsion-énergie," *Mémoires Acad. Roy. de Belgique* **18** (1940), 1-30.
9. A. Sakharov, "Vacuum Quantum Fluctuations in Curved Space and the Theory of Gravitation" *Doklady Akad. Nauk S. S. R.* **177** (1967), 70-71; English trans. in *Sov. Phys. – Doklady* **12** (1968), 1040-1041.

---

(\*) References marked with an asterisk are available in English translation at the author's website: neo-classical-physics.info.

## CHAPTER IV

# NON-RELATIVISTIC, SPINLESS PARTICLES

In this chapter, we shall attempt to merely document the various contributions that were made along the way to the formulation of non-relativistic quantum mechanics for spinless particles as something that also related to the motion of extended matter under the influence of a force of quantum origin. Since that force and its associated dynamical quantities, such as a quantum potential and a quantum stress tensor have a somewhat enigmatic character that one suspects will be the key to a better understanding of the foundations of quantum physics, we shall defer that discussion to the next chapter, in which the author's own thoughts on the matter will be presented.

**§ 1. The time-varying Schrödinger equation.** The approach that we shall take to defining the time-varying Schrödinger equation is that of the canonical quantization <sup>(1)</sup> of the classical expression for the total energy  $E$  of a point mass  $m$  that moves in space (which will generally be  $\mathbb{R}^3$ ) under the influence of a conservative force that is derived from a potential function  $U$ , namely:

$$E = \frac{1}{2m} P^2 + U. \quad (1.1)$$

In this,  $P = P_i(t) dx^i$  is the momentum 1-form that is defined along the trajectory  $x(t)$  of the mass point, which is related to its velocity:

$$\mathbf{v}(t) = v^i(t) \frac{\partial}{\partial x^i}, \quad v^i = \frac{dx^i}{dt} \quad (i = 1, 2, 3) \quad (1.2)$$

by the mechanical constitutive law:

$$P_i = m \delta_{ij} v^j = m v_i. \quad (1.3)$$

One refers to the 1-form:

$$v = v_i dx^i \quad (1.4)$$

as the *covelocity* of the motion.

Of course, we are implicitly assuming that the three-dimensional space in which  $m$  moves has a Euclidian metric  $\delta = \delta_{ij} dx^i dx^j$  defined on it. That will then make:

$$P^2 = \delta^{ij} P_i P_j = (P_1)^2 + (P_2)^2 + (P_3)^2. \quad (1.5)$$

Furthermore:

$$P^2 = m^2 v^2 = m^2 (\delta_{ij} v^i v^j) = m^2 [(v^1)^2 + (v^2)^2 + (v^3)^2]. \quad (1.6)$$

---

<sup>(1)</sup> Canonical quantization means associating the canonical variables  $x^i, p_i$  of Hamiltonian mechanics with operators. Here, we shall mostly be concerned with the momentum variables.

*a. Energy as an operator.* The process of canonical quantization really begins in kinematics, although it is usually introduced in the context of dynamics. That is, as a result of a Fourier transformation of functions that depend upon  $t$  and  $x^i$ , one will find the following pair of associations (<sup>1</sup>):

$$\frac{\partial}{\partial t} \leftrightarrow -i\omega, \quad \frac{\partial}{\partial x^i} \leftrightarrow ik_i, \quad (1.7)$$

in which  $\omega$  is the radial frequency of a wave, and  $k_i$  are its wave numbers in each spatial direction.

We can also anticipate the relativistic theory by generalizing our spatial coordinates to space-time ones  $(t, x^i)$ , and then defining  $x^0 = t$ ,  $k_0 = -\omega$ , so that the associations (1.7) can be consolidated into a single one:

$$\frac{\partial}{\partial x^\mu} \leftrightarrow ik_\mu \quad (\mu = 0, \dots, 3). \quad (1.8)$$

The components  $k_\mu$  define the *frequency-wave-number 1-form*:

$$k = k_\mu dx^\mu = -\omega dt + k_i dx^i. \quad (1.9)$$

If we define the *energy-momentum 1-form* of the point mass  $m$  by extending the momentum 1-form with total energy  $E(t)$ :

$$P(t) = P_\mu dx^\mu = -E(t) dt + p_i(t) dx^i \quad (1.10)$$

then the de Broglie relations for matter waves, i.e.:

$$E = \hbar\omega, \quad P_i = \hbar k_i, \quad (1.11)$$

can be expressed concisely as:

$$P = \hbar k, \quad (1.12)$$

in which  $\hbar = h / 2\pi$ , where  $h$  is Planck's constant.

If one compares (1.9) to (1.10) then one will see that since  $P(t)$  is defined only along the trajectory of the point-mass, while  $k$  is defined on the support of a wave-function, which will be a region of space-time, the only way that (1.12) can make sense is if one essentially assumes that  $\omega$  and  $k_i$  are all constants, which would make the wave a plane wave if the coordinates were rectangular. Clearly, the association of  $k$  with  $P$  would be more satisfying if  $P$  were also defined over a region of space-time that is not just a curve, as well. Hence, one sees that perhaps a continuum model for the mass would be more mathematically convenient, after all. In particular, it would make sense if the mass were distributed over the support of the corresponding wave.

---

(<sup>1</sup>) The minus sign before  $\omega$  comes from using  $e^{-i(\omega t - k_i x^i)}$  in that transform, which represents a travelling plane wave that moves in the direction of  $k^i$ .

If one considers  $k$  to be a representation of the kinematical state of a wave and  $P$  to be a representation of its dynamical state then one can also view the de Broglie relations (1.11) or (1.12) as a type of mechanical constitutive law for wave motion, in which Planck's constant plays a role that is analogous to the role of  $m$  in (1.3). However, if one associates  $k$  with an energy-momentum *density* 1-form  $p$  by way of  $p = \hbar k$  then since  $\hbar$  would need to have different units (viz., energy-time/volume), one might also ask whether the newly-defined  $\hbar$  were still a constant, or some more general function of space.

We shall return to the ideas that were presented in the last two paragraphs later in this chapter.

The de Broglie relations then allow us to put the canonical quantization rules (1.8) into the form:

$$\hbar \frac{\partial}{\partial x^\mu} \leftrightarrow iP_\mu, \quad (1.13)$$

or

$$\frac{\hbar}{i} \frac{\partial}{\partial x^\mu} \leftrightarrow P_\mu, \quad (1.14)$$

which is the way that one usually works with them in conventional quantum mechanics [1]. More to the point, one uses the associations:

$$i\hbar \frac{\partial}{\partial t} \leftrightarrow E, \quad \frac{\hbar}{i} \frac{\partial}{\partial x^i} \leftrightarrow P_i. \quad (1.15)$$

The kinetic energy of the point-mass then gets associated with the linear, second-order partial differential operator:

$$\frac{1}{2m} P^2 = -\frac{\hbar^2}{2m} \left[ \left( \frac{\partial}{\partial x^1} \right)^2 + \left( \frac{\partial}{\partial x^2} \right)^2 + \left( \frac{\partial}{\partial x^3} \right)^2 \right] = -\frac{\hbar^2}{2m} \Delta, \quad (1.16)$$

in which  $\Delta$  is the three-dimensional Laplacian operator.

The total energy equation (1.1) then becomes the operator equation:

$$i\hbar \frac{\partial}{\partial t} = -\frac{\hbar^2}{2m} \Delta + U(t, x^i), \quad (1.17)$$

in which the operator  $U$  means multiplication by a scalar function.

All that one has to do to produce the time-varying Schrödinger equation from the latter operator equation is to apply both sides of the operator equation to a complex-valued wave function  $\Psi(t, x^i)$ , which must then be at least once-differentiable in time and twice-differentiable in space:

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \Psi + U \Psi. \quad (1.18)$$

One can see that virtually everything that one can say about the nature of this equation will depend upon the specific form of the potential function  $U$ . Indeed, the textbook examples that one encounters in most physics courses are mostly concerned with time-invariant potential functions  $U(x^i)$ , which will lead one to the stationary Schrödinger equation. We shall discuss that case in due course, but first we shall discuss the Lagrangian formulation of the time-varying Schrödinger equation, and then we shall introduce the Madelung-Takabayasi form of (1.18).

*b. Lagrangian formulation of the time-varying Schrödinger equation.* One can derive the time-varying Schrödinger equation from a Lagrangian density that takes the form <sup>(1)</sup>:

$$\mathcal{L} = \frac{i\hbar}{2}(\dot{\Psi}\Psi^* - \Psi\dot{\Psi}^*) - \frac{\hbar^2}{2m} \|d_s\Psi\|^2 - U\|\Psi\|^2. \quad (1.19)$$

One sees that (1.18) can then be derived from  $\mathcal{L}$  by annulling the variational derivative of  $\mathcal{L}$  with respect to  $\Psi^*$ :

$$0 = \frac{\delta\mathcal{L}}{\delta\Psi^*} = \frac{\partial\mathcal{L}}{\partial\Psi^*} - \frac{\partial}{\partial t} \frac{\partial\mathcal{L}}{\partial(\dot{\Psi}^*)} - \text{div} \frac{\partial\mathcal{L}}{\partial(d_s\Psi^*)}. \quad (1.20)$$

Similarly, if one varies  $\mathcal{L}$  with respect to  $\Psi$  then one will get the corresponding wave equation for  $\Psi^*$ .

One can give equation (1.20) an “infinite-dimensional  $F = ma$ ” form by introducing the generalized force  $f$  that acts upon the field  $\Psi$  and its generalized momenta  $\Pi^t, \Pi^i$  <sup>(2)</sup>:

$$\begin{aligned} f &= \frac{\partial\mathcal{L}}{\partial\Psi} = -\frac{i\hbar}{2}\dot{\Psi}^* - U\Psi^*, \\ \Pi^t &= \frac{\partial\mathcal{L}}{\partial(\dot{\Psi}^*)} = \frac{i\hbar}{2}\Psi^*, & \Pi^i &= \frac{\partial\mathcal{L}}{\partial(\partial_i\Psi^*)} = -\frac{\hbar^2}{2m}\delta^{ij}\partial_j\Psi^*, \end{aligned} \quad (1.21)$$

and similarly for  $\Psi^*$ :

$$\begin{aligned} f^* &= \frac{\partial\mathcal{L}}{\partial\Psi^*} = \frac{i\hbar}{2}\dot{\Psi} - U\Psi, \\ \Pi^{t*} &= \frac{\partial\mathcal{L}}{\partial(\dot{\Psi})} = -\frac{i\hbar}{2}\Psi, & \Pi^{i*} &= \frac{\partial\mathcal{L}}{\partial(\partial_i\Psi)} = -\frac{\hbar^2}{2m}\delta^{ij}\partial_j\Psi. \end{aligned} \quad (1.22)$$

<sup>(1)</sup> The subscript  $s$  means that one considers only the spatial part of the 1-form. For example,  $d_s\Psi = \partial_i\Psi dx^i$ .

<sup>(2)</sup> One must note that in performing the differentiations, one must treat  $\Psi$  and  $\Psi^*$  as independent variables:

$$\frac{\partial\Psi^*}{\partial\Psi} = \frac{\partial\dot{\Psi}^*}{\partial\dot{\Psi}} = \frac{\partial(\partial_i\Psi^*)}{\partial(\partial_i\Psi)} = 0, \text{ etc.}$$

As one sees, the second set of quantities are all complex conjugates of the first; i.e., complex conjugation commutes with differentiation. (That is because the differentiation is with respect to *real* variables, not complex ones.)

The time-varying Schrödinger equation for  $\Psi^*$  can then be given the form:

$$f = \partial_t \Pi^t + \partial_i \Pi^i, \quad (1.23)$$

and the corresponding equation for  $\Psi$  will be the complex conjugate of this.

The Lagrangian  $\mathcal{L}$  is invariant under the *global* action of the group  $U(1)$ , since the global choice of origin for the phase of the wave function is arbitrary. The corresponding infinitesimal generator of that symmetry takes the form:

$$\delta\Psi = i\alpha\Psi, \quad \delta\Psi^* = -i\alpha\Psi^*. \quad (1.24)$$

The Noether current that is associated with this infinitesimal symmetry is then the vector field whose components are:

$$J^0 = \Pi^t \delta\Psi + \Pi^{t*} \delta\Psi^* = -\hbar\alpha \Psi\Psi^*, \quad (1.25)$$

$$J^i = \Pi^i \delta\Psi + \Pi^{i*} \delta\Psi^* = \frac{i\hbar^2\alpha}{2m} \delta^{ij} (\Psi^* \partial_j \Psi - \Psi \partial_j \Psi^*), \quad (1.26)$$

which are basically  $-\hbar\alpha$  times the expressions that are usually identified with the “probability current” in quantum mechanics, namely:

$$J^0 = \Psi\Psi^*, \quad J^i = \frac{\hbar}{2mi} \delta^{ij} (\Psi^* \partial_j \Psi - \Psi \partial_j \Psi^*). \quad (1.27)$$

Clearly, one can drop an overall constant, non-zero scalar factor from the definition of  $J^\mu$ , since the divergence operator is linear. However, the relative signs and units of  $J^\mu$  do have physical significance.

One can obtain a canonical energy-momentum-stress tensor from the Lagrangian above:

$$T_\nu^\mu = \Pi^\mu \partial_\nu \Psi + \Pi^{\mu*} \partial_\nu \Psi^* - \mathcal{L} \delta_\nu^\mu. \quad (1.28)$$

Its individual submatrices are:

$$T_0^0 = \frac{\hbar^2}{2m} \|d_s \Psi\|^2 + U \|\Psi\|^2 = \mathcal{H}, \quad (1.29)$$

$$T_j^0 = +\frac{i\hbar}{2} (\Psi^* \partial_j \Psi - \Psi \partial_j \Psi^*) = -m J_j, \quad (1.30)$$

$$T_0^i = -\frac{\hbar^2}{2m}(\dot{\Psi} \partial^i \Psi^* + \dot{\Psi}^* \partial^i \Psi), \quad (1.31)$$

$$T_j^i = -\frac{\hbar^2}{2m}(\partial^i \Psi^* \partial_j \Psi + \partial^i \Psi \partial_j \Psi^*) - \mathcal{L} \delta_j^i. \quad (1.32)$$

Note the asymmetry in the components  $T_0^i$  and  $T_j^0$ , while the stress tensor  $T_{ij} = \delta_{ik} T_j^k$  is symmetric.

The trace of this tensor is then:

$$T_\mu^\mu = \mathcal{H} - 3\mathcal{L} = -\frac{3\hbar i}{2}(\dot{\Psi} \Psi^* - \Psi \dot{\Psi}^*) - \frac{2\hbar^2}{m} \|d_s \Psi\|^2 + 2U \|\Psi\|^2. \quad (1.33)$$

The divergence is easier to interpret:

$$\partial_\mu T_\nu^\mu = \partial_\nu U \|\Psi\|^2. \quad (1.34)$$

Its temporal component will then vanish when the external potential is time-invariant, while its spatial components represent the external force (density) that acts upon  $m$ .

## § 2. The Madelung-Takabayasi form of the time-varying Schrödinger equation.

In the same year that the Copenhagen School of quantum mechanics was establishing the statistical interpretation of the quantum wave function  $\Psi$ , Ernst Madelung [2] was proposing an alternative interpretation that he called the “hydrodynamical” interpretation. Although his way of transforming the Schrödinger equation was actually confined to the stationary case, it was subsequently extended to the time-varying case by Takehiko Takabayasi [3], who expanded the scope of its application to quantum mechanics considerably.

*a. The basic transformation.* The essential step in the transformation to the Madelung-Takabayasi form of (1.18) is to introduce polar coordinates  $(r, \theta)$  into the complex plane  $\mathbb{C}$  in which  $\Psi$  takes its values. Thus, one expresses  $\Psi$  in the form:

$$\Psi = R e^{i\theta}, \quad (2.1)$$

in which both  $R$  and  $\theta$  are real-valued functions of  $(t, x^j)$  that are “sufficiently-differentiable.”

It helps to know that:

$$\dot{\Psi} = \left( \frac{\dot{R}}{R} + i\dot{\theta} \right) \Psi, \quad \dot{\Psi}^* = \left( \frac{\dot{R}}{R} - i\dot{\theta} \right) \Psi^*, \quad (2.2)$$

$$d_s \Psi = \left( \frac{d_s R}{R} + i d_s \theta \right) \Psi, \quad d_s \Psi^* = \left( \frac{d_s R}{R} - i d_s \theta \right) \Psi^*, \quad (2.3)$$

and the fact that  $\Psi \Psi^* = R^2$ .

Substituting the polar form of  $\Psi$  into (1.18) will give:

$$i\hbar \left( \frac{R_t}{R} + i\theta_t \right) = -\frac{\hbar^2}{2m} \left[ \frac{\Delta R}{R} - (d_s \theta)^2 + i \left( \Delta \theta + 2 \frac{\langle d_s \theta, d_s R \rangle}{R} \right) \right] + U. \quad (2.4)$$

In this equation, the subscript  $t$  indicates the first partial derivative with respect to  $t$ , and we are using the notation  $\langle \alpha, \beta \rangle = \delta^{ij} \alpha_i \beta_j$  to indicate the Euclidian scalar product of spatial 1-forms.

Since both sides of the equation represent complex functions of  $(t, x^i)$ , one can separate the real and imaginary parts and arrive at a pair of real equations that are equivalent to (1.18):

$$R_t = -\frac{\hbar}{2m} (R \Delta \theta + 2 \langle d_s \theta, d_s R \rangle), \quad -\hbar \theta_t = \frac{\hbar^2}{2m} \left[ (d_s \theta)^2 - \frac{\Delta R}{R} \right] + U. \quad (2.5)$$

One might note that the factor of  $\hbar$  always occurs with the same power as the partial derivatives of  $\theta$ , so one might make the perfectly reasonable replacement:

$$k = d\theta \quad (2.6)$$

with  $k$  defined as above; i.e.:

$$\omega = -\frac{\partial \theta}{\partial t}, \quad k_i = \frac{\partial \theta}{\partial x^i}. \quad (2.7)$$

Of course, the fact that  $k$  must be exact implies that it must also be closed:

$$d \wedge k = 0. \quad (2.8)$$

Later on, we shall examine some attempts to go beyond the constraint that this implies on the original wave function  $\Psi$ .

Since  $\omega$  and  $k_i$  always appear with a factor of  $\hbar$ , one can then replace  $-\hbar \theta_t$  with  $E$  and  $\hbar k_i$  with  $P_i$ , as before, and the system of equations (2.5) will take the form:

$$R_t = -\frac{1}{2m} (R \operatorname{div} \mathbf{P}_s + 2 \langle P_s, d_s R \rangle), \quad E = \frac{1}{2m} \langle P_s, P_s \rangle + U - \frac{\hbar^2}{2m} \frac{\Delta R}{R}. \quad (2.9)$$

There are two essential differences between the  $E$  and  $P_i$  in these equations and the energy and momentum that were defined above. First of all, we were previously talking about functions of time that were defined only along the points of a curve that was followed by a point-mass, but now since  $\theta$  is a function of  $(t, x^i)$ , so are  $E$  and  $P_i$ . Hence,



if we are to still think of them as being associated with energy and momentum, we must no longer think in terms of a point-like mass, but a spatially-extended one. Secondly, the units of  $E$  and  $P_i$  are still those of energy and momentum, resp., not energy density and momentum density, resp., whereas for an extended mass distribution the role of  $m$ ,  $E$ ,  $P_i$  would now be that of the *total* mass, energy, and momentum, resp., which are only meaningful when one passes to the point-mass approximation.

Clearly, we are not dealing with point mechanics anymore, but continuum mechanics. The main challenge to physics is then to identify the nature of the extended matter that is being described by the equations to which we have arrived, and especially, the constitutive properties of that medium, which we shall refer to as a *Madelung medium*.

*b. The conservation of mass.* In order to get more insight into what the mass density function  $\rho$  might be, we start with the fact that:

$$\|\Psi\|^2 = \Psi\Psi^* = R^2. \quad (2.10)$$

In the statistical interpretation of the wave function  $\Psi$ , one thinks of  $\|\Psi\|^2$  as representing a *probability* density function for finding the *point* particle  $m$  in the differential region of space  $dV$  around the point  $x$  at time  $t$ . Of course, that can only come about as long as the wave function  $\Psi$  has been normalized to have “total mass” equal to 1; i.e.:

$$1 = \int \|\Psi\|^2 dV, \quad (2.11)$$

in which the domain of integration is all space, as it pertains to one’s particular problem.

However, this normalization is not necessary in the eyes of the Schrödinger equation itself, but only in the interpretation of the wave function. That leaves open the possibility of giving the space-time function  $\|\Psi\|^2$  or  $R^2$  a different interpretation. Madelung chose to interpret:

$$n = R^2 \quad (2.12)$$

as a *number density*.

One usually introduces number densities in the context of ensembles of a large, but integer, number of point-like masses distributed in space, such as in the kinetic theory of gases or with crystal lattices. Since  $n$  is a real number here, if one assumes that one is dealing with extended matter, not a finite collection of points, then there will be nothing conceptually wrong with saying that the integral  $\frac{1}{N} \int_V n$  represents the *fraction* of all the extended objects that is contained in the spatial volume  $V$ , where  $N$ , namely, the integral of  $n$  over all space represents the total number of distinct objects. One might then relate this to a probability density function by saying that if an object is extended in space then the probability of finding it in a region of space will equal the fraction of the object that is contained in that region.

If the total mass of the extended object is  $m$  then one can define a mass density for it by means of:

$$\rho = mn = mR^2. \quad (2.13)$$

If the motion of this mass density were associated with a velocity vector field  $\mathbf{v}(t, x^i)$ , whose support would then be the same as that of  $\rho$ , then the equation of continuity for the conservation of that mass would be:

$$\rho_t = -\partial_i(\rho v^i) = -\partial_i \rho v^i - \rho \partial_i v^i. \quad (2.14)$$

If we rewrite the first of (2.5) in the form:

$$R_t = -\frac{1}{2m} R \frac{\partial P_i}{\partial x^i} - \frac{1}{m} R_i P^i \quad (2.15)$$

and multiply both sides by  $2mR$  then that will give:

$$2mRR_t = -R^2 \partial_i P_i - 2R \partial_i R P^i,$$

or

$$\partial_t(mR^2) = -\partial_i(R^2 P_i). \quad (2.16)$$

This equation will be consistent with (2.14) iff:

$$\rho = m R^2, \quad R^2 P^i = \rho v^i. \quad (2.17)$$

One can solve the last equation for  $v^i$ :

$$v^i = \frac{1}{m} P^i = \frac{\hbar}{m} \delta^{ij} \partial_j \theta = \frac{\hbar}{m} k^i. \quad (2.18)$$

Thus, we have succeeded in defining a mass density  $\rho$  and a velocity vector field  $\mathbf{v}$  for the motion of an extended mass distribution in space in terms of the kinematical information that was contained in the wave function  $\Psi$  (in the form of  $R$  and  $\theta$ ) and its differential, along with the empirical data that was given by  $m$  and  $\hbar$ .

