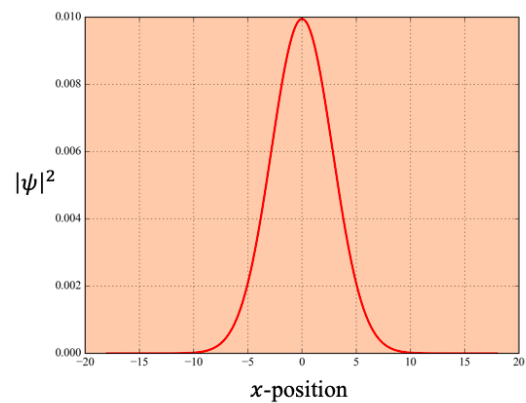
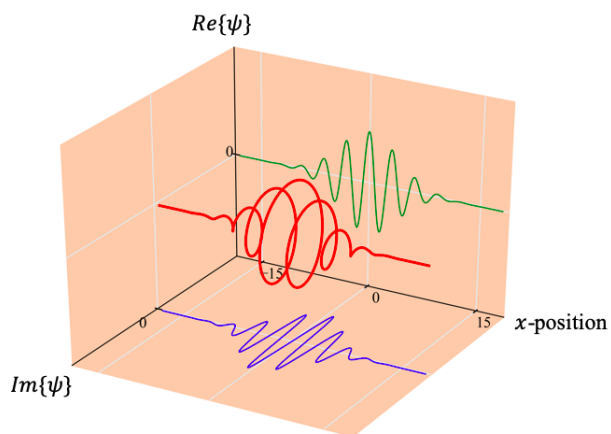


# The Aharonov-Bohm Effect: Are Potentials Real?

The effect of electromagnetic potentials on charged quantum particles

Word Count: 2977



Blackett Laboratory  
Imperial College London  
UK  
January 2021

---

# Introduction

The modern notion of "field" first came to the mind of physicists like Maxwell and Faraday in the 1800s when electromagnetism was studied (1). In contrast to the way Newton dealt with gravitation, electromagnetism was considered as originating from these fluid-like entities that permeate the entire universe called "fields". Fields then became the centre of classical electromagnetism. Nowadays, fields are considered the "fundamental building blocks" of our universe in quantum field theory (2); elementary particles, fundamental forces, and even mass are described by ripples of the corresponding fields (2).

However, the forms of fields in electromagnetism are usually difficult to obtain. The concept of potential was then introduced as a tool of representing the fields. We would focus mainly on the magnetic field and its corresponding vector potential. The magnetic vector potential,  $\vec{A}$ , was defined so that its curl gives the magnetic field. However, in classical electrodynamics, only fields exert forces on particles; potentials are but just mathematical constructs to represent the field (3, 4). No matter how the potentials were configured, as long as the fields were kept at zero, a classical particle would not be affected at all<sup>1</sup>. The question then arose: Are potentials real entities or merely mathematical constructs?

The story started in 1900 when Planck hypothesized that the radiation energy from a black body came in discrete packets (5). The theory of quantum mechanics was brought into place soon after. Quantum mechanics abandoned the concept of force, but made energy and momentum the defining properties. However, the energy and momentum of a particle were closely related to the potentials! Although we have seen that charged particles could not be affected by Newtonian forces without any electromagnetic field, their quantum properties could vary if potentials were defined. It has been well

established that charged particles could have their quantum behaviours altered by the magnetic potential  $\vec{A}$  without any local magnetic field (4, 6, 7). This idea revealed the physical reality of the vector potential and was named the Aharonov-Bohm effect.

The research around this effect never stopped since its first discovery. However insignificant this effect might seem, it originated from the most fundamental concepts in physics. In the past few decades, people relentlessly explored many ideas in the quantum theory to give various interpretations of this effect (8–11). More practically, the experiments aimed to demonstrate this effect promoted countless technological advances in electron wave optics and electron microscopy (12). Direct observations of extremely weak magnetic fields generated by magnetised materials were enabled as one of the consequences. The Aharonov-Bohm effect could hence be considered as the ancestry of many theoretical and experimental research areas.

