Physical limits for scaling of integrated circuits

Waldemar Nawrocki

Poznan University of Technology, Faculty of Electronics and Telecommunications, ul. Piotrowo 3A, PL 60965 Poznan, Poland

E-mail: nawrocki@et.put.poznan.pl

Abstract. In this paper we discuss some physical limits for scaling of devices and conducting paths inside of semiconductor integrated circuits (ICs). Since 40 years only a semiconductor technology, mostly the CMOS and the TTL technologies, are used for fabrication of integrated circuits in the industrial scale. Miniaturization of electronic devices in integrated circuits has technological limits and physical limits as well. In 2010 best parameters of commercial ICs shown the dual-core Intel Core i5-670 processor manufactured in the technology of 32 nm. Its clock frequency in turbo mode is 3.73 GHz. A forecast of the development of the semiconductor industry (ITRS 2009) predicts that sizes of electronic devices in ICs circuits will be smaller than 10 nm in the next 10 years. The physical gate length in a MOSFET will even amount 7 nm in the year 2024. At least 5 physical effects should be taken into account if we discuss limits of scaling of integrated circuits.

1. Introduction

Scaling of electronic devices in integrated circuits has technological limits and physical limits as well. Since 40 years only a semiconductor technology, mostly the CMOS and the TTL technologies, are used for fabrication of integrated circuits in the industrial scale. Probably the CMOS technology will be used at least in the next 10-15 years. In 2010 best parameters of commercial ICs shown the dual-core Intel Core i5-670 processor manufactured in the technology of 32 nm. Its clock frequency in turbo mode is 3.73 GHz. The other example is the quad-core Intel Core i7-975 manufactured in the technology of 45 nm. The last forecast of the development of the semi-conductor industry (ITRS 2009) predicts that sizes of electronic devices in ICs circuits will be smaller than 10 nm in the next 10 years [1]. The physical gate length in a MOSFET will even amount 7 nm (see Table 1) in the year 2024. One can notice that the previous forecast (ITRS 2007) gave different values of the predicted length of a physical gate in ICs: 9 nm in 2016 and 4 nm in 2022.

At least 5 physical effects should be taken into account if we discuss limits of miniaturization of integrated circuits:

- spread of doping atoms in a semiconductor material; each dopant would induce a relatively high potential bump;
- quantization of both electrical and thermal conductance in narrow and thin transistors’ channels and in conducting paths;
- propagation time of electromagnetic wave along and across a chip (IC);
- electrostatics; a loss of electrostatic control of the drain current vs the gate voltage;
electron tunneling between a source and a drain inside a MOSFET through a insulation (oxide).

Table 1. Data of integrated circuits (IC) according to the Report of The International Technology Roadmap for Semiconductors (Edition 2009) [1]

| Year     | 2009 | 2012 | 2015 | 2018 | 2021 | 2024 |
|----------|------|------|------|------|------|------|
| Physical gate length in a FET transistor inside of ICs (microprocessor unit – MPU) (nm) | 29   | 22   | 17   | 12.8 | 9.7  | 7.4  |
| Clock frequency (on chip MPU) (GHz) | 5.45 | 6.82 | 8.52 | 10.6 | 13.3 | 16.6 |
| Functionality of IC (number of transistors) (mln) | 2212 | 4424 | 8848 | 17696| 35391| 70782|
| Supply voltage V | 1.0  | 0.9  | 0.81 | 0.73 | 0.66 | 0.60 |
| Dissipated power (cooling on) W | 143  | 158  | 143  | 136  | 133  | 130  |

We discuss spread of doping atoms in a semiconductor material and quantization of both electrical and thermal conductance in nanostructures. We mention only the other physical effects important for scaling of integrated circuits.

2. Spread of doping atoms in a semiconductor material

Classical theories of electrical and thermal conductance assume a huge number of atoms and free electrons. Let’s assume a silicon cube with one side dimension of $a$ and with common doping of $10^{16}$ cm$^{-3}$. In a n-doped silicon cube with the size of $(100 \text{ nm})^3$ there are $5 \times 10^7$ atoms and 10 free electrons at 300 K, but in the Si cube with the size of $(10 \text{ nm})^3$ there are $5 \times 10^4$ atoms and 1% chance only to find one free electron. Free electrons are necessary for electrical conductance as charge carriers. It means not only that classical theories of conductance are not valid for nanostructures. It means that common doping in semiconductor material is not sufficient for electronic devices of nanometric size.