It is worth noting that since the difference between the mass density  $\rho$  and the number density  $n$  is only a multiplicative constant, the number density will also be conserved; i.e.:

$$n_t = -\partial_i(n v^i). \quad (2.19)$$

Since we started out with a wave function, which can be associated with two velocities, namely, a phase velocity and a group velocity, whose components will take the form:

$$v_g^i = \frac{c}{\|k\|} k^i, \quad (2.20)$$

the necessary and sufficient condition for the velocity (2.18) to equal the group velocity is the physically-reasonable constraint that the magnitude of the wave number  $\|k\|$  must be the Compton wave number that is associated with the total mass  $m$ :

$$\|k\| = \frac{mc}{\hbar}. \quad (2.21)$$

One can get further information about the mechanical properties of the extended matter in motion from the differential of the spatial velocity vector field  $\mathbf{v}$ , or even better (since the metric is Euclidian), the covelocity 1-form:

$$v = \frac{\hbar}{m} d_s \theta. \quad (2.22)$$

This also means that  $v$  admits a *velocity potential* <sup>(1)</sup>, or *stream function*, in the form of  $\hbar\theta/m$ .

If one takes the (spatial) differential of this 1-form then the resulting *rate of deformation* can be polarized into a symmetric and an antisymmetric part:

$$d_s v = \dot{e} + \dot{\vartheta} = \frac{1}{2} \dot{e}_{ij} dx^i dx^j + \frac{1}{2} \dot{\vartheta}_{ij} dx^i \wedge dx^j, \quad (2.23)$$

for which:

$$\dot{e}_{ij} = \partial_i v_j + \partial_j v_i = \frac{\hbar}{m} (\partial_i k_j + \partial_j k_i) = \frac{2\hbar}{m} \partial_{ij} \theta \quad (2.24)$$

is the *rate of strain* and:

$$\dot{\vartheta}_{ij} = \partial_i v_j - \partial_j v_i = 0 \quad (2.25)$$

is the *kinematical vorticity*, which will vanish because it is also the exterior derivative of  $v$  which is exact, since  $d_s \theta$  is. Thus, the Madelung medium will be *irrotational*. The vanishing of the vorticity also makes the rate of deformation  $d_s v$  tensor symmetric.

One can also obtain the *rate of dilatation* from the trace of  $\dot{e}_{ij}$ :

$$\dot{\lambda} = \dot{e}_i^i = \text{div } v = \frac{2\hbar}{m} \text{div } k_s = \frac{2\hbar}{m} \Delta \theta. \quad (2.26)$$

The vanishing of this function would be equivalent to the incompressibility of the medium, so we shall generally treat the Madelung medium as being compressible.

We can then express the rate of deformation tensor in the form:

$$\partial_i v_j = \dot{e}_{ij}^0 + \frac{1}{3} \delta_{ij} \Delta \theta, \quad (2.27)$$

in which:

$$\dot{e}_{ij}^0 = \dot{e}_{ij} - \frac{1}{3} \delta_{ij} \Delta \theta \quad (2.28)$$

is the traceless part of  $\dot{e}_{ij}^0$ .

---

<sup>(1)</sup> Of course, if the support of the covelocity 1-form is multiply-connected then the velocity potential will exist only locally.

*c. The balance of energy.* Turning to the first equation (2.9) of the Madelung-Takabayasi equations, which we rewrite as:

$$E = \frac{1}{2m} P_s^2 + U + U_{\hbar}, \quad (2.29)$$

we note that, despite the usual line of reasoning, it would *not* be entirely consistent to call it the balance of energy for the motion of an *extended* mass distribution, since all of the terms have the units of *total* energy, which is more appropriate to the *point*-mass approximation, as is the total mass  $m$ . Hence, we are still talking about the energy balance for a point-mass  $m$  whose total energy is  $E$ , and whose total momentum is  $P_s$ , and which moves in the presence of a conservative force with a force potential  $U$ , along with a somewhat mysterious conservative force whose potential function takes the form:

$$U_{\hbar} = -\frac{\hbar^2}{2m} \frac{\Delta R}{R}. \quad (2.30)$$

When we discuss the Lagrangian density, we shall see that the simple solution to making densities out of totals is to multiply the total quantities by the number density  $n$ , as we did with mass. Although this makes sense for  $E$  and  $P_s$ , one must note that multiplying a potential function by a number density will change its character noticeably.

For now, we shall continue in the manner of the classical researchers in this field.

Because  $U_{\hbar}$  is the only place in the equations in which  $\hbar$  is actually used explicitly (although a factor of  $\hbar$  has been *implicitly* absorbed into the definition of  $P_s$ ),  $U_{\hbar}$  came to be referred to as the *quantum potential* or *Madelung potential*. Indeed, in the classical limit as  $\hbar$  goes to zero, so will  $U_{\hbar}$ . Hence, one suspects that  $U_{\hbar}$  should contain the essence of the difference between classical continuum mechanics and quantum wave mechanics in its structure. One also notes that the expression  $U + U_{\hbar}$  is suggestive of a “loop expansion” in powers of  $\hbar$ , which will give one an “effective potential” in the form of a classical potential plus quantum corrections of increasing order in  $\hbar$ .

Of course, the vanishing of a potential function is not as physically definitive as the vanishing of the force that it defines. The quantum force that goes with  $U_{\hbar}$  is the 1-form:

$$F_{\hbar} = -d_s U_{\hbar} = \frac{\hbar^2}{2m} d_s \left( \frac{\Delta R}{R} \right). \quad (2.31)$$

Once again, this will have the units of force, not force density.

$F_{\hbar}$  will vanish iff  $\lambda \equiv \Delta R / R$  is constant in space, which can be rewritten:

$$\Delta R = \lambda R. \quad (2.32)$$

Thus, *eigenfunctions of the Laplacian will produce no quantum force*. Some elementary examples of such amplitude functions are  $R_0 e^{\sqrt{\lambda}x}$  when one is using Cartesian coordinates and  $R_0 e^{\sqrt{\lambda}r}$  when one is using spherical coordinates.

*d. The balance of momentum.* If one takes spatial gradients of both sides of the energy balance equation (2.29) then the resulting equation will be:

$$d_s E = d_s \left( \frac{P_s^2}{2m} \right) + d_s U + d_s U_h.$$

Since:

$$d_s E = -d_s \left( \hbar \frac{\partial \theta}{\partial t} \right) = -\frac{\partial}{\partial t} (\hbar d_s \theta) = -\frac{\partial P_s}{\partial t},$$

and

$$\partial_i \left( \frac{P_s^2}{2m} \right) = \frac{1}{2m} [2(\partial_i P_j) P^j] = v^j \partial_i P_j = v^j \partial_j P_i,$$

in which the last step will be true because:

$$\partial_j P_i = \hbar \partial_{ji} \theta = \hbar \partial_{ij} \theta = \partial_i P_j,$$

if we put all of this together then we will get the equation:

$$L_v P_i = -\partial_i (U + U_h). \quad (2.33)$$

In this, we have introduced the Lie derivative of the component functions:

$$L_v P_i \equiv i_v dP_i = v^\mu \partial_\mu P_i = \frac{\partial P_i}{\partial t} + v^j \partial_j P_i \quad (\mu = 0, \dots, 3, v^0 = 1). \quad (2.34)$$

This means that the equation (2.33) represents the balance of (total) linear momentum for the motion of the medium when one describes it in the ‘‘Euler picture’’ of continuum mechanics. That is, one describes the time variation of physical functions as they are seen from someone moving with the extended object. Since this is typical of hydrodynamical problems, the classical authors on the Madelung medium have identified the equation as the Euler equation for the motion of the fluid. However, one can use the Euler picture to represent media that are not fluids, as well.

*e. The quantum stress tensor.* Following Takabayasi [3a], one can associate the quantum force with a *quantum stress tensor*  $\sigma_j^i$  that makes:

$$\partial_j \sigma_i^j = n F_h = -n \partial_i U_h. \quad (2.35)$$

It takes the form:

$$\sigma_{ij} = -\frac{\hbar^2}{2m} R^2 \frac{\partial^2 \ln R}{\partial x^i \partial x^j} = -\frac{\hbar^2}{2m} \left[ R \frac{\partial^2 R}{\partial x^i \partial x^j} - \frac{\partial R}{\partial x^i} \frac{\partial R}{\partial x^j} \right]. \quad (2.36)$$

Here is an example of how turning total quantities into densities by multiplying by the number density can change the character of a quantity in other ways. Namely, the force density 1-form:

$$f_h = -n dU_h \quad (2.37)$$

will not necessarily be exact anymore. That is,  $U_h$  will not be precisely a force potential, so  $f_h$  will not be a conservative force, in general. Since:

$$d \wedge f_h = -dn \wedge dU_h = -dn \wedge f_h \quad (2.38)$$

this will vanish iff  $dn$  (i.e., the gradient of  $n$ ) is collinear with  $f_h$ . Of course, this would be reasonable for a spherically-symmetric  $n$  with a radial force acting upon it.

One does see that  $f_h \wedge df_h$  will vanish in any case, so by Frobenius's theorem, the codimension-one differential system that is defined by the annihilating planes of  $f_h$  will be completely integrable. That is, the support of  $f_h$  will admit a foliation by integral surfaces. However, since the annihilating planes of  $f_h$  are the same as those of  $dU_h$ , those integral surfaces will be the level surfaces of  $U_h$ , after all; in other words, the integral surfaces of  $f_h = 0$  will be the equipotentials of  $U_h$ .

Since the off-diagonal components of  $\sigma_{ij}$  – viz., the shears – do not have to vanish, and they are not generally coupled to the rate of deformation tensor, which would suggest viscosity, one sees that calling the Madelung medium a fluid is somewhat imprecise. Hence, we shall simply refer to it as the “Madelung medium” and leave the details of its true mechanical constitutive properties to the results of further analysis.

Indeed, the quantum stress tensor will take the form of a perfect fluid stress tensor:

$$\sigma_j^i = -\pi \delta_j^i \quad (2.39)$$

iff  $n$  takes the form of a Gaussian distribution:

$$n(x) = C e^{-\|\mathbf{x}-\mathbf{x}_\mu\|^2/2\sigma^2}, \quad (2.40)$$

for which, one will find that pressure must be:

$$\pi = \frac{\hbar^2}{m\sigma^2} n. \quad (2.41)$$

This, too, has a certain reasonableness to it, since it suggests that the pressure increases with number density, as well as the degree to which the distribution is mostly found near the mean. One might expect that trying to compress a mass distribution into a smaller region would increase its pressure.

This quantum stress tensor is associated with a (*mean*) *quantum pressure*:

$$\pi_h = -\frac{1}{3}\sigma_i^i = -\frac{\hbar^2}{6m}R^2\Delta(\ln R) = \frac{\hbar^2}{6m}\left[(d_s R)^2 - R\Delta R\right]. \quad (2.42)$$

This quantum pressure can then be positive, negative, or zero, and it will vanish iff:

$$\Delta R = \frac{(d_s R)^2}{R}. \quad (2.43)$$

The equation (2.29) of energy balance can also be regarded as a second-order, nonlinear partial differential equation for the functions  $R$  and  $\theta$  by replacing  $E$ ,  $P_s$ , and  $U_h$  with their expressions in terms of those functions:

$$\hbar d\theta = \frac{\hbar^2}{2m}\left[(d\theta)^2 - \frac{\Delta R}{R}\right] + U. \quad (2.44)$$

Originally, this equation was naturally identified as being of the Hamilton-Jacobi type, although it was later pointed out in a letter to Physical Review by Halpern [4] that, strictly speaking, that was not true. In particular, the functions  $\theta$  and  $R$  are not independent of each other, since they are coupled implicitly by the continuity equation.

*f. Lagrangian formulation of the Madelung-Takabayasi equations.* In order to obtain a Lagrangian density function that will give the Madelung-Takabayasi equation, all that one needs to do is substitute  $Re^{i\theta}$  for  $\Psi$  in the Schrödinger Lagrangian density (1.19), so one also substitutes  $Re^{-i\theta}$  for  $\Psi^*$ . After all of the substitutions have been made, the result will be:

$$\mathcal{L} = -R^2 \left\{ \hbar\dot{\theta} + \frac{\hbar^2}{2m} \left[ (d_s\theta)^2 + \left( \frac{d_s R}{R} \right)^2 \right] + U \right\}. \quad (2.45)$$

This time, we see that since  $n = R^2$ , the effect of multiplying the term in curly brackets by  $R^2$  is to convert the total energy sum inside of it into a total energy density. Meanwhile, the last term in square brackets takes the form of a stress energy density that is associated with the gradient of  $R$ .

The conjugate forces and momenta to the coordinates  $R$  and  $\theta$  are:

$$f_R = \frac{\partial \mathcal{L}}{\partial R} = -2R \left[ \hbar\dot{\theta} + \frac{\hbar^2}{2m} (d_s\theta)^2 + U \right],$$

$$\Pi'_R = \frac{\partial \mathcal{L}}{\partial \dot{R}} = 0, \quad \Pi'_R = \frac{\partial \mathcal{L}}{\partial(\partial_i R)} = -\frac{\hbar^2}{m} \partial^i R. \quad (2.46)$$

$$f_\theta = \frac{\partial \mathcal{L}}{\partial \theta} = 0, \quad \Pi'_\theta = \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = -\hbar R^2, \quad \Pi^i_\theta = \frac{\partial \mathcal{L}}{\partial(\partial_i \theta)} = -\frac{\hbar^2}{m} R^2 \partial^i \theta. \quad (2.47)$$

One confirms that the Madelung-Takabayasi equations can be obtained from this Lagrangian density when one varies it with respect to the basic variables  $R$  and  $\theta$ :

$$0 = \frac{\delta \mathcal{L}}{\delta R} = f_R - \partial_i \Pi'_R - \partial_i \Pi^i_R, \quad 0 = \frac{\delta \mathcal{L}}{\delta \theta} = f_\theta - \partial_i \Pi'_\theta - \partial_i \Pi^i_\theta, \quad (2.48)$$

in which the first equation will be the balance of energy, while the second one will be the conservation of mass.

*g. Noether currents for the Madelung Lagrangian.* In order to find the Noether current that corresponds to the phase-invariance of the wave function  $\Psi$ , one substitutes  $R e^{i\theta}$  for  $\Psi$  and computes the variation from that:

$$\delta \Psi = \delta R e^{i\theta} + i \delta \theta R e^{i\theta} = \left( \frac{\delta R}{R} + i \delta \theta \right) \Psi = i \alpha \Psi. \quad (2.49)$$

This necessarily implies that:

$$\delta R = 0, \quad \delta \theta = \alpha, \quad (2.50)$$

which is reasonable, since only the phase is being varied.

The components of the Noether current that is associated with  $\delta \theta$  are then:

$$J^0 = \Pi'_\theta \delta \theta = R^2 = n, \quad J^i = \Pi^i_\theta \delta \theta = \frac{1}{m} R^2 P^i = n v^i, \quad (2.51)$$

after removing an overall constant factor of  $-\hbar \alpha$ . When one compares these equations to equation (1.27), one will see that one could also simply make the substitution (2.1) in the latter equation and arrive at the present expressions.

This current is actually the number density flux, but, as we pointed out above, if one multiplies it by  $m$  then one will get the mass flux, which will also satisfy the continuity equation.

The canonical energy-momentum-stress tensor:

$$T^\mu_\nu = \Pi^\mu_\theta \partial_\nu \theta + \Pi^\mu_R \partial_\nu R - \mathcal{L} \delta^\mu_\nu \quad (2.52)$$

that is associated with  $\mathcal{L}$  is then ( $S = \hbar \theta$ ):



$$T_0^0 = \Pi_\theta^t \dot{\theta} - \mathcal{L} = \frac{\hbar^2}{2m} [R^2 (d_s \theta)^2 + (d_s R)^2] + R^2 U = \mathcal{H}, \quad (2.53)$$

$$T_j^0 = \Pi_\theta^t \partial_j \theta = -\hbar R^2 \partial_j \theta = -m J_j, \quad (2.54)$$

$$T_0^i = \Pi_R^i \dot{R} + \Pi_\theta^i \dot{\theta} = -\frac{\hbar^2}{m} [R^2 \partial^i \theta \partial_i \theta + \partial^i R \partial_i R], \quad (2.55)$$

$$T_j^i = \Pi_R^i \partial_j R + \Pi_\theta^i \partial_j \theta - \mathcal{L} \delta_j^i = -\frac{\hbar^2}{m} [R^2 \partial^i \theta \partial_j \theta + \partial^i R \partial_j R] - \mathcal{L} \delta_j^i. \quad (2.56)$$

One can check that these expressions are simply the polar forms of (1.29)-(1.32).

The trace of the matrix  $T_\nu^\mu$  is then [compare to (1.33)]:

$$T_\mu^\mu = 3\hbar R^2 \partial_i \theta + \frac{\hbar^2}{m} [R^2 (d_s \theta)^2 + (d_s R)^2] + 4R^2 U. \quad (2.57)$$

The divergence is now [compare (1.34)]:

$$\partial_\mu T_\nu^\mu = n \partial_\nu U. \quad (2.58)$$

**§ 3. Planck's constant as a density.** – Before we go onto the examination of some special cases of the Schrödinger equation, let us pause to introduce a consistent “density-only” formulation of the Madelung-Takabayasi equations. First, we shall discuss the possibility that Planck's constant itself is only the integral of a density over the support of the quantum wave function, just like the total mass, total energy, or total charge.

We pointed out above that there is a fundamental inconsistency in the original Madelung-Takabayasi form of the Schrödinger equation that is based in the fact that since the statistical interpretation of quantum mechanics was directed towards point-like matter as its classical limit, not extended matter, the physical quantities that one was dealing with, such as  $m$ ,  $E$ ,  $P_i$ , were numbers, not functions (viz., total mass, energy, momentum, resp.), even though the kinematical quantities, such as  $\omega$  and  $k_i$ , that they were being coupled to were functions whose supports were contained within that of the quantum wave function  $\Psi$ . Hence, the suggestion was made that one should be properly dealing with densities, not the total values that one only obtains from integrating those densities over the support of the wave function.

The basic density that we started with was the number density  $n = R^2$ , which represents a kinematical variable that related only to the number of particles present in a given space-time region, and indeed, it is entirely possible that only one (extended) particle is present in all of space-time. One notices that the dynamical quantities that get associated with the kinematical state, which we will regard as  $(R, \theta)$  and all of the higher derivatives of these functions that might be relevant to the discussion, start off as total

quantities, not densities, and that the constant  $\hbar$  is invariably involved with the association. The question arises of whether one can convert all of the total quantities to densities in a manner that will permit one to formulate the Madelung-Takabayasi picture solely in terms of the appropriate densities.

In particular, as we pointed out, the basic constitutive association  $P = \hbar k = \hbar d\theta$  is inconsistent because  $P$  represents a total energy-momentum 1-form that is defined only along a curve, while  $k$  is defined over a region that is contained in the support of  $\theta$ , which is contained in the support of  $\Psi$ . Hence, we need to convert  $P$  to an energy-momentum density. However, the units would not be consistent anymore, because  $\hbar$  has the units of energy-time, action, or angular momentum, not densities of those things.

Thus, if we decide that our first attempt at turning totals into densities is to always multiply them by  $n$  then we will have to multiply both sides of the basic constitutive association by  $n$  in order to keep the units consistent. One sees that this prescription will always be consistent with the definition of total quantities as integrals of densities, except that we are implicitly assuming that the various dynamical densities are all proportional to each other as space-time functions, which could only be a first approximation to something more involved. Hence, we shall accept that limitation and proceed heuristically in order to simply explore the effects of making those definitions.

In the case of  $P = -E dt + P_s$ , we define the energy-momentum density 1-form by:

$$p = nP = -\varepsilon dt + p_s, \quad (3.1)$$

which makes the energy and momentum densities take the form of:

$$\varepsilon = nE, \quad p_s = nP_s. \quad (3.2)$$

If we then multiply the right-hand side of the constitutive relation by  $n$ , as well, then we will get:

$$p = \hbar nk. \quad (3.3)$$

Now, since  $k$  is a kinematical quantity, it would make no physical sense to absorb  $n$  into its definition, since the concepts of frequency density and wave number density are physically absurd. However, one could make a better case for absorbing the  $n$  into the definition of  $\hbar$ , which only means that it would no longer be a constant, but a space-time function with units of energy-density-seconds, action density, or angular momentum density. The value of  $\hbar$  (which has, of course, been quite well established as a universal physical constant by countless experiments) then becomes the integral of the density:

$$\hbar(t, x^i) = \frac{1}{2\pi} \hbar(t, x^i) = \hbar n(t, x^i) \quad (3.4)$$

over the spatial support of the wave function  $\Psi$ :

$$\hbar = \int_{\text{supp}(\Psi)} \hbar(t, x^i) dV_s. \quad (3.5)$$

In order for this to make physical sense, the spatial integral would always have to be time-invariant and it would have to be the same for all physical-possible wave functions, which might suggest a limited number of elementary forms that their supports could take. It could also suggest that the differences in the actual values of  $\hbar$  between the various possibilities are immeasurably small, at this point in time.

We can now rewrite the basic constitutive relation in the form:

$$p = \hbar k = \hbar d\theta, \quad \varepsilon = \hbar \omega = -\hbar \partial_t \theta, \quad p_i = \hbar k_i = \hbar \partial_i \theta. \quad (3.6)$$

This also means that  $p$  does not have to be an exact differential form, anymore, even when  $k$  is exact, since its dynamical vorticity is now:

$$\Omega_d = d \wedge p = d\hbar \wedge d\theta, \quad (3.7)$$

which will vanish iff  $d\hbar$  is collinear with  $d\theta$ .

Note that the character of canonical quantization itself will change, since although the kinematical variables  $\omega$ ,  $k_i$  will still go to the same operators, the dynamical ones  $\varepsilon$ ,  $p_i$  will now take the form:

$$\varepsilon = i\hbar(t, x^i) \frac{\partial}{\partial t}, \quad p_i = \frac{\hbar(t, x^j)}{i} \frac{\partial}{\partial x^i}, \quad (3.8)$$

which are still linear operators, but they no longer have constant coefficients.

We can now define the spatial covelocity 1-form  $v$  in terms of only densities, since:

$$v_i = \frac{1}{m} P_i = \frac{1}{mn} n P_i = \frac{1}{\rho} p_i, \quad (3.9)$$

which makes:

$$p_i = \rho v_i, \quad (3.10)$$

which is consistent with the usual way of obtaining a momentum density in non-relativistic continuum mechanics.

