GARY BERNSTEIN

PHYSICS 016: ENERGY, OIL, & GLOBAL WARMING

SPRING 2021

Contents

1

What is Energy?

Energy is the currency of the physical world: nearly everything that happens in our Universe can be described as the transfer of energy from one object to another. This is true for stars and planets; volcanoes and weather; people and cars and light bulbs and motors; and cells and bacteria and subatomic particles. An ecosystem can be described as the transfer of energy from sunlight to plants to herbivores to predators.

Energy is also the basis of post-agrarian human life. Early humans made use of energy generated only by their own bodies (a human generates about 100 W of ***power***, which we'll define later), and wood/dung fires for cooking. Great advances accrued to civilizations that harnessed the energy of draft animals. Modern society was, however, only enabled by the exploitation of fossil fuels, starting with steam engines in 1800's. Currently the most profligate energy users in the world—US citizens—use about 10,000 W of power, about 100× more than it takes to run our bodies, and equivalent having 100 "energy servants." *Each US citizen would need to burn his/her weight in coal every 2–3 days to run their lives!*

Phrase taken from textbook by Wolfson.

Really we burn only about 4.5 kg (= 10 lb) of coal per day per person, and get the rest of the energy from oil, natural gas, nuclear, and a little renewable energy.

We have come to depend on this, and most Americans take cheap energy as their birthright. Is this sustainable? This would mean that we have some means to continue this lifestyle for hundreds of years without incurring huge damage or cost. There are ways in which current energy practices are not sustainable and pose real problems we need to deal with:

- Supplies of some critical energy sources are finite and likely to begin dwindling. Domestic production of oil peaked in the 1970's, then dropped in half by 2005, then began increasing very rapidly, passing the 1970's record in 2013 and now 150% of the 1970's peak—is this sustainable? US natural gas production is also now above the 70's level. The future of world oil production is highly uncertain, even as demand continues to grow rapidly. Coal supplies are good for 100 years perhaps, but that's not forever either.

- Laws of supply and demand, plus political instabilities in regions of oil wealth (or *caused* by oil wealth) will inevitably increase prices well beyond our current comfort zone. International relations have long been highly influenced by energy access.
- Energy production has long exacted environmental and social costs: air pollution, oil spills, strip mining, acid rain, radiation leakage from Chernobyl. The costs of this damage have rarely been included in the price of energy.
- The greatest challenge is now apparent: global warming. Burning of fossil fuels has raised the atmospheric CO_2 concentration to levels having significant effects on climate due to the **greenhouse effect.** Massive relocation of people and resources will be forced by the effects on ocean levels and weather. Again the costs are not included in energy prices, and it is a problem that must be addressed on global scale immediately. How can we produce energy with minimal CO_2 release?

So we first must understand what energy is.

Kinetic energy

THE SIMPLEST KIND of energy is ***kinetic energy,*** or the energy associated with *motion.* It is defined as

$$KE = \frac{1}{2}mv^2, \tag{1.1}$$

where

- m =***mass***, the amount of material that is in motion. In this course we will use standard metric units wherever possible. Mass is measured in ***kilograms (kg)***. 1 kg is 2.2 pounds.[1]
- v =***velocity***, or speed, which we will measure in meters per second (m/s or m s^{-1}).

How much kinetic energy does a car have when travelling at highway speed? A small car might weigh about 2200 lb, and be travelling at 67

---

[1] Technically this is only true on the surface of Earth since pounds measure "weight", not "mass." We won't worry about this difference.

miles per hour (mph). We'll have

$$
\begin{aligned}
m &= (2200\,\mathrm{lb}) \times \frac{1\,\mathrm{kg}}{2.2\,\mathrm{lb}} = 1000\,\mathrm{kg} \\
v &= 67\,\frac{\mathrm{miles}}{\mathrm{hour}} \times \left(\frac{1\,\mathrm{hour}}{3600\,\mathrm{s}}\right) \times \left(\frac{1609\,\mathrm{m}}{1\,\mathrm{mile}}\right) = 30\,\mathrm{m\ s}^{-1} \\
\Rightarrow KE &= \frac{1}{2}mv^2 = \frac{1}{2}\,(1000\,\mathrm{kg})\left(30\,\mathrm{m\ s}^{-1}\right)^2 \\
&= \left(5 \times 10^2\right) \times \left(3 \times 10^1\right)^2\ \mathrm{kg\ m^2\ s^{-2}} \\
&= \underline{\underline{45 \times 10^4\,\mathrm{J}}} = \underline{\underline{450\,\mathrm{kJ}}} = \underline{\underline{0.45\,\mathrm{MJ}}}.
\end{aligned}
$$

In our first numerical example we have used some techniques that will be *extremely* important for this class—or any quantitative work. The first is the use of **scientific notation.**

Scientific notation and metric prefixes

We will deal with some very large (and small) numbers. For example the total 2019 energy consumption of the world was about 584,000,000,000,000,000,000 J. This is cumbersome and unintelligible, so will write this as 5.84×10^{20} J. You should *always* use scientific notation for results that are more than a few thousand or smaller than ≈ 0.001. Please refresh your scientific notation skills if they are rusty (see Wikipedia, Khan Academy, etc.). A few reminders about scientific notation, with the nomenclature that 2.6×10^7 has the *coefficient* or *mantissa* 2.6 and the *exponent* 7:

- When you *multiply* two quantities, you *multipy* the coefficients and *add* the exponents.
- When you *divide* two quantities, you *divide* the coefficients and *subtract* the exponents.
- When you *add* two quantities, you must first *match* the exponents and then *add* the coefficients.
- When you *subtract* two quantities, you must first *match* the exponents and then *subtract* the coefficients.
- In all cases you may need to re-adjust the exponent so that the cofficient is between 1 and 10.
- Your calculator can of course do this for you. You should *have a calculator that can do scientific notation and know how to use it.* Typically to enter something like 2.5×10^{-5} you would type `2.5E-5`, and you'll often see numbers written this way.

The metric system makes scientific notation easier with a set of *standard prefixes* representing all powers of 1000. You are probably familiar with *kilo (k)*$= 1000 = 10^3$, and *milli (m)*$= 0.001 = 10^{-3}$. We obviously need something bigger for the annual world energy usage: we'll whip out *exa (E)*$= 10^{18}$ to write $5 \times 10^{20}\,\mathrm{J} = 500 \times 10^{18}\,\mathrm{J} = 500\,\mathrm{EJ}$.

You should know all of the metric prefixes from micro (10^{-6}) through exa for this course. Here is a helpful table.

Another essential skill is **units conversion**. We will need this nearly every day: to change between standard metric units and many other (useful) units, and also to recognize when unit combinations such as $\mathrm{kg\,m^2\,s^{-2}}$ can be replaced with a simpler unit like *J* above, or when we need to do the reverse.

Units conversion

The units of your calculations are as important as the numerical parts. You should never leave the units out, and you should do all the algebra on the units too. Remember these rules:

- To *add or subtract* two numbers, the units must *match*.
- To *multiply or divide* two numbers, you *multiply or divide* the units.
- If you *square* a number, you also *square* its units. Same for square roots, cubes, or any power.
- If you are taking a *logarithm or exponential* of a number, it should be "dimensionless," meaning that all the units have cancelled out. If you find yourself trying to take the log of a unit, you've probably made a mistake.

While we would like to skip this tedium and get the answer in our heads, it's very easy to get confused with complicated unit mixtures. Here is my suggested strategy for getting the correct answers on *any quantitative calculation:*

1. Write out your formula symbolically, *e.g.* $KE = \frac{1}{2}mv^2$.
2. Figure out what you're solving for, and if it's not already isolated on the left-hand side, do the solution algebraically. *No numbers yet!*
3. Insert known values for symbols, *with their units.* Use the units you are given for each quantity.
4. Gather the numerical part of answer together, use scientific notation to do arithmetic when the answers are more than few thousand or less than ~ 0.001.
5. Also do the arithmetic on the units. If you have an addition/subtraction in your equation, you need to insure that the units match before you add the numbers.
6. Wherever you have mismatched or undesired units, convert them by this technique: multiply by a fraction that has different units on the top and bottom, but is equal to 1 because we've made the top and bottom the same. For example, since we know that $1\,\mathrm{mile} = 1609\,\mathrm{meters}$, I also know that

$$\frac{1609\,\mathrm{m}}{1\,\mathrm{mile}} = 1$$

because any fraction with equal top and bottom is equal to one, and I can multiply my answer by 1 any time I want without changing its meaning. If you've got the fraction right you'll be able to cancel out the undesired units and leave behind just the ones you want.

7. Convert units to simpler form if possible (*e.g.* $\mathrm{kg\,m^2\,s^{-2}} \rightarrow \mathrm{J}$).

Gravitational potential energy

Suppose I hold a bowling ball over my head. I know if I let it go, gravity will cause it to fall toward the floor[2] and in the process will gain speed and hence kinetic energy. Because the raised ball has the *potential* to gain KE from gravity, we can assign it ***gravitational potential energy***. The formula is

$$PE_{\text{grav}} = mgh, \tag{1.2}$$

- m = mass of the object;
- g = the value $9.8\,\mathrm{m\,s^{-2}}$, which gives the strength of gravity near the surface of Earth;
- h = height above the floor.

As the ball drops, PE is converted to KE. *The total energy $E = PE + KE$ stays the same as the ball falls.* We can use this to answer some questions, like: how much KE does the ball have when it hits the floor (or my toe)? How fast is it moving when it hits? Let's assume the bowling ball has mass $m = 6\,\mathrm{kg}$ and I drop it from height $h = 2\,\mathrm{m}$.

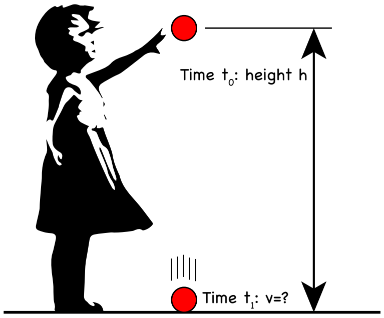


Figure 1.1: A ball is dropped from height h. How fast is it moving when it hits the floor?

Almost all energy problems can be solved by setting up an equation where we set the total energy at the start equal to the total energy at the end. We'll call time t_0 the moment when I release the ball, when it has height $h_0 = h$ and speed $v_0 = 0$. At time t_1 when it hits the floor, it has height $h_1 = 0$ and speed v_1 that we don't know. The constant total energy means

$$\begin{aligned}
E_0 = KE_0 + PE_0 &= KE_1 + PE_1 = E_1 && (1.3)\\
\frac{1}{2}mv_0^2 + mgh_0 &= KE_1 + mgh_1 && (1.4)\\
0 + mgh &= KE_1 + 0 && (1.5)\\
\Rightarrow KE_1 &= mgh = 6\,\mathrm{kg} \times 9.8\,\mathrm{m\,s^{-2}} \times 2\,\mathrm{m} && (1.6)\\
&= 117.6\,\mathrm{kg\,m^2\,s^{-2}} = \underline{\underline{117.6\,\mathrm{J}}}. && (1.7)
\end{aligned}$$

---

[2] Really, the bowling ball is falling toward the center of Earth.

We also can solve for the final speed of the ball:

$$KE_1 = \frac{1}{2}mv_1^2 = mgh \tag{1.8}$$

$$\Rightarrow v_1^2 = 2gh \tag{1.9}$$

$$\Rightarrow v_1 = \sqrt{2gh} = \sqrt{2 \cdot 9.8\,\mathrm{m\,s^{-2}} \cdot 2\,\mathrm{m}} \tag{1.10}$$

$$= \sqrt{39.2}\sqrt{\mathrm{m^2\,s^{-2}}} = \underline{\underline{6.26\,\mathrm{m\,s^{-1}}}}. \tag{1.11}$$

Notice how the units get square-rooted in this solution, just like the numbers do.

Note that in (1.11) we see that the speed of impact does not depend on the mass of the ball (although the energy does)! This fact led Galileo and Einstein to some great physics discoveries.

Conservation of Energy

When the ball is travelling through the air its total energy stays the same. We say that total energy is "conserved," meaning it never changes. The key principle of this course is that **the total energy is always the same, as long as we keep track of all the possible places that energy can go.** This central concept has been tested and verified so often that it is called a "Law" of nature:

> *The Law of Conservation of Energy:* **Energy is never created or destroyed, it is only transformed from one kind to another.**

Be careful not to confuse the physical "law of conservation of energy" with the colloquial usage that "conserving energy" is consuming less energy.

We can state this another way:

> **The Law of Conservation of Energy: The total amount of energy never changes.**

The consequences are obvious. If we want to get a car going to 60 mph, it's going to have KE, and we're going to need to get those Joules from somewhere. If we want an elevator to take us to the top floor, we are gaining PE, and those Joules need to come from somewhere. We continually need to find energy sources to provide the energy that our daily life requires.

Here's a useful analogy: energy is like money. Every bit that you spend has to come from somewhere.[3] You can think of kinetic energy as the cash in your pocket, and potential energy is like a bank account. Additions to one must balance subtractions to another. Also this analogy is useful because we see that KE can never be negative, only zero or positive. You can have no cash in your pocket, but you can never have negative cash!

On the other hand you *can* have negative bank account balances, and likewise you can have negative potential energy. For instance if we dug a hole in the floor and dropped a ball in there, we could have $h < 0$ and $PE_{\mathrm{grav}} < 0$. This is like being in debt to the energy bank.

[3] Unless you're the government or a forger and can print money out of nowhere. Ignore that possibility for this analogy!

You're going to have to give the ball some other kind of energy in order to get back up to the floor where you're "even" with the bank.

Here's a very simple fact that results from conservation of energy: if I drop a ball from some height h, it will never bounce higher than the point I released it. Think about this: it's true because the ball can never have less than the zero KE that it starts with.[4]

Electromagnetic energy and its many guises

Gravity has potential energy because gravitational pull can put an object into motion and change its KE. The attraction or repulsion between two electric charges q_1 and q_2 can also put them into motion, so we know there is also ***electromagnetic potential energy.*** In the simplest case, two charges at distance r_{12} from each other have potential energy

$$PE_{\text{elec}} = \frac{q_1 q_2}{r_{12}}. \tag{1.12}$$

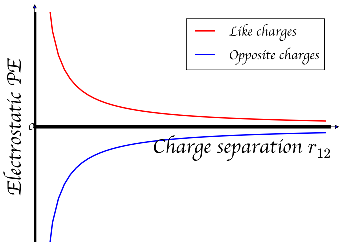


Figure 1.2: The electrostatic PE of like charges is positive and higher as they approach each other. Opposite charges have negative PE and an infinitely deep "well" as they approach.

Notice that when one charge is positive and the other negative, the PE is negative. It gets more negative as they approach—this makes sense, as the charges will speed up and gain KE as they pull together, so the PE must be getting lower (more negative). Opposite charges are "in debt" to the PE bank—you need to add energy to them if you want to pull them apart and set them free, *i.e.* $r_{12} \to \infty$.

We will rarely encounter loose electric charges, and in this class we never use Equation (1.12). But electromagnetic (or EM) energy is critically important because matter is made of microscopic charges, namely electrons and protons, and nearly every rearrangement of matter involves changing the distances between these and hence changing the EM energy levels.

Chemical energy

All of chemistry is the re-arrangement of electrons when atoms come near to each other and bond. Forming new molecules releases or absorbs EM energy. In this case we will call this ***chemical energy***, but remember it's still EM energy. Every chemical reaction releases or absorbs a particular amount of energy. We are mostly concerned with the burning of ***fossil fuels***, which are the remnants of plants and algae that lived millions of years ago and preserved underground.

The fossil fuels are all *hydrocarbons* containing predominately carbon (C) and hydrogen (H) atoms. Energy is released when the C and H atoms bind with oxygen from O_2 in the air. **Coal** is mostly carbon (50-70% for bituminous coal, 90% for anthracite), the rest is H,

---

[4] You can cheat by giving the ball some initial KE, or letting it get energy from some other source.

water, and up to 10% sulfur (S). When coal is burned, about 3×10^7 J are released per kg, mostly through the reaction

$$C + O_2 \rightarrow CO_2. \tag{1.13}$$

Essentially every carbon atom in a fossil fuel yields a CO_2 molecule released into the atmosphere. The increased CO_2 level of Earth's atmosphere causes global warming.

Natural gas is primarily methane, CH_4. When it burns, $5.2 \times 10^7\,\mathrm{J\,kg^{-1}}$ are released from the reaction

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O. \tag{1.14}$$

Notice that natural-gas burning gets energy from its 4 H atoms per carbon yielding water, so natural gas has a much lower CO_2 release per J of energy released. This is one way in which natural gas is "cleaner" than coal. Methane is the simplest hydrocarbon; others have more carbons arranged in chains or rings (ethane, propane, butane, pentane, hexane, heptane, octane...). Petroleum (= crude oil) consists of medium-length liquid hydrocarbon chains that are liquid or semi-solid at room temperature. Crude oil is sold by the barrel (bbl, 42 US gallons) which averages $6 \times 10^9\,\mathrm{J\,bbl^{-1}}$. Before sale, crude oil is **refined** into more homogeneous products, including **gasoline**, which when burned yields about $1.3 \times 10^8\,\mathrm{J\,gallon^{-1}}$, or $4.7 \times 10^7\,\mathrm{J\,kg^{-1}}$. Petroleum has about 2 H per C, between coal and natural gas, so it's between those two for CO_2 release per J of energy.

Explosive TNT releases about $4.2 \times 10^6\,\mathrm{J\,kg^{-1}}$. Note this is a lot less than same weight of gasoline. So why do we think of TNT as being more dangerous?

This is slightly cheating since a kg of gasoline needs roughly a kg of oxygen from the air to burn, but TNT has all the needed elements built in. Even so, gasoline holds much more energy.

Let's do one application of chemical energy: how much gasoline must be burned to accelerate a car from 0 to 67 mph? Earlier we found this needs 0.45 MJ of energy, so we should need

$$0.45\,\mathrm{MJ}\frac{1\,\mathrm{kg\ gas}}{4.7 \times 10^7\,\mathrm{J}} = 0.0096\,\mathrm{kg\ gas} \approx \underline{\underline{10\,\mathrm{g\ gasoline}}}. \tag{1.15}$$

This is about 12 ml or *two teaspoons* of gasoline—a remarkably small amount! Now we see why fossil fuels are so useful!

Actually it takes about 5× this much because of the low *efficiency* of engines. More on this later.

When I jump, or throw a ball, I am creating kinetic energy. Since energy is conserved, this energy must have gotten inside me somehow before I gave it to the ball. That's why we eat food! The energy content of food is typically measured in ***Calories***. A Calorie is just a different unit for energy:

$$1\,\mathrm{Cal} = 4186\,\mathrm{J}. \tag{1.16}$$

Something very confusing: a food *Calorie* is equal to 1000 of a chemist's *calorie*, *i.e.* $1\,\mathrm{Cal} = 1000\,\mathrm{cal} = 1\,\mathrm{kcal}$. Note the use of capital C for the food unit. Outside the US, food is typically labelled in kcal.

There are many units of energy used in different countries and different industries. We must be adept at converting between them. Keep a copy of the conversion table handy.

Unit	Abbrev.	Conversion
British Thermal Unit	BTU	$1055\,\mathrm{J} = 1.054\,\mathrm{kJ}$
Quadrillion (10^{15}) BTU	Quad	$1.054 \times 10^{18}\,\mathrm{J} = 1.054\,\mathrm{EJ}$
kilowatt-hour	kWh	$3.6 \times 10^{6}\,\mathrm{J} = 3.6\,\mathrm{MJ}$
calorie (chemistry)	cal	$4.186\,\mathrm{J}$
Calorie (food)	Cal	$4186\,\mathrm{J} = 4.186\,\mathrm{kJ}$
Tonnes oil equivalent	toe	$4.19 \times 10^{10}\,\mathrm{J} = 41.9\,\mathrm{GJ}$

Table 1.1: Energy conversion factors

Electricity

Electricity is a flow of electrons down a wire. They are basically being pulled by a positive charge at the other end of the wire. This version of EM energy is a very effective means of *transporting* energy from one place to another. A section of the course will be devoted to understanding electricity and how we generate and transport it.

Light

Light is the flow of EM energy across space. We can think of it as a stream of ***photons***, massless particles that travel at the speed of light (duh!) to carry energy from one electric charge to another, whether those electric charges are in the same atom or billions of light years apart in distant galaxies. A section of the course will be devoted to understanding light.

Material deformation

A compressed spring is a classic example of potential energy: it's ready to launch something into motion if you let it go. Sometimes we'll refer to the PE in springs and related things as ***mechanical energy*** since they are parts of macroscopic moving machines. But deep down this is just EM energy, since a coiled spring is storing its energy by squeezing its constituent electrons/protons more tightly together in such a way that they have more EM PE.

Nuclear energy

Particle physicists tells us there are only four "fundamental forces" in nature, *i.e.* only four ways to put a body into motion and hence only four true kinds of potential energy. We have met two of them: gravity and electromagnetism. The other two are forms of ***nuclear energy*** that hold the protons and neutrons together in the nuclei of atoms. While nuclear energy is extremely powerful, it is not encountered in everyday human activities and in fact was not discovered until the

1900's. We will study nuclear energy and its use by humans toward the end of the course.

Thermal energy and heat

According to fundamental physics, once we account for kinetic energy plus gravitational, EM, and nuclear potential energies, we have completed our inventory of energy and should find that every action in the Universe has a balanced energy ledger.

But think of this simple experiment: drop a tennis ball from over your head. This gravitational PE is converted to KE as the ball drops; it bounces and the process reverses. Then it falls again, etc. But on each bounce the ball loses height and eventually comes to rest on the floor, and now has zero KE and zero gravitational PE. Where did its energy go? There have been no chemical reactions, no nuclear reactions, and a negligible amount of energy escaped as light or sound (sound is moving air, a form of KE). We seem to have violated the Law of Conservation of Energy. Which is a bad way to start the course.

The answer is that the missing energy has been transformed into vibrations of the individual atoms and molecules of the ball (and of the floor and air). All the atoms around us are in constant motion and have kinetic energy, but we cannot see it because the atoms are small and are moving in *random directions*. This ball is no longer moving *as a whole*, but its atoms are jostling back and forth in their places. This microscopic form of KE is called ***thermal energy***. The amount of thermal energy inside a material is determined by its ***temperature***. Roughly speaking the average amount of thermal energy per atom or molecule is given by

$$\langle KE \rangle \sim k_B T. \tag{1.17}$$

- The angle brackets indicate that we're talking about the average energy: at any given moment, a given particle may have more or less energy as it jostles about.

- T is the *temperature*. It is essential that we use the ***Kelvin*** temperature scale because we need a scale that registers $T = 0$ at the coldest possible temperature—remember that we cannot have negative KE. The Kelvin scale is related to the more familiar Celsius metric scale and US Fahrenheit by

  $$T\,(\mathrm{K}) = T\,(\mathrm{C}) + 273.15 = \frac{5}{9}\left[T\,(\mathrm{F}) + 459.67\right]. \tag{1.18}$$

  So when the temperature goes up by 1 K, it also goes up by 1°C and 1.8°F. It helps to remember that room temperature is about 300 K.

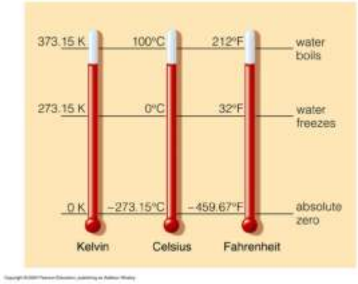


Figure 1.3: Illustration of the Kelvin temperatue scale in comparison to Fahrenheit and Celsius.

- k_B is ***Boltzmann's constant***, $k_B = 1.38 \times 10^{-23}\,\mathrm{J\,K^{-1}}$. This is a very tiny number. However this is the energy *per atom* and there are a *lot* of atoms in any macroscopic bit of matter, so the thermal energy is quite large.

When you put a higher-temperature object in contact with a lower-temperature one, the hotter atoms will transfer some of this microscopic KE to the colder ones. Energy transferred in this way is called ***heat***.

Raising temperature of a substance requires adding energy. The (chemists') calorie is defined to be the amount of heat it takes to raise 1 g of water by $1^\circ\,\mathrm{C} = 1\,\mathrm{K}$. It was not realized until the 1840's that heat was a form of energy. James Joule's demonstration of the equivalence of mechanical energy to heat was one of the key discoveries of ***thermodynamics***, the science of thermal energy and heat. The development of thermodynamics hastened the advent of efficient engines that made the industrial revolution and modern life possible.

So we can't usually see thermal energy, but we can feel it! Once we account for transfers into thermal energy and heat, the conservation of energy is restored.

Power

Power is the *rate* at which energy is consumed or produced:

$$\text{Power} = P = \frac{E}{t} = \frac{\text{Energy}}{\text{time}}, \tag{1.19}$$

$$E = P \times t = \text{Power} \times \text{time}. \tag{1.20}$$

The metric unit for power is the ***Watt (W).***

$$1W = \frac{1\,\mathrm{J}}{1\,\mathrm{s}} = 1\,\mathrm{J\,s^{-1}}. \tag{1.21}$$

It is **critical** that you distinguish energy from power. You will see this mistake made almost every time you read an article about energy. In our analogy of energy to money, power is like your salary, which is the rate at which you accumulate money. Energy is more like your wealth, which is the total amount of money to your name at any given moment.

Continuing with our car example: how much power does it take to go from 0 to 67 mph in 10 seconds? We need to add 0.45 MJ to the car's KE in 10 seconds, so

$$P = \frac{E}{t} = \frac{4.5 \times 10^5\,\mathrm{J}}{10\,\mathrm{s}} = 4.5 \times 10^4\,\mathrm{J\,s^{-1}} = 4.5 \times 10^4\,\mathrm{W} = \underline{\underline{45\,\mathrm{kW}}}. \tag{1.22}$$

If you buy a car in the US, the maximum power of its engine will be given in *horsepower (hp)*. You might guess (correctly) that 1 hp is about the average power that a horse will be able to generate. The conversion is $1\,\text{hp} = 746\,\text{W} \approx 0.75\,\text{kW}$. Our car's engine must be at least 60 hp to accelerate to highway speed in 10 seconds. You would not want to buy this car with an engine whose maximum power is below 60 hp.

Another example: a typical adult human diet contains 2500 Calories per day. What power does it take to run a human?

$$P = \frac{E}{t} = \frac{2500\,\text{Cal}}{1\,\text{day}} \times \frac{4186\,\text{J}}{1\,\text{Cal}} \frac{1\,\text{day}}{24\,\text{hr}} \frac{1\,\text{hr}}{3600\,\text{s}} = 121\,\text{J s}^{-1} = \underline{\underline{121\,\text{W}}}. \quad (1.23)$$

We can't expect a human to give more than $\approx 100\,\text{W}$ of power to any work task for an extended period without giving them substantially more food. Conservation of energy will not allow it. In fact we will see momentarily that the useful power output of a human is substantially smaller than their food input.[5]

When you look at your electricity bill you will see that you are being charged by the kilowatt-hour (kWh). This is a unit of *energy*, not *power*, which makes sense, because the utility should bill you by how much energy you used, not by how quickly you used it. Converting a kWh into Joules is straightforward; just unpack the components of its name:

$$1\,\text{kWh} = 10^3 \times 1\,\text{W} \times 1\,\text{hr} = 10^3 \times \frac{1\,\text{J}}{1\,\text{s}} \times 3600\,\text{s} = 3.6 \times 10^6\,\text{J}. \quad (1.24)$$

A kWh is the energy that you use if you run a 1000 W appliance for one hour. Hairdryers, space heaters, and microwave ovens are roughly 1 kW.

Efficiency

When we build something to convert one type of energy to another, we usually have a purpose. Unfortunately we usually do not convert all of the input energy to the desired form of output energy that serves this purpose. For instance a typical automobile engine only converts about 20% of the gasoline's chemical energy into kinetic energy of the car! Generally we define the ***efficiency*** of an energy conversion to be

$$\text{Efficiency} = \epsilon = \frac{\text{Useful energy out}}{\text{Total energy in}}. \quad (1.25)$$

Energy processes vary wildly in their efficiencies, and we'll learn why. A electric motor converts electric energy into kinetic energy, and can do so with $> 90\%$ efficiency. But an incandescent light bulb's job is to convert electric energy into light, and typically only a few percent of the input energy actually turns into visible light. Clearly,

[5] This is why *The Matrix* doesn't make sense. Maybe not the only reason.

increased efficiency means that you require less energy input, which is a very good way to reduce energy needs and harmful side effects like climate change.

Let's pull everything together for one more example: how much fat would I burn off my body by climbing 20 flights of stairs?

I gain gravitational PE as I climb, and this amount of energy must come from metabolizing fat stored in my body. To get the PE, we'll need my mass—let's guess 80 kg—and the height I climb—let's say 1 flight of stairs is 3 meters. We'll need to know that 1 gram of fat stores 9 Calories of chemical energy. So we might guess that

$$PE_{\text{grav}} = PE_{\text{fat}} \tag{1.26}$$

$$mgh = \text{mass of fat} \times \frac{9\,\text{Cal}}{1\,\text{g fat}} \tag{1.27}$$

$$\Rightarrow \text{mass of fat} = mgh \times \frac{1\,\text{g fat}}{9\,\text{Cal}} \tag{1.28}$$

$$= 80\,\text{kg} \times 9.8\,\text{m s}^{-2} \times 20 \times 3\,\text{m}\frac{1\,\text{g fat}}{9\,\text{Cal}}\frac{1\,\text{Cal}}{4186\,\text{J}} \tag{1.29}$$

$$= \underline{\underline{1.25\,\text{g fat}}}. \tag{1.30}$$

That is a depressingly small amount of weight lost for that exercise! The good news is that we have messed up by forgetting that the human body is *not* 100% efficient at turning food energy into the work that our muscles do. We're about 25% efficient, meaning that

$$\epsilon = \frac{\text{Useful energy out}}{\text{Total energy in}} = \frac{PE_{\text{grav}}}{PE_{\text{fat}}} = 0.25 \tag{1.31}$$

$$\Rightarrow PE_{\text{fat}} = 4 \times PE_{\text{grav}} \tag{1.32}$$

$$\Rightarrow \text{mass of fat} = \underline{\underline{5\,\text{g fat.}}} \tag{1.33}$$

That's 45 Cal burned. For reference, a Wendy's Baconator® is 940 Cal. It takes a lot of exercise to burn a Baconator's worth of calories.

Figure 1.4: 4 MJ of energy. Copyright Wendy's.

Practice Problems:

- Use scientific notation to calculate the US national debt per US citizen.
- When you fill your gas tank, how much are you paying per GJ of energy? What about when you buy electricity at a price of 0.12 per kWh?
- How much energy does a speeding bullet have? How high in the air do you think a bullet goes if you fire it straight up? (look up the mass and speed of your favorite bullet)

- Make an educated guess at the world record pole-vaulting height, knowing that pole vaulters can sprint 100 meters in 10 seconds.

2

Overview of US and world energy usage

How much energy does it take to run US or world civilization each year? We need to know these numbers if we want to evaluate the ability of a given energy source to provide for our needs. We will spend this week examining statistics that address these key questions:

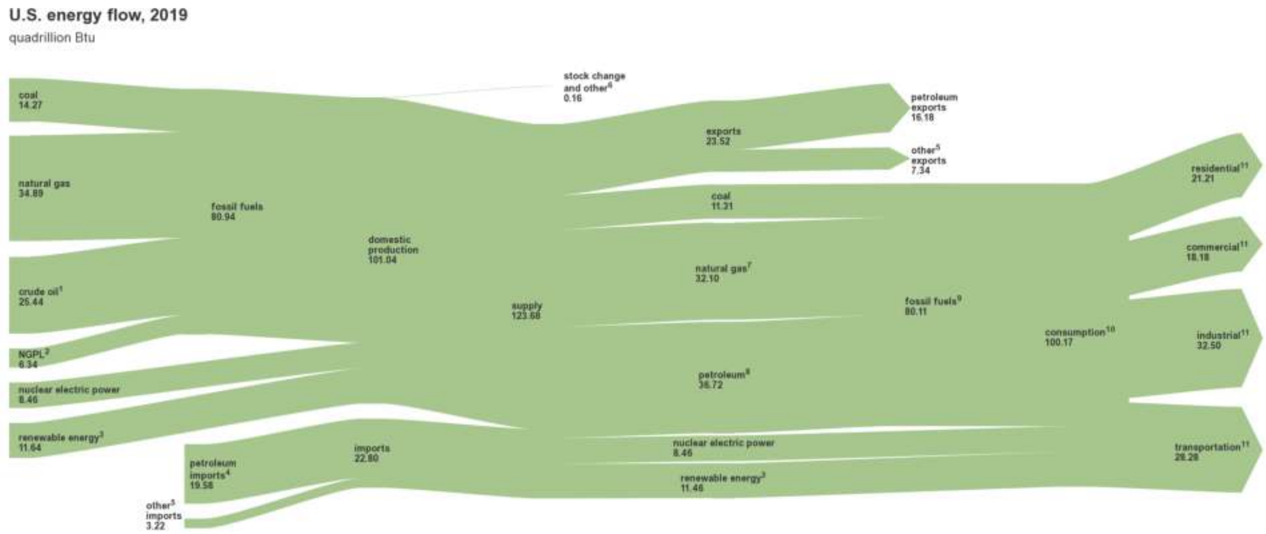
- How much *primary* energy do we use total, and per person?[1]
- What sources do we get the energy from? (coal, gas, nuclear, solar, etc.)
- What are the end users and uses of energy?
- How is energy usage changing with time?
- Which nations are the largest producers of different energy sources, and hold the largest reserves?[2]
- For which energy sources is the US dependent on imports?
- How much CO_2 do we emit, and from which fuel sources and which countries?

Gathering the information to answer these questions is a huge amount of work. We will make use primarily of these compilations:

- The *Monthly Energy Review (MER)* published by the US Department of Energy's Energy Information Agency (EIA). An excellent compilation of up-to-date information on US energy production and consumption. Very little information on the rest of the world though.
- *Key World Energy Statistics* from the International Energy Agency (IEA). Useful statistics per country, but tends to be 1–2 years behind.

[1] ***Primary*** energy means that we count it the first time it comes into human control. For example we would not count electricity generated by burning coal; we count the coal.

[2] ***Reserves*** of a resource are quantities not yet exploited but known with good confidence to exist and be recoverable at current prices.



[1] Includes lease condensate. | [2] Natural gas plant liquids. | [3] Conventional hydroelectric power, biomass, geothermal, solar, and wind. | [4] Crude oil and petroleum products. Includes imports into the Strategic Petroleum Reserve. | [5] Natural gas, coal, coal coke, biomass, and electricity. | [6] Adjustments, losses, and unaccounted for. | [7] Natural gas only; excludes supplemental gaseous fuels. | [8] Petroleum products supplied. | [9] Includes -0.02 quadrillion Btu of coal coke net imports. | [10] Includes 0.13 quadrillion Btu of electricity net imports. | [11] Total energy consumption, which is the sum of primary energy consumption, electricity retail sales, and electrical system energy losses. Losses are allocated to the end-use sectors in proportion to each sector's share of total electricity retail sales. See Note 1, "Electrical System Energy Losses," at the end of U.S. Energy Information Administration (EIA), *Monthly Energy Review* (April 2020), Section 2. See Note 2, "Other Energy Losses," at the end of U.S. Energy Information Administration (EIA), *Monthly Energy Review* (April 2020), Section 2. | Notes: • Data are preliminary. • Values are derived from source data prior to rounding for publication. • Totals may not equal sum of components due to independent rounding.

Sources: EIA, *Monthly Energy Review* (April 2020), Tables 1.1, 1.2, 1.3, 1.4a, 1.4b, 1.4c, and 2.1.

Figure 2.1: 2019 US energy flow, in quads, from the EIA

- *Statistical Review of World Energy* by the BP. Good global statistics, particularly on fossil fuel transfers, but keep in mind this is from an oil company.

- Very nice diagrams of energy flow from Lawrence Livermore National Lab (also part of the DOE), available at `https://flowcharts.llnl.gov`.

These publications favor different units of energy. I will convert everything to **exajoules (EJ)**. Recall that 1 quad is 1.055 EJ.

Total energy use

Figure 2.1 from the EIA summarizes the production and consumption of energy in the US. At center right we see a total US *consumption* for this year of 100.1 quads or 106 EJ. We also see that domestic *production* is 107 EJ, so the US is (for the first time in many decades) a net energy *exporter*. Oddly though, we both import and export a significant amount of petroleum, about 40% of our total usage. Remember this:

> **Total US energy consumption is about 100 EJ per year, roughly equal to production.**

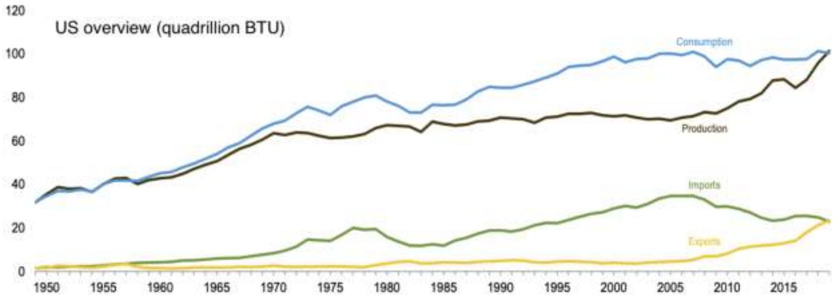


Figure 2.2: History of US primary energy consumption and production since 1950, from Dec 2020 MER.

What this means is that *an energy source (or savings) needs to supply about 10 EJ per year to make a significant impact to the US energy budget.*

Figure 2.2 shows the history of US energy consumption. Here we see that *US energy* ***use*** *climbed continually (since the 1800's) until leveling off over the past 20 years.* There was downward step in the early 1980's after the oil shocks from OPEC boycotts, but growth soon resumed. We also see that *US* ***production*** *levelled off c. 1970, so the US has imported an increasing fraction of its energy supply since the 1970's, until a recent increase in production beginning c. 2007 that has brought net energy imports near zero.*

Let's calculate the ***per capita power consumption*** in the US, knowing that the 2019 population of the US is about 330 million people:

$$\frac{\text{power}}{\text{people}} = \frac{E/t}{3.3\times10^8} = \frac{1.06\times10^{20}\,\text{J}}{3.3\times10^8\times1\,\text{yr}}\,\frac{1\,\text{yr}}{365\,\text{day}}\,\frac{1\,\text{day}}{24\,\text{hr}}\,\frac{1\,\text{hr}}{3600\,\text{s}} \tag{2.1}$$

$$= 1.02\times10^4\,\text{W} = \underline{\underline{10.2\,\text{kW}}}. \tag{2.2}$$

It takes an average of 10 kW of continuous power to run the life of an average US resident! We spend about $1.2 trillion in 2018 to obtain this energy, about 6% of GDP.[3]

World usage

The BP Review gives a 2019 **total world energy usage of 584 EJ.** Figure 2.3 shows that, unlike US usage, world usage continues to climb, by about 1.6% per year for the past decade, and that this increase is almost entirely outside the developed (OECD) countries.

Indeed the US was surpassed in 2010 by China as the world's largest energy energy consumer. Table 2.1 summarizes the energy consumption for the US, China, and the world, according to BP. These are good things to remember:

World energy consumption in 2019 was about 600 EJ.

The USA has just 4.3% of the world population but uses 16% of the world's energy.

[3] See MER Table 1.7

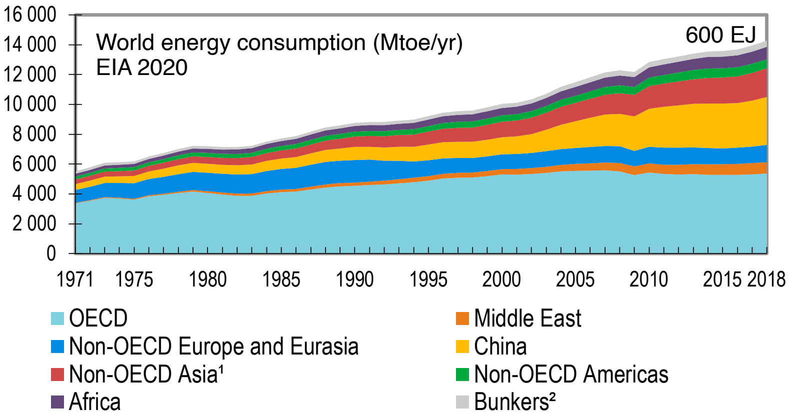


Figure 2.3: World energy consumption, divided into regions. From IEA Summary, 2020.

Country	Consumption (EJ)	Population (billion)	Power per Capita (kW/person)
United States	95	0.33	9.10
China	142	1.40	3.13
World	584	7.7	2.40

Table 2.1: 2019 energy consumption. Energy usage is from BP review, populations are from US Census bureau.

Energy consumption has stabilised in the US and developed countries in the past decade but there is rapid growth in developing countries, and China is now the most energy-consuming country, 50% higher than the US in second place.

The average American uses $3\times$ as much energy as the average Chinese resident and $3.8\times$ as much energy as the average world resident. Among large nations, only Canada has per-capita consumption as high as the US.

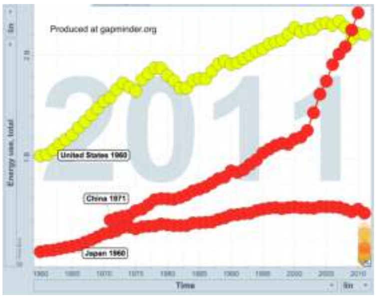


Figure 2.4: China's energy usage passed the US in 2010 and is growing quickly. From the excellent (but outdated) site `http://www.gapminder.org`. China is above 3.3 billion toe as of 2019!

Energy sources

Figure 2.5 plots the primary sources of US energy since WWII. The distribution of energy sources worldwide is similar, except that the dramatic downturn in coal usage and upturn in natural gas usage since 2007 in the US has been much more muted worldwide–global coal consumption has been fairly level for the past 8 years.

The breakdown for the US energy supply in 2019 is:

- 80% from fossil fuels, including
  - 11% from coal;
  - 32% from natural gas;
  - 37% from petroleum.
- 8.5% from nuclear power.
- 2.5% from hydropower.

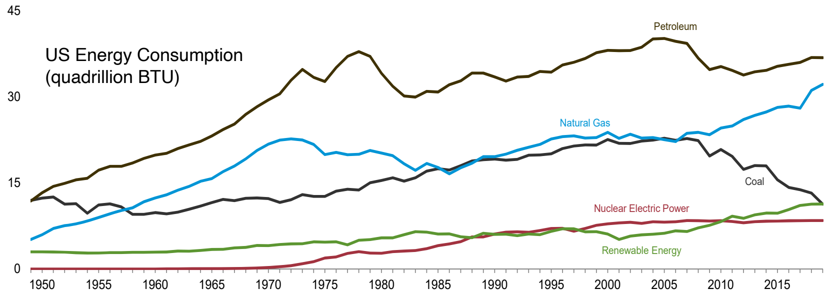


Figure 2.5: History of US energy sources, from the Dec 2020 EIA MER.

- 8.8% from other renewable sources, as detailed in Figure 2.6.

In 2019, US energy consumption from renewable sources (including hydro) surpassed coal for the first time since 1885! But note that a mere 3.6% of the US energy supply came from the potentially sustainable renewable sources of solar and wind power in 2019.

So the key fact on energy sources is:

80% of US energy comes from fossil fuels, with oil the largest fraction (> $\frac{1}{3}$ of all consumption), and coal the smallest and shrinking.

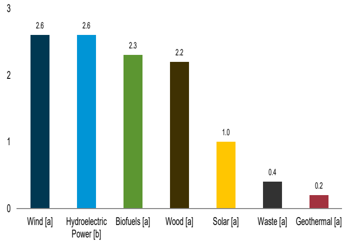


Figure 2.6: Energy from US renewable sources in 2019, in quads, from Dec 2020 MER.

Energy usage

Energy usage is roughly evenly divided into four *sectors*:

Figure 2.7: The usage sectors are at the right of Figure 2.1

- The *residential* sector used 21% of primary energy in 2019. This includes energy consumed in private residences. Recall that "primary" energy includes all the energy needed at the power plants to make the electricity used in the residence.

- The *commercial* sector used 18% of energy in 2019. This is stores and businesses.

- The *industrial* sector used 32% of energy in 2019. This includes factories, mines, refineries, farms, etc.

- The *transportation* sector used 28% of energy in 2019. This is cars, trucks, trains, airplanes, boats—anything that moves!

The distribution of energy use among the sectors worldwide is similar to the pattern in the US.

We can dig a little deeper and ask exactly what purpose the energy is being used for in these sectors. Transportation is pretty obvious, but what about the others?

In the residential sector, we find more than half of the energy used goes to *heating things up and cooling them down.* In fact even more than the pie chart shows, since energy for cooking is under "Other." The remainder of the energy use is mostly lighting and other electrical appliances (TV's, computers, . . .).

In the commercial sector, even more energy goes to heating and cooling (e.g. cooking) than in residences, 69% according to EIA data from 2016.

In the industrial sector, most of the energy goes again to heating things up, and some more to cooling them down. Figure 2.9 is an diagram of energy use in this sector from the EIA. [Actually manufacturing only, not all industry. But close enough.] A smaller chunk of energy goes into electricity to run machinery and electro-chemical processes.

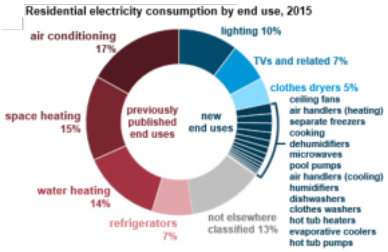


Figure 2.8: From 2015 data in the EIA *Residential Energy Consumption Survey*.

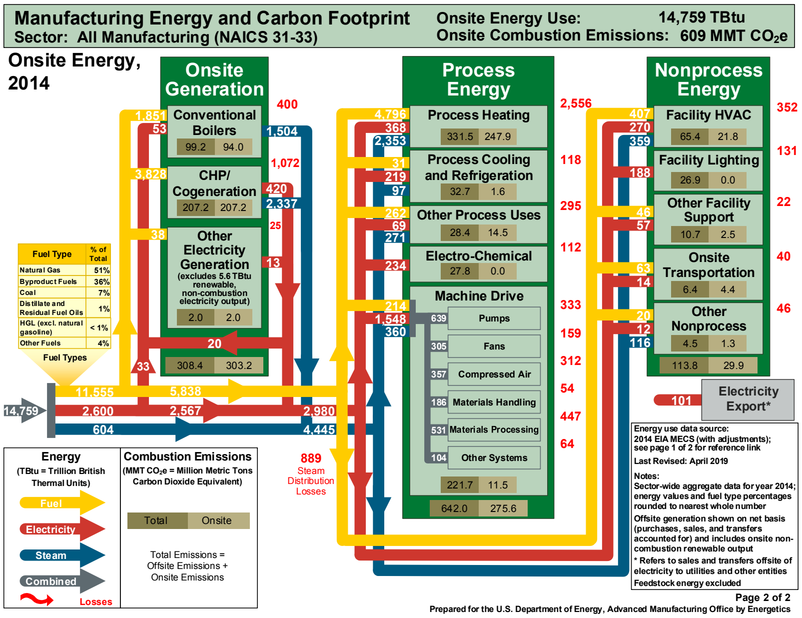


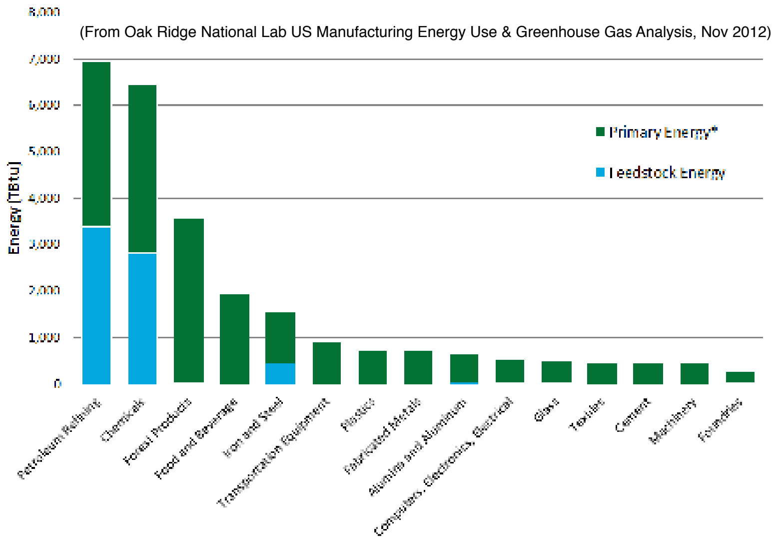
Figure 2.9: Manufacturing energy uses, from EIA 2014 Manufacturing Energy Consumption Survey.

We can also ask which industries use the most energy. This graph shows that (among manufacturing industries again), it is petroleum refining and chemical industries that use the most energy, followed by "Forest Products" (including paper) and "Food and Beverage."

> **The biggest uses of primary energy are running transportation engines, heating things up, and cooling them down.**

What fuels are used for which purposes?

Figure 2.11 beautifully illustrates the flow of energy through US society. Several important things should be noted.


(From Oak Ridge National Lab US Manufacturing Energy Use & Greenhouse Gas Analysis, Nov 2012)

*Primary energy use has been adjusted to exclude the energy value of byproduct fuels derived from feedstock energy sources (e.g., waste gas from LPG feedstock in the chemicals sector); this exclusion avoids double counting feedstock energy. In the case of the petroleum refining sector, the feedstock energy includes energy feedstock used for the production of non-energy products (such as asphalt); it does not include energy feedstock that is converted to other energy products. Therefore, for petroleum refining only, there is no adjustment to the primary energy use for byproducts derived from feedstock energy sources.

Fig. 2.1-2. U.S. manufacturing total primary and feedstock energy

Figure 2.10: Distribution of the 28 EJ of energy used in US manufacturing in 2006.

- Coal is used almost entirely for generation of electricity, which is then distributed to the commercial, residential, and industrial sectors.
- The transportation sector runs almost entirely on oil, and this is the dominant use of oil.
- With the exception of biomass, all of the non-fossil energy sources produce electricity nearly exclusively.

These facts have critical implications. First, if our sustainability worries are greatest for oil, then we *must* talk about transforming the transportation fleet. We can build all the wind power we want, but unless we find a way to use it for transportation, we will continue to need just as much oil.

Second, we must always keep in mind that electricity is *not a primary source of energy.* Making something electric does not make it clean or sustainable unless the electricity is being made from a clean or sustainable primary energy source. At present the majority of electricity is made with fossil fuels.

Third: as we will see later, coal is the most abundant fossil fuel, and also will be the dominant dose of CO_2 into the atmosphere if we keep relying on fossil fuels. If we want to keep that CO_2 out of the atmosphere we must concentrate on electric power plants.

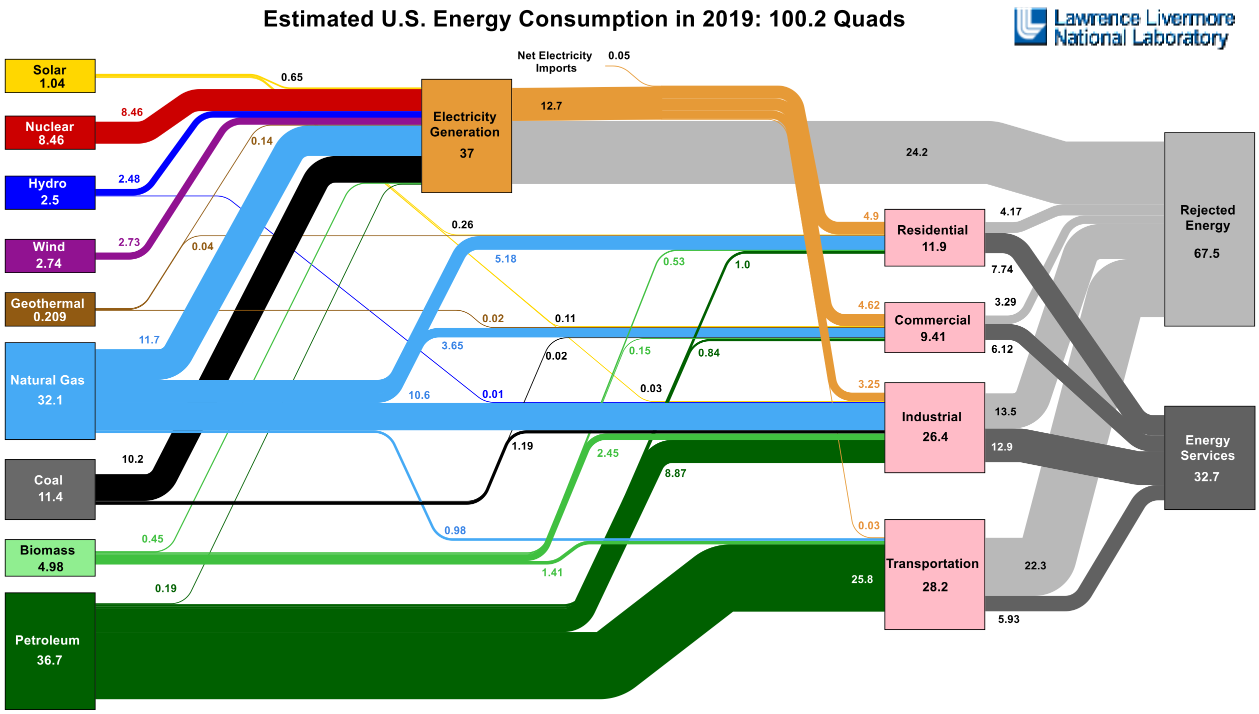


Figure 2.11: LLNL diagram of 2019 US Energy use. On the left are sources of energy, the pink boxes are the sectors where they are used. Keep in mind that generation of electricity is only 1/3 efficient; 25 Quads goes to "Rejected Energy," and the sectors are not being "billed" for that in this graph. So each orange input of electricity should really be tripled for an accounting of primary energy use.