![Figure 1. Free electrons in a silicon cube](image)

In order to keep the conductive properties of the semiconductor material one should apply more intensive doping, eg. $10^{20}$ cm$^{-3}$. However such intensive doping decreases resistivity of the material dramatically from $2 \times 10^{-3}$ Ωm to $10^{-5}$ Ωm, respectively (for n-type Si, at 300 K). Low number of free electrons should be scattered evenly in whole volume of a material.
3. Quantization of electrical conductance in nanostructures

Electric and thermal properties of electronic devices or paths with nanometer sizes are not more described by a classical theory of conductance but by quantum theories. The theoretical quantum unit of electrical conductance $G_0 = \frac{2e^2}{h}$ was predicted by Landauer in his theory of electrical conductance [2]. Parameters characterizing the system are a Fermi wavelength $\lambda_F$ ($\lambda_F = \frac{2\pi}{k_F}$, where $k_F$ is the Fermi wavevector), and a mean free path $\Lambda$. For metals like gold $\lambda_F \approx 0.5$ nm is shorter than free electron path $\Lambda (\Lambda_{Au} \approx 14$ nm). If a length of the system is shorter than the free electron path, the impurity scattering is negligible, so the electrons transport is ballistic. If a wire has outside diameter comparable with the Fermi wavelength $\lambda_F$, and its length $L$ is less than $\Lambda$, the system can be regarded as one-dimensional (1-D), the electron – as a wave, and one can expect quantum effects – see Fig 2.

Figure 2. Conductance quantization in a nanowire (conductor with length $L < \Lambda$ and width $W$ comparable with the length of Fermi wave $\lambda_F$): a) nanowire outline (the third dimension is not considered); b) conductance quantization $G$ versus width $W$

The total electrical conductance of a nanowire is given by formula (1).

$$G = \frac{2e^2}{h} N$$  \hspace{1cm} (1)

where $N$ is the number of transmission channels. For 1-D system, with thickness $H \leq \lambda_F$, $N$ depends on the width of the wire, $N = \text{int} \left( \frac{2W}{\lambda_F} \right)$. For 2-D system, with $H, W \geq \lambda_F$, $N = \text{int} \left( \frac{W \times H}{\lambda_F^2} \right)$, where int $(A)$ means the integer of $A$. However, defects, impurities and irregularities of the shape of the conductor can induce scattering, then conductivity is given by the Landauer equation:

$$G = \frac{2e^2}{h} \sum_{i,j=1}^{N} t_{ij}$$  \hspace{1cm} (2)

where $t_{ij}$ denotes probability of the transition from $j^{th}$ to $i^{th}$ state. In the absence of scattering $t_{ij} = \delta_{ij}$ thus Eq. (2) is reduced to Eq. (1). Figure 1 presents a picture of a path (nanowire) – the constriction in an electrical conductor with dimension $W$ (width), $H$ (thickness) and $L$ (length) $W$.

A set-up for measurements of electrical conductance in nanowires formed by mechanical contact between two microwires is shown in Figure 3. The experimental setup consisted of a pair of metallic wires (they formed a nanowire), a digital oscilloscope, a motion control system (doesn’t show on the picture) and a PC. Instruments are connected in one system using the IEEE-488 interface. There was the resistor $R_p = 1$ kΩ in series to the connected wires. The circuit was fed by the constant voltage $V_s$ and measurements of current $I(t)$ have been performed.
Transient effects of making contact or breaking the contact give time dependent current. The voltage $V_p$ on the resistor $R_p$ was measured with computer controlled oscilloscope. The piezoelectric device is used to control the backward and forward movement of the macroscopic wires between which nanowires are formed. A high voltage amplifier controlled by a digital function generator supplies the piezoelectric device. Both electrodes (macroscopic wires) are made of wire 0.5 mm in diameter. The conductance was measured between two metallic electrodes, moved to contact by the piezoelectric tube actuator. The oscilloscope was triggered by a single pulse. All experiments were performed at room temperature and at ambient pressure. The quantization of electric conductance depends neither on the kind of element nor on temperature. For conductors and semiconductors the conductance quantization in units of $G_0 = 2e^2/h = (12.9 \text{ k}\Omega)^{-1}$ was measured in many experiments.

