1 Introduction

Why does computation require energy?

Because there must be some irreversibility to ensure that calculations go forward (from inputs to outputs) and not in reverse, and because logical erasure necessarily implies dissipation because of the compression of phase-space.

What is a quantum computer?

One that operates on quantum bits that can be in a superposition of many different states simultaneously and that maintain a connection (called entanglement) following an interaction. These properties change the computational order of many important problems, such as reducing factoring from requiring a time that is exponentional to polynomial in the number of bits.

What limits the bit density for semiconductor memory?

Lithography (constrained by the wavelength used to pattern a memory cell, and the resulting yield), electromigration (when too few atoms are used in a wire they move in response to currents), and capacitance (when too few electrons are used, the fluctuation in their number becomes significant).

What limits the bit density in a typical hard disk?

Magnetic domain wall energies, and the head height.

What limits the bit density for optical storage?

The diffraction limit for focusing light, which is proportional to the wavelength.

Why are twisted pairs twisted, and coaxial cables coaxial?

To reduce the generation of unwanted radiation and the sensitivity to interference, and to effectively guide the signal. Twisted pairs are best at low frequencies, and coaxial cables at high frequencies.

Where does electronic noise come from, and how does it limit data rates?

Thermodynamic fluctuations, defect scattering, and finite-size statistics. The capacity of a communications channel grows as the logarithm of the ratio of the energy in the signal and the noise.

What is a liquid crystal, and how does it modulate light?

It is a material that maintains long-range orientational ordering without translational ordering. Under an applied field it is able to rotate the direction of polarization of light, thereby modulating the intensity of the light if the material is enclosed between polarizers.

These questions are examples of the many ways in which familiar devices that detect, transmit, process, store, and deliver information operate surprisingly near fundamental physical limits. The goal of this book is to explore how such devices function, how they can be used, what the limits on their performance are, and how they might be improved. This will require developing familiarity with the physical governing equations for a range of types of behavior, and with the mathematical tools necessary to manipulate these equations. One important aim is to equip the reader to work out quantitative answers to questions such as these.

A note about pedagogy: reading about physics is as satisfying as reading about food or exercise. It can be useful, but there is no substitute for experience solving problems. Each chapter has problems that apply and develop the preceeding ideas, ranging from trivial calculations to open research questions. Since another goal of this book is to help develop problem-solving skills, consulting the supplied answers before a problem is attempted is entirely counter-productive because the real problems that will come after this book don't come with such handy answers.

And a note about epistemology: it is important to keep in mind the distinction between truth and models. I will be describing models for a variety of types of behavior; these are the product of both experimental observations and theoretical inferences. A good model should compactly explain what you already know and allow you to predict new things that you did not know, but it does not necessarily contain any guide to an underlying "truth." Some physicists believe that there is an ultimate "correct" answer that these models are approaching, and some violently disagree, yet all agree on the usefulness of the current set of models and on how to manipulate them. *Truth* and *Meaning* are concepts that one may choose to associate with these models, but their presence or absence does not affect the models' use. At most, they do guide what you choose to think about. This distinction is very important because, when faced with unexpected claims or results, there is a recurring danger of seeing particular models as privileged correct answers rather than being open-minded about judging evidence on its merits. The history of science is littered with conflicts arising from prior beliefs that were stronger than experimental observations.

2 Interactions, Units, and Magnitudes

Modern information technology operates over a spectacular range of scales; bits from a memory cell with a size of 10^{-7} meters might be sent 10^{7} meters to a geosynchronous satellite. It is important to be comfortable with the orders of magnitudes and associated interaction mechanisms that are useful in practice. Our first task will be to review the definitions of important units, then survey the types of forces, and finally look at typical numbers in various regimes.

2.1 UNITS

Many powers of ten have been named because it is much easier to say something like "a femtosecond optical pulse" than "a 0.000 000 000 000 001 second optical pulse" when referring to typical phenomena at that scale (a cycle of light takes on the order of a femtosecond). The dizzying growth of our ability to work with large and small systems pushes the bounds of this nomenclature; data from terabyte storage systems is read out into femtofarad memory cells. It is well worth memorizing the prefixes in Table 2.1.