## Particles? Maybe Waves...

To understand the Aharonov-Bohm effect, we first need a bit of some quantum mechanics. The wave description of particles and the probabilistic nature of their states are probably the most counter-intuitive concepts in quantum mechanics. In 1905, Einstein generalised Planck's idea to explain the photoelectric effect. He interpreted light waves as little packets of energy just like particles, photons (13). But what is so special about light that it can be both waves and particles? This idea then inspired de Broglie in 1924 to propose that if we could interpret light waves as particles, then we could surely describe particles as waves (14)! Unifying their ideas, the wavevector,  $\vec{k}$ , and frequency,  $\omega$ , of the "matter wave" were expressed as

$$\vec{k} = \frac{\vec{p}}{\hbar}, \quad (1)$$

---

<sup>1</sup>A familiar example would be a constant electric potential; the electric field vanishes while the potential doesn't. We could also demand the magnetic potential to be the gradient of some scalar field. Since the curl of a gradient is always zero, we then have the magnetic field to vanish even though the vector potential doesn't

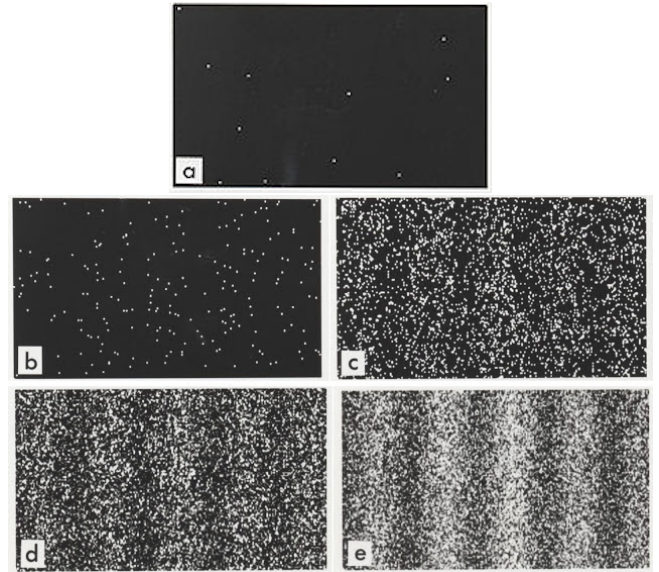
$$\omega = \frac{E}{\hbar}, \quad (2)$$

where  $p$  and  $E$  are the momentum and energy of the particle, and  $\hbar$  is reduced Planck's constant which equals to Planck's constant divided by  $2\pi$  ( $\hbar = h/2\pi$ ). The two equations above will be referred as the de Broglie relations below.

But what exactly does it mean to describe particles as waves? If we had something in our hand, it surely wouldn't feel wavy. Also, waves are going to span a large region of space, but everything we see has definite positions; a cup of coffee for breakfast can't have part of it on the kitchen table and part of it on the moon simultaneously. Actually, we treat these "matter waves" as a representation of probabilities.

The "matter waves" are formally called wave functions with the symbol  $\psi(x, t)$  to represent their amplitudes. The wave function is in general complex-valued, just like functions that describe light waves. Peculiar to wave functions, their norm-squares,  $|\psi(x, t)|^2$ , give the probability density of the particle being at position  $x$  and at time  $t$  (15). In analogy to light waves, the intensity, which is proportional to the norm-square of the amplitude, tells us how likely it is to find a photon there. Hence, the wave functions are just probability amplitudes and their norm-squares give probability densities.

Now we have a basic probabilistic interpretation of wave functions. In optics, amplitudes of light add but not the intensities. Similarly, if we had two wave functions, their amplitudes add directly but not the probability densities; this leads to interference. In Young's double slit experiment, two sources of light waves interfere to form bright and dark patterns on a screen. The same thing could happen if we shoot electrons through the double slit. The electron wave functions would interfere resulting in areas with oscillating probability densities. If we had a screen that produces a bright spot every time an electron hits, we would be able to see the interference from the pattern of spot accumulations (7, 16). Figure 2 shows the result from an experiment of electron interference done by Akira Tonomura *et. al.*



**Figure 2:** The results of electrons building up interference pattern by A. Tonomura *et. al.* in (16). Each bright spot corresponds to one electron hitting the screen. From picture (a) to (e), each included 10, 100, 3000, 20000, and 70000 electrons, respectively.

