Math 321/322/323: Applied Mathematics Notes — Quantum Physics module

Matt Visser
School of Mathematics, Statistics, and Operations Research
Victoria University of Wellington,
New Zealand.
E-mail: matt.visser@msor.vuw.ac.nz
URL: http://www.mcs.vuw.ac.nz/~visser

Version of 26 February 2013; LaTeX-ed February 27, 2013

Warning:

These notes will basically be the textbook for this module.

This is a reading course, so these notes, and various web resources, should be your primary source of information.

I may also assign some specific supplementary reading.

There are still a few rough edges in these notes.

If you find errors, typos, and/or obscurities, please let me know.

Contents

Chapter 1

Overview

This module investigates the *mathematical structure* of quantum physics. So we will be very much emphasizing mathematical features of the theory, and will be particularly interested in seeing how much can be deduced purely by mathematical reasoning without any (or with very little) physics input.

To set the stage, remember the two faces of *any* physical theory:

- Physical theories have a *mathematical structure* that can be investigated purely by logic. This mathematical structure exists independently of whether or not the theory has anything to do with the "real world", and this mathematical structure can be studied and analyzed by pure logic without recourse to experiment.
- Physical theories are *useful* only if they closely reflect what actually happens in the "real world" — it is at that stage of the process that you need to worry about experimental input.

This is important because sometimes the physics results you get tell you as much (or more) about the general mathematical framework you are working in, as they do about the specific physical theory you thought you were investigating. I'll illustrate this point with a few specific examples...

The main topics to be touched on in this module will be:

- the Heisenberg uncertainty principle;
- tunnelling;
- one-dimensional scattering.
- the S-matrix.

In each of these cases I hope to give you some deep insight into how very general and fundamental mathematical tools can be brought to bear on specific physics problems.

Chapter 2

Introduction to quantum physics

- These comments closely parallel the introductory comments I made in the notes on special relativity — *with a few key differences*!
- The comments in this particular chapter are for your cultural edification — I will not set any assignment problems based on this chapter.

First, while the special theory of relativity is the theory for which Albert Einstein is most famous in the public mind, to a physicist, the special relativity is merely *one* of Einstein's many contributions to physics. Einstein, together with Planck, Schrödinger, Heisenberg, *et al.*, were all involved in setting up the very foundations of quantum physics.

2.1 Quantum mechanics

1. Non-relativistic quantum mechanics generalizes Newtonian mechanics to "small" systems. Non-relativistic Quantum mechanics gives a good description of individual atoms and molecules.
2. However, non-relativistic quantum mechanics is limited to small velocities (small compared to lightspeed)[1] and weak gravitational fields.[2]

[1] For instance, non-relativistic quantum mechanics is already not quite good enough for dealing with the innermost electrons of heavy elements such as Uranium. For the outermost electrons, on which all of chemistry is based, non-relativistic quantum mechanics is fine.

[2] Non-relativistic quantum mechanics plus Newton's gravity is quite good enough for almost all terrestrial purposes. It's in the vicinity of black holes that things get dicey — see for instance Stephen Hawking's ideas on black hole evaporation via relativistic quantum effects.

2.2 Quantum technology

During the second half of the twentieth century the major advances in human technological development have been due to our increasing ability to manipulate the effects inherent in the quantum theory. The fundamental equations of the (nonrelativistic) quantum theory were written down in 1925 by Schrödinger, Heisenberg, Bohr, Jordan, and others. Eighty-five years later the Schrödinger equation continues to provide us with hints and ideas for manipulating matter in its quantum aspects. The nonrelativistic Schrödinger equation is intrinsic to understanding the structure of atoms, chemical compounds, and the solid, liquid, gaseous, and plasma states. From a technological perspective the Schrödinger equation is the underpinning for the solid-state diode and the solid-state transistor. These two simple devices are based on quantum physics, but they are designed to produce results at the classical level, where they act as one-way gates and switches. The importance of the transistor is that it gives us technological access to simple, rugged, *small* switches, and lets us replace the bulkier and hotter thermionic valves found in older radios and televisions. There are only so many thermionic valves that can be packed into a small space before they begin to fry each other with the waste heat they generate.

In contrast, transistors can be made smaller, packed together tighter, and generate much less heat (though getting rid of waste heat is still a significant problem). It is quantum physics that underlies modern solid-state electronics and permits us to build the memory chips, central processing units, and multitude of special-purpose integrated circuit chips that underlie modern computer technology (and modern miniaturized televisions, radios, CD players, and other household electronics). Again, even though the basic Schrödinger equation was written down in 1925 it is not a simple direct route from that equation to integrated circuit technology — as generations of physicists and engineers will attest. On the other hand, as indicated above, a rather good case can be made for nonrelativistic quantum mechanics as the *primum movil* behind solid-state electronics.

One thing that should be emphasized in this situation is that transistors are still *classical* devices that function *using* quantum physics, but at a fundamental level transistors do not themselves exhibit such quantum weirdness as the Schrödinger's cat effect. (Is there something rotten in the eigenstate of Schrödinger's cat?) There is currently (2013) a revolution underway using nanoscale and mesoscale technology to try making intrinsically quantum switches, with a view to ultimately making intrinsically quantum computers — this *may* cause a major revolution in computer technology over the next few decades but it is too soon to tell with certainty. Similarly there are efforts underway to apply quantum physics directly to cryptography (code-making and code-breaking), and to the construction of eavesdropper-proof communication links. I mention these possibilities so that you can appreciate that even nonrelativistic quantum physics (let alone relativistic quantum field theory) is a field that has not yet yielded all is secrets, and that there is still a vast potential for technological development hiding in the innocent-looking Schrödinger equation — a potential that many people are eagerly seeking to exploit.

2.3 Quantum field theory

Quantum field theory (QFT), which I *won't* deal with in this module, is the proper tool for describing relativistic quantum physics — the physics of the very small at arbitrary speeds up to the speed of light. As such, quantum field theory is the appropriate description underlying lasers,[3] nuclear physics, "atom smashers", and radioactivity in all its forms.

1. Quantum field theory generalizes both non-relativistic quantum mechanics and special relativity, the physics of the very small and the very fast. This class of theories is quite adequate for doing all of particle physics (the physics of elementary particles such as quarks, leptons, gauge bosons, Higgs, *etc.*.)
2. Quantum field theory, in its current incarnation, is incapable of dealing with quantized gravitational fields in a clean manner.
3. If you ever come up with a really good theory of quantum gravity, publish.

   (Physicists and mathematicians have been struggling with this problem for fifty years now, and progress is fragmentary and haphazard. Look up, for instance, "string models", "loop variables", "spin networks", "spin foams", "lattice quantum gravity", "causal dynamical triangulations", and "Hořava gravity" for more than you ever really wanted to know...)

2.4 Classifying the theories of physics: Superb, Useful, and Tentative

Professor Roger Penrose (Oxford University) has come up with a very interesting classification of scientific theories:

1. **Superb**: The **Superb** theories will always be with us. They are rock solid descriptions of fundamental reality that are simply too useful to discard, and will if nothing else be retained for all time as suitable *limiting cases* of any more fundamental theory that may come along.
   For example, you will never abandon Newtonian physics completely, it's simply too good an approximation to reality in the regime where most engineers are working.

[3] Actually, trying to decide where to classify lasers is a little tricky. The photons in the laser beam are of course moving at the speed of light, which screams "special relativity". But there are approximations where you can treat the atoms that emit the photons non-relativistically, and agree not to ask too many nasty questions about the photons themselves. Plus for good measure, there is such a thing as non-relativistic quantum field theory, suitable for studying condensed matter physics, so the dividing lines between theories are not always as sharp as I have made them appear in this account.

Similarly, Maxwell's electromagnetism is a fantastically good approximation, which is why we still use it even if quantum electrodynamics is "more fundamental".

Warning: Strictly speaking, the word "theory" should really be reserved for theories belonging to the **Superb** classification. Similarly, any talk about the "laws" of physics, should really be reserved for items in the **Superb** category. In internal discussions among experts, these linguistic rules are often violated because we know enough background information to be able to fill in the blanks for ourselves. Unfortunately this can lead to comprehension difficulties when communicating across discipline boundaries, or to the general public.

2. **Useful**: The **Useful** theories are provisional in nature; while it is clear that they capture some aspects of physical reality, there will be nagging details and reasons to suspect they are not the last word. **Useful** theories will almost certainly last a few centuries, and be relatively stable on the one to two decade timescale (and perhaps longer).

   Warning: Strictly speaking the **Useful** theories should not be called theories, but rather "standard models". (As in "standard model of particle physics", or "standard model of cosmology".) For standard models you should really not talk about "laws", but rather use constructions such as "rules of thumb", "empirical relationships", "phenomenologically determined parameters" and the like.

3. **Tentative**: The **Tentative** theories represent "work in progress"; they are likely to suffer severe revision as our ideas (and experimental input) gets better.

   Warning: Strictly speaking you should *never* use the word theory to describe something in the **Tentative** category; words like "model", "concept", or "idea" are more appropriate. Sometimes you should go so far as to use the phrase "toy model" or the word "conjecture". (For instance grand unified theories, which are definitely in the **Tentative** category, should really be called grand unified models, but the acronym GUT sounds so much better than GUM, that GUT has been universally adopted.)[4]

This is not to say that *everything* is in constant flux: the **Superb** theories have pretty much settled into their final forms and we don't expect any really revolutionary surprises there (though there may still be interesting refinements to be extracted). The **Useful** "theories" (standard models) tend to evolve slowly over the lifetimes of most professional physicists, while only the **Tentative** "theories" (conjectures) are in a state of constant upheaval.

[4] One particle theorist has been known to put it this way: Current "Grand Unified Theories" are neither grand, nor unified, nor even theories...

2.5 Superb means "Never to be Discarded"

Quantum physics has been very well-tested, both experimentally and mathematically, it is definitely **Superb** in the sense of Roger Penrose. The mathematical structure of quantum physics has been investigated in elaborate detail (sometimes in baroque detail), and its experimental consequences have been extensively checked (in the parameter ranges where we expect this theory to be valid, *and where we have appropriate technology*).

It is critically important to realize that in a certain sense quantum physics will *never* be discarded — it is simply too useful in the range where we *know* it works. At the very worst, this theory will be superseded by some more complicated "master theory" that must effectively reduce to standard quantum physics in appropriate limits.

2.6 Textual analysis: A warning

Before we go any further, I feel that an important warning is in order: *Never attempt a comparative textual analysis of popular-level physics books* (including these notes) — the results are almost certain to be abject nonsense. By way of example, any serious literary study of Don Quixote will require you to learn the Spanish language — working only from an English language translation is never going to provide deep insight into the work. Similarly, popular level books on physics are inherently limited by the translation difficulties of adopting a natural language at the cost of excising the underlying mathematics. Do not take pretty pictures and verbal descriptions too seriously — they can be dangerously misleading — natural language is a subtle and shifting foundation on which to attempt to build physical theory.

For an example of the problems that can arise purely at the level of English language usage I need merely point out the confusion attendant on use of the word **"paradox"**. In English this word has two primary meanings. Either:

1. an *apparent contradiction* in a logically consistent theory, or
2. a real *logical inconsistency* in a truly inconsistent theory.

(There are also a few more archaic meanings that are not currently of relevance.) Worse, the most likely meaning has shifted over the past few decades.

Problems arise for instance, in the discussion of the famous "twin paradox" of special relativity. Einstein and his contemporaries used the word in the sense of an "apparent inconsistency" (what we might now call a "pseudo-paradox") and certainly did not claim or imply that special relativity was internally inconsistent. Unfortunately, many commentators have fixated attention on the word "paradox" and automatically assumed that the meaning of "real logical inconsistency" was intended, leading to discussions whose results are both pathetic and predictable. [At least half of the yelling and screaming surrounding the issue of the twin paradox in special relativity can be tracked down to not having a good dictionary on hand.] And this is just a simple ambiguity within the English language itself — this is not even a translation difficulty from mathematics to English. (To add to the confusion, don't forget that Einstein's native language was German, not English, and that his early works were written in German, not English.)

Another famous, well infamous, example of the troubles that can be caused by outright mis-translation between natural languages is that of the infamous Martian **"canali"**. Now "canali" is a perfectly good Italian word that has the English meaning of "channels" (naturally occurring channels, with no implication of human or alien intervention). Unfortunately US newspapers of the late 1800's [deliberately] mis-translated this into English as "canals" (implying they were constructed by someone or something). So much for the canals of Mars; they were never more than endless speculation heaped upon a simple mis-translation (and a few highly ambiguous and noisy ground-based visual observations of some things that looked vaguely like channels). Still, John Carter and Barsoom will continue to live on in legend.

Yet another example of places where problems commonly arise is in the discussion of the **"Einstein elevator"**. This is a gedanken-experiment (thought-experiment) devised by Einstein that argues for the complete equivalence between acceleration and an applied gravitational field. (This is the Einstein Equivalence Principle, one of the main principles underlying Einstein gravity, about which I will have more to say next year in Math 464 [differential geometry] and Math 465 [general relativity and cosmology].) More precisely, the Einstein elevator gedanken-experiment argues for the complete equivalence between acceleration and a *homogeneous* gravitational field.

Now all real gravitational fields are inhomogeneous, so the result of the Einstein elevator gedanken-experiment should really be phrased as: *"in any real gravitational field, if one has an elevator that is sufficiently small that inhomogeneities in the gravitational field can be safely ignored, then a person inside the elevator cannot tell the difference between gravity and acceleration"*

This is often shortened for convenience to: *"a person inside an elevator cannot tell the difference between gravity and acceleration"*. Unfortunately I have then seen people who take this shortened version of the Einstein Equivalence Principle too literally. If you take the short version as *the one and only* definition of the Equivalence Principle, and then observe that real gravitational fields are inhomogeneous, than you can *mistakenly* conclude the existence of an internal inconsistency in general relativity. [This mistake is unfortunately rather common.] Of course, what you have really deduced is that the shortened version of the Equivalence Principle is not quite precise enough — going to the long version of the Equivalence Principle removes the problem.

This all came about because in the interests of clarity it is sometimes appropriate to delete some of the qualifying phrases that would otherwise make a popular description or an introductory textbook completely unwieldy and impenetrable — in fact, in the interests of getting any coherent message across I shall occasionally have to resort to such trimming myself. But the reader should be warned that some simplification along these lines is inevitable — and if by determined textual analysis the reader discovers a logical paradox, the paradox is almost certainly a translation difficulty and not a part of the underlying physics. I trust that forewarned is forearmed.

2.7 Filtering out the nonsense

Because the theories and concepts that I am talking about in this course are so far beyond the pale of everyday experience, I think that it would be useful for the student if I were to provide some rules of thumb for filtering out the more extreme crackpot nonsense. (It is unfortunately a truism that nothing attracts the crackpots quite like the words "Einstein" and "relativity", it's like waving a red flag in front of a bull. The word "quantum" is also rather good at attracting the nutters.) Now it is actually rather difficult to give hard and fast rules for detecting crackpot nonsense. Certainly any practitioner in the field can look at a specific document and within sixty seconds can come to a snap decision. Many of rules used in coming to such a conclusion are entirely heuristic and impossible to formalize in all generality. Fortunately however, a certain subset of the rules used by practicing physicists can be more or less formalized: these are the rules associated with the internal consistency of physical theories.