Physical quantities must of course be measured in a system of units; there are many alternatives that are matched to different regimes and applications. Because of their interrelationships it is necessary only to define a small number of fundamental quantities to be able to derive all of the other ones. The choice of which fundamental definitions to use changes over time to reflect technological progress; once atomic clocks made it possible to measure time with great *precision* (small variance) and *accuracy* (small bias), it became more reliable to define the meter in terms of time and the speed of light rather than a reference bar kept at the Bureau International des Poids et Mesures (BIPM, `http://www.bipm.fr`) in Sevres, France. The kilogram is still defined in terms of a platinum–iridium cylinder held at BIPM instead of a fundamental physical process, a source of great frustration in the metrology community. Aside from the difficulty in duplicating it, the accumulation of contaminants on the surface increases the mass by about 1 part in 10^{9} per year, requiring that it be measured only after a special cleaning procedure [Girard, 1994].

The most common set of base defined quantities in use is the *Système International d'Unités* (*SI*) [BIPM, 1998]:

length: *meter* (m)

The meter is the length of path traveled by light in vacuum during a time interval of 1/299 792 458 of a second.

Table 2.1. *Orders of magnitude.*

Magnitude	Prefix	Symbol	Magnitude	Prefix	Symbol
10^{-24}	yocto	y	10^{24}	yotto	Y
10^{-21}	zepto	z	10^{21}	zetta	Z
10^{-18}	atto	a	10^{18}	exa	E
10^{-15}	femto	f	10^{15}	peta	P
10^{-12}	pico	p	10^{12}	tera	T
10^{-9}	nano	n	10^{9}	giga	G
10^{-6}	micro	μ	10^{6}	mega	M
10^{-3}	milli	m	10^{3}	kilo	k
10^{-2}	centi	c	10^{2}	hecto	h
10^{-1}	deci	d	10^{1}	deka	da

mass: *kilogram* (kg)

The kilogram is the unit of mass; it is equal to the mass of the international prototype of the kilogram.

time: *second* (s)

The second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom.

current: *ampere* (A)

The ampere is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed 1 meter apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newtons per meter of length. (See Problem 5.4.)

temperature: *kelvin* (K)

The kelvin, the unit of thermodynamic temperature, is the fraction of 1/273.16 of the thermodynamic temperature of the triple point of water. (Temperatures in degrees Celsius are equal to temperatures in kelvin + 273.15. The triple point is the temperature and pressure at which the liquid, solid, and gas phases of water co-exist. It is fixed at 0.01 °C, and provides a more reliable reference than the original centigrade definition of 0 °C as the freezing point of water at atmospheric pressure.)

quantity: *mole* (mol)

The mole is the amount of substance of a system which contains as many elementary entities as there are atoms in 0.012 kg of carbon 12 (i.e., Avogadro's constant $6.022\ldots\times10^{23}$).

intensity: *candela* (cd)

The candela is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} hertz and that has a radiant intensity in the direction of 1/683 watts per steradian. (The frequency corresponds to the wavelength of 555 nm where the eye is most sensitive, the factor of 683 comes from matching an earlier definition based on the emission from solidifying platinum, and a steradian is the solid angle subtended by a unit

Cambridge University Press
978-0-521-58044-1 - The Physics of Information Technology
Neil Gershenfeld
Excerpt
More information

area on the surface of a sphere with unit radius; see Chapter 11 for more on luminosity.)

From these seven fundamental units many other ones are derived in terms of them, including:

capacitance: *farad* F ($m^{-2} \cdot kg^{-1} \cdot s^4 \cdot A^2$)
The farad is the capacitance of a capacitor between the plates of which there appears a difference of potential of 1 volt when it is charged by a quantity of electricity equal to 1 coulomb.

charge: *coulomb* C ($A \cdot s$)
The coulomb is the quantity of electricity transported in 1 second by a current of 1 ampere.

energy: *joule* J ($m^2 \cdot kg \cdot s^{-2}$)
The joule is the work done when the point of application of a force of 1 newton is displaced a distance of 1 meter in the direction of the force. (Remember that energy equals force times distance.)

force: *newton* N ($m \cdot kg \cdot s^{-2}$)
The newton is that force which, when applied to a body having a mass of 1 kilogram, gives it an acceleration of 1 meter per second squared. (Remember that force equals mass time acceleration.)

