

# Chapter 1

## The Story

Following the work of Thompson (See "The Origins of Dalton's Chemical Atomic Theory"), physicists came to the conclusion that matter was composed of atoms, fused into molecules. In 1897 (see "Thompson's 1906 Nobel Lecture: Carriers of Negative Electricity"), physicist J.J. Thompson argued that all negative electrical charge was carried by particles carrying both mass and electricity. Thus the first elementary particle was discovered. Using cathode rays, he measured the ratio between the unit of electrical charge  $e$  and the mass  $m$  of an electron. In 1913, Millikan found an experiment to measure the electrical charge individually. The mass is on the order of  $9 \times 10^{-31}$  kilograms, and the electrical charge is on the order of  $2 \times 10^{-19}$  Coulombs. The resolution of current measurement technology still leads us to believe that the electron is a *point particle*, i.e. it has no extent, existing in a single point in space.

Accumulated matter, like solid materials or gasses, on averages carries little to no electrical charge, or objects at the human scale would constantly be repulsed from one another. Thus there must exist matter with positive charge. Thompson speculated that electrons were embedded in a uniform sphere of jelly-like positive charge – the *plum pudding model*. However, Rutherford proved in 1903 that the carriers of positive charge must also behave like points; by firing a positively charged helium nuclei (an alpha particle) through a thin metal foil. The metal foil is much heavier than the alpha particle, The even distribution of positive charge lead Rutherford to expect that the helium nuclei would pass through the foil. However, Rutherford instead observed that the helium atoms scattered in various directions – distributed in the same way as would be expected via the Kepler problem, where the trajectories are hyperbolas.

It follows from Rutherford's experiment that the positive charge in the metal

foil must be concentrated in very small regions of space; Rutherford had discovered that electrons existed separately from positively charged nuclei. Each nucleus can have various different types of charge, itself being formed from several more elementary particles; positively charged protons, and chargeless neutrons. We often let  $z$  denote the number of protons in a nucleus. In nature,  $z$  lies between  $\{1, \dots, 92\} - \{43, 61, 85\}$ ; the other values  $\{43, 61, 85, 93, \dots, 109\}$  have all been produced in experiments, but decay very quickly. The values  $\{84, \dots, 92\}$  are also unstable, decaying over time, and producing *radioactive energy*.

Though consisting of several elementary particles, a nucleus itself is not a point mass. Nonetheless, in practice the protons and neutrons are concentrated in a region of radius  $10^{-15}$  meters, which is often much smaller than the scale of the problem under consideration, and so a nucleus can also be treated as a point mass. The mass of a proton is on the order of  $2 \times 10^{-27}$  kilograms, more than 1000 times the mass of an electron. The mass of a naturally occurring nucleus is observed to have mass between  $z$  protons, and  $3z$  protons. Thus a nucleus is in practice much much heavier than an electron, and can be considered infinite for most points, i.e. unmovable relative to the motion of an electron, just like the sun is often regarded as having infinite mass relative to the motion of planets. The protons repel one another, but are held together by the *nuclear force* between themselves, and the neutrons in the nucleus, and is about 100 times stronger than the corresponding electrical force. Thus for most of our purposes it is natural to consider a nucleus as a point particle with infinite mass.

With the plum pudding model thrown out, it was reasonable to try and apply the theory of electrostatics to electrons and nuclei, existing apart from one another; electrostatics tells us that the potential energy between two particles with charges  $q_1$  and  $q_2$  existing at two points  $x, y \in \mathbb{R}^3$  is given by

$$k \frac{q_1 q_2}{|x - y|},$$

where  $k = 9 \times 10^9 Nm^2/C^2$  is Coulomb's constant.

