Adapting Mathematical Methods from Quantum Mechanics to Classical Waves

Max Lein

December 11, 2019

Tohoku University
WPI-AIMR

Contents

Chapter 1

1 Quantum–Wave Analogies

Ever since the inception of modern quantum mechanics in 1926 physicists have relied on *quantum–wave analogies*, i. e. systematic similarities shared between quantum systems and analogous systems for classical waves. The foundation of these is ultimately that *both theories describe waves.* At the inception of modern quantum mechanics in the 1920s, physicists borrowed intuition from electromagnetism and other classical waves, and transferred it to quantum systems. However, in recent years physicists read them the other way around: they would like to propose new experiments with classical waves by taking inspiration from quantum mechanics. **The goal of this lecture series is to show *how to make quantum-wave analogies rigorous,* starting from the fundamental equations.**

$$\left.\begin{array}{c}\textit{Quantum Mechanics}\\ \mathrm{i}\,\partial_t \Psi = H\Psi \\ H = (-\mathrm{i}\nabla - A)^2 + V \\ \text{(Schrödinger equation)}\end{array}\right\} \quad \stackrel{?}{\longleftrightarrow} \quad \left\{\begin{array}{c}\textit{Classical Electromagnetism}\\ \begin{pmatrix}\varepsilon & 0\\ 0 & \mu\end{pmatrix} \frac{\partial}{\partial t}\begin{pmatrix}\mathbf{E}\\ \mathbf{H}\end{pmatrix} = \begin{pmatrix}-\nabla\times\mathbf{H}\\ +\nabla\times\mathbf{E}\end{pmatrix} \\ \text{(dynamical equations)}\\ \begin{pmatrix}\nabla\cdot & \\ & \nabla\cdot\end{pmatrix}\begin{pmatrix}\varepsilon & 0\\ 0 & \mu\end{pmatrix}\begin{pmatrix}\mathbf{E}\\ \mathbf{H}\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix} \\ \text{(constraint equation)}\end{array}\right.$$

The idea here is to systematically rewrite classical wave equations in the form of a Schrödinger equation. This rephrasing rests *solely* on *mathematical similarities* between the two theories, but is *not* meant to imply that the physical interpretation from quantum mechanics carries over. Recasting a classical wave equation as a Schrödinger equation does two things: first of all, it gives us access to all the tools initially developed for quantum systems and allows us to apply them to classical wave equations. Secondly, this presents a unified formalism for a whole class of classical wave equa-

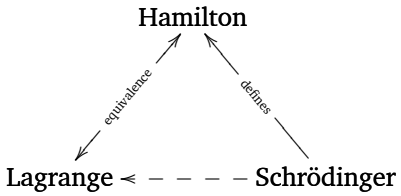

Figure 1.0.1: Many mathematical equations can be rephrased using different mathematical frameworks. That includes, the Schrödinger formalism, the framework of classical Hamiltonian equations of motion and the Lagrangian formalism that defines equations of motion via variational principles.

tions, and the Schrödinger formalism enables comparisons between *different* classical wave equations.

Of course, we could study classical wave equations in other mathematical frameworks (see Figure 1.0.1) such as the Lagrangian or Hamiltonian formalism. Each gives access to a different set of tools (e. g. Noether's theorem in case of Lagrangian mechanics), so for specific questions one may be more advantageous than another. And to study similarities to quantum mechanics, not surprisingly, it is advantageous to rewrite the wave equation in the same language that a quantum theory is described in.

Chapter 2

2 Paradigms of quantum mechanics

The explanation of the photoelectric effect through light *quanta* is the name sake for quantum mechanics. Quantization here refers to the idea that energy stored in light comes in "chunks" known as *photons*, and that the energy per photon depends only on the frequency. This is quite a departure from the classical theory of light through Maxwell's equations (cf. [Jac98]).

2.1 Two archetypical quantum systems

To introduce the main notions and emphasize the physics, we start by way of example.

2.1.1 The Stern-Gerlach experiment

The simplest *bona fide* quantum system is that of a quantum spin, and it can be used to give an effective description of the *Stern-Gerlach experiment* where a beam of neutral atoms with magnetic moment g is sent through a magnet with inhomogeneous magnetic field $B = (B_1, B_2, B_3)$. It was observed experimentally that the beam splits in two rather than fan out with continuous distribution. Hence, the system behaves as if only two spin configurations, spin-up $\uparrow$ and spin-down $\downarrow$, are realized. A simplified effective model neglects the translational degree of freedom and focusses only on the internal spin degree of freedom of the particle. Then the energy observable, the *hamiltonian*, is the 2×2 matrix

$$H = gB \cdot S = H^*$$

which involves the spin operator $S_j := \frac{\hbar}{2}\sigma_j$ defined in terms of Planck's constant $\hbar$ and the three Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ +\mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix},$$

and the magnetic moment g and the magnetic field B. The prefactor of the Pauli matrices are real, and thus, $H = H^*$ equals its adjoint and is a *hermitian* matrix.

For instance, assume $B = (0, 0, b)$ points in the x_3-direction. Then spin-up and spin-down (seen from the x_3-direction) are the *eigenvectors* of

$$H = \begin{pmatrix} +\frac{\hbar g b}{2} & 0 \\ 0 & -\frac{\hbar g b}{2} \end{pmatrix},$$

i. e. $\psi_\uparrow = (1, 0)$ and $\psi_\downarrow = (0, 1)$. The dynamical equation is the *Schrödinger equation*

$$\mathrm{i}\hbar \frac{\partial}{\partial t}\psi(t) = H\psi(t), \qquad \psi(0) = \psi_0 \in \mathcal{H}. \tag{2.1.1}$$

The vector space $\mathcal{H} = \mathbb{C}^2$ becomes a Hilbert space if we equip it with the scalar product

$$\langle \psi, \varphi \rangle_{\mathbb{C}^2} := \sum_{j=1,2} \overline{\psi_j}\, \varphi_j.$$

Moreover, the hermitian matrix H can always be diagonalized, and the eigenvectors to distinct eigenvalues are orthogonal. The complex-valued *wave function* ψ encapsulates probabilities: for any $\psi \in \mathbb{C}^2$ normalized to $1 = \|\psi\|_{\mathbb{C}^2}$, the probability to find the particle in the spin-up configuration is

$$\mathbb{P}\big(\mathsf{S}_3 = \uparrow \mid \psi\big) = |\psi_1|^2 = \big|\langle \psi_\uparrow, \psi \rangle\big|^2$$

since $\psi_\uparrow = (1, 0)$. The above notation comes from probability theory and means "the probability of finding the random observable spin S_3 in the spin-$\uparrow$ configuration $+\frac{\hbar}{2}$".

