Ultra Low Power Bioelectronics

Fundamentals, Biomedical Applications, and Bio-inspired Systems

Chapter 26

RAHUL SARPESHKAR
Massachusetts Institute of Technology
Energy harvesting and the future of energy

Nature uses only the longest threads to weave her patterns, so that each small piece of her fabric reveals the organization of the entire tapestry.

Richard P. Feynman

Energy surrounds us, is within us, and is created by us. In this chapter, we shall discuss how systems can harvest energy in their environments and thus function without needing to constantly carry their own energy source. The potential benefits of an energy-harvesting strategy are that the lifetime of the low-power system is then not limited by the finite lifetime of its energy source, and that the weight and volume of the system can be reduced if the size of the energy-harvester is itself small. The challenges of an energy-harvesting strategy are that many energy sources are intermittent, can be hard to efficiently harvest, and provide relatively low power per unit area. Thus, energy-harvesting systems are usually practical only if the system that they power operates with relatively low power consumption.

We shall begin by discussing energy-harvesting strategies that have been explored for low-power biomedical and portable applications. First, we discuss the use of strategies that function by converting mechanical body motions into electricity. A circuit model developed for describing energy transfer in inductive links in Chapter 16 is extremely similar to a circuit model that accurately characterizes how such mechanical energy harvesters function. Thus, tradeoffs on maximizing energy efficiency or energy transfer are also similar. Energy harvesting with RF energy is discussed extensively in Chapters 16 and 17, so we shall not discuss it in this chapter. Then, we discuss the use of thermoelectric strategies that function by converting body heat into electricity. A fundamental thermodynamic principle limits the energy efficiency of a ‘heat engine’, whether in an internal combustion engine in a car, in a refrigerator, or in a thermoelectric device powered by body heat. The limiting efficiency is called the Carnot efficiency. The Carnot efficiency and models of heat flow from the body will help us understand the limits of operation of thermoelectric energy harvesting.

This book has largely discussed ultra-low-power systems at relatively small spatial scales in biomedical and in bio-inspired systems, mostly in the $10^{-12}$ W to $10^{-2}$ W range. In this final chapter, we shall see that principles of low-power design are also relevant to systems at large spatial scales with gigantic power...
consumption, e.g., a 40 kW gasoline-powered car moving at 30 mph, which if operated for 1 hour each day leads to an average power consumption of 1.67 kW.

The average human being on Earth consumes 2.5 kW of power such that our planet’s current aggregate power consumption is roughly 15 TW. The average power consumption of people in richer countries is higher than that in poorer countries. For example, the average person in the United States consumes 10.4 kW.\(^1\) We have been able to sustain such power consumption thus far largely because the 46,400 J/g energy density of gasoline, the 53,600 J/g energy density of natural gas, and the 32,500 J/g energy density of coal, and their relative abundance, have enabled us to burn energy at a profligate rate. In comparison, a well-optimized lithium-ion battery for portable applications operates at 650 J/g. Gasoline is currently cheaper per liter than bottled water in the United States.

For every kWh of oil, natural gas, or coal that is consumed, 250 g, 190 g, and 300 g, respectively, of CO\(_2\) is dumped into our atmosphere.\(^2\) This means that 5.5 tons of CO\(_2\) is generated on average per person per year, increasing CO\(_2\) levels by \(~2.5\) ppm (parts per million) per year today.\(^3\) The accumulation of CO\(_2\) has increased the atmospheric concentration from 280 ppm in pre-industrial times to \(~390\) ppm today.\(^4\) The pace of CO\(_2\) emissions is expected to increase significantly as India, China, and other developing nations output more CO\(_2\). For every ppm increase in CO\(_2\), the average Earth temperature appears to rise due to a greenhouse effect.\(^3\) Many climatologists believe that there will be serious and irreversible consequences to world climate, partly due to positive-feedback loops, if the CO\(_2\) concentrations increase significantly beyond 550 ppm.

The profligate burning of fossil fuels will lead to their inevitable extinction, which is not only catastrophic for energy and climate reasons, but also because they are quite useful for making several materials like plastics cheaply. Due to the need for minimizing fossil-fuel CO\(_2\) emissions that impact climate change and due to the exhaustion of these fossil-fuel energy sources, our planet will need to function increasingly on renewable energy sources. These sources include solar power, wind power, hydroelectric power, wave power, tidal power, geothermal power, and biofuels. Since the areal power densities of these sources are relatively small, it is imperative that our power consumption be reduced. Most of our power consumption arises from transportation, heating, electricity usage, and material-synthesis costs.

We discuss how electric cars, powered by batteries driving motors, enable improvements in transport energy efficiency, i.e., energy consumed per person-km, over those of gasoline-powered cars. We shall discuss an equivalent circuit for a car, which will allow us to draw on principles of low-power design in electronics to understand how power consumption in cars can and is being reduced. We shall compare the energy efficiency of advanced electric cars versus cheetahs, the fastest land animals on earth. Even though legged locomotion is significantly less

\(^1\) Interestingly, the average national per-capita income of a person in K\$/ divided by 4 is a good predictor of that nation’s average per-person power consumption in kW.
efficient than wheels on flat terrains, we shall see that animals have impressively good transport energy efficiency when compared with even highly energy-efficient electric cars.

We will focus on two renewable sources that are likely to be very important in our future, namely, solar photovoltaics and biofuels. The basic principles of phototransduction described in Chapter 11 will be useful for understanding how solar photovoltaic cells function. We shall delve deeper into phototransduction in this chapter to understand the limits of solar-cell efficiency. Solar photovoltaic sources are important at small scales, e.g., for solar photovoltaic cells that power portable and biomedical applications, and also at large scales, e.g., for 300 MW electric generators. Solar energy is widely viewed as the most important renewable energy source because of its relatively high power density and ubiquitous presence [5]. We shall discuss some challenges in making solar electricity generation cost effective. We conclude by discussing biofuels, which are created by plants storing the energy of sunlight in chemical bonds through the process of photosynthesis. Biofuels represent an energy-dense method for the storage and distribution of solar energy. Such biofuels could be useful in cars and in implantable biomedical systems in the future.

### 26.1 Sources of energy

Figure 26.1 shows six common sources of energy that we can harvest. We have discussed RF energy harvesting in near-field systems in Chapter 16 for biomedical implants and in far-field systems for cardiac monitoring in Chapters 17 and 20. In general, ambient RF energy from cell phones and wireless devices in the environment may be harvested. Implantable biomedical systems can potentially harness the energy of blood flow or the energy of airflow during respiration to function; work in this area is just beginning. Ultra-low-power outdoor monitoring applications can exploit potential differences between two points on a tree trunk, which can vary by a few hundreds of mV, to operate [6]. In this chapter, we shall primarily focus on inertial-motion, heat, and solar energy harvesting.

![Figure 26.1. A typical energy-harvesting architecture.](image)
At first, we shall only focus on harvesting at small scales for low-power biomedical applications.

Noninvasive and implanted biomedical systems can harvest the energy in the inertial motions of the bodily limbs or the head to which they are attached to function. Small ultra-low-power implanted systems that are attached to the heart or to the lungs can harness the mechanical energy of heart or lung motions to operate. The body is maintained at nearly 37°C while the environment is usually below this temperature. Therefore, the flow of heat energy from the body towards its surround can be harnessed in a thermoelectric device to operate electronics attached to the body. For example, the micropower EKG or PPG amplifiers discussed in Chapter 20 can be powered in such a fashion. Solar cells attached to the body can power noninvasive electronics attached near them as they now power watches. In general, several energy sources are intermittent, e.g., solar energy is only available during the day, mechanical energy is only available during motions of the body, and RF energy may only be available when there is a wireless device in the environment. The intermittency of the energy implies that there is need for storage of the harvested energy, e.g., in a battery or in a large capacitor as shown in Figure 26.1. The energy-storage system serves to smooth energy fluctuations such that power is always reliably available to the load. To prevent residual power-supply fluctuations output by the energy-storage system from affecting the electronics that it powers, and to ensure that there is good impedance matching for maximum or energy-efficient transfer of power to the load, a regulation and impedance-matching stage is usually necessary. For example, the load may need high voltage and low current while the energy harvester inherently provides low voltage and high current. Thus, a dc-to-dc up-converter from the output of the energy-storage element to the load may be necessary. In the RF antenna-based energy-harvesting system that we discussed in Chapter 17, the energy harvester is an antenna, and the energy-storage and impedance-matching functions are combined in the charge pump and in the capacitors of the pump. In general, to reduce the variability in available energy and to gather more energy, several energy sources can be simultaneously harvested and stored.

Before we begin with our discussion of inertial-motion mechanical-energy harvesting, we shall digress briefly to explain how electrical circuit models of mechanical systems are constructed.

### 26.2 Electrical circuit models of mechanical systems

The electrical equivalents of Newton’s three laws of motion in mechanical systems are as follows:

1. Newton’s first law:
   
   *Every body continues in a state of rest or in its state of motion unless it is acted on by a force.*
Electrical equivalent:

*Every capacitor holds its charge unless it is charged or discharged by an electrical current.*

2. Newton’s second law:

\[ F = \frac{mdv}{dt} \]  

(26.1)

*F* is the force, *m* is the mass, and *v* is the velocity of the moving or stationary mass.

Electrical equivalent:

\[ I = \frac{CdV}{dt} \]  

(26.2)

*I* is the current, *C* is the capacitance, and *V* is the voltage on the capacitor.

3. Newton’s third law:

*For every action, there is an equal and opposite reaction.*

Electrical equivalent:

*In any two-terminal electrical element, whether active or passive, dependent or independent, linear or nonlinear, the current flowing into one terminal on the element is equal to the current flowing out of the other terminal of the element.*

In the formulation above, current is analogous to a force, capacitance is analogous to a mass, and voltage is analogous to a velocity. The electrical equivalent of Newton’s third law is such that it is automatically satisfied and represented in any circuit. Mutual interactions between two bodies are represented as a floating current between two nodes such that one of the currents through the two-terminal element creates a sink current on the node that it is attached to while its paired current creates a source current on the node that it is attached to. Thus, Newton’s third law is nothing more or less than stating that a floating current source between two nodes may always be represented as a grounded sink current at one node and a grounded source current at the other node. The automatic and natural representation of Newton’s third law by a circuit makes electrical representations of mechanical systems powerful because one is relieved from the burden of having to constantly keep track of symmetric pushing and pulling between bodies. Furthermore, force balancing is also automatic. Since the voltage on a capacitor stops changing when all the currents flowing towards (or away from) it sum to zero, Kirchhoff’s current law is the law of force balance. Vector forces require 3D electrical circuits because the electrical analogies of mechanical systems hold separately for each of the *x*, *y*, and *z* components of force and velocity. For example, Figure 17.1 shows how circuit descriptions of Maxwell’s equations conceptually represent vectors.