illuminance: *lux* lx ($cd \cdot m^{-2}$)
The lux is equal to an illuminance of 1 lumen per square meter.

inductance: *henry* H ($m^2 \cdot kg \cdot s^{-2} \cdot A^{-2}$)
The henry is the inductance of a closed circuit in which an electromotive force of 1 volt is produced when the electric current in the circuit varies uniformly at a rate of 1 ampere per second.

luminous flux: *lumen* lm (cd)
The lumen is the luminous flux emitted within a unit solid angle of 1 steradian by a point source having a uniform intensity of 1 candela.

magnetic flux: *weber* Wb ($m^2 \cdot kg \cdot s^{-2} \cdot A^{-1}$)
The weber is the magnetic flux which, linking a circuit of 1 turn, produces in it an electromotive force of 1 volt as it is reduced to zero at a uniform rate in 1 second.

magnetic flux density: *tesla* T ($kg \cdot s^{-2} \cdot A^{-1}$)
The tesla is the magnetic flux density given by a magnetic flux of 1 weber per square meter.

power: *watt* W ($m^2 \cdot kg \cdot s^{-3}$)
The watt is the power which gives rise to the production of energy at the rate of 1 joule per second.

pressure: *pascal* Pa ($m^{-1} \cdot kg \cdot s^{-2}$)
The pascal is the pressure of 1 newton per square meter.

potential: *volt* V ($m^2 \cdot kg \cdot s^{-3} \cdot A^{-1}$)
The volt is the difference of electric potential between two points of a conductor carrying a constant current of 1 ampere, when the power dissipated between these points is equal to 1 watt.

Table 2.2. *Selected conversion factors.*

1 dyne $(\mathrm{gm}\cdot\mathrm{cm}\cdot\mathrm{s}^{-2})$	=	1×10^{-5} N
1 erg $(\mathrm{gm}\cdot\mathrm{cm}^2\cdot\mathrm{s}^{-2})$	=	1×10^{-7} J
1 horsepower (hp)	=	745.7 W
1 atmosphere (atm)	=	101325 Pa
1 ton (short)	=	2000 pounds
	=	907.18474 kg
1 electron volt (eV)	=	$1.602176462\times10^{-19}$ J
1 amu	=	1.66053873×10^{-27} kg
1 ångstrom (Å)	=	1×10^{-10} m
1 fermi (fm)	=	1×10^{-15} m
1 parsec (pc)	=	3.085678×10^{16} m
1 mile (mi)	=	1609.344 m
1 foot (ft)	=	0.3048 m
1 inch (in)	=	0.0254 m
1 liter (L)	=	0.001 m^3
1 pound (lb)	=	0.45359237 kg
1 pound-force (lbf)	=	4.44822 N

resistance: *ohm* Ω $(\mathrm{m}^2\cdot\mathrm{kg}\cdot\mathrm{s}^{-3}\cdot\mathrm{A}^{-2})$

> The ohm is the electric resistance between two points of a conductor when a constant difference of potential of 1 volt, applied between these two points, produces in this conductor a current of 1 ampere. (These derivative definitions of the volt and ohm have more recently been replaced by fundamental ones fixing them in terms of the voltage across a *Josephson junction* and the resistance steps in the *quantum Hall effect* [Zimmerman, 1998], and capacitance may be defined by counting electrons on a *Single-Electron Tunneling* (*SET*) device [Keller *et al.*, 1999].)

It is important to pay attention to the units in these definitions. Many errors in calculations can be caught by making sure that the final units are correct, and it can be possible to make a rough estimate of an answer to a problem simply by collecting relevant terms with the right units (this is the subject of *dimensional analysis*). Electromagnetic units are particularly confusing; we will consider them in more detail in Chapter 5. The SI system is also called *MKS* because it bases its units on the meter, the kilogram, and the second. For some problems it will be more convenient to use *CGS* units (based on the centimeter, the gram, and the second); MKS is more common in engineering and CGS in physics. A number of other units have been defined by characteristic features or by historical practice; some that will be useful later are given in Table 2.2.