The supplies of fossil fuels

The fossil fuels have finite supplies. It took millions of years to make them and Earth will not be making new batches any time soon. Table 2.2 gives the best rough estimates that I can find for the US and world total economically-recoverable reserves of the three fossil fuels. Reserve estimates are notoriously unreliable, so we will discuss this in greater detail in the chapter on fossil fuels, but I include this summary here to help give the big picture.

> **Domestic reserves of oil and natural gas are only $\approx$10 years' worth of current production rates, even with the recent increases in reserves and production due to the advent of hydrofracturing.**
>
> **Globally, the oil and NG supplies are better at 40–50 years.**
>
> **Most of the energy in fossil fuel reserves is in coal, particularly in the US. At current production rates, the lifetimes are > 100 years.**

Keep in mind: the 100-year+ coal lifetime is for current usage rates. If the whole world increases its power consumption to US per-capita levels, then even without population growth this energy supply drops to 25 years! And if we are relying on coal for *all* our energy needs instead of just the current fraction, then this coal supply would be only **10 years' worth!**

	United States			World		
Fuel	Reserves	Production (2018)	Lifetime (years)	Reserves	Production (2018)	Lifetime (years)
Coal	251 GT 6300 EJ 151 GTC	0.61 GT 15 EJ 0.36 GTC	365	1055 GT 26,400 EJ 637 GTC	6.56 GT 164 EJ 3.9 GTC	132
Natural Gas	420 tcf 450 EJ 6.3 GTC	29 tcf 31 EJ 0.44 GTC	14.3	6950 tcf 7,500 EJ 104 GTC	138 tcf 149 EJ 2.07 GTC	51
Petroleum +NG Liquids	61 Gbbl 373 EJ 7.2 GTC	5.6 (7.5) Gbbl 34 (46) EJ 0.66 (0.88) GTC	11 (8)	1730 Gbbl 10,600 EJ 205 GTC	35 Gbbl 214 EJ 41 GTC	50

Table 2.2: Fossil fuel reserves and production rates for 2018, from *BP Statistical Review of World Energy 2019*. For each fuel, we give values in terms of quantities (tons, cubic feet, and barrels), energy content (EJ), and carbon content (GTC). For US petroleum, values based on annual *consumption* rather than production are in parentheses. For petroleum, values include tar sands.

Also keep in mind that new resources will be discovered. But it seems clear that it is foolish to assume that fossil fuel resources will suffice to supply growing world demand through your lifetime.

Who has it?

Figures 2.12 show which countries/regions hold the oil, natural gas, and coal reserves. You will not be surprised to see Saudi Arabia, Iran, Iraq, and Kuwait near the top of the oil reserves list, but what about Venezuela and Canada in first and third places? This is because the the data include the oil available from ***tar sands,*** an unconventional source of oil that is currently only marginally economically competitive with normal oil drilling methods or hydrofracturing. We will discuss this more in Chapter 7.

Interestingly, the US, Russia, and Saudi Arabia produced roughly 6.2, 4.2, and 4.2 billion barrels of oil in 2019, respectively, even though their reserves are vastly different: 69, 107, and 298 billion barrels, respectively (all these numbers from BP). Clearly these countries have different incentives for getting the oil out of the ground today vs the future!

The natural gas story is similar in that we see the US, Russia, and the Middle Eastern nations high on the list of reserves, but the US was the largest *producer* of natural gas in 2018 (by a lot) even though its *reserves* are well below the top 4 countries. Thus the US is pulling its resources out of the ground much more rapidly than any other major producer.

The coal chart shows that the US is the "Saudi Arabia of coal"! Note that the largest-population countries, India and China, appear on the coal chart but not on the oil or gas charts, so most of the people of the world would need to use coal if they did not wish to

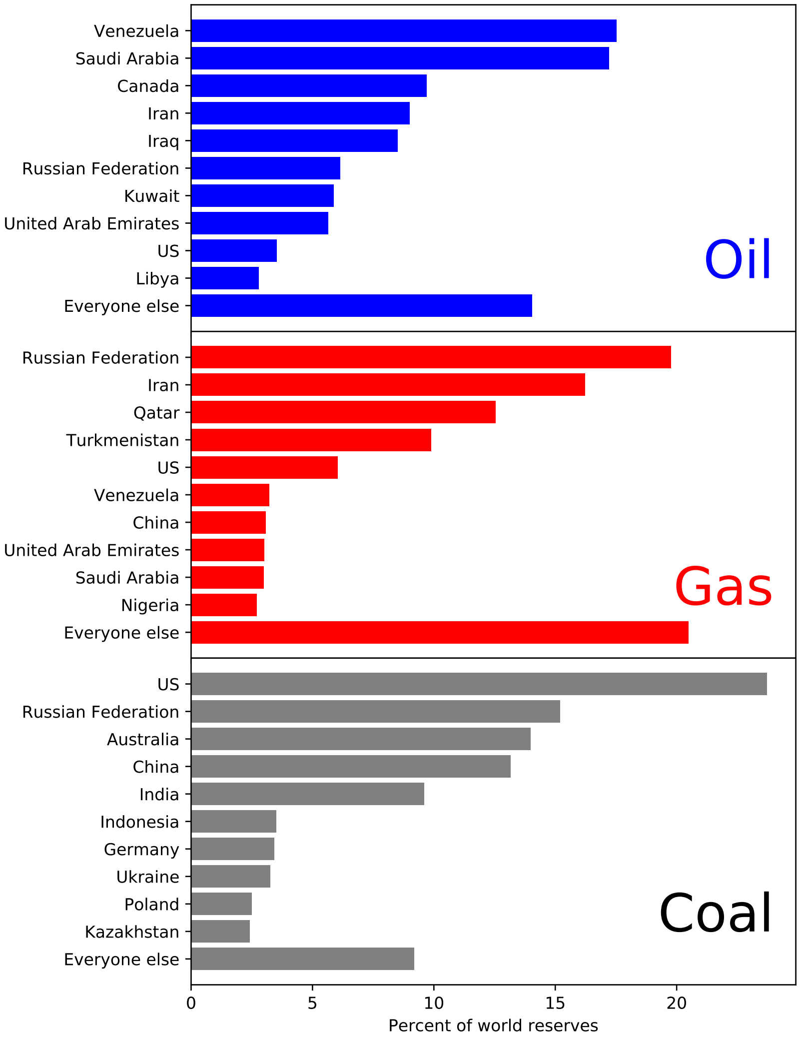


Figure 2.12: The top 10 countries in terms of reserves of oil, gas, and coal. Plotted by GMB from data in 2019 BP review.

import all of their fossil fuels. This, plus the overall abundance of coal, make it very likely that these coal resources will be exploited.

Finally, we ask: for which sources of energy is the US currently self-sufficient, and for which does it rely on exports? We can see in Figure 2.13 that crude oil dominates US energy imports. Only in the past few years have imports been less than half of US oil consumption. Right now the US imports about 20 EJ of crude oil about 17 EJ (which seems a little crazy, but see Chapter 7 leaving our net energy imports near zero. Natural gas is basically self-sufficient, and we are net exporters of a small amount of coal. So if you are worried about US energy independence, your concern is oil, and you are less concerned than you would have been 10 years ago.

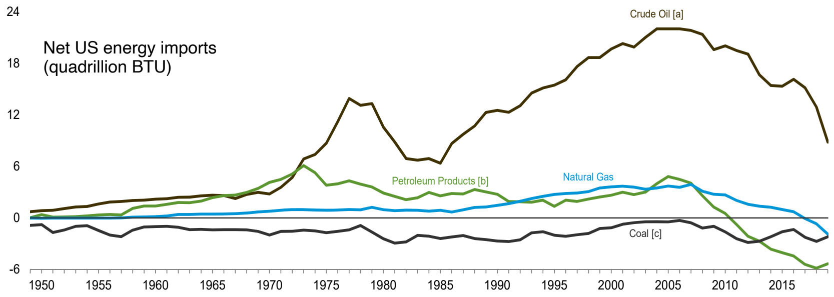


Figure 2.13: Net energy imports to the US through 2019, from the Dec 2020 MER.

CO_2 *emission*

Energy consumption generates an astounding quantity of CO_2. In 2019, the EIA reports US CO_2 emissions (from energy use) to total 5.2 *billion* tons. That is approximately 15,800 kg of CO_2 per person per year, or roughly *your weight in* CO_2 *every 1–2 days!* The good news is that the US CO_2 emission rate has actually declined 15% from its 2007 peak. Figure 2.14 shows that CO_2 emission for the *developed* world has been steady for the past decade, but the *total* world CO_2 emission continues to climb rapidly due to increased energy use in China and other developing nations. World CO_2 emissions are growing at about 1.1% per year for 2008–2018, so while still increasing, the growth rate is slower than it has been in any non-recession period in modern times. As CO_2 emission follows fossil-fuel use, it remains true that US residents emit several times more CO_2 per year than the average world citizen.

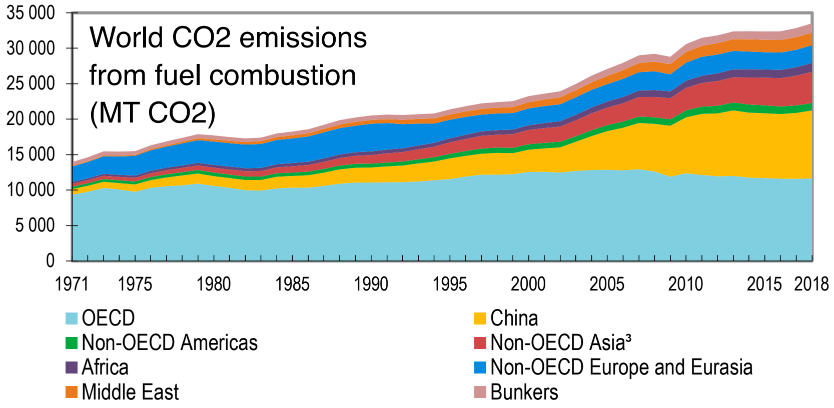


Figure 2.14: Annual world CO_2 emissions, divided into regions. From the 2020 IEA statistical summary

.

Half of recent increases in global CO_2 emission are attributable to increased coal-burning in China. Figure 2.15 shows the rapidly rising emissions from China, which passed the US in 2006 to become the most global-warming-friendly nation. Note that heavier reliance on coal in China than in the US meant that it became the world CO_2-emission leader a few years before it became the world energy-consumption leader. Chinese CO_2 emissions have grown an average of 2.6% per year over 2009–2019.

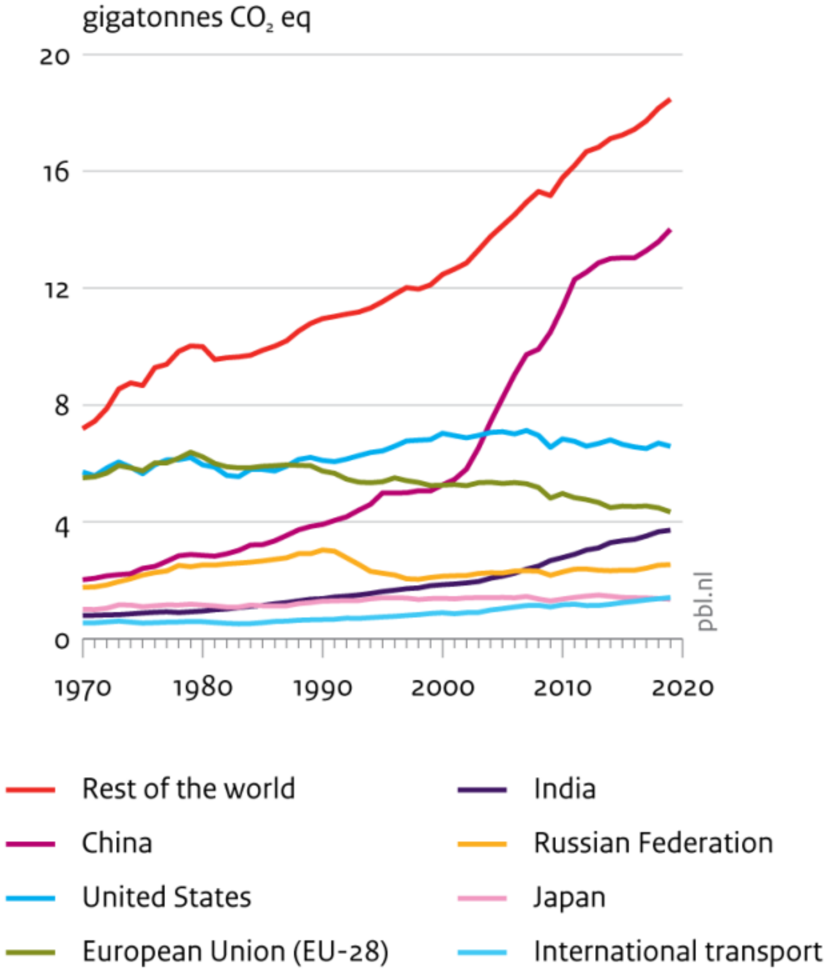


Figure 2.15: CO_2 emission history through 2019 for the largest-emitting countries plus the EU. From 2020 *Trends in Global CO_2 Emissions* report from the PBL Netherlands Environmental Assessment Agency.

Within the US, we have another excellent figure from LLNL showing the flow of carbon through our economy. Compare Figure 2.16 to 2.11 showing the flow of energy. The obvious differences are that (a) the renewables and nuclear do not contribute to carbon flow, and (b) coal has a bigger, and natural gas a lesser, impact on CO_2 than oil per Joule by nature of its chemical composition.

Transportation and coal power plants account for roughly 60% of US carbon emissions. The fraction is higher in more coal-reliant countries such as China.

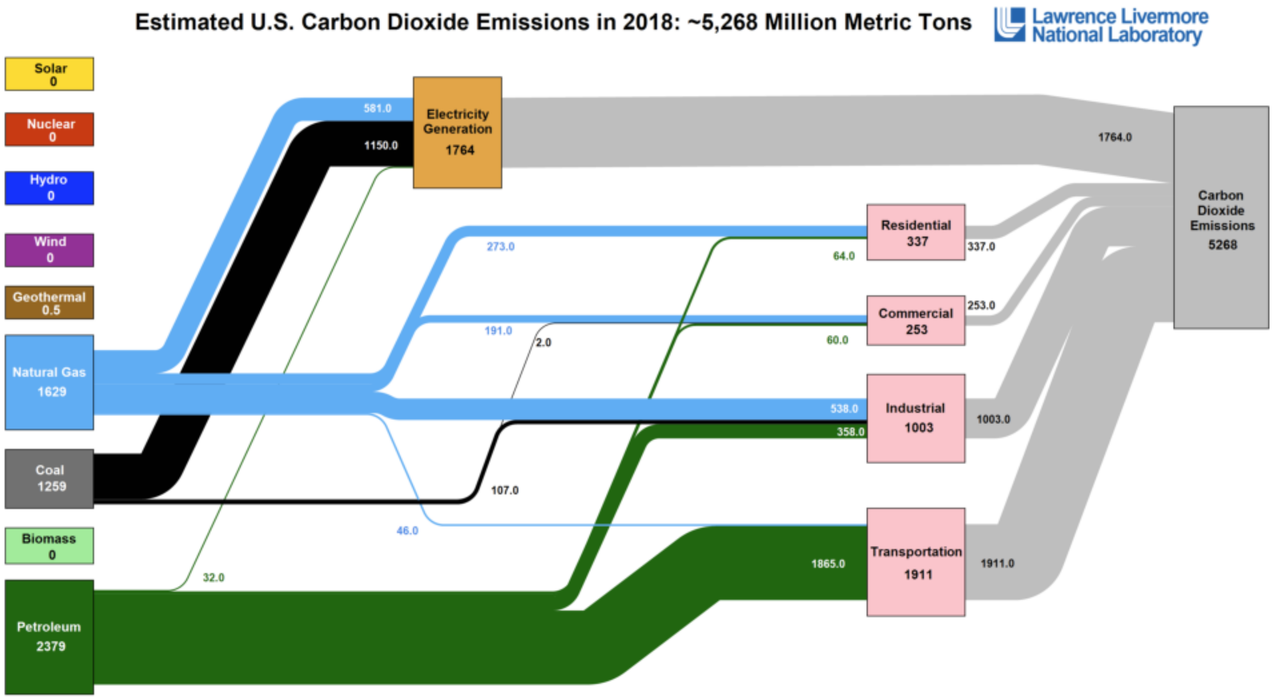


Figure 2.16: LLNL diagram of 2018 US carbon flows. Keep in mind that the carbon emissions during electricity generation are not being assigned to the end-users of the electricity in the pink boxes.

We close this chapter by showing the mean global surface temperature of Earth over the past 140 years. The distinct rise in this diagram is attributable to human CO_2 emission at very high confidence. This of course is another reason why the current energy supply situation is not sustainable, especially if other nations approach the US energy consumption rate. We can see that 2020 is tied for the warmest year of the 140 years with with accurate records. **Every one** of the past 20 years has been warmer than **all but one** of the preceding 120 years. There is no serious doubt that this is the result of fossil-fuel burning, as we will examine in detail in Chapter 6.

Practice Problems

- If a new energy source is to provide $> 10\%$ of US energy needs, how many Watts of power must it produce, on average, per square meter of the continental US? The lower 48 states have a combined area of about 8 million km^2.
- The average sunshine falling on each square meter of the continental US is about 180 W. Is solar energy a potential candidate for provide 10% (or all) of US energy?
- How long will *all* of known fossil-fuel reserves be able to meet cur-

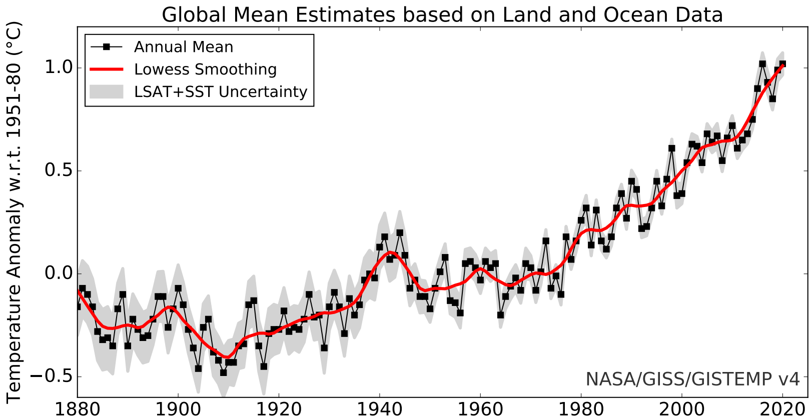


Figure 2.17: Global mean temperatures through 2020, from estimates. From NASA GISS

rent global energy demand? What if all of the world's population were to immediately raise their power consumption to US levels?

- Explain a critical advantage that biofuels might have over all the other renewable energy sources in the quest to replace all fossil-fuel usage.
- A debate question: there is no reason that the US should try to reduce its carbon output since China is now the largest (and growing) CO_2 emitter.

3

Transfer of Heat and Money

HEATING AND COOLING account for most non-transportation energy use, so our first priority should be to understand what determines how much energy we need to use for these purposes. We will see that it is straightforward to drastically reduce the energy needs for space heating and air conditioning. You can even build a house that requires *no heating system at all*, even in a cold climate. But we choose to continue using more energy than is needed. To see why, we will make a short detour into finance and study the tradeoff between spending money now (*e.g.* on more house insulation) vs later (*e.g.* on energy for heat).

Specific and latent heat

The amount of thermal energy stored within a material is determined by its temperature (and also other variables, like the pressure, which we will encounter later). If you want to raise the temperature of some mass m of material by an amount ΔT, then the we need to add an amount of energy

$$Q_{\text{in}} = m\,c\,\Delta T, \tag{3.1}$$

where c is the ***specific heat*** of the material. Liquid water has an unusually high specific heat:[1]

$$c_{\text{water}} = 4186\,\text{J}\,\text{K}^{-1}\,\text{kg}^{-1}. \tag{3.2}$$

How much energy does it take to make a hot bath? Let's guess that you are filling a big tub that is 2 m long by 0.4 m wide with water 0.3 m deep, and you want the bath temperature to be 40° C. The water comes into the house at room temperature, 20° C, so $\Delta T = +20^\circ\,\text{C} = 20\,\text{K}$.[2] To get the mass of our water, we have to remember that the ***density*** of water is $\rho = m/V = 1000\,\text{kg}\,\text{m}^{-3}$.[3] So our energy

---

[1] Recall that this is 1000 calories, or 1 kcal, or 1 Calorie, because the calorie was initially defined by the heat needed to raise the temperature 1 g of water by 1 K.

[2] Recall that a *change* of 1° C is also a *change* of 1 K. Also note there is no degree sign used with K.

[3] The kg was originally defined as the mass of 1 liter of water.

need will be

$$\begin{aligned} Q_{\text{in}} &= mc\,\Delta T = \rho V c \Delta T && (3.3)\\ &= \frac{1000\,\text{kg}}{1\,\text{m}^3}\,(2\,\text{m} \times 0.4\,\text{m} \times 0.3\,\text{m})\,\frac{4186\,\text{J}}{\text{K kg}}\,20\,\text{K} && (3.4)\\ &= 2.01 \times 10^7\,\text{J} = \underline{\underline{20.1\,\text{MJ}}}. && (3.5) \end{aligned}$$

You need to burn about 1/2 gallon of gasoline to heat this water, using many times more energy than the kinetic energy of a speeding car! We see that there is a lot of energy stored in thermal energy of everyday materials. Thermal energy can be a good way to store energy, or to remove it. Water makes a very good cooling system for machines, since it conveniently has a very high specific heat and can flow through pipes.

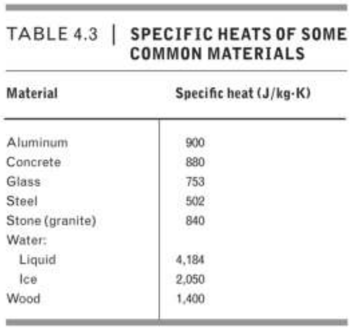


TABLE 4.3 | SPECIFIC HEATS OF SOME COMMON MATERIALS

Material	Specific heat (J/kg·K)
Aluminum	900
Concrete	880
Glass	753
Steel	502
Stone (granite)	840
Water:	
Liquid	4,184
Ice	2,050
Wood	1,400

Figure 3.1: Wolfson's Table 4.3

Latent heat

Like most substances, water undergoes ***phase changes***: it can be a liquid, or solid, or gas. The amount of thermal energy is different for each phase even at the same temperature, as the molecules are in different arrangements. To melt or boil water, you therefore need to add energy, without a temperature change. The energy needed to change the phase of a material is called the ***latent heat.*** To melt water, the latent heat is $334\,\text{kJ kg}^{-1}$. To boil water, the latent heat is an enormous $2.26\,\text{MJ kg}^{-1}$. Note it takes 500× more energy to boil water than to raise its temperature by 1 K! So a good way to relocate energy is to use it to boil water, then transport the steam somewhere. Letting the steam cool and condense will release the energy. Most electric power plants use the fuel to generate steam, then the steam is used to make electricity. This is also why being exposed to steam causes very severe burns (or scalds), because so much heat is released into your skin when the steam condenses to water in contact with your skin.

This is how UPenn heats its buildings: a central boiler makes steam, which travels through pipes to all the buildings. Condensing steam heats the air in the buildings. You can see steam vents in the sidewalks around campus.

Higher and lower heating value

When I look up the amount of energy released by burning 1 kg of methane, I sometimes find 55.5 MJ, sometimes 50.0 MJ. The former is sometimes called the ***higher heating value (HHV)*** and the latter the ***lower heating value (LHV).*** What's the difference? When we burn CH_4 we get CO_2 and H_2O out, at high temperature. The HHV assumes that you extract all the energy from these products as they cool back to the original temperature, and this is the true answer to the question "how much chemical energy was released when we burned the methane?" The LHV assumes that the steam produced is lost, and you do not retrieve the latent heat that was in this water. I will try to stick with the HHV, but be on the lookout for calculations that use (or should use) the LHV.

Heat transfer

Specific and latent heat can be used to estimate the energy needed in industrial processes that require *e.g.* melting some metal, or heating ingredients up to some high temperature. But we still have to wonder: why do we need to buy energy to heat our house? After all, we like to keep the interior of our house at a constant temperature, so its thermal energy should never change, and we should never need to add or remove energy.

The reason is because thermal energy *will naturally flow from hotter to colder materials.* The transferred energy is then called ***heat,*** so for instance *heat transfer* will occur from our warm indoors to the cold winter outdoors. So we must keep running (the heater) just to stay in place (at the same indoor temperature).

In summer the situation is reversed: heat is leaking *into* the house, and we will need a way to continually *remove* energy from the indoors to keep the house from warming up. For both heating and cooling, it is clear that we want to reduce the heat transfer across the walls of the house to address the biggest residential uses of energy.

Conduction

The transfer of thermal energy between two places can occur by three modes. ***Conduction*** occurs when hotter atoms or molecules collide with their colder neighbors and thereby speed them up a bit. The heat moves *through* a material without the material itself moving. An example would be sticking one end of a metal rod into a fire and feeling the other end warm up. Another example is the heat leaking out of your house through a solid wall.

The rate of heat conduction across a slab of material is

$$P = \frac{Q}{t} = \frac{kA\,(T_h - T_c)}{d} = \frac{kA\Delta T}{d} \tag{3.6}$$

$$= \frac{A\Delta T}{R}, \tag{3.7}$$

$$R \equiv \frac{d}{k}. \tag{3.8}$$

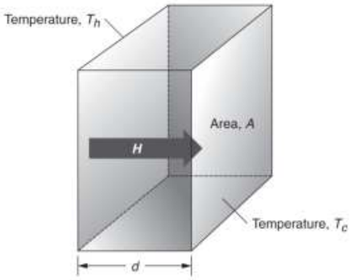


Figure 3.2: Heat power H flowing across a slab of material. Figure 4.3 from Wolfson.

- Q is the amount of energy in heat form transported in time t, giving a power P.
- A is the *area* of the surface facing the hot/cold sides.

- d is the *thickness* of the slab.
- $\Delta T = T_h - T_c$ is the temperature drop from hot to cold sides.
- k is the ***thermal conductivity*** which is a characteristic of the material making up the slab.
- $R = d/k$ is called the ***R-value*** of the slab, and gives the total effectiveness of the slab in blocking heat.

We see sensible things: the bigger the area of the slab (*e.g.* a wall), the more power conducts across it. The bigger the temperature difference, the more heat travels across it. And if we reverse the temperatures, the heat flow will reverse also. The thicker the slab, the *less* heat conducts.

Material	Thermal Conductivity ($\text{W m}^{-1}\text{ K}^{-1}$)
Air	0.026
Styrofoam	0.029
Fiberglass	0.042
Wood	0.11
Water	0.61
Glass	0.8
Concrete	1.0
Granite	3.4
Steel	46
Aluminum	240

Table 3.1: Thermal conductivities of some common materials. Values from Wolfson Table 4.1

Table 3.1 lists the thermal conductivities of some materials. We notice that metals have *high* conductivity and should be used when you want to carry a *lot* of heat (such as the bottom of a frying pan), but are very bad choices for housing insulation! Stone is less conductive, and wood even less, but clearly Styrofoam and fiberglass with lower k are much better choices for slowing heat transfer. But the lowest conductivity of all in the table is air—which suggests that we should be making our walls out of air to save energy, which we know is a very wrong idea.

Convection is the transfer of heat by the flow of warm material to colder locations. A central heating system is an example of "forced" convection, whereby the heat in the furnace warms up air nearby, then we blow this warm air to cool parts of the house using a fan. Similarly we cool engines by flowing water around them. "Natural" convection happens too: in a pot of water on the stove, the fire heats the water at the bottom. For air and (usually) water, the material expands when heated, and will become less dense and float to the top. Colder material will flow down to take its place. The heat is therefore transferred through the whole pot by convection of water. Wind is just a form of convection: the air is moving because some parts of the atmosphere are being warmed more than others.

When convection occurs, it is typically much more effective at heat transfer than conduction. Of course this is why we do not have walls built of air: a slight breeze, or other convection will replace our nice warm inside air with cold outside air.

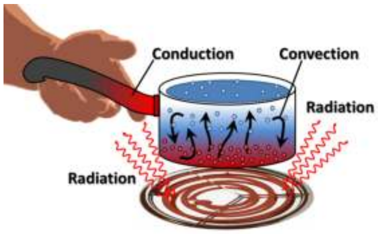


Figure 3.3: Heat transfer mechanisms - old Microsoft clip art.

How can we stop the evil of convection? We have to stop the ability of material to flow. One way is to make the walls out of a solid instead of water or air, and of course we need some solid walls to hold the roof up. The best insulators, like styrofoam, are made mostly of air, but the air is trapped in a plastic foams cells to prevent it from flowing. Fiberglass works the same way. Hair and feathers exist to trap air and insulate warm-blooded animals.

Radiation is the transfer of energy from hot to cold surfaces by exchange of light (= electromagnetic energy). Since light travels across empty space, this is the only way to transfer heat across a vacuum, and is for instance how heat gets from the surface of the Sun to Earth. It is also the only way for Earth to *lose* the energy it gets from the Sun: by radiating heat to empty space, which is very cold (3 K!). Radiation is the key to understanding the greenhouse effect and global warming, and we will cover it in a later chapter. It is usually a smaller contribution to a building's heat losses than conduction and convection. But we can use radiation to our advantage in a good house design, for instance by admitting the Sun's light in winter time to warm the house.

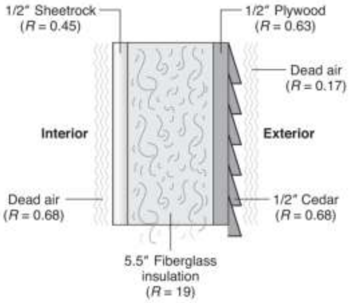


Figure 3.4: A wall typically has layers of different material. Add up the R-values for all the layers to get the total R for the wall. This wall would have an R-value of ≈ 22, nearly all coming from the fiberglass layer. Figure 4-6 Wolfson.

Calculating heat transfer

We want to calculate the *total* power being lost to the outdoors in winter because this tells us how quickly we need to get new energy to replace it and stabilize the house temperature. Assume ΔT difference between indoor and outdoor temperatures.

First consider a single wall that has area A_1 and an R-value R_1. Recall from Equation (3.8) that $R = d/k$ measures the ability of a wall to *block* heat transfer: higher R is a better insulator. This equation will tell us the rate of heat loss across the wall. We get high R by a combination of thickness (which costs and weighs more) and choosing a good insulation material. There are a few things to note about R-values:

- What if the wall has many layers of different material? We add up the heat-blocking ability of each layer, so the total R is the sum from each layer.

- While Equation (3.7) came from the conduction formula, most references will bundle the conduction, convection, and radiation transfer together into one R that you can use in this equation.

- Metric units of R-values should be $\mathrm{m^2\,K\,W^{-1}}$. If you look at a product listing it will usually leave the units off and just say something like "R-11." In the US, R-1 means

$$R = 1^\circ\,\mathrm{F\,ft^2\,hr\,BTU^{-1}} = 0.175\,\mathrm{m^2\,K\,W^{-1}}. \tag{3.9}$$

  So if you take the published R-value and multiply it by 0.175, you'll have it in metric units (of course the formula works if you use other units, as long as you make them all match properly before you're done).

Here is an example: how much heat leaks out a $3\,\text{m} \times 5\,\text{m}$ wall that is insulated to R-10 when it's 65° F inside and 47° F outside?

First we recognize that $\Delta T = 18^\circ\,\text{F} = 10\,\text{K}$. The heat loss is then

$$P = \frac{A\Delta T}{R} = \frac{3\,\text{m} \times 5\,\text{m} \times 10\,\text{K}}{10 \times 0.175\,\text{m}^2\,\text{K}\,\text{W}^{-1}} \tag{3.10}$$

$$= \underline{\underline{86\,\text{W}}}. \tag{3.11}$$

What would the heat loss be if we increased the insulation to R-20? What if we lowered the indoor thermostat to 62° F?

To get the total heat loss for a building, we need to analyze all the surfaces of our residence that divide indoors from outdoors—appropriate roofs and floors as well as exterior walls. Now divide these surfaces up according to their R-values. For instance we might have area $A_1 = 40\,\text{m}^2$ that is $R_1 = 10$, and an area of $A_2 = 10\,\text{m}^2$ of windows that are $R_2 = 4$, etc. The total heat loss will be

$$P_{\text{tot}} = \frac{A_1\Delta T}{R_1} + \frac{A_2\Delta T}{R_2} + \ldots = \Delta T \sum_i \frac{A_i}{R_i}.. \tag{3.12}$$

The building as a whole loses heat like a single slab whose A/R is equal to the sum of the A_i/R_i of each segment i of the building envelope. This equation can also be used for calculating the heat that an air conditioner will have to remove from a house. Or, when applied to the walls of your refrigerator, the heat that it will need to continuously remove from its interior.

Heating and cooling costs are also increased by ***infiltration,*** which is the leakage of outside air into the building. This is a form of convective heat transfer. There are spaces around doors, window frames, foundations, etc., where air can leak. If we know the rate at which air leaks in, we can use the specific heat of air to calculate the extra heat. Filling these gaps with sealant and weather stripping can reduce this heat transfer substantially and cost-effectively.

Annual heating costs

What is the total annual cost of heating your home? Is it cost-effective to build a house with better insulation and lower infiltration? To upgrade a house's insulation?

First let's calculate the total annual heat loss from the house. In a time interval Δt_i when the outside temperature is ΔT_i colder than inside, the heat lost will be

$$Q_i = P_i \Delta t_i = (\Delta T_i \Delta t_i)\,\frac{A}{R}. \tag{3.13}$$

For example it is common to take $\Delta t_i = 1\,\text{day}$ and get ΔT_i by subtracting the thermostat setting from the average outdoor temperature for the day.

Now we want to know the total heat lost over a year, Q_{tot}. To do this we just add up the heat from each day $i = 1, 2, 3 \ldots$, remembering to skip the days when its warmer outside than inside, since we won't need any heat for those days (we will put these days into the cooling calculation later).

$$Q_{\text{tot}} = \frac{A}{R}\left(\Delta T_1 \Delta t_1 + \Delta T_2 \Delta t_2 + \ldots\right) \tag{3.14}$$

$$= \frac{A}{R}\sum_i \Delta T_i \Delta t_i = \frac{A}{R}\left(\Delta T\,\Delta t\right)_{\text{tot}}. \tag{3.15}$$

The A/R part of the equation describes the building; the $(\Delta T\,\Delta t)_{\text{tot}}$ part describes the weather at the location. Weather statistics for the US are compiled by the National Centers for Environmental Information. Figure 3.5 shows this value for locations in the US. The $(\Delta T\,\Delta t)_{\text{tot}}$ values are usually given in ***degree-days***, in this case the units being $^\circ$ F day. For example: if the mean outside temperature were 0° C $=$ 32° F on every day during January, and we had our thermostat set for 20° C $= 68^\circ$ F, then during this month we would accumulate a total of $31 \times 20 = 620$ degree-days in K or 1116° F-days. The value for Philadelphia is about 5500° F-days $= 3050$ K-days per year. If I moved my house to Anchorage, Alaska, my heat consumption would double!

The annual *cost* of the heat you'll buy is

$$\text{cost} = Q_{\text{tot}} \frac{Q_{\text{purchased}}}{Q_{\text{tot}}} \frac{\text{cost}}{Q_{\text{purchased}}} \tag{3.16}$$

$$= \frac{A}{R}\left(\Delta T\,\Delta t\right)_{\text{tot}} \frac{\text{cost of energy}}{\epsilon} \tag{3.17}$$

- Q_{tot} is the heat that leaked out of the house, which is of course the amount you'll need to put back into the house to maintain constant temperatures. We can use Equation (3.15) to get it.
- ϵ is the efficiency of the furnace, which is the fraction of the total fuel energy input $Q_{\text{purchased}}$ that is actually delivered to the house.
- The *cost of energy* is the amount you pay for your heating fuel per unit energy that it supplies.

A furnace has $<$ 100% efficiency because some of the heat from burning the fuel goes out the flue instead of into the house. A decent gas or oil furnace will have $\epsilon > 0.9$.

Table 3.17 gives the average US retail cost of the most commonly used residential heat sources as of April 2018. An important thing to notice is that *natural gas is currently far less expensive per J than oil, and electricity costs the most per unit energy.* Natural gas prices are very low right now, making it the cheapest residential energy source.

Now we have all the ingredients to estimate the annual costs of heating. Consider a single R-2 window that is 32×36 inches or $A = 0.75\,\text{m}^2$ installed in a Philadelphia house with gas heating. What are

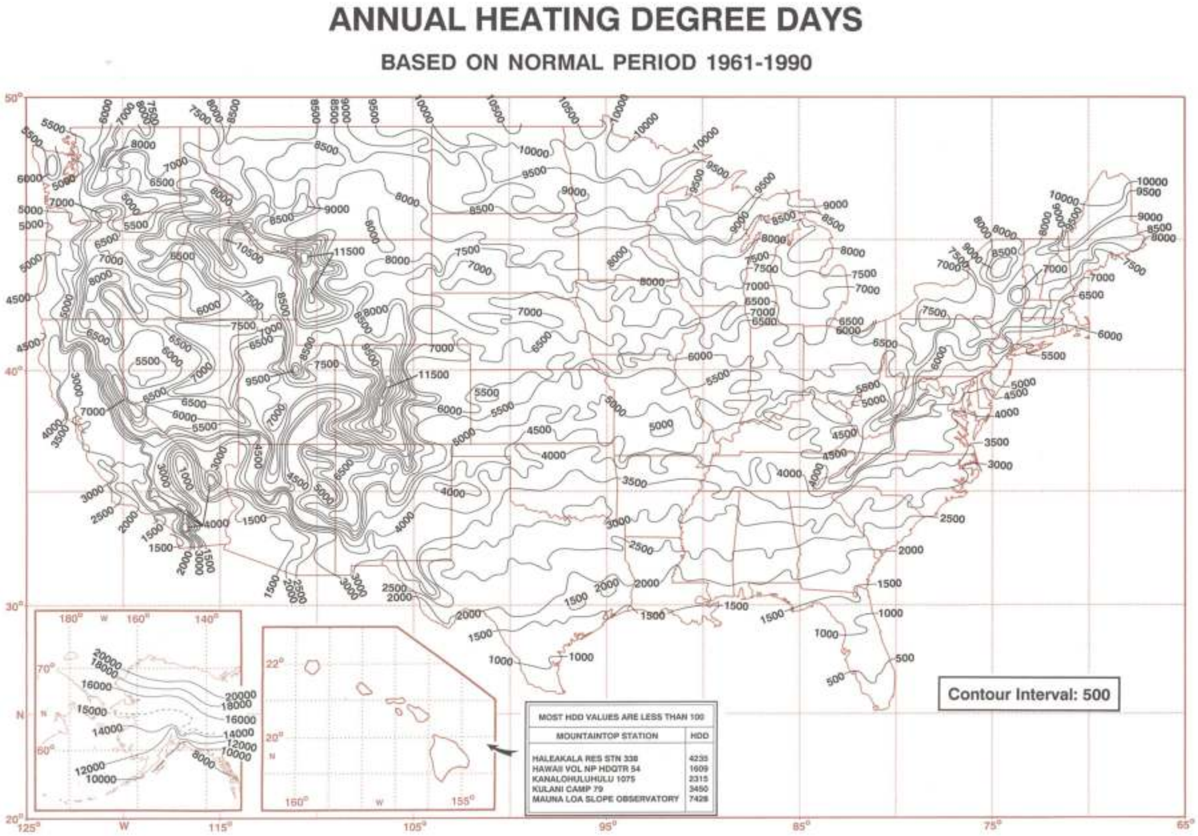


Figure 3.5: Average annual heating degree-days in the US, from the National Climate Data Center via Wikipedia. [The NCDC site no longer is capable of producing this map, unfortunately.]. Values are given in ° F-days, and assume an indoor temperature of 65° F.

Table 3.2: Average costs of residential heating fuels for 2019, from the Dec 2020 EIA MER.

Fuel	Cost	Energy content	Cost of Energy ($/GJ)
Natural Gas	$10.51/kcf	1.10 GJ/kcf	9.55
Fuel Oil	$2.27/gal	0.146 GJ/gal	15.55
Electricity	$0.130/kWh	3.6 MJ/kWh	36.11

we paying per year for the heat that leaks out this window? Let's assume our furnace has efficiency of 90%.

$$\text{cost per year} = \frac{A}{R}\,(\Delta T\,\Delta t)_{\text{tot}}\,\frac{\text{cost of energy}}{\epsilon} \tag{3.18}$$

$$= \frac{0.75\,\mathrm{m}^2}{2\times 0.175\,\mathrm{m}^2\,\mathrm{K\,W}^{-1}}\left(\frac{3050\,\mathrm{K\,day}}{\mathrm{yr}}\right)\frac{\$9.55\,\mathrm{GJ}^{-1}}{0.9}\,\frac{24\,\mathrm{hr}}{1\,\mathrm{day}}\,\frac{3600\,\mathrm{s}}{1\,\mathrm{hr}} \tag{3.19}$$

$$= \underline{\underline{\frac{\$5.99}{\mathrm{yr}}}}. \tag{3.20}$$

If we were in an oil-heated house, the cost would be about 1.6× higher, ≈ \$9.75/yr.

Now we're given an opportunity to buy a better-insulated R-4 window instead of R-2. This would cut our heat loss, and hence our energy bill, in half, saving us \$3.00 per year, or \$4.88/yr in the oil-heated house. How much extra would you be willing to pay for this better window? The window will probably be in the house for 40 years, for a total savings of about \$120, but maybe you want to make your up-front money back in just 10 years, so you would only be willing to spend an extra \$30.[4] Maybe if you were willing to pay money to reduce your CO_2 emission and future global warming, the better window would be worth more to you or to society as a whole, but right now you pay no direct cost for contributing to climate change.

Future money vs today money

Buying a higher-R window is a choice to spend money now in exchange for paying *less* money later. In other words you are trading future dollars for present-day dollars. Most people are very foolish and will choose a short-term gain over a bigger long-term gain. Let's examine the tradeoff between ***present value (PV)*** and ***future value (FV)*** of money.

Simple interest

A bank will give you \$1 today if you promise to pay $\$(1+rT)$ a time T from now, where r is the ***interest rate.*** Conversely you might invest \$1 today and expect to get $\$(1+rT)$ back later, where r would now be the rate of return on your investment. Normally we expect $r > 0$, that is to say, a dollar today is worth more than a dollar in the future. We would write

$$FV = (1+rT)PV. \tag{3.21}$$

[4] We are neglecting here the additional money you would save on air conditioning. More on that later.

Note that r is a *rate* of interest, *e.g.* we will typically assume something like 5% *per year* or $r = 0.05\,\text{yr}^{-1}$.

Compound interest

Usually the interest on an investment is ***compounded***. A $1 investment compounded annually at 5% per year is worth $1.05 after one year, $\$(1.05)^2$ after two years, $\$(1.05)^3$ after three years, etc. The more general formula for interest compounded every T_0 interval, leading to $N = T/T_0$ interest payments over the length of the loan, is

$$FV = (1 + rT_0)^{T/T_0} = (1 + rT/N)^N \,. \tag{3.22}$$

Compounding	Value
None (simple)	$2.00
Annual	$2.65
Monthly	$2.712
Daily	$2.7181
Continuous	$2.7183

Table 3.3: Value of a $1 investment at 5%/yr interest rate after 20 years, for different rates of compounding.

Compounding yields a bigger return than simple interest, especially as the interest rate and the time of the investment grow, because you are earning interest on your interest. Compounding each month is better than each year; what happens if we compound every day, or every hour or millisecond? It doesn't keep getting more and more lucrative; calculus veterans will recognize that with infinitely frequent compounding, $N \to \infty$ in Equation 3.22 becomes

$$FV = e^{rT} PV; \qquad PV = e^{-rT} FV. \tag{3.23}$$

We will use this equation of ***continuous compounding*** to relate future expenditures to up-front ones. You should review the exponential function e^x and its inverse, the natural logarithm $\ln x$. This equation is an example of ***exponential growth*** whereas simple interest is ***linear growth***. In linear growth, you *add* a fixed amount to your value every year. In exponential growth, you *multipy* your value by a fixed amount each year. Exponential growth always exceeds linear growth at the same positive growth rate. In fact exponential always *eventually* surpasses linear growth even if the interest rate is lower for the exponential case!

Exponential growth occurs in many important contexts besides finance. A very common one is population growth: if people (or plants, or bacteria) produce, on average, more than one offspring per individual, then the total population will grow exponentially. For instance the world population is currently increasing at about 1.1% per year. If this continues, when will the population double from its current 7.6 billion to 15.2 billion? We can solve for the time T that it will take to double the present value of population:

$$FV = e^{rT} PV = 2 \times PV \tag{3.24}$$

$$\Rightarrow e^{rT} = 2 \tag{3.25}$$

$$\Rightarrow rT = \ln 2 \tag{3.26}$$

$$\Rightarrow T = \frac{\ln 2}{r} = \frac{0.693}{0.011\,\text{yr}^{-1}} = \underline{\underline{63\,\text{yr}}}. \tag{3.27}$$

Naively we would have guessed it would take $100\%/1.1\% = 91$ years for population to increase by 100%. Exponential growth is faster. Of course this assumes that the 1.1% growth rate stays the same. In fact the fertility rate of the world has gone down significantly in the past few decades; it was closer to 2% per year in the mid-1900's and is expected to slow further in the future. The growth rate of Earth's population is an important factor in predictions of energy consumption and CO_2 emission.

Annuities

Equation 3.23 tells us how to exchange a single payment today with a single payment at some time T in the future. More common in our class are choices between a single payment today and a *continuous stream of payments* from today until some time in the future. For example buying extra insulation saves you money on utilities every month until the house is abandoned. Buying a solar panel costs money up front but reduces your electric bill every month until the panel dies. When I bought my house, the bank gave me a big pile of cash up front to pay for the house in exchange for my promise to send them money every month for $T = 30\,\text{yr}$. This mortgage is just one example of an ***annuity,*** a continuous stream of payments. Pensions are another form of annuity.

Let's define the annuity A as the rate at which I will pay (or recieve) money, *e.g.* $A = \$2000\,\text{month}^{-1}$ for a mortgage payment. What would be a fair present value PV to receive today in exchange for paying out the annuity over time T starting today?

Let's say we'll make N payments spread over the time T. Each payment will be an amount AT/N. Payment number i has $FV_i = AT/N$ and is made at time $T_i = iT/N$. According to (3.23), the present value of that payment is

$$PV_i = e^{-rT_i} FV_i = \frac{AT}{N} e^{-irT/N} \tag{3.28}$$

$$\Rightarrow PV = \frac{AT}{N}\left(e^{-rT/N} + e^{-2rT/N} + e^{-3rT/N} + \cdots + e^{-NrT/N}\right). \tag{3.29}$$

The second line is the total present value of the all the payments in the annuity contract. You might recognize this as a geometric series, and there is an easy formula to express the sum. Also we can again assume that payments are frequent $(N \to \infty)$ and get a straightforward formula:

$$PV = \frac{A}{r}\left(1 - e^{-rT}\right); \qquad A = \frac{r\,PV}{1 - e^{-rT}}. \tag{3.30}$$

This is the **annuity equation** and we will use it whenever we want to trade up-front costs against a stream of future savings.

Here's an example: if you want to get \$300,000 to buy a house, and you find a 30-year mortgage that has 4% per year interest, what will your monthly mortgage payment be? We solve for A:

$$A = \frac{r\,PV}{1-e^{-rT}} = \frac{\$3\times10^5\times0.04\,\mathrm{yr}^{-1}}{1-\exp\left(-0.04\,\mathrm{yr}^{-1}\times30\,\mathrm{yr}\right)} \tag{3.31}$$

$$= \frac{\$1.2\times10^4\,\mathrm{yr}^{-1}}{1-e^{-1.2}} = \$17,172\,\mathrm{yr}^{-1}\frac{1\,\mathrm{yr}}{12\,\mathrm{month}} \tag{3.32}$$

$$= \underline{\underline{\$1431\,\mathrm{month}^{-1}}}. \tag{3.33}$$

You can check this against any online mortgage calculator! Notice the depressing fact that you will be making 360 payments for a total of \$515,165 over the loan lifetime, well above the principal of \$300,000 on the loan—a vivid illustration of the fact that money in the future is worth less than money in hand today.

Now let's get back to our window. We found that the better-insulated window would save our natural gas customer \$3.00/yr. If we expect this window to last 40 years before being replaced again, then the fuel savings should be compared to the present value of an annuity of this amount. We need to choose an interest rate that we could expect to receive on a 40-year annuity. Let's guess 5%/yr. Then the present value of the fuel savings is

$$PV = \frac{A}{r}\left(1-e^{-rT}\right) = \frac{\$3.00\,\mathrm{yr}^{-1}}{0.05\,\mathrm{yr}^{-1}}\left(1-e^{-0.05\,\mathrm{yr}^{-1}\cdot40\,\mathrm{yr}}\right) = \underline{\underline{\$51.87}}. \tag{3.34}$$

If the better window costs this much more than the cheap window, it's a break-even proposition. Of course we have assumed that the cost of natural gas will be the same for the next 40 years, and that we could have obtained 5% interest on an alternative investment with this \$50.

There are a few things to notice about this result:

- If we assume the window will last *forever*, $T\to\infty$, then the present value increases only slightly to \$60.00. In other words, the standard financial formulae assign very little value to savings that occur many decades in the future. This is one reason why it is hard to make an economic argument for actions that will prevent catastrophic climate change that might be 50–100 years away. Standard finance literally discounts the value of the future generations' problems.

- Our naive guess of the value of the savings would have been $PV = AT = \$3.00\,\mathrm{yr}^{-1}\times40\,\mathrm{yr} = \$120$. This simple "payback time" calculation overestimates the value of the future savings. Your calculus skills might show that this is the answer you would get from assuming $r = 0$ interest rate. It shows that these PV calculations can be quite sensitive to the interest rate choice, if you are dealing with long-term costs and benefits.

- The fuel-oil-using homeowner should be willing to pay about 1.6× as much for the same window upgrade!

Inflation

Inflation is when the price of goods increases over time. We have *not* taken inflation into account in the above formula. Inflation degrades the true value of a dollar in the future, so it *reduces* the PV of our energy-saving window. On the other hand, inflation also means that fuel prices are likely to increase, so our utility savings in the future will be higher than we have guessed, which will *increase* the PV of the good window. In fact these two effects will exactly cancel each other out as long as fuel prices rise at the same rate as general inflation. The one thing we must do is use an *inflation-adjusted* interest rate r_{real} in our annuity formula (sometimes called the *real* interest rate or real rate of return on investment) which would be

$$r_{\text{real}} = r - \text{rate of inflation}. \tag{3.35}$$

Real interest rates are currently near historic lows: inflation for the calendar year 2020 was just 1.4%[5] while the best interest rate on a 2-year consumer certificate of deposit (CD) is about 0.8%, so real interest rates are *negative* (-0.6%) for consumers investing money. (Not surprisingly, the real interest rates for consumers *borrowing* money are substantially higher and positive). This makes all of this annuity calculation seem unimportant. Low real interest rates will not persist, however, through the decades of time we need to consider.

Air conditioning

Everything that we have calculated about the bills for buying energy to heat your home in winter applies equally well for the energy bills that you will need to pay for air conditioning to *remove* heat from your home in the summer. We just need to change the names of a few things.

- Now we use $\Delta T = T_{\text{outside}} - T_{\text{inside}}$ and the heat is flowing *into* the house.
- To calculate the total heat leaking in and the cost of removing it, we use Equation (3.17) but we'll need to put in the ***cooling degree-days*** into $(\Delta T\, \Delta t)_{\text{tot}}$ by getting historical records of the days when

[5] According to the Consumer Price Index (CPI-U) from the US Bureau of Labor Statistics.

