Report
Fundamentals of Semiconductor Devices

Nikolaus Hammler*
Hirose-Lab
The University of Tokyo

February 19, 2010

Contents

*niki@hammler.info

1 Task 1 - The Momentum Operator

In this task it should be shown that

$$\hat{\mathbf{v}} = \frac{\hat{\mathbf{p}}}{m} \tag{1}$$

i.e. the classical definition. $\mathbf{v}$ is defined as the derivation of location:

$$\mathbf{v} = \frac{d}{dt}\mathbf{r} \tag{2}$$

Proof. My proof is based on this definition:

$$\frac{d}{dt}\mathbf{r} = \frac{d}{d\,t}\int \psi^*(\mathbf{r}) \cdot \mathbf{r} \cdot \psi(\mathbf{r})\,d\mathbf{r} \tag{3}$$

$$= \int \frac{d}{d\,t}\left(\psi^*(\mathbf{r}) \cdot \mathbf{r} \cdot \psi(\mathbf{r})\right)\,d\mathbf{r} \tag{4}$$

$$= \int \mathbf{r}\frac{d}{d\,t}\left(\psi^*(\mathbf{r})\psi(\mathbf{r})\right)\,d\mathbf{r} \tag{5}$$

Using the product rule to solve the differential:

$$\frac{d}{dt}\mathbf{r} = \int \mathbf{r}\left(\psi^*(\mathbf{r})\frac{d}{d\,t}\psi(\mathbf{r}) + \psi(\mathbf{r})\frac{d}{d\,t}\psi^*(\mathbf{r})\right)\,d\mathbf{r} \tag{6}$$

$$= \int \psi^*(\mathbf{r})\mathbf{r} \cdot \underline{\frac{d}{d\,t}\psi(\mathbf{r})} + \psi(\mathbf{r})\mathbf{r} \cdot \underline{\frac{d}{d\,t}\psi^*(\mathbf{r})}\,d\mathbf{r} \tag{7}$$

These can now be replaced in the following way: Let us look at the definition of Schroedingers equation:

$$H\psi = j\hbar\frac{\partial}{\partial t}\psi \quad \text{with} \tag{8}$$

$$H = -\frac{\hbar^2}{2m}\Delta + V(r) \tag{9}$$

$$\Rightarrow \frac{\partial}{\partial t}\psi = \frac{H}{j\hbar}\psi \tag{10}$$

Now the same can be done for the conjugated version:

$$\frac{\partial}{\partial t}\psi^* = \left(\frac{H}{j\hbar}\psi\right)^* = -\frac{H^*}{j\hbar}\psi^* \tag{11}$$

and with the important observation that $H^* = H$ finally

$$\frac{\partial}{\partial t}\psi^* = -\frac{H}{j\hbar}\psi^* \tag{12}$$

Now equations 10 and 12 can be inserted into equation 7:

$$\frac{d}{dt}\mathbf{r} = \int \psi^*(\mathbf{r})\mathbf{r}\cdot\frac{H}{j\hbar}\psi(\mathbf{r}) + \psi(\mathbf{r})\mathbf{r}\cdot\frac{H}{-j\hbar}\psi^*(\mathbf{r})\,d\mathbf{r} \tag{13}$$

Now the Hamiltonian H (equation 9) is replaced:

$$\frac{d}{dt}\mathbf{r} = \int \psi^*(\mathbf{r})\mathbf{r}\Big(\frac{1}{j\hbar}\frac{-\hbar^2}{2m}\Delta + \underbrace{\frac{1}{j\hbar}V(r)}_{0}\Big)\psi(\mathbf{r}) + \psi(\mathbf{r})\mathbf{r}\Big(\frac{1}{j\hbar}\frac{-\hbar^2}{2m}\Delta + \underbrace{\frac{1}{j\hbar}V(r)}_{0}\Big)\psi^*(\mathbf{r})\,d\mathbf{r} \tag{14}$$