2.7.1 The two faces of physical theory

It is extremely important to realize that physical theories have two main attributes that are logically distinct from one another. Physical theories must be *both* internally consistent, *and* an accurate reflection of experimental reality. To discuss the first aspect, consistency, a physical theory must be formalized as some well-defined mathematical structure, some set of equations and mathematical rules that interrelate various mathematical symbols in some way. If this mathematical structure is internally inconsistent then the theory has already failed without a single experiment being performed. The second aspect is the extent to which this mathematical structure represents reality. The various mathematical symbols appearing in the equations must be asserted to correspond to some in-principle-measurable experimental quantities. A successful physical theory is one that is mathematically consistent *and* that accurately predicts/explains/retrodicts a suitably large class of experimental results.

But note one very important point: the internal logical and mathematical consistency of the theory is not decided by experiment — consistency is purely an issue of mathematics and logic and can be settled once and for all without recourse to experiment. (It is extremely rare for a physical theory to become in any way well-established and then later fail some internal consistency checks, there are simply too many physicists and mathematicians working on problems and checking each other's results.) Experiment only comes in at the second stage — no matter how beautiful or internally consistent a physical theory is it is simply not useful unless it is an accurate description of how the real universe works. (Sometimes we add qualifying phrases — such as "this theory works well in thus and so a range of parameters, but is known not to accurately reflect nature if one goes outside this range of parameters".)

Very Important Point: It is absolutely critical to realise that there is an enormous difference between being "wrong" and being a "crackpot" — more on this later.

2.7.2 Rules based on mathematical consistency

When it comes to detecting psycho-ceramics [crack-pottery] in quantum physics, the most basic guiding rule is fortunately very simple:

Rule 1 *If you run across someone who claims that the mathematical structure of quantum mechanics is internally inconsistent, then you can safely ignore them: they are wrong.*

Issues of internal mathematical consistency in quantum mechanics boil down to verifying the existence of suitable Hilbert spaces and suitable self-adjoint operators on these Hilbert spaces. (This is undergraduate-level mathematics, at least in principle.)

Unfortunately, when you start digging deeper, things very rapidly become much more subtle:

Rule 2 *If you run across someone who claims that the metaphysical foundations [epistemology and ontology] of quantum mechanics are internally inconsistent, be prepared to take such claims with several kilograms of salt (a pinch of salt is not sufficient).*

The philosophical foundations and metaphysics of quantum mechanics is an area where reasonable people can and do still differ — fortunately mucking around with the epistemology and ontology *does not change the predictions of the experimental results* — which are in excellent agreement with empirical reality.

Rule 3 *If you run across someone who claims that the metaphysical foundations [epistemology and ontology] of quantum mechanics can change the predictions for an experimental result, then you can safely ignore them: they are wrong.*

While essentially all physicists agree on what the rules of quantum mechanics are, there is very little agreement on what it all means. Interpretational issues are still murky (despite many loud, forceful, and mutually incompatible claims to the contrary). The major interpretational schemes are:

- Copenhagen interpretation: Bohr's original interpretation — this is the official party line.
- Everett's relative state interpretation: This has metamorphosed over the years into the Everett–DeWitt "many-worlds" interpretation.
  Popular among the high priests of quantum cosmology.
- Bohm's hidden variable interpretation: This one is a wild card. It works well for non-relativistic systems but does not seem to have a clean special relativistic extension, let alone a clean general relativistic extension. Highly non-local. Not really popular with anybody except a small cadre of dedicated converts.
- Cramer's transactional interpretation.
- Decoherence (Gell–Mann, Hartle, Zurek, *et. al.*).
- The "shut-up-and-calculate" non-interpretation: Extremely popular when teaching.
- Various variations on these themes. [Consistent histories (in the sense of quantum mechanical consistent histories, as opposed to time-travel consistent histories — don't ask...), pilot waves, *etc.*.]

The single most important thing to know about these various interpretations of quantum mechanics is that they are *all* compatible with experiment. Applied to ordinary quantum mechanics these interpretations are in fact experimentally indistinguishable.

Many physicists (and presumably a few interested outsiders as well) will be utterly mortified by the realization that there is a consistent hidden variable interpretation of ordinary (non-relativistic) quantum mechanics that is experimentally indistinguishable from the more usual interpretations. (This is not what you are generally taught in undergraduate or graduate school). If you want to get further into this, read David Bohm's 1952 papers and the commentary by John S. Bell (*Speakable and Unspeakable in Quantum Mechanics*). Note in particular that Bell's theorem deals only with *local* hidden variable theories.

Warning: Do not confuse Bohm's 1952 papers with his later work on "implicate order". These are rather different theories, and you do not automatically have to accept "implicate order" if you accept Bohm's 1952 "ontological interpretation".

Since the various interpretations of quantum mechanics are experimentally indistinguishable, they are by definition metaphysical constructs of no physical relevance. So why bring these issues up in a mathematics course? The various metaphysical interpretations of ordinary quantum mechanics become physically interesting *if and only if* one tries to alter or extend quantum mechanics in some fashion. Different metaphysical interpretations lead to different ideas of what a "natural" extension of quantum mechanics might be.

That is: An "interpretation" of quantum mechanics has no physical consequences and is merely an additional layer of metaphysics added to the theory. If one ever finds a situation where different "interpretations" of quantum mechanics lead to different physical results then one has *not* "interpreted" quantum mechanics — rather one has *modified* quantum mechanics to produce a *different* physical theory.

Warning: The relevant scientific literature is vast, inconclusive, and internally contradictory. The quality of the contributions is also extremely variable.

One of the most exciting aspects of all these interpretational issues is that the relevant experimental technology is getting better very quickly. Experimentalists are getting much better at constructing and manipulating mesoscopic (*i.e.*, reasonably large) quantum systems, and keeping them intact long enough to start to do interesting experiments. This is the same underlying technology that is driving the current interest in "quantum computing" and "quantum cryptography". A very important side-effect of these investigations is that it seems that at least some of the ideas related to foundational and interpretational issues in quantum mechanics are on the verge of moving out of the realm of philosophy and directly into the realm of experimental physics.

2.7.3 Parameterized post classical formalism?

To get a feel for how the situation in quantum physics differs from the way many of the theories of classical physics have been tested, consider this analogy: When discussing experimental tests of general relativity it is extremely useful to invoke the so-called parameterized post-Newtonian (PPN) formalism. The PPN formalism is a general phenomenological framework for describing weak-field gravity that contains a number of adjustable parameters. Depending on the values of these parameters, the PPN formalism describes the weak-field limits of Newtonian gravity, Einstein gravity, and many other *a priori* reasonable alternatives to ordinary general relativity. It is then a matter of experiment to measure the various parameters in the PPN formalism to see whether or not Einstein gravity is the theory that describes physical gravity. It is. (For a popular-level survey of the experimental situation, see Cliff Will's book, *"Was Einstein Right?"*.)

So far, no similar unified "parameterized post-classical" (PPC) formalism exists for studying quantum mechanics. What would be desired is a phenomenological model with many adjustable parameters. One choice of parameters should correspond to classical Newtonian mechanics. Another choice of parameters should correspond to ordinary quantum mechanics. Present tests of quantum mechanics are pretty much constructed on an *ad hoc* basis — a quantum mechanics calculation is performed and the results checked experimentally. The closest we currently have to a PPC formalism is the general complex of ideas associated with the Bell inequalities. Studying the metaphysical interpretations of ordinary quantum mechanics *might* give us a hint on how to flesh out a PPC formalism.

2.7.4 High weirdness in quantum field theory

It is when you move beyond (non-relativistic) quantum mechanics to (relativistic) quantum field theory that things really start to get messy:

Rule 4 *If you run across someone who claims that the mathematical structure of relativistic quantum field theory is internally inconsistent, then you will need to ask (and get answers to) a long string of highly technical questions to establish what exactly is being claimed.*[5]

There is little doubt in the community that quantum field theory gives a highly accurate description of what is really going on but there are a few niggling highly technical matters that might give reasonable people reason for pause — fortunately for the quantum physics community the relevant questions involve such highly abstruse mathematical issues that crackpot infestation has to date been kept to a minimum. (The relativity community has not been as lucky: Because special relativity can be formulated in terms of high-school algebra, anyone and everyone who is capable of getting high-school algebra wrong believes that they have something to say about special relativity.)

Rule 5 *If you run across someone who claims that the standard model of particle physics [$SU(3) \times SU(2) \times U(1)$] is not an accurate description of reality, then you will need to ask (and get answers to) a somewhat shorter string of somewhat less technical questions to establish what exactly is being claimed.*

The standard model of particle physics [$SU(3) \times SU(2) \times U(1)$] is our current best (though still provisional) model for elementary particles and their interactions. Crudely speaking, the particles in question are quarks and leptons, interacting via forces mediated by gluons, intermediate vector bosons, and photons, with the Higgs mechanism thrown in to get particle masses and symmetry breaking to come out right. While the standard model

[5] For the initiated I will mutter the incantations "Haag's theorem", and "triviality of $\lambda\, \phi_4^4$". Physicists tend not to worry about Haag's theorem, there's probably ways around it, though it sends mathematicians into paroxysms of apoplexy. Triviality on the other hand, worries even the physicists, and probably means that the Higgs mechanism of the standard model of particle physics is not truly fundamental, but is only an "effective" low-energy description of some deeper reality.

of particle physics has been tested to impressively high energies, it is widely [though not universally] expected that revisions to the model will be needed by the time accessible energies reach 10 TeV or so. Reasonable people can disagree on these issues (sometimes violently so), without going over the edge into crack-pottery.

Rule 6 *If you run across someone who claims that the logical structure of quantum gravity is mathematically inconsistent; ignore them, they are* **not even wrong**.[6]

The problem in this case is more subtle: as of 2013 we simply do not have a fully successful theory of quantum gravity — so attempts to show that a non-existent theory is internally inconsistent is more than a trifle premature. What is probably going on is that someone has stumbled across one of the many candidate models for quantum gravity, taken the model a little too seriously, and then misinterpreted a defect in the model as an intrinsic feature of quantum gravity.

2.7.5 The rough guide to crackpot filtering

The alert reader will have noticed that all this discussion of how to filter out potentially strange and peculiar physics I have not actually defined what a crackpot is. This is partly because there is no really generally agreed upon definition (though everyone will recognize one when they run across one). Crack-pottery is associated more with a style of argument and a style of presentation than it is with the actual content. It is important to realize that people can be wrong without being crackpots, and that crackpots can accidentally be right on some issues while still remaining crackpots — crack-pottery can be loosely characterized as:

1. an inability to mentally separate the logical structure of a physical theory from issues of experimental evidence, and

2. the inability to dispassionately assess the experimental evidence, generally coupled with overwhelming arrogance [and often, unfortunately, some form of mental disease].

A very rough-and-ready guide to crackpot detection has now been circulating in the internet for a few years. The crackpot index (see the website) was developed as a humorous attempt to summarize some of the rules of thumb derived from bitter experience in the flamewars infesting the internet newsgroup sci.physics. This internet newsgroup is so

[6] The pejorative phrase "not even wrong" is generally attributed to Wolfgang Pauli. In Pauli's view, an attitude shared by most physicists, there is no crime *per se* in being wrong — it's when ideas are so ill thought out that they are impossible to assess that you are better off keeping your mouth zipped until you have a firmer grasp on things.

heavily infested with crackpot drivel that very few (zero?) professional physicists are willing to put up with the personal abuse that generally results from giving straightforward non-crackpot answers to honest questions from genuinely curious non-experts. I shall leave it as an exercise to the reader to obtain internet access and make their own judgments. (For that matter, I should also warn readers with internet access that if you go to any of the standard internet search engines and type in the word relativity, your hits will be about 50 percent crackpot nonsense.)[7]

You should of course, not take the final score obtained from the crackpot index too seriously. A high crackpot index merely indicates that there *might* be a problem with the document, but there may be extenuating circumstances. Likewise a low crackpot index does not guarantee that the document is correct. Unlike the relatively rigid rules I provided earlier in this chapter, the crackpot index should be used only as a rough guide.

The key issues in avoiding a high crackpot index are:

1. Think your proposal through carefully and check it for internal consistency.
2. Make sure your proposal is compatible with current experimental data.
3. Don't *ever* try to claim that classical mechanics, special relativity, general relativity, or quantum mechanics are internally inconsistent.
4. Don't try to claim that any other presently accepted theory is internally inconsistent unless you have *very* good evidence presented in a *very* clear and convincing manner.

If any of these suggestions is violated you should be very suspicious of the author's claims.[8]

2.8 Last words

To wrap up this introductory chapter, permit me to summarize what you should have learned:

1. Physics theories can be quality graded (**Superb**/ **Useful**/ **Tentative**) with the **Superb** theories being so well verified by experiment that any direct attack on them is simply quixotic.
2. Quantum physics and general relativity are two of the **Superb** theories. In the rest of these notes I will describe quantum physics in a little more detail, and you should then have a basic understanding of what this theory entails.

[7]Exercise: Use a search engine to look up the crackpot index by John Baez.

[8]Exercise: Use a search engine to look up the essay "How to become a bad theoretical physicist" by Nobel prizewinner Gerhard 't Hooft. In a simiar vein, there is Warren Siegel's effort "Are you a quack?".

3. When judging strange and exotic claims and unfamiliar physics, try to look first for issues of internal mathematical consistency, secondly for compatibility with present experiment, and only then should you worry about the details of the "new physics".

Chapter 3

Heisenberg uncertainty principle

The Heisenberg uncertainty principle is such a basic aspect of quantum physics that it is at first a little scary to realise that the uncertainty principle intrinsically has very little to do with quantum physics itself. In fact:

- The first 90% of the Heisenberg uncertainty principle is actually rather basic and fundamental mathematics — the mathematics underlying Fourier series and Fourier transform theory.
- The remaining 10% of the Heisenberg uncertainty principle is *extremely elementary* quantum physics — amounting merely to invoking the de Broglie/ Einstein relations between (energy) $\leftrightarrow$ (frequency) and (momentum) $\leftrightarrow$ (wavenumber).

Now these claims will make many physicists choke, though many others will think these comments are utterly "obvious", so I'd better make a good job of justifying them.

3.1 Fourier transforms and signal theory

Suppose we are interested in some signal $s(t)$. This signal might be the pressure at a certain sensor, the voltage at a certain place on a wire, or the current through a given wire.

(Or it *might* be the value of the Schrödinger wavefunction at a certain place — but the signal does not *have* to be a quantum signal, classical signals are quite good enough for the point I want to make.)

Now Fourier transform the signal into frequency space:

$$\mathcal{F} : s(t) \mapsto \tilde{s}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} s(t)\ \exp[-i\omega t]\ \mathrm{d}t; \tag{3.1}$$

$$\mathcal{F}^{-1} : \tilde{s}(\omega) \mapsto s(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \tilde{s}(\omega)\ \exp[+i\omega t]\ \mathrm{d}\omega. \tag{3.2}$$

Some of you will never have seen Fourier transforms yet, don't panic:

- Simply take the integral above as the *definition* of the Fourier transform operator $\mathcal{F}$ which maps functions $s(t)$ to functions $\tilde{s}(\omega)$.
- Note that $\mathcal{F}$ is a linear operator (on the space of all functions).