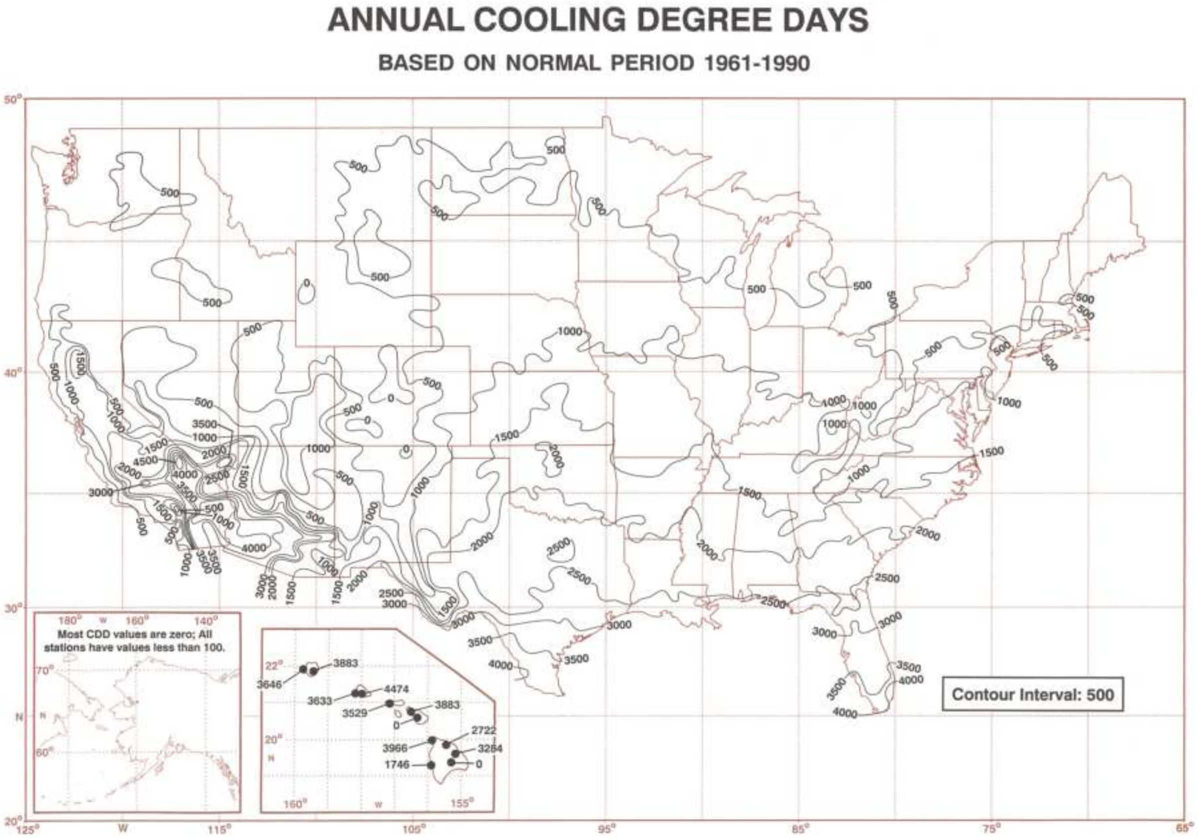


Figure 3.6: Average annual cooling degree-days in the US, from National Climate Data Center. Values are given in ° F-days, and assume an indoor temperature of 65° F.

it is warmer than room temperature outside. Figure 3.6 shows the cooling map of the US. Notice Philadelphia has about 1000° F-days of cooling, well below its 5500 heating degree-days. But in my old home of Tucson, AZ, it's quite the opposite.

- We also don't really want the efficiency of the air conditioner. Instead we want something called the ***coefficient of performance (CoP)*** of the air conditioner to go where ϵ was:

  $$CoP = \frac{\text{heat removed}}{\text{electricity input}}. \tag{3.36}$$

  The best air conditioners these days have $CoP \approx 6$. Does it violate the Law of Conservation of Energy to have $CoP > 1$? No—we are not *producing* more energy than we buy, we are just moving energy from indoors to outdoors. That's a pretty good trick to do, which we will investigate more in the next chapter. Notice that using $CoP = 6$ in place of $\epsilon \approx 0.9$ in our cost equation will lower the amount (and cost) of the energy we need to buy, compared to heating.

- The "fuel" in this case must be electricity, with its higher cost per Joule than natural gas, which raises the cost of A/C compared to heating. Again we'll see why in the next chapter.

Reducing heating bills

Is it possible to substantially reduce the energy needs for heating and cooling buildings? Reduce the CO_2 output from these jobs? Is it cost-effective? From our cost Equation (3.17), many strategies for lower cost and lower emissions are clear:

- Move to a different climate to reduce the heating degree-days $(\Delta T\, \Delta t)_{\text{tot}}$. But be careful not to raise your cooling degree-days too much in the process! More practically, you can reducing your degree-days by changing T_{inside} instead of T_{outside}, namely by lowering your thermostat in winter and raising it in summer. You should estimate the fraction of your heat and A/C usage you can expect to eliminate this way in your climate.

- Reduce A, the area of the exterior-facing walls. This means having a smaller dwelling (per person). Another great efficiency: live in an apartment building, where some of your walls do not face the exterior. This is one of the reasons that city living is *much* more energy-efficient than country living.

- Increase R by adding insulation and eliminating low-R weak spots in your building envelope. This means building in or retrofitting more insulation into the walls, ceilings, and floor of the living spaces, and using less-conductive materials.

- Increase R of the windows too. High-R windows use several tactics: first, they have double- or triple-glazing with air gaps between the glass panes that act as insulation. Filling these gaps with argon gas instead of air lowers the conductivity and convection further. Special coatings on the glass can also reduce the amount of unwanted heat transfer through radiation.
- Reduce infiltration (which is convective heat loss) by sealing up cracks and gaps in the building envelope.
- Improve the efficiency ϵ of the heater or the CoP of the air conditioner. Modern heaters are usually already at $\epsilon > 0.9$ so there isn't much to gain here, but A/C CoP's continue to improve.
- Switch your heating fuel to natural gas if it is available. This will lower your bills, but also lower your CO_2 output compared to oil, since gas emits less CO_2 per Joule. Electric heat is a very poor choice for CO_2 reduction—recall that we burn 3 J of fuel for each J of electricity, so the CO_2 consequences of that J are tripled unless the electricity is coming from nuclear or renewable power plants.
- In summer time, don't generate heat indoors that you'll then need to remove. Less baking, for example.
- Use *passive solar* heating strategies that will admit the radiative heat from the sun in winter when you want it, and avoid it in summer when you do not. One example that works for Northern buildings is to put more windows on the south-facing side of the house, since that is the side of the sky where the sun spends most of its time. Another strategy is to plant deciduous trees near the house which will shade the sun in summer but let the light through in winter. A light-colored roof will reduce the heat absorbed into the house in the summer but have little effect in winter.

It may not surprise you to learn that it is perfectly feasible to build a house that requires no furnace at all, even in Northern US or European climates. This is done primarily by using well-insulated walls and attention to infiltration. In the US, with very low natural gas prices, this is not likely to be cost-effective. My *total* utility bill was only $\approx$ \$2000 for all of 2020. If the 2020 level were to continue to hold, then the annuity formula shows that the present value of an upgrade that would reduce my utility bills to *zero, forever* would be about \$50,000 if I assume 4% real interest rate on my loans or investments. I should not spend more than this for improved efficiency or

A zero-infiltration house is not a good idea: with no outside air coming into the building, the residents would suffocate. A milder issue is that the air will get stale and smell bad, and might accumulate some harmful substances like radon. So we need to introduce outside air; but we can be clever and warm the incoming air with the outgoing air using a "heat exchanger," to have very little net heat loss.

insulation, at least for financial reasons. Some major caveats to this calculation: (1) my utility bills have been low lately because of unusually warm winters and low natural gas prices. They could go up. (2) No amount of insulation will reduce my utility bills to zero, since we'll still need fuel for cooking and electricity to run our appliances. (3) Interest rates fluctuate also.

It is certainly true, however, that the typical US home has many opportunities for improved thermal performance that are cost-effective. Many detailed studies suggest that home heating bills can be reduced by 50% compared to typical construction practices used for older homes. There is a huge difference in the cost of *building* a house to be energy-efficient and the cost of *retrofitting* a house to be energy-efficient. The latter is typically several times more expensive. My house, for example, was built in 1952 and has *no insulation* in the walls! Energy was cheap back then. . . . Unfortunately it would be very expensive to install insulation in my brick and plaster walls, and this would not be cost-effective. On the other hand there were substantial improvements in efficiency and comfort from other changes, such as better attic insulation/ventilation, sealing, etc., that were cost effective after I bought the house.

Houses last many decades, so we are stuck for a long time with every inefficient building that is constructed, and we should really value high efficiency in every new building. The problem is that home *builders* do not have to pay long-term energy bills, so they have incentive to build cheaper, inefficient houses, if they care at all either way. In economics-land, buyers would pay more for houses constructed to use less future energy, but in the real world very few buyers think about costs beyond the purchase price, and even if they thought about energy costs, they would not be able to evaluate the choices made by the builder. This market failure means that buildings use much more energy than they need to. This is a prime candidate for government intervention, in the form of building codes that mandate higher-efficiency construction practices. Codes and practices have improved substantially in the past 30 years. There are *e.g.* also regulations that have raised the minimum CoP that an air conditioner must have to be sold in the US.

In summary: we could eliminate nearly half of residential and commercial energy consumption by greatly boosting the energy efficiency of buildings, although that would be very expensive. But even a partial effort in building heating and cooling could easily cross the threshold of having a $> 10\%$ impact on energy use. This would, however, take concerted willpower, in the form of government building standards, and will take decades to phase in since the housing stock lasts a long time.

Practice problems

- You want to make ice cubes ($T = -5^\circ$ C) from room-temperature water. Which step requires removal of the most energy from the ice: (a) cooling the water down to freezing; (b) freezing the water; or (c) cooling the ice down to the final temperature?
- What is a "condensing furnace" and would you want to have one?
- Why does goose down make a good winter coat?
- See if you can find prices for single vs double-paned windows at Lowe's, and decide which you would choose to build with.
- I spend about $800 per year on natural gas to heat my home in Philadelphia. Estimate how much money I would save by lowering my thermostat setting from 68° F to 63° F in winter.

4

Thermodynamics

When we burn fossil fuels, the chemical energy turns into thermal energy (heat). If heat is what we want (home furnace, hot water heater, cooking), then that's great, we're done. But what if we want to move something (like a car or the shaft of an electric generator)—how do we take random microscopic KE and use it to move something big? A machine that takes thermal energy as input and converts it to some non-thermal form like kinetic energy is called a ***heat engine.*** We use them all the time, in electrical power plants and car engines. But we have seen that these do not work very well—their efficiencies are shockingly low. Why?

How do we *remove* heat from something? In other words how do ***refrigerators*** or air conditioners work? And if they are removing heat from indoors, why should we be feeding them electrical power? Shouldn't they be able to run on the thermal energy they remove?

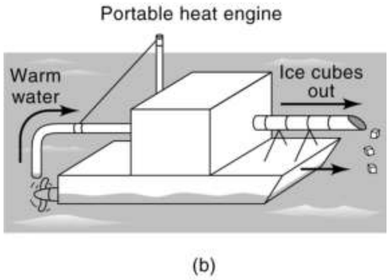


Figure 4.1: A great idea for a boat engine, from Hinrichs Figure 4-21b, ©2006, Thompson Higher Ed

If energy never disappears, why do we need to keep getting more of it? For example if we have kinetic energy in a car, and it gets turned into heat when we hit the brakes, why can't we just recover the heat and use it to get the car moving again? We should never need any more gasoline!

Even in winter there is plenty of thermal energy in the outdoor air. Why can't we just extract some of this energy and add it to the indoor air to keep the indoors warm? Such a device would be a ***heat pump*** and it seems should save us from buying heating fuel. Or run the machine in Figure 4.1 that can take thermal energy out of sea water to run its propellor, ejecting a stream of ice cubes? This would not violate the Conservation of Energy Law but it sure would save fuel.

These questions are the domain of ***thermodynamics***, the study of heat. The basic concepts of thermodynamics were developed in the 1800's, when people first started industrializing with steam engines and these questions became critically important. The answer to all these questions involves ***entropy***, which is the most difficult concept we will encounter from physics in this course. But it's essential if we want to understand why we keep needing to find new energy.

Pressure and engines

How can we turn heat into mechanical forces and motion? We need to have the randomly-moving atoms push on something. A gas consists of rapidly moving atoms. Each time one of them bounces off a wall, it gives the wall a little push. Add all these little pushes together and you get a strong, coherent ***pressure*** against the wall. The pressure of air at sea level is equivalent to the weight of 15 lbs per square inch, or over 10 tons per square meter! Though this pressure does not just push down; it is pushing outward on every surface containing the gas. Air pressure is exerting hundreds or thousands of tons of force against each wall of the room you are in—why doesn't the wall collapse under all this stress? Because there is air on the other side of the wall, pushing back with an exactly equal force. Hence we rarely notice this air pressure.

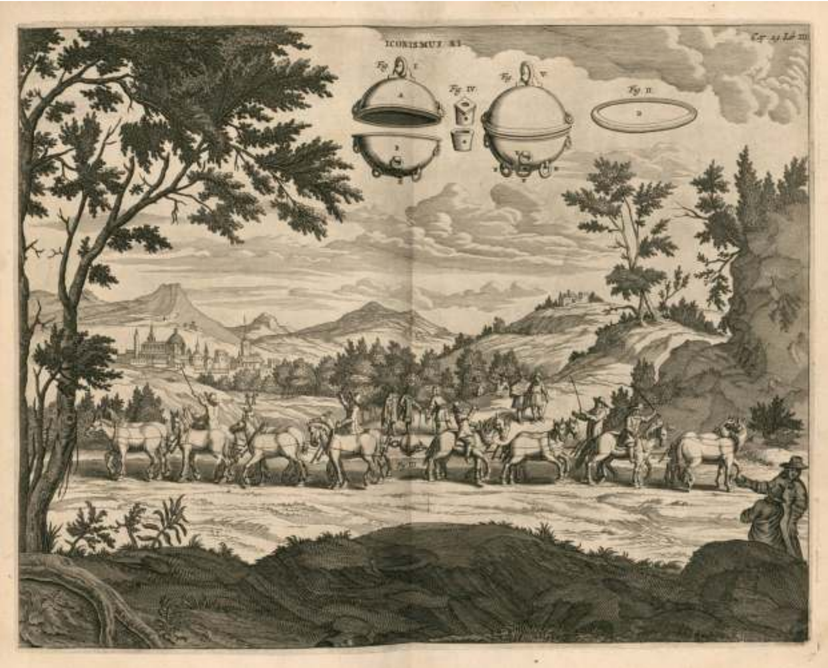


Figure 4.2: Engraving by Caspar Schott (1672) of the Otto von Guericke's famous 1656 "Magdeburg experiment" in which he evacuated the air between two hemispheres. Two teams of horses could not pull the hemispheres apart, vividly demonstrating the strength of air pressure.

Pressure becomes noticeable—and we can use it to move things—if we can create an *imbalance* of pressure, so there is a net push toward the side of a wall with lower pressure. If we arrange for this to occur, the strength of the pressure force can be stunning. Heat engines typically create motion by raising or lowering the pressure on one side of a piston, with normal air pressure on the other side.

The pressure in a gas can be raised by adding heat to the atoms. You may have learned the *ideal gas law* in chemistry or physics class:

$$P = \frac{N}{V} k_B T, \tag{4.1}$$

which says that the pressure will rise when we either squeeze more molecules (N) into less volume (V), or we raise the temperature T. Adding heat to a fixed volume of gas will raise its temperature, hence its pressure. Alternatively, boiling water forces many water molecules into the gas volume, making high-pressure steam.

The first practical heat engine, invented by Thomas Newcomen around 1710, turns heat into motion as follows:[1] Heat is used to make steam, which fills the space in a cylinder behind a rising piston. Then cold water is sprayed into the cylinder, which condenses the steam back into liquid, leaving a near-vacuum behind with very low pressure. Now the air pressure will push the piston back down the cylinder. The back-and-forth motion of the piston can be used to do useful ***work***.

"Work" has a precise technical meaning: it is the amount of non-thermal energy W that we get out of the heat engine. If the engine was given an amount Q_{in} of input heat, then the *efficiency* of a heat engine is

$$\epsilon = \frac{W}{Q_{\text{in}}}. \tag{4.2}$$

If the engine is $<$ 100% efficient, conservation of energy demands that there is some waste energy (heat) Q_{out} emitted by the engine, with

$$Q_{\text{out}} = Q_{\text{in}} - W. \tag{4.3}$$

In an ***internal combustion engine (ICE),*** the fuel is burned *inside* the cylinder, heating up the air in the cylinder and increased pressure. The exhaust gas is released after it has expanded and the piston is at its high point. The thermal energy in the hot exhaust gas is the waste energy.

Newcomen's engine was used to pump water out of coal mines. In fact the only place it was useful was at a coal mine, because it took so much fuel to run it! Newcomen's engine had very low efficiency. Many improvements to the steam engine were made (notably by James Watt) to increase its efficiency and eventually make it feasible for use in factories and railway engines. Vexingly, the efficiency could never be made anywhere close to 100%.

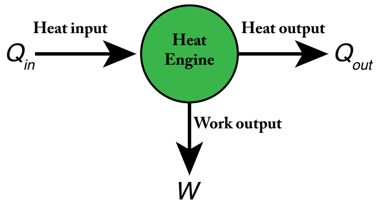


Figure 4.3: Schematic diagram of a heat engine.

---

[1] See GIF movie at `http://en.wikipedia.org/wiki/Newcomen_atmospheric_engine`.

Air conditioners and refrigerators

How can we *remove* thermal energy from air (or our food, or anything else)? Conservation of energy demands that this energy go somewhere. A simple strategy is to have our air give its thermal energy to something else via heat transfer. We simply put our air in contact with something *colder*. For example most of Penn's campus building are air conditioned using a supply of chilled water that circulates through campus. The room air is blown around pipes holding the cold water flow in a *heat exchanger*.[2]

So immediately you will ask: how do you make a supply of something *colder* than the ambient temperature? If you are not lucky enough to have a lake of cold water handy in the summer, you need to build yourself a ***refrigerator***, a device that can lower the temperature of a material somehow.[3] We need to turn some of its thermal energy into non-thermal energy. The most commonly used strategy is to take a gas and let it *expand* into some lower-pressure region. Recall that we used this strategy in our heat engine to extract energy (work) from a gas. Well if we extract work energy, we have necessarily reduced the thermal energy, and we will end up with lower temperature if we play the game well.

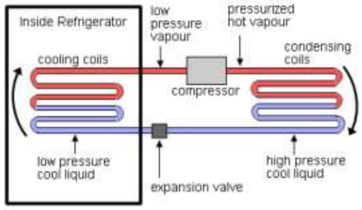


Figure 4.4: Diagram of how a refrigerator works, from this site.

So a feasible air conditioner would work like this:

1. Start with a pressurized container of gas (or liquid)—the *refrigerant*—at our room temperature T_C.
2. Let the gas (or liquid) expand into a pipe. Its temperature will drop.
3. The gas can now absorb some heat Q_C from the room. We now have achieved our goal!
4. But unless we have an infinite pressurized gas supply, we are stuck when the gas runs out. So next we need to re-compress our refrigerant. Compressing the gas will require us to *input* some non-heat energy, which we'll again call the "work" W. This is typically coming from electricity used to run the compressor motor.
5. When we compress the gas, it will heat up again. We let the gas give up some of its heat, an amount Q_H, to air or water that circulates around the pipes. Don't do this indoors, because that would just be putting back *into* the house all the energy that we were trying to *remove* from it. So this radiator should be outdoors, where the temperature T_H is warmer than the indoors.
6. Bring the gas back inside and start again at (1).

---

[2] Note that a car's radiator is a heat exchanger in the opposite direction: its job is to transfer heat from hot water circulating in the engine to cooler air that is flowing by it as you drive.

[3] I'll use the terms air conditioner and refrigerator interchangably here since they work the same way, they just cool different places.

The air conditioner usually keeps the refrigerant in a closed loop, with the net effect that heat Q_C is absorbed from the indoor air at temperature T_C, heat Q_H is emitted into the outdoor air at temperature T_H, and an amount W of non-heat energy is taken in (usually as electricity) to run the compressor. The Law of Conservation of Energy demands that

$$Q_H = Q_C + W \tag{4.4}$$

meaning that *more* heat is ejected to the warm spot than is removed from the cold spot. The *coefficient of performance (CoP)* of the refrigerator is defined as

$$CoP = \frac{Q_C}{W}. \tag{4.5}$$

A higher CoP is better because it means we are getting more cooling done for less electric energy we need to buy.

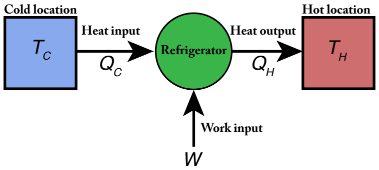


Figure 4.5: Schematic diagram of a refrigerator, air conditioner, or heat pump.

Refrigerators sold in the US must be labelled with their CoP, although this is given a different name—***Seasonal Energy Efficiency Ratio (SEER)***—and this value comes in unusual units, with Q_C given in BTU and the W in Wh. The conversion is

$$\begin{aligned} CoP &= SEER \frac{\text{BTU}}{\text{Wh}} \times \frac{1055\,\text{J}}{1\,\text{BTU}} \times \frac{1\,\text{Wh}}{3600\,\text{J}} && (4.6)\\ &\approx \frac{SEER}{3.5}. && (4.7) \end{aligned}$$

Air conditioners and household refrigerators both do the job of moving heat from someplace cold to someplace warmer, opposite the "natural" direction of heat flow. Another device that does the same thing is a ***heat pump,*** which can be used to move heat from cold outdoors to warmer indoors in winter time. A heat pump is an alternative to a furnace. If $CoP > 1$, as is usually the case, it would seem like a better choice than a furnace, because you only need to buy $W = Q_c/CoP$ worth of energy for the heat pump, but you need to buy all of the Q_C for a furnace. On the other hand a furnace can burn natural gas, which is much cheaper per Joule than the electricity that the heat pump needs, so the cost balance can tip either way. One good thing about a heat pump is that the same device is also your air conditioner, if you just swap the direction that the refrigerant flows when summer comes.

We could save a huge amount of energy if we could build an air conditioner / heat pump with very high CoP, since we would barely need to buy any new energy to heat or cool our home. Why can't we just get $CoP = \infty$ and get free heating and cooling?

Entropy

The questions we have raised were prominent in the minds of 19th-century physicists. They realized that not all energy is equally useful. A body containing energy also has a measure of *quality* or *usefulness* of its energy called ***entropy.*** Here is a working definition of entropy for us:

- When you *add* an amount Q of thermal energy to a material at temperature T, you *increase* its entropy by an amount
$$\Delta S = \frac{Q}{T}. \tag{4.8}$$
- When you *remove* an amount Q of thermal energy from a material, you *decrease* its entropy by Q/T.
- *Work* (non-thermal energy) that you extract from (or add to) a material does not change its entropy.

With this definition of entropy we can state the key to all our mysteries:

> **The Second Law of Thermodynamics: the total entropy of the Universe can never decrease. It can only increase or stay the same.**[4]

We can hypothesize any machine or process, and then add up all the entropy changes ΔS that it would cause anywhere in the universe. If the total change of entropy is $\Delta S < 0$, then this process is *physically impossible.* If it has $\Delta S > 0$, then the process is possible, but is ***irreversible***, because the machine that could undo it would be able to reduce total entropy, in violation of the 2nd Law. Only processes with $\Delta S = 0$ are ***reversible.***

A few examples will help. First: can we make a machine that takes kinetic energy as input and as output passes heat into a body at temperature T? Let's call this a "heater," as illustrated at the top of Figure 4.6. What change of entropy does this cause to the Universe? When we put work into the machine, we neither create nor destroy entropy. But when the energy gets put into the material as heat Q, we create entropy $\Delta S = Q/T$. The total change of entropy is *positive* so this is a physically realizable machine. In fact it's easy and common: a fire turns chemical energy into heat; friction turns kinetic energy into heat; an electric space heater turns electric energy into heat.

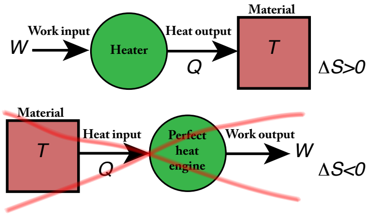


Figure 4.6: Schematic of a "heater" device and a perfect heat engine.

This process is *irreversible* according to the second law. The reverse machine would be able to take heat out of a material and turn all of it into non-heat energy. This would be our dream machine of a 100% efficient heat engine with no waste heat. When it ran, it would:

---

[4] What happened to the First Law of Thermodynamics? It just says that heat is energy, and energy is conserved. This is not news to us, but the idea that heat is energy was new in the 1800's and got "Law" status.

- Reduce the entropy of the material by $\Delta S = -Q/T$ when it removed heat from it.
- Make no further change to entropy when it output the energy as non-thermal work W.

The total change in entropy from running this machine would be negative, $\Delta S = -Q/T$*, and this is forbidden by the Second Law.* Therefore one cannot build a perfect heat engine. You can't build the "ice boat" from Figure 4.1. While friction can bring a sliding object to rest on a table, you will never see the heat in the table spontaneously set the object back into motion.

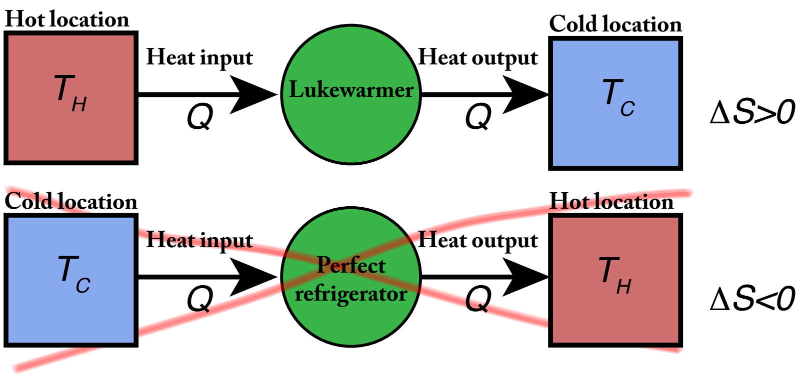


Figure 4.7: Two more machines: the "lukewarmer" transfers heat from a hot place to a cold place. The "perfect refrigerator" transfers heat from cold to hot places without needing any electric energy as input.

Figure 4.7 illustrates another use of the 2nd law, where we transfer of heat from one body to another. The "lukewarmer" machine moves an amount Q of heat *out* of a hot material and *into* some cold stuff. The total change of entropy is the sum of the ΔS for the two materials:

$$\Delta S = \Delta S_H + \Delta S_C = -\frac{Q}{T_H} + \frac{Q}{T_C} = Q\frac{T_H - T_C}{T_H T_C} > 0. \tag{4.9}$$

Entropy increases since $T_H > T_C$, and the 2nd Law says this is ok. Not surprising: all we need to do is put the two materials in contact, and we know heat will conduct from the hot to the cold one. But this process is *irreversible* and the reversed machine, which would be a perfect refrigerator, is impossible. The perfect refrigerator would have the following effect on the entropy of the Universe:

$$\Delta S = \Delta S_H + \Delta S_C = +\frac{Q}{T_H} - \frac{Q}{T_C} = Q\frac{T_C - T_H}{T_H T_C} < 0. \tag{4.10}$$

We can now see that there is a hierarchy of the usefulness of energy, with the energy being *most useful* when it takes the *least entropy* to extract it. It can be useful as thinking of the entropy as a kind of price tag attached to each Joule of energy.

- *Non-thermal* forms of energy that require zero entropy (for our purposes) to extract, (i.e. they have no entropy "cost") and can be used as input to any process—we can't end up with less entropy than the zero we started with.

- Thermal energy at *high temperature* is middle ground as extracting it produces some entropy Q/T. A machine that would turn this energy into purely non-thermal energy is impossible. But we can still use this energy by turning it into...

- Thermal energy at *low temperature* (ambient, basically) is essentially useless, because the entropy that it costs to extract it (Q/T_C) is higher than the entropy we would gain put putting that heat Q anywhere. This is the highest-entropy form that the energy can take; we cannot convert this energy to any form that has higher entropy, so any process that extracts that heat would end up reducing the entropy of the universe, and hence there are no physically possible uses for this energy.[5]

Now we understand why we keep needing to buy new primary energy: each time we use energy we increase its entropy, narrowing our options of how we could use those Joules. Once this energy has been converted to ambient-temperature heat, it is no longer possible to turn it into kinetic energy or another useful form. This is "waste heat."

We have taken the Second Law as a fact, but it can be derived from elementary laws of physics. Entropy is sometimes described as randomness or disorder, and the second law is the statement that the Universe is very unlikely to transform spontaneously into a highly ordered state from a disordered one. A classic example is to imagine a frictionless billiard table with a divider down the middle, all the balls bouncing on one side. When we remove the barrier, the balls will spread over the whole table. What are the odds that they'll spontaneously end up all on one side again? Very small! The laws of probability say that the half-empty state is unlikely, or a "low entropy" state, while nearly 50-50 division of balls is more likely, "high entropy." With the divider removed, system is very likely to spontaneously move to the higher-entropy (less ordered state), but extremely unlikely to evolve the other way. Maybe with 5 balls this might happen occasionally, but imagine 10^{23} billiard balls all being on one side—it's so unlikely as to be impossible. But that's how many molecules would be in a bottle of compressed gas. Once you open the bottle, the gas molecules spread throughout the room. You are increasing entropy, and that gas will never spontaneously go back into the bottle.

---

[5] Unless you find some reservoir of material with below-ambient temperature.

The Second Law says two bodies in contact will move toward the same temperature. If the two bodies stay in contact, it is highly unlikely—impossible, for practical purposes—that all the atoms in one body will spontaneously have low KE while all the atoms in the other body develop high KE. This is in concordance with our 2nd-Law deduction that heat will always flow in the direction that equalizes temperatures.

While this "disorder" explanation of the 2nd Law can be helpful, it is easy to get the wrong answer about entropy by using this qualitative meaning.[6] The mathematical statement $\Delta S \geq 0$ is the real Law that we should use to predict behavior.

Carnot efficiency

We have seen that a perfect heat engine would reduce entropy and be forbidden, but we know that heat engines exist. The key is to recognize that an engine's waste heat produces positive ΔS, which can counteract the negative ΔS when the heat is removed from the source. This is only true if *the waste heat is put in a place with lower temperature than the heat source.* Let's analyze the entropy balance of the less-than-perfect heat engine:

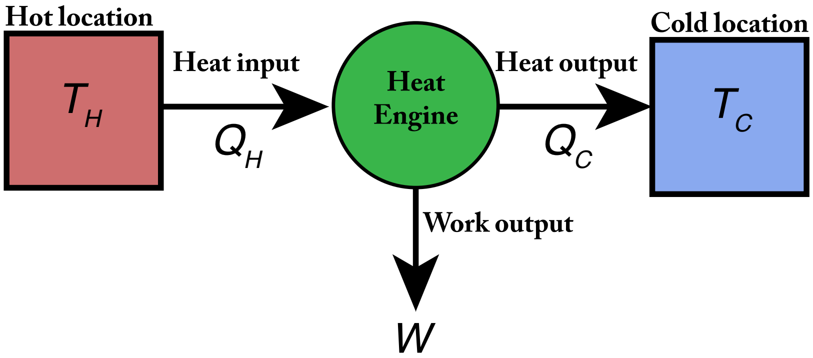


Figure 4.8: A heat engine takes energy from a source of temperature T_H, produces some work W, and sends the rest as waste heat to a place at cooler temperature T_C.

$$0 \leq \Delta S = -\frac{Q_H}{T_H} + \frac{Q_C}{T_C} = -\frac{Q_H}{T_H} + \frac{Q_H - W}{T_C} \tag{4.11}$$

$$\Rightarrow 0 \leq Q_H \left(T_H - T_C\right) - W T_H \tag{4.12}$$

$$\Rightarrow \epsilon = \frac{W}{Q_H} \leq \frac{T_H - T_C}{T_H} = 1 - \frac{T_C}{T_H}. \tag{4.13}$$

The Second Law of Thermodynamics demands that no heat engine can have an efficiency higher than the right-hand side of (4.13), which is called the ***Carnot efficiency***, after Sadi Carnot, who first postulated it in 1824, well before the Second Law was formulated.

---

[6] An example of an erroneous attempt to use the Second Law: it is often claimed that the Second Law forbids evolution from occurring.

We want to get $\epsilon \to 1$. To do this we need a large ratio of the output "heat sink" temperature to the source temperature. A heat engine needs to run on a temperature *difference.* In practice T_C is going to be no lower than the ambient temperature, near 290 K. If we have a coal furnace that heats gas up to temperature 580 K, then highest efficiency we could ever get from an engine using this hot gas would be $\epsilon = 1 - (290/580) = 50\%$. This is of course the *best* you could do—real engines will be less efficient than this, sometimes substantially.

We now understand why electricity generation and transportation have low efficiencies: because both are trying to turn the heat energy of burning fossil fuels into non-heat energy. Electricity is the most expensive energy form because it has zero entropy, which means that producing it from burning fuels is necessarily inefficient and expensive. On the other hand, it is a zero-entropy source of energy and hence 1 J of electricity is more useful than 1 J of natural gas or oil because we can do more with it at high efficiency.

Best possible refrigerators

We can also analyze the refrigerator in Figure 4.5 for its energy balance. We need

$$0 \leq \Delta S = +\frac{Q_H}{T_H} - \frac{Q_C}{T_C} = +\frac{Q_C + W}{T_H} - \frac{Q_C}{T_C} \tag{4.14}$$

$$\Rightarrow 0 \leq W T_C - Q_C\left(T_H - T_C\right) \tag{4.15}$$

$$\Rightarrow CoP = \frac{Q_C}{W} \leq \frac{T_C}{T_H - T_C} \tag{4.16}$$

Note that we can easily have $CoP > 1$. In fact the CoP can get larger and larger when T_H and T_C are near each other. When you want to move heat "uphill" against a bigger ΔT, you need to put in more work. A heat pump, therefore, will work better (higher CoP) if it's only mildly cold outside.

Cogeneration

Waste heat from electrical plant is usually dumped to a river. But this "waste" heat is usually in a substance that is above room temperature, so it could be used as heat for buildings, or "process heat" in industry. It is efficient to combine electrical generation with such a use for the heat in a "cogeneration plant." A difficulty is that we want our power plants sited far away from where we live. Cities used to pipe steam produced in power plants to buildings around the city, which would use it to warm the inside air ("district heating"), but it is less common now. Penn gets its heat from steam generated at Grays Ferry by Veolia Energy. They get some of this heat as the byproduct of running an electrical generator. More common is for a factory that has generators or high-temperature processes to use the "waste" heat for heating its buildings and offices.

Practice problems

1. Why do your ears "pop" when you go to high altitude in a plane or car? Why is deep diving dangerous?
2. If you use 1 GJ of natural gas in your oven in summer, how much will this cost you in total? Don't forget the electricity you need to buy to remove this energy from your house with your $SEER = 17.5$ air conditioner.
3. Which of these machines would you invest in?
    - A new type of generator that burns natural gas and converts it into electricity with 90% efficiency because it has no moving parts and no friction.
    - A WiFi access point that needs no batteries or electricity power cord because it creates electricity from the thermal energy of room air.
    - A wristwatch that needs no batteries because it generates electricity from your body heat.
    - A machine that creates electricity by taking heat from the top of the ocean (typically 19° C) and dumps it to the bottom of the ocean (typically 6° C).
4. Estimate how much it costs you in electricity every time you open the refrigerator door (let's assume that all the interior air at 2° C is replaced by room-temperature air, and your refrigerator has $CoP = 4$.

5
Light and atoms

The EIA reports that the US used 0.8 EJ of electricity for lighting in the commerical plus residential sectors (in 2020) and another 0.20 EJ for industrial lighting (in 2014). Recalling that each EJ of electricity takes $\approx 3\,\mathrm{EJ}$ of primary energy to generate, we see that lighting accounts for about 3% of primary energy use in the US per year. In this chapter we investigate the nature of light and how light bulbs work, since this is one of the larger uses of primary energy behind transportation and heating/cooling. Lighting is a success story in reducing energy use and we will see how this is achieved—the above numbers are significantly lower than in earlier years, due to technological advances.

Studying the nature of light will lead us also to an understanding of how CO_2 emission changes the temperature of the planet.

The electromagnetic spectrum

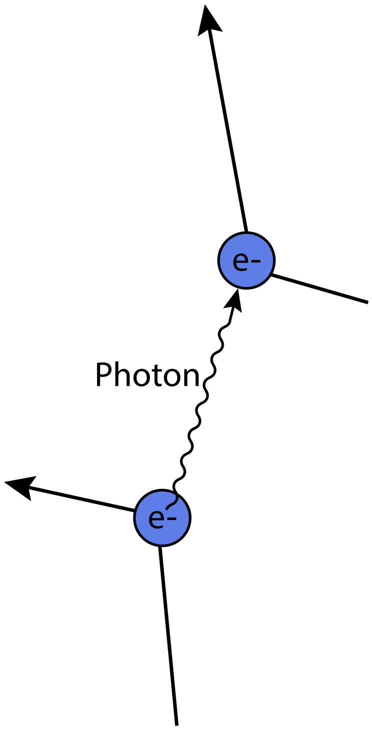


Figure 5.1: Illustration of an electron, e^-, giving up some of its kinetic energy in the form of a photon, which is then absorbed by another electron potentially very far away. Photons are typically labelled with γ and the wavy lines mark their paths.

The $\odot$ is the astronomical symbol for the Sun.

Light is a stream of ***photons***, which are little packets of pure electromagnetic energy. They travel in straight lines at the speed of light (duh!), $c = 3 \times 10^8\,\mathrm{m\,s^{-1}}$, which is 186,000 miles per *second.* Photons are emitted by one electric charge (usually electrons) and absorbed by another, perhaps after travelling across space for billions of years.

The cells in your retina can be the destination for photons, which you will perceive as brightness in the direction the photon came from. A bright (dim) spot will be emitting more (less) *total power* of photon stream toward your eye. The photon power incident per unit area on a surface is called the ***flux***, $f = P/A$. For example, the flux of light incident on Earth from the Sun is called the ***solar constant***,

$$f_\odot = 1360\,\mathrm{W\,m^{-2}}. \tag{5.1}$$

This means that if you were to take a 1-square-meter window and point it toward the Sun, 1360 J of light energy would pass through it every second. This is the value above Earth's atmosphere; on a cloudless day, only about 1000 W of this power will make it to Earth's surface.

Flux describes brightness but of course we also perceive *color* of the light. What distinguishes *e.g.* red light from blue light? The color is determined by *the energy of the individual photons*, whereas the brightness is the *total power* in the stream of photons.

Color	Photon Energy
blue	5.0×10^{-19} J
red	2.5×10^{-19} J

Table 5.1: Energies of visible-light photons

The light from the Sun consists of many different energies of photons, or many different colors. What we perceive as white light is actually a mix of all the colors. If you pass the Sun's light through a prism, it breaks up into a rainbow of colors, arranged in order of increasing energy per photon from red to blue. Isaac Newton first demonstrated that with a second prism, you can combine all these colors of light back into a single stream we perceive as white.

Photons are not restricted to having energy in the narrow range $2.5 - 5.0 \times 10^{-19}$ J of the rainbow; a photon can have almost any amount of energy. What does a photon with 7×10^{-19} J of energy look like? Your eye cannot see this, but we call this ***ultraviolet (UV)*** light. A photon with 1×10^{-19} J would be ***infrared (IR)*** light, also invisible but with very real energy and effects. Figure 5.2 shows that ***radio, microwave, x-ray, and γ-ray*** radiation are all just photons of different energies (colors), making up the ***electromagnetic spectrum.***

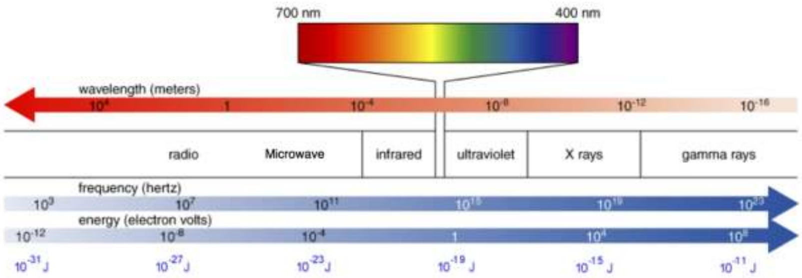


Figure 5.2: The electromagnetic spectrum, showing the names given to photons in different energy ranges. The visible colors are just a narrow slice of energy. We will not be concerned with the "wavelength" and "frequency" labels in this course. Taken from the Bennett et al. astronomy textbook.

Thermal radiation and the temperature of Earth

As illustrated in Figure 5.1, photons are created by charged particles (electrons or protons). If an electron wants to emit a photon[1] with $E_\gamma = 10^{-18}$ J, conservation of energy demands that the electron must have at least 10^{-18} J of energy to give up, probably in the form of KE. A box of electrons and protons can and will, therefore, emit light, called ***thermal radiation*** or ***blackbody radiation,*** that is controlled by the *temperature* of the particles, because T determines how much energy the particles have available to turn into photons.

There are two rules describing thermal radiation. The first is

---

[1] The greek letter γ (gamma) is typically used to denote photons.

Wien's Law, which states the the *typical* energy $\langle E_\gamma \rangle$ of a photon emitted obeys

$$\langle E_\gamma \rangle \approx k_B T. \tag{5.2}$$

Recall that $k_B T$ is the typical thermal (kinetic) energy of each particle. It makes sense that this will also be roughly the typical energy of emitted photons.

The second rule for thermal radiation is the ***Stefan-Boltzmann law***, which states that the *total* thermal radiation power leaving the surface of some object (the flux) is

$$f = \sigma T^4, \tag{5.3}$$

$$\sigma = 5.67 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}. \tag{5.4}$$

The quantity σ is called the Stefan-Boltzmann constant. Figure 5.3 shows how much energy is emitted at a given photon energy from surfaces of different temperatures.

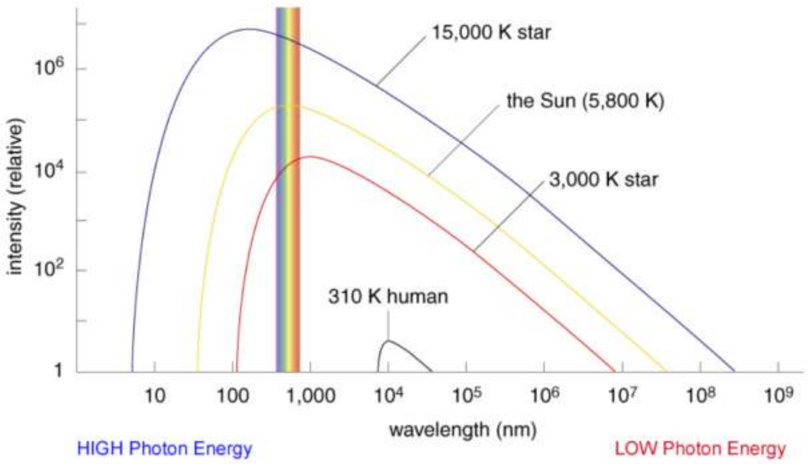


Figure 5.3: The amount of thermal (blackbody) radiation emitted by surfaces at different temperatures. Note the backwards x axis: the rainbow marks the energies of visible photons, with the higher-energy blue photons on the left. To the left of the rainbow is UV and x-ray light; to the right of the rainbow marker would be IR, microwave, and radio radiation. Taken from the Bennet et al. astronomy textbook.

The two thermal-emission formulae above have the following effects:

- As the surface warms up, its *dominant* photon energy moves toward the blue, high-energy end of the spectrum. The Sun, with a surface at 5800 K, looks yellow because this is the peak of its thermal emission curve. At 3000 K, a cooler star, or perhaps a furnace, has thermal emission peaking at the red end of the spectrum, so it will look red. A hotter surface, say 10,000 K, peaks in the blue or UV, so to your eye it will look bluish-white. Contrary to popular convention, *blue-hot is hotter than red-hot.*

Do you think it is a coincidence that the human eye detects a narrow range of photon energies that includes those most abundantly produced by the Sun?

- As the surface becomes warmer, the *total* amount of emitted light grows rapidly because of the T^4 in equation (5.3). A 10,000 K surface emits $2^4 = 16\times$ more radiation power than a 5000 K surface.

- A hotter surface emits more power at *every* color than a cooler one. So the hot star emits more red light than the cool star, even though the cool star's favorite photon is red. The electrons in the hot star have more energy and are more prolific producers of every type of photon.

Figure 5.4: A person sneaking around on a golf course, as seen by an infrared camera.

A surface needs to be at temperature near 1000 K or hotter to emit any photons visible to the human eye. Cooler objects do, however, emit thermal radiation, they just do so at lower photon energies that the human eye cannot see. Your body, at $\approx 310\,\mathrm{K}$, emits primarily *infrared* radiation—you glow in the dark! If you had infrared vision, you would see your skin glowing. Of course the rest of the room glows too; but since your body is warmer than room temperature, you glow a little brighter than your surroundings. An infrared camera is clearly useful for the military looking for enemy soldiers at night. Spy planes with infrared cameras can tell which cars and airplanes have been used recently since they will be warmer. There are many more benign uses of infrared cameras: they can measure the temperature of surfaces, for instance showing where heat is leaking out of a building, lighting the way to spots needing better insulation.

Figure 5.5: Thermal infrared image of a house in cold weather. The lighter regions are warmer and indicate areas with more heat loss, for instance due to missing insulation. From *Introduction to Thermography Principles*.

Incandescent light bulbs

The standard ***incandescent*** light bulb works through simple thermal radiation emission. The *filament* of the light bulb is heated to several thousand Kelvin using electricity so that its thermal radiation is visible light. Halogen bulbs have hotter filaments than standard incandescents: this is why they look blue/white instead of the "warm" yellow/red glow of standard incandescents, and are also much more intensely bright.

There are two tricks to making an incandescent bulb. First, the filament must not melt at high temperatures. They are typically made of tungsten, which has the highest melting point of any element at 3695 K. Second, the filament should not burn up: this is accomplished by putting the filament inside an oxygen-free container. Bulbs are usually filled with nitrogen. That's why the light burns out if you break the glass bulb.

Even if it doesn't melt or burn, the filament will evaporate, and eventually break. If you try to make the filament too hot, it will evaporate quickly and the bulb's lifetime becomes short. Halogen lamps use special tricky chemistry to constant re-deposit the evaporated metal back onto the filament, so they can be run at higher temperatures without "burning" out.

The big problem with incandescent bulbs (aside from their short lifetimes) is that they are horribly inefficient. In Figure 5.3 you can see that the 3000 K star (or filament) emits most of its photons in the IR. These photons take up power but they do not help you read at night. You feel these photons as radiated heat. Plus the hot filament conducts and convects a lot of heat to the bulb. Incandescent lamps turn a very small fraction of their energy into visible light and generate a large amount of useless, even dangerous, heat.

Caution

Wien's Law is of great use for thermography, and also for astronomers, because it lets us judge the temperature of things simply by looking at their color, without need to actually touch the target. We need to be careful however: the Moon also looks yellowish, but its surface is *not* at 6000 K like the Sun. The Moon looks yellow because it is *reflecting* visible light that was emitted by the Sun. The Moon does emit its own light but only in the infrared. Wien's Law only applies to the thermal emission of a surface, not the reflected light from other sources. That's why blue pants do not have to be at 10,000 K.

Earth's radiative equilibrium

Because Earth sits in the vacuum of space, it can gain or lose energy only through radiation. Earth *absorbs* energy that is radiated by the Sun. The Earth presents a circular face to the sun of area $A = \pi R_E^2$, where R_E is the radius of Earth (about 6400 km). So the power that hits Earth is

$$P = fA = f_\odot \pi R_E^2. \tag{5.5}$$

Not all of this light is absorbed: a fraction $a = 0.30$, called the ***albedo***, is reflected back into space, *e.g.* by clouds. Therefore the total power *absorbed* by Earth is

$$P_{\text{abs}} = (1-a) f_\odot \pi R_E^2. \tag{5.6}$$

If the Earth just absorbed this power, it would heat up, and would have melted billions of years ago. However the Earth, like all other objects, also *emits* thermal radiation from its entire surface according to Equation 5.3. Since the total surface area of the spherical Earth is $A = 4\pi R_E^2$, the total power *emitted* by Earth is

$$P_{\text{emit}} = fA = 4\pi\sigma T_E^4 R_E^2, \tag{5.7}$$

where T_E is the average temperature of Earth's surface.

We say the Earth is in ***radiative equilibrium*** when it is emitting exactly as much power as it is absorbing. This must be what happens: if the Earth were absorbing more than it emitted, it would heat up. As T_E rises, P_{emit} would rise, until it matches P_{abs}. Likewise if $P_{\text{emit}} > P_{\text{abs}}$, the Earth would cool until equality were restored. Earth will always, therefore, choose T_E so that $P_{\text{abs}} = P_{\text{emit}}$, which gives

$$(1-a) f_\odot \pi R_E^2 = 4\pi\sigma T_E^4 R_E^2, \tag{5.8}$$

$$\Rightarrow T_E^4 = \frac{(1-a) f_\odot}{4\sigma} \tag{5.9}$$

$$\Rightarrow T_E = \sqrt[4]{\frac{(1-a) f_\odot}{4\sigma}} \tag{5.10}$$

$$= \sqrt[4]{\frac{(1-0.3)\cdot 1360\,\text{W m}^{-2}}{4\cdot 5.67\times 10^{-8}\,\text{W m}^2\,\text{K}^{-4}}} = \underline{\underline{255\,\text{K}}}. \tag{5.11}$$

This answer is clearly wrong, however; only the poles are as cold as this and the average temperature of Earth's surface is more like 295 K. The formula works for the Moon but goes wrong for Earth and Venus. We must have missed something important.

Atoms and light

Figure 5.6 gives a close-up view of the spectrum of light coming from the sun. There are dark spaces, ***absorption lines***, where light from a specific color is *missing*. To understand why, we need to learn a key fact from quantum mechanics, the rules that govern the behavior of atoms, molecules and light.

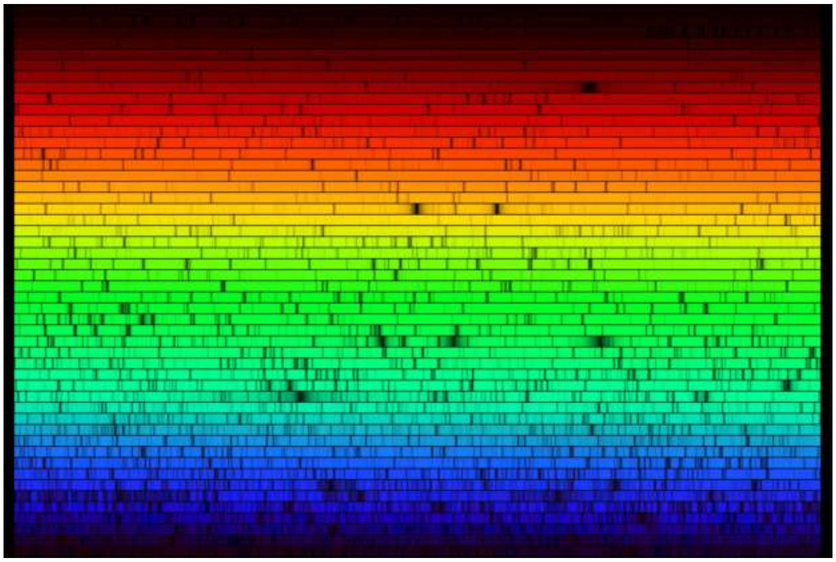


Figure 5.6: The spectrum of the sun (spread out and arranged into lines like the text of a book). Courtesy NOAO.

The simplest atom, hydrogen, has an electron with charge $-e$ orbiting a distance r from a proton—the nucleus—with a charge $+e$. Recall that this atom will have an electric potential energy of $PE = -e^2/r$. The electron could:

- Spontaneously *emit* a photon. Conservation of energy demands that the outgoing photon energy be compensated by moving the electron *closer* to its nucleus, lowering the PE to a more negative value.

- *Absorb* a passing photon, in which case the electron must move *farther* from its nucleus to make its PE less negative. If the photon energy is high enough, $E_\gamma > e^2/r$, then the electron can be set free of its nucleus ($r \to \infty$) with some kinetic energy to spare.

It would seem that the atom could emit or absorb a photon of any energy at any time. In fact there is a mystery: why doesn't the atom just keep emitting photons, getting closer and closer to its nucleus and getting infinitely deep into electric PE "debt"?

In the early 1900's, quantum mechanics was developed as a new theory of microscopic behavior that explains this conundrum as follows:

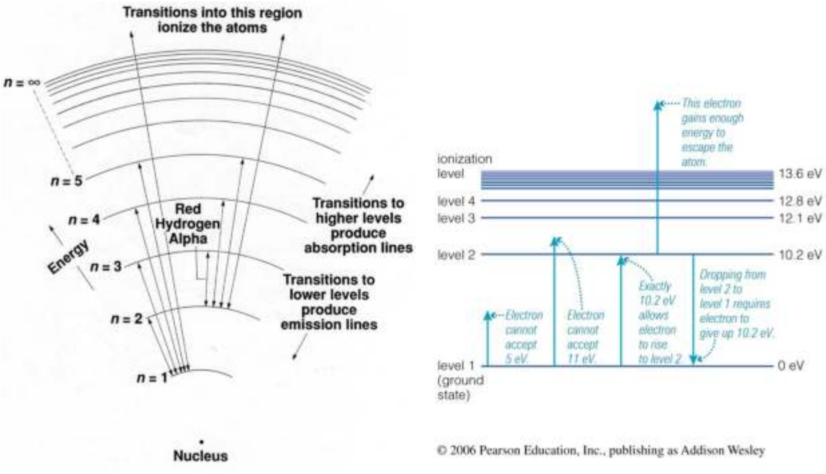
- The electron's orbit is ***quantized,*** meaning that it is allowed to exist in only a particular set of orbits. We call each orbit a "state" of the atom and we could number them $n = 1, 2, \ldots$.