2.1.2 A spinless, non-relativistic particle moving in $\mathbb{R}^d$

The second exemplary quantum system describes a non-relativistic particle of mass m subjected to an electric field generated by the potential V. The classical Hamilton *function* $h(q, p) = \frac{1}{2m}p^2 + V(q)$ is then "quantized" to

$$H = h\big(\hat{x}, -\mathrm{i}\hbar\nabla\big) = \frac{1}{2m}\big(-\mathrm{i}\hbar\nabla\big)^2 + V(\hat{x}) = H^*$$

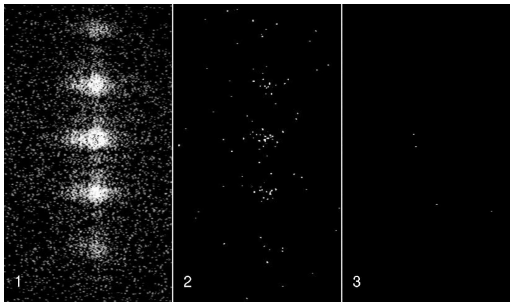

Figure 2.1.1: Images of a low-intensity triple slit experiment with photons (taken from [Cro08]).

by replacing momentum p by the momentum *operator* $\mathsf{P} = -\mathrm{i}\hbar\nabla$ and position q by the multiplication operator $\mathsf{Q} = \hat{x}$.[1] The hamiltonian is now an operator on the Hilbert space $L^2(\mathbb{R}^d)$ whose action on suitable vectors ψ is

$$(H\psi)(x) = -\frac{\hbar^2}{2m}(\Delta\psi)(x) + V(x)\,\psi(x).$$

Quantum particles simultaneously have wave and particle character: the Schrödinger equation (2.1.1) is structurally very similar to a wave equation. The physical constant $\hbar$ relates the energy of a particle with the associated wave length and has units [energy $\cdot$ time]. The particle aspects come into play when one measures outcomes of experiments: consider a version of the Stern-Gerlach experiment where the intensity of the atomic beam is so low that single atoms pass through the magnet. If the modulus square of the wave function $|\psi(t,x)|^2$ were to describe the intensity of a *matter* wave, then one expects that the two peaks build up slowly, but *simultaneously*. In actuality, one registers single impacts of atoms and only if one waits long enough, two peaks emerge (similar to what one sees in a low-intensity triple slit experiment in Figure 2.1.1). This is akin to tossing a coin: one cannot see the probabilistic nature in a few coin tosses, let alone a single one. Probabilities emerge only after repeating the experiment often enough. These experiments show that $|\psi(t,x)|^2$ is to be interpreted as a *probability distribution*, but more on that below.

Pure states are described by wave functions, i. e. complex-valued, square integrable

[1] To find a consistent quantization procedure is highly non-trivial. One possibility is to use Weyl quantization [Wey27; Wig32; Moy49; Fol89; Lei10]. Such a quantization procedure also yields a formulation of a semiclassical limit, and the names for various operators (e. g. position, momentum and angular momentum) are then justified via a semiclassical limit. For instance, the momentum operator is $-\mathrm{i}\hbar\nabla$, because in the semiclassical limit it plays the role of the classical momentum observable p (cf. e. g. [Lei10, Theorem 1.0.1] and [Lei10, Theorem 7.0.1]).

functions. Put more precisely, we are considering $L^2(\mathbb{R}^d)$ made up of equivalence classes of functions with scalar product

$$\langle \varphi, \psi \rangle = \int_{\mathbb{R}^d} \mathrm{d}x \, \overline{\varphi(x)} \, \psi(x)$$

and norm $\|\psi\| := \sqrt{\langle \psi, \psi \rangle}$. In physics text books, one usually encounters the *bra-ket* notation: here $|\psi\rangle$ is a state and $\langle x|\psi\rangle$ stands for $\psi(x)$. The scalar product of $\phi, \psi \in L^2(\mathbb{R}^d)$ is denoted by $\langle \phi|\psi\rangle$ and corresponds to $\langle \phi, \psi \rangle$. Although bra-ket notation can be ambiguous, it is sometimes useful and is also used in the mathematics literature.

The fact that $L^2(\mathbb{R}^d)$ consists of *equivalence classes* of functions is only natural from a physical perspective: if $\psi_1 \sim \psi_2$ are in the same equivalence class (i. e. they differ on a set of measure 0), then the associated probabilities coincide: Physically, $|\psi(t,x)|^2$ is interpreted as the *probability to measure a particle at time t in (an infinitesimally small box located in) location x*. If we are interested in the probability that we can measure a particle in a region $\Lambda \subseteq \mathbb{R}^d$, we have to integrate $|\psi(t,x)|^2$ over Λ,

$$\mathbb{P}\big(X(t) \in \Lambda \mid \psi(t)\big) = \int_\Lambda \mathrm{d}x \; |\psi(t,x)|^2 \,. \tag{2.1.2}$$

If we want to interpret $|\psi|^2$ as *probability* density, then the wave function has to be *normalized*, i. e.

$$\|\psi\|^2 = \int_{\mathbb{R}^d} \mathrm{d}x \; |\psi(x)|^2 = 1.$$

This point of view is called *Born rule*: $|\psi|^2$ could either be a *mass* or *charge density* – or a *probability density*. To settle this, physicists have performed the double slit experiment with an electron source of low flux (cf. Figure 2.1.1). If $|\psi|^2$ were a density, one would see the whole interference pattern building up slowly. Instead, one *measures* "single impacts" of electrons and the result is similar to the data obtained from experiments in statistics (e. g. the Dalton board). Hence, we speak of *particles*.