  [More precisely: On some suitable class of functions where the integral actually converges.]
- *If it is helpful to you*, you might want to think of the signal $s(t)$ as an infinite-dimensional vector, with one distinct vector component for each distinct value of t. The Fourier transform can then be thought of as an (infinity)×(infinity) matrix. If this point of view confuses you, forget it.
- It is a *theorem* that in the second line above $\mathcal{F}^{-1}$ really is the inverse of the operator $\mathcal{F}$, and this really is the matrix inverse of the (infinity)×(infinity) matrix $\mathcal{F}$.
- Feel free to search on Google and Wikipedia for more information regarding Fourier transforms.

Theorem 1 (Fourier inverse) *Let us define the Fourier transform by*

$$\mathcal{F} : s(t) \mapsto \tilde{s}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} s(t)\ \exp[-i\omega t]\ \mathrm{d}t. \tag{3.3}$$

Then the inverse Fourier transform is

$$\mathcal{F}^{-1} : \tilde{s}(\omega) \mapsto s(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \tilde{s}(\omega)\ \exp[+i\omega t]\ \mathrm{d}\omega. \tag{3.4}$$

We will not prove this result — we will merely use it.

Definition 1 *The Dirac delta function is* defined *by*

$$f(t) = \int_{-\infty}^{+\infty} f(t')\ \delta(t-t')\ \mathrm{d}t'; \tag{3.5}$$

$$g(\omega) = \int_{-\infty}^{+\infty} g(\omega')\ \delta(\omega-\omega')\ \mathrm{d}\omega'. \tag{3.6}$$

From the Fourier inversion theorem, and the definition of the Dirac delta function, it is easy to show:

Theorem 2 (Dirac delta function) *The Dirac delta function satisfies*

$$\delta(t - t') = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-i\omega[t - t'])\ d\omega, \tag{3.7}$$

or equivalently

$$\delta(\omega - \omega') = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-i[\omega - \omega']t)\ dt. \tag{3.8}$$

Here the Dirac delta function is defined *by*

$$f(t) = \int_{-\infty}^{+\infty} f(t')\ \delta(t - t')\ dt'; \tag{3.9}$$

$$g(\omega) = \int_{-\infty}^{+\infty} g(\omega')\ \delta(\omega - \omega')\ d\omega'; \tag{3.10}$$

We will not prove this result — we will merely use it.

You should prove this result as one of the assignment problems.

If we were being very careful, I would of course admit that the Dirac delta function is not really a function — it is a *functional* defined so as to act on a suitable set of well-behaved functions (for example, functions of compact support or functions of exponential falloff at infinity) — the resulting Schwartz distribution theory is a way of making standard delta-function manipulations completely rigorous.

Let's now assume that the signal $s(t)$ is peaked at time $\bar{t}$ and has a width Δt. Similarly let's assume that the Fourier transformed signal $\tilde{s}(\omega)$ is peaked at angular frequency $\bar{\omega}$ and has a width $\Delta\omega$. (I do not at this stage need to define these notions more precisely.)

If I now re-scale the time coordinate and the signal according to the prescription

$$\mathcal{T} : s(t) \to s_\kappa(t) = s(\kappa\ t), \tag{3.11}$$

then even without knowing the precise definition of Δt it makes sense to demand that for the re-scaled signal

$$\Delta t_{[s_\kappa]} = \kappa^{-1}\ \Delta t_{[s]}. \tag{3.12}$$

Furthermore it is easy to calculate the Fourier transform of $[\mathcal{T} s_\kappa](t)$ and see that

$$\tilde{s}_\kappa(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} s_\kappa(t)\ \exp[-i\omega t]\ dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} s(\kappa\ t)\ \exp[-i\omega t]\ dt = \frac{\tilde{s}(\omega/\kappa)}{\kappa}. \tag{3.13}$$

Then even without knowing the precise definition of $\Delta\omega$ it makes sense to demand

$$\Delta\omega_{[\tilde{s}_\kappa]} = \kappa \; \Delta\omega_{[\tilde{s}]}. \tag{3.14}$$

That is, squeezing the signal in the time direction by a factor κ forces it to spread in the frequency domain by the same factor κ — and this will be true for *any* sensible definition for Δt and $\Delta\omega$. Consequently

$$\Delta t_{[s_\kappa]} \; \Delta\omega_{[\tilde{s}_\kappa]} = \Delta t_{[s]} \; \Delta\omega_{[\tilde{s}]} = C(s, \tilde{s}), \tag{3.15}$$

where $C(s, \tilde{s})$ is some functional independent of κ, and which depends at most on the "shape" of the signal $\{s(t), \tilde{s}(\omega)\}$. This is already enough to tell you, with minimal calculation, that precisely localizing the signal in time t will force its Fourier transform to be de-localized in angular frequency ω, and vice versa.

We can sharpen these comments to obtain a rigorous inequality, by adding a little technical structure and by more precisely defining Δt and $\Delta\omega$. Start by defining

$$\bar{t} = \frac{\int_{-\infty}^{+\infty} \; t \; |s(t)|^2 \; \mathrm{d}t}{\int_{-\infty}^{+\infty} \; |s(t)|^2 \; \mathrm{d}t} \tag{3.16}$$

and

$$(\Delta t)^2 = \frac{\int_{-\infty}^{+\infty} \; (t - \bar{t})^2 \; |s(t)|^2 \; \mathrm{d}t}{\int_{-\infty}^{+\infty} \; |s(t)|^2 \; \mathrm{d}t} \tag{3.17}$$

and analogously

$$\bar{\omega} = \frac{\int_{-\infty}^{+\infty} \; \omega \; |\tilde{s}(\omega)|^2 \; \mathrm{d}\omega}{\int_{-\infty}^{+\infty} \; |\tilde{s}(\omega)|^2 \; \mathrm{d}\omega} \tag{3.18}$$

and

$$(\Delta\omega)^2 = \frac{\int_{-\infty}^{+\infty} \; (\omega - \bar{\omega})^2 \; |\tilde{s}(\omega)|^2 \; \mathrm{d}\omega}{\int_{-\infty}^{+\infty} \; |\tilde{s}(\omega)|^2 \; \mathrm{d}\omega} \tag{3.19}$$

Now these are *definitions*, there is nothing here that has to be *proved*. The ultimate justification for these definitions is that they are both natural and useful. In particular:

- These definitions are plausible, and do the sensible thing under rescaling the time and frequency domains.
- You can think of $\bar{t}$ as an estimate of the "arrival time" of the non-trivial part of the signal $s(t)$, and Δt as an estimate of the amount of time over which the signal is non-trivial.
- You can think of $\bar{\omega}$ as the "average frequency" contained in the signal, and $\Delta\omega$ as an estimate of the amount of the "frequency spread".

- These definitions lead to useful results.
- The reason for the presence of $|s(t)|^2$ and $|\tilde{s}(\omega)|^2$ as useful weighting factors can be motivated by considering:

  Theorem 3 (Parseval's theorem)

$$\int_{-\infty}^{+\infty} |s(t)|^2 \, \mathrm{d}t = \int_{-\infty}^{+\infty} |\tilde{s}(\omega)|^2 \, \mathrm{d}\omega. \tag{3.20}$$

  We will not prove this result — we will merely use it.

  You should prove this result as one of the assignment problems.

I want to emphasise that everything is perfectly classical, there is not a $\hbar$ in sight.

Nevertheless, one can prove the following *classical uncertainty relation* as an utterly rigorous *theorem* of Fourier transform theory.

Theorem 4 (Classical uncertainty relation)
Let $s(t)$ be an arbitrary signal (classical or quantum), and let $\tilde{s}(\omega)$ be its Fourier transform.

Let Δt and $\Delta\omega$ be as defined above.

Then

$$\Delta\omega \times \Delta t \geq \frac{1}{2}. \tag{3.21}$$

Proof: The proof is deferred for now. You can find it a few pages further along in these notes, or at several places on the internet, (use Google), or you can adapt various proofs of the Heisenberg uncertainty relations from physics textbooks, by stripping out all the irrelevant $\hbar$'s and dropping all irrelevant references to energy and momentum.

You will effectively prove this result, in stages, as assignment 1.

There is of course nothing special about the use of time and frequency as our Fourier transform variables, we could just as easily work with position and wave-number. That is, consider a position-dependent signal $s(x)$ and its Fourier transform

$$\mathcal{F} : s(x) \mapsto \tilde{s}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} s(x) \, \exp[-ikx] \, \mathrm{d}x \tag{3.22}$$

$$\mathcal{F}^{-1} : \tilde{s}(k) \mapsto s(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \tilde{s}(k) \, \exp[+ikx] \, \mathrm{d}k \tag{3.23}$$

Then, suitably defining Δx and Δk (and I leave this step as an exercise), we have the second classical theorem below:

Theorem 5 (Classical uncertainty relation)
Let $s(x)$ be an arbitrary position-dependent signal (classical or quantum), and let $\tilde{s}(k)$ be its Fourier transform into the wave-number domain.

Let Δx and Δk be defined in a manner analogous to Δt and $\Delta\omega$ above.

Then

$$\Delta k \times \Delta x \geq \frac{1}{2}. \tag{3.24}$$

These two classical uncertainty relations are basic theorems of fundamental mathematics — no physics assumptions have been made. These two theorems also encode about 90% of what is referred to as the "Heisenberg uncertainty relations".

Here's a nice application of the classical uncertainty theorem: Since for any signal $s(t)$ we have

$$\Delta\omega \times \Delta t \geq \frac{1}{2}, \tag{3.25}$$

ask what happens if I try to force B bits per second down a communications channel?

With B bits per second going by, each bit has to be localized in time to within roughly $\Delta t \lesssim 1/(2B)$ seconds. But then

$$\Delta\omega \gtrsim \frac{1}{2\Delta t} \gtrsim \frac{1}{2/(2B)} \gtrsim B \tag{3.26}$$

That is, when you Fourier transform any old signal containing B bits per second of information, it must spread out a distance of at least B in frequency space.

Definition 2 (Bandwidth — physical definition)
For any signal $s(t)$ the bandwidth is defined as the difference between the maximum frequency in the signal and the minimum frequency in the signal.

(For pragmatic reasons it is common practice to approximate *this* max–min *definition of bandwidth by either FWHM, that is "full width at half maximum" for the Fourier transformed signal $\tilde{s}(\omega)$, or by some suitably weighted standard deviation $\Delta\omega$.)*

The first classical uncertainty theorem then says:

$$(\text{bit rate}) \lesssim (\text{bandwidth}) \tag{3.27}$$

In fact, this theorem is so basic and fundamental that in computer science it is now common to use this as the *definition* of bandwidth:

Definition 3 (Bandwidth — computer science definition)
Bandwidth is the maximum possible rate at which you can force bits through a communications channel.

I hope this has convinced you of the importance and fundamental nature of the classical uncertainty theorem. Remember — there's not (yet) a $\hbar$ in sight...

3.2 The de Broglie and Einstein relations

Quantum physics enters once you adopt the de Broglie hypothesis and Einstein's quantization hypothesis:

Hypothesis 1 (de Broglie waves)
All particles exhibit wavelike properties, and the relation between the energy/momentum of the particle and the angular-frequency/wavenumber of the associated wave is:

$$E = \hbar\, \omega; \qquad \text{and} \qquad p = \hbar\, k. \tag{3.28}$$

The de Broglie hypothesis is the converse of Einstein's quantization hypothesis (an extension of Einstein's analysis of the photoelectric effect, which led to the concept of the photon). In its most general formulation:

Hypothesis 2 (Einstein quantization)
All wavelike excitations exhibit particle-like properties and the relation between the angular-frequency/wavenumber of the wave and the energy/momentum of the associated particle is:

$$E = \hbar\, \omega; \qquad \text{and} \qquad p = \hbar\, k. \tag{3.29}$$

Once we have these relations between angular frequency and energy, and wavenumber and momentum, we can convert the classical uncertainty relations above (relevant for arbitrary signals), into the quantum Heisenberg uncertainty relation (relevant now for specifically quantum excitations).

3.3 The Heisenberg uncertainty principle

Let the signal we are interested in be $\psi(t)$, the value of the Schrodinger wavefunction at some particular point. We can still define $\bar{t}$ and Δt as for a classical signal above. Since in quantum physics $|\psi(t)|^2$ has the physical interpretation of being proportional to a probability density, we now have the interpretation that the particle passes by roughly at time

$$\bar{t} \pm \Delta t \tag{3.30}$$

(That is, in a statistical ensemble of identically prepared particles, the average arrival time is $\bar{t}$, and the standard deviation in arrival times is Δt.)

We can still define $\bar{\omega}$ and $\Delta\omega$ as for the classical signal. The new twist is that, in view of the Fourier transform formula,

$$\psi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \tilde{\psi}(\omega) \; \exp[+i\omega t] \, \mathrm{d}\omega \tag{3.31}$$

the signal can be viewed as a superposition of waves, each of which via the Einstein quantization hypothesis, corresponds to a particle-like excitation. The average energy of these particle like excitations is

$$\bar{E} = \hbar \, \bar{\omega} \tag{3.32}$$

and the uncertainty in the energy is

$$\Delta E = \hbar \, \Delta\omega \tag{3.33}$$

(In a statistical ensemble of otherwise identically prepared particles this would be the standard deviation of the measured energies.) Then, since the classical theorem tells us

$$\Delta\omega \times \Delta t \geq \frac{1}{2}, \tag{3.34}$$

adding quantum physics results in

$$\Delta E \times \Delta t \geq \frac{\hbar}{2}, \tag{3.35}$$

which is now the quantum Heisenberg uncertainty relation. Note that in this way of setting things up ΔE and Δt have very simple physical interpretations:

- ΔE is simply the standard deviation in measured energies.
- Δt is simply the standard deviation in arrival times.

Similarly for the position-momentum uncertainty relation. We have

$$\bar{p} = \hbar \, \bar{k} \tag{3.36}$$

for the average momentum, and

$$\Delta p = \hbar \, \Delta k \tag{3.37}$$

for the standard deviation in measured momentum.

Then, since the classical theorem tells us

$$\Delta k \times \Delta x \geq \frac{1}{2}, \tag{3.38}$$

adding quantum physics results in

$$\Delta p \times \Delta x \geq \frac{\hbar}{2}, \tag{3.39}$$

which is the other quantum Heisenberg uncertainty relation.

Viewed in the light above the Heisenberg uncertainty principle is closely related to signal processing theory, and in particular to Nyquist's sampling theorem and Shannon's theorem on the relationship between channel capacity and bandwidth. With hindsight, given what was already known about Fourier transforms in 1925, it should have been "obvious" that the de Broglie relations automatically and immediately lead to the Heisenberg uncertainty principle. That was not the way Heisenberg's insight was historically achieved, at least partly because signal processing theory had not yet been developed to the level it subsequently was.

(In fact it was one of John Wheeler's graduate students [yes, the same chap that wrote your special relativity textbook], Claude Shannon, who was then largely responsible for laying the mathematical and physical foundations of signal theory.)