In the formulation above, capacitance is a mass. If

\[ F = k \int vdt \]  

\[ I = \frac{1}{L} \int Vdt, \]  

(26.3)
then the reciprocal of an inductance represents a spring stiffness, or equivalently inductance represents a compliance. Mechanical damping is represented by a conductance:

\[ F = \eta \nu \]
\[ I = GV \]

(26.4)

Thus, a resonant mechanical mass-spring-damper system acted on by a force is represented by a parallel LCR resonator sourced by a current.

Frequently, a dual version of Equations (26.1), (26.2), (26.3), and (26.4) is used to represent mechanical systems by an electrical equivalent: force is represented by a voltage, velocity is represented by a current, mass is represented by an inductance, damping is represented by a resistance, and compliance is represented by a capacitance. In this analogy, a resonant mechanical-spring-damper system acted on by a force is represented by a series LCR resonator sourced by a voltage. Both forms are mathematically equivalent. However, one form is often more intuitive than the other and one should always work with a form that is the most intuitive. For example, in purely mechanical systems composed of interacting solids, if the equivalence described by Equations (26.1), (26.2), (26.3), and (26.4) is used, a parallel mechanical geometry maps to a parallel electrical topology, and a series mechanical geometry maps to a series electrical topology; the dual analogy flips parallel mechanical geometries to series electrical topologies and vice versa and is less intuitive. In contrast, in mechanical systems involving fluids, if we represent pressure by voltage and volume velocity by current, parallel fluid geometries map to parallel electrical circuits and series fluid geometries map to series electrical circuits; thus, the dual analogy is more intuitive for fluids. In piezoelectric electromechanical devices, forces cause charge displacements and voltages cause mechanical displacements. Thus, for reasons of symmetry, in piezoelectric devices, it is more natural to represent force by a voltage and velocity by a current.

### 26.3 Energy harvesting of body motion

Mechanical energy harvesting has been performed with three kinds of devices, namely, electromagnetic, electrostatic, and piezoelectric. An electromagnetic device converts flux changes induced by mechanical motion into an electrical voltage as in hydroelectric generators. If the voltage across a sensing capacitance is fixed, an electrostatic device, e.g., like the MEMS capacitance discussed in Chapter 8, converts capacitance changes due to mechanical displacements into charge changes. Electrostatic devices also convert capacitance changes into voltage changes if the charge on the sensing capacitance is fixed. A force imposed on a piezoelectric device causes mechanical deformation and charge changes within it. The charge changes manifest as a voltage across the piezoelectric device’s electrical capacitance. An exhaustive review of energy harvesting with all three kinds of
devices may be found in [7]. Work in [8] has shown that models for electromechanical energy harvesters are mathematically identical across all three classes of devices. Therefore, for reasons of brevity, we shall focus primarily on piezoelectric energy harvesters.

In all such passive devices, the presence of mechanical-to-electrical transduction implies that there is also correspondingly electrical-to-mechanical transduction in the reverse direction. The presence of transduction in both directions, each of which affects the other, leads to a feedback loop in the device. For example, electrical generators or electromagnetic energy harvesters don’t just convert mechanical motion to an electric voltage. Their operation causes a ‘back torque’ in addition to the mechanical torque driving the generator because the electric voltage that is generated also causes the generator to behave like a motor. In electrostatic devices, mechanical displacements lead to voltage or charge changes and also changes in the attractive electrostatic force between the capacitor plates. Piezoelectric devices are no exception. Applied force induces charge motions within the piezoelectric, which manifest as a voltage across their electrical capacitance. Applied voltage induces mechanical displacement changes, which manifest as a ‘back force’ across their mechanical compliance. Figure 26.2 (a) reveals a two-port electromechanical circuit that represents the functioning of a piezoelectric device. Figure 26.2 (b) reveals the feedback loop that represents this two port.

In Figure 26.2 (a), for convenience, we operate with current, which is the derivative of charge, such that all device characterizations can be done in terms

![Figure 26.2. A circuit description of a piezoelectret in (a) and a feedback block diagram that represents this circuit in (b).](image-url)
of voltage and current rather than in terms of voltage and charge. The mechanical compliance of the piezoelectret is represented by $C_m$. The electrical capacitance is represented by $C_e$. The mechanical force is represented by the voltage $F_{rcm}(s)$ and the electrical voltage by $V_e(s)$. The piezoelectric forward-and-back coefficients from the mechanical to the electrical domain and vice versa are typically symmetric and represented by $d$. The inertial input to the piezoelectric device is represented by a Norton velocity source of value $Vel_m(s)$ in parallel with an output admittance $G_{in}(s)$. The output admittance is often due to an inertial mass $M$ with impedance $M_s$, i.e., $G_{in}(s) = 1/(M_s)$. The output voltage of the piezoelectric device drives an electrical load with effective admittance $G_{eff}(s)$. Some algebra maps the two-port of Figure 26.2 (a) to the feedback loop of Figure 26.2 (b). For maximum efficiency, or maximum power transfer, we should configure $G_{in}(s)$ such that it is resonant with $C_m s$, and arrange $G_{eff}(s)$ such that it is resonant with $C_e s$. The alert reader will then immediately notice that Figure 26.2 (a) is exactly the dual circuit of the resonant mutual-impedance link that we described in Chapter 16. The similarity of Figure 26.2 (a) to its dual circuit in Figure 16.2 and the similarity of Figure 26.2 (b) to the feedback loop in Figure 16.3 are striking: in mapping Figure 16.2 to its dual version in Figure 26.2 (a), we simply exchange voltage for current, impedance for conductance, and a series two-port circuit for a parallel two-port circuit. The piezoelectric coefficient $d$ is analogous to the mutual inductance $M$, $C_m$ is analogous to $L_1$, $C_e$ is analogous to $L_2$, and we can define a coupling coefficient $k$ given by

$$k^2 = \frac{d^2}{C_m C_e} \quad (26.5)$$

identical to that defined in other treatments [9]. Therefore, we can exploit the analysis discussed in Chapter 16 to analyze piezoelectrets since the mathematics is virtually identical. We can define a reflected admittance $G_{reff}(s)$ analogous to the reflected impedance of Chapter 16 that is given by

$$G_{reff}(s) = -\frac{d^2 s^2}{C_e s + G_{eff}(s)} \quad (26.6)$$

This admittance is reflected from the electrical domain to the mechanical domain and appears in parallel with $G_{in}(s)$. From Chapter 16, for maximal energy efficiency, the resonance in the electrical and mechanical domain must both occur at the same optimal $\omega = \omega_m$. For maximal energy efficiency in the ‘primary’ mechanical domain, the reflected conductance must be much greater than the conductive portion of $G_{in}(s)$. For maximal energy efficiency in the ‘secondary’ electrical domain, the conductive portion of $G_{reff}(s)$ must be much greater than an effective parasitic conductance, $G_e$, that is in parallel with $C_e$, and which represents electrical losses. Note that $G_e$ is not shown in Figure 26.2 (a) since it is a parasitic. To determine the maximum overall energy efficiency, we can define an effective quality factor $Q_1$ for the resonator in the mechanical domain and an unloaded quality factor $Q_2$ for the resonator in the electrical domain.
Then, Equations (16.35) and (16.36) in Chapter 16 apply exactly to piezoelectric energy harvesters and describe their energy efficiency. From [8], since the mathematics of through-and-across variables is very similar for electromagnetic and electrostatic energy harvesters, we can analyze other motion-energy harvesters through the equations of Chapter 16 as well. Through variables are analogous to generalized current variables while across variables are analogous to generalized voltage variables. Later in this chapter, we shall discuss the operation of electric motors, which are electromagnetic energy generators (harvesters) that operate in reverse. This discussion will further illustrate the similarity between different kinds of electromechanical devices.

In Chapter 16, since we had a required load power consumption in the secondary and we wanted to ensure that the reflected power consumption in the primary due to this load was minimal, we focused on optimizing energy efficiency. In many energy-harvesting situations, energy efficiency may not be as important as maximizing energy transfer, i.e., getting as much absolute energy out of the harvester as possible, even if it means that a large fraction of energy is wasted. For example, in a resistive-divider circuit composed of a voltage source with a source impedance $R_S$ driving a load impedance $R_L$, energy efficiency is maximized when $R_L \gg R_S$; maximum energy is transferred from the source to the load when $R_L = R_S$, where the energy efficiency is only 50%. In this case, in the terminology of Chapter 16, it can be shown that for maximal power transfer

$$Q_{opt}^L = \frac{Q_2}{1 + k^2 Q_1 Q_2}$$

The power dissipated in the electrical load $P_e$ at this optimal value is related to the power dissipated at the mechanical input with no reflected load ($d$ or $k = 0$), $P_m$, according to

$$\frac{P_e}{P_m} = \frac{1}{4} \left( \frac{k^2 Q_1 Q_2}{k^2 Q_1 Q_2 + 1} \right)$$

Piezoelectric harvesters for wireless sensor networks are described in [10] and have generated 180–335 $\mu$W in 1 cm$^3$ of volume. They can be adapted for use in the noninvasive medical-monitoring systems described in Chapter 20. A piezoelectric energy harvester that scavenges energy from compression of the shoe sole has been able to generate 0.8 W of electrical power [11], [12]. Attempts to generate large amounts of electrical power from body motions, however, create a significant reflected electrical load on the mechanical side such that the metabolic effort needed to generate electrical power is consciously felt by the user. Since only 25% of the chemical oxidative energy of glucose is output as useful mechanical work by the body, even a highly efficient energy harvester at 31% can lead to a metabolic load to the body that is 12 times greater than the energy being harvested. One innovative effort to reduce such metabolic loading on the body uses an electromagnetic energy harvester placed on a knee brace slightly above the knee that harvests energy only during leg decelerations. It helps the leg to
decelerate by serving as an effective energy-harvester ‘brake’ to slow the leg at the end of a leg-extension movement. The metabolic load on the muscle is then reduced on average compared with the condition when an energy-harvesting brake is not present [13]. In the proof-of-concept design, the weight of the device, however, increased the mean metabolic load of walking by 20%. The device successfully generated 5 W of power. An electrostatic generator meant to harness ventricular wall motions of the heart generated 36 $\mu$W with simulated heart motions, sufficient to power a cardiac pacemaker. However, it was too big to implant and test directly on the heart [14]. In general, devices less than 1 cm$^3$ in volume are unlikely to generate more than 1 mW of power from body motions [7], but for many low-power applications like we have discussed in Chapter 20, 100 $\mu$W is more than adequate. One challenge in the field is that it is easy to make small devices that have high resonant frequencies but most of the power spectrum of motion energy is below 100 Hz. Furthermore, if the body motion is far in excess of the maximal motions possible in a small device, resonant amplification is not necessarily an advantage. Non-resonant conversion strategies are being investigated [15].

