The Nature of the Electron

Don Lincoln, Fermilab, Batavia, IL

Winston Churchill once said of Russia that it was a riddle wrapped in mystery inside an enigma. Were the British Bulldog a physicist, he might have been talking of something other than our Slavic comrades. He might have been talking about an electron.

It might seem weird to think of an electron as mysterious. After all, introductory physics instructors everywhere teach students about charge and current and have them do labs to calculate the particle’s charge-to-mass ratio. If the particle is easy enough to understand that you can present it in technical detail to first-year physics students, how mysterious can it be?

Of course, you have probably now turned your thoughts to the quantum electron and Schrödinger’s equation and whether cats are alive or dead, and perhaps you are beginning to remember that the electron hasn’t told us all of its tales. And the story of quantum mechanics is now nearly a century old. Science has moved on and researchers have abandoned the simpler quantum representations of the electron for the more advanced relativistic quantum mechanics and now quantum field theory. Like Salome in Richard Strauss’s opera, the electron seems to be doing a dance of seven veils and the dance is not yet over. The electron still has a few secrets to reveal.

This article will mostly cover the more modern representations of the electron. A brief reminder of the electron of classical physics and early quantum theories will begin the story, but these will be used simply as a diving board as we plunge into the story of the modern electron. And, as we will see, that story is very deep.

The early electron

Before we knew of electrons, we knew of cathode rays (Fig. 1). Cathode rays were emitted when two metal plates were enclosed in a vacuum-filled tube in which two metal plates were placed. When a high electrical potential was applied between the two metal plates, it appeared that some sort of rays emanated from the cathode (negatively charged plate) and headed towards the positive plate. (While most readers will know that a filament is often used to create cathode rays, thermionic emission is simply more efficient and not essential.) Some of these cathode rays would hit the glass envelope, causing it to glow and revealing their presence. Later experiments added phosphorescent material to the glass to enhance the glow.

From their discovery in 1869 through much of the following three decades, scientists studied the nature of cathode rays and established that they had negative electrical charge. It was in 1897 that British physicist J.J. Thomson (Fig. 2) determined that cathode rays actually were composed of charged particles that he called corpuscles. He further postulated that not only were corpuscles constituents of atoms, they were the only massive constituents; the remainder was imagined to be a massless, positively charged, field.

Thomson’s experiments showed that they had a very large charge-to-mass ratio, suggesting that corpuscles either had a mass similar to that of a hydrogen atom and very large charge, or they had a charge comparable to an ionized hydrogen atom and a very small mass. While Thomson’s measurements were ambiguous on this conundrum, earlier measurements by German physicist Philipp Lenard on how cathode rays penetrated gas suggested that the low mass solution was more likely. Of course, we now know that the electron and proton have equal electrical charge and very different masses. The mass of the proton is about 1836 times more than an electron.

While Thomson discovered the electron, the name was coined by G. Johnstone Stoney in 1891 to describe the unit of charge observed in experiments that generated electrical currents in chemicals. Thomson’s model of the atom, consisting of negatively charged and massive electrons embedded inside a massless, positively charged field, was called the “plum pudding” model of an atom. But it didn’t last long.

It was Thomson’s student, New Zealand born physicist Ernest Rutherford, who in 1911 disproved Thomson’s model of an atom and replaced it with his idea of a nuclear atom, consisting of a concentrated positive charge in the center of the atom, with electrons orbiting at relatively large distances. Conceptually one can envision this model as analogous to a solar system, with a positively charged nuclear sun and negatively charged electron planets.