2.2 The mathematical framework of quantum mechanics

To identify the structures common to all physical theories, let us study quantum mechanics in the abstract. Physical theories usually consist of *three ingredients:*

(1) A notion of *state* which encodes the configuration of the system,

(2) a notion of *observable* which predicts the outcome of measurements, and

(3) a *dynamical equation* which governs how the physical system evolves.

We have to identify the notions of *states*, *observables* and *dynamical equations* in *Schrödinger* and *Heisenberg picture*. Here, Schrödinger and Heisenberg picture refer two equivalent formulations of the dynamics where on the one hand one can evolve states or on the other develop observables in time.

2.2.1 Quantum states

Pure states $\psi \in \mathcal{H}$ are normalized elements of a *complex Hilbert space* $\mathcal{H}$. Moreover, if ψ and $\psi' = \mathrm{e}^{\mathrm{i}\vartheta}\,\psi$ differ only by a total phase, then they represent the *same physical state* — just like one can measure only energy *differences*, only phase *shifts* are accessible to measurements. Consequently, *wave functions themselves are not measurable quantities in experiment* as a matter of principle.

The normalization condition $\|\psi\| = 1$ is key to having a *probabilistic interpretation* of quantum mechanics — the norm square of the wave function is just the total probability.

Another way to think of pure states is as *orthogonal projections*

$$P_\psi\,\varphi := |\psi\rangle\langle\psi|\,\varphi = \langle\psi,\varphi\rangle\,\psi,$$

where the phase ambiguity no longer matters as the orthogonal projections

$$P_\psi\,\varphi = \langle\psi,\varphi\rangle\,\psi = \langle\mathrm{e}^{\mathrm{i}\vartheta}\,\psi,\varphi\rangle\,\mathrm{e}^{\mathrm{i}\vartheta}\,\psi = P_{\psi'}$$

associated to φ and $\varphi' = \mathrm{e}^{\mathrm{i}\vartheta}\,\varphi$ coincide. Here, one can see the elegance of bra-ket notation vs. the notation that is "mathematically proper".

Definition 2.2.1 (Pure quantum state) *A pure quantum state is represented by a normalized element $\psi \in \mathcal{H}$ up to a global phase, i. e. if $\vartheta \in \mathbb{R}$, then ψ and $\mathrm{e}^{\mathrm{i}\vartheta}\,\psi$ represent the same physical state.*

Generalizations of this concept exist in the form of *density operators* or states defined as linear functionals on C^*- or von Neumann algebras.

2.2.2 Quantum observables

Quantities that can be *measured* are represented by selfadjoint (hermitian in physics parlance) operators F on the Hilbert space $\mathcal{H}$ (such as $L^2(\mathbb{R}^d)$ or $\ell^2(\mathbb{Z}^d, \mathbb{C}^N)$), i. e. special linear maps

$$F : \mathcal{D}(F) \subseteq \mathcal{H} \longrightarrow \mathcal{H}.$$

Here, $\mathcal{D}(F)$ is the *domain* of the operator since *typical* observables are not defined for all $\psi \in \mathcal{H}$. *This is not a mathematical subtlety with no physical content*, quite the

contrary: consider the observable energy, typically given by

$$H = \frac{1}{2m}(-\mathrm{i}\hbar\nabla)^2 + V(\hat{x}),$$

then states in the domain

$$\mathcal{D}(H) := \left\{\psi \in L^2(\mathbb{R}^d) \;\middle|\; H\psi \in L^2(\mathbb{R}^d)\right\} \subseteq L^2(\mathbb{R}^d)$$

are those of *finite energy*. For all ψ in the domain of the hamiltonian $\mathcal{D}(H) \subseteq L^2(\mathbb{R}^d)$, the energy expectation value

$$\langle \psi, H\psi \rangle < \infty$$

is bounded. Well-defined observables have domains that are *dense* in $\mathcal{H}$. Similarly, states in the domain $\mathcal{D}(\hat{x}_j)$ of the jth component of the position operator are those that are "localized in a finite region" in the sense of expectation values. Boundary conditions may also enter the definition of the domain: as seen in the example of the momentum operator on $[0, 1]$, different boundary conditions yield different momentum operators.

The set of possible outcomes of measurements of H is its spectrum

$$\sigma(H) := \{z \in \mathbb{C} \;|\; H - z \text{ not invertible}\},$$

namely the set of complex numbers z so that $H - z$ is not invertible. These can include eigenvalues, but also other types of spectra are possible if H is defined on an infinite-dimensional Hilbert space.

The energy observable is just a specific example, but it contains all the ingredients which enter the definition of a quantum observable:

Definition 2.2.2 (Quantum observable) *A quantum observable F is a densely defined, selfadjoint operator on a Hilbert space. The spectrum $\sigma(F)$ is the set of outcomes of measurements.*

Physically, results of measurements are real which is reflected in the selfadjointness of operators (cf. [RS72, Chapter VIII]), $H^* = H$, and one can show that spectra of self-adjoint operators are necessarily subsets of the reals. Typically one "guesses" quantum observables from classical observables: in $d = 3$, the angular momentum operator is

$$L = \hat{x} \times (-\mathrm{i}\hbar\nabla).$$

In the simplest case, one uses Dirac's recipe (replace x by $\hat{x}$ and p by $-\mathrm{i}\hbar\nabla$) on the classical observable angular momentum $L(x, p) = x \times p$. In other words, *many quantum*

observables are obtained as quantizations of classical observables: examples are position, momentum and energy. Moreover, the *interpretation* of, say, $L = \hat{x} \times (-\mathrm{i}\hbar\nabla)$ as angular momentum is taken from classical mechanics.

Quantum mechanics is fundamentally a statistical theory: given a quantum state ψ and a quantum observable F, we can define fundamental notions from statistics such as the *expectation value*

$$\mathbb{E}_\psi(F) := \langle \psi, F\psi \rangle \tag{2.2.1}$$

and the *variance*

$$\mathrm{Var}_\psi(F) := \mathbb{E}_\psi\Big(\big(F - \mathbb{E}_\psi(F)\big)^2 \Big), \tag{2.2.2}$$

which quantifies the deviation from the mean.