26.4 Energy harvesting of body heat

When heat flows from a hot body at temperature $T_{\text{high}}$ to a cold body at temperature $T_{\text{low}}$, some of the heat energy can be harnessed to perform useful work. For example, the internal combustion engine in a car burns gasoline fuel in a controlled fashion, which releases energy primarily as heat. A fraction, i.e. $\sim25\%$, of this heat energy is exploited to perform useful mechanical work. A large fraction, i.e., $\sim75\%$, of it is wasted as heat from the radiator to the surround. In the body, energy in glucose molecules is first converted to energy in many smaller energy-carrying molecules called adenosine tri-phosphate (ATP) at nearly 50\% efficiency within our cells. The ATP molecules serve as universal energy currency throughout the cell and power various activities in the cell that perform useful work. For example, ATP powers electricity generation across all cell membranes in all cells of the body and also powers the contractions of muscle cells. The efficiency of energy conversion from ATP to useful work is nearly 50\%. Thus, the overall efficiency from fuel to useful work in the body is also $50\% \times 50\% = 25\%$ as in a gasoline engine. Hence 75\% of the energy in the food that we eat is converted to heat energy. This heat energy is used to maintain the body at an internal 37 $^\circ$C temperature significantly higher than the external temperature, at say 22 $^\circ$C, and compensates for heat lost from the body to the environment.

With temperature analogous to voltage, and heat flow analogous to current, a circuit for thermoelectric generation is as shown in Figure 26.3 [16]. Each resistance in Figure 26.3 is described by an Ohm’s law equation of the form

$$\Delta T_x = I_{\text{heat}}R_x$$

(26.9)
where $\Delta T_x$ is the temperature drop across the resistance $R_x$, and is measured in units of $^\circ C$. The heat flow, $I_{heat}$, is measured in units of W/m$^2$. From Equation (26.9), the thermal resistance is then measured in units of $m^2K/W$. The $R_{thermoelec}$ element is built with a cascade of several BiTe Seebeck-effect thermopiles that each provide about 0.2 mV/$^\circ C$ of output voltage. Such thermopiles are built by bringing two dissimilar metals together at two junctions, one at the hot side and one at the cold side. A series stack of several of these thermopiles is necessary to develop voltages of $\sim$1 V. For example, the recent design described in [16] used $158 \times 8 \times 4 = 5056$ of these devices in series to develop nearly 0.7 V. What determines the values of $R_{body}$, $R_{air}$, and $R_{thermoelec}$, and thus the value of $I_{heat}$?

Since 0.75 of the body’s resting power dissipation of 81 W is dissipated over $\sim$2 m$^2$ surface area, the net average heat flow out of the body may be expected to be $\sim$30 W/m$^2$ under resting or sleeping conditions. Under normal conditions, where the power dissipation averages to 125 W–150 W, it has been measured to be $\sim$60 W/m$^2$ or $\sim$6 mW/cm$^2$ when the ambient temperature is 28 $^\circ C$ [17]. At steady state, the heat flow out of the body must be matched by the heat that it generates to ensure that the body does not heat up or cool down. Not surprisingly, the value of the body’s effective thermal resistance, $R_{body}$, is altered via blood-vessel dilation and constriction and other feedback mechanisms to ensure that the body’s temperature is maintained. There is variance in the heat flow at different positions in the body. For example, the relatively hot blood in the radial artery on the underside of the forearm, where watches are worn, is only separated from the ambient air by a $\sim$7 mm layer of skin without any heat-insulating muscle. Thus, the heat flow in this region of the body is around 100 W/m$^2$ or 10 mW/cm$^2$. The thermal resistance of the body, which is $\sim$500 cm$^2K/W$, is reduced in this region to $\sim$100 cm$^2K/W$.

The thermopile resistance, $R_{thermopile}$, is determined by the heat-conduction properties of the thermopile material, the cross-sectional area of the legs that join together to create its junctions, and its length. Larger cross-sectional areas and smaller lengths lead to lower resistances. The dependence of thermopile heat resistance on geometry is similar to that of electrical resistances except that the heat conductivity $\kappa$ plays the role of the electrical conductivity $\sigma$. In Figure 26.3, $R_{thermoelec}$ must have its geometry designed such that it is comparable to $R_{body} + R_{air}$. It is hard to make it significantly larger than this value without making devices too long or cross-sectional areas too thin, since the value of $R_{air}$ is typically quite high. For example, commercially available thermopiles have $R_{thermopile}$ at $\sim$50 cm$^2K/W while $R_{air}$ is $\sim$1000 cm$^2K/W$. The value of $R_{air}$ is

![Figure 26.3. A circuit model for body heat loss.](image-url)
determined by radiative and convective losses in air. The presence of wind lowers $R_{\text{air}}$ by promoting heat exchange and increasing heat flow. A value of $R_{\text{thermopile}}$ that is significantly smaller than $R_{\text{body}} + R_{\text{air}}$ will not have much temperature dropped across it, and lead to a loss in sensitivity.

The device described in [16] achieved $250 \, \mu W$ of power extraction with a heat flow of 20 mW/cm$^2$ across a 6 cm$^2$ wristwatch-sized device with an ambient air temperature of 22°C. The output voltage into a matched load was 0.93 V. Given that we have 120 mW of heat flowing into our device, why are we only able to extract 250 $\mu W$? A fundamental limit known as the Carnot efficiency limits the amount of power that can be extracted from a thermoelectric device.

Figure 26.4 shows what is termed as a 'heat engine', i.e., a system that generates useful work $W$ as heat flows from a 'hot reservoir' at temperature $T_{\text{high}}$ to a 'cold reservoir' at temperature $T_{\text{low}}$. An amount of heat, $\Delta Q_{\text{high}}$, flows out of the hot reservoir, some of it generates useful work $W$, and the rest, $\Delta Q_{\text{low}}$, flows into the cold reservoir. By energy conservation, it is necessary that

$$\Delta Q_{\text{high}} = \Delta Q_{\text{low}} + W$$ (26.10)

The reservoirs are assumed to have so many degrees of freedom in which to absorb heat energy, i.e., a high heat capacity, such that their temperature barely changes with the modest amount of heat drawn out of or poured into them. But what determines how much of $\Delta Q_{\text{high}}$ ends up as useful work $W$, rather than wasted heat $\Delta Q_{\text{low}}$?

The fundamental second law of thermodynamics in physics states that the amount of disorder in the world, measured by 'entropy', can only have a net increase or remain the same. It is based on the fact that disordered and highly random system states where energy is distributed equally amongst many degrees of freedom are statistically significantly more likely than ordered system states where energy is concentrated amongst a few degrees of freedom. In fact,
temperature is just a measure of the average random thermal energy per degree of freedom in a system. A degree of freedom represents the voltage on a capacitor, the current in an inductor, the position of a spring, or the velocity of a mass in a physical system. (Degrees of freedom are discussed in Chapter 7 in the context of the equipartition theorem.) Heat flows from a hot high-temperature body to a cold low-temperature body because the random thermal energy tries to redistribute itself equally amongst all degrees of freedom in both the hot body and the cold body. Since the hot body has more average thermal energy per degree of freedom than the cold body, there is a net thermal energy flow from the hot body to the cold body as the energy redistribution occurs. When a small amount of heat $\Delta Q_{\text{low}}$ flows into a heat reservoir of temperature $T_{\text{low}}$, its entropy is defined to increase by $\Delta Q_{\text{low}}/T_{\text{low}}$ since the number of accessible states in the reservoir increases as more energy is poured into it, leading to more uncertainty about its state, and more disorder. When a small amount of heat $\Delta Q_{\text{high}}$ flows out of a heat reservoir of temperature $T_{\text{high}}$, its entropy is defined to decrease by $\Delta Q_{\text{high}}/T_{\text{high}}$ since the number of accessible states in the reservoir decreases as energy leaves it, leading to less uncertainty about its state, and less disorder. Since the net entropy change must be nonzero by the second law

$$-\frac{\Delta Q_{\text{high}}}{T_{\text{high}}} + \frac{\Delta Q_{\text{low}}}{T_{\text{low}}} \geq 0$$  \hspace{1cm} (26.11)

Some algebra on Equations (26.10) and (26.11) then reveals that

$$W \leq \Delta Q_{\text{high}} \left( 1 - \frac{T_{\text{low}}}{T_{\text{high}}} \right)$$  \hspace{1cm} (26.12)

Thus, the maximum efficiency of the heat engine, that is the fraction of heat $\Delta Q_{\text{high}}$ that is converted to useful work $W$, is limited to a maximum value known as the Carnot efficiency, $C$,

$$C = 1 - \frac{T_{\text{low}}}{T_{\text{high}}}$$  \hspace{1cm} (26.13)

The Carnot efficiency sets limits on the efficiencies of steam engines, plane engines, car engines, and on our thermoelectric harvester as well. If the body is at $T_{\text{high}} = 37 \, ^\circ \text{C}$, and the ambient temperature is at $T_{\text{low}} = 22 \, ^\circ \text{C}$, the maximum possible efficiency for the thermoelectric harvester is given by

$$C_{\text{thrmhrv}} = 1 - \frac{273 + 22}{273 + 37}$$  \hspace{1cm} (26.14)

$$= 0.0484$$

Thus, if 120 mW flows into a thermoelectric energy harvester, the best we can hope to do is extract 5.8 mW of power. It is not atypical for an experimental system to operate at 10% of the limiting possible Carnot efficiency, which is only achievable at infinitely slow operation. The system described in [16] achieves nearly 4% of the Carnot limit but it is one of the best systems reported thus
far. Its delivered power density of \( \sim 0.2 \text{ W/m}^2 \) is in excess of what an average 10\%-efficient solar cell might deliver in indoor environments. An efficient charge pump for such thermal harvesters is described in [18].