The terms in equation 14 are zero because they are imaginary and describe a potential. But potential must always be real, so the terms become zero. After cancelling minus and $\hbar$:

$$\frac{d}{dt}\mathbf{r} = \int \left(\psi^*(\mathbf{r})\mathbf{r}\frac{-\hbar}{2mj}\Delta\psi(\mathbf{r}) + \psi(\mathbf{r})\mathbf{r}\frac{\hbar}{2mj}\Delta\psi^*(\mathbf{r})\right)\,d\mathbf{r} \tag{15}$$

$$= \frac{\hbar}{2mj}\left(\int \psi^*(\mathbf{r})\mathbf{r}\Delta\psi(\mathbf{r})\,d\mathbf{r} + \int \psi(\mathbf{r})\mathbf{r}\Delta\psi^*(\mathbf{r})\,d\mathbf{r}\right) \tag{16}$$

Following identity is used now:

$$-\psi^*\Delta\psi + \psi r\Delta\psi^* = -\nabla(\psi^*\nabla\psi - \psi r\nabla\psi^*) + (\nabla\psi^* r)\nabla\psi - (\nabla\psi r)\nabla\psi^* \tag{17}$$

in order to obtain

$$\frac{d}{dt}\mathbf{r} = \frac{\hbar}{2mj}\Big[-\int \nabla(\psi^*(\mathbf{r})\nabla\psi(\mathbf{r}) - \psi(\mathbf{r})\mathbf{r}\nabla\psi^*(\mathbf{r}))\,d\mathbf{r} \tag{18}$$

$$+\int (\nabla\psi^*(\mathbf{r})\mathbf{r})\nabla\psi(\mathbf{r})\,d\mathbf{r} \tag{19}$$

$$-\int (\nabla\psi(\mathbf{r})\mathbf{r})\nabla\psi^*(\mathbf{r})\,d\mathbf{r}\Big] \tag{20}$$

Using the theorem of Gauss:

$$\int_V \nabla\psi\,dr = \int_{\partial V} df\,\psi \tag{21}$$

the first (volume) integral can be converted to a surface integral:

$$\frac{d}{dt}\mathbf{r} = \frac{\hbar}{2mj}\Big[-\int d\mathbf{f}(\psi^*(\mathbf{r})\nabla\psi(\mathbf{r}) - \psi(\mathbf{r})\mathbf{r}\nabla\psi^*(\mathbf{r})) \tag{22}$$

$$+\int(\nabla\psi^*(\mathbf{r})\mathbf{r})\nabla\psi(\mathbf{r})\,d\mathbf{r} \tag{23}$$

$$-\int(\nabla\psi(\mathbf{r})\mathbf{r})\nabla\psi^*(\mathbf{r})\,d\mathbf{r}\Big] \tag{24}$$

The wave function must be squared-integrable, i.e.

$$\int_{-\infty}^{\infty}|f(x)|^2\,dx < \infty \tag{25}$$

Therefore the surface integral must vanish. With further simplification I get:

$$\begin{aligned}
\frac{d}{dt}\mathbf{r} &= \frac{\hbar}{2mj}\left[\int(\nabla\psi^*(\mathbf{r})\mathbf{r})\nabla\psi(\mathbf{r})\,d\mathbf{r} - \int(\nabla\psi(\mathbf{r})\mathbf{r})\nabla\psi^*(\mathbf{r})\,d\mathbf{r}\right] \\
&= \frac{\hbar}{2mj}\left[\int(\mathbf{r}\nabla\psi^*(\mathbf{r}) + \psi^*(\mathbf{r})\nabla\mathbf{r})\nabla\psi(\mathbf{r})\,d\mathbf{r} - \int(\mathbf{r}\nabla\psi(\mathbf{r}) + \psi(\mathbf{r})\nabla\mathbf{r})\nabla\psi^*(\mathbf{r})\,d\mathbf{r}\right] \\
&= \frac{\hbar}{2mj}\int\left(\psi^*(\mathbf{r})\nabla\mathbf{r}\nabla\psi(\mathbf{r}) - \psi(\mathbf{r})\nabla\mathbf{r}\nabla\psi^*(\mathbf{r})\right)\,d\mathbf{r} \\
&= \frac{\hbar}{2mj}\int\left(-\nabla(\psi^*(\mathbf{r})\psi(\mathbf{r})) + 2\psi^*(\mathbf{r})\nabla\psi(\mathbf{r})\right)\,d\mathbf{r}
\end{aligned}$$