- Each state has a particular potential energy E_n (which will generally be negative). These are called the ***energy levels*** of the atom. At any given moment the atom must have one of the energy levels $E_1, E_2, \ldots$. The state with the lowest energy level E_1 is called the ***ground state*** and has the electron closest to its nucleus.

- The atom *can only absorb or emit photons that contain exactly the right amount of energy to go from one energy level to another.* Thus the atom will absorb and emit only specific colors of light, with photon energies $E_\gamma = E_n - E_m$ for a transition from state n to m. For photon energies other than this specific set of ***spectral line*** energies, the atom is essentially transparent, and can neither emit light nor absorb passing light.

Every atom or molecule, not just hydrogen atoms, has a well-defined set of states, which we can derive using the formal mathematics of quantum mechanics. These rules make some very interesting—and invariably correct—predictions:

- Since there is a lowest energy level, the ground state, an atom left in isolation will stop emitting light once it falls down into the



© 2006 Pearson Education, Inc., publishing as Addison Wesley

Figure 5.7: At left, illustration of the different electron orbits allowed by quantum mechanics for a hydrogen atom. Closer orbits to the nucleus have lower (more negative) energy. Photons can be emitted or absorbed by moving between these states. At right, an illustration that photons can only be emitted or absorbed if their energy exactly matches the difference to a new energy level. Left and right from the Impey and Bennet astronomy textbooks, respectively.

ground state. It cannot emit light again until something comes along—a photon, or a collision with another atom or electron—and gives the atom energy to get to a higher state. The higher states are called ***excited states.***

- The thermal radiation we described above, where all energies of light are emitted, is a characteristic of charges that are free to occupy any energy, such as a free-floating electron. This is called ***continuum emission***

- If there is a cloud of atoms between us and the continuum source, these atoms will selectively absorb the photons at their quantum-allowed energies, hence creating the dark absorption lines in the spectrum of a source like the Sun.

- Each atom or molecule will have a characteristic set of photon energies that it will absorb, based on its arrangement of energy levels. These spectral lines will serve as a fingerprint for that atom or molecule; we can tell which atoms/molecules are present without having to be there to collect them. This is the key to "remote sensing." Figure 5.6, for example, tells us that the Sun is comprised of about 3/4 hydrogen, 1/4 helium, and $< 1\%$ of all the other elements combined.

- We can add energy to atoms or molecules to put them into excited states. For instance we can run an electric current of fast-moving electrons through some gas or solid; the electrons will transfer some of their KE to the atoms when they collide. The excited atoms/molecules can then release light when as they fall back into the ground state. This light will come out only in the specific

Figure 5.8: Emission-line "fingerprints" of three common chemical elements. From Bennett et al textbook.

colors required by the energy levels and we will see ***emission lines***. The emission lines' colors will bear the same fingerprint of their source as the absorption lines.

Emission-line light bulbs

We can by clever choice of our atoms and molecules perform many useful manipulations of light. A red shirt does not get its color by emitting light at a temperature of 3000 K. Instead it contains materials that preferentially absorb blue light and reflect red light. All the photons originate from the Sun or a light bulb.

We can greatly improve on the efficiency of the incandescent bulb by building devices that emit visible light without the infrared light and heat emission associated with thermal radiation. The simplest is a **neon light**, whose pretty orange glow adorns many a bar and tattoo parlor. This is a tube filled with neon gas. Electricity is used to run electrons down the length of the tube at high speed. They collide with neon atoms and put them into excited states. The neon atoms then emit the characteristic orange light of the emission lines of neon. Other gases make other colors.

Sodium vapor lamps and **mercury vapor lamps** work on a very similar principle with different light-emitting elements. Because sodium and mercury are solid or liquid at room temperature, these lamps must be warmed up and create gas atoms before they can emit their full brightness. But once running, they can produce a lot of light very efficiently by having an abundance of strong emission lines in the visible range. These are the kinds of lights you'll typically see on highways or in stadiums and arenas.

Fluorescent lamps are a little more complex: inside the tube is a gas vapor (including a little mercury) that generates UV light when we run electricity. On the inside walls of the glass tube we place molecules called **phosphors** which can absorb the UV to go into an excited state, but then re-emit the energy in multiple lower-energy, visible-range photons by stopping at intermediate energy levels on their way back to the ground state. You can mix different phosphors on the bulb wall to obtain a more pleasing light to mimic the color of the Sun.

Light-emitting diodes (LEDs) use tricks of solid-state physics to create energy levels in a solid material instead of a gas of atoms. This makes them more robust and compact as there are no glass tubes to hold gas. They also last longer. LED bulbs have gotten substantially brighter and cheaper in recent years, so are replacing incandescent and fluorescent lamps in more and more applications.

Ionizing radiation

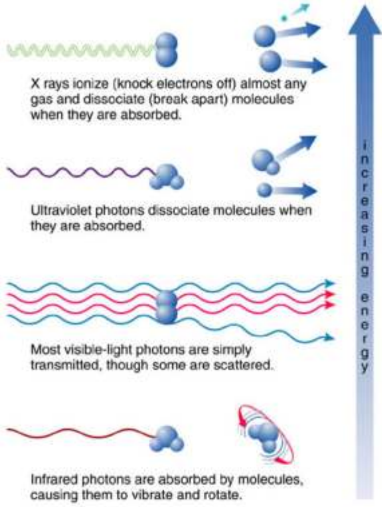


Figure 5.9: Effect of different energy ranges of photons on atoms and molecules. © Pearson Education

If we give an atom a photon with energy $E_\gamma > -E_n$ to an atom in a state with (negative) energy level E_n, the electron can absorb the photon, pay off its electric PE "debt" to the nucleus to be set free to wander away, with the remainder of the energy turned into KE for this free electron. The atom left behind has a net positive charge and is called an ***ion***. Photons with enough energy to rip an electron off a typical atom are called ***ionizing radiation,*** and of course these occupy the upper end of the electromagnetic spectrum: UV, x-rays, and γ-rays (gamma rays).

Notice that once the electron is free, it can have any amount of KE, so we are no longer restricted to specific photon energies to move between energy levels. Ionization can make use of any photon with energy above a certain threshold.

Similarly every molecule is two atoms bonded together, and if you add enough energy you can break the molecule apart ("dissociation"). Hitting the atom with a sufficiently energetic photon will break the molecule and change its chemical behavior. The photons capable of breaking molecules are essentially the same as the ionizing ones, although the threshold for damage depends on the specific atom or molecule.

Ionizing radiation is a health hazard because it can disrupt the biological function of the complex molecules in living tissue. *This is why ultraviolet light gives you sunburn while visible light does not.* When a molecule absorbs a visible photon, it gets excited but doesn't break. Very shortly thereafter, before another visible photon can come along, the molecule will de-excite, turning the photon energy into thermal energy. An intense beam of visible (or any color) of light can cause a *burn* as we put enough power into the tissue to raise its temperature. The damage from ionizing photons is very different: every *single* ionizing photon can break a molecule when it hits, disrupting the tissue function rather than cooking it.

The DNA molecules in your cells are particularly critical. If we break or rearrange them, we've changed the genetic code of the cell. This is a ***mutation*** and ionizing radiation is ***mutagenic.*** The mutation will usually cause the cell (or its descendant) to die. Less commonly an unlucky mutation will lead the cell to divide out of control and become cancerous. UV, x-ray, and gamma-ray radiation are therefore also ***carcinogenic.*** It is critical to realize that the lower-energy photons—radio, microwave, infrared, and visible light—*are not cancer-causing!* Claims in the 1990's that cell-phone use causes brain cancer were met very skeptically by anyone with basic scientific knowledge, because cell phones communicate with a small amount of microwave radiation to/from the towers. These microwave photons cannot break DNA bonds. Likewise power lines do not generate ionizing radiation. The threat of skin cancer from UV light is, however, very real. Exposure to X-rays and gamma rays, which penetrate deep into the body, increases risk for many kinds of cancer. The more such photons your body absorbs, the more random mutations you generate, and the more likely it becomes that a cancerous one is produced and thrives.

On rare lucky occasions, the mutated cell will perform some function better and the organism has evolved to a new state.

This is not to say that microwave radiation is harmless: we know you should not dry your cat in the microwave oven. Microwave ovens cook food by bombarding them with huge numbers of microwave photons, which are absorbed by water and fat molecules in the food. Over 1 kW of microwave energy can be absorbed in your food and is turned into heat. The temperature rises rapidly, and this would damage any living creature. It is the *total power* that is dangerous, not the creation of mutations. Compare this to the $\approx 1\,\text{mW}$ of microwave power that a cellphone broadcasts, which produces much less heat than your body is already producing internally. Clearly the dose makes the poison when it comes to non-ionizing radiation!

Spectral lines in Earth's atmosphere: ozone and UV

We know that the Sun emits UV light and that UV light can destroy biological molecules. Survival of life on Earth depends on blocking these photons before they reach the surface. Luckily there is a level in the upper atmosphere containing substantial quantities of ozone (O_3), a molecule that has many strong spectral lines in the UV part of the spectrum. The ozone molecules hence absorb most of the UV from the Sun and act to shield us.

Paradoxically the ozone not only blocks the UV light, but it would not exist without UV light either because it's a highly reactive molecule, which is easily destroyed in reactions with other molecules. Ozone must be constantly created to restock the upper atmosphere's supply. Plants fill our atmosphere with O_2 molecules. A UV photon can split an $O_2 \rightarrow 2O$ and then $O_2 + O \rightarrow O_3$ can occur. This is why

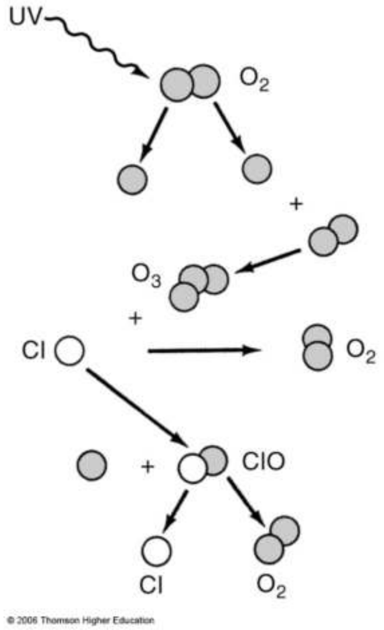


Figure 5.10: Processes for creation of ozone in the stratosphere and destruction by chlorine. ©2006 Thompson, from the Hinrichs textbook.

the ozone is only found in the upper atmosphere, since there is little UV available to produce it at lower altitudes. UV photons can also break up O_3, so there is a complex interaction between ozone and UV light. Note that the ozone layer requires O_2 to be in the atmosphere, which is produced by plants. So how did life evolve if there was initially no ozone layer to protect it? Liquid water blocks UV light effectively, so life could safely evolve in the oceans. Life could not take root on land until photosynthesis "polluted" the atmosphere of Earth with O_2.

Most glass also absorbs UV light, so you don't get a sunburn inside your house or car if the windows are shut.

Humans nearly destroyed Earth's ozone layer inadvertently. The culprit are chemicals called ***chlorofluorocarbons (CFCs)***, which are extremely useful because they are exceptionally stable and inert. They make, for example, excellent refrigerants, and were the standard working fluid for air conditioners and refrigerators for several decades (one brand name is Freon). In the 1970's, chemists Sherwood Rowland and Mario Molina began to wonder what would happen to all the CFC's being released. The stability of the CFC molecules makes them survive many decades in the atmosphere, so they would eventually drift up to the stratosphere, where exposure to UV light would be able to break them apart. Rowland & Molina also realized, however, that the chlorine atoms released in this process would be act as very potent catalysts to turn O_3 back into O_2, disrupting the current balance between O_3, O_2, and UV in the stratosphere. They calculated that then-current CFC production rates would lead to substantial degradation of the ozone layer.

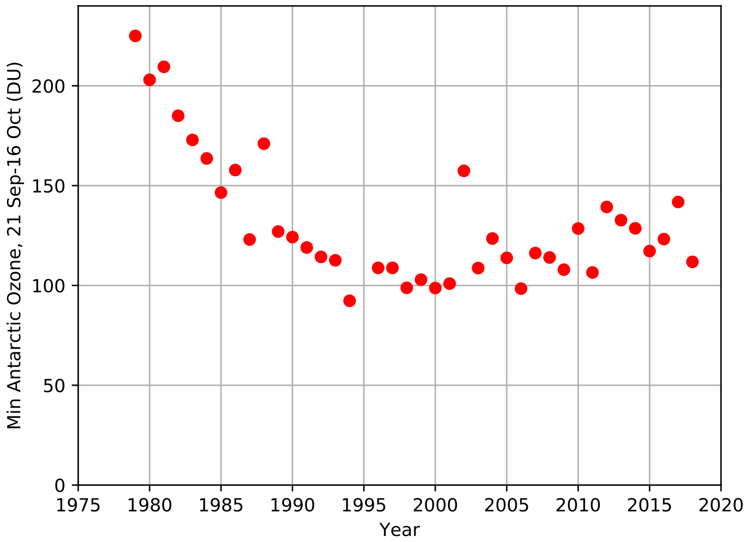


Figure 5.11: The dots show the lowest measured ozone level in the Sep/Oct Antarctic atmosphere each year. The Montreal Protocal went into effect in 1989. From NASA/GSFC.

By 1987 it was clear that dramatic depeletions of ozone were occurring over the South Pole every year, a new phenomenon. Careful research confirmed that this was due to CFC emission. Remarkably quickly, a series of international agreements was signed, beginning with the 1989 Montreal Protocol, to severely curtail the manufacture and use of CFCs. Figure 5.11 shows that this action has succeeded in arresting the accumulation of chlorine in the stratosphere and the resultant decline in Antarctic ozone. Because CFCs are very long-lived, it will take roughly 100 years for the ozone to return to pre-industrial behaviors. Rowland and Molina were awarded the 1996 Nobel Prize in Chemistry for their work. The ozone levels are slowly recovering (Figure 5.11), as the atmospheric chlorine dissipates, but it will take several decades to recover.

Spectral lines in Earth's Atmosphere: the greenhouse effect

Recall that our calculation of Earth's surface temperature based on thermal radiation predicted a mean temperature of 255 K, but we do not live on an ice planet. The is due to the presence of ***greenhouse gases*** that have strong spectral lines in the *infrared* portion of the EM spectrum. Earth absorbs energy from the Sun, which at a temperature of 5800 K is primarily emitting in the visible region of the spectrum. Our eyes tell us that Earth's atmosphere is highly transparent to visible light (very few absorption lines) except for clouds, so most of the Sun's energy reaches Earths' surface.

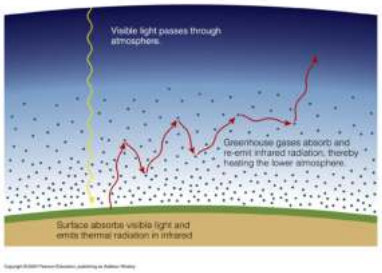


Figure 5.12: Illustration of the heat-trapping effect of greenhouse gases. ©Pearson Education

Earth must re-radiate this energy back to space to maintain a stable equilibrium. But since Earth's surface is 20× cooler than the Sun's, Wien's Law says Earth is radiating its energy primarily as IR light. Earth's atmosphere is $> 99.9\%$ N_2, O_2 and Ar gas, which are highly transparent to both visible and IR light. But in the last 0.1% are gases such as CO_2, H_2O, and methane (CH_4) which have strong spectral lines in the IR. These molecules intercept the IR light leaving Earth's surface. Then the molecules will re-emit the energy as IR photons, but in a random direction, *so that much of the intercepted energy is directed back toward Earth's surface.* These spectral lines hence inhibit the release of Earth's energy back to space while leaving the incoming solar energy largely undisturbed. The atmosphere acts as a kind of one-way insulating blanket around Earth's surface.

The consequence of blocking some of the outgoing power is that Earth's surface warms up. As it warms, Equation (5.3) tells us that it will emit *more* infrared light. Eventually it warms to the point where the escaping fraction of IR light carries the same power as the incoming solar light. Equilibrium is restored, but with Earth's surface temperature higher than our simple calculation suggested.

This higher surface temperature due to blocking of IR light is called the ***greenhouse effect.*** Earth would be an uninhabitably cold planet without it! On the other hand, you can have too much of a good thing. The surface of Venus has an average temperature of 470° C, hot enough to melt lead! Yes, Venus is closer to the Sun than Earth, but only by 25%, which leads to only a few degree rise above Earth's temperature. The problem on Venus is a runaway greenhouse effect: it has an enormously thick atmosphere of almost pure CO_2, which keeps its surface over 500 K warmer than the radiative equilibrium temperature would be.

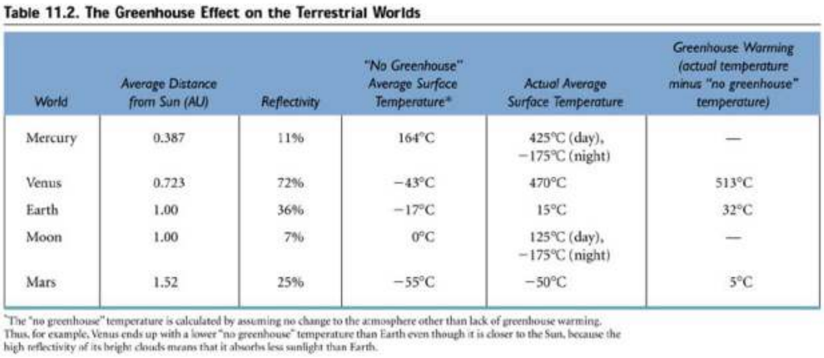


Table 11.2. The Greenhouse Effect on the Terrestrial Worlds

World	Average Distance from Sun (AU)	Reflectivity	"No Greenhouse" Average Surface Temperature*	Actual Average Surface Temperature	Greenhouse Warming (actual temperature minus "no greenhouse" temperature)
Mercury	0.387	11%	164°C	425°C (day), −175°C (night)	—
Venus	0.723	72%	−43°C	470°C	513°C
Earth	1.00	36%	−17°C	15°C	32°C
Moon	1.00	7%	0°C	125°C (day), −175°C (night)	—
Mars	1.52	25%	−55°C	−50°C	5°C

*The "no greenhouse" temperature is calculated by assuming no change to the atmosphere other than lack of greenhouse warming. Thus, for example, Venus ends up with a lower "no greenhouse" temperature than Earth even though it is closer to the Sun, because the high reflectivity of its bright clouds means that it absorbs less sunlight than Earth.

Figure 5.13: The strength of the greenhouse effect in raising temperatures above the radiative equilibrium "no-greenhouse" level on some of our Solar System neighbors. Earth and especially Venus are warmer because of their atmospheres containing IR spectral lines. ©Pearson Academic.

The physics of light and atoms makes this result straightforward: if the CO_2 content of the atmosphere goes up, less radiation will escape Earth's surface and Earth will start have a net gain in energy. This additional energy has to go somewhere. One place for it to go is to gradually warm the Earth's surface, which eventually will restore the equilibrium between incoming and outgoing radiation power. Direct sampling of the atmosphere shows that its CO_2 content has risen $1.5\times$ higher in the past 50 years, and is higher than it has been for many thousands of years. We can expect global warming, which we will discuss in detail in a later chapter.

Light bulb roundup

Let us get back to our initial question about energy use for lighting: how can/did we produce the light we need or want using less input energy? We have seen that using spectral-line physics we can improve the fraction of our power that goes to producing visible light. The visible-light output of a bulb is measured in ***lumens***. This is an unusual unit in that it contains biology as well as physics. It takes into account the higher efficiency of the human eye for some colors compared to others. If the light is a particular shade of green, where the human eye is most sensitive, then 1 W of light energy is equivalent to 683 lumens. Let's call this 100% efficiency.

Living in pure green light would not, however, be acceptable to most people, so we want to produce a full visible spectrum of light that will look close to white. Since the eye is less efficient at using red and blue light than green, we necessarily would get less perceived brightness from even an ideal white-light source: 242 lumens per Watt, or about 37% efficiency, is the best possible performance for near-white bulbs.

Lamp type	Luminous efficacy (Lumens per Watt)	Efficiency
Candle	0.3	0.04%
100W tungsten incandescent	17.5	2.6%
Tungsten quartz halogen	24	3.5%
Compact fluorescent	46–75	4.4–7.3%
Commercial LED	100-175	15–25%
High-pressure sodium	85–150	12–22%
Perfect sun-colored	251	37%

Table 5.2: Luminous efficacy and efficiency of some typical bulbs. The "efficacy" is the perceived light output (lumens) per W of input power. The "efficiency" is the ratio of this efficacy to a stream of pure green light, to which the eye is most sensitive. Values taken from Wikipedia "Luminous efficacy" article, 2019.

From the Table above we can see that fluorescent (CFL), LED, and sodium lamps, which work using emission lines, need as little as 1/10 as much electric power as an incandescent to produce the same brightness, enabling an enormous savings. So why would anyone ever use an incandescent bulb? Most people like to buy them because they are *cheaper,* of course.

If, however, you conduct a ***lifecycle analysis*** of the choice between an incandescent bulb and a CFL or LED, you find that the later savings on electricity from the latter more than make up for their higher purchase price compared to incandescents. And of course they use less electricity and hence produce less CO_2. Furthermore the CFLs and LEDs last far longer since they do not have hot filaments to evaporate away. Therefore you save more by not having to buy them as often (or, in a commercial setting, pay someone to go around replacing bulbs).

Let's do an example life-cycle analysis. Table 5.3 gives the specifications of an incandescent and this LED bulb that produce nearly the same light output. Let's ask what is the *total* cost to the consumer of operating a lamp for 15,000 hours with either bulb.

Bulb type	Incandescent	LED
Power use (W)	100	15
Light output (lumens)	1690	1600
Lifetime (hours)	750	15,000
Purchase cost	\$0.40	\$1.60

Table 5.3: Specificiations of incandescent vs LED bulbs.

- First consider purchase price. A single LED will last this whole time (which is over 10 years if you use the bulb for 4 hours per night!). This will cost \$1.60 (in an 8-pack from Home Depot) But you will need to buy $15,000/750 = 20$ incandescents over this much use for a total cost of \$8.00, which is already more than the

LED.

- Now consider the operating costs of the bulb. If I pay \$0.13/kWh for electricity, then over this lifetime I will spend on electricity:

$$1.5 \times 10^4\,\text{hr} \times 100\,\text{W} \times \frac{\$0.13}{\text{kWh}} = \$195 \qquad \text{(Incandescent)} \qquad (5.12)$$

$$1.5 \times 10^4\,\text{hr} \times 15\,\text{W} \times \frac{\$0.13}{\text{kWh}} = \$29.25 \qquad \text{(LED)} \qquad (5.13)$$

- The total cost for this life cycle is hence \$203 for the incandescent vs \$33 for the LED—*choosing the incandescent because it was \$1.20 cheaper at the store cost you \$170 over the next 10 years!*. Plus the LED would result in 6 times lower CO_2 emissions.

This is a "no-brainer" choice in favor of energy conservation and lower CO_2 emissions. The commerical world has essentially completely converted to fluorescents and LEDs from incandescents, and the residential customers are following, spurred on by US government regulations—in fact it is now illegal in the US to sell bulbs with as low efficiency as a standard incandescent, and you'd need a black market contact to purchase the above bulb, if you were so insistently foolish. The Trump administration, in September 2019, has announced it was cancel plans (dating from 2007) to extend the minimum-efficiency rules to many more types of light bulbs, thus managing to impede the impressive reductions in energy usage and CO_2 emission that the switch away from incandescent bulbs has already given us, while at the same time costing consumers a substantial sum of money. The Biden administration placed this action up for official review (along with many other energy-related regulations) on his first day in office. Many people dislike CFL bulbs because of their color or convenience issues (e.g. ex-President Trump complained that new bulbs made him look bad). But LED bulbs have improved rapidly, and do not have the cosmetic drawbacks of CFL's, so are driving both incandescent and fluorescent bulbs out of all except a few niche markets. Since some LED bulbs claim lifetimes up to 50,000 hours, you may have to explain to your children what it means to "change a light bulb"!

In class we will be able to observe the spectra of different light bulbs. We will see that CFLs and LEDs manage to look "white" even though they only emit only a few colors of light, not the whole visible spectrum. This is because we have just three different types of light receptors in our eyes, which are sensitive to the high, middle, and low-energy ranges of the visible spectrum. With the right mix of just red, green, and blue light you can get your retina to respond the same way as for any mixture of visible colors. This is why television and computer monitors are "RGB" displays that look true-color even though if you look closely you can see that they have only three different colors of dots.

Practice problems

- Which would you rather be exposed to: 10 J of microwaves, 1 MJ of visible light, or 1 Joule of x-rays?
- Find the cost of an LED bulb and a CFL bulb and conduct a life-cycle cost comparison.
- Can you think of any circumstances where an incandescent bulb would be the more economical choice?
- Explain why CFL and LED bulbs are even extra-economical if you're using them in summer.
- What survival strategies could you adopt if Earth's ozone layer were to be depleted?
- Do you think the US and other governments would agree to the Montreal Protocols if they were proposed today? What's the difference between solving the ozone/UV problem and solving the CO_2/IR climate-change problem?

6

Climate Change

The addition of CO_2 to Earth's atmosphere will decrease the radiation power escaping to space, disturbing the equilibrium between incoming and outgoing power and increasing the *net* flow of energy to Earth. This ***radiative forcing*** is a consequence of well-known laws of physics, and will result from any ***greenhouse gas (GHG)*** that has strong infrared spectral lines, including methane (CH_4), nitrous oxide (N_2O), ozone (O_3), CFCs, and even water vapor. The burning of fossil fuels releases CO_2 into the atmosphere, and other human activities affect the level of GHGs as well. Estimate as of 2011 are that human activity has caused $f = 2 \pm 1\,\mathrm{W\,m^{-2}}$ of radiative forcing flux in 2011 compared to a typical pre-industrial year (the forcing due to greenhouse gasses is certainly above this $2\,\mathrm{W\,m^{-2}}$ now) . In one year, the total extra energy on Earth is

$$
\begin{aligned}
E = fAt &= 4\pi R_E^2 f t && (6.1)\\
&= 4\pi \left(6.3 \times 10^6\,\mathrm{m}\right)^2 \left(2\,\mathrm{W\,m^{-2}}\right)(1\,\mathrm{yr}) && (6.2)\\
&= 3 \times 10^{22}\,\mathrm{J} = 30{,}000\,\mathrm{EJ!!} && (6.3)
\end{aligned}
$$

This is 50 times as much as all direct human energy consumption! We must ask: *what are the consequences of all this additional solar energy being accumulated at Earth's surface?* Energy is conserved, so it must go somewhere. Either it can be stored on Earth as thermal energy (primarily in the oceans); or it must be radiated back into space by raising surface temperatures to increase the strength of outgoing radiation. There is ample evidence that both are occurring.

If we think this global climate change is detrimental to our well being, now or in the future, then we should consider reducing or eliminating CO_2 emissions from fossil fuel burning by either cutting back their use, or capturing the CO_2 from combustion before it gets into the atmosphere. Before we make the huge investments that

Other activities besides fossil-fuel burning lead to radiative forcing, such as clearing of forests, or concrete manufacturing. In this course we'll confine our discussion to fossil-fuel burning.

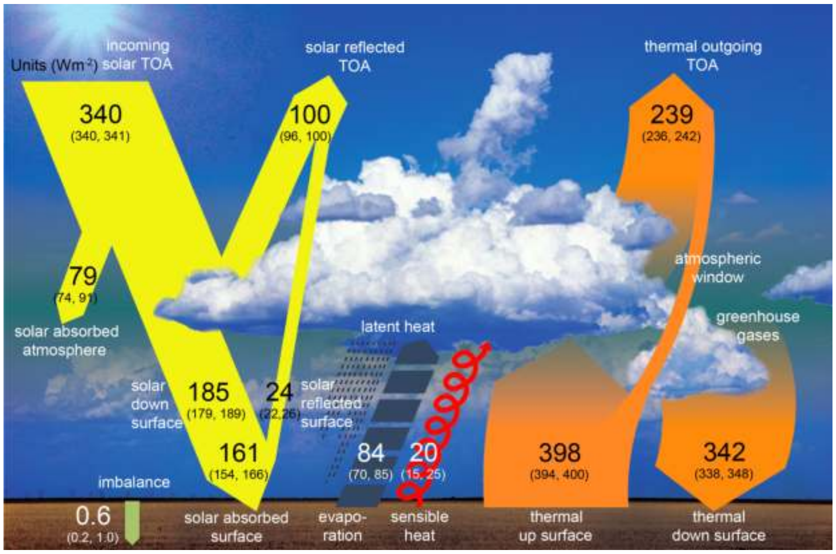


Figure 6.1: The flow of the radiation power to/from space. Note how much power the greenhouse gases block and send back to the surface. AR5 WG1 Figure 2.11.

would be required to reduce our need for fossil-fuel energy, we need to ask these questions:

1. Is human activity causing a significant rise in GHG levels in the atmosphere? In particular, is CO_2 rising due to fossil fuel use?
2. Are global temperatures rising?
3. Is the global temperature rise *caused* by the anthropogenic radiative forcing?
4. Would *reductions* in future emissions reduce future temperature increases or other detrimental effects of the radiative forcing?
5. Will it be less costly to reduce emissions than to live with the consequences of "business as usual" fossil-fuel usage? Here we must be careful to use a broad definition of "cost" to include costs borne by the entire population of Earth (human or otherwise), not just energy producers, including future residents. It can be very difficult to express all of these costs as dollar values.

What would you consider the dollar value of the Great Barrier Reef? The cost of submerging most of Florida?

We need to find an answer of "yes" to all of these questions to justify the expensive transition away from the fossil fuels that have made modern life possible (or to implement carbon capture).

While the basic physics of more $CO_2 \rightarrow$ more net solar heating is straightforward and undeniable, there are many processes on Earth that complicate prediction of the subsequent consequences. For example, higher CO_2 in the atmosphere will lead to more CO_2 being dissolved into the oceans, reducing the amount of radiative forcing,[1]

---

[1] Absorption of CO_2 into the oceans in turn causes ***acidification***, which we may find as harmful as the temperature rise it is ameliorating.

and higher temperatures might increase plant growth rates, lowering the CO_2 levels. These are examples of ***negative feedback***, whereby the consequences of a disturbance tend to reduce the original disturbance. A thermostat is a good example of negative feedback: if the temperature goes down, the heater is turned on, with the result that the room temperature is very stable. Negative feedbacks can reduce the impact of CO_2 emission on atmospheric GHG content and/or reduce the resultant surface temperature.

Conversely there are *positive feedback* processes: higher temperatures can mean less ice and snow on Earth, which means that shiny icy surfaces are replaced with darker, sun-absorbing ground. This leads to *higher* net energy input from the Sun. Or warming may thaw permafrosts, releasing methane that was locked into the soil. Methane is a highly potent GHG and this would speed up the warming. A very important positive feedback is that warmer air holds more water vapor, which is a very strong greenhouse gas (stronger than CO_2!). Positive feedbacks hence tend to cause highly unstable behavior that can run away and become irreversible even if the original disturbance is removed. In this case, "positive" is not a good thing!

Questions 1 and 2 can be answered relatively straightforwardly by careful observation of the present and past conditions on Earth. But the complexity of the carbon and climate feedbacks makes it difficult to answer questions 3 and 4. Realistic computer models of Earth's response can provide guidance. The last question is the most difficult to answer, not least because it requires us to define our values and costs, which is outside the realm of science—it is literally a value judgment. For example if you value the current year or financial quarter but have no concern for the state of the world a decade or a century hence, then you will always choose business as usual over decarbonization.

In this chapter we will rely heavily on information from the *Fifth Assessment Report (AR5)* of the Intergovernmental Panel on Climate Change (IPCC). These reviews of the state of climate, and our knowledge of it, have been issued by consensus of hundreds of the world's leading climate scientists and economists since concerns over global warming first became serious in the late 1980's. Two more recent assessments are the 2017 Fourth Climate Science Special Report (CSSR) of the *US Global Change Research Program* and Climate Change Evidence & Causes produced by the US National Science Foundation and UK Royal Society. You are highly advised to read the latter and the executive summary of the former!

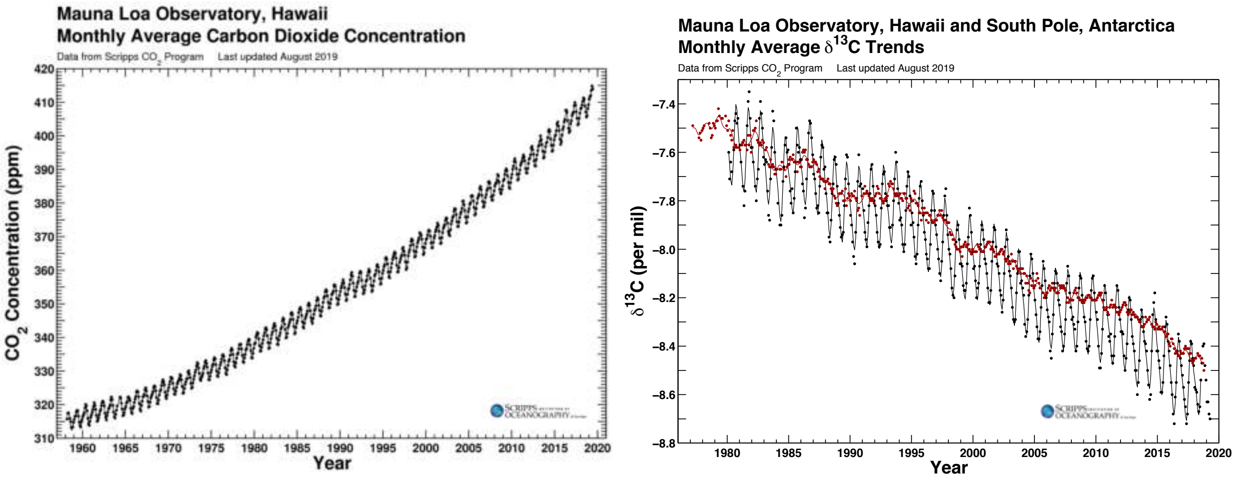


Figure 6.2: Left shows concentration of CO_2 in Earth's atmosphere as measured at Mauna Loa. Right shows the relative ratio of ^{13}C to ^{12}C in the atmosphere. Both from Scripps CO_2 Program.

Anthropogenic carbon

We can measure directly the concentration of CO_2 in Earth's atmosphere. These measurements have been made continuously since c. 1960, with the results shown in Figure 6.2. We can determine the history of the CO_2 content of Earth's atmosphere by measuring *fossil air.* In Greenland and the Antarctic, snow falls each year and never melts; the layers of snow are compressed into ice layers that can be counted down from the surface like tree rings. Each annual ice layer contains tiny bubbles of air trapped in the snow layer from that year. Incredibly, we can take direct measurements of CO_2 concentration going back *over 800,000 years!* The measurements show:

- The average CO_2 concentration in 2018 was 408 parts per million (ppm),[2] meaning that out of every million molecules in the atmosphere, 408 of them are CO_2. This has risen rapidly from 315 ppm since direct monitoring started in 1958.
- Between the end of the last ice age, $\approx 10,000$ yr ago, and the start of the industrial revolution c. 1800, the CO_2 level in the atmosphere was steadily between 260–280 ppm. *The 50% rise in* CO_2 *level since industrialization is more rapid than at any time since the end of the ice age, i.e. since human civilization has existed.*
- The CO_2 level has been moving between about 175–280 ppm over the past 100,000 years, as Earth moves between ice ages and "interglacial" periods. *The current* CO_2 *level is substantially higher than at any time in the past 800,000 years.*
- Geological evidence (not shown) demonstrates that Earth's CO_2 level and temperatures have been well above current levels at various epochs in the distant past, for instance in the "carboniferous"

---

[2] Trends in Atmospheric CO_2, from NOAA.

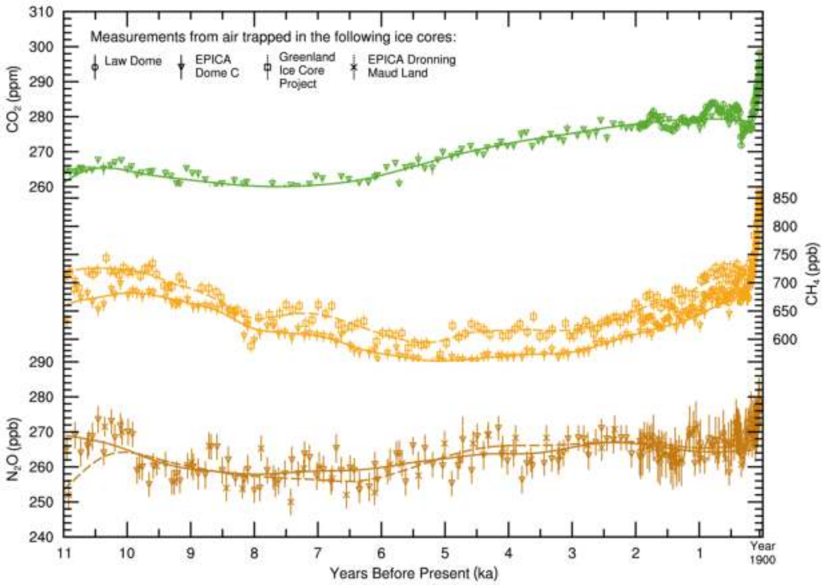


Figure 6.3: The atmospheric levels of CO_2, methane, and nitrous oxide as determined from ice core samples since the end of the last ice age 11,000 years ago. Note that the present-day value of 407 ppm CO_2 is far above the top of the plot. From AR5 WG1 Figure 6.6.

era more than 100 million years ago. This was the time when most of plants grew that now make up our fossil fuels. This of course is well before our species walked the Earth.

The data leave no doubt that levels of CO_2 (plus methane and other GHGs) have risen rapidly in the past century to levels well above any time in the past 10,000 years. Can we attribute this to fossil-fuel burning, or is this just some random fluctuation in CO_2 levels attributable to some non-human processes? We can argue first that a random process would have only 1% chance of showing up during the same century, out of the 100 centuries since the ice sheets receded, as the era when we burned all the fossil fuels. It would be especially odd that such a rapid rise would occur when no rises even close to this large have ever occurred.

There are however two other lines of argument that lead inexorably to the conclusion that fossil-fuel burning is primarily responsible for the CO_2 increase since industrialization. There are three ***isotopes*** of carbon: 99% of carbon is ^{12}C with 6 protons and 6 neutrons; 1% is ^{13}C with 7 neutrons; and 1 in a trillion carbon atoms in the atmosphere are ^{14}C with 8 neutrons.[3] The photosynthesis process has a very slight preference for absorbing ^{12}C over ^{13}C, so that the carbon making up plants (and animals which eat the plants) have a slightly lower ratio of ^{13}C to ^{12}C than the CO_2 in the air. Figure 6.2 shows that the fraction of ^{13}C in the air's CO_2 has been dropping, which tells us that the new CO_2 being added to the atmosphere *must*

[3] The ^{14}C is radioactive and will decay into a different nucleus in a few thousand years, so it is absent from fossil fuels that have been buried for millions of years.

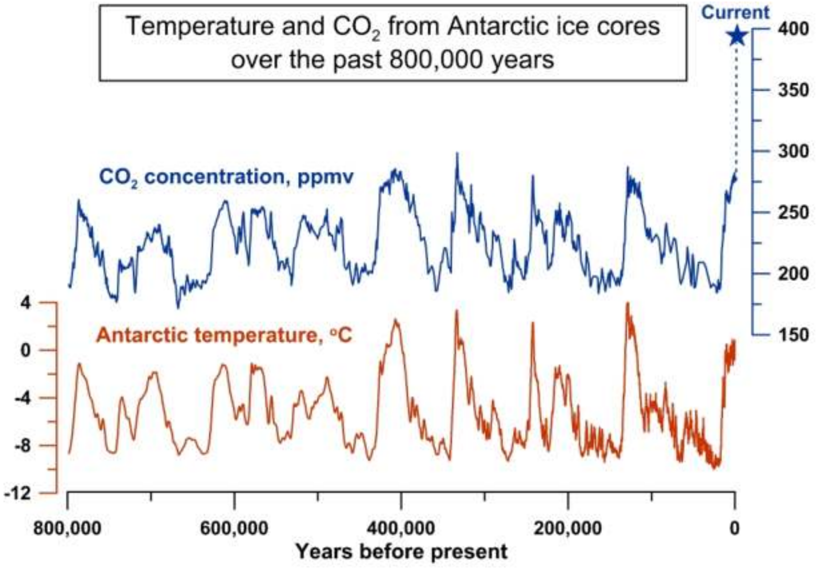


Figure 6.4: The CO_2 concentration of the atmosphere (blue) and relative global mean temperature (orange) for the past 800,000 years as derived from Antarctic ice-core samples obtained by the *EPICA* collaboration. Figure from this site.

be coming predominantly from the remains of plants. This of course is exactly what fossil fuels are.

Another test we can apply to the fossil-fuel theory is to ask whether the amount of carbon being added to the atmosphere since industrialization is similar to the amount of carbon in all the fossil fuels that have been burned. For this test we need to know the mass of carbon it takes to raise the atmospheric CO_2 by 1 ppm. We start with the knowledge that the total mass of Earth's atmosphere is about 5.1×10^{18} kg.[4] We also need to know that the average carbon atom has a mass of 12 *atomic mass units (amu)*, while an average nitrogen and oxygen atom have masses of 14 and 16 amu, respectively. About 3/4 of Earth's air molecules are N_2, with mass of 28, and nearly all the other 1/4 are O_2, with mass of 32, so the average molecule has mass 29 amu. If one molecule per million is CO_2, the mass in the carbon is

$$1\,\text{ppm} = 10^{-6} \times 5.1 \times 10^{18}\,\text{kg} \times \frac{12\,\text{amu}}{29\,\text{amu}} = \underline{\underline{2.2 \times 10^{12}\,\text{kg C} = 2.2\,\text{GTC.}}} \tag{6.4}$$

In the past 5 years the atmospheric CO_2 level has risen by 2.4 ppm per year. We conclude that an average of **5.3 GTC** are being added to the atmosphere each year. Now we can take the global fossil-fuel usage statistics for each fuel, multiply them by the amount of carbon released when burning that fuel, and obtain the total carbon *emissions.* In 2018, worldwide carbon emissions were **9.2 GTC.** [5] We find that *fossil fuel emissions of carbon are sufficient to explain the increase in carbon in the atmosphere.* In fact we now have a puzzle in that *nearly*

[4] You can figure this out if you know that the atmospheric pressure of 15 pounds per square inch arises because there are 15 pounds of air above every square inch of Earth's surface.

GTC stands for "gigatonnes carbon," where 1 tonne is a metric ton of 1000 kg. It is important to know that the mass of the CO_2 *molecule* is 44 amu, but only 12 amu are in the carbon. So the total mass of the *carbon dioxide* is $44/12 = 3.67\times$ larger than the mass of the carbon. Some publications will give you the GTC, others will give GT of CO_2, so be careful. You may also see the unit of Petagrams carbon (PgC); 1 PgC=1 GTC.

[5] From *BP Statistical Review of World Energy 2019*.

half the carbon emissions are disappearing from the atmosphere. The AR5 WG1 report[6] estimates that the total increase in atmospheric CO_2 from 1750–2011 is 240 GTC. The world emissions from fossil fuels over that time is estimated at 375 GTC,[7] so 36% of the carbon emitted has disappeared from the atmosphere.

The carbon cycle

The atmosphere and fossil fuels are just two of the ***reservoirs*** of carbon on Earth. Figure 6.5 illustrates that quantities of carbon equal to or vastly exceeding the amount in the atmosphere exist in the vegetation, soils, oceans, and rocks of Earth's surface layers. These reservoirs are constantly exchanging carbon quantities well in excess of the fossil-fuel-burning outputs. But over 10,000 years the carbon flows have been in a near balance that has been disturbed by our relatively rapid release of the fossil fuel reservoir of carbon. Since we've released an amount of carbon nearly equal to the entire carbon content of the atmosphere, the concentration has risen sharply. If, however, this excess carbon were spread evenly through the much larger ocean and geological carbon reservoirs, it would represent a small change. Eventually the fossil-fuel carbon will flow to these other reservoirs and atmospheric levels will drop. We will want to know the ***characteristic timescales*** for flows between the reservoirs, to get an idea of how long the anthropogenic GHGs will stay in the atmosphere and influence infrared cooling.

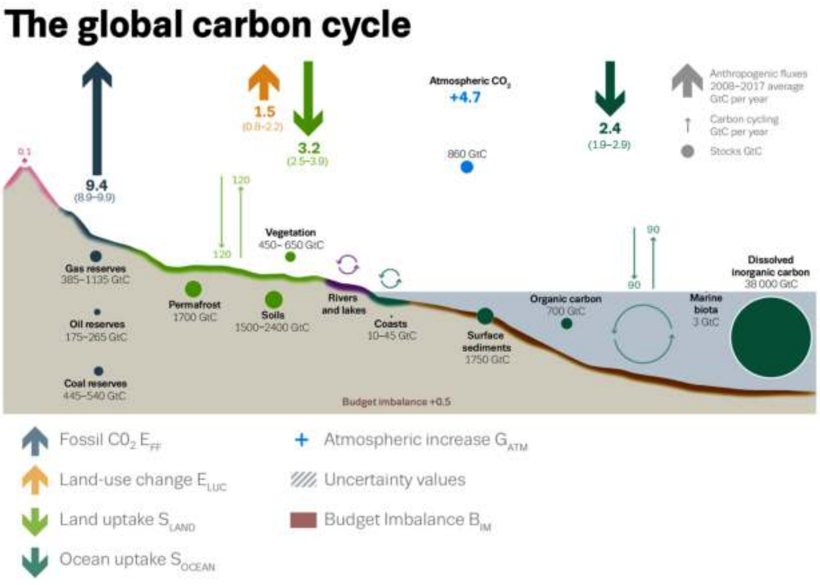


Figure 6.5: The carbon reservoirs in Earth's atmosphere and surface are given in GTC and as the size of the circles, with recent annual changes to/from the atmoshere given in the bold arrows. The thin arrows label the natural flow rates in GTC per year. From Le Quéré et al. 2018.

There are numerous flows and feedbacks in the Earth's carbon

---

[6] Section 6.3.2.3

[7] Table 6.1 of AR5 WG1.

system, and it is not simple to predict what will happen to carbon emitted into the atmosphere. Important reservoirs are:

- *Vegetation,* and the organic materials in *soils,* hold several thousand GTC of carbon. The characteristic timescale for exchange of carbon between the atmosphere and vegetation is years to decades—think of the time it takes for a plant to live and die and decay.

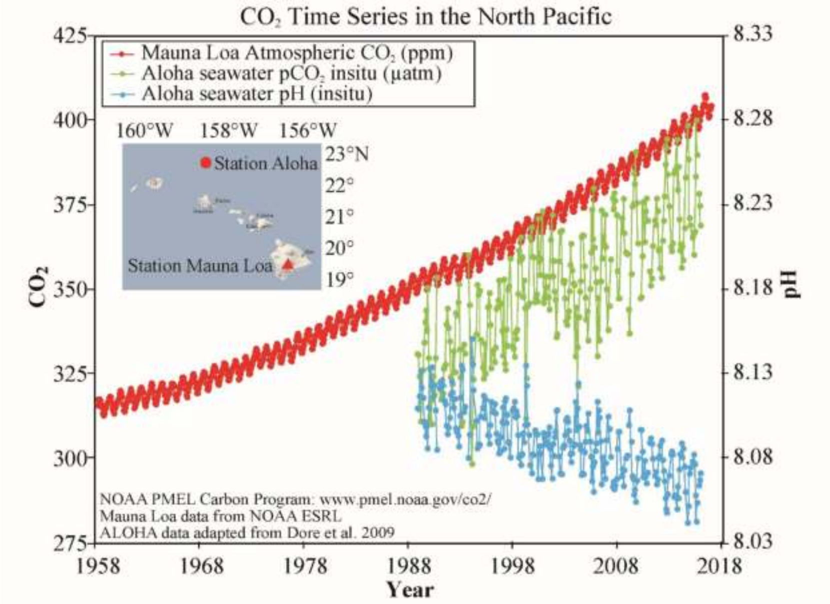


Figure 6.6: Measured Pacific Ocean CO_2 content increase and pH decrease shown alongside atmospheric CO_2increase. From Unesco

- The *oceans* contain nearly 40,000 GTC of dissolved CO_2, dwarfing the amount in the atmosphere. Increasing the atmospheric concentration of CO_2 causes more to be dissolved into the water. The surface ocean will re-equilibrate with the air on a timescale of decades, but it takes 10^3 years for the surface ocean water to fully mix with the deep oceans and equilibrate the CO_2 levels. As water dissolves CO_2, it become more acidic. Figure 6.6 shows measurements of average ocean pH over the past 20 years, confirming the expected drop as excess CO_2 dissolves into the oceans.

- *Seafloor sediments* are formed when CO_2 dissolved in the oceans reacts to form calcium carbonate ($CaCO_3$) that sinks to the ocean floors. Shellfish and corals, for example, build their shells from $CaCO_3$. It takes thousands of years for this process to exchange carbon between the oceans and the seafloor sediments. These sediments, when compressed, form limestone.

- *Geologic* processes that take millions of years can cycle the sediments into limestone and silicate rocks that cycle through Earth's crust and undergo weathering at the surface and even processing through volcanoes.

Computer models of the carbon cycle suggest that, of carbon emitted into the atmosphere today, roughly 60% will be removed within 100 years through uptake into vegetation and surface oceans.[8] But the other 40% will disappear only over thousands of years as circulation into the deep ocean and sediments is much slower. It is estimated to take over 10,000 years to remove $> 90\%$ of the carbon that we are emitting now. Hence current fossil-fuel burning has very long-term impacts on the greenhouse effect.

It is believed that geological feedbacks have helped to stabilize Earth's temperature over the billions of years it has existed. But it can take hundreds of millions of years for these very slow adjustments to occur after a disturbance. It is interesting to consider the case of Venus: although it was probably created with similar carbon content as Earth, it has no oceans, and the carbon that is in Earth's oceans and sediments forms a very thick CO_2 atmosphere on Venus instead. This causes an enormous greenhouse effect: the surface of Venus is hot enough to melt lead!

The annual releases of CO_2 due to fossil-fuel burning and land-use changes are shown above the zero line in Figure 6.7. Below the zero line are shown the estimated destinations for these carbon releases: absorbed in the oceans, on land, or left in the atmosphere.

Other GHGs

Other greenhouse gases have their own reservoirs and transitions in Earth's surface layers. The gas having the greatest impact on Earth's infrared emissions is water vapor, not CO_2. Not only do H_2O molecules have strong IR spectral lines, but water vapor condenses into droplets to make clouds, and the clouds reflect sunlight back to space and hence have a cooling effect. It would be impossible to understand climate without accounting for water vapor.

Water does not appear in the charts of GHG given by the AR5 and other climate studies. The reason is that its cycle time is very short—water emitted into the atmosphere comes back as rain within a few days. The amount of water in the atmosphere is therefore not determined by its emission rate; climate scientists treat water as part of the *feedback* processes, not as part of the *forcing.* Understanding how water vapor and clouds respond to the radiative forcing of other GHGs is one of the difficulties of climate science.

After water and CO_2, methane is the strongest GHG. Per molecule, methane is a much stronger infrared blocker than CO_2. But CH_4 is less chemically stable than CO_2 and is removed from the atmo-

[8] AR5 WG1 Box 6.1.

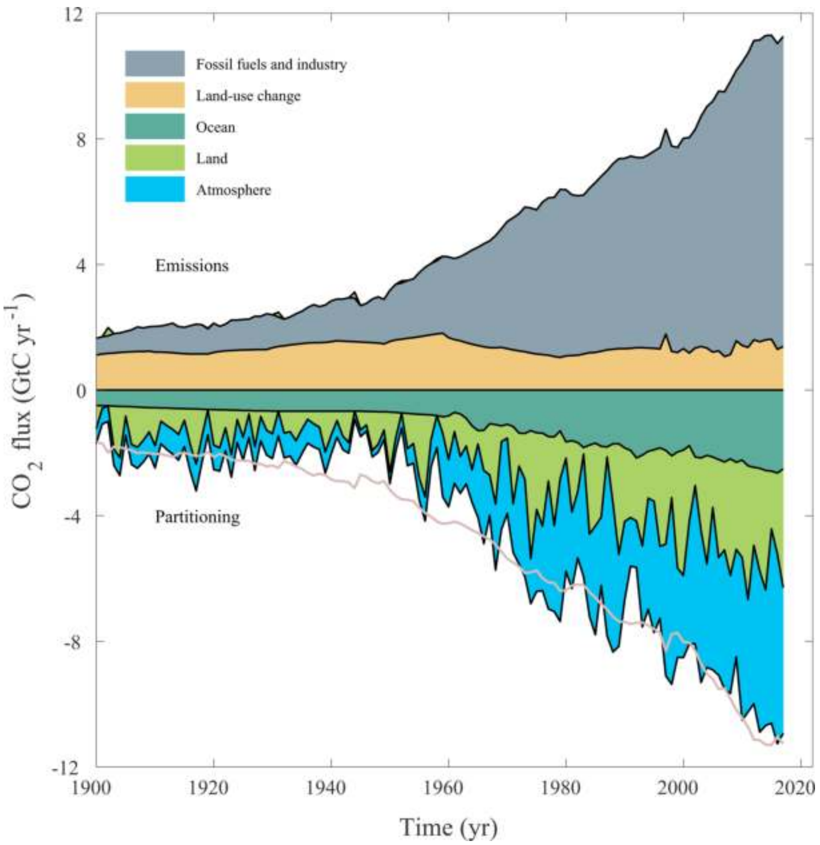


Figure 6.7: History of annual emission and absorption of CO_2 since 1870, from Le Quéré et al. (2018).

sphere in about 10 years, as opposed to the 100–10,000-yr lifetime of CO_2. The *global warming potential (GWP)* or total greenhouse effect of a molecule of methane compared to CO_2 therefore depends on how far into the future you will continue to count the damage. If we count 20 years out, methane is 86× more damaging than CO_2 on a per-molecule basis. On 100-year time scales, this value drops to 34.[9] Either way, *it is clearly much better for the environment to burn methane than to let it escape into the atmosphere.* CFCs have global warming potentials over 5000 times CO_2. The CFCs released before full implementation of the Montreal Protocols will continue to contribute to the greenhouse effect as well as ozone depletion as they slowly decline over approximately 100 years.