Hence, quantum observables, selfadjoint operators on Hilbert spaces, are *bookkeeping devices* that have two components:

(1) The *set of possible outcomes* of measurements, the spectrum $\sigma(F)$, and

(2) the *statistics* contained in the spectral measure, which quantifies how often and under what conditions a possible outcome occurs.

2.2.3 Time evolution

The time evolution is defined through the *Schrödinger equation*,

$$\mathrm{i}\hbar \frac{\partial}{\partial t}\psi(t) = H\psi(t), \qquad \psi(t) \in \mathcal{H},\ \psi(0) = \psi_0,\ \|\psi_0\| = 1. \tag{2.2.3}$$

Alternatively, one can write $\psi(t) = U(t)\psi_0$ with $U(0) = \mathbb{1}_{\mathcal{H}}$. Then, we have

$$\mathrm{i}\hbar \frac{\partial}{\partial t}U(t) = HU(t), \qquad U(0) = \mathbb{1}_{\mathcal{H}}.$$

If H *were* a number, one would immediately use the ansatz

$$U(t) = \mathrm{e}^{-\mathrm{i}\frac{t}{\hbar}H} \tag{2.2.4}$$

as solution to the Schrödinger equation. If H is a selfadjoint operator, this is *still true*, but takes a lot of work to justify (2.2.4) rigorously if the domain of H is not all of $\mathcal{H}$ (the case of unbounded operators, the *generic* case).

As has already been mentioned, we can evolve either states or observables in time and one speaks of the Schrödinger or Heisenberg picture, respectively. In the Schrödinger picture, pure states evolve according to

$$\psi(t) = U(t)\psi_0$$

while observables remain fixed.

As a last point, we mention the conservation of probability: if $\psi(t)$ solves the Schrödinger equation for some selfadjoint H, then we can check at least formally that the time evolution is unitary and thus preserves probability,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\big\|\psi(t)\big\|^2 = \frac{\mathrm{d}}{\mathrm{d}t}\big\langle\psi(t),\psi(t)\big\rangle &= \big\langle \tfrac{1}{\mathrm{i}\hbar}H\psi(t),\psi(t)\big\rangle + \big\langle\psi(t),\tfrac{1}{\mathrm{i}\hbar}H\psi(t)\big\rangle \\ &= \frac{\mathrm{i}}{\hbar}\Big(\big\langle\psi(t),H^*\psi(t)\big\rangle - \big\langle\psi(t),H\psi(t)\big\rangle\Big) \\ &= \frac{\mathrm{i}}{\hbar}\big\langle\psi(t),(H^*-H)\psi(t)\big\rangle = 0.\end{aligned}$$

We see that the condition $H^* = H$ is the key here: selfadjoint operators generate unitary evolution groups. As a matter of fact, there are cases when one *wants* to violate conservation of proability: one has to introduce so-called *optical potentials* which simulate particle creation and annihilation.

The time evolution $\mathrm{e}^{-\mathrm{i}\frac{t}{\hbar}H}$ is not the only unitary group of interest, other commonly used examples are *translations* in position or momentum which are generated by the momentum and position operator, respectively (the order is reversed!), as well as rotations which are generated by the angular momentum operators.

Chapter 3

3 Electromagnetism in Matter

One of the fundamental forces of nature is the electromagnetic interaction between charges. On the macroscopic scale, these interactions can be described by *electromagnetic waves* $\big(\mathbf{E}(t,x),\mathbf{H}(t,x)\big) \in \mathbb{R}^6$, consisting of the electric field $\mathbf{E}(t,x)$ and the magnetic field $\mathbf{H}(t,x)$. These fields are composed of various frequency components

$$\big(\widehat{\mathbf{E}}(\omega),\widehat{\mathbf{H}}(\omega)\big) = \big(\mathcal{F}^{-1}(\mathbf{E},\mathbf{H})\big)(\omega) := \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathrm{d}t\, \mathrm{e}^{+\mathrm{i}t\omega}\,\big(\mathbf{E}(t),\mathbf{H}(t)\big). \tag{3.0.1}$$

Typically, a given wave consists only of frequencies from a given range, e. g. radio waves, microwaves, infrared light, visible light, ultraviolet light, X-rays or γ rays; often, this frequency range is not specified by the frequency f, but the (vacuum) wavelength $\lambda = {}^{c}\!/_{f}$ obtained by dividing the speed of light $c \simeq 3\cdot 10^8\,{}^{m}\!/_{\mathrm{s}}$.

Matter consists of microscopic charges (e. g. atomic cores and electrons), and depending on the wavelength, the interaction with electromagnetic waves can be quite complicated. Depending on the wavelength different types of structures are visible to the electromagnetic waves: X-rays can be used to probe the structure of crystal lattices by interacting with the atomic nuclei, visible light interacts mostly with the valence electrons and so forth.

A more fundamental first principles description would have to include the charges as a dynamical degree of freedom, but this is usually too complicated to be of practical use and often not necessary. So instead, one commonly adopts *effective equations* for the electromagnetic field only. Here, the properties of the medium enter through *constitutive relations*

$$\big(\mathbf{D}(t,x),\mathbf{B}(t,x)\big) = \big(\mathcal{W}(\mathbf{E},\mathbf{H})\big)(t,x) \tag{3.0.2}$$

that connect the so-called *auxiliary fields*, the electric displacement **D** and the magnetic induction **B**. *By their very definition* these quantities stand on the left-hand side of

Maxwell's dynamical equations,

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\mathbf{D}(t)\\ \mathbf{B}(t)\end{pmatrix} = \begin{pmatrix}+\nabla\times\mathbf{H}(t)\\ -\nabla\times\mathbf{E}(t)\end{pmatrix} - \begin{pmatrix}\mathbf{J}(t)\\ 0\end{pmatrix} \tag{3.0.3}$$

and the *constraint equations*,

$$\begin{pmatrix}\nabla\cdot\mathbf{D}(t)\\ \nabla\cdot\mathbf{B}(t)\end{pmatrix} = \begin{pmatrix}\rho(t)\\ 0\end{pmatrix}. \tag{3.0.4}$$