Again applied the theorem of Gauss (equation 21):

$$\int_V -\nabla(\psi^*(\mathbf{r})\psi(\mathbf{r}))\,d\mathbf{r} = -\int_{\partial V} d\mathbf{f}\psi^*(\mathbf{r})\psi(\mathbf{r}) \tag{26}$$

and for the same reason as before, the surface integral vanishes leaving only the second integral:

$$\begin{aligned}
\frac{d}{dt}\mathbf{r} &= \frac{\hbar}{2mj}\int 2\psi^*(\mathbf{r})\nabla\psi(\mathbf{r})\,d\mathbf{r} \\
&= \frac{1}{m}\int\psi^*(\mathbf{r})\frac{\hbar}{j}\nabla\psi(\mathbf{r})\,d\mathbf{r} \\
&= \int\psi^*(\mathbf{r})\underbrace{(-\hbar j\nabla)}_{\mathbf{p}}\psi(\mathbf{r})\,d\mathbf{r}
\end{aligned}$$

$$\boxed{\frac{d}{dt}\mathbf{r} = \frac{1}{m}\int\psi^*(\mathbf{r})\underbrace{(-\hbar j\nabla)}_{\mathbf{p}}\psi(\mathbf{r})\,d\mathbf{r} = \frac{1}{m}\langle\psi|\hat{\mathbf{p}}|\psi\rangle = \frac{1}{m}\hat{\mathbf{p}} = \hat{\mathbf{v}}}$$

□

2 Task 3 - Floating Gate MOSFET

2.1 The Surface Potential

In this task it should be shown that the surface potential is given by

$$\phi_F = \frac{C_1 V_G + Q_F}{C_1 + C_2} \tag{27}$$

The equivalent circuit diagram for the floating gate MOSFET can be drawn as in figure 1:

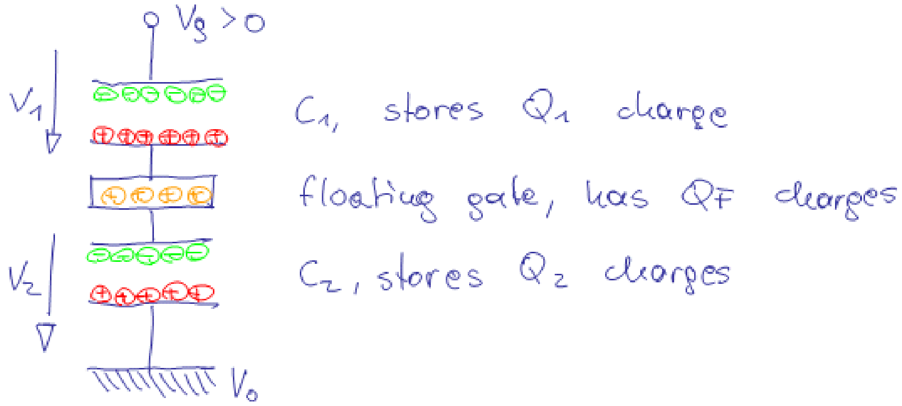

Figure 1: Floating gate

Proof. The derivation begins using the **charge conservation theorem** with Q_F as the reference point. As no charges are destroyed or generated, the following equation must hold:

$$Q_F = +Q_1 + (-Q_2) = Q_1 - Q_2 \tag{28}$$

Now replace the charges using the relation $Q = C \cdot U$:

$$Q_F = C_1 V_1 - C_2 V_2 \tag{29}$$

where V_1 and V_2 can be expressed with the potential ϕ_F on the floating gate:

$$V_1 = \phi_F - V_G \tag{30}$$

$$V_2 = V_0 - \phi_F = -\phi_F \quad \text{as } V_0 = 0 \Rightarrow \text{grounded} \tag{31}$$