What would happen in optics if one of the slits was blocked by a piece of glass? Glass's high refractive index would delay the light wave from that slit and hence produce a phase lag. This additional relative phase between light waves through the two slits would hence shift the entire interference pattern (17). The same thing could happen if electrons instead of photons were involved in such an interference experiment. The Aharonov-Bohm effect is just a result of the magnetic vector potential acting like the glass in optics. The vector potential would change the "refractive index" of the space which the electron wave functions travelled through and thus shift the electron interference pattern.

## Who Touched My Electrons?

Suppose we had the setup for Young's double slit experiment with electrons. In addition, a very long solenoid was placed between the two slits which brought in a magnetic field as shown in figure 3. The magnetic field was strictly confined within the solenoid so that electron waves experienced no field at all in their way. One would expect nothing to change as if the magnetic field it wasn't there since it was out of the electron waves' reach. In reality, although the

magnetic field was confined, the magnetic vector potential extended through all space. Since  $\vec{B} = \nabla \times \vec{A}$ ,  $\vec{A}$  could look like concentric circles as in figure 3(a) (3, 4, 18). It turned out that the vector potential sneakily altered the momenta of electron waves and hence changed their phases.

## History of development

This effect of electromagnetic potentials on electrons was named after Yakir Aharonov and David Bohm following their paper published in 1959 (4). Astonishingly, they were not the first people who discovered it. Raymond E. Siday and Werner Erenberg first predicted this effect in 1949 (6) but it did not attract much attention. Aharonov and Bohm did not know about their discovery until after the 1959 paper was published; however, Siday and Erenberg were given the credit in a following paper on this effect in 1961 by Aharonov and Bohm (19).

The prediction was initially very controversial within the community (20). Even after the first few experimental confirmations, some still argued that it was a consequence of the magnetic fields not being strictly confined. Some field leaked out to the paths of electrons and shifted the interference pattern (21). It was not until much later in 1986, this effect was eventually demonstrated to high precision with a strictly confined magnetic field (22).

## A little bit of maths...

We shall follow a similar mathematical route as Aharonov and Bohm did but with some semiclassical simplifications. Using the de Broglie relations mentioned before (Eq. 1 and Eq. 2), we could express electron wave functions as

$$\psi(\vec{x}, t) = \psi_0 e^{i(\vec{k} \cdot \vec{x} - \omega t)}, \quad (3)$$

where  $\psi_0$  is the amplitude,  $\vec{x}$  is the position vector, and  $t$  is time elapsed, just like a travelling plane wave in optics<sup>2</sup>. We see that the phase of this wave function depends on  $\vec{k}$  and  $\omega$  which were related to the momentum and the energy of the electron.

We have seen before that the vector potential circles around the magnetic field in the right-handed direction. The picture might look very familiar because that is exactly the trajectory of an electron gyrating around a uniform magnetic field! Also if we look at the dimension of  $\vec{A}$ , it looks like the dimension of "momentum per charge". These could give us some hints: just like how we regarded the electric potential as electric potential energy per charge, the vector potential could be also treated as something like "potential momentum per charge". Adding it up to the kinetic momentum, we get the "total momentum"<sup>3</sup> of the electron as