3.4 Classical operators and commutators

A classical operator is simply a linear mapping that takes a signal $s_1(t)$ to a new signal $s_2(t)$. That is, an operator is a linear mapping on the function space of all possible signals. Two very important linear mappings are:

$$t : s(t) \to t\, s(t) \tag{3.40}$$

$$\partial_t : s(t) \to \partial_t s(t) \tag{3.41}$$

corresponding to "multiplication by t" and "differentiation with respect to t". Note that we can evaluate the commutator

$$[\partial_t, t] \tag{3.42}$$

by looking at its action on arbitrary signals:

$$[\partial_t, t]s(t) = (\partial_t\, t - t\, \partial_t)s(t) = s(t) + t\, \partial_t s(t) - t\, \partial_t s(t) = s(t). \tag{3.43}$$

That is

$$[\partial_t, t] = \mathbf{I} \tag{3.44}$$

Note that this operator statement is purely a mathematical statement about how the two processes of differentiation and "multiplication by the variable you are differentiating with respect to" commute with each other. There is — as yet — no quantum physics in this relation. Similarly in position space we have

$$[\partial_x, x] = \mathbf{I} \tag{3.45}$$

Now let's see how these classical operators interact with the process of Fourier transformation.

We have

$$\mathcal{F}\left[ts(t)\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t\, s(t)\;\exp[-i\omega t]\;\mathrm{d}t = i\frac{\mathrm{d}}{\mathrm{d}\omega}\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} s(t)\;\exp[-i\omega t]\;\mathrm{d}t \tag{3.46}$$

$$= i\frac{\mathrm{d}}{\mathrm{d}\omega}\tilde{s}(\omega) = i\frac{\mathrm{d}}{\mathrm{d}\omega}\mathcal{F}[s(t)], \tag{3.47}$$

and similarly (integrating by parts)

$$\mathcal{F}\left[\partial_t s(t)\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \partial_t s(t)\;\exp[-i\omega t]\;\mathrm{d}t = i\omega\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} s(t)\;\exp[-i\omega t]\;\mathrm{d}t \tag{3.48}$$

$$= i\omega\;\tilde{s}(\omega) = i\omega\mathcal{F}[s(t)]. \tag{3.49}$$

More formally we can write this as:

$$\mathcal{F}\circ t = i\frac{\mathrm{d}}{\mathrm{d}\omega}\circ\mathcal{F}; \qquad\qquad \mathcal{F}\circ\partial_t = i\omega\circ\mathcal{F}. \tag{3.50}$$

There are similar statements that can be made about the inverse Fourier transform process $\mathcal{F}^{-1}$. Namely:

$$\mathcal{F}^{-1}\circ\omega = -i\partial_t\circ\mathcal{F}^{-1}; \qquad\qquad \mathcal{F}^{-1}\circ\frac{\mathrm{d}}{\mathrm{d}\omega} = -it\circ\mathcal{F}^{-1}. \tag{3.51}$$

Similar results hold for a signal that depends on position (which is Fourier transformed to wavenumber). We collect these results in a theorem regarding classical operators.

Theorem 6 (Classical operator theorem — time-frequency domain)
Acting on the function space of all time-domain signals $\mathcal{S} = \{s(t)\}$ the classical linear operators

$$t : s(t) \to t\;s(t) \tag{3.52}$$

$$\partial_t : s(t) \to \partial_t s(t) \tag{3.53}$$

satisfy the classical commutation relation

$$[\partial_t, t] = \mathbf{I} \tag{3.54}$$

and the intertwining relations

$$\mathcal{F}\circ t = i\frac{\mathrm{d}}{\mathrm{d}\omega}\circ\mathcal{F}; \qquad\qquad \mathcal{F}\circ\partial_t = i\omega\circ\mathcal{F}. \tag{3.55}$$

$$\mathcal{F}^{-1}\circ\omega = -i\partial_t\circ\mathcal{F}^{-1}; \qquad\qquad \mathcal{F}^{-1}\circ\frac{\mathrm{d}}{\mathrm{d}\omega} = -it\circ\mathcal{F}^{-1}. \tag{3.56}$$

Theorem 7 (Classical operator theorem — space-wavenumber domain) *Acting on the function space of all space-domain signals $\mathcal{S} = \{s(x)\}$ the classical linear operators*

$$x : s(x) \to x\; s(x) \tag{3.57}$$

$$\partial_x : s(x) \to \partial_x s(x) \tag{3.58}$$

satisfy the classical commutation relation

$$[\partial_x, x] = \mathbf{I} \tag{3.59}$$

and the intertwining relations

$$\mathcal{F} \circ x = i \frac{\mathrm{d}}{\mathrm{d}k} \circ \mathcal{F}; \qquad \mathcal{F} \circ \partial_x = ik \circ \mathcal{F}. \tag{3.60}$$

$$\mathcal{F}^{-1} \circ k = -i\partial_x \circ \mathcal{F}^{-1}; \qquad \mathcal{F}^{-1} \circ \frac{\mathrm{d}}{\mathrm{d}k} = -ix \circ \mathcal{F}^{-1}. \tag{3.61}$$

Some simple theorems (that I will leave to the reader to actually prove, they are part of the homework exercises) are the shifting and modulation theorems (and their analogues in the space-wavenumber domain):

Theorem 8 (Shifting theorem)

$$\mathcal{F}[s(t+t_0)](\omega) = \exp(+i\omega\; t_0)\; \mathcal{F}[s(t)](\omega) \tag{3.62}$$

You will prove this as part of assignment 1.

Theorem 9 (Modulation theorem)

$$\mathcal{F}[\exp(-i\omega_0\; t)\; s(t)](\omega) = \mathcal{F}[s(t)](\omega + \omega_0) \tag{3.63}$$

You will prove this as part of assignment 1.

Now introduce some notation:

Definition 4 (Inner product)

$$\langle s_1, s_2 \rangle \equiv \int_{-\infty}^{+\infty} s_1^*(t)\; s_2(t)\; \mathrm{d}t \tag{3.64}$$

$$\langle \tilde{s}_1, \tilde{s}_2 \rangle \equiv \int_{-\infty}^{+\infty} \tilde{s}_1^*(\omega)\; s_2(\omega)\; \mathrm{d}\omega \tag{3.65}$$

Definition 5 (Hermitian operator) *A Hermitian operator H satisfies:*

$$\langle s_1, Hs_2 \rangle = \langle Hs_1, s_2 \rangle. \tag{3.66}$$

Definition 6 (anti-Hermitian operator) *An anti-Hermitian operator A satisfies:*

$$\langle s_1, As_2 \rangle = -\langle As_1, s_2 \rangle. \tag{3.67}$$

Once you have defined this inner product, and these notions of Hermiticity and anti-Hermiticity, it is easy to check that the operator t is Hermitian (self-adjoint), while ∂_t is anti-Hermitian (anti-self-adjoint).

You will prove this as part of assignment 1.

Remember that a Hermitian *matrix* satisfies

$$[(M)^*]^T = M, \tag{3.68}$$

that is, the matrix is equal to the transpose of its complex conjugate.

Similarly, an anti-Hermitian *matrix* satisfies

$$[(M)^*]^T = -M, \tag{3.69}$$

that is, the matrix is equal to minus the transpose of its complex conjugate.

These notions extend naturally to (infinity)×(infinity) matrices, and thence to operators on function spaces.

(The distinction between Hermitian operators and self-adjoint operators is a highly technical one that has to do with the details of the precise function space you are working on: square-integrable functions? functions of compact support? whatever? *Do not worry about it* for the purposes of this module.)

This anti-Hermitian property for the derivative operator ∂_t is why physicists prefer to work with the operator $i\partial_t$, because $i\partial_t$ is Hermitian due to the presence of the extra factor i, but this is merely a matter of taste and convenience.

Furthermore (and this might be a little more surprising the first time you see it), in terms of this inner product the Fourier transform process is *unitary* linear operator on signal-space.

Definition 7 (unitary operator) *A uinitary operator U satisfies:*

$$\langle s_1, Us_2 \rangle = \langle U^{-1}s_1, s_2 \rangle. \tag{3.70}$$

Remember that a unitary *matrix* satisfies

$$U^{-1} = U^\dagger = (U^*)^T \tag{3.71}$$

This notion extends naturally to (infinity)×(infinity) matrices, and thence to operators on function spaces.

The equivalent statement for the Fourier transform process (and inverse Fourier transform process) is that

$$\mathcal{F}^{-1} = \mathcal{F}^\dagger = (\mathcal{F}^*)^T \tag{3.72}$$

or equivalently (but more formally)

$$\langle \tilde{s}_1, \mathcal{F} s_2 \rangle = \langle \mathcal{F}^{-1} \tilde{s}_1, s_2 \rangle. \tag{3.73}$$

You will prove this as part of assignment 1.

This is now a quite sufficient quantity of Fourier transform theory to do the job we are interested in.

3.5 Proof of the classical uncertainty relation

Let $s(t)$ be an arbitrary classical signal. In view of the shifting theorem and modulation theorem there is no loss in generality in taking $\bar{t} = 0$ and $\bar{\omega} = 0$. (You should easily be able to convince yourself this is true. Try it. In fact, it's a homework exercise.)

Now consider the quantity

$$Q(a,b) = \langle (a\;t + b\;\partial_t) s, (a\;t + b\;\partial_t) s \rangle \tag{3.74}$$

where a and b are both real.

Since this is the inner product of a vector with itself we know that $Q(a,b) \geq 0$. On the other hand, since t is Hermitian and ∂_t is anti-Hermitian

$$Q(a,b) = \langle s, (a\;t - b\;\partial_t)(a\;t + b\;\partial_t) \rangle \tag{3.75}$$

That is

$$Q(a,b) = \left\langle s, \left[a^2\;t^2 - b^2\;\partial_t^2 + ab(t\;\partial_t - \partial_t\;t)\right] s \right\rangle \tag{3.76}$$

which simplifies to

$$Q(a,b) = \left\langle s, \left(a^2\;t^2 - b^2\;\partial_t^2 - ab\;[\partial_t, t]\right) s \right\rangle. \tag{3.77}$$

That is

$$Q(a,b) = \left\langle s, \left(a^2\;t^2 - b^2\;\partial_t^2 - ab\;\mathbf{I}\right) s \right\rangle \tag{3.78}$$

so that

$$Q(a,b) = a^2 \langle s, t^2 s\rangle - b^2 \langle s, \partial_t^2 s\rangle - ab \langle s, s\rangle \tag{3.79}$$

But by definition (remember, we have already shifted the time coordinate so that $\bar{t} = 0$)

$$(\Delta t)^2 = \frac{\langle s, t^2 s\rangle}{\langle s, s\rangle} \tag{3.80}$$

and with a few elementary steps (remember $\bar{\omega} = 0$, and the rest can be obtained from Parseval's theorem and simple integral manipulations)

$$(\Delta\omega)^2 = \frac{\langle \tilde{s}, \omega^2 \tilde{s}\rangle}{\langle \tilde{s}, \tilde{s}\rangle} = -\frac{\langle s, \partial_t^2 s\rangle}{\langle s, s\rangle} \tag{3.81}$$

The net result is that

$$Q(a,b) = \left\{a^2 (\Delta t)^2 + b^2 (\Delta\omega)^2 - ab\right\} \langle s, s\rangle \tag{3.82}$$

Consequently for all real a, b we must have:

$$a^2 (\Delta t)^2 + b^2 (\Delta\omega)^2 - ab \geq 0 \tag{3.83}$$

That is

$$[a\ \Delta t - b\ \Delta\omega]^2 + 2ab \left[\Delta t\ \Delta\omega - \frac{1}{2}\right] \geq 0 \tag{3.84}$$

Since this must hold for *arbitrary* real a, b we can *in particular* choose $a = \Delta\omega$ and $b = \Delta t$, to deduce

$$\Delta t\ \Delta\omega \geq \frac{1}{2}. \tag{3.85}$$

This is what we set out to prove.

Notice that as promised, this is purely a theorem about mathematics — in particular about Fourier transform theory and signal processing, with not a single $\hbar$ in sight.

3.6 Comments

Though the math here is straightforward, even trivial, there is a very important point here — a lot of what goes on in quantum mechanics is unnecessarily mystical because people go out of their way to make it look mysterious. In particular, much of what falls under the rubric of "uncertainty" is a purely classical phenomenon that is not really very mysterious at all. If you really want mystery, focus on the de Broglie hypothesis and the Einstein quantization hypothesis — they are in many ways both the *essential physical core* and the *mathematically trivial part* of the Heisenberg uncertainty relations.

Of course it's possible to extend the considerations of this section in many ways:

- Classical extensions of these uncertainty relations focus on different and better ways of estimating (or defining) Δt and $\Delta\omega$.
- You could try to generalize beyond Fourier transforms, maybe to Bessel transforms, Mellin transforms, or wavelet transforms — I'm not really sure what's known or not known on those topics and that might make a nice little research project.
- You might also want to search on the net for "generalized uncertainty relations".

Chapter 4

Tunnelling

Tunnelling is another one of those things you most likely first learn about when studying quantum mechanics — and by and large most physicists are taught that tunnelling is an intrinsically quantum phenomenon. Of course this is complete and utter nonsense; tunnelling is a *wave* phenomenon that occurs whenever you are dealing with wavelike excitations.

The only thing "quantum" about the tunnelling phenomenon is the the de Broglie hypothesis. Whenever you are in a regime where the "wavelike" aspects of your "particles" are important, then you have the possibility of encountering the tunnelling phenomenon.

But tunneling can occur in situations where the "particle" aspects of your "waves" are completely negligible. That is, there are purely classical situations in wave physics where *classical tunnelling can and does occur.*

Three elementary situations where classical tunnelling occurs (and I'm sure there are many others) are:

- Sound waves penetrating a window.
- Radio waves penetrating a building.
- Frustrated total internal reflection.

Two of these situations are so simple as to be almost trivial, but "frustrated total internal refraction" is a nice physical result that merits some attention.

4.1 Sound penetrating a barrier

We have all heard sound waves penetrating a window. And once you stop to think about it the physics is trivial: Sound waves in air are vibrations in air pressure. The pressure fluctuations impinge on the window and set it vibrating. These vibrations travel through the glass and excite pressure fluctuations on the far side of the window. No atoms penetrate the glass, but the wave proceeds — it may be refracted and attenuated, but it will get through.

4.2 Radio penetrating a barrier

We have similarly all had the experience of turning on a radio indoors, without external antenna, which makes it clear that radio waves can penetrating a significant quantity of building material. And once you stop to think about it the physics is again trivial: Radio waves have relatively long wavelengths which lets them diffract around hills or buildings — but equally well the long wavelength implies a significant "skin depth" when encountering a building. The radio waves may be refracted and attenuated, but a certain amount of electromagnetic energy will get through. In the visible range the analogous phrase is "optical depth".

The fraction of electromagnetic radiation getting through is

$$\exp\left[-\int \sigma \, \mathrm{d}z\right] \tag{4.1}$$

where $1/\sigma$ has units of length and is called the "skin depth" while the dimensionless integral $\int \sigma \, \mathrm{d}z$ is referred to as the "optical depth". For a homogeneous slab you can often approximate this as

$$\text{optical depth} = \int \sigma \, \mathrm{d}z \approx \frac{\text{physical depth}}{\text{skin depth}} \tag{4.2}$$

These two examples are sufficiently elementary that I think you should now be convinced:

- Tunnelling is fundamentally a wave phenomenon!