The kinetic energy density can be expressed as:

$$\left(\frac{1}{2}mv^2\right)n = \frac{1}{2} \rho v^2 = \left(\frac{1}{2m} P^2\right)n = \frac{1}{2\rho} p^2. \quad (3.11)$$

These expressions are also consistent with the classical ones.

Although we have been using  $\hbar$  up to now to embody the quantum aspects of the Madelung medium, we shall eventually find that it is more convenient to work with one-half that quantity, which we denote by:

$$\eta = \frac{\hbar}{2} n, \quad (3.12)$$

and refer to as the *dilatation potential* of the wave function, so its differential:

$$\pi \equiv d\eta = \frac{\hbar}{2} dn = \left( \frac{\hbar}{2} \partial_i n \right) dt - \left( \frac{\hbar}{2} \partial_i n \right) dx^i, \quad (3.13)$$

which will then have the dimensions of either an energy density, a momentum flux, or a pressure, shall be called the *dilatation pressure* for the wave function  $\Psi$ .

Clearly, the dynamical vorticity of  $\pi$  will vanish:

$$\Omega_{\pi,d} = d_{\wedge} \pi = 0. \quad (3.14)$$

The rate of strain of the integral curves of the vector field that is associated with  $\pi_s$  is then:

$$\dot{\epsilon}_{ij} = 2(\partial_i \pi_j + \partial_j \pi_i) = \hbar \partial_{ij} n = 2 \partial_{ij} \eta, \quad (3.15)$$

so the rate of dilatation will be:

$$\dot{\lambda}_d = \dot{\epsilon}_i^i = \hbar \Delta n = 2 \Delta \eta = 2 \operatorname{div} \pi. \quad (3.16)$$

We can also define a *specific dilatation pressure* by way of:

$$v = \frac{1}{\rho} \pi = \frac{\hbar}{2} \frac{dn}{n} = d \left( \frac{\hbar}{2} \ln n \right). \quad (3.17)$$

Since this covelocity 1-form clearly admits a covelocity potential function:

$$\eta_k = \frac{\hbar}{2} \ln n, \quad (3.18)$$

its kinematical vorticity will also vanish:

$$\Omega_{v,k} = d_{\wedge} v = 0. \quad (3.19)$$

Meanwhile, its rate of strain will take the form:

$$\dot{\epsilon}_{ij} = 2(\partial_i v_j + \partial_j v_i) = \frac{\hbar}{m} \partial_{ij} \ln \eta, \quad (3.20)$$

so the rate of dilatation will be:

$$\dot{\lambda}_k = \dot{\epsilon}_i^i = \frac{\hbar}{m} \Delta \ln \eta. \quad (3.21)$$

As we shall see,  $\pi$  is just as fundamental as the quantum potential, since we can also use it to eliminate the explicit mention of  $\hbar$ .

The external force that acts upon an extended, deformable object as a body force must take the form of a force density 1-form  $f$ , in order for it to be consistent with  $\rho a$ , where  $a$  is the co-acceleration 1-form  $a = a_i dx^i$  <sup>(1)</sup>. However, if one multiplies an exact 1-form by a non-constant function then the resulting 1-form will not generally be exact, anymore. Hence, the result of multiplying a conservative force field by a density will generally be a non-conservative force density.

Upon closer inspection, one sees that it is not necessary for a force density to be conservative, even in the case of such forces as gravitation and electrostatic attraction and repulsion. That is because the force density will usually take the form  $\lambda dU$  in such cases, not  $dU$ . For instance, the density of gravitational force that acts upon a distributed mass of density  $\rho$  will be  $\rho g$ , where  $g$  is the spatial 1-form that corresponds to the gravitational acceleration at each point. Similarly, the density of electrostatic force that acts upon a distributed charge of density  $\sigma$  will be  $\sigma E$ , where  $E$  is the spatial electric field strength 1-form. Since both  $g$  and  $E$  are exact 1-forms, there are only two cases in which the combined force density 1-form ( $\rho g$  or  $\sigma E$ , resp.) is also exact:

1. The density is constant in space.
2. The gradient of the density is collinear with the applied force.

Thus, we conclude that the best way to convert forces into force densities is to multiply by the relevant density, and accept that the resulting 1-form might no longer be closed or exact.

Since Takabayasi's expression for the quantum stress tensor  $\sigma_j^i$  produced a force density upon taking the divergence:

$$\partial_j \sigma_i^j = -n \partial_i U_h, \quad (3.22)$$

it is not necessary to multiply by  $n$ , and after converting it to density form, we will get:

$$\sigma_{ij} = -\frac{\hbar^2}{m} \eta \partial_{ij} \ln \eta = -\frac{1}{2} \eta \epsilon_{ij}. \quad (3.23)$$

Thus, the Takabayasi stress tensor is a close relative to the rate of strain of the vector field  $\mathbf{v}_s$ .

The quantum potential energy density  $nU_h$  has the density form:

$$nU_h = \frac{1}{2} \rho v_s^2 - \frac{\eta}{\rho} \Delta \eta = \frac{1}{2} \rho v_s^2 - \frac{1}{2} \eta \dot{\lambda}_k = -\frac{1}{2} \rho v_s^2 - \eta \operatorname{div} \mathbf{v}. \quad (3.24)$$

Its corresponding quantum force density will then be:

---

<sup>(1)</sup> Non-relativistic acceleration will have a vanishing temporal component, since the temporal component of a non-relativistic velocity will always be 1.

$$f_{\hbar} = -n d_s U_{\hbar}, \quad (3.25)$$

although its explicit expression is quite elaborate and rather difficult to interpret physically.

**§ 4. The Madelung-Takabayasi equations in density form.** – We first summarize the definitions that we have introduced in a table that we shall use consistently in what follows.

Table IV.1. Definitions of continuum-mechanical quantities in terms of quantum-mechanical ones.

Number density	$n$	$R^2$
Mass density	$\rho$	$m n = m R^2$
Energy density	$\varepsilon$	$nE = -\rho\dot{\phi}$
(Co)velocity potential	$\phi$	$\frac{\hbar}{m}\theta = \frac{S}{m}$
Covelocity 1-form	$v$	$d_s\phi = \frac{\hbar}{m}d_s\theta = \frac{\hbar}{m}k$
Momentum density 1-form	$p_s$	$\rho v = n\hbar k_s$
Dilatation potential	$\eta$	$\frac{\hbar}{2}n = \frac{\hbar}{2}R^2 = \frac{\hbar}{2m}\rho$
Dilatation pressure	$\pi_s$	$d_s\eta = \frac{\hbar}{2}d_s n = \hbar R dR$
Specific dilatation pressure	$v_s$	$\frac{\pi_s}{\rho} = \frac{\hbar}{2m} \frac{d_s n}{n} = \frac{\hbar}{m} \frac{d_s R}{R}$

One can derive some useful relations from this table:

$$\pi_s = \rho v_s, \quad v_s = d_s \left( \frac{\hbar}{2m} \ln n \right), \quad d_s \rho = \frac{2m}{\hbar} \pi_s. \quad (4.1)$$

The time-varying Madelung-Takabayasi equations previously took the form:

$$E = \frac{1}{2m} P_s^2 + U + U_h, \quad \partial_t n = -\operatorname{div}(n \mathbf{v}_s). \quad (4.2)$$

In order to convert the first equation (viz., energy balance) into density form, one needs only to multiply both sides by  $n$ , while if one multiplies both sides of the second equation by  $m$  then it will become a dynamical equation, not a kinematical one. The net result of both operations is the set of dynamical equations:

$$\varepsilon = \frac{1}{2\rho} p_s^2 + n(U + U_h), \quad \partial_t \rho = -\operatorname{div}(\rho \mathbf{v}), \quad (4.3)$$

which contain only densities and no explicit mention of  $\hbar$ .

Previously, the Euler equation that corresponded to the first equation in (4.2) took the form:

$$L_{\mathbf{v}} P_i = -\partial_i (U + U_h). \quad (4.4)$$

If one takes into account the fact that the second equation in (4.3) also takes the form:

$$L_{\mathbf{v}} n = -n \operatorname{div} \mathbf{v} \quad (4.5)$$

and the fact that:

$$L_{\mathbf{v}} (nP_i) = (L_{\mathbf{v}} n)P_i + n L_{\mathbf{v}} P_i = -n P_i \operatorname{div} \mathbf{v} + n L_{\mathbf{v}} P_i, \quad (4.6)$$

so

$$n L_{\mathbf{v}} P_i = L_{\mathbf{v}} p_i + p_i \operatorname{div} \mathbf{v}, \quad (4.7),$$

then the Euler equation will now take the form:

$$L_{\mathbf{v}} p_i + p_i \operatorname{div} \mathbf{v} = -n \partial_i (U + U_h), \quad (4.8)$$

or, if one prefers:

$$\frac{\partial p_i}{\partial t} + \frac{\partial}{\partial x^j} (v^j p_i) = -n \partial_i (U + U_h). \quad (4.9)$$

Previously, the time-varying Madelung-Takabayasi Lagrangian took the form:

$$\mathcal{L} = -R^2 \left[ \hbar \dot{\theta} + \frac{\hbar^2}{2m} (d_s \theta)^2 + U \right] + \frac{\hbar^2}{2m} (d_s R)^2. \quad (4.10)$$

With the substitutions above, one can now give it the form:

$$\mathcal{L}(\phi, \eta, \dot{\phi}, v_i, \dot{\eta}, \pi_i) = -\frac{1}{2} \rho (v^2 + v_s^2) + n (E - U), \quad (4.11)$$

which now suggests that the quantum contribution to the motion can involve the momentum, as well as the potential energy.

For a solution to the Madelung-Takabayasi equations (4.3),  $\mathcal{L}$  will take the form:

$$\mathcal{L} = -\frac{1}{2}\rho v_s^2 + nU_{\hbar} = -\rho v_s^2 - \eta \operatorname{div} \mathbf{v}, \quad (4.12)$$

in which we have used (3.24) in order to replace  $nU_{\hbar}$ .

The canonical forces and momenta that are associated with  $\mathcal{L}$  are then:

$$f_{\phi} = \frac{\partial \mathcal{L}}{\partial \phi} = 0, \quad \Pi_{\phi}^t = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = -\rho, \quad \Pi_{\phi}^i = \frac{\partial \mathcal{L}}{\partial v_i} = -\rho v^i, \quad (4.13)$$

$$f_{\eta} = \frac{\partial \mathcal{L}}{\partial \eta} = \frac{2}{\hbar} \left[ -\frac{1}{2}m(v^2 - v_s^2) + E - U \right], \quad \Pi_{\eta}^t = \frac{\partial \mathcal{L}}{\partial \dot{\eta}} = 0, \quad \Pi_{\eta}^i = \frac{\partial \mathcal{L}}{\partial \pi_i} = -v^i. \quad (4.14)$$

The Euler-Lagrange equations take the immediate forms:

$$0 = f_{\phi} - \partial_t \Pi_{\phi}^t - \partial_i \Pi_{\phi}^i = \dot{\rho} + \operatorname{div} \mathbf{p}_s,$$

$$0 = f_{\eta} - \partial_t \Pi_{\eta}^t - \partial_i \Pi_{\eta}^i = \frac{2}{\hbar} \left[ -\frac{1}{2}m(v^2 - v_s^2) + E - U \right] + \operatorname{div} \mathbf{v}.$$

The first of these is the conservation of mass, while the second one needs to be multiplied by  $\eta$  in order to give it its final form:

$$\varepsilon = \frac{1}{2}\rho(v^2 - v_s^2) + nU - \eta \operatorname{div} \mathbf{v} = \frac{1}{2}\rho v^2 + n(U + U_{\hbar}). \quad (4.15)$$

The density form of the conserved current that is associated with the global phase invariance of the original quantum wave equation changes only by being multiplied by the constant factor  $m$ , so one will now have:

$$\mathbf{J} = (\rho, \rho v^i) = \rho (1, v^i). \quad (4.16)$$

One can also put the energy-momentum-stress tensor into density form as [compare (2.53) to (2.56)]:

$$T_0^0 = \frac{1}{2}\rho(v^2 + v_s^2) + nU, \quad (4.17)$$

$$T_j^0 = -\rho v_j = -p_j, \quad (4.18)$$

$$T_0^i = \varepsilon v^i - \dot{\eta} v_s^i, \quad (4.19)$$

$$T_j^i = -\rho(v^i v_j + v^i v_j) - \mathcal{L} \delta_j^i, \quad (4.20)$$

so its trace will be [compare (2.57)]:

$$T_{\mu}^{\mu} = -3\varepsilon + \rho(v^2 + v_s^2) + 4nU. \quad (4.21)$$



The divergence of  $T_\nu^\mu$  will then be:

$$\partial_\mu T_\nu^\mu = n \partial_\nu U, \quad (4.22)$$

which is identical to (2.58).

**§ 5. The stationary Schrödinger equation.** Actually, most of the solved problems that one studies in basic non-relativistic quantum mechanics do not involve the time-varying Schrödinger equation, which is usually treated only perturbatively, but the time-invariant – or *stationary* – Schrödinger equation. The wave functions that define its solutions will then take the form of standing waves. Of course, whether or not one can use that form of the Schrödinger equation depends entirely upon the nature of the potential function  $U(t, x^i)$ , and in particular, it is only when it does not depend upon  $t$  that one can turn to the stationary form of the Schrödinger equation.

If one does, in fact, have a time-invariant potential function  $U(x^i)$  then one can assume that the wave function  $\Psi(t, x^i)$  has the standing-wave form:

$$\Psi(t, x^i) = T(t) \psi(x^i), \quad (5.1)$$

in which the variable  $t$  has been separated from the variables  $x^i$ .

If one then substitutes this form for the wave function  $\Psi$  into the time-varying Schrödinger equation (1.18) then the result will be:

$$i\hbar \frac{\dot{T}}{T} = - \frac{\hbar^2}{2m} \frac{\Delta \psi}{\psi} + U.$$

Since the left-hand side of this equation is a function of only  $t$ , while the right-hand side is a function of only  $x^i$ , that situation can happen only if both functions are constant. If we call that separation constant  $E$  then we will get a pair of differential equations for the functions  $T$  and  $\psi$ :

$$\dot{T} = -\frac{iE}{\hbar} T, \quad -\frac{\hbar^2}{2m} \Delta \psi + U \psi = E \psi. \quad (5.2)$$

The first equation can be solved immediately and gives a solution of the form:

$$T(t) = e^{-iEt/\hbar}, \quad (5.3)$$

whereas if one introduces the *Hamiltonian* operator:

$$H = -\frac{\hbar^2}{2m} \Delta + U \quad (5.4)$$

then the second equation in (5.2) will take on the *eigenvalue* form:

$$H\psi = E\psi. \quad (5.5)$$

Thus, the spatial wave functions that define the envelope of the standing wave will be eigenfunctions of the Hamiltonian operator  $H$  that are associated with eigenvalues  $E$ , which will represent the total energy (i.e., kinetic + potential) of the motion of the standing wave.

One can derive the stationary Schrödinger equation from a Lagrangian, just as one can in the time-varying case. This time, the Lagrangian will be:

$$\mathcal{L} = \frac{\hbar^2}{2m} \|d_s \psi\|^2 + (U - E) \|\psi\|^2, \quad (5.6)$$

which differs from the time-varying expression (1.19) by an overall sign and the replacement of the time-varying contribution with  $E \|\psi\|^2$ .

The generalized forces and momenta now take the form:

$$f^* = \frac{\partial \mathcal{L}}{\partial \psi^*} = (U - E)\psi, \quad \Pi^{i*} = \frac{\partial \mathcal{L}}{\partial \dot{\psi}^*} = 0, \quad \Pi^{j*} = \frac{\partial \mathcal{L}}{\partial (d_s \psi^*)} = \frac{\hbar^2}{2m} \delta^{ij} \partial_j \psi, \quad (5.7)$$

along with their complex conjugates.

Since the wave function  $\psi$  is a purely spatial in this case, the Euler-Lagrange equations will represent an equilibrium condition for the shape of its envelope, and one will get the equation for  $\psi$  by annulling the variational derivative of  $\mathcal{L}$  with respect to  $\psi^*$ :

$$0 = \frac{\delta \mathcal{L}}{\delta \psi^*} = f^* - \partial_i \Pi^{i*}. \quad (5.8)$$

Under the infinitesimal phase transformation:

$$\delta \psi = i\varepsilon \psi, \quad \delta \psi^* = -i\varepsilon \psi^*, \quad (5.9)$$

the components of the associated Noether current will now take the form:

$$J^0 = 0, \quad J^i = \frac{\hbar}{2mi} \delta^{ij} (\partial_j \psi \psi^* - \psi \partial_j \psi^*) \quad (5.10)$$

when we drop the overall scalar factor of  $2\hbar\varepsilon$ . In form, this current differs from the time-varying analogue (1.27) by only by the vanishing of  $J^0$ .

The energy-momentum-stress tensor will then take the form:

$$T_0^0 = -\mathcal{L}, \quad T_j^0 = T_0^j = 0, \quad (5.11)$$

$$T_j^i = \frac{\hbar^2}{2m} (\partial^i \psi^* \partial_j \psi + \partial^i \psi \partial_j \psi^*) - \delta_j^i \mathcal{L}, \quad (5.12)$$

and its trace will be:

$$T_{\mu}^{\mu} = -\frac{\hbar^2}{m} \|d_s \psi\|^2 + 4(E - U) \|\psi\|^2. \quad (5.13)$$

The main formal differences between the stationary form of  $T_{\nu}^{\mu}$  and the time-varying one that is described in equations (1.29) to (1.32) amount to the vanishing of the contribution to  $T_0^0$  from  $\Pi^i \dot{\Psi} + \Pi^{i*} \dot{\Psi}^*$ , the vanishing of  $T_j^0$  and  $T_0^i$ , and a change of sign in the first term of  $T_j^i$ , which comes from the change in sign of the kinetic term in  $\mathcal{L}$ .

The divergence of  $T_{\nu}^{\mu}$  is the covector whose components are:

$$\partial_{\mu} T_0^{\mu} = -\partial_t \mathcal{L} = 0, \quad \partial_i T_j^i = -\partial_j U \|\psi\|^2, \quad (5.14)$$

which reflects the fact that the Lagrangian density is time-invariant, and it has changed sign.

### § 6. The Madelung-Takabayasi form of the stationary Schrödinger equation.

Under the assumption that the potential function  $U$  is time-invariant, one now introduces polar coordinates into the complex plane, as before, and expresses the spatial eigenfunction  $\psi$  in the form:

$$\psi(x) = R(x) e^{i\theta(x)}. \quad (6.1)$$

If one substitutes this into the stationary Schrödinger equation (5.5), in which  $H$  is defined by (5.4), then the ultimate result of the separation of the real and imaginary equations will be the pair of equations for the functions  $R$ ,  $\theta$ :

$$E = \frac{P_s^2}{2m} + U + U_h, \quad 0 = R \operatorname{div} \mathbf{P}_s + 2\langle P_s, d_s R \rangle, \quad (6.2)$$

in which we are using the same notations as above.

The only things that have changed in the energy balance equation are that now the left-hand side is a constant, and the potential functions in the right-hand side are functions of  $x$ , but not  $t$ .

In order to put these equations into their density form, we multiply the first one by  $n$  and the second one by  $mR$ , which will make then take the form:

$$\mathcal{E} = \frac{1}{2} \rho v^2 + n(U + U_h), \quad 0 = \operatorname{div} \mathbf{p}_s. \quad (6.3)$$

If one takes spatial gradients of both sides of the first one then the resulting equilibrium equation will take the form:

$$\mathbf{L}_{\nu} p_i = -n \partial_i (U + U_h), \quad (6.4)$$

in which the Lie derivative now takes the form:

$$\mathbf{L}_v p_i = v^j \partial_j p_i. \quad (6.5)$$

The second equation of (6.3) says that the Madelung medium is dynamically incompressible in the stationary case. In the event that  $\rho$  is constant, it would also be kinematically incompressible. However, the quantum potential would necessarily vanish for such a mass distribution, and one would not be dealing with a quantum motion.

Since:

$$\rho(x) = m R(x)^2, \quad v(x) = \frac{\hbar}{m} d_s \theta(x), \quad (6.6)$$

the basic flow variables are no longer functions of  $t$ , but only  $x$ , so one can think of the motion in question as being a *steady* motion. Consistent with that, the first equation in (6.3) can be regarded as a form of *Bernoulli's theorem* on the steady flows in fluids, in which one can regard the first term on the right as dynamic pressure, as well as a kinetic energy density, and the quantum potential can play the double role of an internal pressure. Since Madelung was dealing with only the stationary Schrödinger equation to begin with, these facts led him to conjecture that the medium in question was a fluid, although, as we pointed out before, later work by Takabayasi showed that the stress tensor was inconsistent with that form of matter.

One can similarly give a Lagrangian form for the stationary Madelung-Takabayasi equations by substituting  $\psi = R e^{i\theta}$  in the stationary Schrödinger Lagrangian density, which will give:

$$\mathcal{L} = R^2 \left[ \frac{\hbar^2}{2m} (d_s \theta)^2 + U - E \right] + \frac{\hbar^2}{2m} (d_s R)^2, \quad (6.7)$$

in which  $E$  is a constant, this time.

This Lagrangian density can also be expressed in density form as:

$$\mathcal{L}(\phi, \eta, v_i, p_i) = \frac{1}{2} \rho (v^2 + v_s^2) + n(U - E), \quad (6.8)$$

which has a physical reasonableness to it that stems from its transparent simplicity.

For a solution to the equations (6.3),  $\mathcal{L}$  will then take the form:

$$\mathcal{L} = \frac{1}{2} \rho v_s^2 - n U_h = \rho v_s^2 + \eta \operatorname{div} \mathbf{v}, \quad (6.9)$$

which differs from the corresponding expression for the time-varying case (4.12) by an overall sign change.

The canonical forces and momenta that are associated with  $\mathcal{L}$  are then:

$$f_\phi = \frac{\partial \mathcal{L}}{\partial \phi} = 0, \quad \Pi_\phi^i = \frac{\partial \mathcal{L}}{\partial v_i} = \rho v^i, \quad (6.10)$$

$$f_\eta = \frac{\partial \mathcal{L}}{\partial \eta} = \frac{2}{\hbar} \left[ \frac{1}{2} m (v^2 - v_s^2) + U - E \right], \quad \Pi_\eta^i = \frac{\partial \mathcal{L}}{\partial \pi_i} = v^i. \quad (6.11)$$

The equilibrium equations for the functions  $\phi$  and  $\eta$  are then obtained by varying  $\mathcal{L}$  with respect to  $\phi$  and  $\eta$  and take the form:

$$0 = \operatorname{div} \mathbf{p}_s, \quad 0 = \frac{2}{\hbar} \left[ \frac{1}{2} m (v^2 - v_s^2) + U - E \right] - \operatorname{div} \mathbf{v}, \quad (6.12)$$

although if one multiplies the second one by  $\eta$  then it can be given the density form:

$$\varepsilon = \frac{1}{2} \rho (v^2 - v_s^2) + nU - \eta \operatorname{div} \mathbf{v}_s = \frac{1}{2} \rho v^2 + n(U + U_h), \quad (6.13)$$

which is consistent with the first of (6.3).