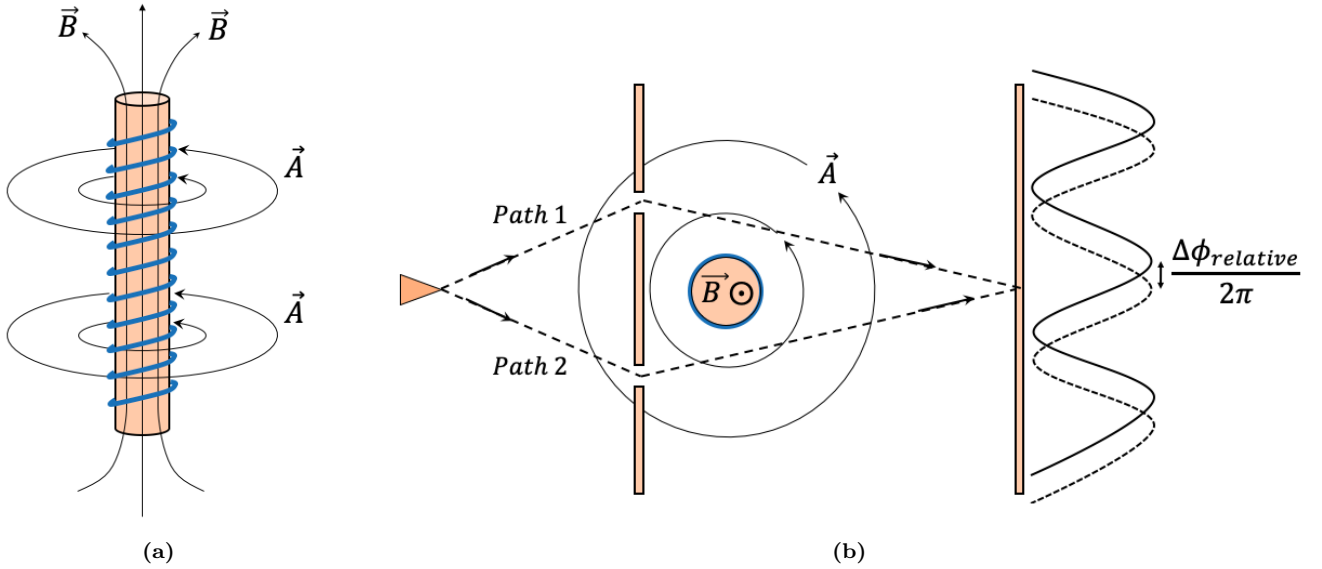
$$\vec{p} = m\vec{v} + q\vec{A}, \quad (4)$$

where  $m$  is the mass of electron,  $\vec{v}$  is the velocity of electron, and  $q$  is the electron charge (24). The vector potential provides an extra term in the momentum of electrons; if the vector potential vanished, we would be only left with the kinetic momentum as usual.

Now we re-examine the expression of electron wave functions in Eq. 3. Since the momentum now depends on  $\vec{A}$ , the wavevector  $\vec{k}$  would also depend on  $\vec{A}$ . To quantify this result, let's first consider the phase change of an electron wave after turning the vector potential on. For a wave where its wavevector was changed by  $\Delta\vec{k}$ , its phase would differ from the original unchanged wave by  $d\phi = \Delta\vec{k} \cdot \vec{dx}$  after it travelled a small distance  $\vec{dx}$ . The total change in phase was therefore just a line integral that sums up all the little contributions of phase difference from the starting point to the end point

<sup>2</sup>Usually the wave function of an electron is a wave packet so that its probability density is localised in some region. For simplicity, we will just use the plane wave solution here.

<sup>3</sup>The "total momentum" is technically called canonical momentum. The canonical momentum arises in Lagrangian and Hamiltonian mechanics as a generalised momentum derived from the Lagrangian (23). Also, the "potential momentum per charge" interpretation of the vector potential is not a formal definition but just used here to present the idea. The mathematics of quantum mechanics were mostly inherited from Lagrangian and Hamiltonian mechanics so we need the canonical momentum here for this quantum system.



**Figure 3:** (a) A solenoid generates magnetic field inside the windings. The simplest choice of the vector potential would circle around the solenoid as shown with its curl giving back the magnetic field. (b) After placing the solenoid into the electron double slit experiment, path 1 would have a phase lag while path 2 would have a phase lead. The interference pattern would then undergo a shift of  $\Delta\phi_{\text{relative}}/2\pi$  fringes.

as

$$\Delta\phi = \int_{\vec{x}_i}^{\vec{x}_f} d\phi = \int_{\vec{x}_i}^{\vec{x}_f} \Delta\vec{k} \cdot \vec{dx}, \quad (5)$$

where  $\Delta\phi$  is the phase difference of an electron wave after its wavevector was changed. Thereafter, we could invoke the first de Broglie relation (Eq. 1) to express that change in wavevector by the extra factor in momentum after we turned on a vector potential as

$$\Delta\phi = \int_{\vec{x}_i}^{\vec{x}_f} \frac{\Delta\vec{p}}{\hbar} \cdot \vec{dx} = \int_{\vec{x}_i}^{\vec{x}_f} \frac{q\vec{A}}{\hbar} \cdot \vec{dx}. \quad (6)$$