For these equations to be consistent, the current density $J(t)$ and the charge density $\rho(t)$ need to satisfy *local charge conservation*,

$$\partial_t\rho(t) + \nabla\cdot\mathbf{J}(t) = 0. \tag{3.0.5}$$

3.1 Maxwell's equations for linear, lossless dielectrics

Mathematically, choosing a medium amounts to specifying constitutive relations. Many media are to very good approximation linear, i. e.

$$\mathcal{W}\big(a_1\,(\mathbf{E}_1,\mathbf{H}_1) + a_2\,(\mathbf{E}_2,\mathbf{H}_2)\big) = a_1\,\mathcal{W}(\mathbf{E}_1,\mathbf{H}_1) + a_2\,\mathcal{W}(\mathbf{E}_2,\mathbf{H}_2)$$

holds for arbitrary linear combinations. Moreover, when we consider electromagnetic waves from a *narrow frequency* range, we can often *neglect dispersion*. For such time-independent media the constitutive relations take the form

$$\big(\mathbf{D}(t,x),\mathbf{B}(t,x)\big) = \big(\mathcal{W}(\mathbf{E},\mathbf{H})\big)(t,x) = W(x)\,\big(\mathbf{E}(t,x),\mathbf{H}(t,x)\big) \tag{3.1.1}$$

where the 6×6-matrix-valued function

$$W(x) = \begin{pmatrix}\varepsilon(x) & \chi^{EH}(x)\\ \chi^{HE}(x) & \mu(x)\end{pmatrix}$$

is sometimes called *material weights*. It consists of 3×3 blocks, the *electric permittivity* ε, the *magnetic permeability* μ, and the bianisotropies χ^{HE} and χ^{EH}.

Simple, lossless dielectrics are characterized by the following

Assumption 3.1.1 (Material weights for a simple, lossless dielectric) *We assume that the material weights $W\in L^\infty\big(\mathbb{R}^3,\mathrm{Mat}_{\mathbb{C}}(6)\big)$ satisfy:*

(a) The medium is lossless, i. e. the function $x\mapsto W(x) = W(x)^*$ takes values in the hermitian 6×6 *matrices.*

(b) The material is time-reversal symmetric*, i. e. it has no* bianisotropy*, i. e.* $\chi^{HE} = 0 = \chi^{EH}$*, and the material weights are* real*, i. e.* $W(x) = \overline{W(x)}$.

(c) The medium is a positive index medium*, i. e. there exist constants* $0 < C_{\text{min}} \leq C_{\text{max}} < \infty$ *so that the 6 (necessarily real) eigenvalues* $\{w_1(x), \dots, w_6(x)\}$ *of* $W(x)$ *satisfy the inequality*

$$0 < C_{\text{min}} \leq w_1(x), \dots, w_6(x) \leq C_{\text{max}} < \infty.$$

For simple dielectrics, the auxiliary fields can be eliminated from the equations, and Maxwell's equations then read

$$\begin{pmatrix} \varepsilon(x) & 0 \\ 0 & \mu(x) \end{pmatrix} \frac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} \mathbf{E}(t) \\ \mathbf{H}(t) \end{pmatrix} = \begin{pmatrix} +\nabla \times \mathbf{H}(t) \\ -\nabla \times \mathbf{E}(t) \end{pmatrix} - \begin{pmatrix} \mathbf{J}(t) \\ 0 \end{pmatrix}, \tag{3.1.2a}$$

$$\begin{pmatrix} \nabla \cdot \varepsilon \mathbf{E}(t) \\ \nabla \cdot \mu \mathbf{H}(t) \end{pmatrix} = \begin{pmatrix} \rho(t) \\ 0 \end{pmatrix}. \tag{3.1.2b}$$

It is these equations we will study in the remainder of the lecture, although our arguments can certainly be generalized to include more general media.

One last note on the assumption that W is L^∞ rather than continuous. This is quite natural as many media are obtained by alternating two or more different materials, e. g. a dielectric and air. On the relevant length scale of the waves, the material weights change discontinuously from, say, $(\varepsilon_{\text{dielectric}}, \mu_{\text{dielectric}})$ to $(\varepsilon_{\text{air}}, \mu_{\text{air}})$.

3.2 The notion of states, observables and dynamics

Now that we have set the stage by introducing the relevant mathematical equations, let us turn our focus on how to interpret them physically. As with quantum mechanics, let us turn our focus on how to interpret them physically. As with quantum mechanics, we need to identify what states are, how measurable quantities are represented mathematically and summarize what the dynamical equations are.

While much of what we say is much more general, we shall always suppose the material weights are linear, non-dispersive and satisfy Assumption 3.1.1.

3.2.1 States

While electromagnetism allows for much more variety as far as the space of electromagnetic field is concerned, one very common assumption is that we only study states of *finite field energy*

$$\mathcal{E}(\mathbf{E}, \mathbf{H}) := \int_{\mathbb{R}^3} \mathrm{d}x \begin{pmatrix} \mathbf{E}(x) \\ \mathbf{H}(x) \end{pmatrix} \cdot W(x) \begin{pmatrix} \mathbf{E}(x) \\ \mathbf{H}(x) \end{pmatrix} \overset{!}{<} \infty,$$

where $\Phi(x) \cdot \Psi(x) := \sum_{j=1}^{6} \phi_j(x)\,\psi_j(x)$ is the usual dot product on $\mathbb{R}^6$. Put another way, we are interested in the *real* Hilbert space

$$L^2_W(\mathbb{R}^3, \mathbb{R}^6) = \Big\{(\mathbf{E}, \mathbf{H}) : \mathbb{R}^3 \longrightarrow \mathbb{R}^6 \;\Big|\; \mathcal{E}(\mathbf{E}, \mathbf{H}) < \infty\Big\}/\sim$$

endowed with the *energy scalar product*

$$\big\langle (\mathbf{E}, \mathbf{H}), (\mathbf{E}', \mathbf{H}')\big\rangle_W := \int_{\mathbb{R}^3} \mathrm{d}x \begin{pmatrix}\mathbf{E}(x)\\ \mathbf{H}(x)\end{pmatrix} \cdot W(x) \begin{pmatrix}\mathbf{E}'(x)\\ \mathbf{H}'(x)\end{pmatrix}. \tag{3.2.1}$$

Note that the weighted Hilbert space $L^2_W(\mathbb{R}^3, \mathbb{R}^6)$ coincides with $L^2(\mathbb{R}^3, \mathbb{R}^6)$ as *Banach* spaces (i. e. as complete, normed vector spaces), meaning that they agree as vector spaces and the two norms are equivalent.