Figure 6.8 summarizes the estimates of the total change in radiation balance between Earth and space due to human activity since 1750—including albedo changes as well as GHG effects. At present it is estimated that between 1.1 and 3.3 W m^{-2} of additional power were being absorbed by Earth in 2011.

---

[9] According to AR5.

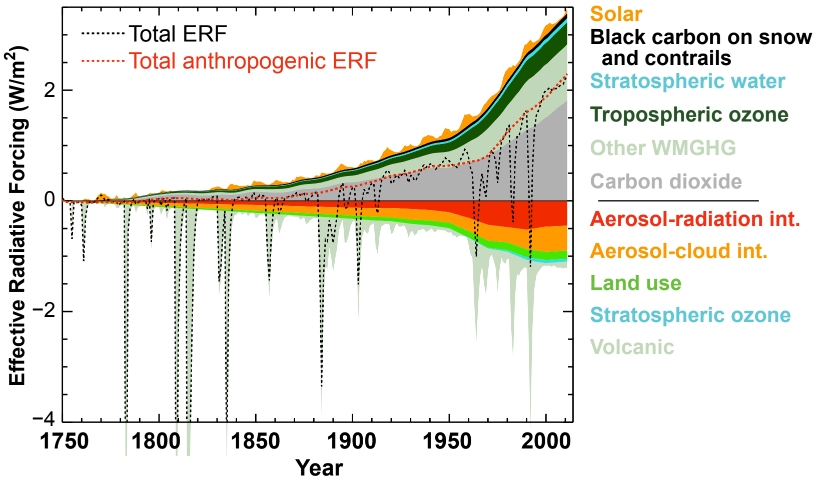


Figure 6.8: History of changes in radiative forcing since 1750, from CSSR Figure 2.6. The heating effects are above the zero line, the cooling effects are below, and the net change is the dashed black line. The red dashed line is the anthropogenic part.

The climate record

Has the climate of Earth changed measurably as the atmospheric CO_2 levels have climbed 40% above pre-industrial values? This is again easy to answer, in principle, by simply measuring such things as the global temperature, sea level, ocean acidity, and ice cover. The complicating factor is that *all of these quantities are changing constantly due to purely natural phenomena.* We must distinguish between ***weather*** and ***climate.*** Weather is the daily, variable conditions at particular locations on Earth. It is well known that weather is variable and chaotic, meaning that it is very hard to predict the weather more than a few days in advance at any given site. Because there is so much natural variation in the weather, it is unwise to attribute single events in specific locales, or a single hot or cold year, to anthropogenic effects.

Climate, on the other hand, is the long-term *average* or trend of the weather. The longer the span of time, and the larger the geographic area over which we average the weather, the more predictable it becomes. Even the global mean temperature has random year-to-year variation atop long-term trends, particularly from the El Niño phenomenon in the Pacific. Some variation occurs, even in the decade-by-decade mean temperatures, without any identifiable forcings. We can track the anthropogenic forcings; and there are also natural forcings, such as volcanic eruptions (which put aerosols into the upper atmosphere that tend to reflect sunlight into space and cool Earth) and small changes in the brightness of the Sun or Earth's distance from the Sun (which can be tracked with astronomical data).

With these caveats in mind, let us examine the climate record for evidence of changes since industrialization that are in excess of the variations that occur naturally. We begin with Figure 6.9, giving the mean global surface temperature (including land *and* the oceans that cover 75% of the globe!) since 1880. This is about the time when thermometers had sufficient accuracy and distribution around the world to make a meaningful global average. We can see that

- **The past decade is the warmest in the past 130 or more years, with a mean temperature roughly** 1.0° **C (**1.8° **F) warmer than the 1880–1920 average,** with $0.017^\circ\,\mathrm{C\ yr}^{-1}$ average warming rate for the past 50 years.
- The warmest 20 years on record all occurred in the past 22 years!

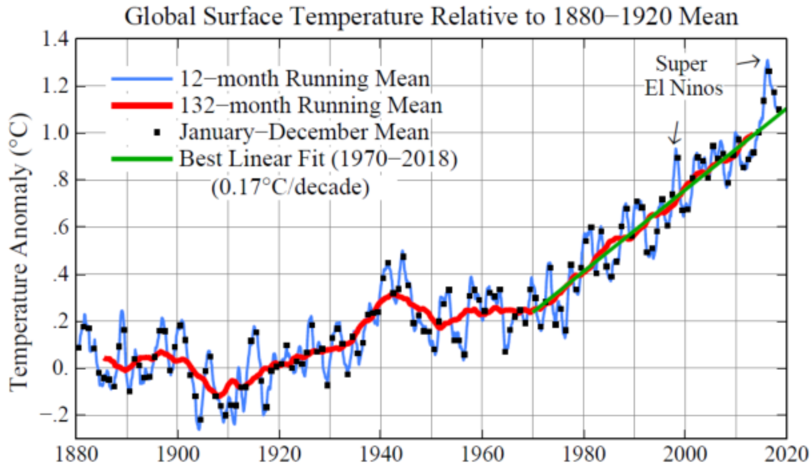


Figure 6.9: Mean annual global surface temperature from instrumental records, 1880–2018, relative to the 1880–1920 mean. From Columbia Universeity/GISS.

As a good example of the difference between weather and climate, consider that may climate-change deniers were claiming that global warming stopped after 1998, because if one looks *only* at the years 1998–2013, it is difficult to see an upward trend. This argument was, not surprisingly, defective in several respects. First, we know that year-to-year natural variations in mean temperature will be caused by weather. The El Niño phenomenon, when there is a warmer southern Pacific ocean that releases more heat into the atmosphere, is the largest source of fluctuations, and 1998 was a record-strong El Niño year. If you start your test on a very high weather fluctuation, then you will reduce the trend you see in the following years. This kind of "cherry-picking" of data with random fluctuations causes biases in your results.

Second, we have to recognize that the "noise" of weather fluctuations can obscure long-term climate trends for quite a few years. A decade of flat global temperatures no more disproves global warming science than a few good spins on the roulette wheel disproves the casino's long-run advantage at the game. We need to look at decades-long records to discern climatic trends.

Third, the "hiatus" decade does not change the laws of physics: the CO_2 in the atmosphere must produce additional net radiation forcing, and this additional energy must go somewhere. Measurements of ocean temperatures, (Figure 6.10) show that the heat content of the *ocean* continued to rise steadily through this decade. Because the specific heat of the ocean is so much larger than that of the atmosphere, a fluctuation in ocean/air transfer can slow the atmosphere's warming, but not the ocean's.

Indeed the El Niño returned, and the five years 2014–2018 have been the five warmest on record, with 2016 a full 0.3°C hotter than the previous record. 2017 was the warmest La Niña year on record (opposite of El Niño). The warming trending has most definitely not stopped!

Figure 6.10 shows the ocean gaining about 10^{22} J of heat per year. This is similar to (but short of) the amount of excess energy arriving at Earth from the extra radiative forcing in Equation (6.3).

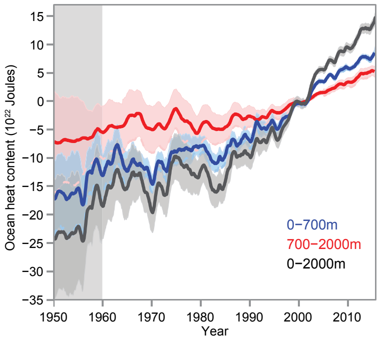


Figure 6.10: The estimated changes in the heat content of the oceans in the past 65 years. CSSR Figure 13.1.

Is this century's warming unique? While direct temperature measurements are not reliably available prior to c. 1850, we can reconstruct historical temperatures using ***proxy*** data that are influenced by temperatures in past eras, such as tree rings and corals. Figure 6.11 summarizes what is know about the past millenium's global mean temperatures. The width of the gray band indicates that there is of course substantially larger uncertainty in the proxy data than in the direct measurements. Nonetheless it is clear that **the most recent decade(s) are warmer than any in the past millenium, with a low probability exception of a period during the "medieval climate anomaly."** AR5 reports that some regions appear to have been as warm at that time as today; but unlike the present warming trend, it did not span the whole globe—the South did not experience the same level of medieval warming.

As with the CO_2 record, we can reconstruct an 800,000-year temperature history using ice cores. Most of the oxygen on Earth is ^{16}O, with 8 protons and 8 neutrons. A small fraction is ^{18}O with 10 neutrons. Water containing the heavier oxygen does not evaporate as readily so rain water is a little bit lighter than the ocean water. When the Earth is cold and more light oxygen gets locked into ice sheets, the oceans and the rain and snow get a little richer in ^{18}O. The ^{18}O content of each year's snowfall found in the ice cores is therefore a good indicator of global temperature. The lower plot of Figure 6.4 shows the nearly million-year temperature history from Antarctic ice cores. We note:

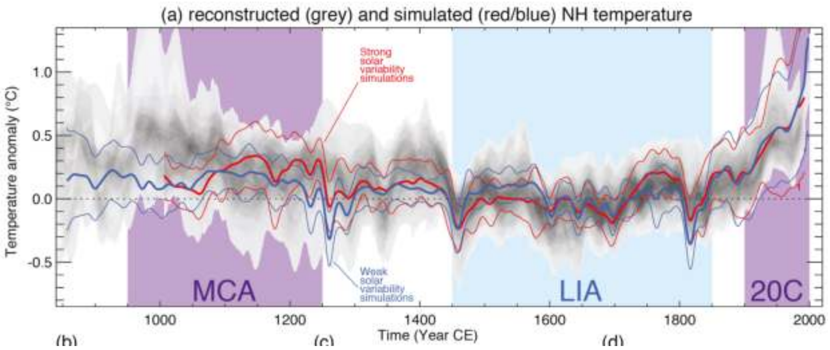


Figure 6.11: Gray band shows reconstructed mean *Northern Hemisphere* surface temperatures over the past millenium using various proxies. The most likely reconstructions are the darker grays. The medieval climate anomaly (MCA, a.k.a. medieval warm period), the little ice age (LIA), and the 20th century (20C) are marked. The thin red lines contain 90% of the results of running computer models on past climate given what is known of the histories of Earth's orbit, volcanic explosions, etc., with the thick red line showing the median model. Blue lines are another set of models. All data are expressed as anomalies from their 1500–1850 mean and smoothed with a 30-year filter. Note that the anomaly today is estimated at 0.8° C above the 1850–1900 period. From AR5 WG1, Fig 5.8.

- The ice age cycles of Earth have typical temperature fluctuations of $\approx 8^\circ$ C associated with them, which are largely in synchrony with variation in the CO_2 content of the atmosphere.
- Some of the interglacial periods may have been warmer than the average of the current interglacial "holocene" epoch of 10,000 years since the end of the last ice age.
- The temperature record shows no excursions more than $\approx 3^\circ$ C *above* the mean holocene temperature.

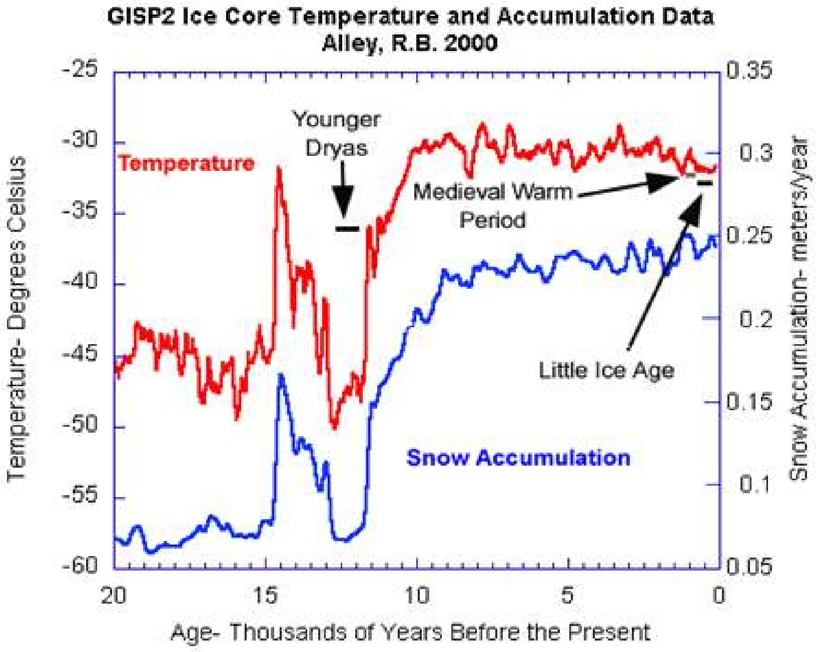


Figure 6.12: Greenland ice core ^{18}O concentrations converted into temperature records for the past 20,000 years. Notice the extremely large and rapid temperature changes occuring toward the end of the ice ages, and the very stable climate characterizing the current interglacial period. From R. B. Alley 2000, *Quaternary Science Reviews*, **19** 213.

Figure 6.12 gives a closer look at the period since the late days of the last ice age. What we see is that the past 10,000 years' climate—

the entire period when human civilization arose—has been *extremely stable compared to the ice ages, or even to previous interglacial periods.* Human agriculture, for instance, has never faced rapid global climate change. A temperature rise of 4 K as might occur during our century would be unprecedented in the post-ice-age ice core record.

We can conclude that global surface temperatures have been climbing nearly every decade since industrialization, and mean temperature are certainly warmer in this decade than any other in ≈ 150 years, and probably warmer any other in 1000 years. The climate has shown intense variation at times earlier than the past 10,000 years, with 8 K being the difference between an ice age and current conditions. The climate has not been more than a few K warmer than current conditions in the past 800,000 years.

Other indicators

There is good evidence of current climate change beyond the mean surface temperature. Figure 6.13 shows that both heat content of the oceans and the mean sea level have been rising for as long as good records are available—though we do not have clear data for thousands of years as we do for temperature, so it is harder to place the current changes into long-term historical context.

Sea level rise is caused by (a) the melting of ice that sits on land (melting of floating ice such as the Arctic sea does not change sea level). The largest ice masses are on Greenland and Antarctica, which if melted entirely would raise sea levels by 7 meters and 61 meters, respectively. But as important to the current sea level rise of $\approx 3\,\mathrm{mm\ yr}^{-1}$ is probably (b) the expansion of water volume when it warms. A large fraction of Earth's population lives within a few meters of sea level, so this is potentially a very damaging component of climate change, especially as the highest water level of severe storms increases.

There are numerous other clear indications of climate change over the past 10–100 years: reductions in the polar ice coverage, recession of glaciers in temperature regions worldwide, and shifts in the ranges of species and weather patterns.

Climate modeling

It is much easier to establish that global temperatures are rising than to establish that the rise is *caused* by anthropogenic forcing. It could just be coincidence that large CO_2 emissions occured in the same half-century when global mean temperatures rose most quickly. There is after all a 1-in-13 chance that the past decade would have

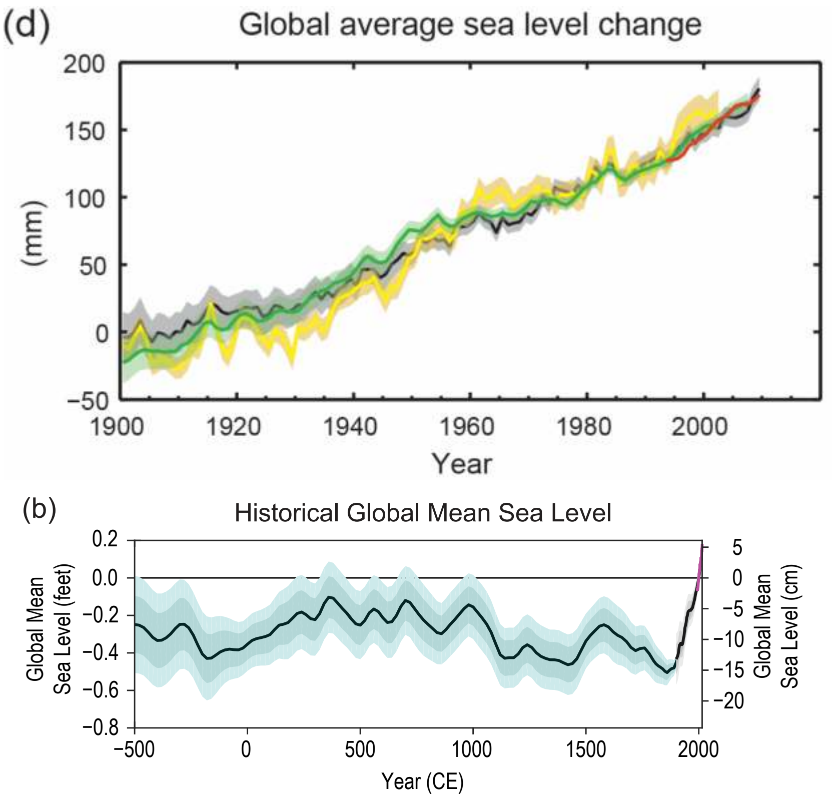


Figure 6.13: The estimated change in mean sea level in the past century (top, from AR5 WG1 Figure SPM.3) and past 2500 years (from CSSR Figure 12.2).

randomly ended up the warmest since 1880. The coincidence becomes more farfetched however, if we believe that current decade is the warmest in 1000 or more years ($<$ 1% chance of being random). Nonetheless we would like to find evidence of causation that is deeper than the rarity of a natural temperature fluctuations as high as the current one.

The classic scientific way to answer this would be to conduct an experiment whereby we make two Earths, identical except that one of them has anthropogenic carbon emission while the other does not. We then watch each of them for a few centuries and see how their temperatures evolve. In fact it would be better to have a few dozen or hundred of Earths because we know there is a random component to temperatures, so we would like to find the *typical* behavior in each case and not be fooled by an unusual string of weather on one of our test Earths.

Since we do not have the time or money to conduct such an experiment, we need another approach: we can produce *virtual* Earths by creating ***global circulation models (GCMs)*** which attempt to calculate the evolution of the atmosphere, oceans, ice, and vegetation under the influence specified inputs of GHGs or radiative forcings. As Fig-

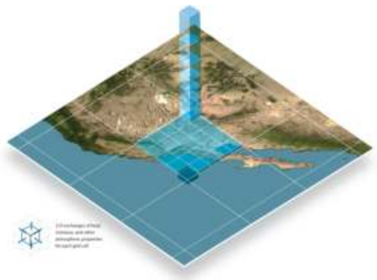


Figure 6.14: Illustration of how a climate model is constructed. From National Center for Atmospheric Research.

ure 6.14 illustrates, these models work by dividing Earth's surface into a 2-dimensional grid. Above each grid square we divide the atmosphere into a set of cells, and we keep track of the state of the air in each cell: its temperature, pressure, wind speed, water/cloud content, upward and downward radiation, etc. The laws of physics can be used to determine how each cell's contents will evolve in time, *e.g.* the pressure will set the air into motion, energy must be conserved as the temperature and radiation change, etc. Each land-surface cell is also described by some numbers giving its temperature, vegetation content, albedo, etc. We can make some educated guesses about how the vegetation will respond to the atmospheric conditions it sees. For grid squares over water we can create a downward stack of cells describing the ocean at various depths—its temperature, salinity, dissolved CO_2 level, flow rates, etc. Again the laws of physics and chemistry, and some biological guesswork, create equations that specify the evolution of each cell in response to its surrounding cells. Complex models can also be used to track Earth's flows of carbon and other GHGs.

Clearly we can make more and more detailed models as the power of our computers increase. We can never build a computer that tracks every plant, air molecule, and photon on Earth, so every GCM must make some approximations about the behavior of constituents too small to be explicitly tracked. Furthermore we do not fully understand some things like cloud formation and vegetation response, so we should never take the results of these models as gospel. "Garbage in, garbage out" is a phrase to keep in mind.

We therefore must ***validate*** the climate models. How can we build trust in the models? One good way to do this is to ask *whether the models correctly "predict" past climate behavior if they are given the past radiative forcings as input.* We can do detailed checks against the real Earth, checking that climate on different continents and layers of ocean and atmosphere are consistent with real-life observations, particularly over the past few decades when we have measured Earth's properties in much more detail. We can also see if models of the carbon cycle can reproduce *e.g.* the fraction of emitted CO_2 that remains in the atmosphere.

Once we have a computer model we trust, we can perform the experiment of turning off the anthropogenic radiative forcings, to establish causality. Figure 6.15 shows the results of this exercise in the AR5. Many climate models were constructed by different labs, using different formulae for the "subgrid" phenomena, and the red bands show the range of results that they give when fed the radiative forcings for the past 50 years or so. This includes GHG emissions but also natural events like volcanic activity and solar fluctuations. Re-

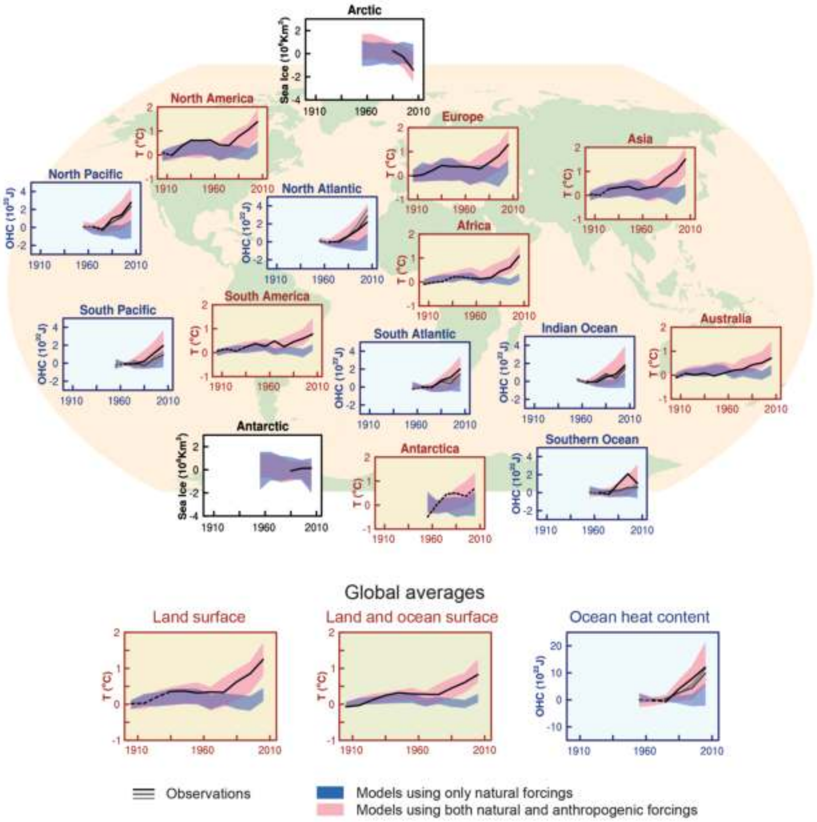


Figure 6.15: Comparison of the mean surface temperatures "predicted" by climate models run on the past 50 years for each continent and ocean, and for global averages. In each case the blue region is the range of models in which anthropogenic forcing is omitted; the red range includes anthropogenic and natural changes; and the black line is the actual recorded temperature history. AR5 WG1 Figure SPM6.

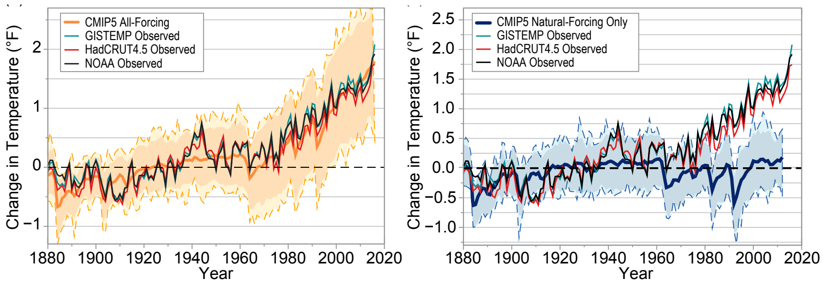


Figure 6.16: At left, a comparison of three estimates of the true global mean surface temperature (green, magenta, and black lines) to the "postdictions" of the CMIP5 climate model (orange line, with the orange band showing the range of uncertainty). The model matches reality well. At right, the same model is run (blue) with the anthropogenic forcings turned off. The last 50 years' temperature rise is not present in this case.

member that the models differ, but also we expect random variations due to weather that will be manifested in the models as well.

The black lines show what really happened: in every continent and ocean and ice cap, reality runs right through the middle of the GCM outputs, giving us confidence that these models have useful power in simulating Earth's climate.

The blue bands in the diagram are the results of running the models with anthropogenic forcings turned off. **For most of the globe, the real climate records of the past 50 years are inconsistent with the natural-fluctuation-only models.** Note at center bottom that the 0.8 K rise in global mean temperature is absent in the natural-only Earth. This is our best evidence that our GHG emissions are the cause of the climate changes observed to date. Figure 6.16 shows the this result for newer data and models.

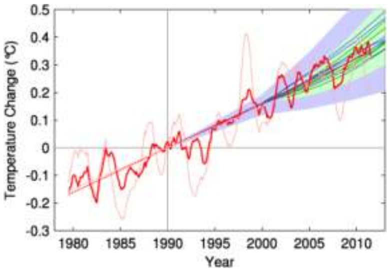


Figure 6.17: The blue and green wedges show the range of climate forecasts issued in the Third and Fourth Assessment Reports, respectively. The forecasts begin as early as 1990. The red line shows the actual global mean temperature record (in red) after corrections for volcanic eruptions and El Niño events. Good agreement is seen after adjustment for short-term natural phenomena. From Rahmstorf, Foster, & Cazenave, 2012, *Environ. Res. Lett.* **7**, 044035.

It is easier to predict the past than to predict the future! The real test of the GCMs is whether they can predict climate that has not occurred yet. Figure 6.17 shows the result of an analysis of whether the climate models featured in the Third and Fourth Assessment reports (which were predicting climate in years 1990 and later) correctly predicted the behavior to 2012. The authors of this study adjusted the real-life climate behavior over these 20 years for natural fluctuations (such as El Niños) that the original predictions could not anticipate. With these adjustments, the real climate (heavy red line) wanders around the middle of the blue wedge of predictions made in AR3. Likewise AR4 models cluster around the actual behavior post-2000. While this is a limited period of time for tests, it is good to see that even earlier, less powerful GCMs were on the right track.

Future scenarios

With believable global climate models in hand, we can estimate what future climate will be. We need to first assume a scenario for future CO_2 emission (and other GHGs). The AR5 adopts a series of *representative concentration pathways (RCPs)* which are illustrated in Figure 6.18. In Table 6.1 we summarize the properties of and predictions from these RCPs. The most optimistic (RCP2.6) assumes that we have now reached the peak of CO_2 emission at ≈ 9 GTC/yr, and that emission of carbon is essentially eliminated by 2070. In this scenario atmospheric CO_2 concentration leavels off at about 500 ppm very soon and begins a slow decline. The most pessimistic RCP8.5 assumes continued growth in global CO_2 emission at about 2% per year until 2060, rising through the end of the century to over 25 GTC/yr and pushing the CO_2 concentration to 900 ppm at century's end, more than triple the pre-industrial level.

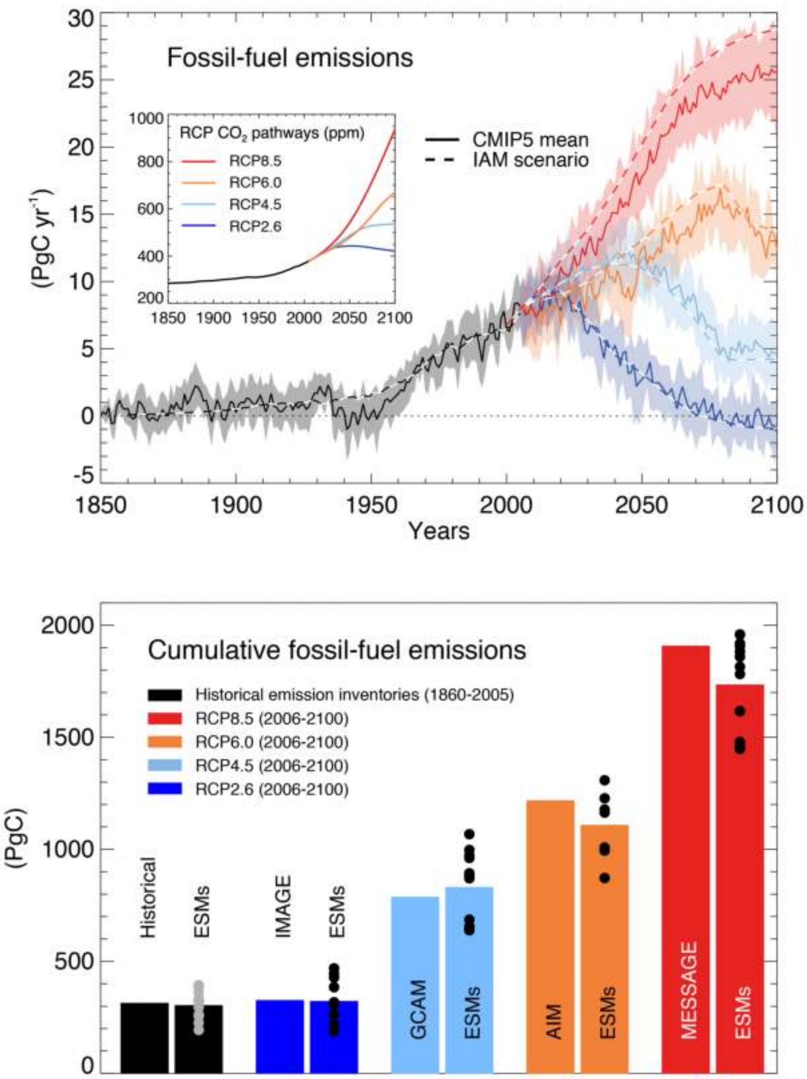


Figure 6.18: The four "representative concentratin pathways (RCPs)" modeled by the IPCC AR5. At the top inset are the atmospheric CO_2 concentrations assumed through the end of the century, and the emission rate per year that would be required to produce these CO_2 levels. At bottom are the total amounts of carbon emitted through 2005, and then the amounts emitted from 2006–2100 in each RCP. The best scenario has emission rates declining sharply starting in 2020. The worst RCP has CO_2 emission continuing to grow at 2% per year through 2060 and then levelling off.

The lower panel of Figure 6.18 gives estimates of the total cumulative carbon emissions from 2006–2100 that it would take to put Earth onto each of the RCP trajectories of atmospheric CO_2. To stay in the optimistic RCP2.6, the total carbon emission for 2012–2100 must be 270 GTC, well *below* the estimated 375 GTC emitted through 2011.[10]. Indeed by emitting $\approx$ 9 GTC/yr for 2013–2019, we've already emitted another ≈ 55 GTC, using up 20% of that allowance in just 6 years.

The outputs of the climate models for each of the RCPs are summarized in Figure 6.19 and Table 6.1. **It is clearly true that future CO_2 emission will enormously influence future climate change.** Other noteworthy results:

---

[10] AR5 WG1 Table 6.1

Scenario	RCP2.6	RCP4.5	RCP6.0	RCP8.5
Carbon emission				
Total 2012–2100 (GTC)	270	780	1060	1685
Peak rate (GTC/yr)	9	12	15	>25
Peak year	2015	2045	2080	>2100
2100 Concentration (ppm)	420 ↓	540 →	670 ↑	940 ↑
Mean surface temperature				
2046–2065 (° C)	1.0 (0.4–1.6)	1.4 (0.9–2.0)	1.3 (0.8–1.8)	2.0 (1.4–2.6)
2081–2099 (° C)	1.0 (0.4–1.6)	1.8 (1.1–2.6)	2.2 (2.2–3.1)	3.7 (2.6–4.8)
>2300	1.1	2.4	3.7	7.6
Mean sea level				
2046–2065 (m)	0.24 (0.17–0.32)	0.26 (0.19–0.33)	0.25 (0.18–0.32)	0.30 (0.22–0.38)
2081–2099 (m)	0.40 (0.26–0.55)	0.47 (0.32–0.63)	0.48 (0.33–0.63)	0.63 (0.45–0.82)

Table 6.1: Predicted future climate change in different AR5 carbon scenarios. The temperatures are sea levels are given relative to the mean 1986–2005, so they are *atop* the changes that have already occurred. The ranges in parentheses bound the results of 90% of the models. From AR5 WG1 Tables SPM.3, TS.1

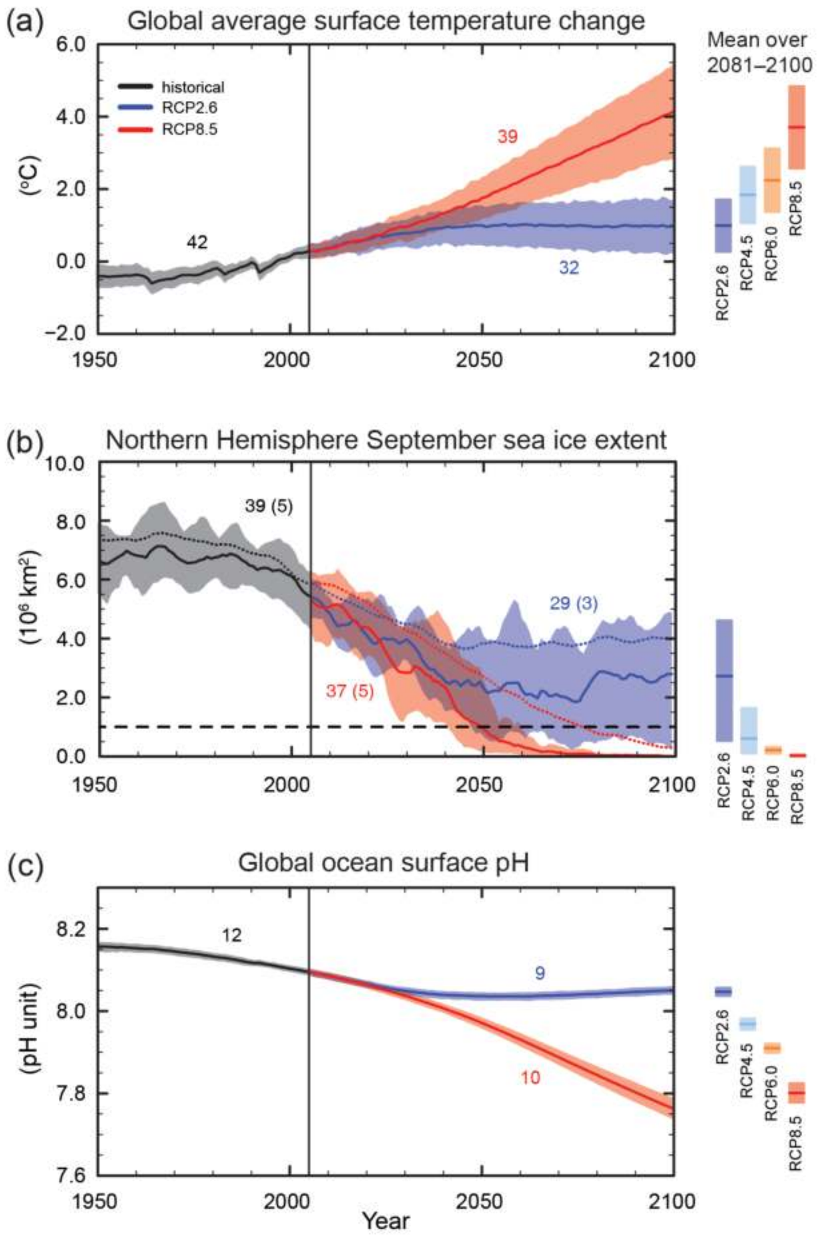


Figure 6.19: The mean and range of model predictions through the century under the 4 AR5 scenarios for future carbon emission (RCP's), giving (a) mean surface temperature; (b) Arctic summer sea ice; (c) ocean pH. The predictions vs time are shown for the most/least favorable scenarios. From AR5 WG1 Figure SPM.7.

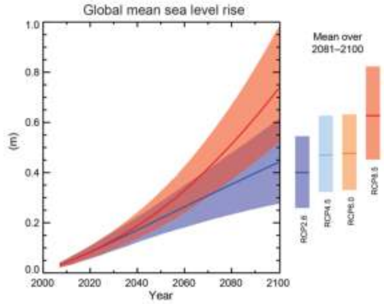


Figure 6.20: Average forecasted mean sea level rise through the century, plus range of model predictions, under the 4 AR5 scenarios for future carbon emission (RCP's). The predictions vs time are shown for the most/least favorable scenarios. From AR5 WG1 Figure SPM.9.

- Substantial warming is already "locked in" for centuries, even if we immediately ramp down CO_2 emission, since it takes centuries to remove CO_2. Figure 6.21 gives the longer-term forecasts for the 4 RCPs. In the "business as usual" scenario, less than half the total warming is accrued before 2100, and the ultimate temperature rise of 8 K (5–10 K range) is comparable to the difference between ice age and the current climate! Except that it goes the other direction (warmer) into temperature regimes that have not been seen in a million years!

- Sea levels are predicted to rise an additional 0.4 to 0.6 meters this century atop the 0.17 m rise in the past century. [Sea level continues to rise past 2100 unless CO_2 emissions go down soon.]

- In the business-as-usual scenarios, the Arctic icepack will be completely melted every summer after about 2050.

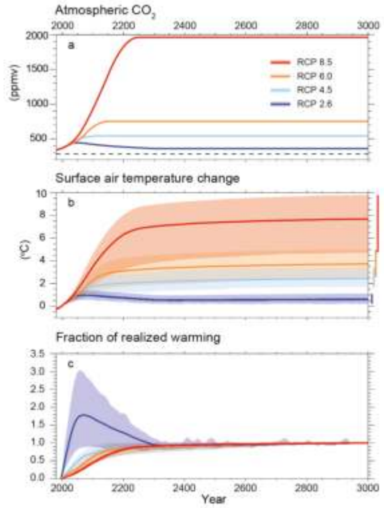


Figure 6.21: Climate model forecasts of the AR5 RCP's into the far future. From AR5 WG1 Figure 12.43.

- RCP4.5 is close to the commonly used ***doubling scenario*** in which atmospheric CO_2 climbs to twice the pre-industrial level of 270 ppm. To avoid passing this milestone, we need to limit future CO_2 emissions to 780 GTC atop the past emissions of about 375 GTC. If we stick to doubling atmospheric CO_2, global temperature rise this century will be be roughly another 1.8 K atop the 0.8 K already accrued 0.8 K that has already occurred, *i.e.* we are about 1/3 of the way through our temperature rise if we are 1/3 of the way through our carbon emissions.

- The "business as usual" RCP8.3 reaches the ***tripling scenario*** for CO_2 before the end of the century.

- Figure 6.22 shows that the expected temperature anomaly closely tracks the cumulative CO_2 emission in all of the scenarios; they just differ in whether and when we reach a certain level of emissions. **Each 1000 GTC of cumulative emission is estimated to cause a temperature rise of about 2 K this century.** The higher-emission scenarios lock in further temperature rise in future centuries.

The value judgment

The final question, the balance of the cost of reducing CO_2 emission vs the cost of avoidable climate change, is the most difficult to answer, and we will not take any close look at it in this course.

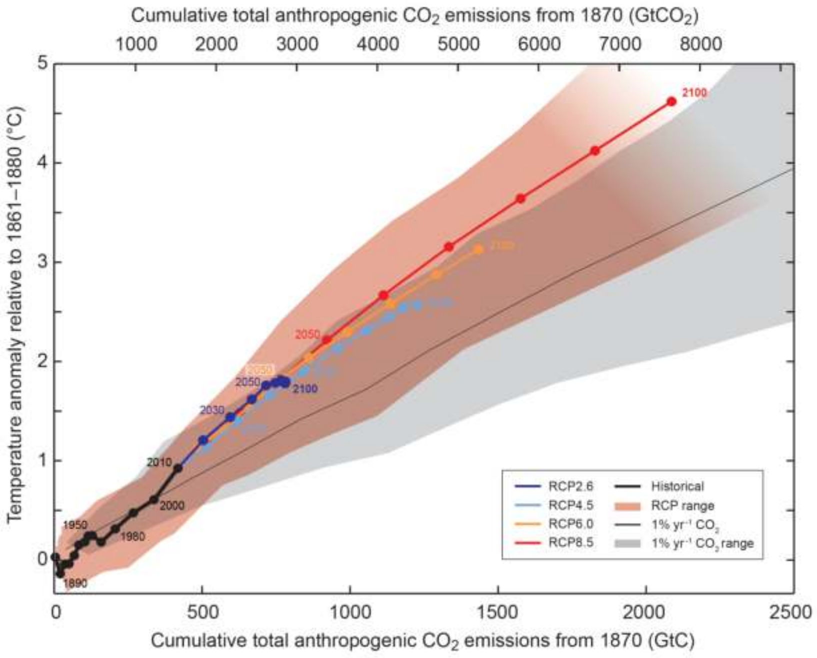


Figure 6.22: Average forecasted global mean temperature is plotted vs cumulative GHG emission for models of the 4 AR5 scenarios for future carbon emission (RCP's). While the different RCPs take different amounts of time to reach a given cumulative emission (as marked by the dots and year labels, the temperatures are similar functions of carbon emission. The black line shows the rise above pre-industrial levels to date. The brown region depicts the range of different model results for a given scenario—uncertainties are significant. From AR5 WG1 Figure SPM.10.

A small indication of the potential damage from climate change: in Florida alone, $145 billion worth of property lies within 1 meter of sea level, including 300,000 homes, and would become uninhabitable by 2100 in business-as-usual scenarios.[11]. The total US population within 1 meter of current high-tide levels is 3.7 million.[12] But the more serious damage would occur in low-lying countries such as Bangladesh that do not have the economic resources for prevention of or adaptation to higher sea levels.

Two or 3 degrees mean temperature change does not seem like much to worry about, given that we have 10 K temperature swings on many individual days. But a degree is a lot for climate. Figure 6.23 gives some intuitive feel (for Northeastern US residents) by showing that a high-emissions scenario would basically change Massachusetts summers into South Carolina summers. Also recall that an 8 K temperature *drop* was enough to *cover most of North America in glaciers.*

There have been several attempts to study the cost questions:

- The UK 2006 Stern Review estimated that climate change would reduce global GDP by 5% per year, whereas reducing GHG emissions would cost roughly 1% of global GDP.

- The AR4's median estimate of the cost of holding CO_2 emission to the doubling scenario was that it would reduce world GDP by

---

[11] See `http://sealevel.climatecentral.org`.

[12] Straus, Ziemlinski, Weiss, & Overpeck 2012, *Environmental Research Letters.*

about 0.6% by 2030.

- A May 2013 update to the snappily titled report *The Social Cost of Carbon for Regulatory Impact Analysis under Executive Order 12866* was prepared by several US executive-branch agencies. They estimate the cost of damage from each additional ton of emitted carbon due to changes in agricultural productivity, human health, flood risk, and "ecosystem services" such as water supplies. An example result is that each TC emitted in 2020 would cause $44–$240 of future damage as we change the discount rate from 5–2.5%, and that upper-end estimates are near $440/TC. Note that the dollar value depends heavily on discount rate since much damage is decades in the future. Also this price on carbon rises as time goes on.

- The 2014 *New Climate Economy Report* funded by a consortium of 7 national governments concludes that "investing in a low-carbon economy is a cost-effective form of insurance against climate risk." They estimate that US$6 trillion will be spent on energy infrastructure in each of the next 15 years, and that the additional capital costs for a low-carbon infrastructure would add US$270 billion to this, much or all of which might be recouped by lower operating costs.

Practice problems

- Using natural gas instead of coal to produce a needed number of Joules reduces CO_2 emissions and is hence beneficial to the climate. But methane that escapes unburned into the atmosphere is much more damaging than CO_2. What fraction of natural gas from a well would have to leak into the atmosphere to make the climate damage outstrip the gain from replacing coal? Assume a 20-year time period for the climate damage calculation.

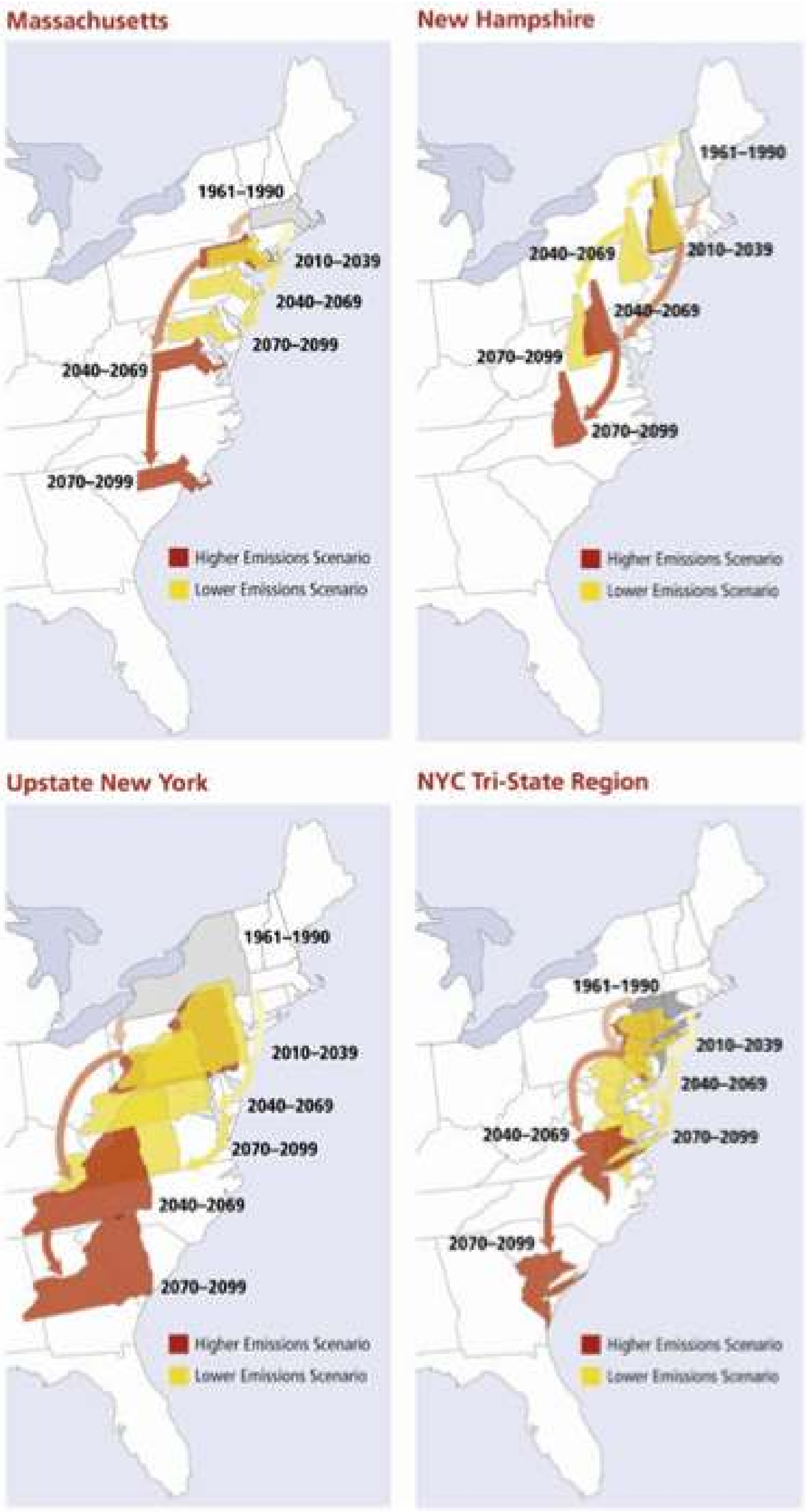


Figure 6.23: Illustration of the change in average summer "heat index" of the Northeast states under the predictions of some mid-2000's climate models. The lower-right panel illustrates, for example, that the higher-emissions scenario predicts that the mean summer heat index of Philadelphia and vicinity would be, in the last 30 years of the century, similar to the 1960's temperatures in Georgia. The "higher" and "lower" emissions scenarios to the highest and lowest of the AR5 RCPs. From the *Climate Change in the US Northeast,* 2006 report of the Union of Concerned Scientists.

7

Fossil fuels

FOSSIL FUELS ARE A WONDERFUL GIFT to us from nature. ***Photosynthetic*** bacteria and plants use the Sun's light energy to break the chemical bonds of CO_2 and produce *carbohydrates* and free oxygen, which contain more chemical potential energy. Photosynthesis in plants generally yields the transformation

$$6CO_2 + 6H_2O + \gamma \rightarrow C_6H_{12}O_6 + 6O_2. \tag{7.1}$$

Normally the chemical energy stored in the output molecule (glucose) is released by reversing this reaction in ***respiration***, with the energy used by the plant, or by some animal that eats the plant, or by bacteria that cause the plant matter to decay after its death. As long as this cycle continues, the net carbon impact on the atmosphere is zero, and the solar energy is basically converted into "body" heat of organisms at Earth's surface.

If, however, organic matter somehow escapes being decomposed before getting buried by sediment, then its chemical energy and carbon are sequestered underground. The heat and pressure underground transform the complex organic molecules in the simpler hydrocarbons (*e.g.* by removing much of the oxygen from the compounds and driving off the water content). After millions of years of sequestration and processing, we are left with the ***fossil fuels***.

Fossil fuels are incredibly cheap, practical, and abundant. We have built our society around them for good reasons. Here are just a few amazing fossil-fuel facts:

- You can buy a gallon of gasoline for much less than most Americans are willing to pay for bottled water.

- When you fill your tank at the gas station, you can put 2 GJ of energy in your tank in just a few minutes, meaning that the power flowing into your tank is $\approx 10\,\mathrm{MW}$, the same as an electric power plant that serves thousands of homes.

- World oil production in 2019 was about 95 *million* barrels per *day*, generating over *$5 billion per day* of wholesale revenue. Over *8.1 billion tonnes* of coal was mined in 2019.[1]

- The global investment in fossil-fuel infrastructure is estimated to be several trillion dollars.

Fuel	Energy Content		CO_2 Output	US Cost		
		(MJ kg^{-1})	(kg CO_2 GJ^{-1})			($/GJ)
Coal	25 GJ/tonne	25	88	wholesale:	$45/tonne	1.91
Crude oil	6.12 GJ/bbl	43	71	wholesale:	$53/bbl	8.66
Gasoline	0.13 GJ/gal	47	67	retail:	$2.33/gal	17.90
Natural gas	1.08 GJ/kcf	55	51	wholesale:	$2.68/kcf	2.54
				retail:	$11.07/kcf	10.25
Electricity				wholesale:	$0.022/kWH	6.05
				retail:	$0.133/kWh	36.94

Table 7.1: Summary of fossil-fuel energy, CO_2, and current costs. Note that coal energy content and cost per tonne vary with the grade of coal. Retail costs are Jan 2021 (gasoline) or Nov 2020 (electricity,NG) residential averages from EIA. Wholesale coal is avg. Nov 2020 cost to electric power plants, from EIA. Wholesale electricity is rough average peak price at PJM interchange for 2018. Wholesale oil, NG costs are late Jan 2021 West Texas Intermediate, Henry Hub, respectively (before Texas deep freeze). Costs can vary widely across regions and time, only a crude estimate is given here.

In this Chapter we'll look at how these fossil fuels were produced, the processes and side effects of extracting them from the ground, and estimates of the recoverable reserves available for future production in the US and the world. First we will look at coal, then petroleum and natural gas. Then we will look at some non-traditional forms of fossil fuel that have or might become economical and widely used as oil prices rise. Table 7.1 provides a summary of the energy, CO_2, and costs of the major fossil fuels in the US.

Coal

In bogs, decay of vegetation is inhibited by water that is too acidic and devoid of nutrients, so that dead vegetation accumulates at the bottom faster than it decays. This forms a layer of ***peat*** that is many meters deep in many locations, and can be used as a fuel.

During the *carboniferous era* roughly 300 million years ago, trees and ferns grew in great abundance in the swamps of this warm, humid climate. Plant growth was so rapid that dead plants were buried before they could decay, so enormous peat layers were formed, which were later buried under layers of sediment. The pressure and temperature compress approximately 20 feet of peat into 1 foot of high-carbon-content coal. Coal is found with varying carbon and heat content depending upon its geologic history, with lignite→bituminous→anthracite being the progression to higher-grade, more valuable resources. On average, the elemental content of coal is something like $CH_{0.8}(SNO)_x$. The $x < 0.1$ is undesirable sulfur

[1] *BP Statistical Review of World Energy 2020*

(causes acid rain when burned), nitrogen, or oxygen (do not yield energy in combustion), and many other trace elements. Remember that any dirt or minerals that were mixed in with the dead-plant layers as they settled will be fossilized into the coal.