### 26.5 Power consumption of the world

Mackay estimates the average power consumption of an affluent British citizen today in his book [1]. If we adapt his units of 1 kW h/day to simple kW units with the conversion factor 1 kW h/day = 41.67 W, we find that this consumption may be broken down as shown in Table 26.1.

The costs of Table 26.1 are estimated for an affluent British citizen. An average British citizen actually consumes 5.2 kW, an average European citizen consumes 5.46 kW, while an average American citizen consumes nearly 10.4 kW. The world average is 2.5 kW with great variance across nations. Since there are

<table>
<thead>
<tr>
<th>Item</th>
<th>Power consumption</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Car usage</td>
<td>1.67 kW</td>
<td>30 mph at 30 mpg for ( \sim 1 ) hour at 10 kW h/liter for gas with 3.8 liters = 1 gallon. Or equivalently, the cost of an average 42 kW car driven for ( \sim 1 ) hour each day.</td>
</tr>
<tr>
<td>2. One transatlantic flight per year on a Boeing 747</td>
<td>1.25 kW</td>
<td>Such planes operate at 0.14 mpg but amortize this cost over ( \sim 400 ) passengers such that they effectively operate at ( \sim 60 ) mpg per person. The power consumption of a Boeing 747 is ( \sim 150 ) MW.</td>
</tr>
<tr>
<td>3. Heating</td>
<td>1.540 kW</td>
<td>Not important in some geographical areas.</td>
</tr>
<tr>
<td>4. Material synthesis energy costs</td>
<td>2.08 kW</td>
<td>It costs energy to manufacture appliances.</td>
</tr>
<tr>
<td>5. Electric lighting</td>
<td>0.167 kW</td>
<td>Estimated for an average home.</td>
</tr>
<tr>
<td>6. Electric gadgets</td>
<td>0.208 kW</td>
<td>Washers, dryers, cell phones, etc.</td>
</tr>
<tr>
<td>7. Material transport</td>
<td>0.500 kW</td>
<td>Trucking and transportation costs to move materials.</td>
</tr>
<tr>
<td>8. Food</td>
<td>0.625 kW</td>
<td>This energy cost in food only tracks industrial energy flows associated with food, not the natural embedded energy in food. For example, it costs energy to transport food, and to maintain animals to be used later as food.</td>
</tr>
<tr>
<td>9. Defense</td>
<td>0.167 kW</td>
<td>These national costs are amortized per person.</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>8.207 kW</strong></td>
<td>Does not include the cost of imported goods, which bear their own energy costs, at 1.667 kW.</td>
</tr>
</tbody>
</table>
approximately 6 billion people on our planet today, the power consumption of the world is 15 TW. The electricity consumption of the world is 2 TW. However, since typical generating stations burn fossil fuels like coal to generate electricity and are only 40% efficient, the actual power consumption due to electricity use is 5 TW. We notice that a large fraction of the power consumption of the world revolves around transportation, heating, and electricity costs. This book has already discussed principles for lowering power in electrical systems. Now, we shall discuss some principles for the design of low-power transportation systems of the future.

### 26.6 A circuit model for car power consumption

Figure 26.5 shows a circuit model of a car that is useful for understanding factors that affect its power consumption. We shall use current to represent force and voltage to represent velocity in accord with Equations (26.1), (26.2), (26.3), and (26.4). Thus, mass is represented by a capacitance, mechanical damping by a conductance, and mechanical compliance by an inductance. A chemomechanical dependent force $i_{ENG}$ due to the burning of fuel along with a Norton-equivalent mechanical admittance $G_{ENG}$ represents the characteristics of the engine power source. The fuel-to-mechanical work efficiency is typically 25% such that the $i_{ENG}v_{ENG}$ power output by the engine requires $4i_{ENG}v_{ENG}$ power to be extracted from the chemical fuel. The motions of the engine are periodic. The engine force is conveyed via gears to provide force to the car wheels. The transformer in Figure 26.5 represents a lossless gearbox (and transmission) that performs an impedance transformation. The reflected admittance of the secondary wheel-and-road side to the primary engine side must be such that most of the power output by the engine is dissipated in the reflected admittance, not in $G_{ENG}$, which is usually the case. As the characteristics of the impedance in the secondary change with flat, uphill, or downhill road conditions, the gear ratios are changed such that this efficiency is preserved.

![Figure 26.5](image-url)

**Figure 26.5.** Equivalent circuit of a car showing losses due to air drag, rolling friction, braking, and chemical-to-mechanical energy conversion.
The mechanical current $i_{CAR}$ charges the capacitance $M_{CAR}$ to a voltage $v_{CAR}$, which represents the car’s velocity. When $M_{CAR}$ is charged, the car accelerates in accord with Newton’s second law. The car has three force currents $i_{DRAG}$, $i_{ROLL}$, and $i_{BRK}$ respectively that oppose $i_{CAR}$ and attempt to decelerate the car. When no brake is applied, $i_{BRK}$ is 0. The brake current is shown as a switched resistance in Figure 26.5. When the car is moving at a steady velocity and no brake is applied, $i_{CAR}$ must balance $i_{DRAG}$ and $i_{ROLL}$ such that there is no charging or discharging current on the capacitor and the velocity of the car is maintained at a constant value.

The drag force $i_{DRAG}$ is due to viscous air resistance caused by the fluid moving past the car. It can semi-empirically be represented by the current through a quadratic conductance [19]:

$$i_{DRAG} = \frac{1}{2} \rho C_D A_{CAR} v_{CAR}^2$$

(26.15)

The parameter $A_{CAR}$ is the effective cross-sectional area of the car, which must be kept small to reduce air drag. That is why most natural creatures and artificial transport mechanisms that move efficiently, e.g., trains, are architectured to be long and thin such that $A_{CAR}$ is small within a given volume constraint. We can estimate $A_{CAR}$ as 3 m$^2$ for a 1.5 m high and 2 m wide car. The parameter $C_D$ is called the coefficient of drag and is typically near 0.3 in most cars and lower in streamlined racing cars. The parameter $\rho$ is the density of air, which is 1.3 kg/m$^3$. The mass of the car, $M_{CAR}$, is typically one tonne, i.e., 1000 kg.

The force $i_{ROLL}$ is the force due to ‘rolling friction’ in the car. It is due to the fact that the tires slightly distort and recover shape when they move, and some of the energy in the tires and tire bearings is dissipated. The force $i_{ROLL}$ is semi-empirically represented by [19], [20]

$$i_{ROLL} = C_r M_{CAR} g$$

(26.16)

with $C_r = 0.013$ for a relatively smooth road and average car tires. The parameter $C_r$ is relatively invariant with speed.

The braking force $i_{BRK}$ dissipates the car’s kinetic energy as heat when the brake is applied and the braking resistance is switched to ‘ground’, i.e., to zero velocity as shown in Figure 26.5. Just as the discharge current in digital CMOS design dissipates the $(1/2)CV^2$ capacitive energy stored in a voltage node as heat, the braking force dissipates the $(1/2)M_{CAR}v_{CAR}^2$ kinetic energy stored in the mass of the car as heat. As in adiabatic CMOS design, discussed in Chapter 21, the switched braking energy can be partially recovered and stored on an ultra-capacitor, flywheel, or battery and then used to provide energy back to the car when it is time to accelerate. Hybrid cars recover 50% of the switching energy through such ‘regenerative braking’ strategies. Equation (21.50) in Chapter 21 showed that, in a high-quality-factor system, the switching energy can be reduced by as much as $2\pi/Q$ where $Q$ is the quality factor of the system. In Figure 26.5, we have represented the regenerative storage abstractly by a mass $M_{RGN}$ although any form of storage may be used.
Lots of braking corresponds to a high activity factor or lots of switching in electronic systems. A high maximum velocity of the car corresponds to a large $V_{DD}$ in electronic systems. The $i_{\text{DRAG}}$ and $i_{\text{ROLL}}$ forces correspond to static ‘leakage’ currents in electronic systems. One leak current increases quadratically with voltage (the drag current) and one leak current is a constant voltage-independent current like subthreshold leakage current in digital systems (the roll current). We can therefore draw upon principles learned in low-power digital design to reduce static energy and dynamic switching energy for a given distance of transport. But what is a good metric for energy-efficient transportation?

As we discussed in Chapter 21, the energy per cycle of operation, $E_{TOT} = P_{TOT} (1/ \nu_{av})$, is used to characterize the energy efficiency of digital systems. Transport is rarely periodic such that ‘a cycle of operation’ does not make sense in transportation systems. However, the analogy to speed of operation, $f_{clk}$, is the average velocity of travel, $v_{av}$. Thus, the average power consumption divided by the average velocity of travel might be a good metric. If the total time of travel is denoted by $t_{trv}$,

$$I_{merc} = \frac{P_{av}}{v_{av}} = \frac{P_{av} t_{trv}}{v_{av} t_{trv}} = \frac{E_{TOT}}{d_{TOT}}$$

Thus, our metric inspired from electronics is mathematically equivalent to the metric that is actually used to characterize energy efficiency by transportation engineers: the energy consumed, $E_{TOT}$, divided by the total distance traveled, $d_{TOT}$. It is pleasing that metrics used in completely different fields are intuitively similar.

Suppose the average distance between braking stops is $d_{stp}$ and that we can recover a fraction $\alpha_{\text{regen}}$ of the kinetic energy wasted during braking. Also suppose that the time spent to accelerate to cruising velocity $V_{MX}$ or to brake to 0 is negligible compared with the time spent traveling at this cruising velocity, i.e., we have a square-wave-like profile in our velocity. Then, from Equations (26.15), (26.16), (26.17), and Figure 26.5, the metric for transport energy efficiency is given by

$$I_{merc} = \left( \frac{(1 - \alpha_{\text{regen}}) M_{\text{CAR}} V_{MX}^2}{2 d_{stp}} + \frac{1}{2} \rho C_D A_{\text{CAR}} V_{MX}^2 + C_r M_{\text{CAR}} g \right)$$

In Equation (26.18), $I_{merc}$ is a current, which implies that we are energy efficient for a given distance of travel if we can minimize the average force or ‘thrust’ required during this travel. Equation (26.18) also reveals that, to be energy efficient, $d_{stp}$ and $\alpha_{\text{regen}}$ should be high such that we do not brake often and that we recover most of the energy when we do, respectively; that $M_{\text{CAR}}$ should be low to minimize rolling friction and energy dissipated during braking; that the frontal area of the car, $A_{\text{CAR}}$, and the drag coefficient, $C_D$, should be minimized by having the car created in a tear drop shape; and that $V_{MX}$ should be low to minimize drag and braking forces. If the car were an electrical system, we would state that, for low-power operation, the activity factor should be small and recycling efficiency should be near 1 such that switching energy is minimized; that capacitances should
be small; that leakage currents should be optimized to be small; and that the power-supply voltage, $V_{DD}$, of the system should be small.