Definition 3.2.1 (Electromagnetic states) *Electromagnetic states are electromagnetic fields* $(\mathbf{E}, \mathbf{H})$ *with finite field energy that satisfy the constraint* (3.1.2b)*, i. e. in the absence of charges* $\rho = 0$ *the field is an element of*

$$\mathcal{H}_{\mathbb{R}} := \Big\{(\mathbf{E}, \mathbf{H}) \in L^2_W(\mathbb{R}^3, \mathbb{R}^6) \;\Big|\; \nabla \cdot \varepsilon\mathbf{E} = 0 = \nabla \cdot \mu\mathbf{H}\Big\}.$$

When the waves are constrained to a domain $\Omega \subseteq \mathbb{R}^3$, the relevant vector space is $L^2_W(\Omega, \mathbb{R}^6)$ instead and we need to choose boundary conditions on the surfaces (such as those for a *perfect electric conductor*). All of our arguments can be adapted.

3.2.2 Observables

Measurable quantities are those that can be computed from the electromagnetic field distribution, i. e. we put forth the following

Definition 3.2.2 (Electromagnetic observables) *In classical electromagnetism observables are functionals of the fields,*

$$\mathcal{F} : \mathcal{D}(\mathcal{F}) \subseteq \mathcal{H}_{\mathbb{R}} \longrightarrow \mathbb{R},\ (\mathbf{E}, \mathbf{H}) \mapsto \mathcal{F}(\mathbf{E}, \mathbf{H}).$$

The simplest example are components of the electric and magnetic fields themselves,

$$\delta^{\mathbf{E}}_{j,x}(\mathbf{E}, \mathbf{H}) := E_j(x),$$
$$\delta^{\mathbf{H}}_{j,x}(\mathbf{E}, \mathbf{H}) := H_j(x).$$

This is in contrast to quantum mechanics where as a matter of principle, the quantum wave function is not directly accessible to measurements!

Many electromagnetic observables are *quadratic* in the fields, including the *local energy density* at the point x,

$$e_x(\mathbf{E},\mathbf{H}) := \frac{1}{2}\begin{pmatrix}\mathbf{E}(x)\\ \mathbf{H}(x)\end{pmatrix}\cdot W(x)\begin{pmatrix}\mathbf{E}(x)\\ \mathbf{H}(x)\end{pmatrix},$$

where W are the material weights from the constitutive relations, or the total field energy in the volume $V \subseteq \mathbb{R}^3$,

$$\begin{aligned}\mathcal{E}_V(\mathbf{E},\mathbf{H}) &:= \frac{1}{2}\int_V \mathrm{d}x \begin{pmatrix}\mathbf{E}(x)\\ \mathbf{H}(x)\end{pmatrix}\cdot W(x)\begin{pmatrix}\mathbf{E}(x)\\ \mathbf{H}(x)\end{pmatrix}\\ &= \int_V \mathrm{d}x\, e_x(\mathbf{E},\mathbf{H}) = \frac{1}{2}\big\langle 1_V(\mathbf{E},\mathbf{H}),\, 1_V(\mathbf{E},\mathbf{H})\big\rangle_W^2,\end{aligned}$$

where $1_V(x) = 1$ when $x \in V$ and 0 otherwise.

Connected to the energy density is the *Poynting vector*

$$\mathcal{S}_x(\mathbf{E},\mathbf{H}) := \mathbf{E}\times\mathbf{H},$$

since this is the current that enters the *field energy conservation law* for linear, lossless, dispersionless media,

$$\partial_t e_x\big(\mathbf{E}(t),\mathbf{H}(t)\big) + \nabla\cdot\mathcal{S}_x\big(\mathbf{E}(t),\mathbf{H}(t)\big) = -\mathbf{J}(t)\cdot\mathbf{E}(t). \tag{3.2.2}$$

3.2.3 Dynamics

The dynamical equations are (3.1.2a) and incorporate the constitutive relations (3.1.1), which describe the properties of the medium.

Charge conservation 3.2.2 and the constraint equation (3.1.2b) enter as consistency conditions and are needed to uniquely fix the solution to the dynamical equations. We will see why in Section 3.3.2.

3.3 Fundamental properties

Lastly, let us discuss some fundamental properties of Maxwell's equations.

3.3.1 Some fundamental discrete symmetries

Maxwell's equations have a lot of discrete and continuous symmetries, and media will *selectively break symmetries of the vacuum Maxwell equations*. Wigner [Wig39] derived

the vacuum Maxwell equations are *the* relativistic equations for a massless spin-1 particle, and so we have a complete understanding of continuous symmetries. These then lead to various conserved quantities such as field energy and total angular momentum.