However, electrons orbiting a nucleus was clearly an incorrect model. We know this because conventional elec-

Fig. 1. Apparatus like the Crookes’ Tube shown here allowed 19th-century physicists to study cathode rays as a first attempt to understand what we now call an electron. (Wikimedia Commons, https://upload.wikimedia.org/wikipedia/commons/b/bf/Crookes_tube_two_views.jpg)

Fig. 2. Physicist J.J. Thomson discovered the electron in 1897, opening the way to our modern technological society. (Burndy Library, courtesy AIP Emilio Segré Visual Archives)
tromagnetic theory indisputably predicted that the constant radial acceleration of the charged electron would result in the emission of electromagnetic radiation, causing the electron to lose energy and spiral into the nuclear center in about 16 picoseconds. This is obviously a fatal flaw, as it means that atoms are inherently unstable. The first step in overcoming this difficulty occurred in 1913, when Danish physicist Niels Bohr postulated that electrons could only exist in atoms in fixed orbits of specific radii. This heuristic model was improved upon when French physicist Louis de Broglie postulated in his 1924 PhD thesis that electrons were actually waves. The fixed orbits postulated by Bohr could be better understood if the various atomic orbitals were actually standing waves of electrons.

When hearing about de Broglie’s insights, Dutch physicist Peter Debye made an offhand comment that if electrons were waves, then they had to be governed by a wave equation. This comment prompted Austrian physicist Erwin Schrödinger to create what is now called the Schrödinger equation, which describes the behavior of matter under the influence of a potential. Physical situations are solved by applying this equation and both boundary and continuity conditions. Perhaps the solution most familiar to physicists is the application of the Schrödinger equation to the hydrogen atom. This results in the wave functions corresponding to the s, p, d, and f orbitals that are commonly taught in introductory chemistry classes.

The physical meaning of the wave function has long been debated. The most common interpretation is called the Copenhagen interpretation, which was devised in the time period 1925–1927 by Niels Bohr and German physicist Werner Heisenberg. The name comes from the location of Bohr’s institute, which Heisenberg was visiting. In essence, the wave function, when multiplied by its complex conjugate, gives the probability for an electron to have a particular energy or location. Prior to detection, the electron is thought to have a myriad of possible energies and locations, and it is only through detection that the wave function “collapses,” and a specific value for the variable under investigation is determined for the electron.

**Developing a modern theory of electrons**

Schrödinger’s equation was a tremendous step forward in our understanding of the nature of an electron, but it had its limitations. The first issue was that it neglected the way in which the intrinsic magnetism of the electron interacted with a classical electromagnetic field. This limitation was first addressed in 1927 by Austrian physicist Wolfgang Pauli when he expanded Schrödinger’s formulation to include this consideration. However, even with this important extension, both Schrödinger’s and Pauli’s equations were explicitly non-relativistic. Given that the electron is one of the least massive subatomic particles and certainly the least massive one in the atom, it seemed clear that a relativistic treatment was crucial.

So these were problems that needed to be solved, and it is via this effort that our modern understanding of the electron began to develop. In 1928, British physicist Paul Dirac wrote down what we now call the Dirac equation. It is a fully relativistic wave equation. It was a breathtaking achievement. In addition to the fairly mundane marriage of quantum mechanics and relativity, perhaps the most famous (and unanticipated) consequence of the equation was that it predicted the existence of antimatter, which was experimentally verified a few years later. However the physical consequences of the equation were even deeper and more fundamental than that. For instance, the wave functions of Dirac theory are actually vectors consisting of four complex numbers. These four components are now viewed as a spin +1/2 electron, a spin –1/2 electron, a spin +1/2 positron (the antimatter version of the electron), and a spin –1/2 positron. Essentially the existence of spin is now understood to be a necessary consequence of the marriage of quantum mechanics and special relativity.

However, Dirac’s theory is not without its difficulties. Because the theory is mathematically complex, it is solvable only using perturbation theory. While theory gives physically sensible results when the leading order in the perturbative expansion was computed, it gives completely nonsensical results when the calculation is done in the next order of the expansion. In fact, predictions for some physical quantities become infinite. Obviously this indicates a fundamental sickness in the theory, but since the calculations make sensible and reasonably accurate predictions for many measurements, such as the Lamb shift described below, it is also clear that there is value in Dirac’s equation.