The quantization of electric conductance depends neither on the kind of metal nor on temperature. However, the purpose of studying the quantization for different metals was to observe how the metal properties affect the contacts between wires. For nonmagnetic metals, the conductance quantization in units of $G_0 = 2e^2/h = 7.75 \times 10^{-5} \text{ [A/V]} = (12.9 \text{ k}\Omega)^{-1}$ was previously observed for the following nanowires: Au-Au, Cu-Cu, Au-Cu, W-W, W-Au, W-Cu. The quantization of conductance in our experiment was evident. All characteristics showed the same steps equal to $2e^2/h$. We observed two phenomena: quantization occurred when breaking the contact between two wires, and quantization occurred when establishing the contact between the wires. The characteristics are only partially reproducible; they differ in number and height of steps, and in the time length. The steps can correspond to 1, 2, 3 or 4 quanta. It should be emphasised that quantum effects were observed only for some of the characteristics recorded. The conductance quantization has been so far more pronouncedly observable for gold contacts. Figure 4 and 5 show example plots of conductance vs. time during the process of drawing Au- and Cu-nanowire, respectively, for the bias voltage $V_s = 0.42 \text{ V}$ [3].

Figure 3. A system for measurements of conductance quantization: an electrical circuit (left) and a piezoelectric actuator for precise moving of electrodes (right).

Figure 4. Conductance quantization in gold nanowires: a time plot (left) and a histogram from 6000 consecutive formations of a nanowire (measurements made by M. Wawrzyniak, PTU, Poland).
Let’s consider a silicon path with length $L = 20 \, \text{nm}$ ($L < L_{Si}$), width $W = 4 \, \text{nm}$, thickness $H = 4 \, \text{nm}$ and $\lambda_{F} = 1 \, \text{nm}$. Such path is a nanowire with the conductance: $G_{P} = (2e^{2}/h) \times N$; $N = \text{int} \left( \frac{W \times H}{\lambda_{F}^{2}} \right) = 16$; $G_{P} = 7.75 \times 10^{-5} \times 16 = 12.4 \times 10^{-4} \, \text{[A/V]}$. A resistance of the path $R_{P} = 1/G_{P} = 800 \, \Omega$, thus it is surprisingly high. An electrical capacity of a path is 2 pF/cm [4], so for $L = 20 \, \text{nm}$ $C_{P} = 4 \times 10^{-18} \, \text{F}$. The path inside an IC forms a low-pass RC filter with the upper frequency $f_{u} = 1/(2\pi R_{P} C_{P}) \approx 50 \times 10^{12} \, \text{Hz}$. Thus the path-nanowire filters of signals.

4. Quantization of thermal conductance in nanostructures

It is generally known that limits for speed-up of digital circuits, especially microprocessors, are determined by thermal problems. Both electrical $G_{E}$ and thermal $G_{T}$ conductance of a nanostructure describe the same process: electron transport in nanostructures. Therefore there are several analogues between the two physical quantities. Electron transport in a nanowire does two effects: an electrical current $I = G_{E} \times \Delta V$ and a heat flux density $Q_{D} = G_{T} \times \Delta T$, where $G_{E}$ – electrical conductance of a sample, $\Delta V$ – difference of electrical potentials, $G_{T}$ – thermal conductance of a sample, $\Delta T$ – temperature difference.

$$G_{E} = \sigma \times A/l, \quad G_{T} = \lambda \times A/l \quad (3)$$

where $\sigma$ – electrical conductivity, $\lambda$ – thermal conductivity, $l$ – length of a sample (e.g. nanowire), $A$ – area of a cross-section of a sample.