Similarly, the Noether current that is associated with the phase-invariance of the function  $\phi$  will take the form [compare the complex expressions (5.10)]:

$$J^0 = \Pi_\phi^t \delta \phi = 0, \quad J^i = \Pi_\phi^i \delta \phi = \rho v^i, \quad (6.14)$$

if we ignore an overall constant factor.

The density form of the energy-momentum-stress tensor is now:

$$T_\nu^\mu = \rho (v^\mu v_\nu + v^\nu v_\mu) - \mathcal{L} \delta_\nu^\mu, \quad (6.15)$$

whose individual submatrices are [compare (5.11) and (5.12)]:

$$T_0^0 = -\mathcal{L}, \quad T_0^i = T_j^0 = 0, \quad T_j^i = \rho (v^i v_j + v^j v_i) - \mathcal{L} \delta_j^i, \quad (6.16)$$

and its trace will take the form [compare (5.13)]:

$$T_\mu^\mu = -\rho (v^2 + v_s^2) + 4(\varepsilon - nU). \quad (6.17)$$

The divergence of  $T_\nu^\mu$  is now:

$$\partial_\mu T_0^\mu = -\partial_t \mathcal{L} = 0, \quad \partial_i T_j^i = -n \partial_j U, \quad (6.18)$$

which is simply the real form of the complex result (5.14).

**§ 7. Coupling to an external electromagnetic field.** – Many of the most important solved problems and experimental results in quantum mechanics were concerned with the motion of charges and magnetic dipoles in an external electromagnetic field. We shall

model such a field by a 2-form on a four-dimensional space-time, although that anticipates the relativistic formulation of the problem, to some extent. That 2-form is the *Minkowski electromagnetic field strength 2-form*:

$$F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu = c dt \wedge E_i dx^i + \frac{1}{2} B_{ij} dx^i \wedge dx^j. \quad (7.1)$$

Note that we are then modeling the magnetic field strength as being *intrinsically* a spatial 2-form  $B = \frac{1}{2} B_{ij} dx^i \wedge dx^j$ , while the usual spatial vector field  $\mathbf{B}$  that one encounters is the (spatial) Poincaré dual of that 2-form using spatial volume element  $dx^1 \wedge dx^2 \wedge dx^3$ , so the components of  $\mathbf{B}$  would be:

$$B^i = \frac{1}{2} \epsilon^{ijk} B_{jk}. \quad (7.2)$$

*a. Coupling the electromagnetic field to the wave function.* If the extended object whose total mass is  $m$  is also associated with a total charge  $Q$  then coupling the motion of that charge to the external electromagnetic field  $F$  is usually achieved by means of “minimal electromagnetic coupling,” which means that the (total) energy-momentum 1-form <sup>(1)</sup>  $P = -\mathcal{E} c dt + P_s$  gets replaced with:

$$P = \tilde{P} + \frac{Q}{c} A = -(\tilde{\mathcal{E}} + Q\phi) c dt + \left( \tilde{P}_i + \frac{Q}{c} A_i \right) dx^i, \quad (7.3)$$

in which:

$$\tilde{\mathcal{E}} = -\hbar \partial_t \theta, \quad \tilde{P}_s = \hbar d_s \theta. \quad (7.4)$$

We shall think of  $\tilde{\mathcal{E}}$  as the *generalized total energy* and  $\tilde{P}$  as the *generalized total momentum*, while  $\mathcal{E}$  and  $P_s$  represent the actual values of those physical quantities.

The 1-form:

$$A = -\phi c dt + A_s \quad (7.5)$$

represents a choice of *electromagnetic potential 1-form*.  $A$  then satisfies the constraint that:

$$F = dA = \frac{1}{2} (\partial_\mu A_\nu - \partial_\nu A_\mu) dx^\mu \wedge dx^\nu \quad (x^0 = ct). \quad (7.6)$$

Of course,  $A$  is not unique, since any other 1-form that differs from  $A$  by a closed 1-form will produce the same  $F$  under exterior differentiation, and that ambiguity in  $A$  is referred to as the *gauge invariance* of  $F$ , while a choice of  $A$  is a choice of gauge.

If one puts  $F$  into time + space form:

$$F = -(c^{-1} \partial_t A_i + \partial_i \phi) dx^0 \wedge dx^i + \frac{1}{2} (\partial_i A_j - \partial_j A_i) dx^i \wedge dx^j \quad (7.7)$$

then one can say that:

---

<sup>(1)</sup> In order to avoid confusion, from now on, the notation,  $E$  will refer to the spatial 1-form  $E_i dx^i$ , while the total energy will be denoted by  $\mathcal{E}$ .

$$E_i = -\frac{1}{c} \partial_t A_i - \partial_i \phi, \quad B_{ij} = \partial_i A_j - \partial_j A_i. \quad (7.8)$$

The associated covelocity 1-form now takes the form:

$$v = \tilde{v} + \frac{Q}{mc} A = \frac{1}{m} \left( \tilde{P} + \frac{Q}{c} A \right), \quad (7.9)$$

so:

$$\tilde{v} = \frac{\hbar}{m} d_s \theta \quad (7.10)$$

becomes the “effective velocity” that is associated with the generalized momentum.

That means that, although  $\tilde{v}$  is (kinematically) irrotational, the velocity field  $v$  will not be irrotational, and in fact:

$$\Omega_k = d \wedge v = \frac{Q}{mc} F. \quad (7.11)$$

If we put the kinematical vorticity into time + space form:

$$\Omega_k = (\partial_t v_i - \partial_i v_0) dt \wedge dx^i + \frac{1}{2} (\partial_i v_j - \partial_j v_i) dx^i \wedge dx^j \quad (7.12)$$

then we will get the following pair of equations from (7.11):

$$\partial_t v_i - \partial_i v_0 = \frac{Q}{mc} E_i, \quad \partial_i v_j - \partial_j v_i = \frac{Q}{mc} B_{ij}. \quad (7.13)$$

Note that if  $\|B\|$  is the Euclidian norm of the 2-form  $B$  then the expression  $Q \|B\| / mc$  will be twice the Larmor frequency of the charge  $Q$  in the magnetic field  $B$ .

*b. The minimally-coupled, time-varying Schrödinger equation.* The minimally-coupled energy-momentum 1-form becomes the pair of operators:

$$E = i\hbar \frac{\partial}{\partial t} - Q\phi, \quad P_i = \frac{\hbar}{i} \frac{\partial}{\partial x^i} + \frac{Q}{c} A_i = \frac{\hbar}{i} \nabla_i, \quad (7.14)$$

in which we have defined:

$$\nabla_i \equiv \frac{\partial}{\partial x^i} + \frac{iQ}{\hbar c} A_i. \quad (7.15)$$

The linear, first-order differential operator  $\nabla_i$  basically amounts to the “spatial covariant derivative” operator for the  $u(1)$ -connection that is defined by regarding  $\frac{iQ}{\hbar c} A_i$  as a

connection 1-form that takes its values in the Lie algebra  $\mathfrak{u}(1)$  of  $U(1)$ , namely, the imaginary line through 1; its complex conjugate is then:

$$\nabla_i^* = \frac{\partial}{\partial x^i} - \frac{iQ}{\hbar c} A_i. \quad (7.16)$$

The time-varying Schrödinger equation for a point charge  $Q$  that is coupled to an external electromagnetic field will take the form:

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \delta^{ij} \nabla_i \nabla_j \Psi + Q\phi \Psi. \quad (7.17)$$

It is important to see the subtlety that is associated with evaluating the square of the operator  $\nabla_i$  that comes from the difference between squaring it naively and squaring it when it is applied to the wave function  $\Psi$ :

$$\begin{aligned} \delta^{ij} \nabla_i \nabla_j \Psi &= \delta^{ij} \left( \frac{\partial}{\partial x^i} + \frac{iQ}{\hbar c} A_i \right) \left( \frac{\partial \Psi}{\partial x^j} + \frac{iQ}{\hbar c} A_j \Psi \right) \\ &= \Delta \Psi - \left( \frac{Q}{\hbar c} \right)^2 A^2 \Psi + \frac{iQ}{\hbar c} [(\partial_i A^i) \Psi + 2A^i \partial_i \Psi]. \end{aligned}$$

If one did not evaluate the operator on  $\Psi$  when squaring it then the result would differ by a missing factor of 2 in the brackets.

One can derive (7.17) and its complex-conjugate from the Lagrangian density:

$$\mathcal{L} = \frac{i\hbar}{2} (\Psi^* \dot{\Psi} - \dot{\Psi} \Psi^*) - \frac{\hbar^2}{2m} \|\nabla \Psi\|^2 - Q\phi \|\Psi\|^2, \quad (7.18)$$

which is simply the previous (non-magnetic) Lagrangian density (1.19) with the replacement of spatial partial derivatives with covariant derivatives and the potential energy  $U$  replaced with  $Q\phi$ .

One can also think of the new  $\mathcal{L}$  as the sum:

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_A \quad (7.19)$$

of the original  $\mathcal{L}$  in (1.19), which we now denote as  $\mathcal{L}_0$ , and a Lagrangian density  $\mathcal{L}_A$  that accounts for the interaction of  $Q$  and the external magnetic field:

$$\mathcal{L}_A = -\frac{Q^2}{2mc^2} A^2 \|\Psi\|^2 - \frac{i\hbar Q}{2mc} A_i (\Psi \partial^i \Psi^* - \Psi^* \partial^i \Psi), \quad (7.20)$$



which can also be put into the form:

$$\mathcal{L}_A = A_i j^i \quad (7.21)$$

when we define the current:

$$j^i = -\frac{Q^2}{2mc^2} A^i \|\Psi\|^2 - \frac{i\hbar Q}{2mc} (\Psi \partial^i \Psi^* - \Psi^* \partial^i \Psi). \quad (7.22)$$

The generalized forces and momenta for  $\mathcal{L}$  now take the form of:

$$f^* = \frac{i\hbar}{2} \dot{\Psi} - Q\phi\Psi + \frac{i\hbar Q}{2mc} A^i \nabla_i \Psi, \quad \Pi^{t*} = -\frac{i\hbar}{2} \Psi, \quad \Pi^{i*} = -\frac{\hbar^2}{2m} \delta^{ij} \nabla_j \Psi, \quad (7.23)$$

and their complex conjugates.

Hence, the generalized force and the spatial components of the generalized momentum now include contributions from the interaction of the charge with the external magnetic field, namely:

$$f_A^* = \frac{\partial \mathcal{L}_A}{\partial \Psi^*} = \frac{i\hbar Q}{2mc} A^i \nabla_i \Psi, \quad \Pi_A^{t*} = \frac{\partial \mathcal{L}_Q}{\partial \Psi^*} = 0, \quad \Pi_A^{i*} = \frac{\partial \mathcal{L}_Q}{\partial (\partial_i \Psi^*)} = -\frac{iQ\hbar}{2mc} A^i \Psi, \quad (7.24)$$

in addition to the expressions in (1.22), while the temporal component of the generalized momentum has not changed from its expression in (1.22).

One can check that equation (7.17) will follow from the Euler-Lagrange equations:

$$0 = \frac{\delta \mathcal{L}}{\delta \Psi^*} = f^* - \partial_t \Pi^{t*} - \partial_i \Pi^{i*}. \quad (7.25)$$

The components of the conserved current that is associated with the phase degree of freedom in  $\mathcal{L}$  now take the form:

$$J^0 = \Psi \Psi^*, \quad J^i = \frac{\hbar}{2mi} \delta^{ij} (\Psi^* \nabla_j \Psi - \Psi \nabla_j^* \Psi^*), \quad (7.26)$$

and when we expand the definition of the covariant derivatives, the spatial components will take the form:

$$J^i = \frac{\hbar}{2mi} \delta^{ij} (\Psi^* \partial_j \Psi - \Psi \partial_j \Psi^*) + \frac{Q}{mc} A^i \|\Psi\|^2. \quad (7.27)$$

Hence, the spatial part of the current (1.27) has been altered by an additive term that is proportional to the magnetic potential. We see that the current  $j^i$  that we defined above relates to  $J^i$  by way of:

$$j^i = \frac{Q}{c} J^i - \frac{3Q^2}{2mc} A^i \|\Psi\|^2. \quad (7.28)$$

However, when we compute the conserved current that is associated with the gauge symmetry of  $\mathcal{L}_A$  with respect to  $A$ , we will get a more direct relationship:

$$J_A^i = \frac{\partial \mathcal{L}_A}{\partial A_i} = \frac{Q}{c} J^i. \quad (7.29)$$

The energy-momentum-stress tensor is now:

$$T_\nu^\mu = \Pi^\mu \nabla_\nu \Psi + \Pi^{\mu*} \nabla_\nu \Psi^* - \mathcal{L} \delta_\nu^\mu, \quad (7.30)$$

whose individual submatrices are:

$$T_0^0 = \frac{\hbar^2}{2m} \|\nabla \Psi\|^2 + Q\phi \|\Psi\|^2 \equiv \mathcal{H}, \quad (7.31)$$

$$T_j^0 = \frac{i\hbar}{2} (\Psi^* \nabla_j \Psi - \Psi \nabla_j \Psi^*) = -m J_j, \quad (7.32)$$

$$T_0^i = -\frac{\hbar^2}{2m} \delta^{ij} (\dot{\Psi} \nabla_j^* \Psi^* + \dot{\Psi}^* \nabla_j \Psi), \quad (7.33)$$

$$T_j^i = -\frac{\hbar^2}{2m} (\nabla^i \Psi \nabla_j^* \Psi^* + \nabla^{i*} \Psi^* \nabla_j \Psi) - \mathcal{L} \delta_j^i. \quad (7.34)$$

The only difference between these expressions and (1.29)-(1.32), comes from the minimal coupling of the magnetic potential to the charge.

The trace of  $T_\nu^\mu$  is:

$$T_\mu^\mu = -\frac{3\hbar i}{2} (\Psi^* \dot{\Psi} - \Psi \dot{\Psi}^*) - \frac{2\hbar^2}{m} \|\nabla \Psi\|^2 - 2Q\phi \|\Psi\|^2, \quad (7.35)$$

which is the minimally-coupled expression for (1.33).

The divergence of  $T_\nu^\mu$  splits into:

$$\partial_\mu T_0^\mu = Q \|\Psi\|^2 \dot{\phi}, \quad \partial_\mu T_i^\mu = Q \|\Psi\|^2 E_i + B_{ij} J_A^j. \quad (7.36)$$

The difference between these expressions and (1.34) reflects the fact that the external force includes the Lorentz force of interaction between  $Q$  and the magnetic field, as well as its interaction with the electric one.

Since  $\mathcal{L}$  is the sum of two Lagrangian densities, one can also think of  $T_\nu^\mu$  as the sum  $\overset{\circ}{T}_\nu^\mu + \overset{A}{T}_\nu^\mu$  of two such tensors, the first of which  $\overset{\circ}{T}_\nu^\mu$  relates to the case with no magnetic

field, while the second one  $\overset{A}{T}_\nu^\mu$  can be obtained from  $\mathcal{L}_A$ . Its individual submatrices are then:

$$\overset{A}{T}_0^0 = -\mathcal{L}_A, \quad \overset{A}{T}_j^0 = 0, \quad \overset{A}{T}_0^i = -\frac{iQ\hbar}{2mc} A^i (\Psi \dot{\Psi}^* - \Psi^* \dot{\Psi}), \quad (7.37)$$

$$\overset{A}{T}_j^i = -\frac{iQ\hbar}{2mc} A^i (\Psi \partial_j \Psi^* - \Psi^* \partial_j \Psi) - \mathcal{L}_A \delta_j^i. \quad (7.38)$$

Hence,  $T_\mu^\mu$  has been altered from its previous, non-magnetic expression additively only by the addition of the trace of  $\overset{A}{T}_\nu^\mu$ , namely:

$$\overset{A}{T}_\mu^\mu = \frac{Q}{c} A_i \left( 3J_0^i + \frac{2Q}{mc} A^i \|\Psi\|^2 \right). \quad (7.39)$$

One can also treat the divergence of  $T_\nu^\mu$  as the sum of two terms  $\partial_\mu \overset{o}{T}_\nu^\mu + \partial_\mu \overset{A}{T}_\nu^\mu$ , the first of which describes the contribution from the non-magnetic part of  $\mathcal{L}$  with  $U = -Q\phi$ , and the second of which contributes the Lorentz force between  $Q$  and  $B$ . However, one must be careful not to think that the first divergence still vanishes, as it did in the absence of an external electromagnetic field, since its vanishing followed from the wave equation for  $\Psi$ , which has changed since then.

*c. The Madelung-Takabayasi form of the minimally-coupled, time-varying wave equation.* The Madelung-Takabayasi equations that correspond to (7.17) will now take the form <sup>(1)</sup>:

$$\mathcal{E} = \frac{1}{2m} \bar{p}_s^2 + Q\phi + U_h, \quad \frac{\partial n}{\partial t} + \operatorname{div} \left[ \frac{n}{m} \bar{\mathbf{p}}_s \right] = 0, \quad (7.40)$$

in which:

$$\bar{p}_s = \hbar d_s \theta + \frac{Q}{c} A_s. \quad (7.41)$$

This also takes the form of the previous equations with a minimal coupling (7.3) of the magnetic potential 1-form to the (total) momentum 1-form.

We put (7.40) into density form by multiplying the first one by  $n$  and the second one by  $m$ :

$$\mathcal{E} = \frac{1}{2\rho} p_s^2 + \sigma\phi + nU_h, \quad \frac{\partial \rho}{\partial t} + \operatorname{div} \mathbf{p}_s = 0, \quad (7.42)$$

---

<sup>(1)</sup> We must now replace the symbol  $\phi$  for the velocity potential with  $\zeta$ , since we are now using  $\phi$  for the electric potential.

in which  $\sigma \equiv Qn$  is the electric charge density. These equations are virtually identical to (4.3) with  $nU = \sigma\phi$ . The first of these can also be given the form:

$$\varepsilon = \frac{1}{2} \rho v^2 + \sigma\phi + nU_h. \quad (7.43)$$

The covelocity 1-form  $v$  that is associated with  $p_s (= \rho v)$  takes the form:

$$v = \frac{\hbar}{m} d_s \theta + \frac{Q}{mc} A_s, \quad (7.44)$$

whose kinematical vorticity is then:

$$\Omega_k = d \wedge v = \frac{Q}{mc} (dt \wedge \dot{A}_s + B), \quad B_{ij} = \partial_i A_j - \partial_j A_i. \quad (7.45)$$

The dynamical vorticity that is associated with  $p_s$  then takes the form:

$$\Omega_d = d \wedge p_s = d\rho \wedge v + \rho \Omega_k. \quad (7.46)$$

The spatial part of the first equation in (7.45) gives a useful commutation rule for the derivatives of  $v_i$ :

$$\partial_i v_j = \partial_j v_i + \frac{Q}{mc} B_{ij}. \quad (7.47)$$

In order to go from (7.43), which expresses the balance of energy, to the balance of momentum, one takes the spatial gradients of both sides, while taking into account the fact that:

$$L_v A_i = \frac{1}{c} \frac{\partial A_i}{\partial t} + v^j \frac{\partial A_i}{\partial x^j}. \quad (7.48)$$

This ultimately alters the previous form (2.33) of the Euler equation only by the addition of the Lorentz force:

$$L_v p_i = \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x^i} (v^j p_j) = \sigma \left[ E_i + \frac{1}{c} B_{ij} v^j \right] - n \partial_i U_h. \quad (7.49)$$

In order to get the Lagrangian density for the minimally-coupled case that we are currently considering, we need only replace the previous expression for momentum with the new one:

$$\mathcal{L} = R^2 \left[ \hbar \dot{\theta} + \frac{1}{2m} \bar{p}_s^2 + Q\phi \right] + \frac{\hbar^2}{2m} (d_s R)^2 \quad (7.50)$$

by using the minimal coupling prescription.

The density form of this is:

$$\mathcal{L} = \frac{1}{2}\rho(v^2 + v_s^2) + \sigma\phi - \varepsilon, \quad (7.51)$$

with the understanding that  $v$  includes the magnetic potential 1-form  $A$ .

The canonical forces and momenta that are associated with  $\mathcal{L}$  are:

$$f_\zeta = \frac{\partial\mathcal{L}}{\partial\zeta} = 0, \quad \Pi_\zeta^t = \frac{\partial\mathcal{L}}{\partial\dot{\zeta}} = \rho, \quad \Pi_\zeta^i = \frac{\partial\mathcal{L}}{\partial v_i} = \rho v^i, \quad (7.52)$$

$$f_\eta = \frac{\partial\mathcal{L}}{\partial\eta} = \frac{2}{\hbar} \left[ \frac{1}{2}m(v^2 - v_s^2) + Q\phi + m\zeta \right], \quad \Pi_\eta^t = \frac{\partial\mathcal{L}}{\partial\dot{\eta}} = 0, \quad \Pi_\eta^i = \frac{\partial\mathcal{L}}{\partial v_i} = v^i. \quad (7.53)$$

One can verify that equations (7.42) follow from the Euler-Lagrange equations that are associated with  $\mathcal{L}$ , although the one that is associated with  $\eta$  will take the form:

$$\varepsilon = \frac{1}{2}\rho(v^2 - v_s^2) + \sigma\phi - \eta \operatorname{div} \mathbf{v}, \quad (7.54)$$

which is equivalent to the latter when one recalls (3.24).

Thus, for a solution, one will have that  $\mathcal{L}$  takes the form:

$$\mathcal{L} = \frac{1}{2}\rho v_s^2 - nU_h = \rho v_s^2 + \eta \operatorname{div} \mathbf{v}, \quad (7.55)$$

which differs from (4.12) by an overall sign.

The current that is associated with the arbitrariness of the global choice of initial phase is now:

$$J^0 = \rho, \quad J^i = \rho v^i = p^i, \quad (7.56)$$

in which we have again suppressed an inessential constant factor. In effect, the only thing that has changed since (2.51) is the definition of  $v^i$ .