Shortly after World War II, American physicists Willis Lamb and his doctoral student Robert Retherford used microwaves to determine that there was a small energy difference between the 2P 1/2 and 2S 1/2 electron energy levels in hydrogen (i.e., the n = 2, l = 1 vs. n = 1, l = 0 levels). These two atomic orbital levels are predicted to have the same energy in Dirac theory, because, while the Dirac equation handles the coupling between the orbital angular momentum of the electron and its intrinsic spin, it does so in a way that is sensitive only to the total angular momentum, J, which is equal to ½ in both these states, but not to the individual components of angular momentum, L, or spin, σ, separately (J = L + σ). But Lamb and Retherford showed experimentally that there was an energy difference between these two levels corresponding to a frequency separation of about 1000 MHz, leading to a crisis in quantized electromagnetic theory, which predicted no such frequency separation.
Lamb presented their analysis at a conference in June of 1947 and the result was general confusion by the attendees. It was German physicist Hans Bethe (Fig. 3) who successfully explained the Lamb shift and set in motion the train of thought that culminated in our modern theory of quantum electrodynamics or QED. Bethe’s explanation can be understood in nonmathematical terms in the following way. Essentially what he did was to show that the electric and magnetic fields associated with the vacuum perturb the dominant electromagnetic field set up by the nucleus. Because the dumbbell-shaped P orbitals vanish at the center of the nucleus, while the spherical S orbitals do not, the P orbitals are less affected by the fluctuations of the quantum vacuum arising from the electric field of the nucleus. Because of the difference between the shapes of the two orbitals, the result is a shift in the energy level of the two.

It is perhaps humbling to know that Bethe worked out his solution while on a train ride from New York City to Schenectady (about 3.5 hours according to current-day Amtrak schedules), but such is the nature of genius. However, Bethe’s real genius is shown in how he solved the issue of the infinities that arose in the second order perturbation theory formulation of the Dirac equation.

For years, physicists had been trying to calculate the mass of the electron in part by calculating its interaction with the surrounding electric field. Using conventional relativistic quantum mechanics, the result is infinite. What Bethe did was to argue that there really were two cases to be considered, the electron bound in an atom and a free electron. In both cases, the self-energy of the electron existed, so Bethe argued that the measured mass of the free electron already contained the self-energy of a free electron. So rather than using the calculated value (which was infinite) for the bound electron, he replaced it with the measured value. This technique is now used widely and is called renormalization. Another way to think of this integration is to restrict the maximum possible photon energy to be the rest mass of the electron.

The technique of renormalization is now better understood than it was when Bethe was scribbling in his bouncy train car. Essentially what is done is to assume that the physics at very high energy cannot affect the lower energy phenomena that we can measure now. Mathematically, this is handled by putting in an upper-level cutoff in the integration, but conceptually it involves replacing the infinities with the physically measured parameters. While this seems a bit sketchy (even called hocus pocus by Richard Feynman [Fig. 4], one of the early proponents), it isn’t as silly as it seems. For one thing, it is almost certainly true that there exists undiscovered phenomena that occur at higher energies than we can probe now. This could be supersymmetry, quantum gravity, or something not yet imagined. Whatever it is, it tames the infinities that arise in our current incomplete theory. Thus one way to make progress is to accept the limitations, impose an integration limit, and absorb our ignorance by replacing the unphysical infinities with measured values. Presumably an improved theory will resolve the current mysteries. Because of this mathematically dodgy methodology, the early investigators of what is now called quantum electrodynamics (or QED) viewed the technique as utterly disreputable, but more recent investigation into what is now called renormalization group theory has shown it to be a cornerstone of modern theories. Indeed for a modern theory to be considered successful, it must be considered renormalizable (i.e., to be written so that the infinities can be hidden and, equally important, for the calculation to be independent of the choice of any cutoffs in the integrals).