4.3 Frustrated total internal reflection

This is a beautiful result of classical physical optics that deserves to be much better known than it currently is.

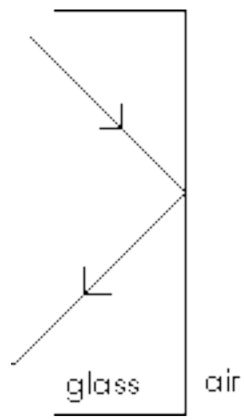

Figure 4.1: Total internal reflection

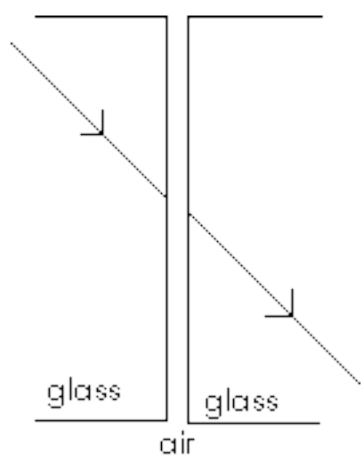

Figure 4.2: Frustrated total internal reflection

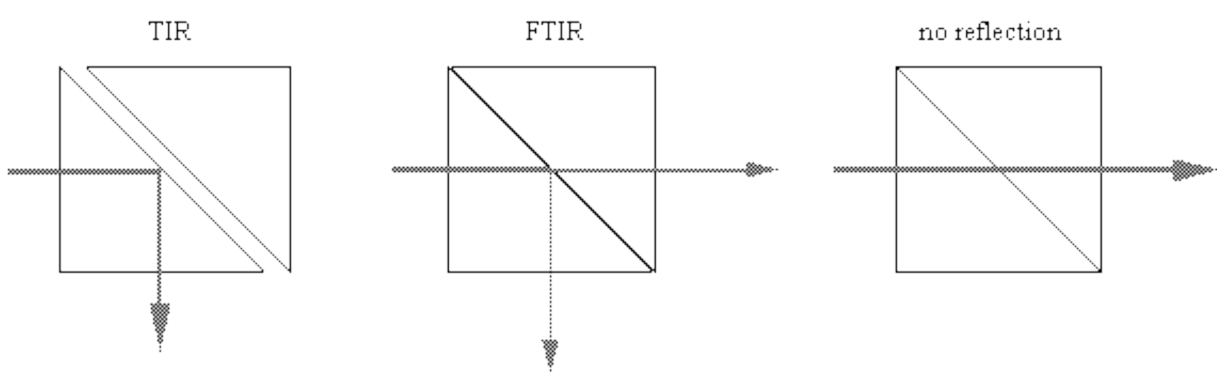

Figure 4.3: Transition from total internal reflection, to frustrated total internal reflection, to transparency

4.3.1 Key idea

Frustrated Total Internal Reflection (FTIR) occurs [for example] when a ray of light travelling through glass strikes an interface at an angle exceeding the critical angle. You would expect it to be totally-internally-reflected at the glass/air interface.

However, if another piece of glass is placed close to (but not touching) the interface, some light will evanescently couple through the thin gap and propagate. Both the reflected and transmitted beams will be affected, depending on the thickness of the gap. In the limit of the gap having zero thickness, the light will continue as if there were no boundary. In the limit of a large gap, more than a wavelength or two, then virtually all the light is internally reflected.

FTIR can also be seen through a experiment that can be done at home. Fill a glass approximately half way full of water. If you try to see objects just outside the glass by looking down through the top of the water and out the side, you cannot see anything because of the total internal reflection at the water–glass interface. However, if you press your fingers tightly against the glass, you can see the whorls of your fingerprints through the glass. In this case the air gap is reduced enough so that the electromagnetic waves can transit into the glass, and from there into the water and air and eventually your eyes.

4.3.2 Reminder: Snell's law

Reminder: Snell's law says

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{4.3}$$

where n_1 and n_2 are the refractive indices of the two media, and θ_1 and θ_2 are the angles between the direction of the light ray (or acoustic ray) and the normal to the interface.

4.3.3 Some technicalities

Here's a few standard definitions that I extracted from "Eric Weisstein's world of physics" (slightly edited for clarity):

> **Frustrated total internal reflection:**
> If an evanescent wave (such as that produced by total internal reflection) extends across a separating medium into a region occupied by a higher index of refraction material, energy may flow across the boundary. This phenomenon is known as frustrated total internal reflection, and is similar to quantum mechanical tunneling or barrier penetration. When transmission across the

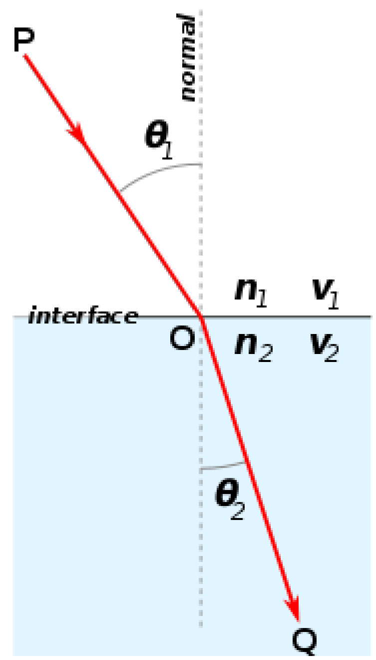

Figure 4.4: Snell's law

boundary occurs in this manner, the "total internal reflection" is no longer total since the transmitted wave comes at the expense of the internally reflected one.

Total internal reflection:
Total internal reflection is the reflection of electromagnetic radiation from the interface of medium with larger index of refraction n_1 with a medium of smaller index of refraction n_2 (with $n_2 < n_1$) when making an angle

$$\theta_1 > \sin^{-1}(n_2/n_1) \tag{4.4}$$

to the normal. Total internal reflection can be used to losslessly redirect a light beam in the direction of its source using a $45° - 45° - 90°$ prism.

However, there is still an electric field in medium n_2, given by

$$E_t(t,x,y,z) = E_2 \ \exp\left\{-k_2 z\sqrt{(\sin\theta_1/\sin\theta_*)^2 - 1}\right\} \ \cos\left(\omega t - k_2 y\frac{n_1}{n_2}\sin\theta_1\right) \tag{4.5}$$

where k_2 is the wavenumber in medium 2 and θ_* is the critical angle

$$\theta_* = \sin^{-1}(n_2/n_1) \tag{4.6}$$

(See, for example, Bekefi and Barrett 1987, p. 477). Note that this field falls off exponentially with distance z from the interface, and propagates along the interface in the y direction. This disturbance is therefore known as a surface wave and has phase velocity

$$v = \frac{\omega}{k_2 \left(\frac{n_1}{n_2}\right) \sin\theta_1} \tag{4.7}$$

Reference:
Bekefi, G. and Barrett, A. H.
"Electromagnetic Vibrations, Waves, and Radiation".
Cambridge, MA: MIT Press, pp. 475-483, 1987.

Evanescent wave:
An electromagnetic wave observed in total internal reflection, undersized waveguides, and in periodic dielectric heterostructures. While wave solutions have real wavenumbers k; k for an evanescent mode is purely imaginary. Evanescent modes are characterized by an exponential attenuation and lack of a phase shift.

Critical angle:
If the angle of incidence of light on a dielectric medium is greater than a critical angle θ_*, then the light experiences total internal reflection instead of refraction. The angle is given by

$$\theta_* = \sin^{-1}(n_2/n_1) \tag{4.8}$$

where is the θ_* angle from the normal, and n_1 and n_2 are the indices of refraction of the original and second media, respectively.

For $n_2 < n_1$, a ray incident at an angle greater than θ_* will undergo total internal reflection. However, it is impossible to satisfy the boundary conditions if there is no transmission, so a surface (or evanescent) wave must be present. Beyond the critical angle, both the reflection and transmission coefficients are complex. If another material is placed near the evanescent wave, frustrated total internal reflection may occur.

4.3.4 Barrier penetration

Note that with a little work, in particular using Snell's law, you can re-write the evanescent wave as:

$$E_t(t,x,y,z) = E_2 \, \exp\left\{-k_1 z\sqrt{\sin^2\theta_1 - \sin^2\theta_*}\right\} \, \cos\left(\omega t - k_1 y \sin\theta_1\right) \tag{4.9}$$

Suppose the slab of material n_2 has thickness L.

- Temporarily ignoring backscatter from either interface, the electric field just *above* the first interface ($z = 0^+$) would be:

$$E_t(t, x, y, 0^+) = E_2 \cos\left(\omega t - k_2 y \frac{n_1}{n_2} \sin\theta_1\right). \tag{4.10}$$

- Again temporarily ignoring backscatter from either interface, the electric field just *below* the second interface ($z = L^-$) would be:

$$E_t(t, x, y, L^-) = E_2 \exp\left\{-k_i L \sqrt{\sin^2\theta_1 - \sin^2\theta_*}\right\} \cos\left(\omega t - k_2 y \frac{n_1}{n_2} \sin\theta_1\right) \tag{4.11}$$

  That is:

$$E_t(t, x, y, L^-) = \exp\left\{-k_1 L \sqrt{\sin^2\theta_1 - \sin^2\theta_*}\right\} E_t(t, x, y, 0) \tag{4.12}$$

- That is, in the approximation where backscatter is neglected, the fraction of electric field that gets through the slab is

$$\exp\left\{-k_1 L \sqrt{\sin^2\theta_1 - \sin^2\theta_*}\right\} \tag{4.13}$$

This sort of exponential suppression of the wave, depending in L the thickness of the intermediate region, should remind you very strongly of quantum mechanical barrier penetration. (Some of you will already have seen the WKB approximation to barrier penetration.)

(Correctly dealing with backscatter from the interfaces is a little tedious, but does not greatly affect the details of the discussion.)

4.4 FTIR in acoustics

Remember that Snell's law also works in acoustics — with the refractive index now being $n = 1/(\text{speed of sound})$. So total internal reflection also works in acoustics, and likewise you would expect FTIR to occur in acoustics.

- Acoustic FTIR should occur in fluid-fluid, fluid-solid, and solid-solid interfaces.
- The fluid-fluid case is likely to be the easiest to analyze (only longitudinal sound, no transverse sound).

4.4.1 TIR in fluid-fluid acoustics

Let's consider total internal reflection of sound waves at a fluid-fluid interface.

- Consider a sound wave in a fluid medium (for example, air, water) where the speed of sound is c_1.

  It is a standard textbook result that the pressure fluctuations are governed by the equation
$$\frac{\partial^2 p}{\partial t^2} = c_1^2 \nabla^2 p. \tag{4.14}$$
  Similarly, in a medium where the speed of sound is c_2 the pressure fluctuations are governed by the equation
$$\frac{\partial^2 p}{\partial t^2} = c_2^2 \nabla^2 p. \tag{4.15}$$

- Now consider a plane wave of angular frequency ω, and wave-vector $\vec{k}$.

  In medium 1 we have
$$\omega^2 = c_1^2 \, ||\vec{k}_1||^2; \tag{4.16}$$
  while in medium 2
$$\omega^2 = c_2^2 \, ||\vec{k}_2||^2. \tag{4.17}$$
  Note that whatever happens to the wave as it crosses from medium 1 to medium 2 the frequency ω must stay the same.

  Exercise: Why? This should be obvious, think about it.

- Deduce
$$c_1 \, ||\vec{k}_1|| = c_2 \, ||\vec{k}_2||. \tag{4.18}$$
  Since we are going to get tired of writing $||\vec{k}||$ all the time, adopt the notation
$$k = ||\vec{k}|| \tag{4.19}$$
  so that
$$c_1 \, k_1 = c_2 \, k_2. \tag{4.20}$$

- Now assume the interface between medium 1 and 2 is the flat plane at $z = 0$. Assume that in medium 1 the plane wave is travelling in the direction
$$\hat{k}_1 = (\sin\theta_1, 0, \cos\theta_1); \tag{4.21}$$
  and that in medium 2 the plane wave is travelling in the direction
$$\hat{k}_2 = (\sin\theta_2, 0, \cos\theta_2). \tag{4.22}$$
  The x-component of $\vec{k}_1$ must equal the x-component of $\vec{k}_2$.

Exercise: Why? This should be obvious, think about it.

That is,

$$(\vec{k}_1)_x = (\vec{k}_2)_x. \tag{4.23}$$

Rearrange this to yield

$$k_1 \ \sin\theta_1 = k_2 \ \sin\theta_2. \tag{4.24}$$

- We have now obtained the two simultaneous equations

$$c_1 \ k_1 = c_2 \ k_2; \tag{4.25}$$

and

$$k_1 \ \sin\theta_1 = k_2 \ \sin\theta_2. \tag{4.26}$$

Solving these linear equations we obtain a version of Snell's law

$$\frac{\sin\theta_1}{c_1} = \frac{\sin\theta_2}{c_2}. \tag{4.27}$$

(The reason this whole discussion is so much simpler than the electromagnetic case is that in fluid acoustics there is only one polarization for the sound waves.)

- Suppose $c_2 > c_1$. Rewrite Snell's law as

$$\sin\theta_2 = \frac{c_2}{c_1} \ \sin\theta_1 \tag{4.28}$$

This implies the existence of a critical angle

$$\theta_* = \sin^{-1}(c_1/c_2) \tag{4.29}$$

such that for $\theta_1 > \theta_*$ there is no physical solution for θ_2 — that is, no refracted ray exists.

(Therefore, $\theta_1 > \theta_*$ corresponds to total internal reflection, which we see can only occur if you are leaving a "slow" medium to enter a "fast" medium.)

- Using the two simultaneous equations

$$c_1 \ k_1 = c_2 \ k_2; \tag{4.30}$$

and

$$k_1 \ \sin\theta_1 = k_2 \ \sin\theta_2. \tag{4.31}$$

it is easy to see that

$$(\vec{k}_2)_z = k_2 \ \cos\theta_2 = k_1 \ \frac{c_1}{c_2} \ \sqrt{1 - \sin^2\theta_2} = k_1 \ \frac{c_1}{c_2} \ \sqrt{1 - \frac{c_2^2}{c_1^2}\sin^2\theta_1} \tag{4.32}$$

finally yielding

$$(\vec{k}_2)_z = k_1 \ \sqrt{\sin^2\theta_* - \sin^2\theta_1} \tag{4.33}$$

What happens to $(\vec{k}_2)_z$ in the situation where $\theta_1 > \theta_*$? It becomes pure imaginary. And note that if $(\vec{k}_2)_z$ is pure imaginary, then $\exp(i[\vec{k}_2]_z \ z)$ is purely real.

- In medium 1 we have by assumption a plane wave of the form

$$p_1 \ \exp\left[-i\{\omega t - \vec{k}_1 \cdot \vec{x}\}\right], \tag{4.34}$$

and in medium 2 a (refracted) plane wave of the form

$$p(t, \vec{x}) = p_2 \ \exp\left[-i\{\omega t - \vec{k}_2 \cdot \vec{x}\}\right]. \tag{4.35}$$

(More precisely, we want the real part of these complex exponentials.)