It's often more relevant to know the value on one quantity relative to another one, rather than the value itself. The ratio of two values X_1 and X_2, measured in *decibels* (*dB*), is defined to be

$$\mathrm{dB} = 20\log_{10}\frac{X_1}{X_2} \quad . \tag{2.1}$$

If the *power* (energy per time) in two signals is P_1 and P_2, then

$$\mathrm{dB} = 10\log_{10}\frac{P_1}{P_2} \quad . \tag{2.2}$$

This is because the power is the mean square amplitude (Chapter 3), and so to be consistent with equation (2.1) a factor of 2 is brought in to account for the exponent. An increase of 10 db therefore represents a increase by a factor of 10 in the relative power of two signals, or a factor of 3.2 in their values. A change of 3 dB in power is a change by a factor of 2.

The name decibel comes from Bell Labs. Engineers there working on the telephone system found it convenient to measure the gain or loss of devices on a logarithmic scale. Because the log of a product of two numbers is equal to the sum of their logs, this let them find the overall gain of a system by adding the logs of the components, and using logarithms also made it more convenient to express large numbers. They called the base-10 logarithm the *bel* in honor of Alexander Graham Bell; multiplying by 10 to bring up one more significant digit gave them a tenth-bel, or a decibel.

Some decibel reference levels occur so commonly that they are given names; popular ones include:

- *dBV* measures an electrical signal relative to 1 volt
- *dBm* measures relative to a 1 mW signal. The power will depend on the (usually unspecified) load, which traditionally is 50 Ω for radiofrequency signals and 600 Ω for audio ones (loads will be covered in Chapter 6). In audio recording, this is also called the *Volume Unit* or *VU*.
- *dBspl*, for *Sound Pressure Level* (or just *SPL*), measures sound pressure relative to a reference of 2×10^{-5} Pa, the softest sound that the ear can perceive. The sound of a jet taking off is about 140 dBspl.

Finally, there are a number of fundamental observed constants in nature that we will use, shown in Table 2.3. In this list the digits in parentheses are the standard deviation uncertainty (see Chapter 3) in the corresponding digits, so that for example the error in the value for G is 0.010×10^{-11} (which, compared to the other constants, is an embarrasingly large uncertainty [Gundlach *et al.*, 1996]).

The speed of light no longer really belongs here, because its value has been defined exactly as part of the SI system. All the others are determined by exquisite metrology experiments. Each fundamental constant can appear in many different types of measurements, and these are done by many different groups, leading to multiple values that unfortunately don't always agree to within their careful error estimates. For this reason, the International Council of Scientific Unions in 1966 formed the *Committee on Data for Science and Technology* (*CODATA*) to do global optimizations over all these data to come up with an internally-consistent set of values. The most recent adjustment was done in 1998, and is available at `http://physics.nist.gov`.

2.2 PARTICLES AND FORCES

The world is built out of elementary particles and their interactions. There are a number of natural divisions in organization, energy, and length that occur between the structure of the nucleus of an atom and the structure of the universe; it will be useful to briefly survey this range in order to understand the relevant regimes for present and prospective information technologies.

Table 2.3. *Selected fundamental constants.*

gravitational constant (G)	=	$6.673(10)\times10^{-11}\ \mathrm{m^3\cdot kg^{-1}\cdot s^{-2}}$
speed of light (c)	=	$2.99792458\times10^{8}\ \mathrm{m/s}$
elementary charge (e)	=	$1.602176462(63)\times10^{-19}\ \mathrm{C}$
Boltzmann constant (k)	=	$1.3806503(24)\times10^{-23}\ \mathrm{J/K}$
Planck constant (h)	=	$6.62606876(52)\times10^{-34}\ \mathrm{J\cdot s}$
$\hbar = h/2\pi$	=	$1.054571596(82)\times10^{-34}\ \mathrm{J\cdot s}$
Avogadro constant (N_{A})	=	$6.02214199(47)\times10^{23}\ \mathrm{mol^{-1}}$
electron mass (m_e)	=	$9.10938188(72)\times10^{-31}\ \mathrm{kg}$
proton mass (m_p)	=	$1.67262158(13)\times10^{-27}\ \mathrm{kg}$
gas constant (R)	=	$8.314472(15)\ \mathrm{J\cdot mol^{-1}\cdot K^{-1}}$
vacuum permittivity (ϵ_0)	=	$10^7/(4\pi c^2) = 8.854188\ldots\times10^{-12}\ \mathrm{F/m}$
vacuum permeability (μ_0)	=	$4\pi\times10^{-7}\ \mathrm{H/m}$