Inserting equations 30 and 31 into equation 29:

$$Q_F = C_1(\phi_F - V_G) - C_2(-\phi_F) \tag{32}$$

$$= C_1\phi_F - C_1 V_G + C_2\phi_F \tag{33}$$

$$= \phi_F(C_1 + C_2) - C_1 V_G \tag{34}$$

$$\phi_F(C_1 + C_2) = C_1 V_G + Q_F \tag{35}$$

and finally solving for ϕ_F:

$$\boxed{\phi_F = \frac{C_1 V_G + Q_F}{C_1 + C_2}}$$

□

2.2 Floating Point Potential as Function of t

In this task the surface potential ϕ_F (equation 27) should be described as a function of t when the input voltage is raised to V_{pp} at time $t = 0$. In equation 27 everything is constant except V_G itself at the floating gate charges Q_F, yielding:

$$\phi_F = \frac{C_1 V_{pp} + Q_F(t)}{C_1 + C_2} \tag{36}$$

The charges change only through the tunneling current which can be described as:

$$I = A \cdot E^2 e^{-\frac{B}{E}} \tag{37}$$

where A and B are constant and E represents the electric field in the substrate. E is dependent on V_G but as this is constant, E itself is constant over time.

Charges over time can be obtained by integrating current over time, so integrating equation 37 yields:

$$Q_F(t) = \underbrace{Q_0}_{0} + \int_0^t I(t)\, dt \tag{38}$$

$$= \int_0^t I\, dt = I \cdot t\Big|_0^t = I \cdot t = tAE^2 e^{-\frac{B}{E}} \tag{39}$$

Q_0 cancels out because the assumption is having no initial charges at $t = 0$. Inserting equation 39 in 36 gives the result

$$\boxed{\phi_F = \frac{C_1 V_{pp} + tAE^2 e^{-\frac{B}{E}}}{C_1 + C_2}}$$

It should be noted however, that E is a function of the oxide thickness. In a theoretical description the field in a capacitor is homogenous, that means that E is piecewise constant over the substrate. A different approach might be to calculate the tunneling current I separately for both capacitors with their respective constant E and insert them together into equation 36.

3 Task 2 - Tunneling in Rectangular Potential

3.1 Calculation of Transmission Coefficient

Attached handwriting.

3.2 Discussion

As far as I understand the task, the following limit should be calculated:

$$\lim_{\hbar \to 0} T(\hbar) \tag{40}$$

Equation 40 clearly goes to zero in this case, if all other parameters are set to a constant value. Unfortunately I can give you no explanation why this should not be the case. But in the last days I did excessive research on this topic. I searched papers, IEEE, spent time at the library, searched google, books and even discussed with friends studying physics in Austria (my home country). But nobody could give me an explanation and there is no source discussing this issue.

From an intuitive point of view it should go to zero as everytime when we have $\hbar$ in our expressions, they need to produce classical results if applied to huge objects outside quantum world. Otherwise there would be something wrong. Concerning the reflection coefficient: For sure, there is always a small probability that a particle will pass this barriere, even if applied to non-quantum world. Although the probability is terrible low and unrealistic. But the signification of a limit is not just setting the value to zero but approaching to it. So if $\hbar$ becomes always smaller, smaller and smaller, the probability will also get smaller, smaller and smaller. But in the limit, the probability need to be zero, although at $\hbar = 0 + \epsilon$ there is a small probability left. Another issue: If the value of equation 40 should not be zero then there are only two possibilities for a limit: Either diverging or a fixed value, maybe dependent on the other parameters. This would mean that that if I approach $\hbar \to 0$ I would always get the same *finite* probability that a particle is passing.

The one source I found which states that equation 40 should not be zero is Wikipedia[1] where tunneling is derived with the semi-classical approach:

> If we take the classical limit of all other physical parameters much larger than Planck's constant, abbreviated as $\hbar \to 0$, we see that the transmission coefficient correctly goes to zero. This classical limit would have failed in the unphysical, but much simpler to solve, situation of a square potential.

I wanted to find the author who posted this to Wikipedia in order to discuss with him what is his source and why this should be the case. I discovered that this entry was made somewhere in 2004 by a 15 years old pupil from the United States. All contact data (e-mail, homepage) disappeared from the web.