As this expression demonstrates, the phase change of an electron wave depends on the local value of the vector potential<sup>4</sup>. As in figure 3(b), path 1 and path 2, as an example of any arbitrary path pairs through the two slits, suffered from the additional phase shifts described by Eq. 6. There was consequently an additional relative phase. This extra relative phase could be given by the difference of two line integrals as in Eq. 6 along path 1 and 2, which then could be

expressed as a loop integral going forward along path 1 and backwards along path two (4, 24):

$$\Delta\phi_{\text{relative}} = \frac{q}{\hbar} \oint_{1-2} \vec{A} \cdot \vec{dx} = \frac{q}{\hbar} \iint_S \vec{B} \cdot \vec{ds}, \quad (7)$$

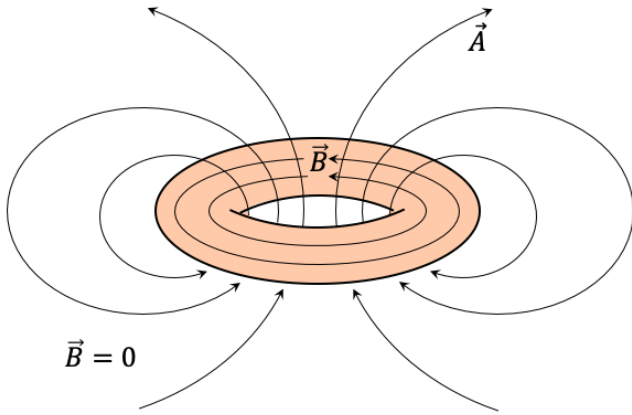
where we used Stokes' theorem to write the loop integral as an integral of the magnetic field over the surface enclosed by the two paths (the magnetic flux). Therefore, the entire interference pattern would be shifted with the number of fringe shift given by  $\Delta\phi_{\text{relative}}/2\pi$  as shown in figure 3(b). Since the relative phase here only depends on the loop integral the arbitrariness of choosing  $\vec{A}$  up to a gradient is preserved.

## Experimental confirmations

Just one year after the prediction by Aharonov and Bohm, Robert G. Chambers first experimentally confirmed the existence of this effect. He used a setup very similar to that we described where a tapered iron whisker was used to provide the magnetic field (26). The

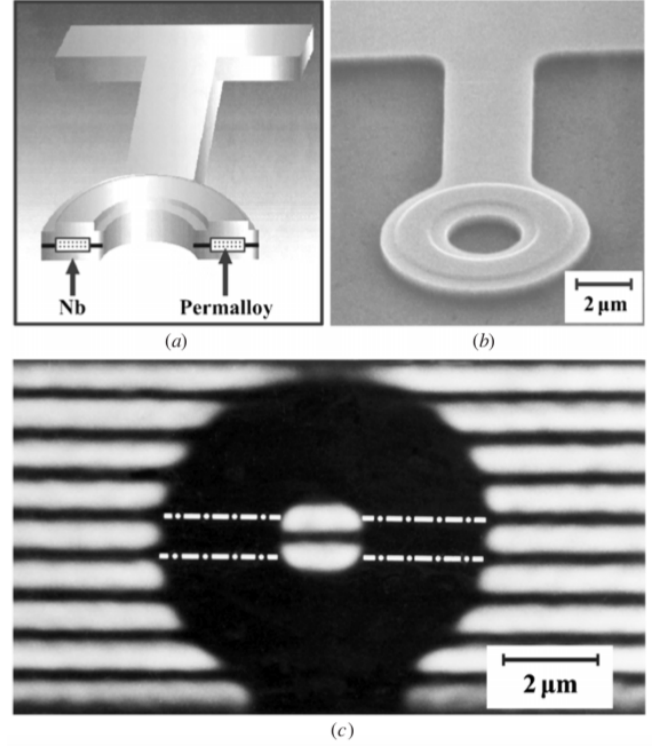
<sup>4</sup>The derivation here (from change in momentum to change in wavevector and eventually to change in phase) is intended to be intuitive but not very rigorous. A more accurate treatment would require the use of Lagrangian and Hamiltonian mechanics to write down the Schrödinger equation for the electron waves; the extra phase would then arise naturally in the solution. The complete derivation could be found in (4, 24, 25), but the physical idea is the same as described here.

experiment was repeated by many other people in the 1960s using different sources of magnetic field such as bar permalloys and solenoids (12, 27). Although the expected fringe shift was observed, some proposed that it might be a consequence of some magnetic field leaking to the path of the electrons since none of the magnetic field sources were infinitely long (21).