The conserved current that is associated with the gauge invariance of  $\mathcal{L}$  is now:

$$J_A^0 = \frac{\partial\mathcal{L}}{\partial\phi} = \sigma, \quad J_A^i = \frac{\partial\mathcal{L}}{\partial A_i} = \sigma v^i, \quad (7.57)$$

which relates to  $J^\mu$  by way of:

$$J_A^\mu = \frac{\sigma}{\rho} J^\mu. \quad (7.58)$$

The individual matrices of energy-momentum-stress tensor  $T_\nu^\mu$  are:

$$T_0^0 = -\frac{1}{2}\rho(v^2 + v_s^2) - \sigma\phi, \quad (7.59)$$

$$T_j^0 = p_j, \quad (7.60)$$

$$T_0^i = -\varepsilon v^i + \dot{\eta} v^i, \quad (7.61)$$

$$T_j^i = \rho v^i v_j + \rho v^i v_j - \mathcal{L} \delta_j^i. \quad (7.62)$$

which differ from the corresponding complex expressions in (7.31)-(7.34) by an overall minus sign.

The trace of  $T_\nu^\mu$  is now:

$$T_\mu^\mu = -\rho(v^2 + v_s^2) - 4\sigma\phi + 3\varepsilon, \quad (7.63)$$

which similarly differs from the complex form (7.35) by an overall sign change.

The divergence of  $T_\nu^\mu$  splits into:

$$\partial_\mu T_0^\mu = -\sigma\dot{\phi}, \quad \partial_\mu T_i^\mu = \sigma(E_i + \frac{1}{c} B_{ij} v^j), \quad (7.64)$$

which is consistent with (7.36), up to overall sign.

One can also represent  $\mathcal{L}$  as a sum  $\mathcal{L}_0 + \mathcal{L}_A$ , in which:

$$\mathcal{L}_A = A_i j^i, \quad (7.65)$$

with:

$$j^i = -\frac{\sigma}{c} \left[ v - \frac{Q}{2mc} A^i \right]. \quad (7.66)$$

This allows one to decompose the currents  $J^\mu$  and  $J_A^\mu$  into sums, although only the spatial part actually splits:

$$J^i = J_0^i + \frac{\sigma}{c} A^i. \quad (7.67)$$

in which we have defined the current:

$$J_0^i = \frac{\sigma}{c} \tilde{v}_j = \frac{\hbar\sigma}{mc} \partial_j \theta, \quad (7.68)$$

which then amounts to the electric current (7.58) without the contribution from  $A$ .

One can also regard  $T_\nu^\mu$  as a sum  $\overset{\circ}{T}_\nu^\mu + \overset{A}{T}_\nu^\mu$ , in which  $\overset{\circ}{T}_\nu^\mu$  is the previous non-magnetic tensor, and the individual submatrices of  $\overset{A}{T}_\nu^\mu$  are:

$$\overset{A}{T}_0^0 = -\mathcal{L}_A, \quad \overset{A}{T}_j^0 = 0, \quad \overset{A}{T}_0^i = \frac{\sigma}{c} \dot{\theta} A^i, \quad \overset{A}{T}_j^i = A^i J_{0,j} - \mathcal{L}_A \delta_j^i. \quad (7.69)$$

The trace of  $T_{\nu}^{\mu}$  is then [compare (7.39)]:

$$T_{\mu}^{\mu} = A^i J_{0,i} - 4\mathcal{L}_A. \quad (7.70)$$

*d. The Madelung-Takabayasi form of the minimally-coupled stationary Schrödinger equation.* – The minimally-coupled stationary Schrödinger equation becomes:

$$\left[ -\frac{\hbar^2}{2m} \left( \frac{\partial}{\partial x^i} + \frac{iQ}{\hbar c} A_i \right)^2 + Q\phi \right] \psi = \mathcal{E} \psi, \quad (7.71)$$

and its Madelung-Takabayasi form will now be:

$$\mathcal{E} = \frac{1}{2} \rho v^2 + \sigma\phi + nU_h, \quad \text{div} \mathbf{p}_s = 0, \quad (7.72)$$

which is consistent with the non-magnetic case (6.3) when one sets  $U = \sigma\phi$ , although the magnetic potential has been absorbed into the definition of  $v$ . Once again, the symbol  $\mathcal{E} = \mathcal{E}n = 2\mathcal{E}\eta/\hbar$  will represent an energy density that is constant in time.

The Euler equation for equilibrium that one gets from this will then take the form of the previous one with an extra force term that would represent the Lorentz force:

$$\frac{\partial}{\partial x^i} (v^i p_j) = \sigma \left[ E_i + \frac{1}{c} B_{ij} v^j \right] - n \partial_i U_h. \quad (7.73)$$

The Lagrangian density takes the form:

$$\mathcal{L} = R^2 \left[ \frac{1}{2m} \bar{p}^2 + Q\phi - \mathcal{E} \right] + \frac{\hbar^2}{2m} (d_s R)^2, \quad (7.74)$$

whose density-only form is:

$$\mathcal{L} = \frac{1}{2} \rho (v^2 + v_s^2) + \sigma\phi - \mathcal{E}, \quad (7.75)$$

in which nothing is a function of time, now. One sees that this is minimally-coupled form of the non-magnetic expression (6.8).

For a solution, one can express  $\mathcal{L}$  in the same form as in the non-magnetic case (6.9):

$$\mathcal{L} = \frac{1}{2} \rho v_s^2 - nU_h = \rho v_s^2 + \eta \text{div} \mathbf{v}, \quad (7.76)$$

so the minimal coupling has not affected that fact.

The canonical forces and momenta that one derives from  $\mathcal{L}$  are:

$$f_\zeta = \frac{\partial \mathcal{L}}{\partial \zeta} = 0, \quad \Pi_\zeta^i = \frac{\partial \mathcal{L}}{\partial v_i} = p^i, \quad (7.77)$$

$$f_\eta = \frac{\partial \mathcal{L}}{\partial \eta} = \frac{2}{\hbar} \left[ \frac{1}{2} m (v^2 - v_s^2) + Q\zeta - \mathcal{E} \right], \quad \Pi_\eta^i = \frac{\partial \mathcal{L}}{\partial \pi_i} = v^i, \quad (7.78)$$

so the Euler-Lagrange equation for  $\eta$  will actually take the form:

$$\mathcal{E} = \frac{1}{2} \rho (v^2 - v_s^2) + \sigma \phi - \eta \operatorname{div} \mathbf{v} = \frac{1}{2} \rho v^2 + \sigma \phi + n U_{\hbar}. \quad (7.79)$$

The conserved current that is associated with phase symmetry is:

$$J^0 = 0, \quad J^i = p^i, \quad (7.80)$$

and the one that is associated with gauge symmetry is:

$$J_A^i = \frac{\sigma}{c} v^i. \quad (7.81)$$

The individual submatrices of the energy-momentum-stress tensor  $T_\nu^\mu$  take the form:

$$T_0^0 = -\mathcal{L}, \quad T_0^i = T_j^0 = 0, \quad T_j^i = \rho v^i v_j + \rho v^i v_j - \mathcal{L} \delta_j^i, \quad (7.82)$$

so its trace is now:

$$T_\mu^\mu = -\rho v^2 - \rho v_s^2 + 4(\mathcal{E} - \sigma \phi). \quad (7.83)$$

Its spatial divergence is:

$$\partial_j T_i^j = \sigma E_i + B_{ij} J_A^j, \quad (7.84)$$

while the temporal divergence  $\partial_\mu T_0^\mu$  vanishes, since  $\mathcal{L}$  is independent of time.

As usual, one can represent the contribution of the external electromagnetic field additively. When one goes over equations (7.65) to (7.70), one will see that the only thing that must be changed in them is to express  $T_\nu^\mu$  in real form.

**§ 8. The introduction of vorticity.** – Since the unavoidable consequence of obtaining a momentum or covelocity 1-form from the exterior derivative of a differentiable function, such as  $\theta$ , is that its exterior derivative (viz., dynamical or kinematical vorticity) must vanish, if one regards the Schrödinger equation that served as a starting point as a linear, spinless approximation to something more physically fundamental then one possible direction of generalization would be to replace the basic



flow that results in the Madelung-Takabayasi formulation with something that has non-vanishing vorticity and see how that might also change the character of the Schrödinger equation in return. Indeed, one might continue along that line of thinking and consider the chaotic (or perhaps stochastic) formation of cascading sequences of vortex-pairs, in the form of turbulence; we shall return to discuss that possibility in the next chapter.

The first possibility was considered independently by Schönberg [5] and Takabayasi [3], and revisited somewhat later by Iwo and Zofia Bialnycki-Birula, along with C. Sliwa [6]. We shall attempt to address some of the issues in this last section.

In the last section, we already saw how the coupling of an external magnetic field to the motion of a charged Madelung medium could result in non-vanishing vorticity. Hence, there is certainly an empirical physical basis for the generalization from an exact energy-momentum or covelocity 1-form to something more involved.

Since we have been regarding the covelocity 1-form  $v = (\hbar/m)d\theta$  as more definitive of continuum mechanics than the total energy-momentum 1-form  $P = mv = \hbar d\theta$ , insofar as the total values of physical observables are more appropriate to the point approximation, while densities are the fundamental observables for continua, we will start with  $v$ , instead of  $P$ . The essential fact about  $v$  is that, like  $P$ , it will be exact when one starts with the traditional form of the Schrödinger equation.

*a. The theory of Pfaff forms* <sup>(1)</sup>. The Pfaff equation is a total differential equation that takes the form  $\alpha = 0$ , where  $\alpha$  is a 1-form on an  $n$ -dimensional differentiable manifold. The solution to such an exterior differential system will be an integral submanifold of the distribution of tangent hyperplanes in space that are spanned by all tangent vectors  $\mathbf{X}$  that are annihilated by  $\alpha$ :  $\alpha(\mathbf{X}) = 0$ . Hence, one first solves the linear-algebraic problem of finding the annihilating hyperplanes of  $\alpha$  and then solves the differential system that they collectively define.

The dimension of an integral submanifold can vary from one up to  $n - 1$ ; integral submanifolds of dimension one will always exist, moreover. The maximum dimension of an integral submanifold for an exterior differential system is called its *degree of integrability*. When that degree is  $n - 1$ , one calls the exterior differential system *completely integrable*. Frobenius's theorem says that complete integrability is equivalent to the vanishing of the Frobenius 3-form  $\alpha \wedge d\alpha$ , which is equivalent to saying that a 1-form  $\eta$  exists such that:

$$d\alpha = \eta \wedge \alpha. \quad (8.1)$$

Generally, the first thing that one does with any 1-form is to put it into one of two *normal forms*:

---

<sup>(1)</sup> The theory of the Pfaff equation has a long tradition, so we shall cite only one modern reference that will suit our purposes, namely, the book by Bryant, Chern, *et al.* [7]. The author has also compiled a number of applications of that theory to physics in [8].

$$\alpha = \begin{cases} d\psi + \sum_{a=1}^{p/2} \lambda_a d\mu^a, & p \text{ even} \\ \sum_{a=1}^p \lambda_a d\mu^a, & p \text{ odd} \end{cases} \quad (8.2)$$

in which the positive integer  $p$  relates to the rank of the first  $k$ -form in the series:

$$d\wedge\alpha, \alpha \wedge d\wedge\alpha, d\wedge\alpha \wedge d\wedge\alpha, \dots$$

to vanish. Note that for any finite-dimensional manifold (of dimension  $n$ ) this sequence must truncate after a finite number of terms, since any  $k$ -form on an  $n$ -dimensional manifold must vanish when  $k$  exceeds  $n$ .

In particular, for  $n = 3$ , the only possible non-zero terms in the sequence are  $d\wedge\alpha$  and  $\alpha \wedge d\wedge\alpha$ . The corresponding possible normal forms for  $\alpha$  are  $d\psi$ ,  $\lambda d\mu$ ,  $d\psi + \lambda d\mu$ , which correspond to  $p = 0, 1, 2$ , resp., for which, the first vanishing  $k$ -forms are  $d\psi$ ,  $\lambda d\mu$ ,  $d\psi + \lambda d\mu$ , resp. The cases that will be of interest to us are the first and third ones.

The only possible dimensions for an integral submanifold of a field of tangent planes are one and two, where one-dimensional integral submanifolds are always possible. One sees that both the first and second normal forms for  $\alpha$  – namely,  $d\psi$  and  $\lambda d\mu$ , resp. – will admit two-dimensional integral submanifolds in the form of the level surfaces of the functions  $\psi$  and  $\mu$ , respectively. From Frobenius's theorem, one knows that since complete integrability is also equivalent to the vanishing of  $\alpha \wedge d\wedge\alpha$ , these will also be the only two possibilities. Hence, the third case of  $\alpha = d\psi + \lambda d\mu$  will no longer be completely integrable, but will admit only one-dimensional integral submanifolds.

*b. Vorticial motions of Madelung media.* If we now apply this analysis to the case of the covelocity 1-form then we will first see that  $v = d\zeta$  (where  $\zeta = \hbar\theta/m = S/m$ , as before) is the possibility that we started with, and which produces irrotational motion, since the kinematical vorticity is  $\Omega_k = d\wedge v$ . If one starts with the total energy-momentum 1-form  $P = dS$  then the dynamical vorticity  $\Omega_k = d\wedge P$  will also vanish, since  $P$  is related to  $v$  by a non-zero multiplicative constant.

The complete integrability of  $v = 0$  would imply that there were integral surfaces whose tangent vectors  $\mathbf{X}$  were annihilated by the 1-form  $v$ :  $v(\mathbf{X}) = 0$ . Since  $v(\mathbf{v}) = v_i v^i = v^2$  is not generally zero, one sees that the velocity vectors will not be tangent to the integral surfaces. In fact, since  $v_i = \delta_{ij} v^j$ , one thinks of the velocity vector field as being normal to the integral surfaces, and refers to the motion as a *surface-normal* motion. As we shall see, one possible set of candidates for integral curves that belong to the integral surfaces are the vortex lines.

Although it was never mentioned by the classical authors, the next possible normal form for a 1-form after exactness is the one that could relate to the energy-momentum density 1-form  $p = \rho v = \rho d\zeta$  as long as the mass density  $\rho$  is non-constant, and its gradient is not collinear with  $v$ . That is because the dynamical vorticity of such a distribution would be  $d\rho \wedge d\zeta = d\rho \wedge v$ , which would vanish iff  $d\rho$  were collinear with  $v$ . For instance, cylindrical or spherical mass distributions whose densities varied only with

distance from the axis or center would generally satisfy this constraint unless their motion were also one of radial expansion. However, if the mass density of a rod that moves in the direction of its longitudinal axis were also variable along only that same axis then it, too, would have vanishing dynamical vorticity.

The case that was considered by Schönberg, Takabayasi, and Bialnycki-Birula was the one that had been previously established in hydrodynamics by Clebsch [9], namely, the treatment of vortical motion by means of the *Clebsch variables* [10]. One then expresses the covelocity 1-form in the form:

$$v = d\zeta + \lambda d\mu. \quad (8.3)$$

Its kinematical vorticity would then be:

$$\Omega_k = d\lambda \wedge d\mu = \frac{1}{2}(\partial_i \lambda \partial_j \mu - \partial_j \lambda \partial_i \mu) dx^i \wedge dx^j, \quad (8.4)$$

while the Frobenius 3-form would be:

$$v \wedge \Omega_k = d\zeta \wedge d\lambda \wedge d\mu. \quad (8.5)$$

If this is non-vanishing – i.e., the motion is not completely integrable – then the three 1-forms  $d\zeta$ ,  $d\lambda$ ,  $d\mu$  will be linearly-independent.

Since we are only concerned with vector spaces (which are orientable), one can associate the 2-form  $\Omega_k$  with a *vorticity vector field*  $\omega_k$  by way of the (inverse of the) Poincaré isomorphism #:  $\Lambda^1 \mathbb{R}^3 \rightarrow \Lambda_2 \mathbb{R}^3$ ,  $\mathbf{v} \mapsto i_{\mathbf{v}} V$  that a choice of spatial volume element  $V$  will define:

$$\omega_k = \#^{-1} \Omega_k = \omega_k^i \partial_i, \quad (8.6)$$

whose components will then be:

$$\omega_k^i = \frac{1}{2} \varepsilon^{ijk} \Omega_{jk} = \frac{1}{2} \varepsilon^{ijk} \partial_j \lambda \partial_k \mu = \frac{1}{2} \nabla \lambda \times \nabla \mu, \quad (8.7)$$

in which we have temporarily introduced the notation of vector calculus in the last expression.

One can then revisit (8.5) with this new concept and see that

$$v \wedge \Omega_k = v \wedge \# \omega_k = v(\omega_k) V. \quad (8.8)$$

Hence, the complete integrability of the exterior differential system  $v = 0$  is equivalent to the vorticity vector field satisfying that exterior algebraic equation, which also means that the vortex lines have to be orthogonal to the path lines or streamlines of the velocity vector field  $\mathbf{v}$  since  $v(\omega_k) = \langle \mathbf{v}, \omega_k \rangle$ .

## References (\*)

1. Some standard textbooks on quantum mechanics are:
  - P. A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford University Press, Oxford, 1982 (first edition, 1930).
  - A. Messiah, *Quantum Mechanics*, 2 vols., Wiley, NY, 1958.
  - L. I. Schiff, *Quantum Mechanics*, McGraw-Hill, 1968.
  - E. Merzbacher, *Quantum Mechanics*, Wiley, N. Y., 1970.
  - C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Quantum Mechanics*, 3 vols., Wiley, 1977.
- 2\*. E. Madelung, “Quantentheorie in hydrodynamischer Form,” *Zeit. Phys.* **40** (1927), 322-326.
3. T. Takabayasi:
  - a. “On the Formulation of Quantum Mechanics Associated with Classical Pictures,” *Prog. Theor. Phys.*, **8** (1952), 143-182.
  - b. “Remarks on the Formulation of Quantum Mechanics Associated with Classical Pictures,” *Prog. Theor. Phys.*, **9** (1953), 187-222.
4. O. Halpern, “A Proposed Re-Interpretation of Quantum Mechanics,” *Phys. Rev.* **87** (1952), 389; response by D. Bohm, “Reply to a Criticism of a Causal Re-Interpretation of the Quantum Theory,” *ibid.*, 389-390.
5. M. Schönberg:
  - a. “A non-linear Generalization of the Schrödinger and Dirac Equations,” *Nuov. Cim.* **11** (1954), 674-682.
  - b. “On the Hydrodynamical Model of the Quantum Mechanics,” *ibid.* **12** (1954), 103-133.
  - c. “Simple Solution of the Generalized Schrödinger Equations,” *ibid.* **12** (1954), 300-303.
  - d. “Vortex Motions of the Madelung Fluid,” *ibid.* (1955), 543-580.
6. I. Bialnycki-Birula, Z. Bialnycki-Birula, and C. Sliwa, “Motion of vortex lines in quantum mechanics,” *Phys. Rev. A* **61** 032110 (2000).
7. R. L. Bryant, S.-S. Chern, *et al.*, *Exterior Differential Systems*, Springer, Berlin, 1991.
8. D. H. Delphenich, “The role of integrability in a large class of physical systems” arXiv:1210.4976.
- 9\*. A. Clebsch, “Ueber die Integration der hydrodynamischen Gleichungen,” *J. f. reine u. angew. Math.* **56** (1859), 1-10.
10. Sir H. Lamb, *Hydrodynamics*, 6<sup>th</sup> ed., Cambridge Univ. Press, Cambridge, 1932; First ed. 1879, reprinted by Dover, Mineola, NY, 1945.

---

(\*) References marked with an asterisk are available in English translation at the author’s website: [neo-classical-physics.info](http://neo-classical-physics.info).

## CHAPTER V

# THE NATURE OF THE QUANTUM TERMS

What one hopes, above all, is that reformulating the wave equations of quantum mechanics in terms that pertain to continuum mechanics will lead to a better intuitive picture of the structure of matter, its interactions, and the time evolution of its states by pointing to analogies between the somewhat-speculative, sub-microscopic world of quantum phenomena and the more familiar world of macroscopic matter. Indeed, one typically expects that natural law has a sufficiently general, universal character that the same fundamental processes and relationships would be going on at every scale of natural phenomena. For instance, an early model of the atom was based on the assumption that electrons orbit the nucleus like planets, although, of course, that model proved to be highly incomplete.

So far, our first hint regarding the key distinction between classical continuum mechanics and the kind that includes “quantum corrections” is the appearance of either a quantum potential, its corresponding quantum force, or its quantum stress tensor. Hence, we suspect that if it is possible to make sense of these notions in more familiar terms then we shall have made some progress in terms of paving a smoother road from classical to quantum physics. Since quantum stress seems to have the most facets to it, we shall use it as our primary focus.

Basically, the gist of this chapter is that when one goes from point-like matter to extended matter, one can consider the way that the geometry of the instantaneous spaces of dynamical states differs from that of the instantaneous kinematical states as a result of the conformal transformation of the metric on the latter space that scalar multiplication by the mass density function represents. In particular, when that function is not constant in space, the shape of its level surfaces plays a role – i.e., its higher derivatives. However, one finds that the real issue is the deformation of tangent frames, not the deformation of the space that they live on, so one is dealing with what the author calls “frame strain,” rather than the metric strain that is customarily addressed by continuum mechanics.

Another aspect of the difference between quantum mechanics and continuum mechanics that bears upon the physics of the models is the presumed incompleteness in the original quantum wave equations. The enlargement of scope that going from quantum mechanics to quantum field theory represents seems to be rooted in the interaction of fields with the quantum vacuum state, so having some way of modeling that state is essential. Hence, we shall also consider the way that various researchers introduced turbulence into the motion of the Madelung medium as a possible source of “quantum fluctuations.”

**§ 1. The geometrical origins of the quantum stress.** Perhaps one of the reasons that the early researchers in the continuum-mechanical models for quantum mechanics imagined that the Madelung medium was a fluid was that at first it did not seem as though there were any obvious way of associating the quantum stress tensor with a

corresponding strain tensor. Since that situation is typical of fluid media, which generally admit no strains (although viscous stresses do get coupled to the rate of strain), it was then assumed that the medium in question was some sort of exotic fluid.

However, more recently, the author wrote a series of articles [1-3] on the role of geometry in Madelung media. In the first one [1], he proposed that the quantum potential could be related to the scalar curvature of a Levi-Civita connection that one associates with a metric that one obtains from deforming the spatial metric by means of a conformal factor that amounted to the mass density function. In the second one [2], he showed that, along the same lines, one could define a metric strain tensor that couples to the quantum stress tensor and is also derived from that same deformation of the metric. In that paper, it was suggested that the coupling of metric strain to quantum stress was more direct if one used the older usage of the words “curvature” and “torsion” that was more consistent with their usage in the context of the Frenet-Serret equations than with their usage in the more modern Riemann-Cartan context. In the third paper [3], the author expanded upon the latter idea in a more general way and showed how the geometry of “teleparallelism” represented a higher-dimensional version of the Frenet-Serret approach to the motion of an orthonormal frame along a curve when one extends to frame fields on surfaces and higher-dimensional objects.