Coal is found in ***seams,*** which are generally horizontal layers that extend over large geographic regions. A typical seam might be 3 m thick, a very rich one 30 m. Coal seams are almost all buried under an *overburden* of rock and soil. If the overburden is less than ≈ 60 m, then it is most economical to extract the coal by *surface mining* in which the overburden is scraped away to expose the seam. This is also known as *strip mining* or *mountaintop removal* depending on the geography. Deeper formations, down to 300 m or more, can be removed by *underground mining.*

Figure 7.1: A piece of coal, and its approximate wholesale value.

Figure 7.2: Left: an underground coal seam being mined (from here). Right: a surface (strip) mining operation. Note the enormous size of the equipment (from this site).

Extraction and use

Coal extraction has many undesirable side effects. Underground coal mining is one of the most hazardous occupations: the Chinese government reported 316 deaths in coal mining accidents in 2019, down significantly from $>$ 4000 per year in 2003. The recent US average is ≈ 12 per year. Coal miners also suffer chronic health problems from lung diseases at elevated rates. There are only about 50,000 people working in the coal production industry in the US, about the same as the number of students at Penn State University's main campus.

Extensive environmental damage is caused by surface mining. While strip-mined areas are required to be "reclaimed" in the US by replanting vegetation, this is definitely not the same as restoration to original state. Mountaintop removal fills in streams and valleys with mine waste. Mine tailings can leach heavy elements into the rainwater that filters through them, polluting the downstream water supply or underground acquifers. Coal also has sulfur in it. When rainwater runs through coal mines, pits, or tailings fields, the sulfur turns into sulfuric acid in the water, which is detrimental to nearby streams and groundwater.

Coal, being solid, cannot be refined to remove impurities before burning. Coal exhaust is therefore the leading single source of many pollutants. The sulfur in coal is burned to SO_2, which creates sulfuric acid when mixed with water. This is the origin of ***acid rain,*** which causes environmental damage downwind from coal-burning power plants and factories. Several government regulations have substantially lessened this problem in the past few decades. New coal power plants are required to install "scrubbers" which remove $> 90\%$ of the SO_2 from stack exhausts. A cap-and-trade program for sulfur emissions has operated successfully since the 1990. Power-plant sulfur emissions in 2019 are > 12 times lower than at the passage of the Clean Air Act amendments in 1990.

Many other elements/compounds in coal that are at least partially released by coal burning or mining, such as arsenic and uranium. While they are tiny fractions of the composition of coal, we are burning a billion tons per year in US alone, so it does not take a high concentration to cause a substantial release into atmosphere. Radiation released by coal burning far exceeds that released by nuclear power industry (save Chernobyl and Fukushima). Emissions of toxic trace elements are greatly reduced by capturing ash particles from the exhaust gases using *electrostatic precipitators* or filters—we want to see only the white water vapor coming from a power plant, not the dark smoke that signals the presence of particulates. Airborne emissions have been greatly reduced through regulatory requirements. Of course these elements are present in higher concentrations in coal ash that is collected, since all of the combustible/gaseous elements have been removed from the coal, so the coal ash itself must be disposed of responsibly. Note that the Trump administration in mid-2020 changed EPA rules on coal plant emissions of mercury and other pollutants so that they ignore the cost to public health of these emissions.

In summary, coal is the dirtiest fuel in several respects beyond its high CO_2 emissions per Joule: its extraction causes the most despoilation of land (except perhaps large oil spills), and its combustion releases numerous pollutants and leaves behind ash. When coal was the principal fuel of the industrial revolution, cities were covered in soot.

Reserves and consumption

Recall that the ***economically recoverable reserves*** of a natural resource are the quantity that is known with confidence to exist underground *and* be extractable profitably in anticipated market conditions. This is very distinct from estimates of the ***total resource,*** which are the total amount that might be present underground. This latter quantity can include more speculative sources and/or deposits that would be unprofitable to extract. Both confidence in the resource size and extraction cost can be judgment calls, and many organizations will have powerful economic incentives to skew reserve estimates up or down. For example, in the US, reserves can be taxable assets, giving some incentive for companies to reduce their estimates; however their stock prices and executive compensation might be higher if they over-estimate reserves.

Also keep in mind that a reserve estimate is *not* the final word on "when we'll run out" of something. Reserves can and must change over time, as

- the resource is mined and reserves decrease;
- new discoveries and exploration increase (or decrease) the amount of confidently identified underground resources;
- scarcity drives up prices of the resource, making it economical to exploit more difficult deposits and increasing reserves;
- extraction technologies (such as hydrofracturing) are developed that lower the cost of extraction;
- social or legal changes, or land-use changes, make it harder or easier to extract a resource.

With this huge pile of caveats in mind, let us take a look at estimates of reserves and the fossil-fuel supply lifetimes that they imply.

Coal reserves are easier to estimate than oil or gas because coal seams spread across large areas: if you find a 10-meter seam in one location, chances are good that you'll find a similar seam in the same geologic stratum several or even hundreds of km away. Also the extraction technologies and costs for coal have not changed rapidly in recent years. Table 7.2 summarizes estimates of the reserves and current production rates of coal and the other fossil fuels. These critical facts are apparent:

- **Coal is the most abundant fossil fuel and holds the majority of the energy in fossil-fuel reserves.**

- **The US holds 23% of the world's total reserves of 1.07 *trillion* tonnes of coal.** Russia holds 15%, and China's estimated reserves are 13%.

- **The world (US) coal supply will last 130 (390) years** *at current consumption rates.* Of course this supply will last much less long if energy consumption rates increase and/or we need to switch use from other scarcer fossil fuels to coal.

- **The 22,000 EJ of energy in coal reserves would fuel the world for 37 years** if we were completely reliant on coal at current usage rates, and less if energy use climbs.

- **If we burn all coal reserves we will release another 530 GTC** into the atmosphere, more than the $\approx$ 450 GTC emitted from all cumulative fossil-fuel use through 2019. This CO_2 emission would be about the amount thought sufficient to **push atmospheric CO_2 to double the pre-industrial level.** The only way out of this would be to either (a) leave the coal in the ground, or (b) capture the CO_2 from its combustion so that it does not reach the atmosphere.

Fuel	United States Reserves	 Production	 R/P Lifetime (years)	World Reserves	 Production	 R/P Lifetime (years)
Coal	250 GT 5586 EJ 134 GTC	0.64 GT 14 EJ 0.34 GTC	391	1070 GT 22,113 EJ 531 GTC	8.13 GT 168 EJ 4.03 GTC	132
Petroleum +NG Liquids	69 Gbbl 422 EJ 10 GTC	6.2 Gbbl 38 EJ 0.91 GTC	11.1	1734 Gbbl 10,612 EJ 255 GTC	35 Gbbl 213 EJ 5.10 GTC	49.9
Natural Gas	455 tcf 464 EJ 6 GTC	32 tcf 33 EJ 0.46 GTC	14.0	7019 tcf 7159 EJ 100 GTC	141 tcf 144 EJ 2.00 GTC	49.8

Table 7.2: Fossil fuel reserves and production rates for 2019, from *BP Statistical Review of World Energy 2020.* For each fuel, we give values in terms of quantities (tons, cubic feet, and barrels), energy content (EJ), and carbon content (GTC). For petroleum, values include tar sands.

Coal has not, in the past, been traded between continents in large quantities, since it is spread around the world and the most populous countries have indigenous supplies. But as of 2012, *China consumes more than half the world's coal,* and as of 2019 was importing $\approx$ 8% of their supply. Chinese coal usage has stopped growing rapidly and has stayed relatively steady since 2011, as has world consumption (*cf. BP 2020*).

Oil and natural gas

The simplest hydrocarbons are the *alkane* series of linear chains of carbon atoms decorated with hydrogen to form 4 bonds per carbon (Figure 7.3). These shortest alkanes are gases at room temperature and pressure; natural gas is 70–90% methane, the simplest hydrocarbon with the highest ratio $H/C = 4$. All can be liquefied at higher pressures for more efficient transport.

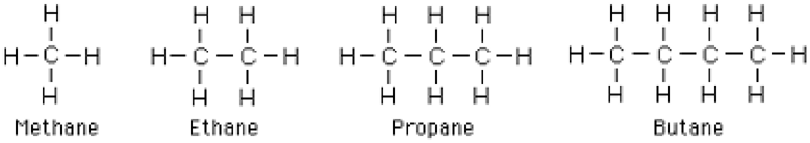


Figure 7.3: The first four alkanes. From http://hyperphysics.phy-astr.gsu.edu/hbase/organic/alkane.html.

The medium-length alkanes—pentane, hexane, heptane, octane, etc. with 5–17 carbons—are liquids at room temperature. The longer chains are solid or semi-solid, like pitch or asphalt. There are many forms of hydrocarbons more complex than the straight-chain alkanes. Petroleum (or crude oil) is a mix of various lengths of carbon chains, and each well has a different mix of molecules. The most valuable oil is called "light sweet" crude, for having short chains and a low sulfur content.

Oil and most natural gas are formed from creatures (mostly plankton) that sink to the bottom of the sea in conditions short of oxygen where decay is inhibited. If they get buried to depths of the "oil window" at 7500–15000 feet, the complex organic molecules are heated to "crack" into shorter chains that are liquid. With longer exposure in deeper, hotter geologic conditions, the average carbon chain length gets shorter and more gas is mixed in with the oil. Much natural gas is *associated gas* that is found in conjunction with oil.

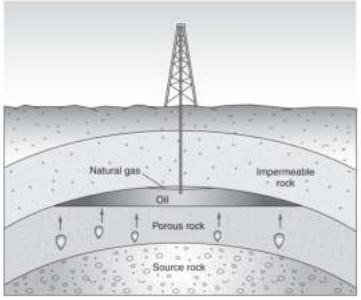


Figure 7.4: Oil and/or gas are formed in the pores of the *source rock*, migrates upwards to a porous *reservoir rock*, and is blocking from reaching the surface by some impermeable *cap rock*. Wolfson Figure 5.2.

The gas and oil are formed in a porous ***source rock.*** Once they are formed, they can flow, and since they are lighter than water they will begin to rise and seep out of the source rock. Oil and gas need to be *trapped* to prevent them from rising to the surface, where they can seep out and evaporate. This requires them to rise into a porous ***reservoir rock*** that is overlain by an impervious ***cap rock*** that has some kind of upside-down-bowl shape that prevents further migration upward.

Only perhaps 10% of oil that is produced at depth was trapped and still sits in a reservoir rock. The rest has seeped away. Perhaps 0.1% of the continents and continental shelves have oil fields under them.[2]. Unlike a coal seam, which will be contiguous over a large geological basin, oil deposits are rarer, scattered, and harder to find.

Extraction and refining

How do you find the oil and get it out? In some places it just seeps to the surface: tar pits and pitch were known to the ancients. In 1859, Edwin Drake's well in Titusville PA was the first to drill specifically for oil, which was found 69 feet down. The market for this "rock oil" was to replace whale oil for cooking, lighting, heating, and lubrication. By 1862, 3 million barrels of oil were produced in Pennsylvania. The introduction of automobiles boosted the oil market at about the time when electric lamps were making the kerosene-lamp market dwindle. By comparison there are now roughly 95 million barrels produced per *day*.

Deciding where to drill for oil and gas is the biggest employment sector for geologists: where might underground traps be holding reservoirs? One can examine surface features, and drill holes to see what the rock layers are like: are there cap rocks overlying porous rocks? In the modern era, one uses seismic techniques to map the underground strata, very similar in concept to ultrasonic medical imaging. But one still has to drill once a potential trap is found to see if there's oil there. Only a small fraction of exploratory ("wildcat") wells develops into a commercially production. Once an oil field has been discovered, further wells drilled in the surrounding area determine the total extent (volume) of the reservoir. This volume, multiplied by the porosity of rock and the fraction of oil vs water yields the estimated oil volume in the field.

Oil and gas are often found in the same reservoir. The natural gas used to be considered a nuisance because it could not be sold, and still cannot be marketed unless there is a pipeline connection to the well. It was usually just burned on site, called ***flaring.***[3] Natural gas that is found somewhere other than an oil-bearing formation is called ***unassociated gas.***

Yet more confusing, the output of a gas well contains contains some medium-length carbon chains that are liquid when separated from the the rest of the gas. These *natural gas liquids* are sometimes included in statistics on oil production, sometimes not. It can be difficult to find a consistent set of statistics.

Once the oil reservoir is discovered, we need to ***extract*** or ***recover*** the oil. Few wells "gush" and then only for a short time, pushed by pressure from gas stored with well. The oil needs to be pumped (pulled) out. The ability to do this depends on how porous the rock is and how viscous the oil is. In a typical oil well, this simple *primary extraction* typically leaves 70% or more of oil in ground! In subsequent *secondary* or *tertiary extraction*, or *enhanced oil recovery (EOR)*, underground interventions allow more of the oil to flow from the

[3] While this seems a huge waste, recall that it is much better to burn it than to let it escape into the atmosphere as methane.

reservoir rock into the well. First, water can be pumped into the well (or an adjacent bore) to push oil through the pores (also recall that oil and gas float on water).[4] Detergents can help oil droplets escape. Gases such as CO_2 can also be pumped into the well to drive out oil. Heat can be applied, *e.g.* by injecting steam, which makes the hydrocarbons less viscous (runnier). Even after these interventions, it is unusual to obtain a *recovery fraction* of more than 50% of the oil in the reservoir. Most of the *reserve growth* in the past 40 years has occurred not by discoveries of new oil fields, but by increases in the recovery fraction thought feasible.

US annual production of oil peaked at 21.5 EJ (about 3 billion barrels) in 1970 and continued to decline fitfully to about 12 EJ/year in 2000-2010, as discoveries of large new fields essentially stopped, and existing fields passed their most productive years, and the oil became more difficult to extract from existing fields. From 1970–2010 the US was increasingly dependent on foreign imports for its petroleum supply. It is hard to overstate how important this was to US politics and policy. The desire of the US populace and politicians to sustain a flow of cheap oil into the country from the Middle East and elsewhere shaped the course of foreign policy, domestic energy policy, and war-making during this time.

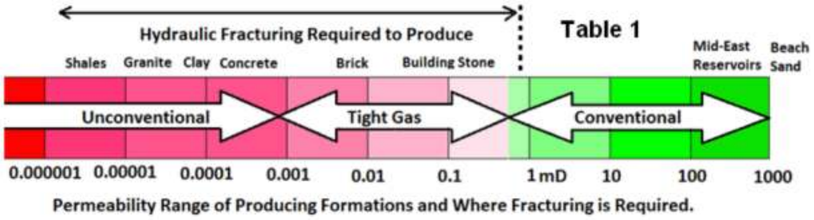


Figure 7.6: Scale of permeability, the porosity of a material, or the amount of fluid that will flow through it under a given pressure. Mideast oilfields are very "loose." Fracking makes it possible to extract oil from less permeable reservoirs. From *Hydraulic Fracturing 101*, G. King, Society of Petroleum Engineers #152596 (2012).

Since c. 2011, US crude oil production has risen rapidly to 27 EJ in 2019. This is primarily attributable to widespread adoption of *hydrofracturing* (a.k.a. *fracking*) to extract oil from "tight oil" formations such as the Bakken (in North Dakota & Wyoming) and the Eagle Ford in Texas, which produced 63% of all US crude oil in 2019. These are shale reservoirs which are not porous enough to yield significant recovery fractions from traditional techniques. In hydrofracturing, water is injected into the well under high pressure to create a network of fractures in the surrounding rock, through which the oil (or gas) can flow to the wellbore. Sand or ceramic materials in the water serve as *proppants* to hold the fractures open after the pressure is removed, and other chemicals in the fracking fluid serve to reduce viscosity, inhibit corrosion, or otherwise enhance the yields.

Hydrofracturing has been tried decades ago, but with little eco-

[4] An increasing ratio of water to oil coming out of a well is a sign that the oil is getting more difficult to extract.

nomic viability, because the fractures reach only small distance away from the wellbore, so a vertical well can only extract oil or gas from a small region as it crosses the reservoir stratum. But development of efficient ***directional*** or ***horizontal drilling*** meant that a wellbore could traverse a long stretch of an oil- or gas-bearing stratum, and multiple fracks along the way could access a much larger quantity of oil or gas. This technology combination, along with a sustained high price for oil, led to large-scale exploitation of shale oil formations that were previously much too "tight" to be economical.

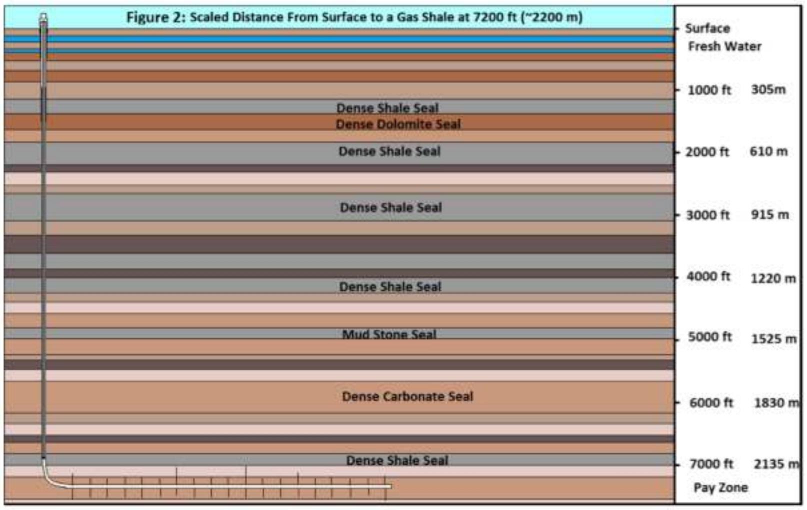


Figure 7.7: Diagram of a typical fracking well, drawn to scale. Note that the gas-bearing layer is far below, and well sealed from, the gas- or oil-bearing stratum. But leakage along the well bore is possible. From *Hydraulic Fracturing 101*, G. King, Society of Petroleum Engineers #152596 (2012).

Since the 1970's, oil exploration has moved to more difficult sites—farther offshore and deeper, smaller reservoirs, less porous "tighter" formations that do not release their oil easily—as the rising price of oil and the exhaustion of "easy" oil has led to willingness to exploit more expensive resources. The average US oil well produces 10–15 barrels per day, in comparison to an average Saudi well's production of over 5000 bbl/day.

Figure 7.11 shows the history of world oil prices, illustrating the spectacular volatility since the 1970's that followed a century of very steady oil prices. The first "oil shocks" in the US occurred in the late 1970's, after US production peaked, and when the OPEC cartel initiated boycotts and pushed up prices. There followed a 20-year period of sustained lower prices until a prolonged rise to record prices through the 2000-2011 period. Unlike the 1970's shocks, this run-up had no particular geopolitical events at its root—just demand growing faster than supply (aside from a drop at the start of the financial crisis). These high crude prices incentivized extraction of more difficult resources, such as deep-water fields and tar sands (described below). This was also the start of the fracking boom in oil in the US. A dramatic price collapse, from ≈ \$100/bbl in June 2014 down to ≈ \$40/bbl in March 2015, occurred as US production from fracking increased while conventional oil producers continued apace, and demand growth in China and elsewhere slowed. This in turn led to a retrenchment in oil exploration, away from the more expensive locations such as the Alaskan north shore and deepwater sites, as well as the less productive fracking sites. Oil prices (and drilling investment) recovered to \$50–60/bbl in late 2019.

But 2020 and 2021 have been wild rides for oil (as for everything else!), as seen in Figure **??**. In April 2020, prices plummeted as demand dropped at the start of the pandemic, while producers kept up their production rate. Major oil markets actually had *negative* prices, meaning you could get paid to take oil off of producers' hands (!), because all storage areas were full. The oil prices have rebounded, and spiked in early February 2021 as a deep freeze and widespread blackouts hit Texas, standing at ≈ \$65/bbl in mid-March 2021. If I could accurately predict what will happen next, I would be a very rich man!

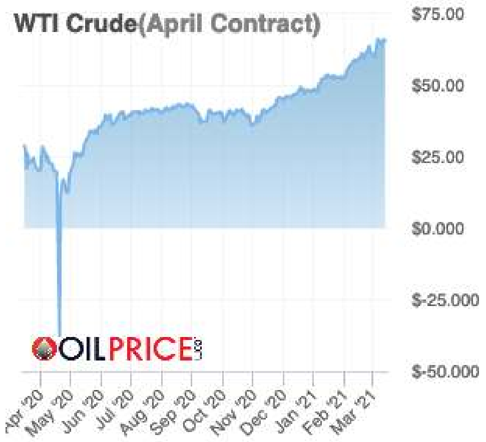


Figure 7.8: West Texas Intermediate crude prices for March 2020–March 2021. These prices are insane!

After crude oil is pumped out of the well, it needs to be ***refined*** before use. This process removes impurities and produces products with consistent compositions for specific uses, such as gasoline, diesel, jet fuel, tar, propane, etc. Chemical engineers can not only separate the different hydrocarbons, but can also break long chains into shorter ones (*cracking*) or build longer chains from short ones (*polymerization*) to produce a mix of petroleum products to meet demand. Just south of UPenn on the way to the airport you will pass the largest oil refinery on the Eastern seaboard. A refinery has been operating at this site since 1866, just after Drake's discovery in Titusville. But after a catastrophic explosion, and multiple bankruptcies, in 2019 the facility closed, and the difficult job of figuring out what to do with that (very dirty) site is a big issue in Philadelphia politics.

As you can tell, refineries are ugly and smelly.[5] It can be difficult to obtain permits and invest the capital to build a new one. The spike in gasoline prices after Hurricane Katrina in 2005 was *not* due to any shortfall in oil production in the US; it was a shortfall in *refining capacity* that drove prices up. Likewise a spike in California gasoline prices in 2012 was due to shutdowns of refineries.

Crude oil and refined products can be transported long distances at low cost using pipelines and tankers, and these liquids are easy to store. There is a truly global market for oil—the product flows to the highest bidder. So it doesn't make much sense to say "we don't buy any of our oil from country X," our energy security really depends on how much we draw from the global supply. One exception to the global marketing of oil was that export of crude oil from the US was banned by law in 1973 in response to the OPEC embargo. This made little difference, however, since US consumption was much higher than its production, so US producers had no incentive to export. Near the end of the Obama administration in 2016, this ban was repealed, since the US fracking oil boom led to crude oil being economical to export in some cases. It is odd that we now export crude oil while still importing substantially more—this occurs because fracked oil is more "heavy" and "sour," and US refiners desire as input a lighter/sweeter mix than the US produces. The US currently currently stores roughly 660 million barrels of crude oil in the Strategic Petroleum Reserve in Lousianna to buffer prices and guard against interruptions in supply.

Reserves, production, consumption

The specialized conditions needed to produce and trap oil mean that it is much harder to estimate reserves, and furthermore the reserve estimates depend on changing forecasts of recovery fractions. To make matters worse, companies and countries intentionally obscure their true reserve estimates for a variety of reasons. Consider for example the official estimates published by Saudi Arabia and Venezuela, the countries currently claiming the largest proven reserves (Figure 7.9).

Apparently during 1980–1987, the Saudis discovered enough new oil to exactly balance the 17 billion barrels they produced in this time, so their reserves stayed exactly flat, then suddenly rose by 100 billion barrels in 1987, and again discoveries have exactly balanced

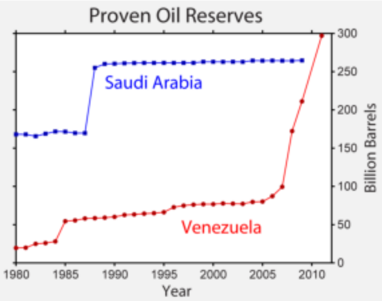


Figure 7.9: Graph from Wikipedia October 2014; data available in *BP Statistical Review of Energy*.

---

[5] New Jersey's reputation as a bad-smelling industrial state is attributable to a refinery that was located at exit 13 of the NJ Turnpike in Newark. It is much cleaner now than it was when I was young.

production since. This is clearly ridiculous. The 1987 jump in claimed reserves was motivated by the OPEC policy that cartel members could sell oil in proportion to their proven reserves. In this period of falling oil prices, several OPEC members suddenly "discovered" large increases in their reserves. We can be similarly skeptical of Venezuela's government oil company claims that oil reserves quadrupled in the space of 4 years to become the world's largest. This isn't totally crazy: remember that reserves must be economically recoverable, and the big rise in oil prices over these years made much of Venezuela's "heavy" oil economial. But meanwhile, the reserves estimate did not come back down when oil prices collapsed. BP's world oil reserve estimates jumped by 130 billion barrels in 1999, and by another 220 billion barrels c. 2010, when Canandian estimated resources of tar sands resources were reclassified as proven oil reserves (see below). Given that much of the military and political history of the past 80 years has been driven by oil supplies, we can assume that many interests besides accuracy drive national and corporate estimates of reserves. All of this suggests we should take the oil reserves estimates with a trillion-barrel grain of salt.

The natural gas story in the US has been hugely influenced by hydrofracturing, in fact the gas fracking boom preceded the oil boom by a few years. Figure 7.10 shows the fracked gas production history in various US geological formations, illustrating the tremendous increase in just 10 years using hydrofracturing+horizontal drilling techniques. Hydrofractured gas is now about half of all US production, and production from just the Marcellus shale underlying Pennsylania, West Virginia, and New York has gone from practically zero in 2010 to $>$ 5 trillion cubic feet (tcf) in 2017, about 20% of all US production. US natural gas reserves estimates have doubled in the past 15 years; it is likely that other countries will see large NG reserves increases as exploitable shale gas formations are developed in the future.

Natural gas is cheaply and efficiently transporting by pipeline where such pipelines have been built, such as most of the US and Canada. Efficient intercontinental transport of natural gas requires cooling it to 112 K to produce ***liquefied natural gas (LNG).*** The infrastructure for global trade in LNG is expanding quickly, but the capacity is too low to have equalized NG prices between continents—US prices are well below those in the rest of the world, as shown in Figure 7.11. While US power plants are moving quickly away from coal to natural gas, coal remains by far the dominant fuel in Chinese power plants. Natural gas has played a central role in geopolitics recently just as oil has for decades: Russia is the 2nd-leading gas producer (in 2016), accounting for 16% of the world's supply, and has

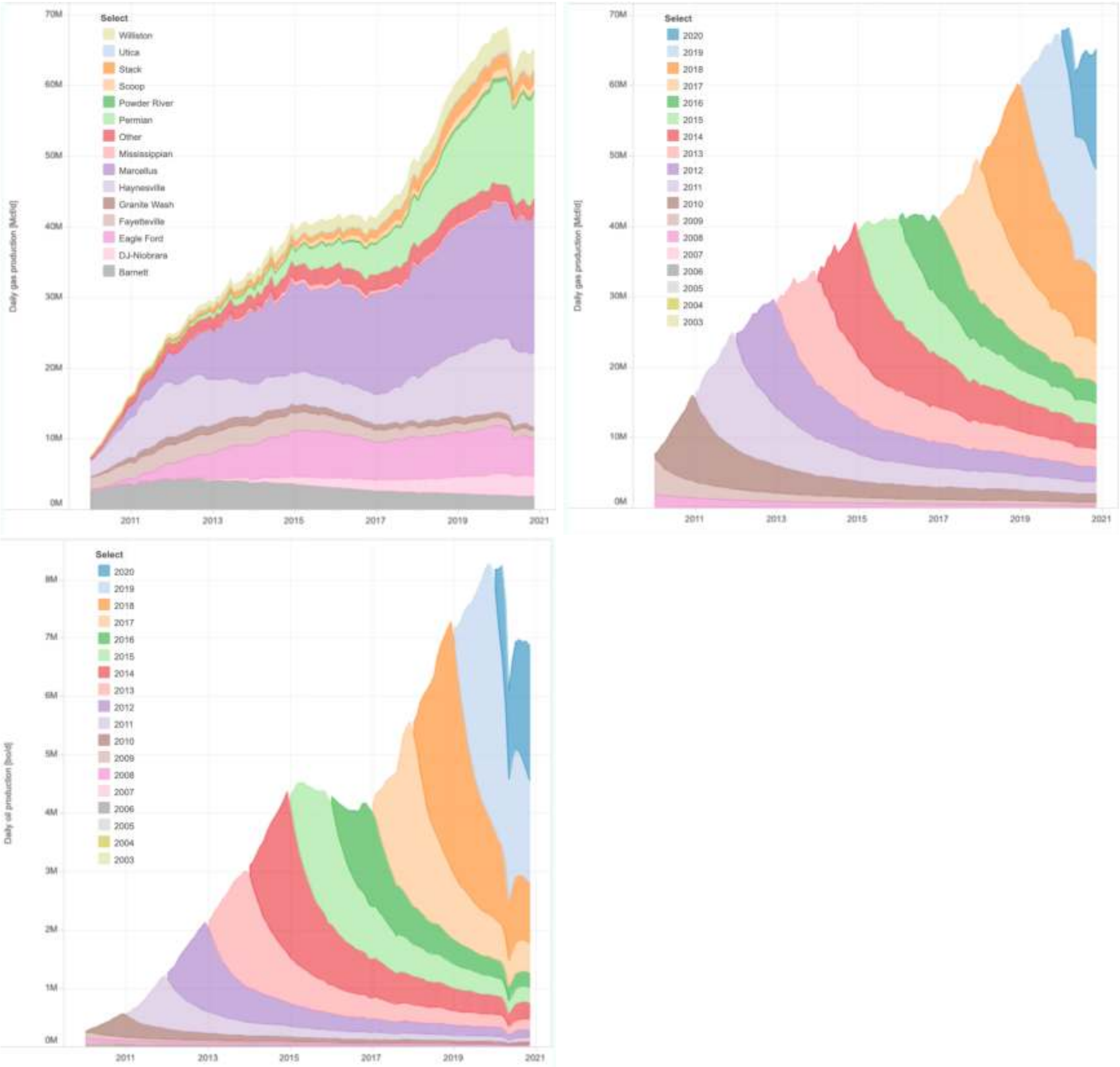


Figure 7.10: Top left: The amount of natural gas produced from shale formations in the US is plotted vs time. Different colors denote distinct geological formations; the large purple wedge is the Marcellus shale that crosses western Pennsylvania, West Virginia, and New York. Top right: the shale gas production shown again, with the colors divided by the year in which the well was drilled. We can see that at any time, most of the production is from wells less than three years old, and well productivity drops by $> 50\%$ after one year. Lower left: same thing for fracked oil. From shaleprofile.com.

used their control over Europe's NG supply to destabilize disfavored governments.

Forging ahead, we consider the BP estimates of proven oil and NG reserves reported in Table 7.2. The key observations about oil and gas are:

- **Globally and in the US, oil and NG reserves hold similar energy, and combined they are less than coal energy reserves.**
- **The great majority of carbon in global reserves is in the coal.**
- **Global oil and NG reserves represent ≈50-year supplies at current consumption rates, compared to >100 yr for coal.**
- **In the US, oil and NG reserves hold only <10–15 yr of current production.**
- US coal reserves are very large by comparison to oil and NG, and by far dominate the carbon content of US fossil fuel reserves.
- The US in 2015 surged past Saudi Arabia to regain its 1970's crown as the largest oil-producing nation, and **as of 2019 was producing (and using!) 18% of the world's oil, with Russia and the Saudis each about 12%.** But it is also true that the great Middle Eastern oil fields still have much higher reserves and are much more porous and productive than the typical US well. **We are burning through our known reserves at a much higher rate than the Middle Eastern oil giants,** suggesting that our current period of lower oil imports will not be long-lived.
- **Total carbon content of fossil-fuel reserves are ≈900 GTC, which in the mean IPCC estimate is sufficient to more than double the pre-industrial atmospheric CO_2 level and produce an additional 2 K temperature rise by the end of the century.**
- **Additional reserves growth will be needed to maintain the current growth rate in energy demand beyond c. 2050 if it is to be continue to be met primarily by fossil fuels.** If additional fossil fuels are discovered and burned, the cumulative global warming will climb >3 K above the pre-industrial mean.
- If reserves are over-estimated, supply of fossil fuels will fail to meet rising demand within about 20 years.

As noted in Chapter 2, oil and gas are much less evenly spread around the world than coal. Roughly 3/4 of conventional oil reserves are in the Middle East. Even though Russia and the US are producing as much oil as Saudi Arabia in 2019, they have much smaller

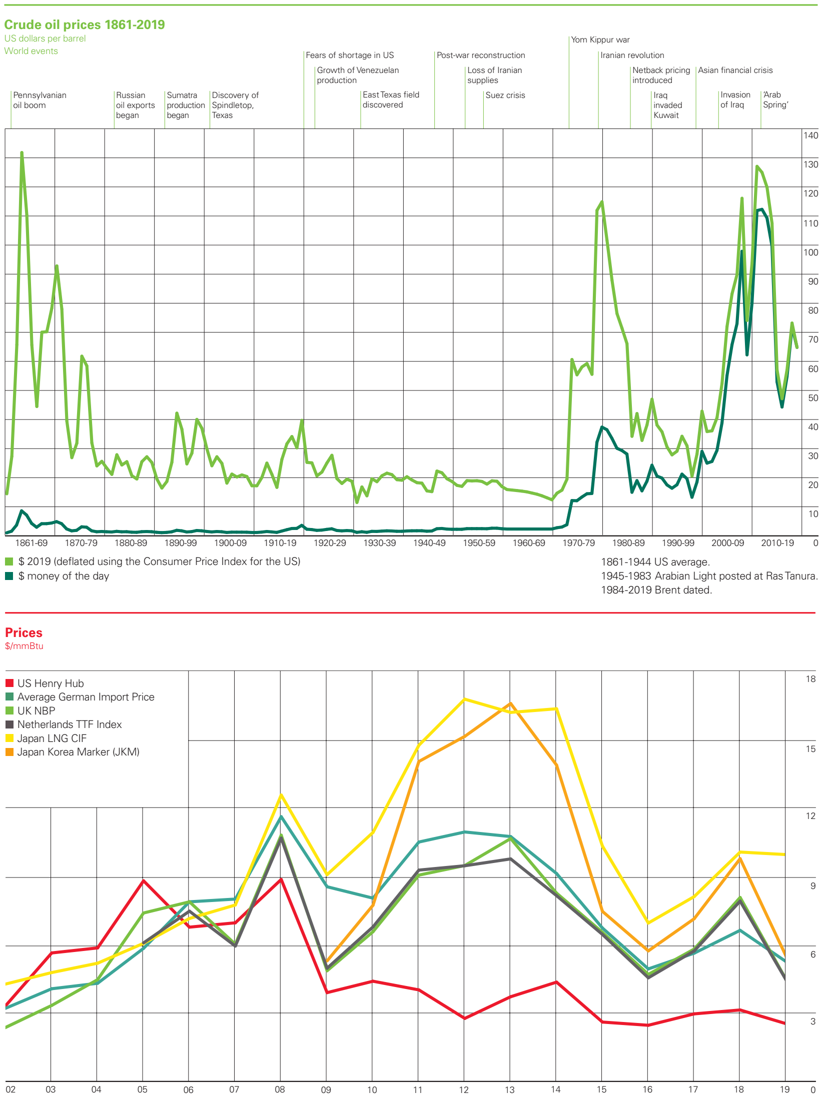


Figure 7.11: The history of crude oil prices (top) and the more recent history of natural gas prices (bottom). Note the dramatic swings in oil prices, and the divergence in US gas prices from the rest of the world after the start of the fracking boom in the US. Both from BP *Statistical Review of World Energy 2020.*

reserves and will deplete their resources sooner at current production. Oil in particular has played a critical role in global military and political affairs for a century. Figure 7.11 shows the price history of oil for the past 150 years. Inflation-adjusted crude oil prices recently spent ≈ 3 years at the highest levels since Drake drilled his first well, then dropped almost three-fold. Previous spikes in oil prices—in the 1970's—were triggered by OPEC boycotts and war-related shutdowns of major producers. But the more recent high prices were not attributable to such supply interruptions; they reflected the growth of demand outpacing supply.

A natural-gas future?

The fracking boom has pushed natural gas prices down substantially in the US and led to an increasing fraction of US energy use—particularly in the electricity sector—to switch to NG from coal. This is beneficial in a CO_2 sense, but detrimental if low NG prices are suppressing investment in zero-carbon nuclear and renewable sources. We should ask whether increased NG dependence is a viable long-term path toward substantially lower CO_2 emission, at least in the US which is currently the best-case scenario for NG. Similarly we might ask whether the fracking-based boom in tight-oil production in the US will continue to supplant imported crude in the US.

The center and right panels of Figure 7.10 illustrate an important point about shale gas wells (and fracked oil as well): their rate of output drops very quickly after drilling. In these graphs we see that more than half of the fracked gas or oil production is coming from wells drilled within the past three years. After one year of operation, the wells are producing less than half as quickly, and a typical well extracts half of all the oil or gas that it will ever get within two years of starting. This means that we have to maintain a high rate of drilling in order to sustain even constant production—much moreso than with conventional wells. The technology is improving, in that drillers are extracting more product per well; but we should expect that that as the best sites are depleted, it becomes harder to maintain high production. Like the Red Queen, we will need to drill more and more just to stay in place at the same production rate.

Can fracking continue to supply larger portions of our transportation (oil) and heating/electricity (gas) energy needs? We see in Figure 7.10 that growth slowed or halted during 2015 and 2016, after the oil price collapse (and sustained low gas prices). Then more rapid growth occurred from 2017 until the pandemic hit. Since fracked wells fade so quickly, the production rate will respond rapidly to the willingness of investors to drill new wells. The peak production rate of shale gas in 2019 was sufficient to supply about 25% of total US primary energy use in . Moving a majority of our energy usage to natural gas would require a massive expansion of the current shale gas production rate, which does not seem feasible, let alone sustainable, to me, as it would require discovery of quantities several times larger than current NG reserves to sustain even half of the US energy economy for just a decade, and a huge increase in the drilling rate.

Unconventional fuels

Reserves of conventional oil appear insufficient to sustain a continually growing market for liquid hydrocarbons through the century. Already there is a substantial premium being paid for energy in the form of oil instead of coal or gas, driven by the need for liquid fuels for transportation. We can ask what other sources of thousands of EJ of liquid hydrocarbons might become economical when conventional oil production inevitably declines and it becomes scarce and expensive. Even longer-term: are there fossil energy sources beyond the coal-oil-gas troika we currently depend on?

Coal-to-liquids

Chemical engineers know how to make hydrocarbon liquids from natural gas (polymerization) on a large scale. Of more interest given the large reserves of coal is *coal-to-liquids (CTL) or liquifaction.*[6] CTL is not a new idea: when Germany was cut off from oil supplies in World War II, they produced over 100,000 barrels per day of liquid fuels from coal. South Africa produced much of its oil supply from coal when they were subject to petroleum boycotts in protest of apartheid. CTL is still in use in South Africa, as the country has abundant coal but little domestic oil production.

Conservation of energy of course requires that the Joule content of the liquid fuel output can be no more than that of the coal input. In fact it will be substantially less, since energy is required to effect the transformation. The Fisher-Tropsch process used for CTL has a theoretical maximum efficiency of 60%.[7] Hence each Joule of the synthetic oil for sale requires >1.67 J of coal to be consumed, and extraction of this energy from the coal will necessarily end up producing CO_2. So the synthetic liquid fuels will have a net CO_2-emission-to-energy ratio that *exceeds* that of coal.[8] In a CO_2 sense, therefore, CTL would be a dirtier fuel than coal. One advantage over coal is that impurities in the coal are removed during the CTL process and so the output liquid would not generate as much sulfur and trace-element pollution as direct coal-burning does.

[6] We will also discuss coal *gasification* in the context of carbon-capture electric power plants later in the course.

[7] A.P. Steynberg, H.G. Nel *Fuel* **83** (2004) 765–770

[8] It is possible, however, that the CTL plants could capture some of the CO_2 that is generated during the transformation process (though clearly not the CO_2 generated with the liquid fuel is burned in transportation engines).

Tar sands

Figure 7.12: A handful of Canadian tar sand. Image from Suncor Energy Inc via `http://ostseis.anl.gov/guide/tarsands/`

.

Unconventional oil is a liquid or semi-liquid hydrocarbon deposit not amenable to extraction by drilling. *Heavy oil* is too long-chained and viscous to pump. One kind of heavy oil resource is ***tar sands:*** sand or sandstone impregnated with up to 18% ***bitumen,*** a mix of long-chained hydrocarbons.

Tar sands are surface or near-surface formations that are probably *unreservoired oil:* the remnants of giant oil formations that had no cap rock formations, and migrated to the surface, where the smaller, lighter carbon chains evaporated for were digested by bacteria, leaving behind the heavier tar/pitch/bitumen.

There are two huge resources of tar sands: the Athabasca (Canada) and Orinoco (Venezuela), claimed to contain 170 and 300 billion bbl of economically recoverable oil, respectively. This is 1/2 as large as the BP estimate of global total conventional-oil reserves. Production in Alberta reached 2.9 million barrels per day in 2019. and Venezuelan heavy-oil production is very low after the country's economic collapse. So tar sands are about 3% of world oil production of 95 million bbl/day.

To get usable oil products, the tar sands are first extracted using open-pit mining. The bitumen then must be separated from the sand, clay, etc. The tar sand is mixed with hot water and agitated (stirred). The bitumen floats to the top of the mixture and is skimmed off. The bitumen is too thick to be transported by pipeline (unless mixed with lighter hydrocarbons) or to be commercially useful, and is very high in sulfur ($> 4\%$), so it must be ***upgraded*** by chemical processing into lighter carbon chains. This ***syncrude*** can then be refined into commercial products. Extraction of tar sands is as environmentally devastating as coal strip mining, but in some sense worse because 2 tons (2000 kg) of tar sands are extracted to yield 1 barrel (140 kg)—so for every kg of fuel, 13 more kg of sand is dug up and discarded.

All of this processing makes oil from tar sands costly. It is costly in a monetary sense –$60–$75/bbl by this Canadian estimate. Clearly there are economically viable deposits at $100/bbl oil prices since production was expanding rapidly at high oil prices. With current oil prices, however, it is less obvious that tar sands development is profitable. We can note, for example, that the economic impetus to build the Keystone XL pipeline (to transport Canadian tar sand oil to the US) has died down quite a bit. Its usefulness as way for politicians to show their (lack of) support for environmental issues remains: Obama delayed it, Trump approved it, and Biden immediately revoked the permits. But its economic viability is no longer very good and it's unlikely to get built (even if approved) while hydrofracturing

Figure 7.13: A bird's-eye view of a tar sands extraction operation in Alberta, Canada. Yuck. Source unknown.

is cheaper than tar sands extraction. Bitumen processing is costly in water and energy as well as in dollars: several barrels of water are needed for each barrel of syncrude. Each J of energy in the final commercial product requires more than 1 J of energy input. In Alberta a good fraction of this energy input actually comes from natural gas (to heat the water for separation, for example). In principle the process can be energy-self-sufficient by using the tar sands as the raw energy input for all the processing, but then of course the final yield is only a fraction of the energy mined. The net carbon impact of tar sands oil is hence *higher than the carbon impact of conventional oil.*

Even ignoring the potentially higher CO_2-per-Joule rate of tar sands oil, tar sands exploitation is a significant danger for climate simply because of the large volume of fossil carbon that it holds, which will significantly worsen climate change if it ends up in the atmosphere. And of course there is tremendous local environmental damage and deforestation at the site of the deposits, and related problems with water usage. Official debate over approval of the Keystone XL pipeline proposed to carry syncrude to the US centers on environmental damage along the route of the pipeline. But this is, environmentally, a minor issue: the real issue is whether we have any power to prevent this additional massive reservoir of carbon from being burned.

Oil shale

Oil shale is a fine-grained sedimentary rock infused with ***kerogen,*** a carbon/hydrogen-bearing, waxy material. Kerogen is "baby oil"—

organic material that has not yet spent time at the depth where pressure and heat would convert it to oil. It is nonetheless a reservoir of large quantities of chemical potential energy, and with sufficient processing it can be liberated from its rock and cracked into liquid fuels.

Figure 7.14: A piece of oil shale. To confuse matters, *oil shale* neither holds oil nor is shale. And it is different from "shale oil," which is the tight oil coming from *e.g.* the Bakken formation.

Most of the world's known oil shale lies in the Green River formation in Utah, Colorado, and Wyoming, the remnants of two large tropical lakes of 50 million years ago that collected 1000 meters of organic-rich sediment. There are estimates that the Green River formation contains 2 *trillion bbl* of oil equivalent in shale layers more than 3 m thick. The is nearly double the proven reserves of conventional oil! But unlike coal-to-liquids and tar sands, there has never been commercial-scale extraction of oil shale, and I am not aware of any concerted present effort to develop this resource.

Getting useful oil out of oil shale is even more difficult and messy than tar sands. The oil shale needs to be heated to 500° C to release the hydrocarbons, which then have to be processed into usable liquid fuels. Producing 1 million bbl/day of fuel from oil shale would leave behind 500 million tons of rock and require 3 million bbl of water every day! To make matters worse, the oil shale *expands* when heated, so it won't fit into the hole you dug it out of! And this is a semiarid region where water is already scarce.

The oil shale is clearly not in the picture for energy supplies in the next decade or so. But if we become desperate enough for liquid fuels in the more distant future, we could see this resource being exploited, probably at substantial environmental and financial cost.

Methane clathrates

One other unconventional fossil fuel is worth mention: ***methane hydrate*** is molecule is formed when methane exists in water under high pressure. The water molecules form a cage around the methane, making this a ***clathrate*** in chemistry lingo. The methane clathrates form solid nodules that are spread around the sea floor and frozen into permafrost around the world. When brought up to sea level, the clathrates disintegrate and methane is freed.

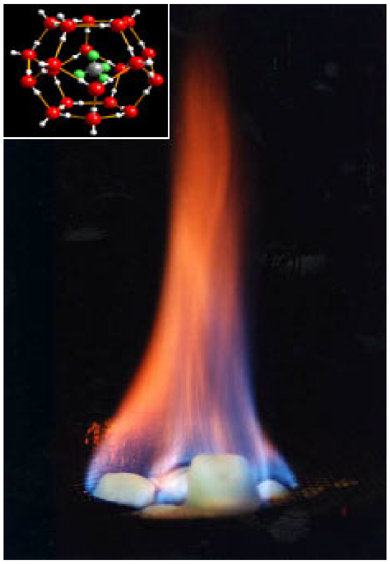


Figure 7.15: A molecular diagram of a methane hydrate clathrate in which water cages a methane molecule; and a picture of burning chunks of "ice," which are really clathrate nodules. From the Naval Research Laboratory.

The quantities of carbon in ocean-bottom clathrates are estimated to equal or surpass the proven reserves of conventional natural gas. But they have never been exploited commercially—there are some serious difficulties, one being that the resource is large but spread over a vast area (the ocean floors) so it would be very expensive to collect them. You would need roaming collection vessels, which means you could not easily ship the gas back by pipeline. We should probably hope that methane hydrates are never exploited, not only because they would boost atmospheric CO_2 by yet another large increment, but because the seafloor ecosystems would probably be devestated if we got in the business of digging up large swaths of ocean bottom to pick up the clathrate nodules.

Figure 7.16: A methane clathrate nodule pulled up from the sea floor (at left). The clathrates break down a surface pressure (at right). Source: ???

Total resource estimate

Estimating *resources* is even speculative than estimating reserves. Table 7.3 gives the estimates used by the IPCC for the AR5 report. We see total resources estimated at perhaps 5–7× in energy, and 10× as much carbon, as we have spoken of being in the reserves. The great majority of the additional energy and carbon are in the form of coal, presumably harder to extract than current coal reserves. If we burn and emit CO_2 from any significant fraction of these additional resources, we are destined for the worst of the IPCC climate scenarios.

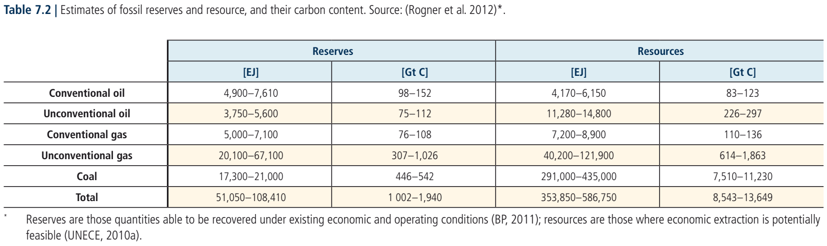


Table 7.2 | Estimates of fossil reserves and resource, and their carbon content. Source: (Rogner et al. 2012)*.

	Reserves		Resources	
	[EJ]	[Gt C]	[EJ]	[Gt C]
Conventional oil	4,900–7,610	98–152	4,170–6,150	83–123
Unconventional oil	3,750–5,600	75–112	11,280–14,800	226–297
Conventional gas	5,000–7,100	76–108	7,200–8,900	110–136
Unconventional gas	20,100–67,100	307–1,026	40,200–121,900	614–1,863
Coal	17,300–21,000	446–542	291,000–435,000	7,510–11,230
Total	51,050–108,410	1 002–1,940	353,850–586,750	8,543–13,649

* Reserves are those quantities able to be recovered under existing economic and operating conditions (BP, 2011); resources are those where economic extraction is potentially feasible (UNECE, 2010a).

Table 7.3: Estimates of energy and carbon stored in world energy *reserves*, from Rogner et al. (2012) as presented in Volume III of the AR5.

Practice problems

- How long will China's proven reserves of coal be able to meet their consumption if they do not import coal? Consider both the current consumption rate, or continued growth at 8% per year.

- When BP publishes their estimate of world reserves of natural gas in 2030, do you expect it to be higher or lower than this year's estimate?

8

Hydropower and electricity

Hydropower harnesses the gravitational potential energy of a flow of falling water. Electricity is very similar, conceptually: we capture the electric potential energy of a stream of negatively-charged electrons "falling" toward positive charges. In this chapter we will gain an understanding of the physical concepts of these two forms of energy, and look at the global usage and potential of hydropower. In the next chapter we will look closer at the economics and practicalities of the fossil-fuel-based electricity industry.

Hydropower

Principles

Hydropower extracts the gravitational potential energy of water descending from the mountains. Usually it goes into kinetic energy of water, then friction against the banks. But in an ***impoundment dam*** system, as illustrated in Figure 8.1, we create a lake behind a dam, then we can open or shut a hole (*penstock*) in the dam to have the water fall where and when we like. When we let water flow through the dam, what is the *power* in the stream of water? Recall that power is the rate of energy use or release, and the energy here is the gravitational PE of the water, so

$$P = \frac{E}{t} = \frac{mgh}{t} = \left(\frac{m}{t}\right) \times (gh) = (\text{current}) \times (\text{head})\,. \tag{8.1}$$

The ***current*** is the rate at which water flows through the dam. The ***head*** describes the energy released per kg of water. Often the head is given as h instead of gh. This is the vertical distance from the place where water disappears (the top of the reservoir) to the place where it exits the dam. The exiting water will have this gravitational potential energy converted to kinetic energy of moving water. A ***turbine***

How did the water get this potential energy?

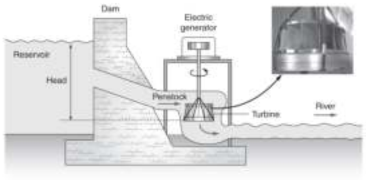


Figure 8.1: Schematic of an impoundment dam. Wolfson Figure 10.6

placed at the exit can transfer the water's energy to a rotating shaft—it's the opposite of a ship's propellor, which transfers kinetic energy in a rotating shaft into moving water. In a ***hydroelectric*** facility, the turbine shaft is connected to a ***generator*** which converts the mechanical energy into electric energy.

Later in the chapter we will investigate how the generator works.

Figure 8.2: A photo (left) and engineering drawing (right) of the Grand Coulee Dam. Source: ???

Consider the largest hydroelectric facility in the US: the Grand Coulee Dam in Washington (Figure 8.2). The engineer's drawing shows that the surface of a full reservoir is about h =100 m above the turbine (elevation 1290 ft vs 938 ft, 352-foot head). The average flow of the Columbia river is 110,000 cubic feet of water per second. Since 1 cf of water has mass 28.3 kg, the Grand Coulee has an average power available of

$$\begin{aligned} P &= (\text{current}) \times (\text{head}) && (8.2)\\ &= 1.1 \times 10^5\,\text{ft}^3\,\text{s}^{-1}\,\frac{28.3\,\text{kg}}{\text{ft}^3} \times 9.8\,\text{m s}^{-2} \times 100\,\text{m} && (8.3)\\ &= 3.0 \times 10^9\,\text{W} = \underline{\underline{3.0\,\text{GW}}}. && (8.4) \end{aligned}$$

Over the course of one year (8640 hours), the Grand Coulee liberates a total of 3.0 GW × 8640 h = 26.3 TWh of gravitational potential energy. They report generating an average of 21 TWh of electricity per year. Therefore the *efficiency* of conversion of the gravitational energy into electricity is $21/26.3 = 80\%$.[1] This is very high compared to the 35% efficiencies that we've been quoting for coal power plants. Why is hydropower so much more efficient? Because there is no conversion from heat into non-heat energy, so we never confront the Carnot limit. The turbine and the generator are each typically >90% efficient.