The measured pressure is after all a real number — but it is a very useful trick to write it as complex exponential:

$$p = A \exp(i\varphi); \qquad p_{\text{physical}} = Re(p) = Re(A)\cos\varphi - Im(A)\sin\varphi$$

- In medium 2 the pressure fluctuation can be more explicitly written (in terms of θ_1 and other quantities measured in medium 1) as:

$$p(t, \vec{x}) = p_2 \ \exp\left[-i\left\{\omega t - ||\vec{k}_1||\left(\sin\theta_1 \ x + \sqrt{\sin^2\theta_* - \sin^2\theta_1} \ z\right)\right\}\right] \tag{4.36}$$

But we still need to determine the amplitude p_2.

- To do this we need to recognize that there will also be a certain amount of reflection from the interface, so that medium 1 will also contain a reflected wave

$$p_1^R \ \exp\left[-i\{\omega t - \vec{k}_1^R \cdot \vec{x}\}\right], \tag{4.37}$$

where

$$\hat{k}_1^R = (\sin\theta_1, 0, -\cos\theta_1) \tag{4.38}$$

and

$$\vec{k}_1^R = k_1 \ \hat{k}_1^R = k_1(\sin\theta_1, 0, -\cos\theta_1) \tag{4.39}$$

represents a ray that is now moving back downwards.

- That is, in medium 1 the total pressure fluctuation is

$$p(t, \vec{x}) = p_1 \ \exp\left[-i\{\omega t - \vec{k}_1 \cdot \vec{x}\}\right] + p_1^R \ \exp\left[-i\{\omega t - \vec{k}_1^R \cdot \vec{x}\}\right]. \tag{4.40}$$

- At the interface we must satisfy the boundary conditions

$$p(t, x, y, 0^-) = p(t, x, y, 0^+) \tag{4.41}$$

and

$$\partial_z p(t, x, y, 0^-) = \partial_z p(t, x, y, 0^+) \tag{4.42}$$

If the first of these conditions is not satisfied then there is an infinite pressure gradient across the interface. (Infinite force $\Rightarrow$ infinite acceleration.) If the second of these conditions is not satisfied, then the wave equation cannot be satisfied at the interface. (Because $\nabla^2 p$ will contain a delta-function contribution).

- The first boundary condition implies

$$p_1 + p_1^R = p_2 \tag{4.43}$$

while the second implies

$$(\vec{k}_1)_z \left[p_1 - p_1^R\right] = (\vec{k}_2)_z \; p_2 \tag{4.44}$$

- To simplify things, let's agree to write

$$\kappa = (\vec{k})_z = \vec{k} \cdot \hat{z} \tag{4.45}$$

so that the second boundary condition becomes

$$\kappa_1 \left[p_1 - p_1^R\right] = \kappa_2 \; p_2 \tag{4.46}$$

- Rearrange the first boundary condition to give

$$p_1^R = p_2 - p_1 \tag{4.47}$$

and substitute into the second boundary condition

$$\kappa_1 \left[2p_1 - p_2\right] = \kappa_2 \; p_2 \tag{4.48}$$

- Rearrange

$$2\kappa_1 p_1 = [\kappa_2 + \kappa_1] p_2 \tag{4.49}$$

- That is

$$p_2 = \frac{2\kappa_1}{\kappa_1 + \kappa_2} \; p_1, \tag{4.50}$$

and consequently

$$p_1^R = \frac{\kappa_1 - \kappa_2}{\kappa_1 + \kappa_2} \; p_1. \tag{4.51}$$

- But we already have

$$\kappa_1 = k_1 \; \cos\theta_1 \tag{4.52}$$

and

$$\kappa_2 = k_2 \; \cos\theta_2 = k_1 \sqrt{\sin^2\theta_* - \sin^2\theta_1} \tag{4.53}$$

so that

$$p_2 = \frac{2 \; \cos\theta_1}{\cos\theta_1 + \sqrt{\sin^2\theta_* - \sin^2\theta_1}} \; p_1 \tag{4.54}$$

which implies that in medium 2

$$p(t, \vec{x}) = p_1 \; \frac{2 \; \cos\theta_1}{\cos\theta_1 + \sqrt{\sin^2\theta_* - \sin^2\theta_1}} \times \exp\left[-i \left\{\omega t - k_1 \left(\sin\theta_1 \; x + \sqrt{\sin^2\theta_* - \sin^2\theta_1} \; z\right)\right\}\right] \tag{4.55}$$

This finally is our complete solution for the pressure fluctuation in medium 2, including the effects of the backscattered reflected wave in medium 1.

- What fraction of the acoustic *power* is transmitted through the interface? It is a standard result that for plane waves the power flux in the $\hat{z}$ direction is proportional to
$$P \propto \mathbf{Re}\left(\frac{p^*\,p}{c}\quad\left[\hat{k}\cdot\hat{z}\right]\right) = \mathbf{Re}\left(\frac{p^*\,p}{c}\;\cos\theta\right) \tag{4.56}$$
Hence the transmitted power is
$$P_2 \propto \mathbf{Re}\left(\frac{p_2^*\,p_2}{c_2}\;\cos\theta_2\right) \propto \mathbf{Re}\left(p_2^*\,p_2\;k_2\cos\theta_2\right) = \mathbf{Re}\left(p_2^*\,p_2\;\kappa_2\right) \tag{4.57}$$
while the incident power is
$$P_1 \propto \mathbf{Re}\left(p_1^*\,p_1\;\kappa_1\right) \tag{4.58}$$

- If $\theta_1 > \theta_0$, so that total internal reflection occurs, then κ_2 is pure imaginary, and $P_2 \to 0$. In general the transmission coefficient is
$$T = \frac{P_2}{P_1} = \frac{\mathbf{Re}\left(p_2^*\,p_2\;\kappa_2\right)}{\mathbf{Re}\left(p_1^*\,p_1\;\kappa_1\right)} = \left|\frac{2\kappa_1}{\kappa_1+\kappa_2}\right|^2\quad\mathbf{Re}\left(\frac{\kappa_2}{\kappa_1}\right) = \frac{4\;\mathbf{Re}\left(\kappa_1\kappa_2\right)}{\left|\kappa_1+\kappa_2\right|^2} \tag{4.59}$$

- We can simplify this to
$$T = \frac{4\;\cos\theta_1\;\mathbf{Re}\left(\sqrt{\sin^2\theta_* - \sin^2\theta_1}\right)}{\left|\cos\theta_1 + \sqrt{\sin^2\theta_* - \sin^2\theta_1}\right|^2} \tag{4.60}$$

- Note that this has all the sensible limits. For $\theta_1 > \theta_*$ we have $T \to 0$, for $c_1 = c_2$ we have $T \to 1$. We can similarly calculate the reflection coefficient
$$R = \frac{P_1^R}{P_1} = \frac{\mathbf{Re}\left([p_1^R]^*\;p_1^R\right)}{\mathbf{Re}\left(p_1^*\;p_1\right)} = \left|\frac{\kappa_1-\kappa_2}{\kappa_1+\kappa_2}\right|^2 \tag{4.61}$$
and verify that $T + R = 1$.

- Note that the reflection and transmission coefficients are *very* similar to quantum mechanical scattering on a one-step potential — a problem that many of you will have seen before.

Exercise: Look up quantum scattering from a one-step potential.

Find formulae for the transmission and reflection coefficients.

Compare them with the above.

Build a suitable translation table for the (two-fluid) $\leftrightarrow$ (one-step potential) problems.

4.4.2 FTIR in fluid-fluid acoustics

- Now introduce a *third* medium, with speed of sound c_3. Assume a third planar interface at $z = L$, this now being an interface between medium 2 and medium 3. (That is, we are considering a three-fluid layered system). It is now trivial to show that

$$\frac{\partial^2 p}{\partial t^2} = c_3^2 \, \nabla^2 p, \tag{4.62}$$

  and for a plane wave

$$\omega^2 = c_3^2 \, ||\vec{k}_3||^2. \tag{4.63}$$

  We now have

$$c_1 \, ||\vec{k}_1|| = c_2 \, ||\vec{k}_2|| = c_3 \, ||\vec{k}_3||; \tag{4.64}$$

  and

$$||\vec{k}_1|| \; \sin\theta_1 = ||\vec{k}_2|| \; \sin\theta_2 = ||\vec{k}_3|| \; \sin\theta_3. \tag{4.65}$$

- In medium 3 we must have a transmitted wave

$$p(t, \vec{x}) = p_3 \; \exp\left[-i\{\omega t - \vec{k}_3 \cdot \vec{x}\}\right], \tag{4.66}$$

  where p_3 is to be determined.

  In medium 2 we now have both transmitted and reflected waves

$$p(t, \vec{x}) = p_2 \; \exp\left[-i\{\omega t - \vec{k}_2 \cdot \vec{x}\}\right] + p_2^R \; \exp\left[-i\{\omega t - \vec{k}_2^R \cdot \vec{x}\}\right], \tag{4.67}$$

  where p_2 and p_2^R are to be determined.

  In medium 1 we still have

$$p(t, \vec{x}) = p_1 \; \exp\left[-i\{\omega t - \vec{k}_1 \cdot \vec{x}\}\right] + p_1^R \; \exp\left[-i\{\omega t - \vec{k}_1^R \cdot \vec{x}\}\right], \tag{4.68}$$

  where p_1 is treated as given, and p_1^R is to be determined.

- We now need to apply boundary conditions at *both* interfaces. At $z = 0$ we find

$$p_1 + p_1^R = p_2 + p_2^R \tag{4.69}$$

$$\kappa_1\{p_1 - p_1^R\} = \kappa_2\{p_2 - p_2^R\} \tag{4.70}$$

  while at $z = L$ we have

$$p_2 \; \exp[i\kappa_2 L] + p_2^R \; \exp[-i\kappa_2 L] = p_3 \; \exp[i\kappa_3 L] \tag{4.71}$$

$$\kappa_2 \left\{p_2 \; \exp[i\kappa_2 L] - p_2^R \; \exp[-i\kappa_2 L]\right\} = \kappa_3 \, p_3 \; \exp[i\kappa_3 L] \tag{4.72}$$

  This gives us 4 linear equations for the 4 unknowns p_1^R, p_2, p_2^R, p_3 in terms of the known variables p_0, κ_1, κ_2, κ_3 and L. Hence we can solve for p_3 (this will be part of the homework, used Maple or Mathematica, or do it by hand) and so determine:

$$p_3 = p_1 \; F(\kappa_1, \kappa_2, \kappa_3, L) \tag{4.73}$$

- The transmission coefficient will be

$$T = \frac{P_3}{P_1} = \frac{\mathbf{Re}\,(p_3^* \, p_3 \, \kappa_3)}{\mathbf{Re}\,(p_1^* \, p_1 \, \kappa_1)} = |F(\kappa_1, \kappa_2, \kappa_3, L)|^2 \quad \mathbf{Re}\left(\frac{\kappa_3}{\kappa_1}\right) \tag{4.74}$$

which expresses everything in terms of the unknown function $F(\kappa_1, \kappa_2, \kappa_3, L)$.

- In the special case $n_3 = n_1$ we also have $\kappa_3 = \kappa_1$ and things simplify a little:

$$T = \frac{P_3}{P_1} = \frac{\mathbf{Re}\,(p_3^* \, p_3 \,)}{\mathbf{Re}\,(p_1^* \, p_1 \,)} = |F(\kappa_1, \kappa_2, L)|^2 \,. \tag{4.75}$$

- If we now explicitly calculate the function $F(\kappa_1, \kappa_2, \kappa_3, L)$, you will find reflection and transmission amplitudes that are completely equivalent to that of a quantum mechanical particle incident on a two-step potential.

4.5 Comments

- Tunnelling is a wave phenomenon, not intrinsically a quantum phenomenon.

- Of course, since it is a wave phenomenon, regardless of the detailed physical situation you are likely to wind up with very similar partial differential equations [PDEs] to solve — typically second-order.

- Once you apply various symmetries, you are likely to be able to reduce things to a one-dimensional ordinary differential equation [ODE], very often of the form:

$$\psi''(x) + K^2(x) \, \psi(x) = 0 \tag{4.76}$$

In quantum mechanics $K(x)$ will depend on the potential $V(x)$ the particle is placed in, whereas for electromagnetism $K(x)$ will depend on the local refractive index. Then the actual mathematical problems to be solved are very similar.

- Where does quantum physics enter? As soon as you adopt the de Broglie hypothesis then particles have wavelike aspects. But, as we have just seen, any wave process, classical or quantum, exhibits tunnelling. So, as soon as you adopt the de Broglie hypothesis, quantum mechanical particles must exhibit tunnelling.

Chapter 5

One-dimensional scattering

Scattering theory in one space dimension is a lovely subject that is mathematically simple, physically transparent, and still contains numerous interesting results. We will be interested in the Schrodinger equation

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi(x) + V(x)\;\Psi(x) = E\;\psi(x). \tag{5.1}$$

in situations where the potential $V(x)$ is zero outside of a finite interval — mathematically we are looking at potentials of compact support.

In any region where the potential is zero we simply need to solve

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi(x) = E\;\psi(x). \tag{5.2}$$

for which the two independent solutions are

$$\exp(\pm ikx); \qquad k = \frac{\sqrt{2mE}}{\hbar} \tag{5.3}$$

or more explicitly

$$\exp\left\{\pm i\frac{\sqrt{2mE}}{\hbar}\;x\right\} \tag{5.4}$$

To the *left* of the potential we have

$$\psi_L(x) = a\exp(ikx) + b\exp(-ikx) \tag{5.5}$$

while to the *right* of the potential we have

$$\psi_R(x) = c\exp(ikx) + d\exp(-ikx) \tag{5.6}$$

Even without knowing anything more about the potential $V(x)$, the linearity of the Schrodinger ODE guarantees that there will be *some* 2×2 matrix M (typically called a transfer matrix) such that

$$\begin{bmatrix} c \\ d \end{bmatrix} = M \begin{bmatrix} a \\ b \end{bmatrix} \tag{5.7}$$

This transfer matrix relates the situation to the *left* of the potential with the wave-function to the *right* of the potential. We might use this formalism, for instance, to think about the propagation of electrons down a wire (approximately one-dimensional) with $V(x)$ used to describe various barriers placed in the path of the electron.

(Similar matrices also occur in classical optics, where they are referred to as "Jones matrices".)

The components of the transfer matrix M will be some horrible nonlinear function of the potential $V(x)$, but by linearity of the Schrodinger ODE these matrix components must be independent of the parameters a, b, c, and d. In some particularly simple situations we may be able to calculate the matrix M explicitly, but in general it will be a complicated mess. Nevertheless we may be able to prove some general theorems about this matrix, and that is what this chapter is all about.