Indeed, a similar sort of coupling of strain to stress was described by Kelvin and Tait [4] in the context of the stresses that are produced by the bending and twisting of beams. Basically, one must realize that one is coupling stress to the *connection* on the deformed space, not to its curvature and torsion in the Riemann-Cartan sense, which is more closely related to the integrability of parallel translation.

Hence, we shall start from that viewpoint in order to suggest how the basic principle at work is the idea that there is a difference between the geometry of the space of kinematical states and that of the space of dynamical states. At the most elementary level, if the kinematical state is described by a velocity vector  $\mathbf{v} = v^i \partial_i$  then the speed  $v$  that is associated with it will be the square root of:

$$v^2 = \langle \mathbf{v}, \mathbf{v} \rangle = \delta_{ij} v^i v^j. \quad (1.1)$$

The Euclidian scalar product that is defined on the tangent vectors will then be referred to as the *kinematical metric*. One can associate it with a corresponding scalar product on the cotangent vectors that gives the same value of  $v^2$  to the covelocity covector:

$$v^2 = \langle v, v \rangle = \delta^{ij} v_i v_j \quad (v_i = \delta_{ij} v^j). \quad (1.2)$$

This is still a kinematical metric, but when one goes from covelocity to (linear) momentum, in its simplest form, one sees that there is a significant difference between point mechanics, for which a potentially-time-varying mass scalar  $m(t)$  is associated with only a single curve  $x(t)$  in the space of motion, and the motion of an extended material object, for which a mass density  $\rho(t, x)$  that is associated with congruence of curves  $x(t, u)$  in space; here, the symbol  $u$  generically refers to the parameters that define the shape of the object in some reference state, such as its initial state. When one defines the momentum density 1-form  $p = p_i(t, x) dx^i$  in the usual way:

$$p = \rho v \quad (p_i = \rho \delta_{ij} v^j), \quad (1.3)$$

one will see that one can define a scalar product on the support of  $\rho$ , namely:

$$\rho v^2 = \langle v, v \rangle_\rho = \rho \delta^{ij} v_i v_j, \quad (1.4)$$

that associates the covelocity 1-form with (twice) the kinetic energy density of motion. The difference between this scenario and that of point mechanics is that one can differentiate  $\rho(t, x)$  with respect to space, as well as time, while  $m(t)$  has only a time derivative.

The resulting metric on tangent vectors:

$$g_{ij}(t, x) = \rho(t, x) \delta_{ij} \quad (1.5)$$

takes the form of a conformal transformation of the original Euclidian metric, and its corresponding metric on cotangent vectors will have components:

$$g^{ij} = \frac{1}{\rho} \delta^{ij}. \quad (1.6)$$

Note that both the metrics  $g_{ij}$  and  $g^{ij}$  can be defined only where the density  $\rho$  is non-zero. Hence, the “dynamical” metric is defined only where the matter itself is defined, but not in the complementary space. It then pertains to the geometry of the “world-tube” that is swept out in space-time by the motion of the object.

A conformal transformation of a metric, which takes the general form  $g \mapsto \bar{g} = \Omega^2 g$ , where  $\Omega$  is a function of position on the relevant manifold, can also be defined by a local homothety of the tangent spaces:

$$\bar{g}(\mathbf{u}, \mathbf{v}) = g(\Omega \mathbf{u}, \Omega \mathbf{v}). \quad (1.7)$$

Thus, the function  $\Omega$  also plays a fundamental role geometrically.

In the case of dynamics,  $\Omega$  will be equal to  $\sqrt{\rho}$ , which we know is also proportional to the function  $R$  that is at the basis for the Madelung medium, along with  $\theta$ , i.e.:

$$\Omega^2 = \rho = m R^2. \quad (1.8)$$

We first denote the open set of space-time points at which  $\rho \neq 0$  by  $\Sigma_\rho$  in order to specify the manifold on which the *dynamical metric*  $g = g_{ij} dx^i dx^j$  is defined. Although it is tempting to then go to the *support* of  $\rho$ , which is the closure of  $\Sigma_\rho$ , nonetheless, since there might be limit points of  $\Sigma_\rho$  at which  $\rho$  is zero, if one expects to divide by  $\rho$  then one cannot use those points, anyway. Hence, we can then treat the Riemannian manifold  $(\Sigma_\rho, g)$  as the *dynamical space* for the motion of the extended object under scrutiny, and the three-dimensional slices  $\Sigma_\rho(t)$  for constant values of the time dimension-parameter will be called *instantaneous dynamical spaces*.

As long as  $\rho$  is not spatially constant, the spatial derivatives of  $g$  will generally be non-zero and equal to:

$$g_{ij,k} = \frac{\partial g_{ij}}{\partial x^k} = \frac{\partial \rho}{\partial x^k} \delta_{ij}, \quad g_{ij,k,l} = \frac{\partial^2 g_{ij}}{\partial x^k \partial x^l} = \frac{\partial \rho}{\partial x^k \partial x^l} \delta_{ij}. \quad (1.9)$$

Hence, one sees that only the spatial derivatives of  $\rho$  contribute in an essential way.

**2. The Riemannian geometry of the instantaneous dynamical spaces.** – If one wishes to take the Riemannian route to differential geometry, as in [1], then one can use (1.6) and (1.9) to obtain the components of the Levi-Civita connection that is associated with the metric on the instantaneous dynamical space  $\Sigma_\rho(t)$ :

$$\Gamma_{jk}^i = \frac{1}{2} g^{il} (g_{lj,k} + g_{lk,j} - g_{jk,l}) = \frac{1}{2\rho} (\rho_{,k} \delta_j^i + \rho_{,j} \delta_k^i - \rho_{,l} \delta^{il} \delta_{jk}). \quad (2.1)$$

The parallel translation of a vector  $\mathbf{u}(s) = u^i(s) \partial_i$  along a curve  $x(s)$  in  $\Sigma_\rho(t)$  whose velocity vector field is  $\mathbf{v}(s)$  will then imply that:

$$0 = \nabla_{\mathbf{v}} \mathbf{u} = \left( \frac{du^i}{ds} + \Gamma_{jk}^i v^j u^k \right) \frac{\partial}{\partial x^i}; \quad (2.2)$$

i.e., that:

$$0 = \frac{du^i}{ds} + \Gamma_{jk}^i v^j u^k = \frac{du^i}{ds} + \frac{1}{2\rho} [(u^j \rho_{,j}) v^i + (v^j \rho_{,j}) u^i - (v_k u^k) \delta^{ij} \rho_{,j}], \quad (2.3)$$

which can be given the component-free form:

$$0 = \frac{d\mathbf{u}}{ds} + \frac{1}{2\rho} [(\mathbf{v}\rho)\mathbf{u} + (\mathbf{u}\rho)\mathbf{v} - \langle \mathbf{u}, \mathbf{v} \rangle \text{grad}\rho]. \quad (2.4)$$

In particular when  $\mathbf{u} = \mathbf{v}$ , one will get the geodesic equations:

$$0 = \frac{dv^i}{ds} + \Gamma_{jk}^i v^j v^k = \frac{dv^i}{ds} + \frac{1}{\rho} (v^j \rho_{,j}) v^i - \frac{1}{2\rho} v^2 \rho_{,j} \delta^{ij}, \quad (2.5)$$

which can be given the component-free form:

$$0 = \frac{d\mathbf{v}}{ds} + \frac{1}{\rho} (\mathbf{v}\rho)\mathbf{v} - \frac{1}{2\rho} v^2 \text{grad}\rho. \quad (2.6)$$



One notices that  $d\rho$  always seems to appear in conjunction with  $1/\rho$ , which suggests that  $d(\ln \rho)$  might be a somewhat more fundamental expression. If we express  $\rho = \Omega^2$  in the form  $e^{2\lambda}$  then we can set:

$$\lambda = \frac{1}{2} \ln \rho, \quad d\lambda = \frac{1}{2\rho} d\rho, \quad (2.7)$$

and the components of the Levi-Civita connection will take the form:

$$\Gamma_{jk}^i = \lambda_{,k} \delta_j^i + \lambda_{,j} \delta_k^i - \lambda_{,l} \delta^l \delta_{jk}. \quad (2.8)$$

This will then make the equation of parallel translation take the form:

$$0 = \frac{d\mathbf{u}}{ds} + (\mathbf{u}\lambda) \mathbf{v} + (\mathbf{v}\lambda) \mathbf{u} - \langle \mathbf{v}, \mathbf{u} \rangle \text{grad } \lambda, \quad (2.9)$$

while the geodesic equation will now be:

$$0 = \frac{d\mathbf{v}}{ds} + 2(\mathbf{v}\lambda) \mathbf{v} - v^2 \text{grad } \lambda. \quad (2.10)$$

In order to compute the Riemann curvature of the connection that we have defined, it is simplest to first form the connection 1-form:

$$\Gamma_j^i = \Gamma_{jk}^i dx^k = d\lambda \delta_j^i + \lambda_{,j} dx^i - \lambda^i dx_j. \quad (2.11)$$

The curvature 2-form that this defines can be obtained from the Cartan structure equation for curvature:

$$\mathfrak{R}_j^i = \frac{1}{2} R_{jkl}^i dx^k \wedge dx^l = d \wedge \Gamma_j^i + \Gamma_k^i \wedge \Gamma_j^k. \quad (2.12)$$

That will give:

$$\begin{aligned} R_{jkl}^i &= (\lambda_{,j,k} - \lambda_{,j} \lambda_{,k}) \delta_l^i - (\lambda_{,j,l} - \lambda_{,j} \lambda_{,l}) \delta_k^i + (\lambda_{,l}^i - \lambda^i \lambda_{,l}) \delta_{jk} - (\lambda_{,k}^i - \lambda^i \lambda_{,k}) \delta_{jl} \\ &\quad + (\lambda_{,m} \lambda^{,m}) (\delta_l^i \delta_{jk} - \delta_k^i \delta_{jl}). \end{aligned} \quad (2.13)$$

One can then find the Ricci curvature 1-forms from this by contraction:

$$\mathfrak{R}_i = R_{ij} dx^j = R_{ikj}^k dx^j, \quad (2.14)$$

in which:

$$R_{ij} = -(\lambda_{,i,j} - \lambda_{,i} \lambda_{,j}) - [\Delta\lambda + (d\lambda)^2] \delta_{ij}. \quad (2.15)$$

If we raise a lower index then we will get the matrix of the Ricci transformation:

$$R_j^i = g^{ik} R_{kj} = e^{-2\lambda} \delta^{ik} R_{kj} = -e^{-2\lambda} \{ \lambda_{,j}^i - \lambda^i \lambda_{,j} + [\Delta\lambda + (d\lambda)^2] \delta_j^i \}. \quad (2.16)$$

If we contract this then we will get the scalar curvature <sup>(1)</sup>:

$$\mathfrak{R} = R_i^i = -4 e^{-2\lambda} [\Delta\lambda + \frac{1}{2}(d\lambda)^2]. \quad (2.17)$$

If we identify  $\lambda$  with  $\ln\sqrt{m} + \ln R$ , where  $R$  is the modulus of the Schrödinger wave function, as before, then:

$$\Delta\lambda = \frac{\Delta R}{R} - (d\lambda)^2,$$

or:

$$\frac{\Delta R}{R} = \Delta\lambda + (d\lambda)^2 = -\frac{1}{4} e^{2\lambda} \mathfrak{R} + \frac{1}{2} (d\lambda)^2. \quad (2.18)$$

Hence, the quantum potential can be expressed in the form:

$$U_h = -\frac{\hbar^2}{2m} \frac{\Delta R}{R} = \frac{\hbar^2}{8m} [\rho \mathfrak{R} - 2(d\lambda)^2]. \quad (2.19)$$

As we shall see, the second term on the right-hand side will no longer appear when we go to the relativistic form of things. Nonetheless, we already see that we are close to identifying the quantum potential with a scalar multiple of the scalar curvature of the deformed metric on the instantaneous dynamical space  $\Sigma_\rho(t)$ .

**3. Clifford-Weitzenböck geometry** <sup>(2)</sup>. – The big problem with differential manifolds, as compared to affine spaces, in the eyes of kinematics is that in the absence of a globally-defined action of the translation group that would allow one to not only move points of space to other points, but (by differentiation) also tangent and cotangent vectors in the tangent and cotangent spaces, as well, one has no way of defining the acceleration of a curve. That is because one usually starts with a finite difference of two velocity vectors that are tangent to finitely-separated points, and unless one has some way of identifying the vectors in the two tangent spaces, such a difference would be ill-defined.

The solution, of course, is to define some form of parallelism, if only between tangent objects. Previously, we discussed Riemannian geometry, in which one defines a connection on either tangent vectors or tangent frames that allows one to define the parallel translation of such objects along curves, and in so doing, one can get isomorphisms of the tangent spaces along such a curve. However, parallel translation is generally curve-dependent then, so one cannot expect to arrive at such a thing as a

---

<sup>(1)</sup> We will use the somewhat non-standard notation of  $\mathfrak{R}$  for the scalar curvature, instead, of  $R$ , which we are already using for the modulus of the quantum wave function.

<sup>(2)</sup> For more details on this subsection and the next, see the author's paper [3].

globally-defined isomorphism of all tangent spaces that would allow one to define “distant parallelism” or *teleparallelism* – viz., the parallelism of finitely-separated tangent vectors or frames <sup>(1)</sup>.

In order to define a global isomorphism of any two tangent spaces on a manifold, it is necessary and sufficient that one should be able to define a global frame field on it. One then calls the manifold *parallelizable*.

Frame fields always exist locally about any point of a differentiable manifold. For instance, any coordinate chart  $(U, x^i)$  defines a *natural frame field*  $\{\partial_i, i = 1, \dots, n\}$  in the form of the directional derivatives  $\partial_i = \partial / \partial x^i$  with respect to each coordinate, as well as a *reciprocal coframe field*  $\{dx^i, i = 1, \dots, n\}$  that is defined by the coordinate differentials. Thus, they are related by the fact that:

$$dx^i(\partial_j) = \delta_j^i. \quad (3.1)$$

The basic property of natural frame fields and coframe fields is that they are *holonomic*, which is expressed by the dual conditions:

$$[\partial_i, \partial_j] = 0, \quad d \lrcorner dx^i = 0. \quad (3.2)$$

However, extending such a local construction to a global one is obstructed by both topology and geometry.

Topologically, it is necessary for all of the Stiefel-Whitney classes of the bundle of linear frames to vanish; indeed, Stiefel was explicitly addressing the problem of parallelizability when he first defined such classes [6]. One finds that even such “homogeneous” spaces as spheres do not generally admit global frame fields. Indeed, the only ones that do have dimensions 0, 1, 3, and 7, which are also make them the underlying manifolds of certain Lie groups. In fact, not only are all Lie groups parallelizable, but one can even think of parallelizable manifolds as, in a sense, “almost Lie groups;” i.e., Lie groups without the multiplication.

Geometrically, a necessary condition for any linear connection to admit a parallel vector field (even locally) is that its curvature must vanish. Indeed, one can even characterize the Riemann-Cartan usage of the term “curvature” as something that obstructs the integrability of parallel-translation, in the sense of either the path-independence of that process or the existence of solutions to the system of equation for a parallel vector field  $\mathbf{u} = u^i \partial_i$ :

$$0 = \nabla u^i = du^i + \Gamma^i_j u^j. \quad (3.3)$$

Although the traditional picture of local constructions in the elementary theory of differentiable manifolds usually gives the impression that the open subsets upon which “local” constructions are defined are usually “sufficiently small” neighborhoods of points, actually, when it comes to the extension of local frame fields to global ones (or not), one often finds that the opposite problem of finding a maximal open subset upon

---

<sup>(1)</sup> The author has compiled a collection [5] of English translations of selected papers on the early attempts by Einstein, Mayer, and others to use the geometry of parallelizable manifolds as the basis for the unification of Einstein’s theory of gravitation and Maxwell’s theory of electromagnetism, including a translation of Stiefel’s thesis on the topological obstructions to parallelizability.

which the definition is still valid can have a solution that amounts to the complement of a finite point set. For instance, a two-dimensional sphere can admit a frame field that is defined everywhere except for a single point. In fact, Stiefel was defining his  $\mathbb{Z}_2$  *homology* classes by first triangulating such sets of singular points and showing that the resulting complex defined a  $\mathbb{Z}_2$ -cycle. Thus, one should not always think of local frame fields as being defined on sufficiently small open subsets, as opposed to global fields with singularities. (The author has been referring to this situation as *singular teleparallelism*.)

Since parallelizable manifolds are “a heartbeat away from Lie groups,” we shall first examine how one defines a global frame field on a Lie group  $G$ . Basically, one chooses any frame  $\{\mathbf{e}_i, i = 1, \dots, n\}$  in the tangent space  $T_e G$  at the identity  $e \in G$  (which is usually identified with the Lie algebra  $\mathfrak{g}$  of  $G$ ) and either left-translates or right-translates it to each other point  $g \in G$  to obtain a frame  $\mathbf{e}_i(g)$  in  $T_g G$ . More precisely, one first defines the Lie group isomorphisms of left and right translation by  $g$ :

$$L_g : G \rightarrow G, g' \mapsto gg', \quad R_g : G \rightarrow G, g' \mapsto g'g,$$

and differentiates them both at the identity element  $e$ . Since they are both diffeomorphisms, their differential maps will be linear isomorphisms from  $T_e G$  to  $T_g G$ . When one repeats this construction for each element  $g$ , one will get two global frame fields  $\mathbf{e}_i(g)$  on  $G$ , as well as two isomorphisms of  $T_e G$  with every other tangent space. As long as  $G$  is non-Abelian, these two frame fields and isomorphisms will generally be distinct. We shall then distinguish the two types of parallelism that come from left and right translation as being *left* and *right* parallelism, as well.

There is nothing special about any particular choice of frame at the identity in the eyes of parallelism. Let us define two vectors  $\mathbf{v}$  and  $\mathbf{v}'$  in the tangent spaces at two distinct elements  $g$  and  $g'$ , resp., in  $G$  to be *parallel* iff their components with respect to the respective frames  $\mathbf{e}_i(g)$  and  $\mathbf{e}_i(g')$  are equal; i.e.:

$$\mathbf{v} = v^i \mathbf{e}_i(g), \quad \mathbf{v}' = v^i \mathbf{e}_i(g'). \quad (3.4)$$

If one chooses another frame  $\bar{\mathbf{e}}_i(e) = \mathbf{e}_j(e)h_i^j$  at the identity then one can also extend  $\bar{\mathbf{e}}_i(e)$  to every other  $g \in G$  by left and right translation and one finds that:

$$\bar{\mathbf{e}}_i(g) = \mathbf{e}_j(g)h_i^j \quad (3.5)$$

for every  $g \in G$ . Since the (invertible) matrix  $h_i^j$  will be the same for every  $g$ , the components of the two vectors will transform by the same matrix. As a result, two vectors that are tangent to two elements of  $G$  will be parallel with respect to the frame field  $\mathbf{e}_i(g)$  iff they are parallel with respect to  $\bar{\mathbf{e}}_i(g)$ . Thus, one can say that a certain form of parallelism is associated with an equivalence class of global frame fields that are related by the same matrix at each point.

In their paper on the geometry of simple Lie groups [7], Élie Cartan and Jan Schouten pointed out that the two aforementioned types of parallelism were actually distant relatives to the two types of parallelism that William Kingdon Clifford had defined previously [8] in the context of projective geometry. Furthermore, Roland Weitzenböck [9] had treated similar issues in his book on the theory of invariants <sup>(1)</sup>. Thus, we shall refer to the geometry of parallelizable manifolds as *Clifford-Weitzenböck geometry*.

The two types of frame fields on a Lie group can also be called *left-invariant* and *right-invariant*, respectively, as can the vector fields that comprise them. Similarly, their reciprocal coframe fields  $\{\theta^i, i = 1, \dots, n\}$  [so  $\theta^i(\mathbf{e}_j) = \delta_j^i$ ] will also be left-invariant or right-invariant, respectively.

A fundamental property of left or right-invariant frame fields on Lie groups is the fact that one will have the dual conditions:

$$[\mathbf{e}_i, \mathbf{e}_j] = c_{ij}^k \mathbf{e}_k, \quad d\theta^i = -\frac{1}{2} c_{jk}^i \theta^j \wedge \theta^k, \quad (3.6)$$

in which the  $c_{ij}^k$  are the *structure constants* of  $G$  relative to that global frame field. The fact that they are constants is due to the fact that the frame field is assumed to be invariant under left or right translation, since otherwise they would be more general functions on  $G$ .

The second set of equations in (3.6) is referred to as the *Maurer-Cartan equations*. One can characterize them as an integrability condition for the coframe field  $\theta^i$  to be left- (or right-) invariant.

One can think of the invariant coframe field  $\theta^i$  as also being a 1-form on  $G$  with values in its Lie algebra  $\mathfrak{g}$ , since if  $\mathbf{X} = X^i \mathbf{e}_i(g)$  is a tangent vector at  $g$  then  $\theta^i(\mathbf{X}) = X^i$  can be associated with the tangent vector to the identity  $X^i \mathbf{e}_i(e) = \theta^i(\mathbf{X}) \mathbf{e}_i(e)$ . The 1-form  $\theta$  on  $G$  with values in  $\mathfrak{g}$  that this defines will then be called the *Maurer-Cartan 1-form* on  $G$  (mod  $\mathbf{e}_i$ ). If  $G$  is a matrix group (i.e., its elements can be represented by invertible matrices and its multiplication takes the form of matrix multiplication), so the coordinate functions for the element  $g$  are  $g_b^a(g)$ , then Maurer-Cartan 1-form for the left-invariant and right-invariant frame fields will take the forms:

$$\omega_b^a = \tilde{g}_c^a dg_b^c, \quad \omega_b^a = dg_c^a \tilde{g}_b^c, \quad (3.7)$$

resp., in which the tilde denotes the matrix inverse.

If one takes the exterior derivative of the Maurer-Cartan equations then one will get the Bianchi identity:

$$0 = c_{il}^m c_{jk}^l + c_{jl}^m c_{ki}^l + c_{kl}^m c_{ij}^l, \quad (3.8)$$

which is dual to the Jacobi identity:

---

<sup>(1)</sup> Amusingly, Roland Weitzenböck was apparently such a Francophobe that he could not resist disguising an acrostic in the opening sentences of the preface to that book. The first letters of the first words of the first 21 sentences actually spell out “Nieder mit den Französern!” (Down with the French!). That also accounts for his activities in the 1930’s during the rise of Third Reich.

$$0 = [\mathbf{e}_i, [\mathbf{e}_j, \mathbf{e}_k]] + [\mathbf{e}_j, [\mathbf{e}_k, \mathbf{e}_i]] + [\mathbf{e}_k, [\mathbf{e}_i, \mathbf{e}_j]], \quad (3.9)$$

which is true for any Lie algebra and expresses the fact that a general Lie algebra will not be associative.