This story starts with quantum mechanics, the laws that govern things that are very small. Around 1900 Max Planck was led by his inability to explain the spectrum of light from a hot oven to propose that the energy of light is quantized in units of $E = h\nu = hc/\lambda$, where ν is the frequency and λ is the wavelength; $h = 6.626\ldots\times10^{-34}\ \mathrm{J\cdot s}$ is now called *Planck's constant*. From there, in 1905 Einstein introduced the notion of massless photons as the discrete constituents of light, and in 1924 de Broglie suggested that the wavelength relationship applies to massive as well as massless particles by $\lambda = h/p$; λ is the *de Broglie wavelength*, and is a consequence of the *wave–particle duality*: all quantum particles behave as both waves and particles. An electron, or a photon, can diffract like a wave from a periodic grating, but a detector will register the arrival of individual particles. Quantum effects usually become significant when the de Broglie wavelength becomes comparable to the size of an object.

Quantum mechanical particles can be either *fermions* (such as an electron) or *bosons* (such as a photon). Fermions and bosons are as unlike as anything can be in our universe. We will later see that bosons are particles that exist in states that are symmetric under the interchange of particles, they have an integer spin quantum number, and multiple bosons can be in the same quantum state. Fermions have half-integer spin, exist in states that are antisymmetric under particle interchange, and only one fermion can be in a particular quantum state. Spin is an abstract property of a quantum particle, but it behaves just like an angular momentum (as if the particle is spinning).

Particles can interact through four possible forces: *gravitational*, *electromagnetic*, *weak*, and *strong*. The first two are familiar because they have infinite range; the latter two operate on short ranges and are associated with nuclear and subnuclear processes (the characteristic lengths are approximately 10^{-15} m for the strong force and 10^{-18} m for the weak force). The electromagnetic force is so significant because of its strength: if a quantum atom was held together by gravitational forces alone (like a miniature solar system) its size would be on the order of 10^{23} m instead of 10^{-10} m. The macroscopic forces that we feel, such as the hardness of a wall, are transmitted to us by the electromagnetic force through the electrons in our atoms interacting with electrons in the adjoining atoms in the surface, but can be much more simply described in terms of fictitious effective forces ("the wall is hard").

Cambridge University Press
978-0-521-58044-1 - The Physics of Information Technology
Neil Gershenfeld
Excerpt
More information

All forces were originally thought to be transmitted by an intervening medium, the long-sought *ether* for electromagnetic forces. We now understand that forces operate by the exchange of spin-1 gauge bosons – the photon for the electromagnetic interaction (electric and magnetic fields), the $W^{\pm}$ and Z^0 bosons for the weak interaction, and eight gluons for the strong interaction (there is not yet a successful quantum theory of gravity). *Quantum ElectroDynamics* (*QED*) is the theory of the quantum electromagnetic interaction, and *Quantum ChromoDynamics* (*QCD*) the theory of the strong interaction. The weak and electromagnetic interactions are united in the *electroweak theory*, which, along with QCD is the basis for the *Standard Model*, the current summary of our understanding of particle physics. This amalgam of experimental observations and theoretical inferences successfully predicts most observed behavior extremely accurately, with two important catches: the theory has 20 or so adjustable parameters that must be determined from experiments, and it cannot explain gravitation. *String theory* [Giveon & Kutasov, 1999], a reformulation of particle theory that starts from loops rather than points as the primitive mathematical entity, appears to address both these limitations, and so is of intense interest in the theoretical physics community even though it is still far from being able to make experimentally testable predictions.