The Grand Coulee dam has a generating ***capacity*** of 6.8 GW. This means that there are enough turbines and generators installed to produce this much electric power when the penstock gates are opened all the way. But notice that the *average* power produced is 21 TWh/8640 h = 2.4 GW, which is well below the maximum pos-

[1] In fact our estimate is a little bit too low, the efficiency is a bit higher than this.

sible. There is not enough water going down the river for the dam to produce its capacity of 6.8 GW all the time. For any generating facility we will want to know the ***capacity factor***[2] which we define as

$$\text{capacity factor} = f = \frac{\text{average power output}}{\text{maximum power output}} = \frac{2.4\,\text{GW}}{6.8\,\text{GW}} = 35\%. \tag{8.5}$$

Why build a hydroelectric plant that can produce three times more power than is available from the river on average? It would have been much cheaper to install 1/3 as many turbines. But **the electric power grid cannot store a significant amount of energy. Electricity needs to be supplied to the grid to match whatever is being consumed at any moment.** This fact is critical to understanding electricity generation. Hydroelectric power is highly ***dispatchable,*** meaning that we can turn the power on and off quickly to meet changing demand—we just need to open and close valves on the penstock. This makes hydro very valuable to the utility companies.

Grand Coulee therefore runs at full power only when electricity demand is high—*e.g.* on hot summer afternoons when many air conditioners are running—and at these times more water is leaving the reservoir through the dam than is flowing into it from the upstream river. At low-demand times, the penstocks can be closed, and the water (and its energy) just collects in the reservoir.

Hydropower has many characteristics we seek in a sustainable energy source:

- No CO_2 or other pollutants are emitted by its operation. There is CO_2 emitted in producing the vast quantities of concrete in the dam, but when this is amortized over the lifetime energy production of the dam, the CO_2 emission per GJ is very low.
- The resource (rain) is renewable.
- Electricity, the highest-quality form of energy, is produced directly, at high efficiency.
- The electricity can be created when needed.

Sounds great! So why not generate all of our electricity with hydro? First, the number of appropriate sites are limited: you need a river with substantial flow; you need the river to flow through a canyon that will contain the reservoir that forms behind the dam. The river's course must be steep enough that we can build a tall dam to generate the most power from the current. Geography allows only so many sites.

Second, dams are not environmentally benign. Upstream of the dam, a large area gets flooded. The Three Gorges Dam in China

---

[2] Capacity factor is not the same thing as efficiency, do not confuse them!

flooded a 375-mile-long area the size of Singapore and displaced millions of residents. The riparian habitat is replaced by a lake, changing the oxygenation of the water and the types of species that can thrive. The dam itself blocks the upstream migration of fish, and the trip through the turbines may not be so good for fish headed downstream. Turbines can be designed to reduce fish kill, and fish ladders can be built to allow some upstream migration, but the dam will still make migration harder. Downstream from the dam, the natural flow of water is very different from water that is released on a schedule driven by electricity demand. For instance shutting off the river at times of low demand is clearly not so good for downstream species. The Grand Canyon gets daily "scouring" by heavy water releases from the upstream Glen Canyon dam at daily electricity demand peaks. This is very different from the seasonal variations that used to exist, changing the nature of the canyon ecosystem.

Another issues with dams is that they block the downstream flow of silt. This depletes downstream riverbeds of silt. It also accumulates in the reservoir behind the dam, reduces its capacity and eventually rendering dams unusable, so on the time scale of centuries, hydropower is not sustainable.

Hydropower history and potential

The energy of falling water has been harnessed since at least 85 BC. It was used to run grain mills or lumber mills for centuries, using water wheels. Before steam engines were practical in 1800's, early industrialization occurred from mills and machinery plants run from shafts turned by river water. The invention of practical steam engines moved industry from water power to coal power.

The first *hydroelectric* facility in US was built in 1882, and hydro facilities were one of the primary sources of electricity. The largest US facilities were built during the Depression (*e.g.* Grand Coulee, Hoover). The completion of the Grand Coulee Dam in 1942 was a critical advantage in World War II, as its prodigious output went toward electricity-intensive military industry: smelting the aluminum that went into many Boeing aircraft, and producing plutonium at the Hanford Reservation.

Hydroelectric facilities, unlike other electric power stations, are typically publicly owned in the US, which makes sense considering that (a) they are very capital-intensive projects, requiring financing that private companies would have trouble generating; (b) private ownership of rivers is rare; (c) the dams have many consequences and purposes beyond electricity generation—flood control, irrigation, navigation, and environmental concerns. The largest electricity generation stations in the world are hydroelectric: the largest of all is China's Three Gorges Dam, with a total capacity of 22.5 GW (a typical nuclear power plant will generate 1 GW, with coal and gas facilities less than this).

Hydroelectric production in the US was 287 TWh (1.03 EJ) in 2019, essentially unchanged since 1970. The best potential sites in the US are already in use for hydropower, and environmental concerns have become more prominent in the cost-benefit analyses for newer dams, so we should not expect hydropower output to increase significantly in the US in the future. As US electric production has grown faster than hydropower, the fraction of US electricity generated by hydro has dropped from 30% in 1950 to 7.0% in 2019.[3] Nonetheless hydro was for many years the largest renewable energy source in the US (and world) by a large margin. In the US, 2019 was the first year that more electricit was generated with wind power than hydro. It is also perhaps the least expensive source of electricity in the US, since the capital costs of dam-building were paid off many years ago, and the operating costs of dams are low since of course there is no fuel to pay for.

Figure 8.3: One of the 32 turbines installed in the Three Gorges Dam. Each generates 700 MW of electric power. Source: ???

The 1.03 EJ of energy produced by hydro are valuable zero-entropy electricity. In accountings such as the Livermore energy flow chart in Figure 2.11, hydro is credited at 2.62 EJ of "primary" energy, since it displaces this much fossil-fuel energy that would be needed to produce the equivalent electricity output in an average power plant of 39% efficiency. Be careful of this when looking at statistics. Usually the statistics of electricity production are given in kWh or TWh, and do *not* have this fictitious factor of 1/0.39 added into their energy.

Table 8.1: Hydroelectricity production and capacity for world and top producing countries, from the IEA *Key World Energy Statistics 2020*.

Country	Production (TWh 2018)	Capacity (GW 2018)	Percent of Total 2018 Electricity
China	1232	352	17.2
Canada	386	81	59.0
Brazil	389	105	64.7
USA	317	103	7.1
Russia	193	51	17.3
Norway	140	33	95.0
India	151	49	9.6
Japan	88	50	8.4
World	4325	1293	16.2

Note that the 2019 yearly output of 287 TWh is equivalent to an *average* hydropower output of 33 GW. The *capacity* of US hydropower is 103 GW,[4] so the typical *capacity factor* of US hydro plants is 32%. Hydro gives ≈10% of our capacity (and we import hydroelectricity

[3] EIA *Monthly Energy Review* Feb 2021, Table 7.2a

[4] The IEA gives a capacity of 102 GW, but the DOE says 80 GW — it's not clear to me which is more accurate.

from Canada as well), more than its fraction of *production*.

Figure 8.4 shows that the saturation of hydroelectric power over the past 40 years in the US is characteristic of the entire developed world. New hydropower facilities are, however, being built in the developing world, particularly in Asia and South America—although hydro production is not growing as quickly as total electricity growth, so hydro produces a shrinking *fraction* of world electricity.

Hydro resources and usage vary widely among nations according to their geography. China produced 28% of all the world's hydropower in 2016. Another 25% came from Brazil, Canada, and the US. Some countries with abundant rainfall and mountains obtain more than half their electricity from hydro: Brazil, Norway (95%!!), and Canada. The global average of 16% is more than double the US fraction.

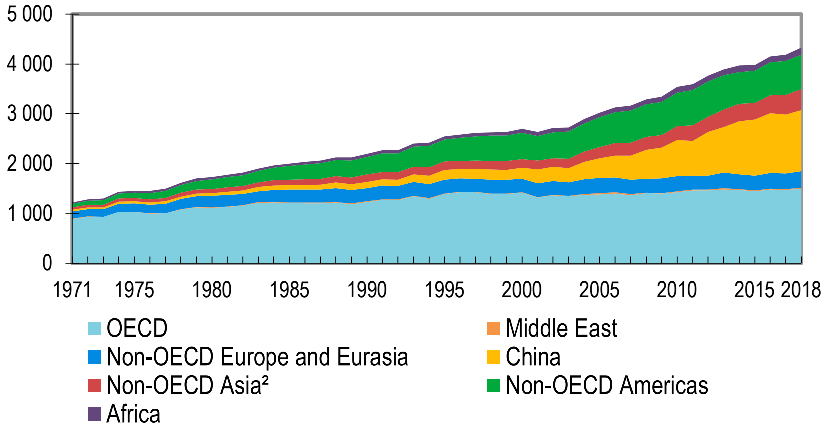


Figure 8.4: History of hydroelectricity production for selected regions and countries. From the IEA *Key World Energy Statistics 2020*.

Is hydroelectricity capable of or likely to provide a much larger fraction of global energy needs than it does now? We note first that 2018 hydro production of 15.6 EJ represents only 6.8% of global energy use even if we credit this energy with a factor of 1/0.39. There are reasons to doubt that this fraction will grow or even be maintained unless total world energy use stops growing. There are significant untapped hydropower resources in underdeveloped countries (*e.g.* an estimated 40 GW on the Congo River). But it seems common for energy use grow faster than hydro production once a country reaches a modest level of development. Furthermore there is greater recognition now of the environmental and societal damage caused by large dams, and the track record for societal benefit of large hydro projects in third-world countries is not good. There may be more emphasis henceforth on renewable wind and solar electricity that can be deployed quickly and locally as opposed to large-scale hydroelectric projects.

Electricity

The charger for my laptop has this cryptic writing on it: "Input: 100–240 VAC 1.5 A 50–60 Hz, Output: 20 VDC, 4.25 A max." Before we examining the national- and global-scale energy and environmental issues associated with electricity production, we will want to understand the basic lingo and physics of electricity.

Volts, amps, ohms, and watts

Electricity is the flow of electric charges—usually electrons, which have negative charge. It is very much like the flow of water, except the motion is driven by electric forces on charges instead of the pull of gravity on masses. Figure 8.5 shows the basic idea:

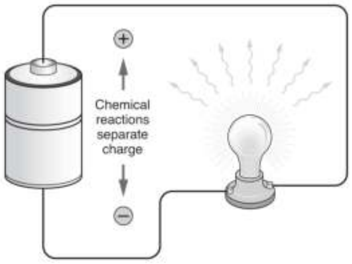


Figure 8.5: A simple electrical circuit. Wolfson Figure 3.3.

- Some device (battery or generator) separates the positive and negative charges in some atoms or molecules. Note that it takes input energy of some kind to perform this separation since we are raising the electric potential energy.

- A metal wire creates a path for electrons to flow from the negative end of the source to the positive end. Any electron that makes this journey will have lower electric potential energy at the end of the trip.

- The light bulb (or some other appliance) is placed along the wire to convert this electric potential energy into some other desired form—just as the turbine in a hydroelectric dam taps the energy of the flowing water. Electricity is the *flow of electrons* which we use to **transport energy** from the source (battery) to the user (light bulb).

- The total power that can be extracted from the electron flow is

$$\begin{aligned} P = \frac{\text{energy}}{\text{time}} &= \frac{\text{enegy}}{\text{electron}} \times \frac{\text{electrons}}{\text{time}} && (8.6)\\ &= \text{Potential} \times \text{Current} && (8.7)\\ &= V \times I. && (8.8) \end{aligned}$$

- The ***current*** is rate of electrons making the trip down the wire. It is measured in ***amperes (A)***, or usually just "amps."[5]

- The ***electrostatic potential*** or just "potential" is the energy each electron gives up if it makes the trip from the negative to the positive end of the battery. It is analagous to the head of a hydropower

[5] In case you are curious: 1 ampere is 6.25×10^8 electrons per second. Also, a technicality is that we think of the current flowing from + to −, opposite to the electron flow, because the electrons have negative charge.

facility—it is helpful to think of the electrons as flowing "downhill" along the wire. We measure potential in ***volts (V)*** and consequently the potential is often referred to as the ***voltage.*** The amp and volt units are chosen such that a 1 A current traversing a 1 V potential will release 1 Watt of power:

$$1\,\mathrm{V} \times 1\,\mathrm{A} = 1\,\mathrm{W}. \tag{8.9}$$

If we were to cut the wire partway, the electrons on their journey would pile up at the $-$ side of the break. These electrons would then repel later electrons. So the current in the wire *stops quickly unless there is a complete path from the $-$ to $+$ end of the source.* This path is called a ***circuit***. Of course a light switch works by opening and closing a break in the electrical circuit, to control whether current can go to the lamp.

Generally a battery will produce a certain voltage V, and the amount of current that flows will depend on the ***resistance*** of the connecting circuit. The resistance describes how easy it is for electrons to get through the wires/appliances that make up the circuit. The lower the resistance (or the higher the ***conductance***) of the circuit, the more current will flow under to "push" of the voltage. Mathematically we have

$$I = \frac{V}{R} \tag{8.10}$$

$$\Rightarrow P = IV = I^2R. \tag{8.11}$$

Going back to our water-flow analogy: a fat pipe has *low* resistance to water flow—you get water faster when you suck on a fat straw than a skinny one. You have three ways to affect the resistance of a wire:

1. Make it *fatter* to produce *lower* resistance. A lower-resistance wire will be more expensive since it will take more metal to make a fatter wire.

2. Make it *shorter* to produce *lower* resistance. A longer path to travel has higher R, *i.e.* it takes more voltage to push the same electric current a further distance.

3. Make the material out of a *more conductive material* to produce *lower* resistance. Metals are good conductors of electricity, with silver and copper being best. Glass, rubber, most plastics, and air are poor electrical conductors (good ***insulators***) and are used where you want to block electricity flow. Very pure water is a good insulator, but water with dissolved ions in it (like the water in our body) is a much better conductor.

Note that the criteria for electrical conductance/resistance are very similar to those we encountered for thermal conductance in Equation (3.7).

Resistance is measured in ***ohms* (Ω).** We have

$$1\,\text{V} = 1\,\text{A} \times 1\,\Omega. \tag{8.12}$$

In the US, the two slots of household electric outlets have an electric potential between them of 120 V. If a light bulb is going to use 100 W of power, then the current through the bulb, and the resistance of the bulb, must be

$$I = \frac{P}{V} = \frac{100\,\text{W}}{120\,\text{V}} = 0.833\,\text{A} \tag{8.13}$$

$$R = \frac{V}{I} = \frac{120\,\text{V}}{0.833\,\text{A}} = 144\,\Omega. \tag{8.14}$$

Notice that if the resistance of the circuit is *small*, such as if we put a short piece of metal, like a paper clip, between the two prongs of the wall socket, then the current becomes *large* and the power released becomes *large*. This power will be turn into heat in the wire and it will spark, melt, and/or set a fire—bad things. This is called a ***short circuit***. Building codes demand that your household wiring be protected against this sudden energy release by having a ***circuit breaker*** connected to every wall socket: this is a device that will shut off the voltage (open the circuit) if the current flow exceeds a specified maximum.[6] A typical US household circuit might have a 20 A circuit breaker. This means the maximum electrical power that can be supplied will be

$$P = VI = 120\,\text{V} \times 20\,\text{A} = 2400\,\text{W}. \tag{8.15}$$

This is why most hair dryers, microwaves, etc., use something like 1 kW. Appliances that need more power—clothes driers, air conditioners—must be put on a circuit with a higher current limit.

Going back to my laptop charger: it says it can take input from a circuit with anywhere between 100–240 V. European household wiring provides 220 V at the socket. Almost every place around the world has standard outlet voltages in the range my charger will accept. Since it will take in less than 1.5 A, it draws less than $120\,\text{V} \times 1.5\,\text{A} = 180\,\text{W}$ from a US outlet.

The charger feeds electricity into the laptop. Its output voltage of 20 V is different from the input, and the maximum output current is 4.25 A. so it can deliver up to 85 W of power to the laptop. A device like this that changes electricity from one voltage to another is called a ***transformer.*** Notice that the output *power* of the transformer must be equal to (or less than) the input power, even though the *current or voltage* of the output can be higher. That's because the transformer cannot create energy out of nothing!

Now consider the more complicated ***series*** circuit in Figure 8.6, consisting of three circuits end-to-end. In a series circuit, *the current I*

---

[6] A *fuse* is an older version of the same principle. It's a little wire that will just melt (safely) if too much current goes through it.

is the same in each segment but *the total voltage is the sum of the segments' voltages*, $V = V_1 + V_2 + V_3 \ldots$. Again, think of water flowing downhill through a series of pipes: the total head (vertical drop) is the sum of the drop on each segment.

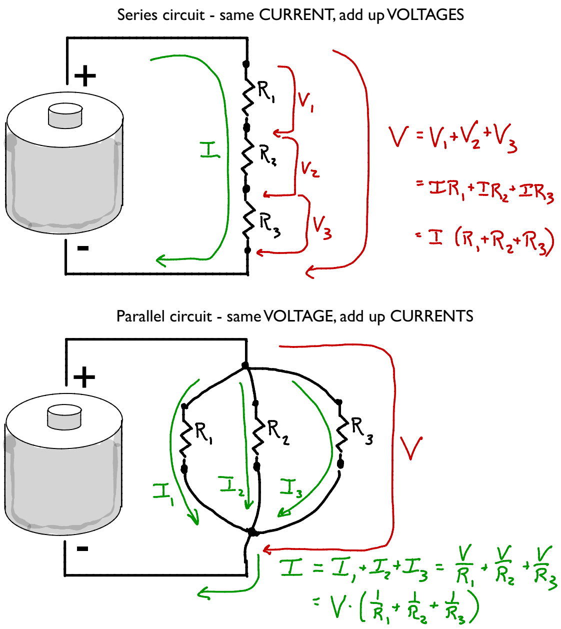


Figure 8.6: Top: illustration of a *series* circuit, with three electrical resistance elements connected end-to-end. They share the same current, but the resistance and voltages of the three sum up to give the total. Bottom: the three resistors are now in *parallel*, so they each carry their own current, and see the same voltage, so the total current is now the sum of the three.

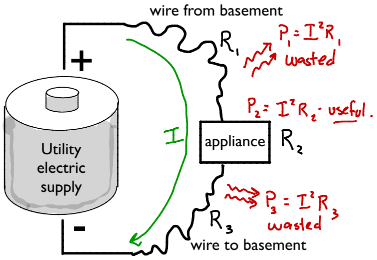


Figure 8.7: Energy is wasted as heat because of resistance in the wiring between the electricity source and its destination appliance, which form a serial circuit.

A common case is illustrated in Figure 8.7: segment 1 is the wire from your house's electric input (the circuit breaker box) to your appliance; segment 2 is your appliance; and segment 3 is the return wire. Our appliance needs some power P, and if the appliance were the *only* thing in the circuit, we'd have $I = P/V$. But there is also resistance in the wires, so there is power generated in the wiring:

$$P_{\text{waste}} = I^2(R_1 + R_3) \approx \frac{P^2}{V^2}(R_1 + R_3) \tag{8.16}$$

$$\Rightarrow \frac{P_{\text{waste}}}{P} = \frac{P}{V^2}(R_1 + R_3) \tag{8.17}$$

This power is (a) wasted, since it turns into heat, and (b) is a potential fire hazard if too much heat is generated in the walls. We want to make it smaller! We can't change the power P or the length of the wires, because presumably our appliance needs to be in a certain place and needs a certain power to do its job. One way is to make the R of the wires smaller, meaning fatter wires, which becomes expensive.[7] The alternative is to *transmit the electricity at higher voltage.* It takes less copper to wire a European house than an American one because of the 220 V standard!

Notice that this gets more and more urgent an issue when we are talking about transporting 100 MW or more from a power plant

[7] You could also switch from copper to silver wire, but that's pricey too.

to cities hundreds of km away, not just transmitting a kW across my basement. The transmission lines on the electric grid operate at 345,000 V or 765,000 V! Distribution would be crippled by transmission losses without this. Even with the current system, about 10% of electric power is lost in transmission.

Figure 8.8: A high-voltage transmission line. Source: ???

Electrical hazards

Why not run our household wiring at 345,000 V? At such high voltages, it's very difficult to prevent electric current from flowing in undesired paths across materials we normally consider insulators. Even air, one of the best insulators, is insufficient at these voltages, as electrons are tempted by the huge potential to leap from one wire to the other. Along the way they knock other electrons off their atoms, creating an avalanche ***breakdown*** of current. Lightning is a naturally occurring form of breakdown. Two wires with such a large potential between them must be kept far apart to avoid breakdown. Clearly this isn't feasible in household wiring.

The second danger from electricity is that you might become part of the circuit by accidentally touching two conductors at different potential. If we run current through your body, then the resistance of your body will generate heat. 100 W of heat might start to hurt you, which for 120 V wall socket would be about 1 amp. But much smaller currents can cause severe injury by disturbing nerve impulses, which are themselves just small electric currents. In particular if current passes through your heart, cardiac rhythms can be interrupted. A 0.1 A current through heart muscle can cause fatal fibrillation. The amount of current, and the part of the body it passes through, determine the amount of damage. 0.005 A or less is generally harmless. And of course the resistance of your body, in combination with the voltage that you touch, determines the current that will pass. Most electronics run on just a few volts, and touching <40 V with dry skin will generally be safe. Wet or broken skin is more conductive, and voltages as low as 20 V can cause injury if immersed. On the other hand, you can experience very high voltages (such as a spark) without injury if the current is tiny. Household line voltage of 120 V or more can easily kill you if you become part of the circuit. Do not experiment with this! Roughly 40 people die per year in the US die from electrocutions from consumer products.

We see that transmission lines require thousands of volts, but we have to lower this to hundreds of volts using a transformer before the electricity enters a home. And this must be lowered again to a few volts for safety in portable devices.

Motors, generators, and transformers

How do we convert electric energy into kinetic energy? This is the job of an electric ***motor.*** Most motors exploit the fact that *an electric current creates a magnetic field* which exerts a push or pull on another current or a permanent magnet. Electric motors do not involve conversion of heat into work, and therefore are not subject to 2nd-Law limitations and can have efficiencies of 95% or higher. Electric motors are great compared to heat engines like internal combustion: there is no fuel or oxygen needed, no exhaust, and very little waste heat generated. Of course someone needs to supply the electric energy.

Another puzzle: how is the electric potential difference created in the first place? A ***battery*** segregates electrons from protons (ions, really) using chemical energy, which of course is just microscopic electric energy, so we're simply rearranging charges. A ***fuel cell*** is just like a battery except it is not meant to be recharged: one can continuously replace the chemicals to generate electricity. Fuel cells are not yet a cost-competitive means of converting chemical energy to on-grid electric energy, but there are companies selling fuel cells as an alternative to diesel generators for commercial backup power.

Except for solar photovoltaics, all grid electricity is produced using ***generators*** which convert kinetic energy into electric potential energy. They are the reverse of motors, and work because *a changing magnetic field produces an electric potential in a loop of wire.* If we rotate a loop of wire between the poles of a magnet, we'll get a voltage across the loop, and can cause current to flow. This can again be done with >95% efficiency.

Most transformers work by using the input electric voltage to produce a current, which makes a magnetic field, which is threaded through another coil of wire to generate an electric potential again. Since only *changing* magnetic fields make electric potential, a transformer *requires that the input voltage be constantly changing.* This is the reason that all large electric grids operate on ***alternating current (AC):*** the voltage (and the currents it causes) change directions. US electric grids run at 60 Hz, meaning they go back and forth 60 times per second. European grids run at 50 Hz. A transformer can change the European 240 V supply into 120 V or vice-versa, but cannot change 60 Hz to 50 Hz.

Most electronic devices don't care about the difference between 50 Hz and 60 Hz supply, though, because they are coverting the AC supply into ***direct current (DC)*** electricity, *i.e.* unchanging voltage, for internal use (this is called ***rectification***). My laptop charger's input is (in the US) 120 VAC, meaning volts alternating current, from the wall socket. The output is 20 VDC, or volts direct current, meaning that it provides a steady output voltage.

In the 1880's there was a "War of the Currents" with Thomas Edison and General Electric on one side—who were advocating use of DC electrical distribution—and AC inventor Nikola Tesla of Westinghouse on the other side. To advocate for DC, Edison staged press events around the country at which they would electrocute horses and dogs with AC to show how dangerous it was. He also "invented" the electric chair that was used to kill a convicted murderer in 1890. But as we have seen, both AC and DC can be lethal. And AC won the war, because it allows construction of transformers and allowed power plants to be more than a few km from the customer. AC motors and generators were also simpler than their DC counterparts in the pre-electronics era. In 1903, after Westinghouse equipment had already been installed to transport electricity from a Niagara Falls hydroelectric power plant to Buffalo, Edison filmed the use of AC electricity to kill a circus elephant that had killed three men. Needless to say, this did not have the desired effect of convincing the world to switch to DC electric distribution.

Practice problems

- What happens if you plug a 100 W US incandescent bulb into a European outlet?
- Why don't the birds that sit on high-tension power lines get electrocuted?
- A 750,000 V high-tension power line carries 100 MW of power to Philadelphia from a power plant 100 km away. Estimate how many ohms of resistance the power line has.

9

Power plants and carbon capture

Electric power plants consume 37% of all primary energy in the US. Owners of the power plants and the electrical transmission system are given the task of insuring that electricity is always available to anyone flipping a switch. Electricity is the ultimate in energy source for convenience, usefulness (low entropy), and cleanliness—at least at the site of use. Power plants are where the dirty work of producing electricity from primary energy sources occurs. In this Chapter we will examine how fossil-fuel power plants work, the economics of building and operating them, and the possibility that CO_2 emitted by power plants could be captured and sequestered, keeping it out of the atmosphere for the thousands of years it would otherwise be affecting the climate.

Slaves to the demand curve

The most important thing to know about power plants is that **there is currently no ability to store electricity on a large scale.** This means that electric power generators must constantly produce just the right amount of electricity to meet the demand of the moment. Demand for electricity has a daily cycle—higher in the daytime when offices are running, and evening when lights and appliances are on at home—and also seasonal variations, with the highest electricity demands on hot summer afternoons when air conditioning reaches its peak. You can get a view of the current US electric demand at this site. Figure 9.1 shows a typical ***demand curve*** for some week at some location. The absence of substantial storage capability on the grid has several important implications:

- Utilities must have access to enough generating *capacity* to provide the highest energy demand that will occur—probably on a very hot summer weekday afternoon. The peak demand in 2020 was

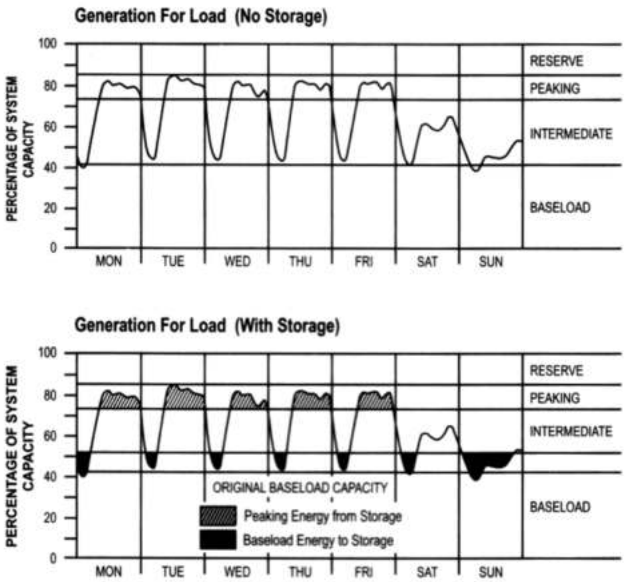


Figure 16.1. Typical weekly load curve of an electric utility. Source: National Academy of Sciences (1976).

Figure 9.1: Top: an example electricity demand curve for a week. Bottom: Illustration of how storage can increase useful base load and decriase peak power needed.

704 GW (with little change since 2005). US utilities have a total electric generating capacity in 2020 of 1.11 TW. At any given time, some facilities are unavailable due to maintenance, malfunctions, obligations for export, etc.

From this EIA site's data.

From the EIA's Electric Power Monthly

- Most generating facilities will be idle for major fractions of their time, *i.e.* will be operating with *capacity factor well below 1.* A generator that is not producing electricity is not making money for its owner! The total amount of electricity generated in the US in 2020 was about 4.01 trillion kWh (14.4 EJ), which works out to an *average* power of 458 GW. But the total generating *capacity* noted above is is 1110 GW so the average capacity factor is 41%. US usage has been fairly level since about 2007 after many decades of continuous growth.

EIA MER Table 7.1

- Power plants that take a long time (days or many hours) to turn on or off are limited to providing ***base load*** power to the grid, the amount of power that is at or below the daily low point of the demand curve. If facilities that cannot adjust their power level are the only ones running are running when demand drops below supply, they will have to "dump" (waste) their electric output somewhere, generating no income for this electricity. Base load plants operate at high capacity factor, >80%.
- Power plants that are *dispatchable* can be used for ***peaking power,***

the power between the low and high points of the demand curve.

- At any given time, the utility wants to generate electricity using its facilities that produce energy at the *lowest marginal cost per kWh, i.e.* the cheapest power sources turn on first. Base load plants should have low cost per kWh, peaking power higher, and the most expensive sources will be turned on only a few afternoons per year.
- In most of the US, there is a minute-to-minute bidding system for electric power. Wholesale electricity prices are much higher at periods of peak demand than at night. A gross example (and a good argument for price regulation) was when the spot price for 1 kWh of electricity rose to $900 in Texas during the 2021 freeze, compared to a normal cost near $3/kWh.
- The ***capital costs*** of a power plant depend on its capacity and the cost per GW of building that particular kind of power plant. Because the revenue depends on the actual electricity generated, those technologies with highest capital costs tend to be economical only if used as base load, with high capacity factors.
- Rarely-used peaking plants, on the other hand, would tend to have *low capital* costs and *high operating* costs, *e.g.* for fuel.
- At periods of high demand electricity prices rise and become very sensitive to the supply. Furthermore, the facilities producing the "last kilowatt" probably do not ever pay back their owner's capital investment and so the generating companies will have little incentive to build them without government intervention to assure that peak demand can be met. Electricity sellers might be happier with a shortage of supply that drives prices very high, so they have incentives to make supply unavailable at such times. Supply manipulation (by Enron and others) is in fact what caused the California electricity shortages of the summer of 2000 and 2001 that ultimately led to Gray Davis being recalled as governor and replaced by Arnold Schwarzenegger.

Coal and especially nuclear power plants have high capital costs, but the cost of fuel is low for coal and nearly zero for nuclear. They are also large, complex facilities that take many hours or days to start up. Hence they are useful almost exclusively for base load.

Hydropower is highly dispatchable so makes a good peaking source. If you have a certain amount of water running through the river, *i.e.* a fixed "fuel" supply, you probably want to sell the electricity it generates at the time when it's most expensive. However we

Base	Peak
High capacity factor	Low capacity factor
Low fuel cost	High fuel cost
High capital cost	Low capital cost
Low dispatchability	High dispatchability
Coal, nuclear, NGCC	NGCC, hydro, oil

probably never want to totally shut off the hydropower output since that would dry up the river downstream.

Oil and natural-gas power plants were the primary sources of peaking power for many years. However the cost per Joule of oil became much higher starting in the 1970's so utilities only use the old oil plants as a last resort, on the hottest days. Natural gas is well suited to peaking power since the gas turbines (see below) can be turned on or off in minutes, and the capital costs of gas plants is much lower than coal because the fuel is clean. There is no ash or pollutants to remove. In the past few years in the US, the price of natural gas has become so low that gas-powered electricity is cheaper than coal power in many cases. This has led to natural gas plants being built to take over base load as well as peaking duties. Older coal plants are being retired—for reasons entirely unrelated to CO_2 emission.

Wind and solar power present a conundrum for electric utilities. They have zero fuel costs of course, all the cost is in the up-front capital. But their capacity factors are limited by the fact that wind and sun are variable. The power output from these renewables is not dispatchable, as we cannot control the sun or wind, and we cannot even fully predict when it will occur. Solar power has the advantage of being strongest at times of high electric demand (hot summer afternoons), but it won't help with the early-evening peak from lighting and appliance demand. If the solar cannot be relied upon, then the utility must have other capacity in place to meet demand on cloudy days. So they are stuck paying capital expenses for *two* plants. Of course wind has the same issues.

The cloudy (low-wind) hours of one solar (wind) site will differ from those at another widely separated site. A portfolio of wind and solar spread across a country or continent provides a steadier total output, ameliorating the dispatchability problem. It is estimated that *up to 20% of electrical power capacity can be in wind and solar plants* without having major problems with the demand curve. Beyond that we would have trouble with *e.g.* having too much electricity in the middle of the night—which means that wind or base-load plants cannot sell their product—or falling short in peak-demand periods without duplicating the renewable capacity with backups.

Utilities (and customers) can save a lot of money by flattening the demand curve, reducing the peak capacity needed and increasing the amount that can be provided by inexpensive base-load plants and/or undispatchable renewables. One method is ***demand management*** whereby the utility encourages customers to reduce usage in peak periods. For instance they may offer a factory cheaper electric rates year-round if they are willing to shut the factory down on peak-demand days. This can be a net win for both parties. Residential customers of some utilities can obtain *demand metering* which charges different rates for night-time and daytime usage. Penn has a deal with their electricity supplier to cut back electric usage at peak periods, which is why my office can get very warm in the summer.

Potential for utility-scale electricity storage

The other option would be to build large-scale, dispatchable energy storage onto the grid. To be economical, the cost of the energy storage must be less than the difference in price between the off-peak electricity that they buy and the peak electricity they sell.

One clever technique is ***pumped storage,*** whereby one has two adjacent reservoirs at different elevations. At peak demand water goes from the upper to the lower reservoir just like a regular hydroelectric power stations. But at night, the turbine/generators are reversed to use electricity from the grid to push water into the upper reservoir, essentially storing the electricity as gravitational potential energy. The total energy that can be stored is determined by the volume of the (smaller) reservoir and the elevation difference between the two (the head). Because both the pumping and the generating phases are ≈90% efficient, the pumped storage facility can return ≈80% of the energy it "borrows" from the grid.

The highest-power facility in the US is the Bath County Pumped Storage Station in Virgina, completed in 1985. It can deliver 3 GW of power for 10 hours. Wikipedia lists ≈70 pumped-storage facilities worldwide with generating capacity $\geq$1 GW in operation as of 2020, with a large number of other projects in planning stations, especially in China. The peak power from pumped storage of perhaps 100 GW is currently a small fraction of the $\approx$ 7.5 TW of global capacity in 2019. For those plants listing total storage capacity, they can typically provide their peak power for 10 hours. Could this be scaled up to provide 10% or more of peak electrical demand and allow expanding wind and solar? In countries with favorable geography, probably yes. MacKay gives a useful analysis for the United Kingdom in his Chapter 26.

There are other potentially economically viable grid-storage technologies. Demonstration plants are in operation for *compressed air storage* whereby energy is stored as pressure in large underground caverns, *e.g.* salt domes. Of course the first thing that one would think of is battery storage. Until very recently this would have been far above a practical price, but battery-backed storage started to be produced at utility scale due to rapidly dropping prices of lithium batteries. cheaper to enable grid-scale storage. Currently (2021), you can buy a residential battery from Tesla with 13.5 kWh usable storage for $\approx$ \$8000 or about \$600/kWh, which is still about 4500× the typical retail cost of buying electricity, \$0.13/kWh. This means that even if renewable electricity were free, and you fully charged and discharged the battery once per day, it would take 4500 days (12.5 years) before renewable+battery-stored electricity were cheaper than grid power from current base-load fossil-fuel plants (even disregarding the discounting of future savings against present value). Current (2020) estimates of costs for lithium battery packs are near \$137/kWh, $> 4\times$ lower than the Powerwall, with some sales at \$100/kWh. A company called Nantenergy claims to be able to produce zinc-air batteries at a cost under \$100/kWh, as do some other companies now building demonstration projects. If prices $\leq$ \$100/kWh become available at utility scale, the economics of non-dispatchable renewable sources will be much more favorable in coming years. Annual additions of battery capacity surpassed 1 GW in 2020 for the first time as several large installations come online. Over a decade of daily charge/discharge cycles, a battery costing \$100/kWh stores electricity for just \$0.027 per kWh of electricity delivered to the grid. This is a very exciting prospect, which we will discuss more when we cover solar and wind power.

Fossil-fuel power plant principles

Coal Plants

Figure 9.2 illustrates the major elements of a modern coal-fired power plant. In the center (1) is the boiler, where powdered coal is blown in with the air needed to provide the oxygen for burning. The chemical energy is converted to heat here, and the flue gases are circulated through pipes filled with water that absorb this energy to produce high-temperature, high-pressure steam. The energy flows to the right, where the steam is used to run ***steam turbines*** (8). Figure 9.3 shows the fan-like structures inside the turbines that use the pressure of the steam to rotate a shaft. The shaft spins a generator (13) to convert the energy to electricity. At the output of the turbine the steam is condensed (and pressure lowered) by circulating around pipes filled with cold water, which also carries away the waste heat. The turbine water is returned to the boiler. While the turbine does not have the pistons that ran early steam engines, they are nonetheless heat engines and must obey the 2nd Law. This means that they (a) must have a place to dump waste heat, and (b) will be more efficient if the input steam temperature (and pressure) are higher and the outputs lower. Getting high efficiency is an engineering challenge since the

Figure A-3.B.1 Schematic of an Advanced, Low-Emissions, Pulverized Coal Unit

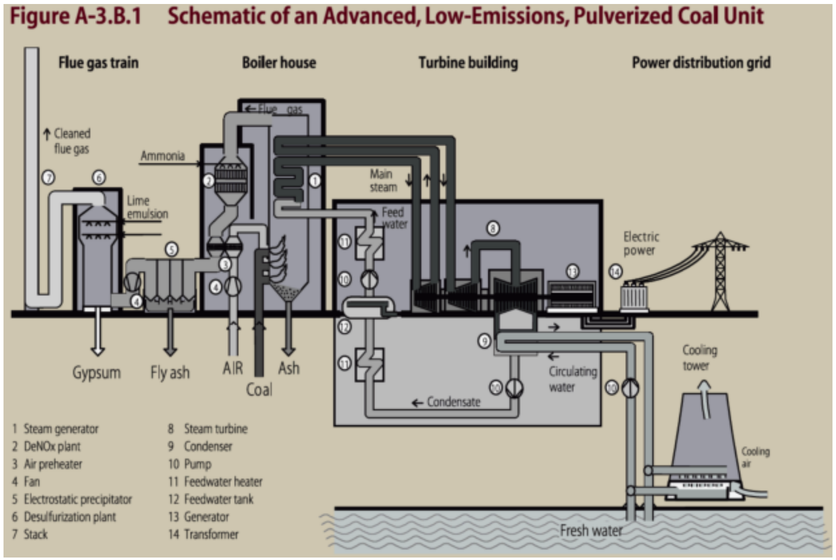


Figure A-3.B.3 Ultra-Supercritical 500 MW_e Pulverized Coal Unit without CO_2 Capture

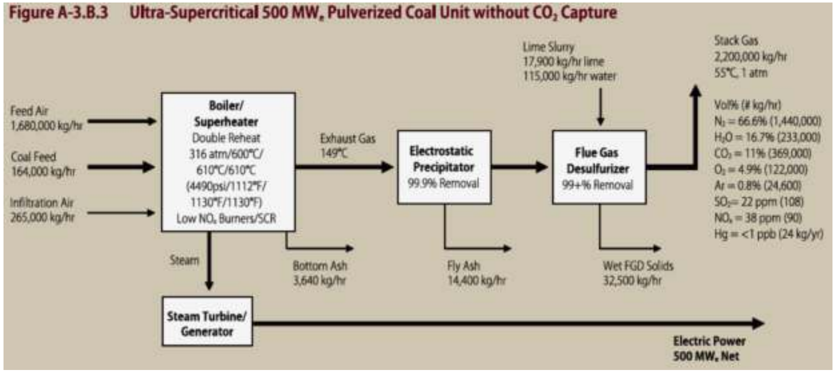


Figure 9.2: Top: schematic diagram of the components of a coal-fired power plant. Bottom: an accounting of the inputs and outputs of a state-of-the-art plant. From the *MIT Report on the Future of Coal*.

materials of the boiler and turbine must withstand high temperatures and pressuring without failing or wearing out. Advances in material fabrication make it possible to build the "ultra-supercritical" (USC) pulverized coal (PC) plant with 600° C steam described in the figure. New high-temperature plants can 43% efficiency in comparison with the 34% typically obtained from subcritical (SC) plants that form the bulk of existing US coal plants.

Figure 9.3: At left: the turbines and generator (rearmost cylinder) for one of the coal units at the Eddystone Generating Station just south of the Philadelphia airport. At right, plant manager Bryan Bennett of Excelon stands in front of a fan that has been removed from one of the turbines. The plant, built in the 1960's, has since been decommissioned. Photos by Gary Bernstein.

To the left of the boiler in Figure 9.2 we follow the waste products of coal burning. The heavier elements locked into the coal form solid particles that remain after burning. Some drop to the bottom of the boiler and are removed (ash) but the rest are suspended in the outflowing flue gas. These "particulates" cause smoke and smog, so modern plants must remove them with an electrostatic precipitator (3). Similarly most newly constructed plants must run the flue gas through chemical solutions that remove nitrous oxides (2) and sulfur dioxide (6) before they reach the atmosphere, as these are local/regional pollutants with significant health and environmental effects. The bottom panel of Figure 9.2 shows that this plant needs *164,000 kg of coal per hour* to provide the energy needed for 500 MW of electric power. The boiler produces 18,000 kg of ash that are collected with remarkable efficiency: only about 15 kg/hr of ash escape the flue. And only ≈100 kg of SO_2 and NO_x escape the flue. So what *does* come out of the smokestack? There are the CO_2 and H_2O that are products of complete hydrocarbon combustion, plus the N_2 that was in the input air and the 20% of the air's O_2 that was not consumed in combustion—over 2000 tons of gas per hour.

Natural Gas Plants

Figure 9.4 shows that natural-gas-fueled power plants do not need a seperate boiler, because the gas lacks the ash and impurities of coal that would damage a turbine. So in fact the NG can be burned *directly in a turbine.* The exhaust leaving this first turbine is still hot and

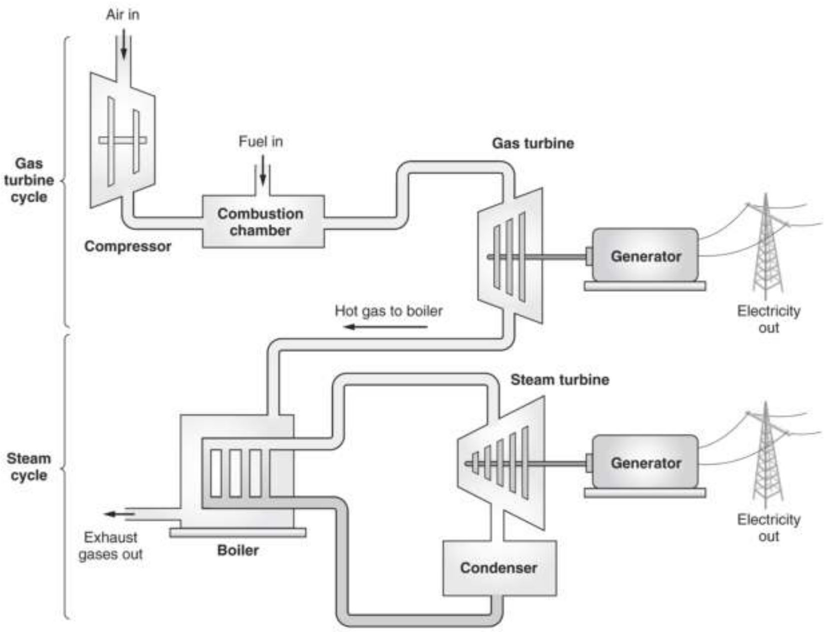


Figure 9.4: Schematic diagram of natural gas combined-cycle power plant.

can be used to generate steam to run a second turbine. This ***natural gas combined cycle (NGCC)*** power plant has a typical efficiency of 55%—much better than a coal plant. It therefore *requires fewer Joules of NG than coal to make a kWh of electricity*, lowering the relative fuel costs of gas plants. The CO_2 emitted per GJ of electricity (which we will write as GJe) is

$$\frac{\mathrm{kg}\,CO_2}{\mathrm{GJe}} = \left(\frac{\mathrm{kg}\,CO_2}{\mathrm{GJ\ fuel}}\right)\left(\frac{\mathrm{GJ\ fuel}}{\mathrm{GJe}}\right) = \left(\frac{\mathrm{kg}\,CO_2}{\mathrm{GJ\ fuel}}\right)\frac{1}{\mathrm{efficiency}} \tag{9.1}$$

$$= \begin{cases} \frac{88\,\mathrm{kg}\,CO_2}{\mathrm{GJ}}\frac{1}{0.34} = \underline{\underline{\frac{259\,\mathrm{kg}\,CO_2}{\mathrm{GJe}}}} & \text{for older coal plant,} \\ \frac{51\,\mathrm{kg}\,CO_2}{\mathrm{GJ}}\frac{1}{0.55} = \underline{\underline{\frac{93\,\mathrm{kg}\,CO_2}{\mathrm{GJe}}}} & \text{for NGCC.} \end{cases} \tag{9.2}$$

So we can *eliminate 2/3 of CO_2 production by replacing an old coal-fired power plant with NGCC!*

On General Electric's product page for gas turbines, it is claimed that their NGCC products have now demonstrated nearly 58% efficiency These turbines have input temperatures of $\approx 1800\,\mathrm{K}$, so Carnot's limit says efficiency could theoretically be $> 80\%$.

GE quotes 64% efficiency: this uses the lower heating value (LHV) of natural gas. We need to divide by 1.11 to get the real thermodynamic efficiency using the higher heating value (HHV), which includes the energy obtained from cooling the water in the combustion exhaust.

The cost of electricity

If you own a power plant (or are deciding whether to build one), you need to know what price you will have to charge for your electricity in order to make a profit. Let's calculate this cost of electricity ***(CoE).***

First we want to calculate the total amount of electric energy E that you will produce in a year. This will be

$$E = f P_{\text{max}} \times (1\,\text{yr}), \tag{9.3}$$

where P_{max} is the *capacity* of your power plant—its maximum possible power output, and f is the *capacity factor*, the fraction of its capacity that is actually produced, on average. The total annual revenue from selling this electricity will be

$$\textit{Revenue} = \textit{CoE} \times E = \textit{CoE} \times f P_{\text{max}} \times (1\,\text{yr}). \tag{9.4}$$

You have several kinds of costs:

- The **capital costs** must be paid when the plant is built, perhaps by selling bonds that you will then pay back to investors at some interest rate r over time T. The up-front cost PV of the plant can be written as $PV = CP_{\text{max}}$, where C is the capital cost per unit capacity, *e.g.* \$2000 per kW. Naturally we expect the cost of a power plant to grow with its capacity, *e.g.* do we need to buy 4 turbines or 3? During the bond payback period we will need to pay an annuity A given by the annuity formula. So our annual payments on capital are

$$\textit{Capital} = A \times (1\,\text{yr}) = \frac{rCP_{\text{max}} \times (1\,\text{yr})}{1 - e^{-rT}}. \tag{9.5}$$

- **Fuel costs** depend not on the capacity, but on the total energy used. If F is the cost per unit energy in the *fuel*, then the amount we pay to produce E in *electricity* is increased by the efficiency ϵ so that

$$\textit{Fuel} = \frac{FE}{\epsilon} = \frac{fFP_{\text{max}} \times (1\,\text{yr})}{\epsilon}. \tag{9.6}$$

- Fixed **Operations and maintenance (O&M)** are annual costs that must be paid to keep the plant working whether it is used or not—paying the employees who run the plant, keeping the machines oiled and repaired, etc.[1] As with the capital costs, we will usually give the O&M rate per unit of capacity. So the total annual costs will be

$$\textit{Ops} = \textit{O\&M} \times P_{\text{max}} \times (1\,\text{yr}) \tag{9.7}$$

Now we are prepared to calculate the CoE by setting annual *Revenue* equal to the total cost $\textit{Capital} + \textit{Fuel} + \textit{Ops}$. We obtain

$$\textit{CoE} = \frac{rC}{f\left(1 - e^{-rT}\right)} + \frac{F}{\epsilon} + \frac{\textit{O\&M}}{f}. \tag{9.8}$$

For this equation we need to be a little careful of the units, because the interest rate r and the $\textit{O\&M}$ rate are usually given per *year*,

[1] There are also *variable* O&M costs that depend on how much the generator is run during the year. But these costs are usually small compared to the fuel costs so we will ignore them, for simplicity.

whereas the CoE is usually desired per kilowatt-*hour*. A factor of $365 \times 24 = 8760$ is required. Here is a version of the formula that works if:

- CoE is desired in dollars per kWh of electric energy,
- C is in dollars per kW of capacity,
- F is the fuel cost in dollars per kWh,
- $O\&M$ is the cost in dollars per year per kW of capacity:

$$CoE = \frac{rC}{8760f\left(1-e^{-rT}\right)} + \frac{F}{\epsilon} + \frac{O\&M}{8760f}. \tag{9.9}$$

We can see here now why we want a high capacity factor to keep electricity cheap, especially for sources with high capital costs.

Type	Capital costs ($/kW)	Fixed O&M ($/kW/yr)	Fuel Cost ($/kWh)	Efficiency
Advanced (USC) coal	3640.	42.	0.0069	39%
Natural gas combined cycle	960.	12.	0.0099	55%
Nuclear	6300.	120.	(small)	N/A
Onshore wind	1846.	26.	0.	N/A
Solar PV	1250.	15.	0.	N/A
Solar PV w/storage	1600.	32.	0.	N/A

Table 9.1: Cost components for power plants estimated 2022–2023, from the EIA Annual Energy Outlook. The USC coal figures are from an older issue. Fuel costs are those paid by US power plants in 2019, from the EIA Monthly Electricity Report. These costs should be taken as indicative since they are estimated and can vary substantially.

Now let's get to the bottom line for coal and natural gas. Table 9.1 gives estimates for construction of state-of-the-art power plants as of 2021 from the EIA. We can apply Equation (9.8) to the coal and NGCC plants to obtain estimates of CoE at present prices, assuming two different capacity factors that might be indicative of base load use and peaking use.

Fuel	CoE, $f = 85\%$ ($/kWh)	CoE, $f = 35\%$ ($/kWh)
Coal	0.055	—
Gas	0.028	0.042
Nuclear	0.070	—
Wind	—	0.047
Solar PV	—	0.031
PV w/storage	0.31 ($f = 50\%$)	—

Table 9.2: Cost of electricty calculated from Table 9.1 inputs. We also assume $r = 5\%\,\mathrm{yr}^{-1}$, $T = 30\,\mathrm{yr}$.

We see that *NGCC is currently cheaper than a new coal plant even for base-load usage.* As a consequence, one would build a new coal plant only if you thought that gas prices were likely to rise relative to coal over the (several-decade) lifetime of the plant. In fact very few new coal plants are being built, and most coal plants in the US are quite old. Indeed at this point *new NGCC provides cheaper baseload power than running an* ***existing*** *power plant*, i.e. one whose capital costs are already paid off. As the coal plants get old and more expensive to keep running, or in need of environmental upgrades, they become candidates for decommissioning. This has been happening over the past few years in the US. Figure 9.5 is revealing: aside from a few plants around 2010, the commissioning of new coal power plants in the US stopped over 20 years ago. Most of US coal capacity (as of 2015) was built before 1990, with the typical plant being 40 years old. The vast majority of new capacity added since 1993 has been gas, with wind power showing up since 2006 and solar more recently. Indeed since 2015, more kWh of electricity are generated by natural gas than by coal in the US.

EIA MER, Table 7.2a

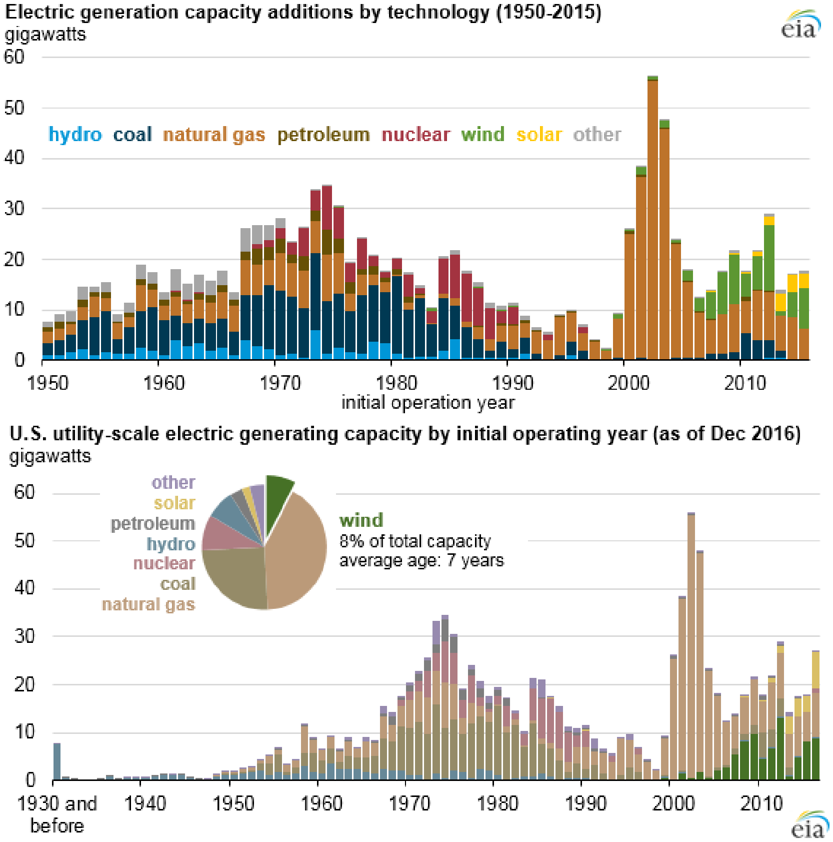


Figure 9.5: From the EIA: the top figure shows how much of 2015's electrical generating capacity of each type of power plant was commissioned in each year. At bottom, only those plants still in operation are shown. The hydro plants are very old, the coal and nuclear plants are old, and almost all new capacity additions in the past 20 years have been gas, wind, or solar.