Equation (26.18) reflects the energy efficiency of the secondary wheel side. If we reflect this back to the primary side, the actual energy efficiency of transport, which is related to the fuel consumed, is given by

$$I_{act} = \frac{4}{\epsilon(V_{MX})} \left( \frac{(1 - \eta_{eng}) M_{CAR} V_{MX}^2}{2 d_{sp}} + \frac{1}{2} \rho C_D A_{CAR} V_{MX}^2 + C_{rr} M_{CAR} g \right), \quad (26.19)$$

where the factor of 4 arises from the 25% fuel-to-mechanical-work efficiency, and the $\epsilon(V_{MX})$ function represents the fact that the efficiency of the engine changes with speed $V_{MX}$. It is usually maximum at an optimal value of $V_{MX}$ due to the characteristics of $G_{ENG}$ and the nature of the fuel-to-mechanical-energy transfer.

The power consumption of a car on the secondary side is just $V_{MX}$ times $I_{act}$, since $I_{act}$ represents the force thrust of the car. Reflecting this power consumption back to the secondary side, and assuming that $\epsilon(V_{MX})$ is 1, for simplicity, we can compute the power consumption of the car to be

$$P_{CAR} = 4 \left( \frac{(1 - \eta_{eng}) M_{CAR} V_{MX}^3}{2 d_{sp}} + \frac{1}{2} \rho C_D A_{CAR} V_{MX}^3 + C_{rr} M_{CAR} g V_{MX} \right) \quad (26.20)$$

Figure 26.6 plots the power consumption, $P_{CAR}$, versus its velocity $V_{MX}$ for $\eta_{eng} = 0$, $M_{CAR} = 1000$ kg, $\rho = 1.3$ kg/m$^3$, $C_D = 0.3$, $A_{CAR} = 3$ m$^2$, $C_{rr} = 0.013$, $d_{sp} = 150$, and $d_{sp} = 5000$. 

![Figure 26.6. Power consumption of a car versus its speed.](image)
and \( g = 9.8 \text{ m/s}^2 \) for two different values of \( d_{sp} = 150 \text{ m} \) (city driving) and \( d_{sp} = 5000 \text{ m} \) (highway driving). The transition of a car from a linear \( V_{MX} \) power consumption to a cubic \( V_{MX}^3 \) power consumption occurs when drag and/or braking forces eventually overwhelm the forces of rolling friction. While Equation (26.20) is only an approximation since \( e(V_{MX}) \) has simply been set to 1, and it ignores several other details of a complex system like a car, it does approximate actual car power consumptions, especially if \( V_{MX} \) is not too small. For example, the power consumption of the car at 30 mph with the parameters that we have chosen is nearly 40 kW. At lower speeds, the higher curve with a smaller \( d_{sp} \) more accurately reflects car power consumption. At higher speeds, the lower curve with a larger \( d_{sp} \) more accurately reflects actual car power consumption. A real car’s power consumption will transition between curves like those in Figure 26.6 as \( d_{sp} \) increases with speed.

The transport energy efficiency of a gasoline car given by Equation (26.19) for the parameters used in Figure 26.6 at \( d_{sp} = 150 \text{ m} \) at 30 mph is nearly 3000 \( \text{N} \). The ‘N’ stands for newtons, a unit of force. The car consumes 3000 \( \text{J} \) of energy to go 1 m or, equivalently, the car burns \( \sim 40 \text{ kW} \) of power to go 13 m/s. The transport energy efficiency of a car is usually optimal at a moderate speed where \( e(V_{MX}) \) is high in Equation (26.19) and the \( V_{MX} \) drag and braking terms are not too large. The transport energy efficiency that we have computed is indeed near to that observed in real cars.

Part of the staggering transport inefficiency of a car arises from the fact that it has to transport its own weight (1000 kg), which is significantly in excess of the weight of its cargo (~65 kg for one person). The presence of just four persons in the car improves its transport efficiency per person by a factor of almost four. Trains take this idea to an extreme and also incorporate other ideas making them highly energy efficient. Trains are significantly more energy efficient per person than cars for six reasons. First, since they rarely stop and brake and have a dedicated track to run on, the braking term in Equation (26.18) is nearly 0, except when the train stops at the end of its journey, so we can set it to 0. Second, they employ a ‘collective’ or ‘parallel’ low-power principle (see Chapters 21 and 22) to amortize transportation costs per person to a small value: if \( N \) persons occupy a train, the effective transport energy efficiency per person is given from Equation (26.18) by

\[
\frac{\text{PE}_{\text{prsn}}} {\text{mtrc}} = \left( \frac{1}{2} \rho C_D \frac{A_{TRN}}{N} V_{MX}^2 + \frac{C_{rr} (M_{TRN} + NM_{prsn})}{N} \right) g
\]

(26.21)

Hence, the fuller a train is, the lower its drag and rolling-friction terms per person, and the more efficient it becomes. Third, their long and lean design with relatively low frontal area reduces drag. Fourth, the coefficient of rolling friction, \( C_{rr} \), for steel on steel is 0.002 rather than 0.013 for a car. Fifth, they can run at an optimal speed where their engine has optimal efficiency. Sixth, they can be run on electricity and
thus have engine efficiencies of over 80%. The overall result is that trains can operate at high speeds with a transport energy efficiency of 58 N per person, about 50 times less than that of a car. In practice, trains are rarely full in developed countries. Nevertheless, in Japan, rail transport operates at 216 N-per-person efficiency, i.e., it has a 14 times better transport efficiency than that of an average car today.

### 26.7 Electric cars versus gasoline cars

Given the high energy efficiency of electric engines in trains, can we build purely electric cars that are more efficient than gasoline cars? Purely electric cars function by using a battery to power a motor, which, after gearing, provides torque to turn the wheels of the car. The battery takes the place of the fuel as the energy source, and the motor takes the place of the car engine. The primary side of Figure 26.5 in a purely electrical car is different from that in a gasoline car. The secondary side of Figure 26.5 is identical in a purely electrical car to that in a gasoline car. From now on, for brevity, we shall refer to purely electrical cars as electric cars. Before we compare the efficiency of electric cars versus gasoline cars, it is useful to understand how a motor works.

Figure 26.7 (a) reveals the equivalent electromechanical circuit of a motor and Figure 26.7 (b) reveals the feedback loop that describes Figure 26.7 (a). In Figure 26.7 (a), $V_m(s)$ represents the input voltage source to the motor with source impedance $Z_{in}(s)$. The motor has an electrical impedance $Z_e(s)$ and a ‘back emf’, $K\Omega_m(s)$, which is proportional to its angular velocity $\Omega_m(s)$. The back emf arises, because, just as in a piezoelectret, mechanical-to-electrical transduction occurs simultaneously with electrical-to-mechanical transduction. Equivalently, all motors generate a back emf because they are also electrical generators. The net current in the motor, $I_m(s)$, generates a torque, $\Gamma(s)$, which drives the mechanical admittance of the motor, $G_m(s)$, and the effective mechanical admittance of the load, $G_{eff}(s)$, to create the motor’s angular velocity, $\Omega_m(s)$. The admittance of the motor, $G_m(s)$, is typically inertial/capacitive with a little loss. To find $G_{eff}(s)$, the effective load on the motor, we need to model an electric car. We can use the model shown in Figure 26.5 to model an electric car. The load $G_{eff}(s)$ is the reflected load of the secondary side of Figure 26.5, which appears, after gearing, as a load to the motor on the primary side. If the torque and angular velocity of the motor are represented by $\tau$ and $\omega$, respectively, and if the transformer, which represents the gears, is lossless,

$$\tau_m\omega_m = i_{CAR}v_{CAR} \tag{26.22}$$

As an aside, note that, if $R$ is the wheel radius on the secondary side of Figure 26.5, we can also choose to parametrize the mechanical output variables of the car by the torque on the wheel and the angular velocity of the wheel, i.e.,

$$\tau_{CAR} = i_{CAR}R \tag{26.23}$$

$$\omega_{CAR} = v_{CAR}/R$$
In Figure 26.7 (a), $K \Omega_m(s)$ is analogous to $I_{ENG}(s)$ in Figure 26.5 and similarly $G_m(s)$ is analogous to $G_{ENG}(s)$. Thus, the overall circuit of an electric car is formed by having the circuit of Figure 26.7 (a) replace the primary side of Figure 26.5 and with $G_{eff}(s)$ replaced by the gears and secondary side of Figure 26.5. Hence, there are two 2-port circuits in an electric car, which are cascaded with one another: The first occurs due to the electromechanical motor two-port circuit of Figure 26.7 (a) and the second occurs due to the simple gearing transformer of Figure 26.5.

The mechanical load of the motor is reflected to the electrical side in Figure 26.7 (a) as an equivalent electrical impedance of value

$$Z_{refl}^{me} = \frac{K^2}{G_{eff}(s) + G_m(s)}$$  \hspace{1cm} (26.24)

which appears in series with $Z_m(s)$ and $Z_e(s)$ in Figure 26.7 (a), i.e., the reflected impedance $Z_{refl}^{me}$ replaces the $K \Omega_m(s)$ dependent generator in Figure 26.7 (a). A physical interpretation of Equation (26.24) leads to the realization that admittance on the secondary side in Figure 26.5 simply transforms to a scaled identical electrical admittance $G_{reff}^{me}(s) = 1/Z_{reff}^{me}(s)$ that is an exact mimic of the mechanical admittance. If $g$ is the gear ratio, greater than 1, that represents the up-conversion of force from the primary to the secondary side in Figure 26.5, and $G_{CAR}(s) = I_{CAR}(s)/V_{CAR}(s)$ is the car admittance, then, from Equations (26.24) and Figures 26.5 and 26.7(a), we get

$$G_{reff}^{me} = \frac{G_m(s)}{K^2} + \frac{1}{g^2 K^2} G_{CAR}(s)$$  \hspace{1cm} (26.25)
Note that Equation (26.25) is only a small-signal frequency-domain characterization of the reflected impedance of the car. However, the feedback loop of Figure 26.7 (b), suggests that even if the $\Gamma \rightarrow \Omega$ or equivalently $i_{\text{CAR}} \rightarrow v_{\text{CAR}}$ transformation is nonlinear due to the nonlinear load characteristics of the car (quadratic drag conductance and fixed rolling-resistance current source in Figure 26.5), we can still represent this relationship as a nonlinear block in an equivalent time-domain feedback loop version of Figure 26.7 (b). We can then reflect this nonlinear block into the electrical domain of the motor by replacing the back emf by a nonlinear $I-V$ block that characterizes the car. That is, in Equation (26.25), we simply use the scaling constants $1/K^2$ and $1/(g^2 K^2)$ on the $I-V$ curves that characterize the motor admittance $g_m$ or the car admittance $r_{\text{CAR}}$ in the time domain respectively, and reflect these summed-and-scaled $I-V$ curves into the electrical domain as a nonlinear $I-V$ element. A small-signal frequency analysis about each large-signal operating point as in transistor circuits, however, is still useful for providing intuition.