Then there are *discrete symmetries* that relate components of **E** and **H**,

$$\begin{aligned} T_1^{\mathbb{R}} &= \sigma_1 \otimes \mathbb{1} : (\mathbf{E}, \mathbf{H}) \mapsto (\mathbf{H}, \mathbf{E}), \\ T_2^{\mathbb{R}} &= \mathrm{i}\sigma_2 \otimes \mathbb{1} : (\mathbf{E}, \mathbf{H}) \mapsto (\mathbf{H}, -\mathbf{E}), \\ T_3^{\mathbb{R}} &= \sigma_3 \otimes \mathbb{1} : (\mathbf{E}, \mathbf{H}) \mapsto (\mathbf{E}, -\mathbf{H}), \end{aligned}$$

which can be conveniently written in terms of the usual Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ +\mathrm{i} & 0 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The operator on right-hand side of Maxwell's dynamical equation (3.1.2a) can be rewritten using the Pauli matrix σ_2,

$$\begin{pmatrix} +\nabla \times \mathbf{H} \\ -\nabla \times \mathbf{E} \end{pmatrix} = \begin{pmatrix} 0 & +\nabla^{\times} \\ -\nabla^{\times} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} = \mathrm{i}\sigma_2 \otimes \nabla^{\times} \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix}.$$

Hence, in the absence of current densities, by the (anti)commutation relations of Pauli matrices the operator on the right-hand side *commutes* with $T_2^{\mathbb{R}}$ and *anticommutes* with $T_{1,3}^{\mathbb{R}}$. So for vacuum where $W = \mathbb{1}$ the two anticommuting symmetries $T_{1,3}^{\mathbb{R}}$ reverse the arrow of time, whereas the commuting symmetry $T_2^{\mathbb{R}}$ leads to a conserved quantity — *helicity*. We emphasize that these symmetry operations also leave the constraint equation (3.1.2b) untouched when $W = \mathbb{1}$, since the two conditions

$$\mathrm{Div}\,(\mathbf{E}, \mathbf{H}) = \big(\nabla \cdot \mathbf{E}\,,\, \nabla \cdot \mathbf{H}\big) = 0 \iff \mathrm{Div}\, T_j^{\mathbb{R}}(\mathbf{E}, \mathbf{H}) = 0$$

are equivalent.

When $W \neq \mathbb{1}$ some or all of these three symmetries are broken. Here, for media with real material weights the deciding factor is whether W commutes with the $T_j^{\mathbb{R}}$,

$$T_j^{\mathbb{R}}\, W \stackrel{?}{=} W\, T_j^{\mathbb{R}}. \tag{3.3.1}$$

Provided this condition is satisfied, then the dynamical equations either possess a proper symmetry ($j = 2$) or a time-reversal-type symmetry ($j = 1, 3$),

$$\begin{aligned} W\, \frac{\partial}{\partial t}\, T_j^{\mathbb{R}} \begin{pmatrix} \mathbf{E}(t) \\ \mathbf{H}(t) \end{pmatrix} &= T_j^{\mathbb{R}}\, W\, \frac{\partial}{\partial t} \begin{pmatrix} \mathbf{E}(t) \\ \mathbf{H}(t) \end{pmatrix} = T_j^{\mathbb{R}} \begin{pmatrix} 0 & +\nabla^{\times} \\ -\nabla^{\times} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{E}(t) \\ \mathbf{H}(t) \end{pmatrix} \\ &= (-1)^j \begin{pmatrix} 0 & +\nabla^{\times} \\ -\nabla^{\times} & 0 \end{pmatrix} T_j^{\mathbb{R}} \begin{pmatrix} \mathbf{E}(t) \\ \mathbf{H}(t) \end{pmatrix}. \end{aligned}$$

We can again show by explicit computation that in the absence of sources the constraints spelled out by equation (3.1.2b) are preserved by $T_j^{\mathbb{R}}$.

These symmetries are well-known in the physics literature — $T_2^{\mathbb{R}}$ is usually referred to as *dual symmetry* and $T_3^{\mathbb{R}}$ as *time-reversal*.

3.3.2 The Helmholtz decomposition and its relation to Maxwell's equations

To better understand the nature of Maxwell's equations (3.1.2) a useful concept is derived from the *Helmholtz decomposition of vector fields*

$$\mathbf{F} = \mathbf{F}_\perp + \mathbf{F}_\parallel = \nabla \times \mathbf{A} + \nabla\chi \in L^2(\mathbb{R}^3, \mathbb{R}^3) \tag{3.3.2}$$

into the curl of a vector potential and the gradient of a scalar potential. Equivalently, we could define the *transverse part* $\mathbf{F}_\perp$ and the *longitudinal part* $\mathbf{F}_\parallel$ by demanding

$$\begin{aligned} \nabla \cdot \mathbf{F}_\perp &= 0, \\ \nabla \times \mathbf{F}_\parallel &= 0. \end{aligned}$$

We can succinctly write this as

$$\mathbf{F}_\perp \in \operatorname{ran}\big(\nabla\times\big) \subseteq \ker\big(\nabla\cdot\big), \tag{3.3.3a}$$

$$\mathbf{F}_\parallel \in \operatorname{ran}\nabla \subseteq \ker\big(\nabla\times\big), \tag{3.3.3b}$$

because on $L^2(\mathbb{R}^3, \mathbb{R}^3)$ (and other spaces like $\mathcal{C}_c^\infty(\mathbb{R}^3, \mathbb{R}^3)$, the vector space of smooth vector fields with compact support) the Laplace equation

$$\Delta\Phi = 0$$

has no solution apart from the trivial solution $\Phi = 0$.

Longitudinal and transversal fields are perpendicular to each other,

$$\big\langle \mathbf{F}_\perp, \mathbf{F}_\parallel \big\rangle = \int_{\mathbb{R}^3} \mathrm{d}x\, \mathbf{F}_\perp(x) \cdot \mathbf{F}_\parallel(x),$$

so that they evolve independently of one another.