**Figure 4:** A toroidal magnet or solenoid would be able to produce magnetic fields that go around inside but vanish outside. The vector potentials, however, would still extend through all a space.

How could we make a perfectly confined magnetic field but with finite materials? The answer is to use a toroidal solenoid or magnet. The magnetic field lines created by such solenoids or magnets only revolve inside the torus without leaking out as shown in figure 4. In 1986, a group of physicists in Japan led again by Akira Tonomura fabricated permalloys into tiny toroidal magnets (12, 22). To further prevent field leakage, the toroidal magnets were covered with a layer of superconductor and a layer of copper so that the magnetic field was perfectly shielded (22). The electron paths through the hole of the torus and those outside had an additional relative phase given by Eq. 7. A shift of fringes inside the hole of the torus was then observed with the expected value as shown in figure 5(c).



**Figure 5:** Structure of the toroidal magnet and the result adopted from (22). (a) A schematic of the toroidal permalloy magnet. Nb is the symbol for a type of superconductor. (b) A scanning electron micrograph of the magnet. (c) The interference pattern was shifted inside the hole of the toroidal magnet where the original dark fringes outside was replaced by bright fringes inside.

### *What does this effect imply?*

The majority of physicists interpreted the Aharonov-Bohm effect as an evidence of the vector potential being a real field. The argument is that although the value of phase difference between the two paths could be expressed as proportional to the magnetic flux, its origin was the vector potential as there was no field along the electron paths (3, 4, 7, 12, 19, 24). Physicists strongly believed in the principle of locality which says that an object could only be affected by its surroundings but not some thing very far away (3, 11, 28). This idea also appeared in special relativity where events could only affect and be affected by events inside the light cone. In the Aharonov-Bohm effect, the principle of locality is saying that since the confined magnetic field was out of the reach of the electrons, the vector potential must be a physical reality which altered the phases.

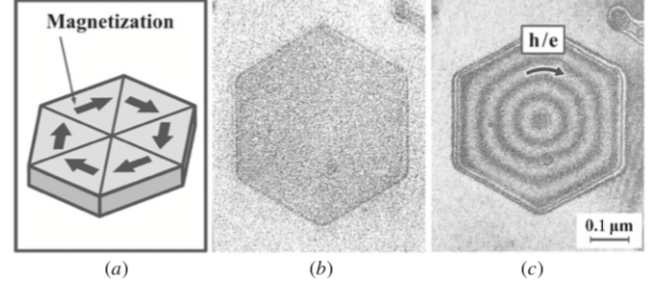


On the other hand, some other opinions came up in the past ten years which challenged the conventional interpretation. Lev Vaidman in 2012 proposed that if the source of the magnetic field, the solenoid, was also treated quantum mechanically, we could get the same phase difference without getting into the discussion of the vector potential (8). He argued that although the magnetic field generated by the solenoid was shielded from the electrons, the electromagnetic field created by the moving electrons were not shielded from the solenoid. Since the solenoid and the moving electrons could interact through this electromagnetic field, they were actually one system. Hence, the solenoid had to be included with the electrons in an overall wave function forming an entangled state (8). This effect is just another non-local effect in quantum mechanics similar to quantum entanglement<sup>5</sup>. Following Vaidman, Philippe Pearle and Antony Rizzi published two papers in 2017 to provide more detailed quantum models to the Aharonov-Bohm effect (9, 10). These three papers gave three different quantum mechanical approaches to the problem but all avoided exploiting the physical reality of the vector potential.