[1] `http://en.wikipedia.org/wiki/Quantum_tunneling`

Apart from that I found dozens of books about quantum mechanics which all discuss the tunneling by the rectangular barrier and most of them do not discuss any problems with the derivation, in particular not what happens if $\hbar \to 0$.

Finally I found a few sources, including books, which clearly states that equation 40 must be zero. As an example I will give the book "Quantum Mechanics - Concepts and Applications" [1, p. 229]:

> Taking the classical limit $\hbar \to 0$, the coefficients (4.66) and (4.67) reduce to the classical result: $R \to 1$ and $T \to 0$.

I also tried thinking about other ideas: Maybe I misunderstand and a different limit is meant? Maybe it has something to do that the rectangular barriere itself is unrealistic? For example when I take the Fourier transform of a rectangular, I will need an infinite amount of sinoids to express the sharp step and there is also an effect called *ringing* which only disappears in the limit.

If I am wrong with my assumptions I would be very interested about the solution of this task.

4 Task 8 - My Opinion About Quantum Mechanics

Let me begin with physics and philosophy. The discovery of quantum mechanics was for sure one of the greatest events in science in the last century. It drastically changed the view about physics. Before everything was thought to be completely deterministic and repeatable. And with proper equipment we have the opportunity to make errors in experiments as small as we like. But quantum physics tells us now that an uncertainity is inherent in physics itself. This does not only concern the theory itself but also opens philosophical questions: Is there any randomness in nature or is everything completely deterministic (as believed before)? Which role has the observing person? For classical theory it does not matter if there is an observing person or not. In quantum physics it suddenly it is an important issue. Related to this is the question of reality: In classical physics we think that each measurement is part of a real physical object and if we measure something it tells us more about the real world. In quantum physics it can be proven that there are measurements where the result can not have been known before the measurement. This leads to the question if everything we can measure is *real* or also: What is reality?

Concerning engineering we have seen that if sizes become very small (in the region of $\hbar$) classical theorems do not hold any more but effects can be explained through quantum mechanics. In the last century the developement of the computer was maybe the most important invention. The first computers were only able to do simple calculations but powerful computers we know from today can be built with the same concept by just using massive amount of switches. Until now the growth of switches (transistors) per area can be described with the well-known *Moore's law*. It is exponential. Even today this law still holds and this means that transistors need to be exceptionally small. The new announced Intel Itanium 2 CPU (*Tukwila*) for example has over 2 billions of transistors in just 700 square millimeters! In this range the transistors are so small that classical field theories do not appropriately describe the behaviour anymore but it is necessary to use quantum mechanics.

A practical example directly related is the use of specific valleys. The drain current is directly dependent on the mobility of electrons and the mobility in turn is dependent on the relaxation time and the the conductive mass in parallel to the MOS interface. So it is very important to have a low conductive mass. *2-fold valleys* have a lower conductive mass than 4-fold valleys. So by designing the transistor in a way that the electron occupancy in the 2-fold valleys are increased makes it possible to build smaller size transistors.

Another very prominent example is Flash memory, a revolution in todays life. Flash memory is based on floating gate transistors where a certain charge is injected in the floating gate in order to store some information. This injection is done by *Fowler-Nordheim tunneling* which would not be possible using classical physics. The development of Flash memory made the rapid development of digital fotography possible and also revolutionized storage in general. Without the invention of Flash memory maybe we would still use floppy disks. Devices whose "intelligence" (firmware) can be upgraded or changed by the user are possible with Flash ROM. Also the rapid progress of mobile devices was only possible with cheap and reliable storage as Flash ROM.

Of course, there are also other fields which make use of quantum mechanic's tunnel effect: The electron miscroscope for example would not exist without the discovery of tunneling.

But when we think of computer engineering we should not only focus on transistors. The usage of transistors can not guarantee Moore's law for a long time. But there are scientists who generalized Moore's law and see it just as a part of a more complex law: A law which states that developement of nature grows exponentially: The beginning was the developement of the universe, the evolution, followed by human intelligence which includes exponential progress in science and developement. If such a law is true it means that even if transistors reach their limit there will be new technology which displaces the old one in order to preserve the law. And I think new technologies will even rely more on quantum effects.