The transition from a Lie group  $G$  to more general parallelizable manifold  $M$  is to replace the concept of left or right invariant frame field  $\mathbf{e}_i(g)$  (which would no longer be meaningful in the absence of group multiplication) with a general frame field  $\{\mathbf{e}_i(x), i = 1, \dots, n\}$ . Any other frame field  $\bar{\mathbf{e}}_i(x)$  on  $M$  will be related to  $\mathbf{e}_i(x)$  by way of:

$$\bar{\mathbf{e}}_i(x) = \mathbf{e}_j(x) h_i^j(x), \quad (3.10)$$

in which  $h_i^j(x)$  is the *transition function* that takes one from one frame field to the other one. It is then a function  $h : M \rightarrow GL(n)$ ,  $x \mapsto h_i^j(x)$  that associates an invertible  $n \times n$  matrix with each point of  $M$ .

Once again, one can define two tangent vectors  $\mathbf{v}$  and  $\mathbf{v}'$  at  $x$  and  $x'$ , resp. to be parallel with respect to  $\mathbf{e}_i$  iff their components with respect to  $\mathbf{e}_i(x)$  and  $\mathbf{e}_i(x')$ , resp., are equal. This situation is the same as it was for Lie groups, as well as the fact that any two distinct tangent vectors  $\mathbf{v}$  and  $\mathbf{v}'$  will be parallel with respect to two frame fields  $\mathbf{e}_i$  and  $\bar{\mathbf{e}}_i$  iff the transition function that takes one to the other is a constant function. Thus, the parallelism that a given frame field defines is common to a whole equivalence class of other frame fields.

A vector field  $\mathbf{v}(x) = v^i(x) \mathbf{e}_i(x)$  is *parallel with respect to  $\mathbf{e}_i$*  (and all other frame fields that differ by a constant transition function) iff its components  $v^i(x)$  are all constant functions. Hence:

$$dv^i = 0. \quad (3.11)$$

Thus, any vector or frame at a single point of  $M$  can be extended to a global parallel vector field or frame field by means of  $\mathbf{e}_i$ .

For the general frame field  $\mathbf{e}_i$  on  $M$ , one will now have the dual equations:

$$[\mathbf{e}_i(x), \mathbf{e}_j(x)] = c_{ij}^k(x) \mathbf{e}_k(x), \quad d \wedge \theta^i = -\frac{1}{2} c_{jk}^i(x) \theta^j \wedge \theta^k, \quad (3.12)$$

in which the  $c_{ij}^k(x)$  are now *structure functions* that do not have to be constant. Indeed, the existence of a global frame field for which they are constant would define a local diffeomorphism to a Lie group manifold whose Lie algebra has those structure constants<sup>(1)</sup>. Upon comparison with (3.6), one sees that the second set of equations is basically a generalization of the Maurer-Cartan equations for a Lie group.

As far as the Jacobi-Bianchi identities are concerned, the Jacobi identity will have the same form as in (3.9), while the Bianchi identity (3.8) will pick up another term that comes from the fact that  $dc_{ij}^k$  does not have to vanish, namely:

---

<sup>(1)</sup> This statement is cited without proof in Singer and Sternberg [10].

$$0 = \mathbf{e}_i c_{jk}^m + c_{il}^m c_{jk}^l + c_{jl}^m c_{ki}^l + c_{kl}^m c_{ij}^l. \quad (3.13)$$

Since a choice of frame field  $\mathbf{e}_i$  defines a notion of parallelism, one naturally wonders if it also defines a linear connection that implies the same parallelism, if only along curves. Indeed, if one goes back to the definition of the parallelism of a vector field  $\mathbf{v}$  (mod  $\mathbf{e}_i$ ), and the condition (3.11) then one will see that  $\mathbf{v}$  is parallel (mod  $\mathbf{e}_i$ ) iff  $dv^i = 0$ . We then define the covariant differential of  $\mathbf{v}$  to be:

$$\nabla \mathbf{v} = dv^i \otimes \mathbf{e}_i, \quad (3.14)$$

and since one usually expects that:

$$\nabla \mathbf{v} = (dv^i + \omega_j^i v^j) \otimes \mathbf{e}_i \quad (3.15)$$

for some connection 1-form  $\omega_i^j$  (which does not transform linearly from one frame to another), we see that we must have  $\omega_i^j = 0$  relative to this frame field.

If  $\mathbf{e}_i = h_i^j \partial_j$  then the transformation of  $\omega_i^j$  to the natural frame will take the form:

$$\bar{\omega}_i^j = \tilde{h}_k^j \omega_i^k h_i^l + \tilde{h}_k^j dh_i^k = \tilde{h}_k^j dh_i^k. \quad (3.16)$$

Thus  $\bar{\omega}_i^j$  is a 1-form with values in  $\mathfrak{gl}(n)$ .

If one compares (3.16) to (3.7) then one will see that this 1-form  $\bar{\omega}_i^j$  is the pull-back of the Maurer-Cartan form on  $GL(n)$  by the function  $h : M \rightarrow GL(n)$ . Hence, one sees that the connection that makes  $\mathbf{v}$  parallel has much in common with the Maurer-Cartan geometry of  $GL(n)$ .

If  $\bar{v}^i = \tilde{h}_j^i v^j$  are the components of  $\mathbf{v}$  relative to the natural frame then one can rewrite (3.14) in the form:

$$\nabla \mathbf{v} = \nabla \bar{v}^i \otimes \partial_i, \quad \nabla \bar{v}^i = d\bar{v}^i + \bar{\omega}_j^i \bar{v}^j. \quad (3.17)$$

Hence, the condition for the parallelism of  $\mathbf{v}$  can be expressed in the familiar form:

$$d\bar{v}^i + \bar{\omega}_j^i \bar{v}^j = 0 \quad (3.18)$$

in the natural frame, while it is simply  $dv^i = 0$  in the anholonomic frame  $\mathbf{e}_i$ .

Similarly, the differential equations for a geodesic  $x(s)$  (whose velocity vector field must be parallel) take the two forms:

$$\frac{dv^i}{ds} = 0, \quad \frac{d\bar{v}^i}{ds} + \bar{\omega}_j^i(\mathbf{v}) \bar{v}^j = 0 \quad (3.19)$$

in the anholonomic and natural frames, respectively.

If we express  $\bar{\omega}_j^i$  in the form:

$$\bar{\omega}_j^i = \bar{\omega}_{jk}^i dx^k \quad (3.20)$$

then the last system in (3.19) will take the form:

$$\frac{d\bar{v}^i}{ds} + \bar{\omega}_{jk}^i \bar{v}^j \bar{v}^k = 0. \quad (3.21)$$

One can obtain the torsion and curvature 2-forms for the teleparallelism connection from the Cartan structure equations:

$$\Theta^i = d \wedge \theta^i + \omega_j^i \wedge \theta^j, \quad \Omega_j^i = d \wedge \omega_j^i + \omega_k^i \wedge \omega_j^k, \quad (3.22)$$

in which we are using a generic coframe field and its associated connection 1-form.

Relative to the coframe field  $\theta^i$  for which  $\omega_j^i = 0$ , these become simply:

$$\Theta^i = d \wedge \theta^i, \quad \Omega_j^i = 0. \quad (3.23)$$

Relative to the natural frame field  $dx^i$  for which  $\omega_j^i = \bar{\omega}_j^i$ , one will get:

$$\Theta^i = \bar{\omega}_j^i \wedge dx^j = -\frac{1}{2}(\bar{\omega}_{jk}^i - \bar{\omega}_{kj}^i) dx^j \wedge dx^k, \quad (3.24)$$

$$\Omega_j^i = d \wedge \bar{\omega}_j^i + \bar{\omega}_k^i \wedge \bar{\omega}_j^k = 0. \quad (3.25)$$

(Proof of last equation:

$$\begin{aligned} d \wedge \bar{\omega}_j^i + \bar{\omega}_k^i \wedge \bar{\omega}_j^k &= d \wedge (\tilde{h}_k^i dh_j^k) + \tilde{h}_k^i dh_m^k \wedge \tilde{h}_l^m dh_l^i \\ &= d\tilde{h}_k^i \wedge dh_j^k + \tilde{h}_k^i dh_m^k \wedge \tilde{h}_l^m dh_l^i \\ &= d\tilde{h}_k^i \wedge dh_j^k + \tilde{h}_l^i dh_m^l \tilde{h}_k^m \wedge dh_j^k \\ &= d\tilde{h}_k^i \wedge dh_j^k - d\tilde{h}_k^i \wedge dh_j^k = 0 \end{aligned}$$

The last expression in (3.24) shows that the components of the torsion 1-form are the antisymmetric parts of the connection components, as usual. That is, if:

$$\Theta^i = \frac{1}{2} \bar{\Theta}_{jk}^i dx^j \wedge dx^k \quad (3.26)$$

then

$$\bar{\Theta}_{jk}^i = \bar{\omega}_{jk}^i - \bar{\omega}_{kj}^i. \quad (3.27)$$

From the first equation in (3.23), and a glance back at (3.12), one can also see that the components of the torsion 2-form in the coframe  $\theta^i$  will be:



$$\Theta_{jk}^i = -c_{jk}^i. \quad (3.28)$$

**4. The Clifford-Weitzenböck geometry of dynamical space.** – We shall assume that topologically our dynamical space takes the form of the embedding in  $\mathbb{R}^4 = \mathbb{R} \times \mathbb{R}^3$  of a world-tube  $\mathcal{T}^4 = (t_0, t_1) \times \Sigma$ , where  $\Sigma$  is diffeomorphic to the interior of the support of  $\rho$  at any time-point  $t$ ; hence, the essential contribution to its topology shall come from  $\Sigma$ . That embedding will take  $(t_0, t_1)$  into the first factor of  $\mathbb{R}^4$  and  $\Sigma$  into the second one. The image  $\Sigma_t$  of  $\Sigma$  under that embedding at time  $t$  shall be referred to as the *instantaneous dynamical space at  $t$* .

In order to apply the preceding section to the case of the Madelung medium, we start by noting that the metric on each instantaneous dynamical space is defined by:

$$g = \rho \delta_{ij} dx^i dx^j = \delta_{ij} (R dx^i)(R dx^j) \quad (R \equiv \sqrt{\rho}). \quad (4.1)$$

Since  $\mathbb{R}^3$  is parallelizable, we can then think of its natural frame field  $\partial_i$  as being defined at all points of each  $\Sigma_t$ , although it is not generally adapted to that submanifold. However, it is a basic result of Stiefel [6] that if that space is compact, orientable, and three-dimensional then it will always be parallelizable. All of those conditions are physically reasonable. For instance, the (conformal) compactification of  $\mathbb{R}^3$  by a point at infinity will produce a 3-sphere, which relates to the study of functions that are constant (e.g., zero) at infinity.

*a. Basic frame for the deformed metric.* – From (4.1), the frame field  $\mathbf{e}_i$  and its reciprocal coframe field  $\theta^i$  that we need to define will take the simple forms:

$$\mathbf{e}_i = \frac{1}{R} \partial_i, \quad \theta^i = R dx^i; \quad (4.2)$$

i.e.:

$$h_j^i = R \delta_j^i, \quad \tilde{h}_j^i = \frac{1}{R} \delta_j^i. \quad (4.3)$$

The connection 1-form that this defines in the natural coframe field will then be:

$$\bar{\omega}_j^i = \tilde{h}_k^i dh_j^k = \left( \frac{1}{R} \delta_j^i \right) dR, \quad \bar{\omega}_{jk}^i = \delta_j^i \left( \frac{1}{R} \partial_k R \right). \quad (4.4)$$

Since:

$$\frac{1}{R} dR = d(\ln R) = \frac{1}{2} d(\ln \rho) = \frac{1}{2\rho} d\rho,$$

we can also say that:

$$\bar{\omega}_j^i = \left( \frac{1}{2\rho} \delta_j^i \right) d\rho. \quad (4.5)$$

If we compare this connection to the Levi-Civita connection that was defined in (2.1) then we will see that they are related by:

$$\Gamma_{jk}^i = \bar{\omega}_{jk}^i + \bar{\omega}_{kj}^i - \frac{1}{3} \bar{\omega}_{ml}^m \delta^{li} \delta_{jk}, \quad (4.6)$$

in which we have used the fact that we can solve the second equation in (4.4) for:

$$\frac{3}{\rho} \partial_k \rho = \bar{\omega}_{mk}^m, \quad \text{i.e.,} \quad \frac{3}{\rho} d\rho = \bar{\omega}_m^m. \quad (4.7)$$

It is important to note that the connection form  $\bar{\omega}_j^i$  takes its values in the Lie algebra of  $(\mathbb{R}, \times)$ , so it will not be a metric connection, as opposed to  $\Gamma_{jk}^i$ , nor will it have vanishing torsion, as we shall see.

If we recall that  $\rho = e^{2\lambda}$  then we can also say that:

$$\bar{\omega}_j^i = \delta_j^i d\lambda = (\delta_j^i \partial_k \lambda) dx^k, \quad (4.8)$$

so

$$\bar{\omega}_{jk}^i = \delta_j^i \partial_k \lambda. \quad (4.9)$$

One then has:

$$d\lambda = \frac{1}{3} \bar{\omega}_i^i. \quad (4.10)$$

Since  $R = e^\lambda$  is a homothety that acts upon coframes on the instantaneous dynamical spaces, one can regard  $\lambda$  as the infinitesimal generator of that homothety, which will then be an element of the Lie algebra of  $(\mathbb{R}, \times)$ . For some purposes, it is more convenient to regard the connection 1-form  $\bar{\omega}_j^i$  as taking the form:

$$\bar{\omega}_j^i = \delta_j^i \bar{\omega}, \quad (4.11)$$

in which the 1-form  $\bar{\omega}$  will then take the elementary form:

$$\bar{\omega} = d\lambda. \quad (4.12)$$

One also finds that the deformed coframe  $\theta^i$  will generally be anholonomic:

$$\Theta^i = d \wedge \theta^i = \frac{1}{2} (\bar{\omega}_{jk}^i - \bar{\omega}_{kj}^i) dx^j \wedge dx^k$$

so:

$$\bar{\Theta}_{jk}^i = \partial_j \lambda \delta_k^i - \partial_k \lambda \delta_j^i. \quad (4.13)$$

Hence, the torsion of  $\bar{\omega}_j^i$  will be non-vanishing as long as the density  $\rho$  is not spatially constant. That also says that the Lie algebra that is defined by the structure functions  $c_{jk}^i(t, x) = -\bar{\Theta}_{jk}^i(t, x)$  at each  $(t, x)$  will not be that of the Lie algebra of  $(\mathbb{R}, +)$ , which is Abelian, so its structure constants would have to vanish. Of course, the curvature of  $\bar{\omega}_j^i$  will vanish identically.

*b. The frame strain that is due to  $\rho$ .* – The type of frame strain that is of immediate interest to us is the simplest of all of them: a dilatation. The function  $\rho(t, x)$  that defines by a mass density on  $\mathcal{T}^4$  acts upon tangent vectors, and therefore tangent frames, as well. Moreover, it is clearly vertical, since the linear transformation that it induces in any tangent space is a homothety – i.e., a scalar multiple of the identity transformation – that projects to the identity transformation on  $\mathcal{T}^4$ . (Of course, there are such things as non-trivial, local diffeomorphisms of  $\mathcal{T}^4$  that differentiate to a homothety in each tangent space, but the dilatation that is defined by  $\rho$  is not due to one of them.)

If we anticipate the ultimate coupling of infinitesimal frame strain to the Madelung-Takabayasi stress tensor then we can rather concisely define that kinematical concept by way of:

$$\bar{\omega} = d\omega \quad \bar{\omega}_{ij} = \frac{\partial^2 \lambda}{\partial x^i \partial x^j}. \quad (4.14)$$

*c. The stress that is associated with the frame strain.* – Now, recall the definition of the Madelung-Takabayasi stress tensor in the form:

$$\sigma_{ij} = \frac{\hbar^2}{4m} \rho \frac{\partial^2 \ln \rho}{\partial x^i \partial x^j} = \frac{1}{2} \hbar^2 n \frac{\partial^2 \lambda}{\partial x^i \partial x^j}. \quad (4.15)$$

We see that this can be expressed as a constitutive law by way of:

$$\sigma_{ij} = \frac{1}{2} \hbar^2 n \bar{\omega}_{ij}. \quad (4.16)$$

Thus, one finds that, once again, the constant  $\hbar$  is being used as part of a constitutive law, and once more multiplied by the number density, although one cannot simply absorb  $n$  into  $\hbar = \hbar n$  as conveniently, since  $\hbar$  is squared, but  $n$  is not.

All of this gives us more confidence about asserting that the quantum stress that appears in the continuum-mechanical formulation of the Schrödinger equation originates in the deformation of the spatial metric into a dynamical metric by means of the mass density.

**§5. The possible role of the quantum vacuum.** One must realize from the outset that the Schrödinger equation is, presumably, only a semi-empirical approximation to

something more fundamental. Thus, the equations that would result from converting it would also have to manifest some of that approximate character. One then must wonder what sort of extension of the Madelung-Takabayasi equations might be appropriate, considering that it would also have to correspond to some extension of the Schrödinger equation, as well.

Some of the extensions that took place historically were to a relativistically-invariant form of the Schrödinger equation (viz., the Klein-Gordon equation), the inclusion of spin (viz., the Pauli equation), and a relativistically-invariant wave equation for spinning particles (viz., the Dirac equation). We shall discuss all of those extensions in the following chapters of this book, but for now, we shall consider the extension of scope that comes from thinking of mechanics as a limiting case of field theory.

In quantum field theory, the most fundamental construction seems to be that of the quantum vacuum state (or probably, state *space*). That is because one thinks of elementary particles as being basically quantum fluctuations or excitations of that vacuum state. In quantum electrodynamics, that typically takes the form of applying creation or annihilation operators to some (non-zero) vacuum wave function at each point of space-time. However, the existence of the spontaneous breaking of the symmetry of that vacuum state can turn the single vacuum state into an orbit of such states under the action of the broken symmetry group. Furthermore, if the QED vacuum state takes the form of the “zero-point field” then one would expect that it is highly unlikely that such a field would be defined uniquely, as opposed to being a more stochastic sort of ensemble of random fields.

Here, one must address the issue of why one introduces stochastic (i.e., probabilistic) considerations into any scientific theory. Basically, probability is the science of ignorance. That is, one turns to probability and stochastic systems when one has only so much information about the states of a complex system and their dynamics <sup>(1)</sup>. Hence, the real issue is the incompleteness of the model, which often takes the form of some fundamental approximation that is usually introduced in order to avoid some intractable complexity.

Some of the forms that such complexity can take are nonlinearity, heat, friction, viscosity, turbulence, vibration, and unmodeled degrees of freedom in a system. In the case of quantum physics, one can say that one is often dealing with the breakdown of the “test particle” approximation, which is also closely related to the “external field” approximation.

Basically, the approximation that is associated with test particles and external fields is that the orders of magnitude between the effect of the test particle and that of the external field are so appreciable that one can ignore the interaction between them. That is, a test particle cannot appreciably alter the external field that it interacts with. However, one can see that although that assumption is perfectly reasonable for a microwave photon from a radar unit interacting with an oncoming motor vehicle on a freeway, it is totally unrealistic when that same photon interacts with an atomic electron.

---

<sup>(1)</sup> This is to be distinguished from the way that engineering models often introduce randomness even when a deterministic model exists, simply because the excess complexity in the model is unnecessary for the achievement of the overall goal of the system. For instance, the flat-Earth approximation for gravity is generally adequate for air-to-air missiles because their flight time is only a matter of seconds.

Thus, one might wish to think of the quantum particle that is being described by the Schrödinger wave function and equation as being an approximation to something that is in interaction with its environment, which one might think of as the space of quantum vacuum states. For now, we shall pass over the equally-important extension of scope that comes from going from a linear wave equation to a nonlinear one, which is also probably very much in the nature of quantum physics, since one expects to be dealing with enormous electric and magnetic field strengths at the atomic and subatomic levels.

The extensions of the Madelung-Takabayasi model for the Schrödinger equations that we shall consider came from Bohm and Vigier, Schönberg, and de Broglie. Since Bohm and Vigier were working independently of Schönberg at roughly the same time, we shall start with the former paper, since it had a more general character.

*a. Bohm and Vigier's inclusions of fluctuations.* In 1954, David Bohm and Jean-Pierre Vigier attempted to enlarge the scope of the Madelung picture in a paper that was published in the Physical Review [11]. Their motivation was based upon the idea that even though one was ultimately trying to replace the statistical interpretation of the quantum wave function with a causal interpretation, nonetheless, one still had to account for the success of the statistical interpretation; that is the spirit of the correspondence principle.

Their approach to that problem was to then assume that one had a large, statistical ensemble of Madelung media moving together in the presence of a “sub-quantum” medium that they interacted with in such a way as to produce random fluctuations of the number/mass density  $\rho(t, x^i)$  and flow velocity  $\mathbf{v}(t, x^i)$ . The possible origins of the fluctuations that they suggested were:

1. Random external disturbances that were transmitted through the boundary of the medium.
2. Nonlinearity in the dynamical equations might lead to instability and turbulence in the flow.
3. Brownian motion of the Madelung media at a “molecular” level.

The Madelung-Takabayasi picture was then assumed to be something that was only supposed to apply to the mean density  $\bar{\rho}(t, x^i)$  and mean flow velocity  $\bar{\mathbf{v}}(t, x^i)$ , except that the continuity equation was presumed to still apply for the corrupted flow variables. Thus, one was assumed to have:

$$\frac{\partial \bar{\rho}}{\partial t} + \text{div}(\bar{\rho} \bar{\mathbf{v}}) = \frac{\partial \rho}{\partial t} + \text{div}(\rho \mathbf{v}) = 0. \quad (5.1)$$

The volume elements of matter were assumed to have some (otherwise unspecified) tendency to move to regions of highest mean density, which amounted to assuming the stability of that equilibrium density. After introducing the further assumption that such a volume element  $dV'$  at  $(t', x'^i)$  would always have a non-zero probability  $K(t, x^i; t', x'^i)$  of

making the transition to a volume element  $dV$  at  $(t, x^i)$ , they proved the theorem that the limit of  $\rho(t, x^i)$  as  $t$  becomes infinite would be  $\bar{\rho}(t, x^i) = |\psi|^2$ .