5.1 Physical interpretation of the transfer matrix M

Let's start with a wave moving in from the left

$$\exp(ikx) \tag{5.8}$$

which then hits the potential, is partially reflected and partially transmitted. In this case, on the *left* of the potential we have

$$\psi_L(x) = \exp(ikx) + r_L \exp(-ikx) \tag{5.9}$$

where r_L is the left-moving reflection amplitude and on the *right* of the potential

$$\psi_R(x) = t_L \exp(ikx) \tag{5.10}$$

where t_L is the left-moving transmission amplitude. This tells us that

$$\begin{bmatrix} t_L \\ 0 \end{bmatrix} = M \begin{bmatrix} 1 \\ r_L \end{bmatrix} \tag{5.11}$$

But since the Schrodinger equation is real, the complex conjugate of any solution is also a solution. So the solution which on the left has the form

$$\psi_L(x) = \exp(-ikx) + r_L^* \exp(+ikx) \tag{5.12}$$

must on the right have the form

$$\psi_R(x) = t_L^* \exp(-ikx) \tag{5.13}$$

and so we also have

$$\begin{bmatrix} 0 \\ t_L^* \end{bmatrix} = M \begin{bmatrix} r_L^* \\ 1 \end{bmatrix} \tag{5.14}$$

These two matrix equations now imply

$$M = \frac{1}{1 - r_L^* r_L} \begin{bmatrix} t_L & -t_L r_L^* \\ -t_L^* r_L & t_L^* \end{bmatrix} \tag{5.15}$$

But by conservation of flux we must have[1]

$$|t_L|^2 + |r_L|^2 = 1 \tag{5.16}$$

so

$$\frac{1}{1 - r_L^* r_L} = \frac{1}{1 - |r_L|^2} = \frac{1}{|t_L|^2} \tag{5.17}$$

whence

$$M = \frac{1}{|t_L|^2} \begin{bmatrix} t_L & -t_L r_L^* \\ -t_L^* r_L & t_L^* \end{bmatrix} = \begin{bmatrix} 1/t_L^* & -r_L^*/t_L^* \\ -r_L/t_L & 1/t_L \end{bmatrix} \tag{5.18}$$

Similarly, consider a wave moving in from the right

$$\exp(-ikx) \tag{5.19}$$

which then hits the potential, is partially reflected and partially transmitted. In this case, on the *right* of the potential we have

$$\psi_R(x) = \exp(-ikx) + r_R \exp(+ikx) \tag{5.20}$$

where r_R is the right-moving reflection amplitude and on the *left* of the potential

$$\psi_L(x) = t_R \exp(-ikx) \tag{5.21}$$

where t_R is the left-moving transmission amplitude. This tells us that

$$\begin{bmatrix} r_R \\ 1 \end{bmatrix} = M \begin{bmatrix} 0 \\ t_R \end{bmatrix} \tag{5.22}$$

Again, since the Schrodinger equation is real, the complex conjugate of any solution is also a solution. So the solution which on the left has the form

$$\psi_L(x) = t_R^* \exp(+ikx) \tag{5.23}$$

[1] In particular, if $r = 0$ then $|t| = 1$. Similarly, if $t = 0$ then $|r| = 1$.

must on the right have the form

$$\psi_R(x) = \exp(+ikx) + r_R^* \exp(-ikx) \tag{5.24}$$

whence

$$\begin{bmatrix} 1 \\ r_R^* \end{bmatrix} = M \begin{bmatrix} t_R^* \\ 0 \end{bmatrix} \tag{5.25}$$

But now these two matrix equations imply

$$M = \begin{bmatrix} 1/t_R^* & r_R/t_R \\ r_R^*/t_R^* & 1/t_R \end{bmatrix} \tag{5.26}$$

Combining the information from left moving and right moving cases we have first that

$$t_L = t_R \tag{5.27}$$

and secondly that

$$\frac{r_R}{t_R} = -\frac{r_L^*}{t_L^*} \tag{5.28}$$

implying

$$r_R = -r_L^* \; \frac{t_L}{t_L^*}; \qquad |r_R| = |r_L| \tag{5.29}$$

Note that we *cannot* in general deduce $r_L = r_R$, indeed in general this is false.

So for *any* potential we have

$$T = |t_L|^2 = |t_R|^2; \qquad R = |r_L|^2 = |r_R|^2 \tag{5.30}$$

implying that the transmission and reflection coefficients are independent on whether or not the particle is incident from the left or the right — and we have *not* made any assumption here about any symmetry for the potential $V(x)$ itself. We conclude

$$M = \begin{bmatrix} 1/t^* & -r_L^*/t^* \\ -r_L/t & 1/t \end{bmatrix} = \begin{bmatrix} 1/t^* & r_R/t \\ r_R^*/t^* & 1/t \end{bmatrix}. \tag{5.31}$$

Note that the M-matrix has the general form

$$M = \begin{bmatrix} \alpha^* & \beta^* \\ \beta & \alpha \end{bmatrix}. \tag{5.32}$$

Exercise: Verify that with this notation

$$|\alpha|^2 - |\beta|^2 = 1. \tag{5.33}$$

Exercise: Explore the relationship between the M-matrix and the so-called Bogoliubov coefficients that you may have run across in other contexts.

5.1.1 Special case: Definite parity

A special case arises if the potential happens to have definite parity

$$V(-x) = \pm V(x), \tag{5.34}$$

then whenever $\psi(x)$ solves the Schrodinger ODE, so does $\psi(-x)$. But this means that the solution whose left and right limits are

$$\psi_L(x) = \exp(+ikx) + r_L \exp(+ikx) \tag{5.35}$$

$$\psi_R(x) = t_L \exp(+ikx) \tag{5.36}$$

gives rise to

$$\psi_L(-x) = \exp(-ikx) + r_L \exp(+ikx) = \tilde{\psi}_R(x) \tag{5.37}$$

$$\psi_R(-x) = t_L \exp(-ikx) = \tilde{\psi}_L(x) \tag{5.38}$$

whence

$$r_R = r_L; \qquad\qquad t_R = t_L. \tag{5.39}$$

So a potential of definite parity (either even or odd) *does* have simple left-right symmetry in the scattering *amplitudes* (not just the scattering *coefficients*). In this case we have

$$(r/t)^* = -r/t \tag{5.40}$$

implying that r/t is pure imaginary — the amplitudes r and t must be 90° out of phase,

$$r = \pm i|r| \, \frac{t}{|t|}, \tag{5.41}$$

and

$$M = \begin{bmatrix} 1/t^* & -r^*/t^* \\ -r/t & 1/t \end{bmatrix} = \begin{bmatrix} 1/t^* & r/t \\ r^*/t^* & 1/t \end{bmatrix}. \tag{5.42}$$

Notation: Remember that for an arbitrary complex number $z = x + iy = re^{i\phi}$.

The modulus is $r = \sqrt{x^2 + y^2}$ and the phase is $\phi = \tan^{-1}(y/x)$.

5.2 Simple examples

Here's a few cases where everything can be solved analytically.

5.2.1 Delta-function potential

Take $V(x) = \Gamma\,\delta(x)$, and consider a wave incident from the left

$$\psi_L(x) = \exp(+ikx) + r\exp(-ikx); \qquad \psi_R(x) = t\exp(+ikx) \tag{5.43}$$

with

$$k = \frac{\sqrt{2mE}}{\hbar} \tag{5.44}$$

Then the wave-function must be continuous at $x = 0$ and integrating the Schrodinger equation across the delta function

$$-\frac{\hbar^2}{2m}\left[\psi'(0^+) - \psi'(0^-)\right] + \Gamma\psi(0) = 0 \tag{5.45}$$

This implies

$$1 + r = t \tag{5.46}$$

and

$$-\frac{\hbar^2}{2m}\left[ik(1-r) - ikt\right] + \Gamma t = 0 \tag{5.47}$$

whence

$$-2ik(1-t) + \frac{2m\Gamma}{\hbar^2}t = 0 \tag{5.48}$$

That is

$$t = \frac{2ik}{2ik + 2m\Gamma/\hbar^2} = \frac{k}{k - im\Gamma/\hbar^2} \tag{5.49}$$

and the transmission coefficient is

$$T = \frac{k^2}{k^2 + (m\Gamma/\hbar^2)^2} \tag{5.50}$$

Note that $T(k=0) = 0$ and that as the momentum increases $T(k \to \infty) \to 1$ smoothly and monotonically. Such simple smooth monotonic behaviour is actually the exception, not the rule.

Note that

$$r = t - 1 = \frac{-im\Gamma/\hbar^2}{k - im\Gamma/\hbar^2} \tag{5.51}$$

so that r and t are 90° out of phase, as expected for this even parity potential.

Let's write

$$k_0 = \frac{m\Gamma}{\hbar^2} \tag{5.52}$$

so that

$$t = \frac{k}{k - ik_0}; \qquad r = \frac{-ik_0}{k - ik_0} \tag{5.53}$$

and we see

$$M = \begin{bmatrix} 1 + ik_0/k & -ik_0/k \\ ik_0/k & 1 - ik_0/k \end{bmatrix} = \mathbf{I} + \frac{ik_0}{k} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \tag{5.54}$$

Exercise: Calculate $\det(M)$ and $\mathrm{tr}(M)$ for this particular potential. ♢

5.2.2 Two delta-function potentials

Now consider the potential

$$V(x) = \Gamma\,[\delta(x-a) + \delta(x+a)] \tag{5.55}$$

consisting of two symmetrically placed delta functions. For a particle incident from the left we now have

$$\psi_L(x) = \exp(+ikx) + r\exp(-ikx); \qquad \psi_R(x) = t\exp(+ikx); \tag{5.56}$$

$$\psi_{inner}(x) = A\exp(+ikx) + B\exp(-ikx); \tag{5.57}$$

where now we also have to consider the inner region between the two delta-functions. Applying the same sort of boundary conditions we now have four equations. From continuity at $x = -a$ we have

$$\exp(-ika) + r\exp(+ika) = A\exp(-ika) + B\exp(+ika) \tag{5.58}$$

while continuity at $x = +a$ implies

$$A\exp(+ika) + B\exp(-ika) = r\exp(+ika) \tag{5.59}$$

Integrating across the delta functions leads to

$$-\frac{\hbar^2}{2m}\{ik[\exp(-ika) - r\exp(+ika)] - ik[A\exp(-ika) - B\exp(+ika)]\} \tag{5.60}$$

$$+\,\Gamma[A\exp(-ika) + B\exp(+ika)] = 0 \tag{5.61}$$

and

$$-\frac{\hbar^2}{2m}\{ik[A\exp(+ika) - B\exp(-ika)] - ik[t\exp(+ika)]\} + \Gamma[t\exp(+ika)] = 0 \tag{5.62}$$

To reduce clutter it is again useful to define the constant

$$k_0 = \frac{m\Gamma}{\hbar^2} \tag{5.63}$$

in which case our four boundary conditions become

$$\exp(-ika) + r\exp(+ika) = A\exp(-ika) + B\exp(+ika) \tag{5.64}$$

$$A\exp(+ika) + B\exp(-ika) = t\exp(+ika) \tag{5.65}$$

$$\begin{aligned}\{ik[\exp(-ika) - r\exp(+ika)] - ik[A\exp(-ika) - B\exp(+ika)]\}\\ = 2k_0[A\exp(-ika) + B\exp(+ika)]\end{aligned} \tag{5.66}$$

$$\{ik[A\exp(+ika) - B\exp(-ika)] - ik[t\exp(+ika)]\} = 2k_0[t\exp(+ika)] \tag{5.67}$$

These are 4 simultaneous linear equations for four unknowns: r, A, B, and t (in terms of the known quantities k, k_0, and a). These can be solved, either by brute force or by Maple or something similar. A tedious little exercise then leads to

$$t = \frac{k^2}{(k - ik_0)^2 + k_0^2\ \exp(4ika)} \tag{5.68}$$

$$r = \frac{2ik_0[k\ \cos(2ka) - k_0\ \sin(2ka)]}{(k - ik_0)^2 + k_0^2\ \exp(4ika)} \tag{5.69}$$

Note that r/t is pure imaginary, in agreement with our general argument regarding definite parity potentials. Furthermore note that

$$R = |r|^2 \propto [k\ \cos(2ka) - k_0\ \sin(2ka)]^2 \tag{5.70}$$

and so $R = 0$ whenever

$$\tan(2ka) = \frac{k}{k_0} \tag{5.71}$$

That is, the system exhibits "transmission resonances" where $T \to 1$ and $R \to 0$. If we work at fixed energy then these resonances occur at equally spaced spatial separation for the two delta functions, namely:

$$a_{\text{resonance}} = \frac{1}{2k}\left\{\tan^{-1}(k/k_0) + n\pi\right\}; \qquad n \in Z \tag{5.72}$$

If instead we hold a fixed and vary k then the location of the resonances is determined by the transcendental equation

$$k_0\ \tan(2ka) = k \tag{5.73}$$

and there is no simple formula explicit for finding $k_{\text{resonance}}(k_0, a, n)$, though there are various graphical and approximation techniques that give useful information. The existence of these "transmission resonances" in one-dimensional scattering is in fact very widespread, it's not specific to this particular example. A brief computation leads to the explicit transmission coefficient

$$T = \frac{k^4}{k^4 + 4k_0^2[k\cos(2ka) - k_0\sin(2ka)]^2} \tag{5.74}$$

which agrees on the location of the transmission resonances. Finally note $T(k \to 0) \to 0$ and $T(k \to \infty) \to 1$. After we look at one more illustrative example we'll return to the issue of obtaining some general theorems governing one-dimensional scattering.

Exercise: Calculate the transfer matrix M for this particular potential.

Now calculate $\det(M)$ and $\mathrm{tr}(M)$ for this particular potential. ♢

5.2.3 Two-step potential

Let $\Theta(\cdot)$ be the Heavyside function (step function) and consider the potential

$$V(x) = V_0\,\Theta(a - |x|) \tag{5.75}$$

which has width $2a$ and is zero for $|x| > a$ and equals V_0 for $|x| < a$. (I call this two-step because it's one step up and then one step down.) Solving the scattering problem for this potential is a standard textbook exercise. To the left and right of the barrier we have

$$\psi_L(x) = \exp(+ikx) + r\exp(-ikx); \qquad \psi_R(x) = t\exp(+ikx); \tag{5.76}$$

while inside the barrier itself

$$\psi_{inner}(x) = A\exp(+ik_0 x) + B\exp(-ik_0 x); \tag{5.77}$$

where k_0 is defined by

$$k_0 = \frac{\sqrt{2m(E - V_0)}}{\hbar} \tag{5.78}$$

and we allow the possibility that k_0 is pure imaginary (whenever $E < V_0$).

The junction conditions at $x = \pm a$ come from the fact that the wavefunction and its derivative must be continuous at those points. They are similar to (but distinct from) the previous example and are easily seen to be

$$\exp(-ika) + r\exp(+ika) = A\exp(-ik_0 a) + B\exp(+ik_0 a) \tag{5.79}$$

$$A\exp(+ik_0 a) + B\exp(-ik_0 a) = t\exp(+ika) \tag{5.80}$$

$$k[\exp(-ika) - r\exp(+ika)] = k_0[A\exp(-ik_0 a) - B\exp(+ik_0 a)] \tag{5.81}$$

$$k_0[A\exp(+ik_0 a) - B\exp(-ik_0 a)] = k[t\exp(+ika)] \tag{5.82}$$

These are 4 simultaneous linear equations for the four unknowns r, A, B, and t (in terms of the known quantities k, k_0, and a). These can be solved, either by brute force or by Maple or something similar. A tedious little exercise then leads to

$$t = \frac{4kk_0\exp(-2i[k - k_0]a)}{(k + k_0)^2 - (k - k_0)^2\exp(4ik_0 a)} \tag{5.83}$$

$$r = \frac{-2i(k^2 - k_0^2)\sin(2k_0 a)\exp(-2i[k-k_0]a)}{(k+k_0)^2 - (k-k_0)^2\exp(4ik_0 a)} \tag{5.84}$$

Again, note that r/t is pure imaginary as expected for a definite parity potential. Note that the reflection coefficient satisfies

$$R = |r|^2 \propto \sin^2(2k_0 a) \tag{5.85}$$

so that the barrier becomes completely transparent whenever

$$\sin(2k_0 a) = 0; \qquad a = \frac{n\pi}{k}; \qquad n \in Z \tag{5.86}$$

another example of "transmission resonance". A brief computation leads to the explicit transmission coefficient

$$T = \frac{4k^2 k_0^2}{4k^2 k_0^2 + (k^2 - k_0^2)^2 \ \sin^2(2ak_0)} \tag{5.87}$$

Note that this agrees on the location of the transmission resonances — at $\sin(2k_0 a) = 0$. Finally note that $T(k \to 0) \to 0$ and $T(k \to \infty) \to 1$.