Quantized thermal conductance in one-dimensional systems (e.g. nanowires) was predicted theoretically by Rego [11] using the Landauer theory. The thermal conductance is considered in a similar way like the electrical conductance. In one-dimension systems are formed conductive channels. Each channel contributes to a total thermal conductance with the quantum of thermal conductance $G_{T0}$. Quantized thermal conductance and its quantum (unit) $G_{T0}$ was confirmed experimentally by Schwab [7]. The quantum $G_{T0}$ of thermal conductance

$$G_{T0} \, [\text{W/K}] = (\frac{\pi^{2} k_{B}^{2}}{3h})T = 9.5 \times 10^{13} T \quad (4)$$

depends on the temperature (10). At $T = 290 \, \text{K}$ value of $G_{T0} = 2.8 \times 10^{-10} \, \text{[W/K]}$. This value is determined for an ideal ballistic transport (without scattering) in a nanowire, with the transmission coefficient $t_{ij} = 100\%$. It means that in all practical cases (for $t_{ij} < 100\%$) the thermal conductance is below the limit given by formula (10). However an analyze of thermal conductance is more complex than electrical conductance because of contribution either phonons or electrons in heat exchange. Quantization of thermal conductance in one-dimensional systems was predicted theoretically by
Greiner [5] and by Rego [6] for ballistic transport of electrons and phonons. Quantized thermal conductance $G_T$ and its quantum (unit) $G_{T0}$ was confirmed experimentally by Schwab [7].

Table 2. Electrical resistance $R$ and thermal conductance $G_T$ of gold nanowire with diameter $D$ ($\lambda_F = 0.5$ nm)

| Diameter $D$ (nm) | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 |
|-------------------|-----|-----|-----|-----|-----|-----|
| Number of channels | –   | 1   | 4   | 9   | 16  | 25  |
| $R = 1/G_E$ (Ω)   | 12,903 | 3226 | 1434 | 806 | 516 | 358 |
| $G_T$ (at 300 K) [$10^{-9}$ [W/K]] | 0.285 | 1.14 | 2.56 | 4.56 | 7.12 | 10.26 |

The thermal conductance of a nanowire is very small. We can compare the thermal conductance in a nanowire with the thermal conductance of Au microwire which connects a silicon part with a metallic terminal (pin) in a transistor. For a microwire with a length of $L = 1.5$ mm and a diameter $D = 25$ μm is the thermal conductance is much larger: $G_T \approx 10^{-4}$ [W/K].

One can obtain a similar (but not the same) value of the quantum of thermal conductance $G_{T0}$ using the Wiedemann-Franz law. The law describes the relation between the thermal conductivity $\lambda$, and electrical conductivity $\sigma$ of a sample for macroscopic objects (5).

$$\frac{\lambda}{\sigma} = \frac{1}{3} \left( \frac{\pi k_B}{e} \right)^2 T = L \times T = 2.35 \times 10^{-8} T$$  \hspace{1cm} (5)

where $e$ – electron charge, $k_B$ – Boltzmann constant, $L$ – Lorenz number, $T$ – temperature.

The Wiedemann-Franz law is valid for the relation between conductances $G_T/G_E$ as well. The value of $G_{T0}$ can be obtained directly from (6).

$$G_{T0} = L \times T \times G_{E0} = \frac{1}{3} \left( \frac{\pi k_B}{e} \right)^2 T \times \frac{2e^2}{h} = \frac{2}{3} \left( \frac{\pi k_B}{e} \right)^2 T \times T = 1.89 \times 10^{-12} T$$  \hspace{1cm} (6)

The value of $G_{T0}$ obtained from the Wiedemann-Franz law is twice larger than this from (4) and it is not correct.

A single nanowire should be consider together with its terminals (Figure 6). They are colled reservoirs of electrons. A single nanowire should be consider together with its terminals. Electron transport in the nanowire is ballistic itself, it means the transport without scattering of electrons and without energy dissipation. The energy dissipation occurs in terminals. Because of the energy dissipation the local temperature $T_{term}$ in terminals is higher then the temperature $T_{wire}$ of nanowires itself. A heat distribution in terminals of a nanostructe should be analyzed.