Their proof was a generalization of the proof of a well-known theorem of Markov processes to the case in which  $K(t, x^i; t', x'^i)$  and  $\bar{\rho}(t, x^i)$  were time-varying. The theorem was essentially equivalent to Boltzmann's  $H$ -theorem in statistical thermodynamics.

*b. Schönberg's inclusion of turbulence.* At roughly the same time as Bohm and Vigier, Schönberg [12] was extending the scope of the Madelung-Takabayasi model along lines that the latter researchers had briefly suggested, namely, the idea that the random fluctuations would take the form of turbulence in the "fluid." By contrast, Schönberg, like Takabayasi, was beginning with the more causal assumption that the irrotational character of the Madelung medium could be easily generalized, so the appearance of a chaotic cascade of vortex-pairs that characterizes the onset of turbulence would be reasonable from that standpoint.

Schönberg's stated objective in writing his paper was to derive an improved version of the Madelung-Takabayasi model from the usual interpretation of the Schrödinger equation. Like Bohm and Vigier, he also consider the Madelung-Takabayasi model to represent the mean motion of a fluid that was corrupted by fluctuations, but he specifically attributed the fluctuations to be due to turbulence, since he had previously been considering the extension of both Madelung-Takabayasi equations and the Schrödinger equations to the case in the Madelung medium was no longer irrotational.

However, Schönberg then diverged from the approach of Bohm and Vigier by drawing upon the statistical theory of turbulence in particular, rather than the theory of Markov processes in general. In fact, he showed that the usual stochastic modeling of turbulence also involves a process of "second quantization" that replaces the dynamical functions (mass density, momentum density, etc.) with operators and then recovers the mean motion by taking the expectation values of the operators in a state, which is essentially Ehrenfest's theorem [13]. The dynamical functions themselves are assumed to be related by the Navier-Stokes equation, which governs the time evolution of viscous fluids. Since, as we showed previously, the Madelung medium is usually inviscid, but turbulence cannot form in inviscid fluids, one sees that turbulence would also require that extension of the scope of the Madelung-Takabayasi model.

If one uses the Heisenberg representation for quantum mechanics then one can find a direct analogue in that representation with the stochastic differential equations of turbulence. One of the intriguing consequences of that is that when one passes to expectation values, one finds that it is known that the stress tensor for the mean motion of a turbulent fluid is not the same as the expectation values of the stress tensor for the turbulent motion. The difference between the two stress tensors is explained by the fact that there will be a certain amount of momentum that is transferred by the turbulence. A consequence of this is that one can then account for the terms of the quantum stress tensor of the Madelung medium as being totally consistent with some standard constructions of the theory of turbulence, namely,  $\frac{\hbar}{2m} \frac{\partial \sqrt{R}}{\partial x^i} \frac{\partial \sqrt{R}}{\partial x^j}$  comes from the

transfer of momentum by turbulence, while  $-\frac{\hbar^2}{4m}\Delta R\delta_{ij}$  represents the elastic tension in the medium.

The introduction of turbulence as a stochastic or thermodynamic background in the relativistic version of the theory would proceed in essentially the same way as in the non-relativistic case, except that one would have to deal with the relativistic theory of turbulence, which is not very well developed, in addition to relativistic thermodynamics. The effect of it would be the same as before, namely, if one regards the current formulation of the Madelung-Takabayasi picture as the “classical ground state” then the turbulence would represent “quantum fluctuations” of that classical ground state.

*c. De Broglie’s hidden thermodynamics.* Louis de Broglie had been one of the first to look into Madelung’s theory, and that study culminated in the publication of his book [14]. Eventually he decided that there was still a fundamental incompleteness in the hydrodynamical interpretation, and gradually learned to accept the statistical interpretation of Born and the Copenhagen School. However, the paper of Bohm and Vigier (the latter of whom had been a student of de Broglie who also wrote a book [15] on the causal interpretation of quantum mechanics) rekindled de Broglie’s interest in the hydrodynamical interpretation, which ultimately resulted in a book on what he was calling the “hidden thermodynamics” of an isolated particle [16].

The basic gist of that concept is that at the quantum level, an isolated point particle will still continually exchange energy and momentum with what de Broglie was calling a “hidden thermostat,” and which corresponded to the “sub-quantum medium” of Bohm and Vigier or what would now be called simply the “quantum vacuum.” He used the analogy of a particle of small mass moving in a warm fluid that is then subject to Brownian motion due to the fluctuations that are produced by the exchange of energy between the motion of the particle and the heat of the environment.

De Broglie was mostly drawing upon some early work that had been done in thermodynamics in the hope of formulating a “mechanical theory of heat,” which predated the more currently standard statistical foundations. The former theory had its roots in the work of Clausius, who had originally introduced the concept of entropy, and later Szily, Helmholtz, and Boltzmann, that largely centered on the attempt to formulate the second law of thermodynamics as being, in some way, analogous to the principle of least action. Typically, one wished to establish an analogy between action and entropy.

Helmholtz began by dividing the (generalized) coordinates of a mechanical system into two types: slowly-varying and rapidly-varying. The former would generally be associated with the external forces and forces of constraint, which were also assumed to be conservative. The latter would represent molecular motions at the quantum level and would thus manifest themselves in the expressions for kinetic energy. Hence, the total energy of the system would consist of potential energy that was due to the slowly-varying variables and kinetic energy that was due to the rapidly-varying ones. Helmholtz managed to obtain the desired analogy only for what he called “monocyclic” systems, which had only one rapidly-varying coordinate. The result that he derived was the theorem that if all of the work  $\mathcal{W}$  that was done upon the system affected only the cyclic coordinate then the force 1-form  $\mathcal{A}\mathcal{W}$  would be such that  $\mathcal{A}\mathcal{W}/KE$  would be an exact 1-form. One sees that when one passes to mean values (so that the temperature  $T$

represents the mean value of the kinetic energy), that exact 1-form will be  $dS$ , where  $S$  is the entropy of the system. (Sometimes this is a definition.)

Boltzmann attempted to extend the result of Helmholtz to polycyclic mechanical systems (i.e., more than one rapidly-varying coordinate). However, he had to restrict the scope to periodic systems, so he could introduce period integrals. His generalization of the Helmholtz result then took the form of saying that variation  $\delta\mathcal{W}$  of the work that was done on the system would relate to the variation of the (Maupertuisian) action over a period:

$$\mathcal{A} = \oint p_k dx^k$$

by way of:

$$\delta\mathcal{W} = \frac{\delta\mathcal{A}}{\tau} = \nu \delta\mathcal{A} = \frac{2}{\tau} \delta(\overline{KE} \tau),$$

in which  $\tau$  is the period of the motion, so  $\nu$  will be its frequency (in Hz).

Starting with these results as a basis, de Broglie then derived some key formulas for the interaction between the isolated particle and the hidden thermostat, namely:

$$h\nu_c = kT, \quad \frac{\mathcal{A}}{h} = \frac{S}{k},$$

in which the thermodynamic variables  $T, S$  relate to the thermostat, while the mechanical variables  $\nu_c, \mathcal{A}$  relate to the particle. The frequency  $\nu_c$  was actually a relative frequency:

$$\nu_c = \nu_0 \sqrt{1 - v^2/c^2},$$

in which  $\nu_0$  was then the rest frequency, and  $v$  was the relative speed.

De Broglie also assumed that the rest mass of the particle  $m_0$  was prone to fluctuations, as a result of the ongoing exchange of energy-momentum with the thermostat, and that fluctuating value was then denoted by  $M_0$ . He then derived another key formula in the form of:

$$S(M_0) - S_0 = -k \frac{M_0}{m_0}.$$

Thus, the fluctuation of the mass of the particle was compensated by a fluctuation in the entropy of the thermostat about  $S_0$ .

Of course, what one actually measures will be the mean value of  $M_0$ , and under the assumption that  $\Delta M_0 = M_0 - m_0$  has a mean of zero, that will be  $m_0$ . Although one leaves open the possibility that the value of  $\Delta M_0$  can be infinite, nevertheless, one would associate that possibility with a correspondingly infinitesimal probability.

Note that when one uses real-world values for the physical constants, one finds that the equivalent temperature of a mass  $m_0$  that is due its rest energy (namely,  $m_0 c^2 / k \sim 10^{37} m_0$ ) can be unimaginably high. For instance, the rest mass of an electron would correspond to a temperature on the order of  $10^{10}$  K. To give that some context, note that



the temperature at which nuclear fusion reactions take place in the Sun is typically of order  $10^7$  K, while the Planck scale of temperature is of order  $10^{32}$  K.

As we just saw, it is probably premature to speak of the nature of the quantum vacuum in advance of the consideration of relativistic considerations. Indeed, one should remember that we are, after all, talking about *wave* mechanics, and the basic Lorentzian metric of relativity is, of course, derived from the most elementary dispersion law for electromagnetic waves. Hence, one would expect that wave mechanics would properly start in a relativistic context, so we shall revisit the question of the quantum vacuum after we have had a chance to discuss the continuum-mechanical form of the relativistic wave equations.

**§ 6. The relationship between the Madelung transformation and the WKB method.** – Starting in the early Nineteenth Century, a method for solving ordinary differential equations of a certain type by successive approximations emerged from some of the work of Joseph Liouville [17]. Unlike the Taylor series method, though, the power series used did not generally converge as one went to more terms in the summation, but tended to improve up to a point and then diverge after that. However, if one truncated the series after a finite number of terms then typically, the approximation would improve as one let the series parameter go to zero. Such a formal power series was then referred to as an *asymptotic expansion*. Other mathematicians contributed to the technique over the next one hundred years or so, including Poincaré, who applied the technique to celestial mechanics [18], Lord Rayleigh [19], who was concerned with a problem in optics, and Harold Jeffreys [20], who was still basically concerned with the theory of ordinary differential equations.

All of that work predated the introduction of the Schrödinger equation into physics, and it was Gregor Wentzel [21], Hendrik Kramers [22], and Léon Brillouin [23] that applied the technique of asymptotic expansions to that particular equation. It then became known amongst physicists as the *WKB method*, although some still insist upon including a “J” for “Jeffreys.” (There are even some who call it the “LR method” for “Liouville-Rayleigh,” according to Heading [24].) A particularly concise and rigorous mathematical treatment of the WKB method was published by George Birkhoff in 1933 [25]. By now, the applications of the technique have expanded to a good number.

In one manifestation (see Messiah [26]), the method starts off the same as the Madelung transformation; that is, one represents the (typically stationary) wave function  $\psi(x^i)$  in the polar form:

$$\psi = \exp\left(\frac{iW}{\hbar}\right) = \exp\left(\frac{iS}{\hbar} + T\right) = R e^{iS/\hbar} \quad (R \equiv e^T). \quad (6.1)$$

Hence, when this is substituted in the stationary Schrödinger equation, one will get the same equations that Madelung obtained when one separates the real and imaginary parts:

$$(d_\nu S)^2 = 2m(E - U) + \hbar^2 \frac{\Delta R}{R}, \quad \text{div}(R^2 \nabla S) = 0. \quad (6.2)$$

So far, no actual approximation has been introduced into the equations. Hence, equations (6.2) are rigorously equivalent to the stationary Schrödinger equation. The WKB method then parts company from the Madelung formulation by treating the function  $W$  as a function of not only  $x^i$ , but  $\hbar^2$  (i.e., an even function of  $\hbar$ ) and expanding it in a power series in  $\hbar^2$ :

$$W = W_0 + \hbar^2 W_1 + \hbar^4 W_2 + \dots \quad (6.3)$$

Indeed, one can also simply expand the  $S$  or  $T$  part of  $W$  (or, equivalently,  $R$ ) in such a power series. For instance:

$$S = S_0 + \hbar^2 S_1 + \hbar^4 S_2 + \dots \quad (6.4)$$

When this is substituted in the first of (6.2), one first computes:

$$d_s S = d_s S_0 + \hbar^2 d_s S_1 + \hbar^4 d_s S_2 + \dots \quad (6.5)$$

In order to square that 1-form (using the spatial metric, of course), we shall truncate the series with the specified terms and then truncate the square after the  $\hbar^4$  term:

$$(d_s S)^2 \approx (d_s S_0)^2 + \hbar^2 2(d_s S_0, d_s S_1) + \hbar^4 [(d_s S_1)^2 + 2(d_s S_0, d_s S_2)] + \dots \quad (6.6)$$

Combining terms with common powers of  $\hbar^2$  then gives:

$$0 = (d_s S_0)^2 - 2m(E - U) + \hbar^2 [2(d_s S_0, d_s S_1) - \Delta R / R] + \hbar^4 [(d_s S_1)^2 + 2(d_s S_0, d_s S_2)]. \quad (6.7)$$

Equating coefficients of powers of  $\hbar^2$  will then give the series of equations:

$$(d_s S_0)^2 = 2m(E - U), \quad 2(d_s S_0, d_s S_1) = \Delta R / R, \quad 0 = (d_s S_1)^2 + 2(d_s S_0, d_s S_2). \quad (6.8)$$

The first one is the classical Hamilton-Jacobi equation for the action function  $S_0$ . Having solved it for  $S_0$  (by Cauchy's method of characteristics), one can then treat the next equation as a linear, first-order partial differential equation for  $S_1$ , which can be solved for  $S_1$ , and so on. Hence, one gets a recursive algorithm that starts with solving the classical Hamilton-Jacobi equation. Note that the so-called "quantum potential" (viz.,  $\Delta R / R$ ) does not appear until the second step, since it is initially multiplied by  $\hbar^2$  even before the power series has been introduced. Indeed, if  $\hbar$  did not appear in the original equation then there would be no recursion from one  $S_n$  to the next, and the successive equations would be independent.

Quantum field theory employs an analogous series expansion in powers of  $\hbar^2$  that they refer to as the "loop expansion," since it involves sorting the perturbation series for the relativistic scattering amplitudes (i.e., the Feynman diagrams) into diagrams with no loops, one loop, etc. Loops correspond to renormalizations, which are often associated vacuum polarization. The zero-loop (or "tree" graph) approximation is then the classical

field-theoretic picture, while subsequent terms in the loop expansion represent “quantum corrections” of increasing precision.

Such a loop expansion is then used to define “effective” field theories, which amount to classical field theories with successive quantum corrections. For instance, one might define an “effective potential” that amounts to a classical potential (e.g., Coulomb), plus a one-loop term, etc. It is also common to define “effective actions” that amount to the classical action plus higher-loop corrections.

Hence, one sees that although the Madelung transformation of the Schrödinger equation involves no actual approximation, nonetheless, from the WKB standpoint, one can think of the quantum potential as a “one-loop” correction to the classical potential  $U$ .

**§ 7. The optical interpretation of quantum mechanics.** – Another compelling analogy between quantum mechanics and the physics that came before it (besides the “hydrodynamical” interpretation), and which did not get nearly as much attention, was mentioned in Messiah [26]. He called it the “optical analogy,” and one sees that it should have an immediate applicability to quantum theory that is based upon the fact that Schrödinger was explicitly trying to make quantum wave mechanics relate to classical (i.e., geometrical) mechanics in the same way that wave optics already related to geometrical optics.

In its essence, wave optics is concerned with the motion of momentary wave-fronts through a space that is associated with non-trivial optical properties (e.g., indices of refraction). Dually, geometrical optics considers the behavior of rays, which are curves that are typically orthogonal trajectories to the motion of the momentary wave fronts.

Since most optical media tend to be time-invariant in their properties, Messiah starts with the stationary Schrödinger equation, introduces the polar decomposition of the stationary spatial wave function  $y = A e^{iS/\hbar}$ , in which we have changed  $R$  to  $A$  to mean “amplitude,” and obtained the Madelung form of the stationary Schrödinger equation, as in (6.2), but with  $A$  in place of  $R$ .

We shall part company with Messiah slightly here, since we only wish to find the partial differential equation for the spatial wave fronts and the system of ordinary differential equations for their orthogonal trajectories; viz., the rays. Hence, we set  $S = \hbar\theta$ , which will put the first equation of (6.2) into the form:

$$\frac{2m}{\hbar^2} (E - U) - (d_s\theta)^2 = -\frac{\Delta A}{A}. \quad (7.1)$$

If we introduce the definitions:

$$k = \omega dt + k_s, \quad \omega^2 = \frac{2m}{\hbar^2 c^2} E, \quad k_s = d_s\theta \quad (7.2)$$

then (7.1) will take the form of a dispersion relation for waves in the spatial medium:

$$k^2 = k_0^2 c^2 \quad (7.3)$$

with

$$k_0^2 c^2 = \frac{2m}{\hbar^2} U - \frac{\Delta R}{R} = \frac{2m}{\hbar^2} (U + U_h). \quad (7.4)$$

Of course, we have encountered this dispersion relation before as an intermediate step between quantum wave mechanics and continuum mechanics.

The momentary wave fronts in space will, of course, be the level surfaces of  $\theta$ . In order to get the orthogonal trajectories to those surfaces, one needs only to define the normal vector field to them, which amounts to the gradient of  $\theta$ :

$$\mathbf{n} = \nabla \theta = (g^{ij} \theta_{,j}) \partial_i. \quad (7.5)$$

Although  $\mathbf{n}$  is proportional to  $\mathbf{p} = \nabla S$ , which is more like a momentum vector field, that factor of proportionality will affect only the parameterization of the orthogonal trajectories, which will be the integral curves  $x(s)$  of the system of ordinary differential equations:

$$\frac{dx^i}{ds} = \nabla \theta(x) \quad \left( \frac{dx^i}{ds} = g^{ij} \frac{\partial \theta}{\partial x^j} \right). \quad (7.6)$$

To the extent that the theory of relativity is rooted in the way that electromagnetic waves propagate through space, it should not be surprising that the optical analogy properly belongs to the realm of relativistic wave mechanics, which we shall also examine in book.

The concept of an asymptotic series has especial relevance to the optical analogy, since the application of asymptotic expansions to optics has been established in the context of diffraction problems. (For instance, see Pauli [27].) In effect, one starts with geometrical optics as the classical solution and adds successive corrections that originate in diffraction effects. The expansion parameter is usually wavelength in that case, so geometrical optics becomes the “small-wavelength” (i.e., high wave number) approximation.

### References (\*)

1. D. H. Delphenich, “The Geometric Origin of the Madelung Potential,” gr-qc/0211065.
2. D. H. Delphenich, “A strain tensor that couples to the Madelung stress tensor,” arXiv:1303.3582.
3. D. H. Delphenich, “The use of the teleparallelism connection in continuum mechanics,” arXiv:1305.3477.
4. Sir W. Thomson (Lord Kelvin) and P. G. Tait, *Treatise on Natural Philosophy*, 2 vols., Cambridge University Press, Cambridge, 1895; available as a facsimile reprint through Elibron Classics, www.elibron.com, 2005.

---

(\*) References marked with an asterisk are available in English translation at the author’s website: neo-classical-physics.info.

5. D. H. Delphenich, ed., *Selected Papers in Teleparallelism*, Neo-classical Press, 2013; available as a free PDF download at neo-classical-physics.info.
- 6\*. E. Stiefel, "Richtungsfelder und Fernparallelismus in n-dimensionalen Mannigfaltigkeiten," *Comm. Math. Helv.*, **8** (1936), 3-51.
7. E. Cartan and J. A. Schouten, "On the geometry of the group-manifold of simple and semi-simple groups," *Proc. Kon. Akad. Wet. Amsterdam* **29** (1926), 803-815.
8. W. K. Clifford, "Preliminary sketch of biquaternions," *Proc. London Math. Soc.* (1) **4** (1873), 381-395; *Mathematical Papers*, pp. 181-200.
9. R. Weitzenböck, *Invariantentheorie*, Noordhoff, Gronigen, 1923.
10. I. Singer and S. Sternberg, "The Infinite Groups of Lie-Cartan, Pt. I (the transitive groups)," *J. Analyse Math.*, **15** (1965) 1-114.
11. D. Bohm and J.-P. Vigièr, "Model of the Causal Interpretation of Quantum Theory in Terms of a Fluid with Irregular Fluctuations," **96** (1954), 208-216.
12. M. Schönberg, "On the Hydrodynamical Model of the Quantum Mechanics," *Nuov. Cim.* **12** (1954), 103-133.
- 13\*. P. Ehrenfest, "Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik," *Zeit. Phys.* **45** (1927), 455-457.
14. L. de Broglie, *Nonlinear wave mechanics*, Elsevier, Amsterdam, 1960.
- 15\*. J.-P. Vigièr, *Structure des micro-objets dans l'interprétation causale de la théorie des quanta*, Gauthier-Villars, Paris, 1956.
- 16\*. L. de Broglie, *La Thermodynamique de la particule isolée: ou thermodynamique cachée des particules*, Gauthier-Villars, Paris, 1964.
17. J. Liouville, "Second Mémoire sur le développement des fonctions ou parties de fonctions en séries dont les divers termes sont assujétis à satisfaire à une même équation différentielle du second ordre, contenant un paramètre variable," *J. Math. pures et appl.* **2** (1837), 16-35.
18. H. Poincaré:
  - "Sur les intégrales irrégulières des équations linéaires," *Acta Math.* **8** (1886) 259-344.
  - *Les méthodes nouvelles de la mécanique céleste*, v. 1, Chap. 3., Dover, Mineola, NY, 1892.
19. Lord Rayleigh, "On the Propagation of Waves through a Stratified Medium, with Special Reference to the Question of Reflection," *Proc. Roy. Soc. A* **86** (1912), 207-226.
20. H. Jeffreys, "On certain approximate solutions of linear differential equations of the second order," *Proc. London Math. Soc.* (2) **23** (1923), 428-436.
- 21\*. G. Wentzel, "Eine verallgemeinerung der Quantenbedingungen für die Zwecke der Wellenmechanik," *Zeit. Phys.* **38** (1926), 518.
- 22\*. H. A. Kramers, "Wellenmechanik und halbzahlige Quantisierung," *Zeit. Phys.* **39** (1926), 828-840.
- 23\*. L. Brillouin, "La mécanique ondulatoire de Schrödinger: une méthode de résolution par approximations successives," *C. R. Acad. sci. Paris* **183** (1926), 24-26.
24. J. Heading, *Phase-Integral Methods*, Methuen and Co., London, 1962.
25. G. D. Birkhoff, "Quantum mechanics and asymptotic series," *Bull. Amer. Math. Soc.* **39** (1933), 681-700.

26. A. Messiah, *Quantum Mechanics*, v. I, John Wiley and Sons, Inc., NY, 1958.
  27. W. Pauli, “On asymptotic series for functions in the theory of diffraction of light,” *Phys. Rev.* **54** (1938), 924-931.
-