**Quantum electrodynamics**

Bethe’s approach was explicitly nonrelativistic, but it was generalized in a Lorentz invariant way in a series of papers by American physicists Richard Feynman and Julian Schwinger and Japanese physicist Sin-Itiro Tomonaga, who shared the 1965 Nobel Prize in physics for their work. Feynman’s technique is based on a series of diagrams and looked very different than Schwinger and Tomonaga’s equation-heavy formulation, but they were shown to be equivalent by British physicist Freeman Dyson. Because of the more intuitive character of the Feynman approach, we restrict the discussion here to this methodology.

All three approaches used perturbation theory to perform QED calculations. Feynman’s brilliance was to be able to construct the quantum electrodynamic calculation as a series of pictures. The pictures represent the individual terms in the perturbative expansion. Each picture can be converted to an equation.

The perturbative expansion can be understood as a series of ever more complex interactions. When an electron scatters from another electron as the two electrons pass by one another, many different interactions can occur. However, the interactions can be ranked by the numerical size of their contributions to the final answer. For instance, the largest term in the expansion is realized when one electron emits a single photon that is absorbed by the other electron, and both electrons can recoil. A less likely scenario is that one of the incoming or outgoing electrons might emit a photon of its own. Even less likely is that two photons are emitted. There are infinite levels of increasing numbers of emitted photons in the scattering.

![Fig. 4. Richard Feynman was one of the architects of modern quantum electrodynamics. By casting the complicated mathematics of QED as a series of intuitive pictures, he allowed generations of particle physicists to understand the essence of a perturbative expansion of the theory. (W. F. Meggers Gallery of Nobel Laureates, Weber Collection, courtesy AIP Emilio Segré Visual Archives)](Image 39x360 to 145x485)
where $S$ is the electron spin, $\hbar$ is the Planck constant divided by $2\pi$, $\mu_B$ is the Bohr magneton, and $g$ is the electron $g$ factor, which is predicted to be identically 2 by Dirac theory. Kusch and his intellectual descendants originally found that the measured magnetic moment differs from the Dirac prediction by about 0.1%. Julian Schwinger was able to first calculate this difference in 1948 using QED. It turns out that QED effects can subtly change the magnetic moment of the naïve Dirac electron because the electron is constantly emitting and reabsorbing photons. When Schwinger took into account the lowest order approximation of electron self-interaction, called the lowest order loop diagram and shown in Fig. 6, he could account for the difference between prediction and measurement. The state of the art has improved over the decades and, as of this writing, the experimental measurement is

$$g_{\text{measured}} = 2.00231930436146 \pm 0.00000000000048,$$

with a predicted value of (using 10th order perturbation theory)

$$g_{\text{predicted}} = 2.00231930436356 \pm 0.0000000000011,$$

with $g_{\text{measured}} - g_{\text{predicted}} = (-1.05 \pm 0.82) \times 10^{-12}$. This agreement between theory and experiment is among the very best ever achieved and is considered a ringing endorsement for the theory of QED.

**Conceptual quantum field theory**

While QED theory has been a spectacular success, it is quite easy to leave a discussion of Feynman diagrams and precision calculations without a fundamental intuition of exactly what an electron is. Thus we now turn to a more holistic discussion. QED is a specific example of what is called a quantum field theory or QFT.

In essence, in a QFT, matter as we ordinarily think of it doesn’t exist. There is no chunk of matter that is an electron, or for that matter, a photon, quark, or any other fundamental particle. Instead, the entire universe is filled with fields—

![Image](image_url)

**Fig. 5. The scattering of two real electrons can be approximated as a series of ever more complex emission patterns. By factorizing the calculation in this way, one can employ perturbation theory to approximate the calculation, and the calculation can be performed to an arbitrary level of precision by including the required number of terms. The relevant expansion variable is determined by the number of vertices with the absorption or emission of a photon, e.g., two in leading order, three in second leading order, four in third leading order, and so on.**

![Image](image_url)

**Fig. 6. The small disagreement between measurements and the Dirac prediction of the magnetic moment of the electron was explained using QED corrections. When an electron is being scattered by a photon (left), one must consider the effect of the electron first emitting a photon, which is then later reabsorbed (right). This “loop” diagram is the dominant contribution to the 0.1% discrepancy.**

This article is copyrighted as indicated in the article. Reuse of AAPT content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP:
128.112.86.85 On: Mon, 21 Mar 2016 23:10:49
electrons, quark fields, photon fields—a field for every type of fundamental (i.e., point-like) particle that exists.