One example of such technology might be a new approach in computation which is based on neural networks and intelligence. Until now, artificial neurons can be built in computers but they have to be "simulated" with many transistors (e.g. on a computer). With floating gate MOSFETs it might be possible to built up neural networks directly in hardware - by using MOSFETs not only as dumb switches but as energy-saving neurons. They are not working in binary mode any more but with continous signals.

Another new technology might be quantum computers which are also based fundamentally on quantum mechanics. Nowadays such computers (except those consisting of very few gates) can not yet be built but the theory suggests the powerful nature of quantum computers: The introduction of a *q-bit* (quantum bit) can store any value as a superposition of the values 0 and 1. When reading out the quantum register the Schroedinger equation collapses and the observer gets a value of 0 or 1. The output is dependent on the coefficients of the linear superposition of the quantum states. This means: A quantum computer can perform calculation with all possible register values simultaniously, as long as the content of the register is not read out! This way a quantum computer can calculate specific operations highly in parallel - if used in an appropriate way. An example which shows this power and also the implications on it is the well known *Shor Algorithm* which makes it possible to solve the integer factorization problem in polynomial time. The algorithm is based on a classical part which reduces the factorization problem to the problem of finding the rank of a multiplicative group of integers modulo n, mathematically: Find the smallest r such that

$$x^r \equiv 1 (\bmod N)$$

This problem in turn is equivalent to finding the period of the signal which is done by the quantum fourier transform. This quantum algorithm can be done in $O(\log_2 N)$. The principle is simple: Initialize a quantum register with a uniform distribution of length $\log_2 N$, then compute the function with a specific quantum circuit but without measuring the state. As long as there is no measurement, the quantum register contains the solution for all values of x. After reading out, the wave function collapses and produces values of 0 and 1 described by their respective probabilities. These probabilities have their highest value at the position of a multiple of r.

Almost all public key cryprography is based on the fact that factorization of an integer in its prime factors takes at least over-polynomial time. So the practical application of Shors algorithm would mean the apocalypse of the global cryptography infrastructure. The one technology which would survive quantum computers are *lattices* but they require a keylength of one megabyte for same security as for example 1024-bit RSA. There are lots of other algorithms which could exploit the power of quantum computers such as the *Grover search algorithm* which can search in an unordered database in $O(\sqrt{n})$.

A third big issue related in computer science is quantum cryptography. Nowadays in cryptography there is always a tradeoff between resources and security. The one proven but also unrealistic encryption scheme is the One-Time pad which needs a key of the same length as the data. All other schemes just rely on the fact that exhaustive keysearch needs too much computational resources (usually exponential time with number of bits). Quantum cryptography is another encryption scheme where it can physically proven that it is secure: It is based on a similar principle as described above: The sender generates a certain q-bit (usually identified by polarization or spin) which is destroyed as soon as information is read out. So the system can detect if an attacker eavesdrop the channel. There are already quantum cryptography systems available commercially.

However it must be noted that the system in the present state has many problems and weaknesses:

- The quantum channel is only used for secure key exchange. The following cryptography is done with classical block ciphers (such as AES)
- The exchange is based on the fact that the sender choses specific polarisations and transmits them to the receiver in a *classical* way. This is needed in order to verify which q-bits are correct (and should be used) and which not. This means that a classical **authentic** channel must exist which usually also relies on classical algorithms.
- The algorithm is probababilistic. The receiver does not know *which* q-bits might have been eavesdropped, he can just compare which q-bits to chose based on the list from above. So the receiver can only calculate the error ratio and detect an eavesdropper based on a specific threshold above which key exchange is restarted.
- For this reason an attacker has also a small chance of success.
- Heavy noise additionally distorts the transmission in the channel.

For those and many other reasons many crytographers consider quantum cryptography not as a solution for classical cryptography. Also, because of the mix with classical algorithms and the higher complexity, quantum cryptographie systems can be attacked in different ways. However, the technology shows what is possible and there might be a lot of ways to improve it.

My examples should show how important quantum cryptography is in microelectronics and computer science in the present time and how important it might be in future.

References

[1] Zettili, Nouredine *Quantum Mechanics - Concepts and Applications*, Second Edition, Wiley