Consider the simplest possible case; the analysis of the Hydrogen atom, with one electron, and one proton. Placing the proton at the origin, and suppose the electron is at a position  $p$ , with momentum  $q$ , the Hamiltonian of the system is given by

$$H(p, q) = \frac{p^2}{2m} - \frac{ke^2}{|x|}.$$

This Hamiltonian is, up to constants, identical to the celestial two body problem; Kepler's analysis thus shows that stable orbits of the electron around the Helium

nucleus are ellipses, with the nucleus at one focus. As the energy of the atom decreases, the electron gets closer and closer to the nucleus.

This analysis, however, results in several paradoxes. Firstly, an orbiting electron about a proton is an accelerating, charged mass. Maxwell theory of electromagnetism tells us that such an object should emit electromagnetic energy away from the mass. As the energy decreases, the electron orbits closer and closer to the nucleus of the atom. But the electron can get arbitrarily close to the proton, since both are point masses, and as this happens the energy of the atom can become arbitrarily negative. But this means that hydrogen atoms would continuously emit an infinite amount of energy over time; matter in the universe would always tend to shrink indefinitely in size, emitting electromagnetic waves as they do so. But we don't observe this happening.

Quantum mechanics fixes this problem. As an electron is forced to get closer and closer to a proton, the *uncertainty principle* tells us the kinetic energy of the electron is forced to increase. Thus the total energy of the hydrogen atom actually increases to  $\infty$  as the electron gets closer and closer to the atom.

# Chapter 2

## The Setup

In any physical theory, we must characterize mathematically the *state* of a system (all information describing the situation of a physical system at a particular time), and the *observables*, the functions of a state, which give ways in which the state of a system can be reduced to quantities that can be observed experimentally. For instance, in Hamiltonian classical mechanics, the state of a system is given by a point in a symplectic manifold  $M$ , and observables given by functions  $f : M \rightarrow \mathbb{R}$ , which should be continuous if we are to correctly measure these observables up to a small degree of error. The observables are then ‘second order’ as they are defined in terms of states, but we can also reverse the situation, describing the observables as the  $C^*$  algebra  $A = C(M)$ . The states then become precisely a *positive* linear functional  $\phi : A \rightarrow \mathbb{R}$  with  $\phi(1) = 1$ . It is natural in the later quantum mechanics to complexify the  $C^*$  algebra  $A$ . Then the observables become the *self-adjoint* elements of  $A$ , and the states the linear functionals  $\phi : A \rightarrow \mathbb{C}$  with  $\phi(1) = 1$  and with  $\phi(X) \geq 0$  if  $X \geq 0$ . The Riesz representation theorem allows us to identify an arbitrary positive linear functional  $\phi : A \rightarrow \mathbb{R}$  such that  $\phi(1) = 1$  with a Borel probability measure  $\mu$  on  $M$ . We then think of an element  $X \in A$  as a *random variable* over the probability space  $(M, \mu)$ , because we then have

$$\mathbb{E}_\phi[X] = \int X \, d\mu = \phi(X).$$

Similarly,

$$\sigma_\phi(X)^2 = \mathbb{V}_\phi(X) = \phi(X^2) - \phi(X)^2.$$

The *pure*, deterministic states  $\phi$  can then be identified from general *mixed states* as those states such that  $\mathbb{V}_\phi(X) = 0$  for all observables  $X$ .

What caused this formulation to fail to explain quantum mechanical phenomena. The most fundamental experimental observation in the theory is the *uncertainty principle*. It is an experimental observation that in any physical system, if  $p : A \rightarrow \mathbb{C}$  and  $q : A \rightarrow \mathbb{C}$  are the position and momentum observables, then for any state  $\phi$ ,

$$2\sigma_\phi(p)\sigma_\phi(q) \geq \hbar,$$

where  $\hbar$  is *Planck's constant*. But there are no two observables  $\phi : A \rightarrow \mathbb{R}$  with this property for all classical states, because  $\sigma_\phi(p) = \sigma_\phi(q) = 0$  for any deterministic state. Thus it appears that the only physically possible states  $\phi$  *must be uncertain* in a suitable sense; this is the *uncertainty principle*.