In small structures a dissipated energy is quite large. For the first step of conductance quantization, $G_E = G_{E0} = 7.75 \times 10^{-5}$ [A/V], and at the supply voltage $V_s = 1.4$ V the current in the circuit $I = 100$ μA ($I = 190$ μA for the second step of quantization). The power dissipation in terminals of nanowires is $P = I^2/G_{E0} = 130$ μW for the first step and $P = 230$ μW for the second step. One ought to notice that the density of electric current in nanowires is extremely high. The diameter of the gold nanowire on the first step of quantization can be estimated to $D = 0.4$ nm, so for $I = 100$ μA the current density $J \approx 8 \times 10^{10}$ [A/cm²].
5. Other physical limits

The channel length $L_E$ of Si MOSFET, limited by a degradation of electrostatic control in the transistor, was analyzed by Frank [4] and Likharev [8]. The minimal channel length $L_E$ (see Fig. 5) depends on thickness of channel $H_{ch}$, thickness of insulation layer $H_i$, dielectric constants of a channel $\varepsilon$ and insulation $\varepsilon_i$ – formula (7) [8].

$$L_E = \left( \frac{\varepsilon H_{ch} H_i}{2\varepsilon_i} \right)^{1/2}$$

Figure 6. Conductance distribution in a nanowire with a ballistic transport

If we take the ratio $\varepsilon_i/\varepsilon \approx 0.3$ (for silicon oxide and silicon), thickness $H_{ch} = 2$ nm, $H_i = 1.5$ nm – the estimated minimal length of channel is $L_E \approx 3$ nm. The channel length $L_E$ can be shorter if a better insulator than silicon oxide (SiO$_2$) would be applied. Following materials are tested as an insulator for MOSFETs: silicon nitride, hafnium oxide and zirconium silicate.

The next limit to miniaturization of electronic devices comes from source to drain tunneling through the potential barrier along the channel. The tunneling effect depends on the channel length $L$ and the supply voltage. Because of tunneling the minimal channel length for silicon device is around 2 nm [8].

High-tech microprocessors have dimensions over 30 mm. A new Intel product the quad-core Intel Core i5-670 has a width $a = 37.5$ mm and a length $b = 37.5$ mm. The pins of the IC are distributed on the bottom plate, along its four sides too. The period of the highest clock frequency (3.73 GHz) is $T_{ck} = 270$ ps. But the propagation time for an electromagnetic wave on the way of $(a + b)$ is: $T_p = (a + b)/v = 75$ mm/2.5×10$^8$ m/s = 300 ps ($v$ – speed of an electromagnetic wave in silicon). Thus, the propagation time of the clock signal is longer than the period of the clock signal.

Figure 7. MOSFET transistor – the simple model
6. Conclusions

Conductance quantization has proved to be observable in a simple experimental setup, giving opportunity to investigate subtle quantum effects in electrical conductivity. The energy dissipation in nanowires takes part in their terminals. Because of the energy dissipation the local temperature in terminals is higher than the temperature of a nanowire itself. According to the state of the art, the minimal length of gate in MOSFET in silicon integrated circuits is around 3 nm (thus – technology of 3 nm!). However, technological limits allow to apply only the 10 nm-technology in the next 10 years.

References
[1] International Technology Roadmap for Semiconductors, 2010, Edition 2009, www.itrs.net/reports
[2] Landauer R 1989 *J. Phys.: Cond. Matter* vol. 1 p 8099
[3] Nawrocki W 2008 *Electron transport in nanosystems*, eds Bonca J, Kruchinin S (Springer: Dordrecht) p 17
[4] Frank D J, Dennard R H, Nowak E, Solomon P M, Taur Y and H.S.P. Wong H S P 2001 *Proc. IEEE* vol. 89 p 259
[5] Greiner A, Reggiani L, Kuhn T and Varani L 1997 *Phys. Rev. Lett.* vol. 78 p 1114
[6] Rego L G C and Kirczenow G 1998 *Phys. Rev. Lett.* vol. 81 p 232
[7] Schwab K, Henriksen E A, Worlock J M and Roukes M L 2000 *Nature* vol. 404 p 974
[8] Likharev K 2003 *Advanced Semiconductor and Organic Nano-techniques*, ed Markoč H. (Amsterdam: Elsevier) chapter 4