## Taking a Picture of the Magnetic Field

One of the most powerful applications of the Aharonov-Bohm effect would be the electron phase microscopy. The electrons were let to diffract through a magnetised object. Due to the Aharonov-Bohm effect, the interference pattern incorporated information about the magnetic field. Eventually, a micrograph illustrating the shape of the magnetic field was reconstructed from the electron interference pattern (30). Thus, we could actually take pictures of the magnetic field lines! This procedure was named electron phase microscopy as the inter-

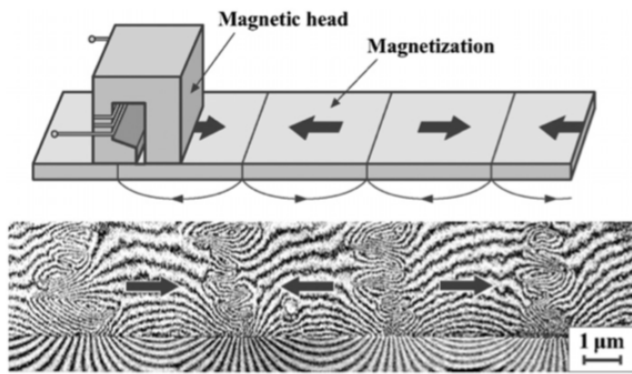
ference pattern, hence the phase information of electron waves, was exploited.



**Figure 6:** The schematic and electron micrographs of a hexagonal cobalt particle adopted from (12). (a) A schematic showing the direction of the magnetic field. (b) An electron micrograph from which we could only see the object's geometry. (c) An electron phase micrograph showing the magnetic field lines. It could be deduced that the fringes representing field lines were separated by steps of  $h/q$  magnetic flux (22).

In the 1980s, A. Tonomura and his team started to research in producing electron phase micrographs for tiny magnetised objects (22, 30). For instance, figure 6 shows the schematic and results for a hexagonal cobalt particle. The interference fringes were along the direction of the magnetic field lines. Micrographs of many differently magnetised objects were made and the most interesting one was probably for a recorded magnetic tape. While making a recording, a magnetic head would manipulate the magnetic field of little segments on the tape into opposite directions to store information as 0s and 1s. In 2008, Tonomura and his team took a micrograph for a recorded tape as shown in figure 7. Knowing the exact magnetic field configurations helped manufacturing magnetic tapes with higher efficiency in terms of storing the recordings.

<sup>5</sup>Quantum entanglement is one of the most well-known quantum "weirdness". Two particles in an entangled state are still correlated even if they were separated far from each other. If the state of one particle was measured, we would immediately know the state of the other. More details on this matter and an explanation of it not violating special relativity could be found in (29).



**Figure 7:** The electron phase micrograph adopted from (12). Segments on the tape were induced with magnetic fields with different orientations as indicated by the arrows. The shape of magnetic field lines were shown by the fringes.

## Conclusion

While the physical significance of the vector potential was clearly demonstrated in the Aharonov-Bohm effect, it also incorporated many other fundamental ideas in physics such

as particle-wave duality, the principle of locality, and even quantum entanglement. In the process of researching into this effect, many side results were also obtained. For instance, the quantisation of magnetic flux<sup>6</sup> was observed when Tonomura and his team was conducting the experiment we previously mentioned to verify the Aharonov-Bohm effect (22). More than half a century after its first discovery, it is still an active area of research today. Some physicists are working on "quantum sensing" which uses the quantum properties of matter to make measurements, just like electron phase microscopy we mentioned before(12, 31); advances in these technologies could appreciably benefit areas like nano-technology, medical imaging and even quantum computing. Some theorists continue to research in subjects related to this effect and its extensions<sup>7</sup> (32–34). More and more ideas would be brought into place and lead us to a better understanding of nature.

## References

- [1] McMullin E. The Origins of the Field Concept in Physics. *Phys Perspect.* 2002 Feb;4:13–39.
- [2] Tong D. Quantum fields. YouTube: The Royal Institute; 2017. [Video]. Available from: [https://www.youtube.com/watch?v=zNVQfWC\\_evq](https://www.youtube.com/watch?v=zNVQfWC_evq).
- [3] Feynman RP, Leighton RB, Sands M. Chapter 15: The Vector Potential. In: *The Feynman Lectures on physics. vol. 2. new millennium edition* ed. 250 West 57th Street, 15th Floor, New York, NY 10107: Basic Books; 2013. p. (15–1) – (15–27).
- [4] Aharonov Y, Bohm D. Significance of Electromagnetic Potentials in the Quantum Theory. *Phys Rev.* 1959 Aug;115:485–491. Available from: <https://link.aps.org/doi/10.1103/PhysRev.115.485>.
- [5] Planck M. The theory of heat radiation. 2nd ed. 1012 Walnut Street, Philadelphia: P. Blakiston's Son & Co.; 1914. Available from: <https://www.gutenberg.org/files/40030/40030-pdf.pdf>.
- [6] Ehrenberg W, Siday RE. The Refractive Index in Electron Optics and the Principles of Dynamics. *Proceedings of the Physical Society Section B.* 1949 Jan;62(1):8–21. Available from: <https://doi.org/10.1088/0370-1301/62/1/303>.