The most fundamental massive particles that we are aware of are the *quarks* and *leptons*. There's no reason to assume that there's nothing below them (i.e., turtles all the way down); there's just not a compelling reason right now to believe that there is. Quarks and leptons appear in the scattering experiments used to study particle physics to be point-particles without internal structure, and are spin-1/2 fermions. The leptons interact through the electromagnetic and weak interactions, and come in pairs: the *electron* and the electron *neutrino* (e^-, ν_e), the *muon* and its neutrino (μ^-, ν_μ), and the *tau lepton* and its neutrino (τ^-, ν_τ). Muons and tau leptons are unstable, and therefore are seen only in accelerators, particle decay products, and cosmic rays. Because neutrinos interact only through the weak force, they can pass unhindered though a light-year of lead. But they are profoundly important for the energy balance of the universe, and if they have mass [Fukuda, 1998] it will have enormous implications for the fate of the universe. Quarks interact through the strong as well as weak and electromagnetic interactions, and they come in pairs: *up* and *down*, *charm* and *strange*, and *top* and *bottom*. These fanciful names are just labels for the underlying abstract states. The first member of each pair has charge +2/3, the second member has charge −1/3, and each charge flavor comes in three colors (once again, flavor and color are just descriptive names for quantum numbers).

Quarks combine to form *hadrons*; the best-known of which are the two *nucleons*. A proton comprises two ups and a down, and the neutron an up and two downs. The nucleons, along with their excited states, are called *baryons* and are fermions. Transitions between baryon states can absorb or emit spin-1 boson hadrons, called *mesons*. The size of hadrons is on the order of 10^{-15} m, and the energy difference between excited states is on the order of 10^9 electron volts (1 GeV).

The nucleus of an atom is made up of some number of protons and neutrons, bound into ground and excited states by the strong interaction. Typical nuclear sizes are on the order of 10^{-14} m, and energies for nuclear excitations are on the order of 10^6 eV (1 MeV). Atoms consist of a nucleus and electrons bound by the electromagnetic interaction; typical sizes are on the order of 1 ångstrom (Å, 10^{-10} m) and the energy difference between states

Cambridge University Press
978-0-521-58044-1 - The Physics of Information Technology
Neil Gershenfeld
Excerpt
More information

is on the order of 1 eV. Notice the large difference in size between the atom and the nucleus: atoms are mostly empty space. Atoms can exist in different *isotopes* that have the same number of protons but differing numbers of neutrons, and *ions* are atoms that have had electrons removed or added.

Atoms can bond to form molecules; bond energies are on the order of 1 eV and bond lengths are on the order of 1 Å. Molecular sizes range from simple diatomic molecules up to enormous biological molecules with 10^6–10^9 atoms. Large molecules fold into complex shapes; this is called their *tertiary structure*. These shapes are responsible for the geometrical constraints in molecular interactions that govern many biochemical pathways. Predicting tertiary structure is one of the most difficult challenges in chemistry.

Macroscopic materials are described by the arrangement of their constituent atoms, and include crystals (which have complete long-range ordering), liquids and glasses (which have short-range order but little long-range order), and gases (which have little short-range order). There are also very interesting intermediate cases, such as quasiperiodic alloys called *quasicrystals* that have deterministic translational order without translational periodicity [DiVincenzo & Steinhardt, 1991], and *liquid crystals* that maintain orientational but not translational ordering [Chandrasekhar, 1992]. Most solids do not contain just a single phase; there are usually defects and boundaries between different kinds of domains.

The atomic weight of an element is equal to the number of grams equal to one mole ($N_A \approx 10^{23}$) of atoms. It is approximately equal to the number of protons and neutrons in an atom, but differs because of the mix of naturally occuring isotopes. 22.4 liters of an ideal gas at a pressure of 1 atmosphere and at room temperature will also contain a mole of atoms.

The structure of a material at more fundamental levels will be invisible and can be ignored unless energies are larger than its characteristic excitations. Although we will rarely need to descend below atomic structure, there are a number of important applications of nuclear transitions, such as nuclear power and the use of nuclear probes to characterize materials.

2.3 ORDERS OF MAGNITUDE

Understanding what is possible and what is preposterous requires being familiar with the range of meaningful numbers for each unit; the following lists include some significant ones:

Time

10^{-43} s: the Planck time (Problem 2.7)
10^{-15} s: this is the period of visible light, and a typical time scale for chemical reactions
10^{-9} s: atomic excitations and molecular rotations typically have lifetimes on the order of nanoseconds, and this is the clock cycle for the fastest computers
10^{-3} s: the shortest time difference that is consciously perceptible by people
10^{17} s: the approximate age of the observable universe