The factors that affect the efficiency of electric cars are the same ones that affect the efficiency of mutual-inductance links that we discussed in Chapter 16. Most of the torque current of the motor in Figure 26.7 (a) should flow through $G_{\text{eff}}(s)$ rather than through $G_m(s)$ to preserve good efficiency in the mechanical output domain. The motor mass is significantly below the car’s mass and the motor damping losses are usually significantly less than the drag and other car losses. Thus, good efficiency in the mechanical domain can be achieved with modest gear ratios in Equation (26.25). Most of the voltage drop of the motor’s input voltage should be dropped across the reflected impedance $Z_{\text{refl}}(s)$ rather than across $Z_m(s)$ or $Z_e(s)$ to preserve good efficiency in the electrical input domain, that is $G_{\text{refl}}(s)$ should be sufficiently small. Thick wiring in the motor windings reduces electrical losses $R_e$, large batteries have low output source impedance, and typically the motor’s electrical inductance $L_e$ is such that the electrical time constant of the motor, $L_e/R_e$, is much less than the mechanical time constant caused by the reflected impedance. Hence, if $K$, which is primarily determined by the amount of flux in the motor, is sufficiently large, $G_{\text{refl}}(s)$ can be made small enough in Equation (26.25) such that efficiency in the electrical domain is excellent. The overall efficiency of the motor is the product of the efficiency in the electrical domain times the efficiency in the mechanical domain. Motors can operate with excellent energy efficiency. Indeed, the Tesla Roadster electric car has achieved efficiencies that average 92%, which are significantly higher than the 25% efficiency of a gasoline car engine [21].

The Tesla Roadster electric car is an impressive engineering feat since it achieves the specifications of a high-performance sports car with excellent transport energy efficiency and zero emissions. Its source of power is a 450 kg lithium-ion battery with 53 kW h capacity (424 J/g) capable of 200 kW output. The car itself weighs 1222 kg, has a peak mechanical power output of 189 kW, accelerates from 0 to 60 mph in 3.9 seconds, has a top speed of 125 mph, and can go 244 miles on a single battery charge. Its battery lifetime is limited to about ~ 500 charge-recharge
cycles (~100,000 miles), and it takes 3.5 hours for a full battery recharge although a full recharge may rarely be necessary. It implements regenerative braking. The lithium-ion battery has several short-circuit protection features including in-built fuses that disconnect it in situations of high temperature and pressure. The battery is architected to be safe even during collisions.

Most importantly, the Tesla Roadster’s transport energy efficiency is ~500 N, about 6 times better than that of an average car. The transport efficiencies of several lighter, lower-range, and low-speed electric cars are not significantly different from that of the Roadster and some are much worse [1]. To be fair to the average gasoline car, though, the Tesla Roadster uses high-grade electric energy, while the average car needs to extract its energy from fossil fuel. Most electricity generating fossil-fuel plants are 30%–40% efficient such that one could argue that the real improvement of a Roadster is a factor of 2× to 2.4×. Nevertheless, the Tesla Roadster does illustrate that direct conversion between high-grade forms of energy, e.g., electrical to mechanical rather than from chemical-to-heat-to-mechanical as in a gasoline car, is efficient. In the future, if electricity is generated in solar plants, such a car could indeed have a zero-emission footprint, especially if it is manufactured in plants using solar electricity as well. Even though the energy density of the lithium-ion battery that was used is 11 times less than that of gasoline, the weight of the car is manageable because the heavy gasoline engine is replaced with an electric motor.

26.8 Cars versus animals

Another impressive example in transportation engineering is the cheetah (*Acinonyx jubatus*), the fastest land animal. Its top sprint speed has been measured to be 30 m/s, i.e., 68 mph [22]. It can accelerate to 68 mph in 3 seconds, faster than a Tesla Roadster, which gets to 60 mph in 3.9 seconds, and faster than most high-performance cars [23], [24]. Since the average cheetah weighs nearly 50 kg, we can estimate that its mechanical power output during this acceleration is 
\[
\frac{1}{2} \cdot 50 \cdot (30^2) / 3 = 7.5 \text{ kW}.
\]
Its rudder-like tail enables it to make incredibly quick turns during its chase of a prey animal. The cheetah uses its spring-like backbone to partly store and regenerate energy in each stride and is airborne for more than half its stride. Its transportation efficiency for aerobic speeds, which can typically be maintained for long distances only if they are less than half the top speed [25], has been measured to be equivalent to 0.14 ml of oxygen consumption per g.km [26]. From the energetics of a glucose or carbohydrate reaction, and from the weight of an average cheetah, these numbers work out to an energy efficiency of 132 N. The cheetah’s transport energy efficiency is 4 times better than that of a highly energy-efficient electric car. What is more impressive is that this transport energy efficiency is achieved even though the cheetah has to make do with a 25% efficient engine (fuel-to-mechanical-work) and that it runs with legs, not as optimal as wheels on flat terrain. For example, the energy efficiency of humans walking at
2.5 mph can be calculated from measurements in [27] to be close to that of a cheetah, 130 N, but humans riding average unoptimized bicycles with wheels can achieve 81 N even though the drag coefficient of such bicycles is extremely poor (0.9 vs. 0.3 in a car). Gazelles and goats have a transport energy efficiency similar to that of a cheetah at aerobic speeds [26]. Thus, many animals including the cheetah achieve a 4 times improvement in transport efficiency over a highly energy-efficient electric car, while operating with at least a \((0.94/0.25) \times (130/81) = 6\) times disadvantage compared to it (inefficient engine, no wheels). Pronghorn antelopes, the prey of cheetahs, have top speeds that are nearly as fast as cheetahs since they need to outrun the cheetah to live [29]. However, unlike sprinting cheetahs, they can maintain a 45 mph speed (20 m/s) for long sustained periods of time with an energy efficiency comparable to cheetahs, to which they are similar in weight. The power consumption of an antelope running at 30 mph is 1.9 kW (130 N \(/C^2\) 13 m/s), about 20 times less than an average gasoline car at the same speed.

Airplanes and birds both operate near the fundamental limits set by the laws of aerodynamics and fly at an optimal speed needed to support their weight and minimize air drag [30]. A bar-tailed godwit bird flying nonstop from Alaska to New Zealand (7,008 miles) at 36 mph on its stored-fat fuel [31] has nearly the same range as the maximal range of a Boeing 747–300 (7,440 miles) flying at 555 mph on its stored jet fuel. The bar-tailed godwit takes 8.1 days for its flight over the central Pacific Ocean and maintains a 9 times increase in its basal metabolic rate. Somehow, it endures sleep deprivation and potential dehydration to complete its heroic journey [31]. There is significantly more room for improvement in land-transport energy efficiency than in air-transport energy efficiency [1]. However, new designs for micro-air vehicles are exploring the use of flapping and flexible wings that are aerodynamically important in small birds but less important for large airplanes and in big birds that glide [32].

Could animal transport inspire the design of cars? It is possible that it will, but it will require deep and insightful knowledge of both animal transport and car engineering to pluck this high-hanging fruit. Clearly, many of the constraints and goals of human transport are different from those of animal transport and we can’t incorporate our energy for high-speed transport within our own bodies like fast animals do. Therefore, it is likely that insights and principles will be useful, not details, just as has occurred in bird-versus-airplane design since the days of the Wright brothers. However, paradoxically, to get these insights, one will need to understand a lot of details in both fields. This principle is true in all of bio-inspired design. For example, in the RF-cochlea section in Chapter 23, we described how the architecture of a biological cochlea inspired the design of an efficient RF

---

2 In contrast, on recumbent bicycles where the supine rider is enclosed in a carbon-fiber-Kevlar-composite shell, the drag coefficient of the overall highly streamlined bicycle can be quite low. Consequently, Sam Whittingham has achieved a top speed > 82 mph on such a bicycle and an energy efficiency near ~41 N [28]. P. Grogan. Sam Whittingham tops 80 mph – on a push-bike. *The Sunday Times*. September 20, 2009.
spectrum analyzer for advanced radio applications. This example required a good knowledge of both cochlear models and of traditional RF design, failing which the algorithmic insight of a cochlear model would have been missed or the experimental performance of the RF design would have been poor.

26.9 Principles of low-power design in transportation

Why do animals have excellent transportation efficiency? One key is that they are light, the analog of having small capacitance to reduce power in electronic design. A car weighs 1000 kg but transports a 65 kg human. If four people use a car, the car’s per-person transport energy efficiency improves by a factor of almost 4. So, we are observing the classic flexibility-efficiency tradeoff of low-power design. If the car is to be flexible enough to handle situations involving transport of up to 4 persons, its efficiency for transporting a single person is degraded. Degrees of freedom needed to maintain flexibility hurt energy efficiency as in electronic systems. In fact, animals pay a price for flexibility as well. In order to have a universal energy currency molecule, ATP, available to power various activities flexibly, they have to suffer the inefficiency of two energy-conversion steps, one from food to ATP, and another from ATP to useful work, rather than one direct step that converts from fuel to electricity, as some bacteria accomplish. All energy-efficient transport vehicles from trains to cheetahs have encoded in the shape of their bodies a long-and-lean structure that minimizes drag. Trains exploit parallelism to achieve better transport efficiency than cheetahs when full (see Equation (26.21)). Animals recycle spring energy in their tendons to improve their transport efficiency just as cars with regenerative braking or adiabatic digital circuits do [29]. The car is also subject to a robustness-efficiency tradeoff as in other low-power systems: heavier cars do better in accidents but are highly energy inefficient. So, there are clear connections between the principles of low-power transportation and the principles of low-power electronic design which we discussed in Chapters 21 and 22. But what is the connection between information and energy in transportation, a connection that led to several power-saving principles in prior chapters?