<sup>6</sup>The quantisation of magnetic flux is a property of superconducting loops. It appeared in this experiment because the toroidal magnet was covered with a layer of superconductor. More information would be found at (22, 25).

<sup>7</sup>These extensions include Shelankov's idea of quantum "force", the relationship between the Aharonov-Bohm effect and the Berry phase, the Aharonov-Casher effect, etc.



- [7] Imry Y, Webb RA. Quantum Interference and the Aharonov-Bohm Effect. *Scientific American*. 1989 Apr;260(4):56–62. Available from: <http://www.jstor.org/stable/24987212>.
- [8] Vaidman L. Role of potentials in the Aharonov-Bohm effect. *Phys Rev A*. 2012 Oct;86:040101. Available from: <https://link.aps.org/doi/10.1103/PhysRevA.86.040101>.
- [9] Pearle P, Rizzi A. Quantum-mechanical inclusion of the source in the Aharonov-Bohm effects. *Phys Rev A*. 2017 May;95:052123. Available from: <https://link.aps.org/doi/10.1103/PhysRevA.95.052123>.
- [10] Pearle P, Rizzi A. Quantized vector potential and alternative views of the magnetic Aharonov-Bohm phase shift. *Phys Rev A*. 2017 May;95:052124. Available from: <https://link.aps.org/doi/10.1103/PhysRevA.95.052124>.
- [11] Aharonov Y, Cohen E, Rohrlich D. Nonlocality of the Aharonov-Bohm effect. *Phys Rev A*. 2016 Apr;93:042110. Available from: <https://link.aps.org/doi/10.1103/PhysRevA.93.042110>.
- [12] Tonomura A. The AB effect and its expanding applications. *Journal of Physics A: Mathematical and Theoretical*. 2010 Aug;43(35):354021. Available from: <https://doi.org/10.1088/1751-8113/43/35/354021>.
- [13] Dirac PAM. Chapter 1: The Principle of Superposition. In: *The Principles of Quantum Mechanics*. 4th ed. Amen House, London, E.C.4: Oxford University Press; 1958. p. 7–14.
- [14] de Broglie L. The interpretation of wave mechanics. *Foundations of Physics*. 1970 Mar;1:5–15. Available from: <https://doi.org/10.1007/BF00708650>.
- [15] Griffiths DJ. Chapter 1: The Wave Function. In: *Introduction to quantum mechanics*. 2nd ed. Pearson Prentice Hall, Upper Saddle River, NJ 07458: Pearson Education, Inc.; 2005. p. 1–18.
- [16] Tonomura A, Endo J, Matsuda T, Kawasaki T, Ezawa H. Demonstration of single-electron buildup of an interference pattern. *American Journal of Physics*. 1989;57(2):117–120. Available from: <https://doi.org/10.1119/1.16104>.
- [17] Damzen M. Optics Lecture 9: Fourier Optics. Imperial College London; 2020. [Lecture].
- [18] Jackson JD. Chapter 5.4-5.5. In: *Classical electrodynamics*. 3rd ed. 111 River Street, Hoboken, NJ 07030: John Wiley & Sons, Inc.; 1999. p. 180–184.
- [19] Aharonov Y, Bohm D. Further Considerations on Electromagnetic Potentials in the Quantum Theory. *Phys Rev*. 1961 Aug;123:1511–1524. Available from: <https://link.aps.org/doi/10.1103/PhysRev.123.1511>.
- [20] Bocchieri P, Loinger A. Nonexistence of the Aharonov-Bohm effect. *Nuov Cim A*. 1978 Oct;47:475–482. Available from: <https://doi.org/10.1007/BF02896237>.
- [21] Roy SM. Condition for Nonexistence of Aharonov-Bohm Effect. *Phys Rev Lett*. 1980 Jan;44:111–114. Available from: <https://link.aps.org/doi/10.1103/PhysRevLett.44.111>.