In the standard theory, this is remedied by replacing the observables of a system with elements of an abstract  $C^*$  algebra  $A$ , and the states with normalized, positive linear functions  $\phi : A \rightarrow \mathbb{C}$ . Each fixed state  $\phi$  then induces an algebra homomorphism  $\Phi$  from  $A$  to the family of random variables in an appropriate probability space, such that  $\mathbb{E}(\Phi(X)) = \phi(X)$ . Thus one can use the spectral calculus to obtain detailed information about the probability distribution of  $\Phi(X)$ , since for any continuous  $f : \sigma(X) \rightarrow \mathbb{C}$ , we have  $\mathbb{E}(f(X)) = \phi(f(X))$ . Note, in particular, that this means that the support of the random variable  $X$  is on  $\sigma(X)$ .

The reason this formulation is useful is that we can theoretically derive the uncertainty principle, provided we are working in a *non-commutative*  $C^*$  algebra  $A$ . Indeed, if  $X$  and  $Y$  are any observables with  $\phi(X) = \phi(Y) = 0$ , we calculate that the matrix

$$M = \begin{pmatrix} \phi(X^2) & (1/2)\phi(i[X, Y]) \\ (1/2)\phi(i[X, Y]) & \phi(Y^2) \end{pmatrix}$$

is positive-semidefinite, since for any  $v = (\alpha, \beta)^T \in \mathbb{R}$ ,

$$v^T M v = \phi(X^2)\alpha^2 + \phi(i[X, Y])\alpha\beta + \phi(Y^2)\beta^2 = \phi((\alpha X - i\beta Y)(\alpha X + i\beta Y)) \geq 0.$$

Thus  $\det(M) = \phi(X^2)\phi(Y^2) - \phi(i[X, Y])^2/4$  is non-negative, which means that

$$2\sigma_\phi(X)\sigma_\phi(Y) = 2\phi(X^2)^{1/2}\phi(Y^2)^{1/2} \geq \phi(i[X, Y]).$$

Thus the uncertainty principle for position and momenta follows immediately if we model these quantities by observables  $p$  and  $q$  with  $[p, q] = -i\hbar$ .

## Chapter 3

# Quantum Information Theory

The simplest unit of information in classical physics is a *bit*, represented by an element of  $\{0, 1\}$ . We can generalize

and the state of a collection of  $n$  bits are represented by an element of  $\{0, 1\}^n$ . From the quantum perspective, a *quantum bit*, or *qubit*, is represented by an element  $\psi = \psi_0|0\rangle + \psi_1|1\rangle$  of a two dimension Hermitian product space with orthonormal basis  $\{|0\rangle, |1\rangle\}$ .

(Gleason)

### 3.1 How can the Hamiltonian operator be Formed From the Classical Hamiltonian?

Thanks to Viktor T. Toth for this explanation on Quora. The Hamiltonian of a system is given by

$$H = p^2/2m + V(q),$$

where  $p$  is the momentum, and  $q$  the position. If we set

$$\psi = e^{(i/\hbar)(p \cdot q - Ht)},$$

then we notice that we can rewrite the Hamiltonian system above as

$$0 = (H - p^2/2m - V(q))\psi = H\psi - \frac{p^2}{2m}\psi - V(q)\psi.$$

Note that

$$H\psi = i\hbar \frac{\partial \psi}{\partial t} \quad \text{and} \quad p\psi = -i\hbar \nabla_q \psi,$$

and so

$$i\hbar \frac{\partial \psi}{\partial t} + \frac{\hbar^2}{2m} \Delta_q \psi - V(q)\psi = 0,$$

or equivalently,

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta_q \psi + V(q)\psi,$$

which begins to look like the Schrödinger equation. In *classical physics*, we restrict ourselves to solutions of this equation of the form  $e^{(i/\hbar)(p \cdot q - Ht)}$ , from which we can extract out the particular positions and frequencies of objects. *Quantum mechanics* begins when we now allow *linear combinations* of solutions of this form, i.e. 'mixed states'.