Exercise: Calculate the transfer matrix M for this particular potential.

Now calculate $\det(M)$ and $\mathrm{tr}(M)$ for this particular potential. ♢

Now that we've seen these illustrative examples, we'll use them as a guide as we return to the issue of obtaining general theorems governing one-dimensional scattering.

5.3 Some general theorems

5.3.1 Translation

What happens to the transfer matrix M if I shift the potential? That is, consider the shift

$$V(x) \to \tilde{V}(x) = V(x-a) \tag{5.88}$$

Then a solution $\psi(x)$ of the original problem becomes a solution $\tilde{\psi}(x) = \psi(x-a)$ of the shifted problem. Look at what this does to $\psi_L(r)$ and $\psi_R(r)$. We deduce

$$\tilde{t} = t; \qquad \tilde{r}_L = \exp(+2iak)\ r_L; \qquad \tilde{r}_R = \exp(-2iak)\ r_R; \tag{5.89}$$

That is, shifting the position of the potential does not affect the transmission amplitude, but does adjust the *phase* of the reflection amplitude. That is, for the transfer matrix

$$M_a = \begin{bmatrix} 1/t^* & -\exp(-2iak)\ r_L^*/t^* \\ -\exp(+2iak)\ r_L/t & 1/t \end{bmatrix} \tag{5.90}$$

or equivalently

$$M_a = \begin{bmatrix} 1/t^* & \exp(-2iak)\; r_R/t \\ \exp(+2iak)\; r_R^*/t^* & 1/t \end{bmatrix}. \tag{5.91}$$

We can also write this in terms of matrix multiplication as

$$M_a = \begin{bmatrix} \exp(-iak) & 0 \\ 0 & \exp(+iak) \end{bmatrix} M_0 \begin{bmatrix} \exp(+iak) & 0 \\ 0 & \exp(-iak) \end{bmatrix}. \tag{5.92}$$

Thus, although in general one cannot deduce $r_L = r_R$ (except when the potential has definite parity), this translation trick is sufficient to guarantee that there will be *some* value of a for which the translated reflection amplitudes satisfy

$$\tilde{r}_L = \tilde{r}_R \tag{5.93}$$

and once this is done, it follows that $\tilde{r}/t$ is pure imaginary, as for potentials of definite parity. If we have adjusted the location of the potential to force $\tilde{r}_L = \tilde{r}_R$ then we shall say that the barrier is in standard position (more precisely, one of its standard positions).

5.3.2 Composition

Suppose now that the potential is the disjoint sum of two (disjoint) components

$$V(x) = V_1(x) + V_2(x) \tag{5.94}$$

with $V_1(x)$ on the left of $V_2(x)$. Then transmission through the compound potential can be considered at two separate processes. (For instance, transmission from the "left" region to the "central" region, described by $V_1(x)$, followed by transmission from the "central" region to the "right" region, described by potential V_2.) Then for the transfer matrices we have

$$M = M_2\; M_1 \tag{5.95}$$

That is: As long as the potentials are disjoint, we can simply multiply the transfer matrices.

Example:
For a single delta-function at the origin

$$M = \mathbf{I} + \frac{ik_0}{k} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \tag{5.96}$$

Translating the delta function to $\pm a$ gives

$$M_{+a} = \mathbf{I} + \frac{ik_0}{k} \begin{bmatrix} 1 & -\exp(-2iak) \\ \exp(+2iak) & -1 \end{bmatrix} \tag{5.97}$$

$$M_{-a} = \mathbf{I} + \frac{ik_0}{k}\begin{bmatrix} 1 & -\exp(+2iak) \\ \exp(-2iak) & -1 \end{bmatrix} \tag{5.98}$$

So for the two-delta function potential we have

$$M = M_{+a} M_{-a}, \tag{5.99}$$

which can easily be computed by multiplying 2×2 matrices.

$$\left\{\mathbf{I} + \frac{ik_0}{k}\begin{bmatrix} 1 & -\exp(-2iak) \\ \exp(+2iak) & -1 \end{bmatrix}\right\} \left\{\mathbf{I} + \frac{ik_0}{k}\begin{bmatrix} 1 & -\exp(+2iak) \\ \exp(-2iak) & -1 \end{bmatrix}\right\}. \tag{5.100}$$

Indeed, simply by looking at the bottom right corner

$$\frac{1}{t} = 1 - \frac{2ik_0}{k} - \frac{k_0^2}{k^2}[1 - \exp(4iak)]. \tag{5.101}$$

That is

$$t = \frac{k^2}{(k - ik_0)^2 + k_0^2 \exp(4iak)}, \tag{5.102}$$

which is exactly the same result we got from directly solving four simultaneous linear equations.

5.3.3 Transmission resonances

We are now ready to prove a general theorem on transmission resonances. Consider an arbitrary potential with compact support, and using the translation property of the reflection amplitude, place the potential in standard position. Call this potential $V_0(x)$, and write the transfer matrix as

$$M_0 = \begin{bmatrix} 1/t_0^* & -r_0^*/t_0^* \\ -r_0/t_0 & 1/t_0 \end{bmatrix} = \begin{bmatrix} 1/t_0^* & r_0/t_0 \\ r_0^*/t_0^* & 1/t_0 \end{bmatrix}. \tag{5.103}$$

where r_0/t_0 is pure imaginary.

Now take a second copy of this same potential and translate it a distance a to the right, producing a potential $V_a(x)$. Then

$$M_a = \begin{bmatrix} 1/t_0^* & -\exp(-2iak)\, r_0^*/t_0^* \\ -\exp(+2iak)\, r_0/t_0 & 1/t_0 \end{bmatrix} \tag{5.104}$$

As long as we shift far enough that the potential $V_a(x)$ does not overlap with the potential $V_0(x)$ we will have

$$M = M_a\, M_0 \tag{5.105}$$

That is

$$M = \begin{bmatrix} 1/t_0^* & -\exp(-2iak)\; r_0^*/t_0^* \\ -\exp(+2iak)\; r_0/t_0 & 1/t_0 \end{bmatrix} \begin{bmatrix} 1/t_0^* & -r_0^*/t_0^* \\ -r_0/t_0 & 1/t_0 \end{bmatrix} \tag{5.106}$$

That is

$$M = \begin{bmatrix} 1/t_0^{*2} + \exp(-2iak)|r_0^2/t_0^2| & -(r_0/t_0^{*2})\{1 + [t_0^*/t_0]\exp(-2iak)\} \\ -(r_0/t_0^2)\{1 + [t_0/t_0^*]\exp(+2iak)\} & 1/t_0^2 + \exp(+2ika)|r_0/t_0|^2 \end{bmatrix} \tag{5.107}$$

By looking at the bottom left element of the compound matrix we see

$$\frac{r}{t} = \frac{r_0}{t_0^2}\left\{1 + \frac{t_0}{t_0^*}\exp(+2iak)\right\} \tag{5.108}$$

Let the phase of t_0 be ϕ_0, that is $t_0 = |t_0|\ \exp(i\phi_0)$.[2] Then

$$r \propto 1 + \exp(2i[\phi_0 + ak]) \tag{5.109}$$

Therefore, we can always make the (compound) reflection coefficient r vanish by picking a value of a such that

$$2[\phi_0 + ak] = (2n+1)\pi \tag{5.110}$$

that is

$$a = \frac{(n+\frac{1}{2})\pi - \phi_0}{k}; \qquad n \in Z \tag{5.111}$$

Note the very general nature of this result — *any* compound barrier constructed out of two copies of the same potential will have transmission resonances as you vary the distance between the two copies of the barrier.

By looking at the bottom right element of the compound matrix we see

$$\frac{1}{t} = \frac{1}{t_0^2} + \frac{r\, r^*}{t\, t^*}\exp(+2iak) \tag{5.112}$$

which can be rearranged to yield

$$t = \frac{T_0 \exp(2i\phi_0)}{1 + (1 - T_0)\exp(2i[\phi_0 + ak])} \tag{5.113}$$

and leads (after a bit of algebra) to the transmission coefficient

$$T = \frac{T_0^2}{T_0^2 + 4R_0\ \cos^2(\phi_0 + ka)} \tag{5.114}$$

The message to take from this is that you can say an awful lot without knowing much about the details of the potential $V(x)$.

[2] Whence, if we were to assume a potential of definite parity, $r_0 = i|r_0|\ \exp(i\phi_0)$. But we have no need, nor any particular desire, to make such an assumption.

5.4 Lessons

- One-dimensional scattering provides a lot of useful examples.
- There is a interesting and general mathematical structure lying behind all of the special case examples you may have run across as examples in other courses.

Chapter 6

The scattering matrix in one-dimension

In the previous chapter we considered the transfer matrix M, and developed some general theorems regarding its behaviour. In this chapter we will briefly consider the related concept of the scattering matrix — the S matrix.

6.1 Physical interpretation of the S-matrix

The idea now is to rephrase the discussion in terms of incoming and outgoing waves instead of left-moving and right-moving waves. To the *left* of the barrier we have:

$$\psi_L(x) = a_L^{in} \exp(ikx) + a_L^{out} \exp(-ikx) \tag{6.1}$$

while to the *right* of the potential we have

$$\psi_R(x) = a_R^{out} \exp(ikx) + a_R^{in} \exp(-ikx) \tag{6.2}$$

Even without knowing anything more about the potential $V(x)$, the linearity of the Schrodinger ODE guarantees that there will be *some* 2×2 scattering matrix S such that

$$\begin{bmatrix} a_R^{out} \\ a_L^{out} \end{bmatrix} = S \begin{bmatrix} a_L^{in} \\ a_R^{in} \end{bmatrix} \tag{6.3}$$

This scattering matrix relates the *incoming* modes (directed *toward* the potential) to the *outgoing* modes (directed *away* from the potential). It is clear that the elements of the S-matrix will be related to the elements of the M matrix, and that many statements made about the transfer matrix can be carried over to the S-matrix. (The physics must ultimately be independent of whether you chose to work with the S-matrix or the transfer matrix; of course some questions might be easier to deal with in one formalism or the other.)

Let's start with a wave moving in from the left

$$\exp(ikx) \tag{6.4}$$

which then hits the potential, is partially reflected and partially transmitted. In this case, on the *left* of the potential we have

$$\psi_L(x) = \exp(ikx) + r_L \exp(-ikx) \tag{6.5}$$

where r_L is the left-moving reflection amplitude and on the *right* of the potential

$$\psi_R(x) = t_L \exp(ikx) \tag{6.6}$$

where t_L is the left-moving transmission amplitude. This tells us that

$$\begin{bmatrix} t_L \\ r_L \end{bmatrix} = S \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{6.7}$$

But since the Schrodinger equation is real, the complex conjugate of any solution is also a solution. So the solution which on the left has the form

$$\psi_L(x) = \exp(-ikx) + r_L^* \exp(+ikx) \tag{6.8}$$

must on the right have the form

$$\psi_R(x) = t_L^* \exp(-ikx) \tag{6.9}$$

and so we also have

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = S \begin{bmatrix} r_L^* \\ t_L^* \end{bmatrix} \tag{6.10}$$

These two matrix equations now imply

$$S = \begin{bmatrix} t_L & -r_L^*(t_L/t_L^*) \\ r_L & t_L \end{bmatrix} \tag{6.11}$$

Similarly, consider a wave moving in from the right

$$\exp(-ikx) \tag{6.12}$$

which then hits the potential, is partially reflected and partially transmitted. In this case, on the *right* of the potential we have

$$\psi_R(x) = \exp(-ikx) + r_R \exp(+ikx) \tag{6.13}$$

where r_R is the right-moving reflection amplitude and on the *left* of the potential

$$\psi_L(x) = t_R \exp(-ikx) \tag{6.14}$$

where t_R is the left-moving transmission amplitude. This tells us that

$$\begin{bmatrix} r_R \\ t_R \end{bmatrix} = S \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{6.15}$$

Again, since the Schrodinger equation is real, the complex conjugate of any solution is also a solution. So the solution which on the left has the form

$$\psi_L(x) = t_R^* \exp(+ikx) \tag{6.16}$$

must on the right have the form

$$\psi_R(x) = \exp(+ikx) + r_R^* \exp(-ikx) \tag{6.17}$$

whence

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = S \begin{bmatrix} t_R^* \\ r_R^* \end{bmatrix} \tag{6.18}$$

But now these two matrix equations imply

$$S = \begin{bmatrix} t_R & r_R \\ -r_R^*(t_R/t_R^*) & t_R \end{bmatrix} \tag{6.19}$$

Combining the information from left moving and right moving cases we obtain the *same* results as we deduced via the transfer matrix. First, we have that

$$t_L = t_R \tag{6.20}$$

and secondly that

$$\frac{r_R}{t_R} = -\frac{r_L^*}{t_L^*} \tag{6.21}$$

implying

$$r_R = -r_L^* \frac{t_L}{t_L^*}; \qquad |r_R| = |r_L| \tag{6.22}$$

Note that we *cannot* in general deduce $r_L = r_R$, indeed in general this is false.

So for *any* potential we have

$$T = |t_L|^2 = |t_R|^2; \qquad R = |r_L|^2 = |r_R|^2 \tag{6.23}$$

implying that the transmission and reflection coefficients are independent on whether or not the particle is incident from the left or the right — and we have *not* made any assumption here about any symmetry for the potential $V(x)$ itself. We conclude

$$S = \begin{bmatrix} t & -r_L^*(t/t^*) \\ r_L & t \end{bmatrix} = \begin{bmatrix} t & r_R \\ -r_R^*(t/t^*) & t \end{bmatrix}. \tag{6.24}$$

Exercise: Show that the S-matrix is unitary. That is, define

$$S^\dagger = [S^*]^T, \tag{6.25}$$

and show

$$S\, S^\dagger = \mathbf{I}; \qquad \text{that is} \qquad S^{-1} = S^\dagger. \tag{6.26}$$

$\diamondsuit$

6.2 Lessons

- The S-matrix represents the same physics as the transfer matrix.
- There is again a deep mathematical structure underlying the specific examples you may have encountered in other courses.
- But I think we have now gone far enough for this particular course....

Chapter 7

Coda

Between these notes and the homework exercises I hope you now have a good feel for the mathematical structure of quantum physics — and hope that you'll be interested in learning more about it.

Cheers
Matt Visser
26 February 2013