The physical variables that change state when we move are our position and velocity. Transportation may be described as an information-processing problem where our state needs to change from $x_0$, our initial position, to $x_f$, our final desired position. To change this state variable from $x_0$ to $x_f$, we need to change the state of another of our variables, our velocity, such that its integral over time is $x_f - x_0$. And, as is true for all systems, it costs energy to maintain state (maintain velocity in spite of drag and rolling friction) and to transform state (the costs of increasing car kinetic energy). We exploit the technology of transport, which can be designed in various topologies (gasoline cars, electric cars, trains), to solve this task. Higher average speeds at which the task is solved lead to higher power consumption. The feedback loop implemented by the visual sensing system of
the driver and his control strategy of the actuating system, i.e., the car, ensures that the car stops and starts at needed positions along the way, with what is usually adequate precision. Thus, the task, technology, topology, speed, and precision costs of a low-power system illustrated in the low-power hand of Figure 1.1 also apply to cars. The car already implements one good principle of low-power design through its use of a feedback loop, i.e., it separates the costs of speed and precision by having an accurate sensor (the driver’s eyes) and control system (the driver’s brain) determine the precision of transport while the actuator determines its speed. The mutual information that is of relevance in a transportation task is that between a desired smooth, relatively fast transport trajectory in the head of the driver and the actual transport trajectory that is achieved. In the future, a trajectory that weighs the costs of carbon emissions will also be important.3

One principle of low-power design that cars can exploit in the future lies in improving the balance between computation costs and communication costs. Cars can wirelessly communicate with traffic lights and with each other such that the transportation of several drivers is more optimal, therefore saving energy. For example, traffic lights could automatically adapt their timing within a reasonable range such that the directions and locations of high flux have lower waiting times than the directions and locations of lower flux. Traffic lights can be coordinated and synchronized like interacting phase-locked loops that receive correction inputs based on traffic flux counts. Traffic lights could also adapt to patterns that are automatically recognized as being due to an accident scenario. The power costs of wireless transmission for relatively short ranges is extremely cheap, especially when compared with the phenomenal power costs of transportation (1 to 10 W versus tens of kW). Furthermore, energy harvesting from LEDs in traffic lights or RF transmissions from incoming cars can provide constant recharging boosts to such systems such that they can be self-powered (see Chapter 17 for a discussion of far-field wireless recharging systems). Car-to-car hopping can be used for longer-range communication, which is significantly more power efficient per unit distance than a non-hopping strategy (\( N(R/N)^2 > R^2 \) for \( N > 1 \) in an \( N \)-hop network). Needless to say, the benefits of such sensor-network schemes will have to be weighed against their costs and ease of implementation within an existing infrastructure. Adaptive traffic lights and car-to-car communication are being researched [33].

We shall now shift gears from discussing how to minimize power consumption to discussing how to generate power. We shall begin with what is likely to be the most important source of the power in our future, solar electricity.

3 Accidents result because of conflicting control algorithms in the heads of different drivers, the imprecision and/or slow reaction times of a drunk-driver’s control algorithm, or the disobedience of traffic rules. Thus, driving precision is strongly determined by feedback loops. The power costs of precision are largely borne by the driver and are relatively small. From [27], D.J. Morton and D. D. Fuller, Human Locomotion and Body Form (New York, NY: Waverly Press, Inc., 1952), they are estimated to be an additional 83 W over the basal 81 W metabolic rate of the driver.
Solar electricity generation

David Goodstein and others have pointed out that the only renewable energy source capable of solely powering our planet at its expected and future power consumption without an incredible use of land area is the sun [5]. The sun transmits 1366 W/m² of power to the Earth when its rays are orthogonal to a location on it [34]. However, 30% of this radiation is reflected back into space, partly by clouds, and about 19% is absorbed directly by the atmosphere [35]. Attenuation factors that vary as the angle of the incident radiation varies throughout the day, the variation in latitude of various places on Earth, the variation in cloud cover in different regions, the complete absence of the sun at night, and the variation in sunshine with seasons cause the power density of the sun to fluctuate over spatial location and over time. The power density of the sun at a particular region on Earth, termed its insolation, is often integrated over the span of a day and expressed in units of kW h/m² per day. Multiplication of the number quoted in kW h/m² per day by 1000/24 yields the average daily insolation in units of W/m². The average annual insolation can range from 100 W/m² in Helsinki, Finland, to 320 W/m² in Inyokern, California, USA. Not surprisingly, there is more variability across seasons at extreme latitudes than at equatorial latitudes. To find the insolation at any latitude and longitude on earth, or at the location where you live, visit [36]. Insolation for various major US cities is available at [37].

Solar photovoltaic cells or photovoltaics that convert solar energy to electricity have efficiency limits that are determined by laws of physics that govern the interaction of light with matter. We shall discuss some of these laws. Losses due to shadowing, due to reflection at the cell surface, due to a loss in light collection area (a loss in ‘fill factor’ as in the imagers described in Chapter 19) can further reduce efficiency. Low-efficiency cells are typically cheap to manufacture while high-efficiency cells are typically expensive to manufacture. The efficiency of solar cells can range from 2% to 40%. Most commercial systems that can be mass manufactured are in the 10%–20% range today. We shall first discuss how solar photovoltaics function and then discuss fundamental limits on their energy efficiency.

In Chapter 11, we discussed how light creates electrons in pn junctions, and how we could exploit such phototransduction to create a photoreceptor. In Chapter 19, we discussed how to build low-power imagers using pn junctions. The basic principles of phototransduction discussed for these applications also apply to solar cells. Solar radiation is largely composed of energy in the visible light and near-infrared regions. Figure 11.2 (a) reveals a pn junction formed by abutting a semiconductor of n-type material with a semiconductor of p-type material. Figure 11.2 (b) reveals the energy diagram that describes such junctions with $E_C$ and $E_V$ representing the minimal and maximal energy of the conduction band and valence band respectively. The depletion region created by the equilibration of

---

4 Nuclear energy from fission sources will eventually also be non-renewable and there has not yet been a breakthrough in nuclear fusion, which could potentially be renewable.
26.10 Solar electricity generation

Drift and diffusion currents at the border between the n-type and p-type regions cause the bending of the band energies. The electric field in the depletion region is such that positive depletion charge in the n-type region, created when electrons diffuse away from the n-type region, raises the potential of the region or equivalently lowers the electron energy in the n-type region. Similarly, negative depletion charge in the p-type region, created when holes diffuse away from the p-type region, lowers the potential of the region or equivalently raises the electron energy in the p-type region. In a junction with zero voltage across it, at thermal equilibrium, the Fermi level $E_F$ is the same at all spatial locations such that the average energy of an electron is the same at all spatial locations. If a photon with energy $h\nu$ greater than the band-gap energy $E_G = E_C - E_V$ is absorbed by the junction, the energy can be used to promote an electron from a low-energy state in the valence band to a higher energy state in the conduction band. The absorption of energy creates a hole in the valence band in the energy state that the electron has come from and an electron in the conduction band in the energy state that the electron has gone to.

The created electrons travel ‘home’ to their native majority-carrier n-type region and the created holes travel home to their native majority-carrier p-type region because these regions represent attractive regions of low energy. Figure 11.2 (b) shows that electrons created in an n-type region do not travel since they are already in a region of low energy. Similarly, holes created in a p-type region do not travel since they are already in a region of low energy. In any case, the absorption of the photon results in the net arrival of an electron in the n-type region and the net arrival of a hole in the p-type region. The arrival of electrons in the n-type region and the arrival of holes in the p-type region effectively results in a floating current source across the junction. Figure 11.2 (c) in Chapter 11 shows that light effectively shifts the $I-V$ curve of the junction due to the presence of the floating current within it. If the junction is in a short-circuit configuration, the current will appear as an enhanced reverse-bias current. If the junction is in an open-circuit mode, the floating current forward-biases the junction to a voltage such that the forward-bias current balances the light-dependent reverse-bias current. Consequently, the open-circuit voltage is at a steady-state value. In Figure 11.2 (c), the short-circuit current corresponds to the value at $V = 0$ while the open-circuit voltage corresponds to the value at $I = 0$.

Suppose the junction can be characterized by the usual equation for a pn junction,

$$I = I_S(e^{\nu V/kT} - 1)$$

where $I$ is the forward-bias current through the junction and $V$ is the forward-bias voltage. The parameter $I_S$ is determined by minority-carrier concentrations, the area of cross-section, and carrier diffusion lengths within the junction [38]. Suppose initially that all photons are of frequency $\nu$ and have an energy $h\nu$. Then, if the probability that a photon of energy, $h\nu$, creates an electron hole pair is $\alpha(\nu)$ and
there are \( I \) photons per second arriving over the collection area of the junction, the current in the junction including the reverse-bias photocurrent is given by

\[
I = I_S(e^{q\nu/kT} - 1) - \alpha(\nu)qI_l
\]  

(26.27)

The short-circuit current is obtained by setting \( V = 0 \) in Equation (26.27). Thus,

\[
I_{sc} = -\alpha(\nu)qI_l
\]  

(26.28)

The open-circuit voltage, \( V_{oc} \), of the junction is given by setting \( I = 0 \) in Equation (26.27). We find that

\[
V_{oc} = \frac{kT}{q} \ln \left( \frac{\alpha(\nu)qI_l}{I_S} + 1 \right)
\]  

(26.29)

The incoming radiation has a power of \( hvI_l \) while the power of the solar cell cannot be greater than \( V_{oc}I_{sc} \). Therefore, the ratio \( (V_{oc}I_{sc})/(hvI_l) \) establishes a crude upper bound on the solar-cell efficiency. From Figure 11.2 (c), since \( I_{sc} \) and \( V_{oc} \) cannot simultaneously be maximal, the actual power output of the solar cell \( IV \) is maximized when \( I < I_{sc} \) and \( V < V_{oc} \). For a given \( I_l \) set by solar insolation, it is clear that the efficiency of the solar cells is maximized when \( \alpha(\nu) \) is maximum, which improves both the open-circuit voltage and short-circuit current, and when \( I_S \) is minimum, which improves the open-circuit voltage. What determines \( \alpha(\nu) \), what is the distribution of photons of a given frequency \( \nu \) in solar radiation, and what is the exact bound on the limit of solar-cell efficiency?