When a field is quiescent, essentially nothing is happening. However, these fields can support quantized vibrations. A quantized vibration is a localized vibration—what has been called a wave packet or a wave in descriptions of 1920s quantum mechanics. And each quantized vibration of the electron field is the physical manifestation of an electron.

The different fields can sometimes interact with one another. For example, when an electron emits a photon, a classical description might be a little ball-like electron moving along and shooting out a photon, which, depending on your mental image, can be envisioned as a ball (e.g., the photoelectric effect) or a little wave packet (e.g., wave theory). However, this picture has an analogue in the QFT paradigm.

In the QFT paradigm, an electron is a localized vibration of the electron field. At the point of emission, the electron field and photon field interact and a localized vibration is set up in the photon field. The photon vibration speeds off at the speed of light, while the electron continues on in the electron field, with its direction, wavelength, and velocity changed. The interaction between the two fields is similar to the resonant behavior of two adjacent tuning forks. If you set one tuning fork vibrating, the other fork will also vibrate.

And this simple picture of vibrating fields applies to all fundamental particles: the quarks and the leptons and also the force-carrying bosons. It does not apply to composite particles like protons and neutrons and all the myriad of mesons and baryons we’ve discovered. Or, more accurately, to describe a composite particle, you need to consider the (often mutually interacting) vibrations of all the fields of the particles’ constituents.

Near the electron

Near an electron, things are complicated. It is often said that an electron is a point-like particle, meaning that it has zero size. This is true, at least to within modern experimental resolution, which sets an upper limit on the size of the electron at $2 \times 10^{-20}$ meters. Given that both (a) QED postulates that the size of the electron is zero and (b) that the data are consistent with that postulate, one can describe what an electron “looks like” under the assumption of zero size. We shall see that the exact meaning of the term “zero size” is subtle.

In the simplest classical picture, the meaning is unambiguous. The electron truly has zero size and it is surrounded with an electrostatic potential that goes as $V = -ke/r$. A more sophisticated classical analysis of the electron yields a size called the classical electron radius, on the order of $10^{-15}$ m, which is simply the size of a sphere with the charge of the electron and with the energy required to assemble it that is equal to the mass energy of the electron.

However, in the QFT paradigm, while a “bare” electron with zero size exists at the center of the electron, it is surrounded by a cloud of both photons and electron/positron pairs. These particles are typically virtual, which means that they don’t have the “right” mass, which specifically means the mass associated with the free particles; more precisely this means that these photons are not massless and the electrons and positrons don’t have a mass of 0.511 MeV/c^2. They can exist because the Heisenberg uncertainty principle allows the existence of particles with the “wrong” mass, if they don’t exist for very long.

The maelstrom of virtual particles grows more frenzied as one gets closer to the core of the electron. In QFT terms, this simply means that the “bare” electric charge agitates the photon field, which in turn disturbs the electron field. The net outcome is a complicated mess and with the “bare” charge shielded by the surrounding cloud of virtual particles. It also means that the effective charge of the electron changes with the energy with which you probe it; as when you shoot a probe particle toward the electron, higher energy probes can penetrate more deeply into the cloud, and both see more of “bare” charge and experience more of the effects of the virtual particles. Thus the electron is now no longer simply a bare charge; it has a complicated radial charge distribution. If one wishes to find its energy in an atom in a specific energy level, one must integrate over its position distribution and the electric field it encounters from the nucleus. The combination of the different shapes of the S and P orbitals, in conjunction with this small charge distribution of the electron, is the basis of the Lamb effect. A complementary effect arises from a similar treatment of the electric field in the neighborhood of the nucleus.