Theorem 3.3.1 (Helmholtz decomposition on $L^2(\mathbb{R}^3, \mathbb{C}^3)$) *(1) The Helmholtz decomposition* (3.3.2) *of vector fields exists and is unique.*

(2) In equations (3.3.3) *we can replace $\subseteq$ with $=$, i. e.*

$$\begin{aligned} \operatorname{ran}\big(\nabla\times\big) &= \ker\big(\nabla\cdot\big), \\ \operatorname{ran}\nabla &= \ker\big(\nabla\times\big). \end{aligned}$$

While the decomposition (3.3.2) is suitable for electromagnetic waves in vacuum where the material weights $W = \mathbb{1}$ are the identity map, there exists a Helmholtz decomposition *adapted to certain electromagnetic media*. To be mathematically precise, we continue to suppose that the material weights satisfy Assumption 3.1.1. Then any electromagnetic field

$$(\mathbf{E}, \mathbf{H}) = \big(\mathbf{E}_\perp, \mathbf{H}_\perp\big) + \big(\mathbf{E}_\parallel, \mathbf{H}_\parallel\big) \tag{3.3.4}$$

can be split into its *transversal part* $\big(\mathbf{E}_\perp, \mathbf{H}_\perp\big)$ defined through the divergence-free condition involving the material weights W,

$$\operatorname{Div} W \big(\mathbf{E}_\perp, \mathbf{H}_\perp\big) := \big(\nabla \cdot \varepsilon \mathbf{E}_\perp \,,\, \nabla \cdot \mu \mathbf{H}_\perp\big) = 0,$$

and a *longitudinal gradient field*

$$\big(\mathbf{E}_\parallel, \mathbf{H}_\parallel\big) = \big(\nabla \varphi^E, \nabla \varphi^H\big) =: \operatorname{Grad}(\varphi^E, \varphi^H).$$

There is a very elegant connection between the space of longitudinal gradient fields

$$\mathcal{G} := \operatorname{ran} \operatorname{Grad} = \Big\{ \operatorname{Grad}(\varphi^E, \varphi^H) : \mathbb{R}^3 \longrightarrow \mathbb{C}^6 \;\Big|\; (\varphi^E, \varphi^H) \in L^2_{\mathrm{loc}}(\mathbb{R}^3, \mathbb{C}^2) \Big\} \subset \mathcal{H}_{\mathbb{R}}$$

and the space of transversal fields

$$\mathcal{J}_W := \ker \operatorname{Div} W = \Big\{ \big(\mathbf{E}_\perp, \mathbf{H}_\perp\big) \in \mathcal{H}_{\mathbb{R}} \;\Big|\; \operatorname{Div} W \big(\mathbf{E}_\perp, \mathbf{H}_\perp\big) = 0 \Big\}$$

through the energy scalar product (3.2.1), namely they are orthogonal complements to one another,

$$\mathcal{J}_W = \mathcal{G}^{\perp_W} := \Big\{ (\mathbf{E}, \mathbf{H}) \in \mathcal{H}_{\mathbb{R}} \;\Big|\; \big\langle (\mathbf{E}, \mathbf{H}), (\mathbf{E}_\parallel, \mathbf{H}_\parallel) \big\rangle_W = 0 \;\; \forall (\mathbf{E}_\parallel, \mathbf{H}_\parallel) \in \mathcal{G} \Big\}.$$

To see this, we exploit the selfadjointness of the weights W and use partial integration to convert Grad into Div,

$$\begin{aligned} \big\langle (\mathbf{E}, \mathbf{H}) \,,\, \big(\mathbf{E}_\parallel, \mathbf{H}_\parallel\big) \big\rangle_W &= \big\langle (\mathbf{E}, \mathbf{H}) \,,\, \big(\mathbf{E}_\parallel, \mathbf{H}_\parallel\big) \big\rangle_W = \big\langle W \, (\mathbf{E}, \mathbf{H}) \,,\, \big(\nabla \varphi^E, \nabla \varphi^H\big) \big\rangle_{L^2(\mathbb{R}^3, \mathbb{R}^6)} \\ &= -\big\langle \operatorname{Div} W \, (\mathbf{E}, \mathbf{H}) \,,\, (\varphi^E, \varphi^H) \big\rangle_{L^2(\mathbb{R}^3, \mathbb{R}^2)} \stackrel{!}{=} 0. \end{aligned}$$

Since this holds for all (φ^E, φ^H), the first argument of the scalar product has to vanish. Consequently, we have the *Helmholtz decomposition adapted to the electromagnetic medium*, which splits

$$L^2_W(\mathbb{R}^3, \mathbb{C}^6) = \mathcal{J}_W \oplus \mathcal{G}. \tag{3.3.5}$$

Theorem 3.3.2 (Helmholtz decomposition adapted to dielectric media) *Suppose the material weights satisfy Assumption 3.1.1. Then the weighted Hilbert space*

$$L^2_W(\mathbb{R}^3, \mathbb{C}^6) = \mathcal{J}_W \oplus \mathcal{G}$$

admits a splitting into $\langle \cdot, \cdot \rangle_W$-orthogonal subspaces composed of divergence-free fields

$$\mathcal{J}_W = \ker\big(\operatorname{Div} W\big) = \operatorname{ran} M^{\mathrm{aux}}$$

and gradient fields

$$\mathcal{G} = \ker M^{\mathrm{aux}} = \operatorname{ran}(\nabla, \nabla).$$

Bibliography

[BS87] M. S. Birman and M. Z. Solomyak. *L_2-Theory of the Maxwell operator in arbitrary domains*. Uspekhi Mat. Nauk **42**, 61–76, 1987.

[Cro08] B. Crowell. *The Modern Revolution in Physics*. Light and Matter, 2008.

[Fol89] G. B. Folland. *Harmonic Analysis on Phase Space*. Princeton University Press, 1989.

[Jac98] J. D. Jackson. *Classical Electrodynamics*. Wiley, 1998.

[Lei10] M. Lein. *Weyl Quantization and Semiclassics*. Technische Universität München. 2010.

[Moy49] J. E. Moyal. *Quantum mechanics as a statistical theory*. Proc. Cambridge Phil. Soc **45**, 211, 1949.

[RS72] M. Reed and B. Simon. *Methods of Mathematical Physics I: Functional Analysis*. Academic Press, 1972.

[Tes09] G. Teschl. *Mathematical Methods in Quantum Mechanics*. Vol. 99. Graduate Studies in Mathematics. American Mathematical Society, 2009.

[Wey27] H. Weyl. *Quantenmechanik und Gruppentheorie*. Zeitschrift für Physik **46**, 1–46, 1927.

[Wig32] E. Wigner. *On the quantum correction for thermodynamic equilibrium*. Phys. Rev. **40**, 749–759, 1932.

[Wig39] E. P. Wigner. *On Unitary Representations of the Inhomogeneous Lorentz Group*. Annals of Mathematics **40**, 149–204, 1939.