Shockley and Queisser provided insight into the limits of solar-cell efficiency in a landmark paper [39]. The analysis in their paper indicates that the limits of solar-cell efficiency in a single-bandgap pn junction receiving 1 sun’s worth of radiation is 31%. We shall summarize the key ideas in an intuitive fashion here. Readers interested in further details should consult [40]. Figure 26.8 (a) shows that, when a photon with energy \( hv > E_G \) creates an electron-hole pair, energy in excess of \( E_G \) is quickly lost as heat such that only an energy equal to \( E_G \) is available as electricity. Thus, while high-energy photons have a high probability of creating an electron hole pair since many possible states are available for their creation, a good fraction of their energy is lost as heat. Lower-energy photons with an energy \( hv \) just greater than \( E_G \) have better efficiencies of energy extraction. Shockley and Queisser assumed that any photon with energy \( hv > E_G \) that created an electron-hole pair would do so with effective energy \( E_G \) and that any photon with energy \( hv < E_G \) would create no electron-hole pair.

Figure 26.8 (b) shows the known 6000 K black-body spectrum of solar radiation, i.e., the solar photon probability distribution for photons of various frequencies \( \nu \). The shaded area to the right of the minimal \( hv_e = E_G \) frequency yields the net fraction of photons in solar radiation that contribute to electric energy generation. Each of the high-energy photons that are represented in this region contributes an energy of \( E_G \) to electric generation and wastes \( hv - E_G \) as heat energy. From the entire probability distribution of Figure 26.8 (b), and the energy \( hv \) of single photons, we can compute the total incoming energy in solar radiation.
From the shaded area in Figure 26.8 (b), we can compute the electric energy generated by the high-energy photons. The ratio of the electric energy to the total incoming radiation energy then yields an ultimate limit for the solar-cell efficiency. Shockley and Queisser showed that their ultimate limit could be attained if the only method for electron-hole pair destruction at 300 K is radiative, i.e., incoming thermal energy at 300 K from the environment creates electron-hole pairs, which then recombine to generate outgoing 300 K blackbody radiation that balances such generation. In this limit, the value of $I_S$ is as low as it can possibly be, and consequently decrease the minority-carrier lifetime $\tau$, and hence, the open-circuit voltage given in Equation (26.29) is reduced. The use of pure semiconductors is thus important for achieving high efficiency, but making pure materials is expensive. The solar cell must be thick enough such that the probability of absorbing a photon is high. Figure 26.8 (c) illustrates an idea for increasing the fraction of photons that contribute to electrical energy in a solar cell. If we have multiple pn junctions made of materials with progressively smaller bandgaps, we can first extract the
energy in the highest-energy photons efficiently, then in those with moderate energy, then in those with the lowest energy, and so on. Such an hierarchical spectral-energy extraction scheme is very much like that in a biological cochlea or in an RF cochlea (see Chapter 23), i.e., it is like a ‘cochlear solar cell’. The overall scheme extracts incoming solar photon energy in all spectral bands such that the area of the shaded region in Figure 26.8 is maximized. It also extracts this energy in a fashion such that little of it is wasted as heat. In theory, such schemes have been shown to be capable of nearly 70% efficiency [41].

Figure 26.9 illustrates an idea for improving solar-cell efficiency in a ‘solar concentrator’. We gather radiation over a large area and focus it into a small area such that the effective intensity of the sun per unit area is increased from ‘1 sun’ to, say, ‘400 suns’. The $I_l/I_s$ ratio in Equation (26.29) then increases, improving open-circuit voltage and thus efficiency. One advantage of the concentrator is that smaller active areas of pure material are needed for the same power, which can minimize cost. However, such schemes have to track the sun to ensure that its radiation does not move off the focal point where the solar cell is located. Solar concentrators have indeed improved efficiencies over simple flat-panel solar cells.

The ultimate limit on solar-cell efficiency occurs when a graded-bandgap or multiple-junction material like that in Figure 26.8 (b) is combined with a concentrator like that in Figure 26.9 to create a structure where it appears that the solar cell is ‘surrounded’ by 6000 K black-body radiation from the sun, on all sides. It absorbs all of this radiation, and then reradiates a small fraction of it as black-body radiation at 300 K. In this limit, the solar cell is just a ‘heat engine’ as shown in Figure 26.4: it absorbs heat from the 6000 K ‘hot reservoir’ of the sun, converts most of it to useful electrical work, and loses some of it as radiated heat to the
surrounding 300 K ‘cold reservoir’. Thus, from the Carnot efficiency limit of Equation (26.13), the best possible efficiency of a solar cell is given by

\[ C_{s} = 1 - \frac{300}{6000} = 95\% \] (26.30)

The best solar cells that have been built today operate very near 40%, with every small percent improvement that is eked out requiring ingenuity and relatively expensive fabrication. The key bottleneck to solar-electricity generation on a large scale in the world today is the implementation of cost-effective strategies that are also efficient.

If \( S \) is the solar insolation in watts per square meter, \( e \) is the efficiency of the solar cell expressed as a fraction, \( C(e) \) is the cost in dollars per square meter of installation of a solar plant, \( N \) is the desired number of years to recoup the installation investment, and \( i \) is the average rate of inflation over \( N \) years expressed as a percentage, then \( S_{\text{cost}} \), the cost in cents per kWh of solar electricity, can be shown to be

\[ S_{\text{cost}} = 11.4 \left( \frac{C(e)}{eSN} \right) \left( 1 + \frac{i}{100} \right)^N \] (26.31)

Hence, if \( e = 0.1, C(0.1) = \$600/m^2, S = 190 W/m^2, N = 30 \) years, and \( i = 3\% \), then \( S_{\text{cost}} = 29 \) cents per kWh. Hence, competing with the cost of generating fossil-fuel electricity at 4 cents per kWh is difficult. However, with increasing research and with increasing economies of scale, \( C(e) \) has been constantly reducing. An important win-win situation can occur if we lower power consumption such that the net change to the user is cost neutral or only results in a modest increase in cost: the electricity costs more but we use less of it such that our overall cost is relatively unchanged.

The intermittency of solar energy implies that we must draw on stores when it is not available, e.g., at night, and replenish these stores during the day. Various storage options are being explored including compressed air, flywheels, and batteries. One promising option may be the use of electrochemical capacitors, sometimes called ultra-capacitors or super-capacitors, which are capable of many cycles of rapid charge and discharge and that have relatively high power densities [42]. Such capacitors are essentially high-surface-area double-layer capacitors, i.e., the Helmholtz capacitors described in Chapter 25.

Ultra-capacitors are complementary to batteries that have fewer cycles of charge and discharge but relatively high energy density. They have already been used for regenerative braking applications in electric cars and hybrid cars. Ultra-capacitors implement short-term storage of energy, and batteries implement medium-term storage of energy. The ultimate in long-term storage of energy is to convert solar energy to a highly energy-dense chemical fuel. Biology has accomplished such storage via the process of photosynthesis in plants for hundreds of millions of years. We shall now briefly discuss biofuels and their importance.
### 26.11 Biofuels

Figure 26.10 illustrates an essential feedback loop between plants and animals. Plants harvest the energy in sunlight to split water into hydrogen and oxygen as part of the process of photosynthesis. The hydrogen is bound with CO\(_2\) to create energy-rich molecules such as glucose (C\(_6\)H\(_{12}\)O\(_6\)) which in turn is often bound up in polymers like starch and cellulose. Plants also generate the oxygen that we breathe. Animals eat plant foods (or eat other animals that eat plant foods) and oxidize C\(_6\)H\(_{12}\)O\(_6\) molecules to water (H\(_2\)O) and carbon dioxide (CO\(_2\)). The energy derived from the process of oxidation generates ATP, which is used to power various energy-consuming processes in animals. Thus, plants are the solar cells and fuel generators for animals. Animals provide raw materials useful to plants.

Research is under way to create biofuels that can power cars in the future by converting grasses and non-edible plants that contain cellulose into biofuels, i.e., to create 'grassoline' [43]. Such fast-growing plants can grow in land areas where normal food crops cannot, use relatively little water, and do not encroach on valuable farm land. Since plants absorb CO\(_2\) in the atmosphere, which is then returned to the atmosphere when the fuel is burned, biofuels are net carbon neutral. The process of converting cellulose in plants to fuel in an economical fashion is technically very challenging and is an active area of research. Scientists are attempting to take inspiration from bacteria and fungi in the guts of cows and in termites respectively to understand how to digest these recalcitrant plants and thus create economical biofuels [44], [45].

If electric cars operate with sufficiently low power consumption in the future, biofuels could potentially power fuel-cell-based batteries in these cars, i.e., directly convert the chemical energy in a fuel to electricity, rather than burning it via a heat engine as in a conventional gasoline-powered car. Chapter 25 contains a discussion of how fuel cells operate. The high energy density of biofuels implies that car batteries can then be lighter, and the use of an electric motor rather than a heavy engine can further lighten the car. It is worth noting that biofuel-based fuel cells have been explored for implantable applications for a long time [46]. A microfluidic fuel cell suitable for implantable applications has recently been described [47]. One challenge in the operation of such biofuel cells has been that...
the enzymes that are used to oxidize the fuel lose their efficacy after some time, making them unattractive in long-term implants or in car applications that may need years of battery operation. Cells solve these problems by constantly degrading and regenerating enzymes needed for various biological processes such that they always maintain their efficacy.

26.12 Energy use and energy generation

Low-power systems can enable sources of energy that would normally be impractical for powering an application to become practical. In inductive links (Chapter 16), in piezoelectric harvesters, in electric motors, and in electric cars, we have seen that a low-power system can improve the energy efficiency of an overall system by altering the effective reflected load seen by the energy source. A system is most energy efficient when there is minimal power transfer but only achieves 50% efficiency at maximal power transfer. In the chapter on batteries (Chapter 25), we saw that a low-power system does not increase battery lifetime merely because of a low-power draw but also because it enables higher energy density, higher efficiency, and lower fade capacity in the battery. In electric cars, the use of a relatively light and efficient electric motor rather than a heavy engine enabled an energy source with significantly lower energy density, i.e., a battery, to become practical for powering a car. The cost effectiveness of solar electricity is improved if electricity consumption can be reduced, thus enabling green electricity rather than ‘red’ electricity. The principles of adiabatic design in Chapter 21, the Shannon limit on the minimum energy needed to compute in Chapter 22, and the Ragone-curve tradeoff between energy density and power density in Chapter 25 all reveal that if you can pull energy out of a source slowly, you can waste less of it and create a higher capacity to store it. The central take-home lesson from these numerous examples is that energy use and energy generation are deeply linked. We must try to optimize them jointly rather than treat them as two separate problems.

References

Energy harvesting and the future of energy


26.12 Energy use and energy generation