This modification of the charge of the electron with increasing probe energy has been experimentally verified to very good precision. There is no indication that the theoretical electron at the core of the whirlwind of photons and virtual particles is composed of smaller objects. (Note I am distinguishing the theoretical particle from the physical electron.) Ongoing measurements continue to look for hints that the core electron might itself be composed of even smaller components. Thus far, all attempts to identify electron constituents have failed.

Remaining mysteries

While it is true that the theory of QED is the most accurate ever devised, even today some mysteries remain. For instance, two parameters of QED theory that cannot be calculated from first principles are the mass and the electrical charge of the electron. Those parameters must be determined empirically and put back into the theory by hand. This is a clear signal that QED is not the final word, but rather is an effective model that arises from a deeper and more fundamental theory.

But if you start thinking more freely, the questions don’t stop there. Here are but a few questions for which we don’t know answers. Why do electrons carry electric charge, but not the strong nuclear charge? What distinguishes the elec-
tron from its cousins the muon and tau lepton (setting aside the obvious answer of mass)? Why is it that the electron and proton have exactly the same amount of electrical charge, up to a sign? What goes on when the electron wave function collapses and localizes the electron? Do electrons ultimately have constituents? Questions without answers are common for researchers and all it means is that our journey to understand the electron is still not quite complete.

Summary

As has been known for nearly a century, the subatomic world is a counterintuitive one, with particle and wave behavior both playing a role. However, we know that the electron is not a classical particle and not a classical wave. It is somewhat described by the wave packet idea of 1920s quantum mechanics, but the reality is more complex and it should really be thought of as a constantly changing nexus of interacting fields. Quantum electrodynamics is a theory that describes the behavior and fundamental properties of the electron with breathtaking precision. The modern picture of the electron is complex, consisting of vibrations of an underlying electron field, surrounded by an equally agitated photon field and even with the vibrations of the fields of other fundamental subatomic particles. The beautiful precision of QED predictions ensures that this image will persist. However no theory is forever. What will our vision of an electron be a century from now?

References


Suggested reading


Don Lincoln is a senior scientist at Fermilab, with 30 years of research experience under his belt and over 1000 refereed publications. Among his more noteworthy scientific achievements are contributing to the discovery of the top quark and the Higgs boson. In addition to his research, he is a tireless popularizer of science. He has written four books, including The Large Hadron Collider: The Extraordinary Story of the Higgs Boson and Other Stuff that Will Blow Your Mind (September 2014). He has written for Scientific American, appeared on the television show “NOVA,” and published many magazine articles, including in The Physics Teacher, for which this is his seventh cover story. You can follow him on Facebook at http://www.facebook.com/dr.don.lincoln. Website: http://drdonlincoln.fnal.gov. lincoln@fnal.gov

Look What’s in The Physics Store!

Preconceptions in Mechanics

This second edition of Charles Camp and John J. Clement’s book contains a set of 24 innovative lessons and laboratories in mechanics for high school physics classrooms that was developed by a team of teachers and science education researchers. Research has shown that certain student preconceptions conflict with current physical theories and seem to resist change when using traditional instructional techniques. This book provides a set of lessons that are aimed specifically at these particularly troublesome areas: Normal Forces, Friction, Newton’s Third Law, Relative Motion, Gravity, Inertia, and Tension. The lessons can be used to supplement any course that includes mechanics. Each unit contains detailed step-by-step lesson plans, homework and test problems, as well as background information on common student misconceptions, an overall integrated teaching strategy, and key aspects of the targeted core concepts. A CD of all duplication materials is included.

Members: $28
Non-Members: $35

Order yours now at www.aapt.org/store

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics

Preconceptions in Mechanics