Another stunning result of Table 9.2 is that *wind and solar power are now cost-competitive or cheaper per kWh than NGCC for intermittent use, and solar with storage is becoming competitive for higher capacity factors.* We will discuss the advent of these zero-carbon alternatives in later chapters.

NGCC plants have other attributes that make them more attractive to utilities than coal:

- The individual generating units can be 10's of MW each instead of 100's of MW for coal. Utilities prefer to add capacity in small increments so that they do not overshoot demand or require as much capital to be acquired.
- NGCC plants are less polluting and smaller, hence easier to find sites for without local opposition. There is no ash to dispose of.
- NGCC generators are "off the shelf" purchases—much easier and faster to install, with much more predictable prices. This makes it easier for a utility to match growing demand, and generate income on investment sooner, and reduce risk.
- If there ever is a tax on carbon emission, the NGCC plants will pay 3× less of it.
- NGCC is dispatchable much more quickly than coal and can fill the time gaps when wind and solar are not productive.
- The US Environmental Protection Agency (EPA) issued its "Clean Power Plan" (CPP) in Aug 2015, which would require each state to meet particular standards in reductions of CO_2 emission from power plants. Standard coal plants are disfavored, of course. The Trump administration EPA replaced the CPP in 2019 with much weaker rules, and in Jan 2021 the newer rules were struck down by a federal court. So the regulatory environment for coal has been unstable.

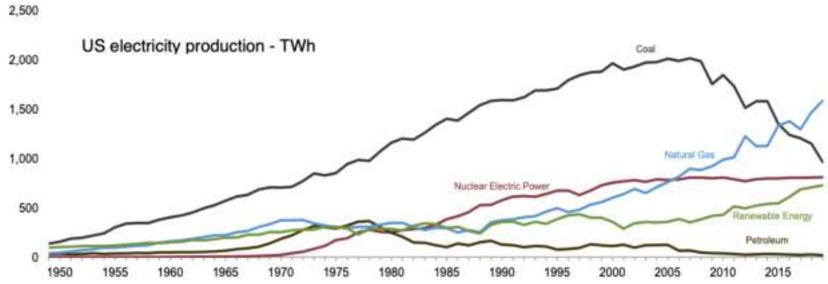


Figure 9.6: US electricity production history from different fuels, in TWh, showing natural gas surpassing coal as the largest electricity source in 2015. In 2020 (not yet on the plot) coal dropped below nuclear and renewable electricity, down 42% from its peak! From EIA MER Feb 2021, Figure 7.2.

For now, natural gas is king of US power generation, and coal's market is shrinking rapidly in the US *for reasons having nothing to do with climate change.* Figure 9.6 shows the precipitous decline of coal-fired electricity in the US. Gas surpassed coal in 2015, and *more electricity was made from* ***each*** *nuclear and renewable fuels in 2020* for the first time in history. In 2017–2021 the Trump administration made many efforts (often successful) to change the regulatory environment to favor coal use, with the Secretary of Energy literally working from a list of changes provided by coal executives. These include not only scrapping the Obama Clean Power Plan, but also forcing electric utilities to purchase coal-fueled electricity at a price premium over natural gas or renewable electricty, and relaxing regulations that require installation of equipment to remove pollutants from coal plant exhaust. Thus these moves are aimed at working against the marketplace to *increase* the cost of generated electricity in order to enable *greater* environmental damage and climate change. Yet the decline in coal electricity generation has continued unabated and no significant new coal capacity is planned.

See this press release in advance of official EIA statistics.

Looking into the future, we see that NG's reign cannot last decades without an enormous increase in proven reserves relative to today. At some point, the price of NG will probably rise, and coal will become cheaper again. Furthermore, the situation outside the US is very different, with natural gas prices being much higher and the potential for future hydrofracturing expansions being much less clear. Developing countries outside the US, have rapidly increased coal usage in recent decades—more than half of increased coal usage is in Chinese power plants. The Chinese coal story is rapidly evolving. After an enormous expansion over the past 2 decades, for some of which China was opening more than 1 GW of new coal capacity *per week*, it has over 1000 GW of coal-fired capacity, more than 3 times that of the US. But the plants were overbuilt and run at only 50% capacity factor. In Jan 2017, the Chinese government announced the cancellation of $>$ 100 GW of coal power plants under construction or planned. The Chinese government has moved to close many existing underutilized, lower-efficiency coal power plants.

New York Times 17 Jan 2017

Center for American Progress, 15 May 2017

Yet most of the Joules in the ground are in coal, so eventually someone will want them and we will burn all we can find, unless (a) we have societal pressures that are stronger than the profits to be made supplying this energy and are effective worldwide, or (b) some renewable source becomes cheaper than coal worldwide.

Carbon capture and sequestration (CCS)

Motivation and context

If we want to keep the 530 GTC in coal reserves and 100 GTC or more in natural gas from going into the atmosphere and blowing past the CO_2 doubling limit, we can either hope that it can be kept in the ground, or we can try to burn it *but keep the carbon from escaping* *into the atmosphere.* ***Carbon capture*** is the removal of CO_2 from combustion products, and ***sequestration*** is storing it somewhere for more the 10,000 years so that it will not build up in the atmosphere.

In an ideal scenario, we would find carbon-free alternatives to all of the current uses of fossil fuels such that these sustainable sources are *cheaper* than coal, gas, or oil, and there will be no incentive to burn the fossil fuels. The most promising sources of 100's of EJ of carbon-free energy per year (solar, wind, nuclear) all produce electricity. So the reduction of fossil fuel use to near zero would require:

1. A massive shift (or reduction) of energy use for transportation and heating to electricity from direct fossil-fuel burning, *and*
2. Replacement of coal and natural gas power plants with renewables that are cheaper, *and*
3. Development of cheap, massive storage of electric energy to allow the not-very-dispatchable renewables to match the demand curve through the days and the seasons.

If we do not accomplish all of these, then we will have to continue burning fossil fuels, and CCS will be an essential technology for greatly reducing CO_2 emissions, which we want to accomplish in the next decade or so to avoid several more degrees of warming.

Clearly the implementation of CCS will impose additional costs on users of fossil fuels. Therefore they would never be installed voluntarily in a market-based system where emitting CO_2 has no cost to the emitter. The use of CCS for future fossil-fuel burning will occur only if regulations require it, or if there is some charge for emitting CO_2 that exceeds the cost of *not* emitting CO_2, *i.e.* by installing CCS. We will therefore be looking at CCS options for their equivalent cost per ton of CO_2 emission that they *avoid.*

CCS is most feasible for large, fixed sources of CO_2 emission, *e.g.* it is far more economical and practical for electric power plants than for home furnaces or for vehicles. The ideas for CCS were first investigated in the context of coal-burning power plants, which 15 years ago were by far the largest source of CO_2 emissions and the obvious target. In the US, coal power plants are rapidly disappearing, undercut by natural gas and, to a lesser degree, wind/solar power. The motivation for research and development of coal CCS has therefore been very much weakened in the US. But the switch to natural gas is less ubiquitous in other nations; and we also have to wonder how long natural gas will be cheap and abundant. So we should keep this technology in mind.

CCS research and development never had great political support in the US. Those who worried about climate change wanted to shut down coal plants and switch to renewables, not to build more coal plants. And the owners of coal mines and coal power plants didn't want to admit that climate change was a problem at all.

More recently, however, a good part of the focus on CCS development in the US has shifted toward implementing it for *natural gas* power plants. NGCC plants inherently emit less CO_2 per kWh of electricity than coal plants, which lessens the incentive for paying to capture it. However, gas plants now produce $2\times$ as much of US electricity than coal. And if we want to aim for carbon neutrality or even drastic reductions in the next $\approx$ 20 years while greatly expanding electrification of heating & transportation, we will need to get rid of the CO_2 emissions from NGCC plants somehow. The non-dispatchable nature of wind and solar electricity means that a large fraction of our power will need to come from gas, until such time as grid-scale storage is a reality.

Is this crazy?

With this in mind, let's look into the practicalities and costs of CCS. The idea of CCS seems preposterous: Figure 9.2 shows 369,000 kg of CO_2 emerging from the stack of a single high-efficiency 500 MWe coal power plant. Now consider that thousands of such units are necessary to provide most of future electrical demand. The idea of an infrastructure to put all that gas somewhere seems impossible. But then we realize that:

- CO_2 *liquifies* at room temperature when subjected to more than $\approx$ 60 times atmospheric pressure. This makes it denser than oil, and we know we can transport huge masses of oil around the globe economically.
- The same power plant takes in 164,000 kg of coal per hour—so we clearly have an economically viable infrastructure for transporting this already in place! Liquid CO_2 would be even easier to ship than coal since we could use pipelines.
- We *already* transport substantial quantities of CO_2 by pipeline or other means! Liquid CO_2 is used in enhanced oil recovery: it is pumped into wells to force oil out. There are many other smaller-scale industrial uses of liquid CO_2.

The transportation of liquified CO_2 over hundreds of miles from a power plant to a sequestration site is essentially a solved problem. So we have two issues to address: first, how to remove the CO_2 from combustion exhaust; second, where we can store it so that it will not leak out any faster than the 10,000-year timescale for reabsorption from that atmospheres into the deep oceans.

Carbon capture strategies

There are two approaches to obtaining a pure supply of CO_2 out of power-plant exhaust. The first is ***post-combustion*** capture, whereby we run the exhaust gas from an otherwise normal boiler or gas turbine through a device containing chemicals that bond to CO_2, as illustrated in Figures 9.7. The challenge is to separate the CO_2 from the nitrogen gas that came into the combustion chamber with the incoming air. There are well-known chemical processes for doing this, for example bubbling the exhaust through a solution of monoethanolamine (MEA) which will bond with the CO_2. The solution can then be removed, and will release the CO_2 when reheated. The MEA is then reused. The reheating ("regenerating") of the MEA consumes a substantial amount of heat, which eats into the energy available for electricity production and reduces the overall efficiency of the plant.

The second approach of interest for carbon capture is the ***oxy-fuel*** technique, whereby the nitrogen is separated from the air *before* the air enters the combustion chamber. If the input air is nearly all oxygen, then the exhaust combustion of coal or natural gas will be nearly entirely water and CO_2. Simply cooling the gas will condense the water, leaving a purified stream of CO_2 with no need for additional energy. We do pay an energy price, however, to run the "air separation unit (ASU)" at the input. Oxy-fueling has advantages: the power plant can often run more efficiently with pure-oxygen fueling. And without nitrogen input, there is no production of the pollutant nitrogen oxide (NO_x).

Carbon capture coal power plants

For this section we will consult the 2007 MIT report on *The Future of Coal*. The co-chair of this study, Ernest Moniz, became US Secretary of Energy in 2013, so until 2017 there was at least one person at DOE who takes the results seriously.

Figure 9.7 shows a block diagram and some statistics for a coal power plant with post-combustion carbon capture. The box labelled "CO_2Capture" in the left panel of Figure 9.7 represents this process. The heat to regenerate the MEA and release the CO_2 is obtained by diverting some of the steam that would otherwise be fed to the turbines. Therefore less electricity is generated, and the overall efficiency of the plant is reduced by 5%. Then the gaseous CO_2 must be compressed to liquefaction for transport. This requires electricity to run compressors, and in fact consumes 10% of all the electricity produced by the plant! As a consequence the overall efficiency of the plant at producing electricity *for the grid* is reduced. The USC coal plant's ef-

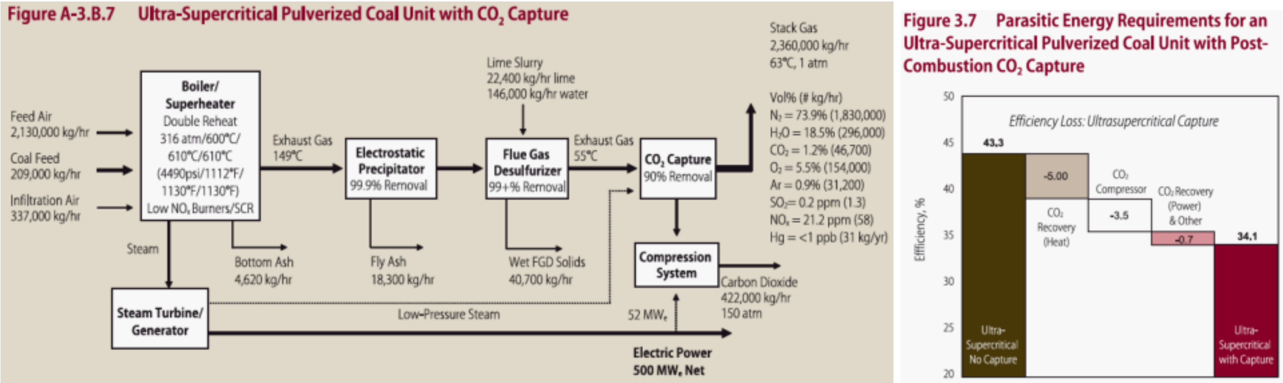


Figure 9.7: Schematic of extraction method for carbon capture, and resultant efficiency losses. From MIT 2007 *Future of Coal*.

ficiency is predicted therefore to drop from 43% without capture to 34% with capture.

Note that the success of CCS is greatly enhanced if we *start* with a high-efficiency plant, because higher efficiency means less coal is being burned which means there is less CO_2 that needs to be captured and less energy is needed to do it!

Implementing CCS therefore raises the CoE in two ways. First, the capital costs are higher because of the additional process required. Second, efficiency is lowered, so more fuel is needed. The good news for the coal miners is that if CCS is implemented more widely, they will sell more coal. When we calculate the reduction in CO_2 emission we must take into account that the lower efficiency means more CO_2 is produced before capture occurs. With this taken into account, the CO_2 released to the atmosphere per net kWhe for USC coal plant with CCS is reduced 87% from the no-CCS output for the same plant.

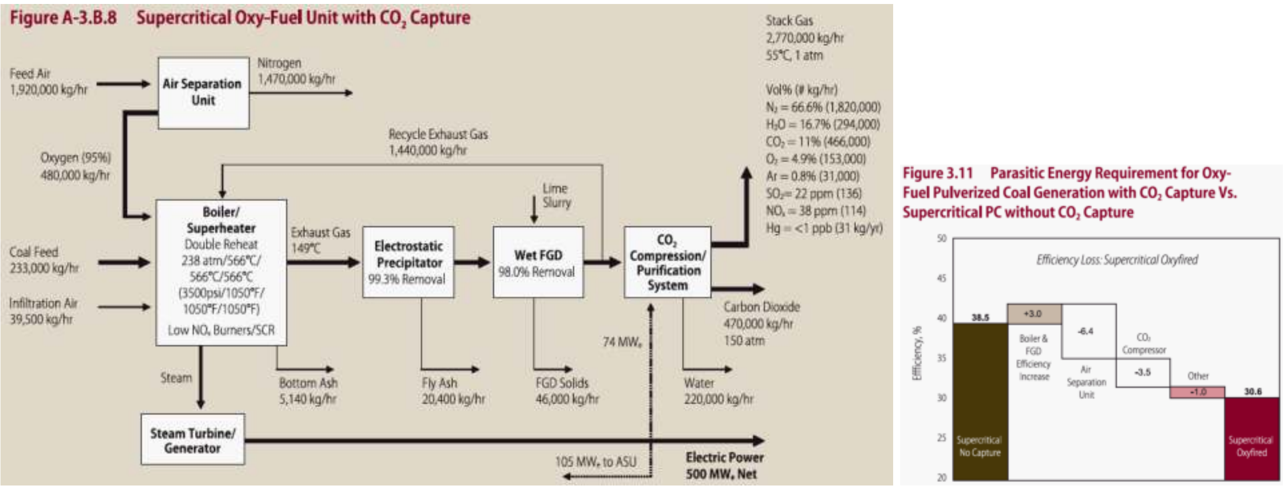


Figure 9.8: Schematic of oxy-fuel method for carbon capture, and resultant efficiency losses. From MIT 2007 *Future of Coal*.

Figures 9.8 illustrate an oxy-fueled coal power plant. Nearly 1/4 of the electricity from the generators must be diverted to run the air separation unit, and also to run the compressor, which lowers the net efficiency.

Another route to CCS coal power is called ***integrated gasification and combined cycle (IGCC)***. Here the coal is not burned directly, but first mixed with steam at high temperature to produce *syngas:* $C + H_2O \rightarrow CO + H_2$. Particulates, sulfur and other pollutants are more easily removed from the gasification process than doing this after combustion. The ***syngas*** can then be burned in a gas turbine just like natural gas, yielding high-efficiency generation as for NGCC. The IGCC process can be fed with pure oxygen from an air separation unit, so it produces nearly-pure-CO_2 exhaust for compression and transport to sequestration.

Attempts to build full-scale CCS coal power plants in the US have not been successful to date. As with any new technology, the first plants will be much more expensive and less reliable than established technologies, and utilities have no incentive to absorb these costs and risks in the absence of a price on CO_2 emission—especially when NGCC is cheaper than coal! The need to have the technology developed in the long term, plus the disincentive to action in the near term, make development of CCS a clear target for government to conduct or subsidize the R&D. The US DOE *FutureGen* project has sputtered in and out of existence as utilities, local governments, and US Presidents and Congress have blown hot and cold on the priority of developing CCS at commercial scale. This project was to retrofit an existing coal plant as oxy-fueled, and institue CCS. The project was suspended in 2015. The 110 MW Boundary Dam plant in Saskatchewan, Canada entered operations in October 2014, becoming the largest power plant to do so, and expecting 90% capture from an amine capture system on a pulverized coal boiler.

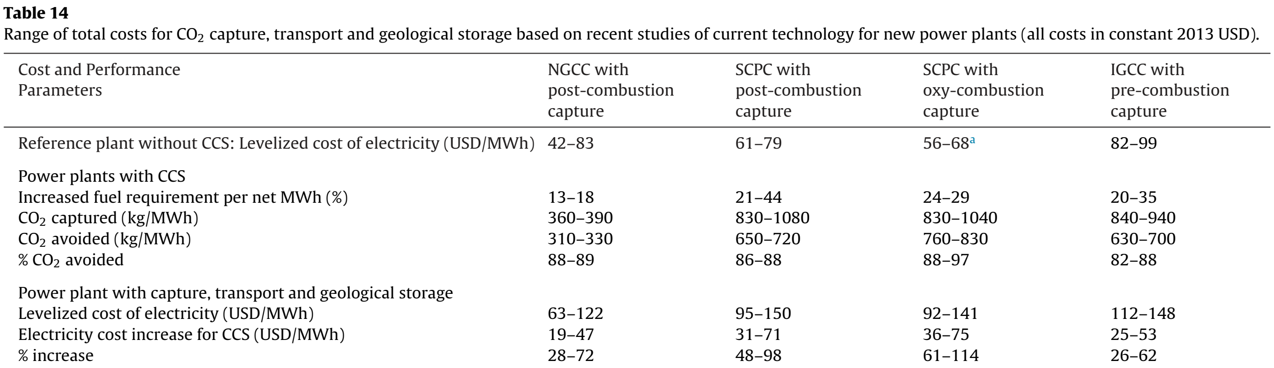


Table 14

Range of total costs for CO_2 capture, transport and geological storage based on recent studies of current technology for new power plants (all costs in constant 2013 USD).

Cost and Performance Parameters	NGCC with post-combustion capture	SCPC with post-combustion capture	SCPC with oxy-combustion capture	IGCC with pre-combustion capture
Reference plant without CCS: Levelized cost of electricity (USD/MWh)	42–83	61–79	56–68[a]	82–99
Power plants with CCS				
Increased fuel requirement per net MWh (%)	13–18	21–44	24–29	20–35
CO_2 captured (kg/MWh)	360–390	830–1080	830–1040	840–940
CO_2 avoided (kg/MWh)	310–330	650–720	760–830	630–700
% CO_2 avoided	88–89	86–88	88–97	82–88
Power plant with capture, transport and geological storage				
Levelized cost of electricity (USD/MWh)	63–122	95–150	92–141	112–148
Electricity cost increase for CCS (USD/MWh)	19–47	31–71	36–75	25–53
% increase	28–72	48–98	61–114	26–62

Figure 9.9: Some estimates of the efficiencies and costs for adding carbon capture and sequestration to fossil-fuel power plants of various types. From Rubin et al, 2015.

While these initial projects are far from cost-competitive with other coal plants, never mind NGCC, the MIT report estimates what the price premium for CCS coal power would be once the technology matures. Figure 9.9 is from a more recent analysis CCS economics by Rubin et al. (2015). Looking in the "SCPC with post-combustion capture" tells us about a Sub-Critical Pulverized Coal plant. Let's look at the optimistic end of the ranges given for various quantities. In the top line: without CCS, the COE for the coal plant is \$61/MWh, or \$0.061/kWh, not far from our estimate in Table 9.2. When the reduced efficiency and increased capital costs are taken into account, the bottom two lines suggest that the COE will be higher by \$0.031/kWh, or 48%. The good news is that the middle lines suggest that the CO_2 output per MWh of electricity sold is reduced by nearly 90%, saving up to 720 kg of CO_2 per MWh. The bad news is that no electricity producer operating on a profit motive would choose to make the more expensive form of electricity! If we want the power-plant owners to choose CCS coal over a non-CCS coal plant, we need to either require this, or to tax the emission of CO_2 to raise the expense of the non-CCS electricity. In order for this incentive to work, the carbon tax would have to be at least

$$\frac{\$31}{\text{MWh}} \times \frac{1\,\text{MWh}}{0.72\,\text{ton}\,CO_2} = \frac{\$44}{\text{ton}\,CO_2}. \tag{9.10}$$

This level of carbon tax would suffice to make it profitable to build a new coal power plant with CCS rather than without it. However we know that for now in the US, utilities don't want to build *any* coal power plants because the COE is much lower for natural gas power. And since NGCC plants emit less CO_2 than coal plants to begin with, there is no carbon tax that would favor the construction of a coal plant (CCS or otherwise) over an NGCC plant.

Where does the CO_2 go?

We need to find a place for our CO_2 which is under high pressure, and where escape of liquid or gaseous CO_2 to the surface is blocked. This sounds just like an oil or gas trap, and in fact *depleted oil and gas wells* are excellent candidates for CO_2 injection. At depths below 800-1000 m, the pressure is high enough to liquefy CO_2. And we already know this is feasible since CO_2 is injected into oil wells for enhanced recovery. But is it feasible to consider capturing most of the CO_2 from coal power plants now, or in an even more coal-dependent future?

A power plant with 1 GWe capacity, 85% capacity factor, and 34% efficiency, and 90% capture will produce roughly 1.7×10^7 kg of CO_2 per day, which when liquified at 2 km depth has a density of $\approx 1100\,\text{kg m}^{-3}$ and will need about 130,000 bbl/day of liquid injected into a well. The biggest injection well of any kind in the world is 40,000 bbl/day, and most are much smaller, so we would need 10's of wells connected to each large power plant by pipelines. Over its 50-year lifetime, a plant like this would need about 2 Gbbl of underground space to store its CO_2.

Roughly 200 Gbbl of oil have been produced in the US so far, so even if just 1/4 of this left behind suitable space for CO_2 sequestration, we could sequester the lifetime CO_2 output of about 25 GW of CCS coal plants. This is certainly enough to get started, but if we ended up heavily dependent on coal again, more space would be needed. Other issues to consider are that oil fields have holes drilled through their cap rocks that will leak the CO_2 into the atmosphere, defeating our goal, unless we know where they all are and plug them well enough to avoid leaks for thousands of years. Record-keeping of drilling may not be good enough. Also, we know that oil is distributed very unequally around the world, and therefore many coal-burning countries would not have many depleted oil and gas wells available for sequestration (such as China).

Fortunately there is another, much larger potential sequestration reservoir: deep saline aquifers. These are water-bearing porous structures that, like shale gas and oil, are far below the fresh-water aquifers from which we obtain drinking and agricultural water. Suitable formations are found on all the inhabited continents. Most estimates of the potential storage volume place it above our 2000 GTCO_2 target, although the estimates vary widely.

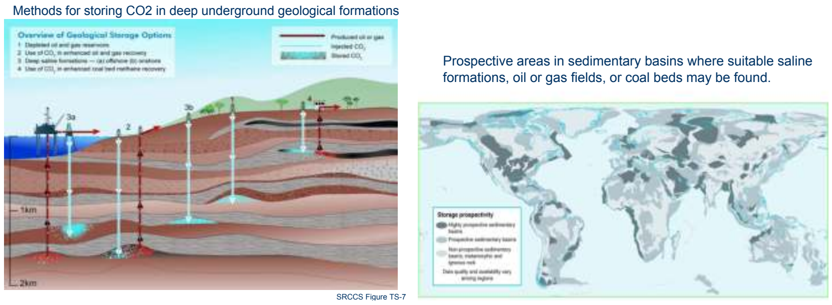


Figure 9.10: Potential sites for CO_2 sequestration and their distribution around the world. From the IPCC *Special Report on Carbon Dioxide Capture and Storage*, (2005).

Deep-ocean storage of CO_2 has also been suggested: CO_2 liquifies at the pressures present at the bottom of the ocean, and it is denser than water, so CO_2 injected to the ocean bottom will form "pools" of liquid on the ocean floor. Obviously the storage volume is practically unlimited. But it's not clear how quickly this CO_2 would remain confined or dissolve into the water, acidifying the vicinity or the entire ocean. Also it would be much more expensive to get the CO_2 to the deep ocean than pipelining it to a fixed set of wells on land.

In any case, the prospects look good for near- and longer-term availability of suitable geologic storage. A large, but not infeasible, new CO_2-handling infrastructure would need to be constructed at considerable, but not outlandish, expense. The estimates for cost of sequestration are smaller than the estimated costs of capture at the power plant.

We conclude that CCS for the bulk of coal usage (at power plants) is likely to be technically feasible at a cost of a few cents per kWh, and economically favorable to building non-CCS coal plants at a carbon tax of perhaps $20–40 per ton CO_2. Furthermore the utilities will need to foresee higher natural gas prices to choose CCS coal ahead of NGCC, even with a carbon tax in place.

It is also clear that significant research and development remains to be done to make CCS technology mature, and that this work will occur only if mandated or subsidized. Thus the days of "clean coal," meaning use of coal that puts less CO_2 into the atmosphere than alternatives, seem possible but far off, as there is currently little incentive or will to advance this technology, at least in the US. Yet the phrase "clean coal" lives on as a meaningless catchphrase used by those who wish to promote coal production, even as they actively oppose spending on development of the necessary engineering improvements.

CCS for natural gas

The same principles of post-combustion or oxy-fueled CCS may be applied to natural-gas-fueled power plants. This is great interest because of the potential need to maintain a fleet of NG plants to make up for the gaps between solar/wind production and the demand curve, in a scenario where we massively expand our production from renewables.

We can now repeat the exercise of asking what level of carbon tax would provide sufficient incentive to build (or retrofit) post-combustion CCS onto a NGCC power plant. Looking at the leftmost data column of Figure 9.9, we find that the optimistic estimate of non-CCS COE for an NGCC power plant is $42/MWh or $0.042/kWh. The addition of CCS is estimated to increase this by at least $19/MWh, while reducing CO_2 emissions by 0.33 tons per MWh. Paralleling Equation (9.10), we find that *a tax of $57 per ton of CO_2 emitted would be needed to incentivize the addition of CCS to NGCC power plants.* A different study places the effective cost of CCS for natural gas at $91 per ton of CO_2 emission avoided (or $41/ton after application of existing tax breaks.)

Running CCS on a natural-gas turbine is easier than for coal in some ways, since the exhaust from NG burning is free of the many chemical contaminants in coal exhaust. But the fact that NG exhaust has less CO_2 in it (due to the H content) also makes it a little more difficult to remove.

This excellent *Vox* article describes the successful first tests of a new type of oxy-fueled natural-gas turbine that is claimed to be capable of producing electricity at 55% efficiency with near-zero CO_2 emission, even accounting for the energy needed for air separation. A full-scale 300 MWe demonstration plant has been announced as under construction for 2022 trials, although this has probably been delayed by the pandemic. If this technique works without capital costs greatly exceeding current NGCC, it could be one of the key elements for reaching carbon neutrality in the next 20 years.

Practice problems

- Estimate the volume of the upper reservoir at the Bath County Pumped Storage Station.
- Estimate the cost of electricity generated from oil, assuming that capital costs are closer to those of natural-gas plants (since little pollution-control equipment needed) but an efficiency closer to a coal plant.
- Under what conditions would a Tesla Powerwall be an economical purchase for a grid-connected residence?

10

Wind power

Figure 10.1: A Dutch windmill (with sails furled).

Roughly 1% of all solar power incident on Earth goes into moving air. Only a small fraction of this wind power is near enough to the surface to be exploitable, but even this fraction is enough to supply much or even all of current human energy needs. Harnessing wind energy is not a new idea. Sailboats go back to 3500 BC or earlier, and 19th-century sailing ships harnessed 5–10 MW of wind power. Wind mills (for grinding grain) appeared as early as 2000 BC in China, and Dutch windmills began in 1100 AD and were built with up to 30 m rotor diameter and 5–30 MW shaft power, primarily for pumping water. But the arrival of steam power replaced wind as the primary source of non-animal power.

Windmills were used in the US Midwest in 1850–1930's to pump water and make electricity at locations off the electric grid. This wind era largely disappeared by the 1930's as even rural areas were put on the electric grid. Interest in wind generation of electricity was reborn in the 1973 oil crisis, and wind farms appeared in prime US locations such as Altamont Pass, California. But by 1996, low fossil-fuel prices had killed the industry and the largest manufacturer (Kenetech) was bankrupt.

The current, and by far the largest, wind-power era began in the late 1990's as concerns over CO_2 emission and energy independence led governments to mandate and/or subsidize the development of the industry. Wind power is now competitive with fossil-fuel electricity, or in some circumstances even cheaper, and in 2019 surpassed hydropower to become the largest source of renewable electricity in the US.

Figure 10.2: 5 MW wind turbines with 126-m rotor diameter. Each blade is about as long as the wingspan of a Boeing 747. This is built by Repower and installed in Germany. note the tractor at lower right for scale. Source.

Physics and engineering of wind turbines

Power and efficiency

How much power is available from extracting the kinetic energy of wind? Consider a turbine whose blades sweep out a circle of diameter D. If the wind speed is v, then the volume V of air that passes through the turbine in time T is $V = \pi D^2/4 \times vT$. The total power in the airstream is

$$P_{\text{air}} = \frac{E}{T} = \frac{\frac{1}{2}m_{\text{air}}v^2}{T} \tag{10.1}$$

$$= \frac{\frac{1}{2}\rho_{\text{air}}Vv^2}{T} \tag{10.2}$$

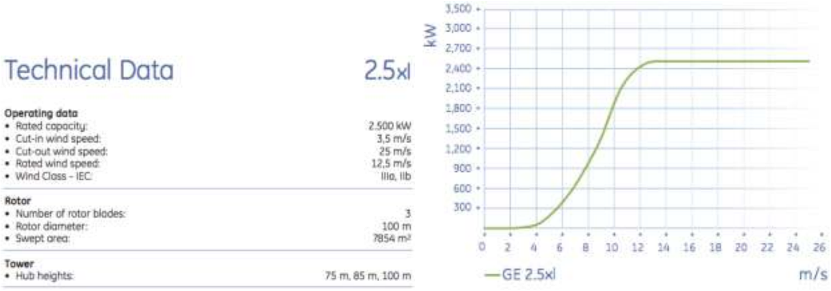
$$= \frac{\frac{1}{8}\rho_{\text{air}}\pi D^2 vTv^2}{T} \tag{10.3}$$

$$= \frac{\pi}{8}\rho_{\text{air}}D^2v^3. \tag{10.4}$$

Here $\rho_{\text{air}} = 1.226\,\text{kg m}^{-3}$ is the density of air at sea level. This density is lower at higher-altitude sites. The turbine will have some efficiency ϵ in extracting this power from the moving air, so the power output of the turbine is

$$P = \frac{\pi}{8}\epsilon\rho_{\text{air}}D^2v^3. \tag{10.5}$$

If a wind turbine had $\epsilon = 100\%$, the air would have $v = 0$ at the back end of the turbine. This would not work since the air has to keep moving to make room for new air to pass through! Under simple assumptions it was shown that the efficiency of a wind turbine must be $\epsilon \leq 59\%$. This is the ***Betz limit.***



Technical Data 2.5xl

Operating data	
• Rated capacity:	2,500 kW
• Cut-in wind speed:	3.5 m/s
• Cut-out wind speed:	25 m/s
• Rated wind speed:	12.5 m/s
• Wind Class - IEC:	IIIa, IIb
Rotor	
• Number of rotor blades:	3
• Rotor diameter:	100 m
• Swept area:	7854 m²
Tower	
• Hub heights:	75 m, 85 m, 100 m

Figure 10.3: Technical specifications (left) and power curve (right) for the GE 2.5xl wind turbine. From the GE brochure for this product.

Figures 10.3 offer a real-life example, for the General Electric model 2.5xl turbine. Its 100-meter rotor diameter is enormous, but in fact nearly all new installations are even larger. GE's largest model in 2021 has an incredible $D = 220$ m rotor diameter, for offshore installation, with 12 MW capacity (see Figure 10.4).

The power curve shows a rapid increase in P as v rises, which is expected from the v^3 that appears in Equation 10.5. We see that at wind speed $v = 10\,\mathrm{m\,s^{-1}}$, the power output will be about 1.85 MW. We can solve Equation 10.5 for ϵ to infer that the 2.5xl operates at an efficiency 38%. Real commercial wind turbines obtain efficiencies as high as 45%. Old Dutch windmills had $\epsilon \approx 5\%$.

Higher-efficiency wind turbines can be built. But it is more important to get the best efficiency per dollar than to get the highest possible efficiency.

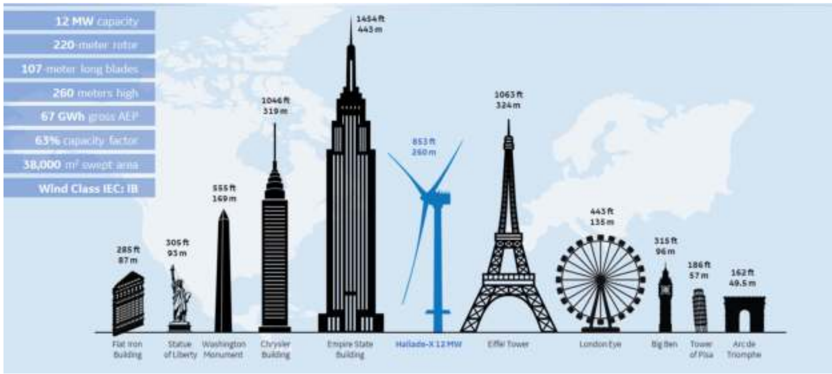


Figure 10.4: GE illustration of the size of their largest wind turbine in 2018, the Haliade-X 12 MW model.

At $v > 11\,\mathrm{m\,s^{-1}}$, the 2.5xl's output stops growing quickly and levels out at 2.5 MW. The turbine's blades are intentionally rotated ("furled") to lower the turbine's efficiency in high winds, because transmitting higher power would damage the turbine and its generator. Generating higher power would require a stronger, more expensive turbine, and this one is designed for sites where speeds in excess of 12 m/s are rare such that its not worth the extra construction expense to exploit them. At wind speeds above the "cut-out speed" of

$v = 25\,\mathrm{m\,s^{-1}}$ listed in the datasheet, the turbine blades are locked into a fixed position to avoid damage. There is also a "cut-in wind speed" of 3.5 m/s, below which the turbine does not bother trying to generate power.

The only time the 2.5xl puts out its full rated 2.5 MW capacity is when $12 < v < 25\,\mathrm{m\,s^{-1}}$. Most of the time the wind is below this level (otherwise we would have built a heftier turbine!), and so the capacity factor is necessarily $f < 1$. Figure 10.5 shows that recently installed real-life wind turbines in the US have capacity factors averaging above 40%. Sites where the wind is very steady can have higher capacity factors since the wind speed stays closer to the maximum that the turbine is designed to extract.

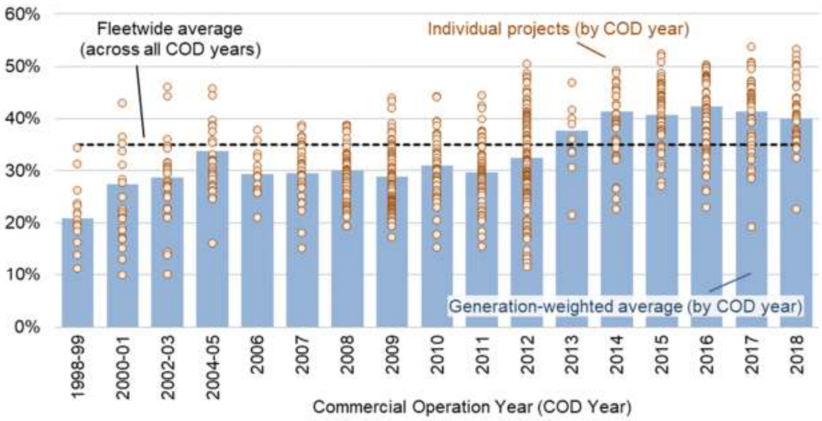


Figure 10.5: Capacity factor of US wind facilities in 2019 vs year of installation. Bars are the average, circles are individual installations. From Lawrence Berkeley National Lab, *Wind Technologies Market Report (WTMR*.

Power density and resource size

Now consider building a wind *farm* of many turbines. Because the turbine saps the air of its energy, neighboring turbines must be $\approx 5D$ away in order to avoid reducing the power of neighbors. This means we require a land area of $A = 25D^2$ per turbine on our wind farm. The total power output of the farm per unit area of

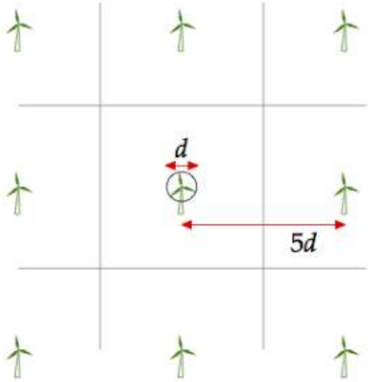


Figure 10.6: In a wind farm, turbines must be spaced by about 5 times the rotor diameter. From MacKay.

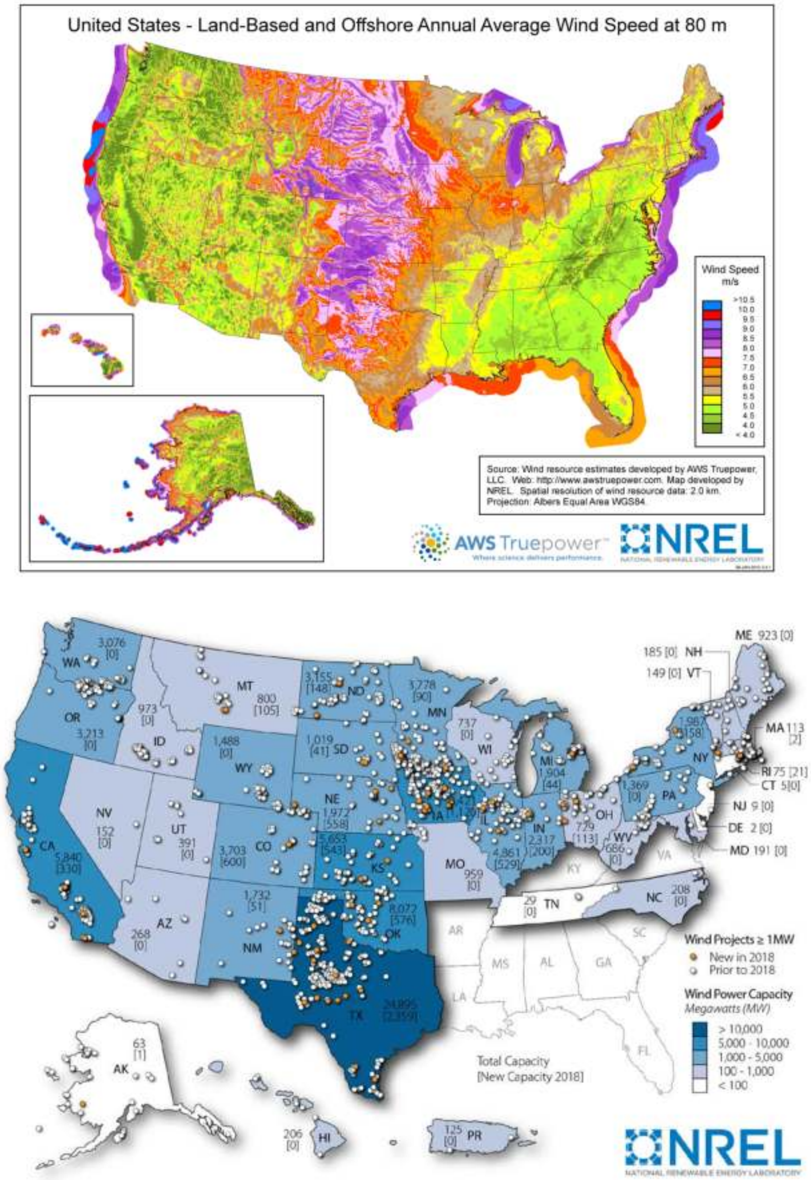
$$\frac{P}{A} = \frac{\epsilon\pi D^2 v^3 \rho_{\mathrm{air}}/8}{25D^2} = \frac{\epsilon\pi\rho_{\mathrm{air}} v^3}{200} \tag{10.6}$$

$$= \underline{\underline{2.6\,\mathrm{W\,m^{-2}}\left(\frac{\epsilon}{0.4}\right)\left(\frac{v}{7\,\mathrm{m\,s}}\right)^3}}. \tag{10.7}$$

In other words a typical turbine farm (40% efficiency) on a good site (steady winds of 7 m/s) generates about 2.6 W of electricity per square meter of land available.

Now we can ask: can wind power make a significant contribution to the US electricity or total-energy budget? The top of Figure 10.7

Figure 10.7: Top: Map of average wind speed at 80 meter altitude, both land and offshore US. Bottom: location of installed wind capacity. From the National Renewable Energy Laboratory (NREL).



Note: Numbers within states represent MegaWatts of cumulative installed wind capacity and, in brackets, annual additions in 2018.

is a map of average wind speed at 80 m altitude across the United States, including the offshore regions. We see substantial stretches of purple ($v > 7.5\,\mathrm{m\ s^{-1}}$) down the midwestern spine of the country, plus very favorable wind conditions offshore and over the Great Lakes. Nearly all of North and South Dakota, for example, have average wind speeds above 7 m/s. The total area of the Dakotas is about $4 \times 10^{11}\ \mathrm{m}^2$, so if we were to cover them with wind turbines we could generate at least

$$P = 2.6\,\mathrm{W\ m^{-2}} \times 4 \times 10^{11}\,\mathrm{m}^2 \approx \underline{\underline{10^{12}\,\mathrm{W} = 1\,\mathrm{TW!!}}} \tag{10.8}$$

This is in fact an *underestimate* of the average power available from Dakotas wind, due to a subtlety worth explaining: because the power formula (10.5) has v^3 in it, the average *power* is more than the power at the average *wind speed*. Consider a simple example of a site that has $v = 2\,\mathrm{m\ s^{-1}}$ half the time and no wind at all the other half. The average wind speed is $v_{\mathrm{avg}} = 1\,\mathrm{m\ s^{-1}}$. But the average power is

$$P_{\mathrm{avg}} = \frac{\pi}{8}\epsilon\rho_{\mathrm{air}}D^2 \times \frac{\left(2\,\mathrm{m\ s^{-1}}\right)^3 + \left(0\,\mathrm{m\ s^{-1}}\right)^3}{2} \tag{10.9}$$

$$= \mathbf{4} \times \frac{\pi}{8}\epsilon\rho_{\mathrm{air}}D^2 \times v_{\mathrm{avg}}^3. \tag{10.10}$$

This is an extreme case. For a site with typical distribution of wind speeds, we have

$$P_{\mathrm{avg}} \approx \mathbf{2} \times \frac{\pi}{8}\epsilon\rho_{\mathrm{air}}D^2 v_{\mathrm{avg}}^3. \tag{10.11}$$

This leading factor of 2 would be closer to 1 in a site with very steady winds.

Getting back to our Dakota calcuation, we see that they can provide an average output power of $P_{\mathrm{avg}} \approx 2\,\mathrm{TW}$—*which is 4 times the average electricity consumption of the entire US!* So the wind resource in just two states[1] is more than sufficient to meet all US demand, and the nationwide wind power resource could provide all the primary energy needed by the US.

We will examine the cost-effectiveness of wind power in a moment, but first let's ask whether there are technical or environmental barriers. Wind power is zero-carbon and uses zero water, of course, once the turbines are installed. There are aesthetic complaints, generally however from people who have never lived near a mountaintop-removal coal mine. Wind installations may have issues with low-frequency noise and with bird/bat deaths. Although the quantitative extent of these problems remains uncertain, they are clearly less severe than the environmental consequences associated with fossil-fuel power. Land being used for wind farms is still useful for farming and industry since the footprints of the towers themselves are small.

[1] Two states whose total population is less than live within Philadelphia's city limits!

A major issue, however, with large-scale wind power is the unpredictability of the power output. As noted earlier, the utility demand curve must be met. It is estimated that supplying more than 20% of electricity with wind power would begin to require serious attention to energy storage and/or demand management. It seems, however, that there are no serious technical barriers to reaching this point. In 2019, six countries (all in Europe) generated $> 20\%$ of their electricity from wind, as have 6 US states.

The financial impact of wind's non-dispatchability is illustrated in Figure 10.8, which shows that electricity from wind turbines sells on the wholesale market for just 60–80% of the average price for all sources of electricity. This "value factor" also is lowest in the regions of the US with the highest fraction of power coming from wind.

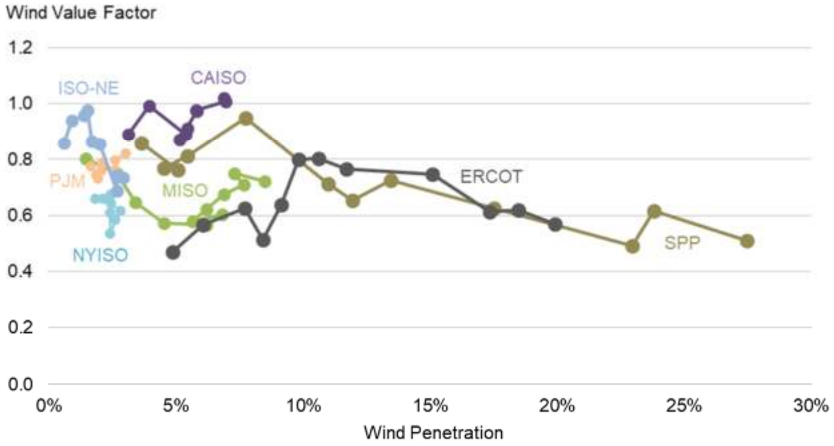


Figure 10.8: The average wholesale price paid for wind power is shown as a fraction of the average price for all electricity. This is shown for different regions of the US—"ERCOT" is Texas and "SPP" is the windy central plains states. This "value factor" drops in the regions with the highest fraction of their electricity coming from wind (a.k.a. "Wind Pentration"). From LBNL WTMR.

Another point apparent from the map is that the wind resource is unequally distributed around the country. The Southeast US is a very poor area for wind power, and the lower map of the locations of wind installations shows that virtually none have been built in this low-yield region. Building out wind power would require increasing the capacity of the US electricity grid to carry power across the country.

The same can be said globally: there is a large total resource, enough to supply all of current electricity demand, though not necessarily at the time and place where it is needed. Some countries are much better endowed with wind than others.

Economics of wind turbines

Since wind power is clearly technically feasible for supplying at least 20% of electric demand, we can ask whether there are economic circumstances that will lead to this happening. The cost of wind electricity is completely driven by the up-front capital costs for construction. We can ask: what capital costs are necessary to make wind power (or any renewable) competitive with the CoE for other sources? We can use Equation (9.8); let's assume and interest rate of 5% per year and a 25-year lifetime for the turbine and the investment:

$$C = f \times CoE \times \frac{1 - e^{-rT}}{r} = \frac{\$2000}{\text{kW}} \left(\frac{f}{0.40} \right) \left(\frac{CoE}{\$0.04\,\text{kWh}^{-1}} \right). \quad (10.12)$$

So, roughly speaking, *any renewable electricity source needs to cost $2000 per kW or less to be grid-competitive.* We saw in Table 9.1 that $0.04/kWh is roughly the CoE for NGCC power used in peaking mode, *i.e.* $f = 35\%$. Since wind power is not dispatchable, it can't compete directly with NGCC at all times and sells at a discount, as shown in Figure 10.8. Prices vary with time and with region, and we are making rough guesses at the financing terms, so we should just consider this a ballpark target rather than an exact value for when wind power becomes economical for utility use.

Figures 10.9 shows the actual costs of commerical-scale US wind installations over the past 15 years. The top panel shows prices for just the wind turbines, and the bottom panel is the total installed cost. In 2019 these averaged about $750 and $1440 per kW, respectively, continuing a gradual long-term decline since about 2008. It's interesting to note that the turbines themselves are only half of the total expenses—installation, site acquisition, and connecting to the power grid are just as much. The technology is mature, in that prices are not dropping dramatically from process improvements, but gradual improvements continue. These prices are well into the range where we estimated that renewable energy becomes economically interesting compared to fossil fuels!

As a consequence there has been substantial growth in US wind power, particularly in the central US region where the resource is strong so net costs per kWh are lower. Figure 10.10 shows that there has been substantial growth in US wind capacity, to 122 GW at the end of 2020, about 10% of total US installed electrical capacity of ≈ 1.2 TW. The EIA reports that wind power generated about 8.5% of all of the 14.4 EJ of electricity in the US in 2020, more than produced by hydropower.

While the overall growth in US wind capacity has averaged ≈ 7 GW/yr since 2007, This capacity growth has been very inconsistent, *e.g.* almost none in 2013, and a record high of 16 GW in 2020. What explains this? Many wind projects have been made attractive to investors by favorable tax treatment or regulations. Many states have ***renewable portfolio standards*** that mandate a certain fraction of the

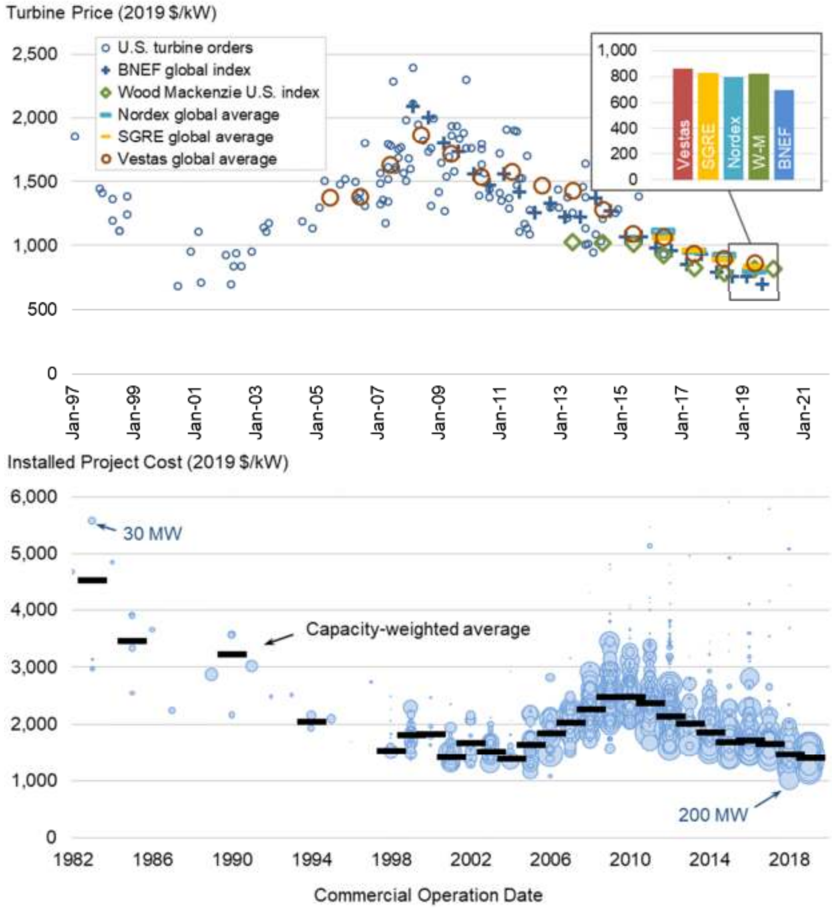


Figure 10.9: Cost per kW capacity of US wind facilities vs year of installation. From the LBNL *WTMR*.

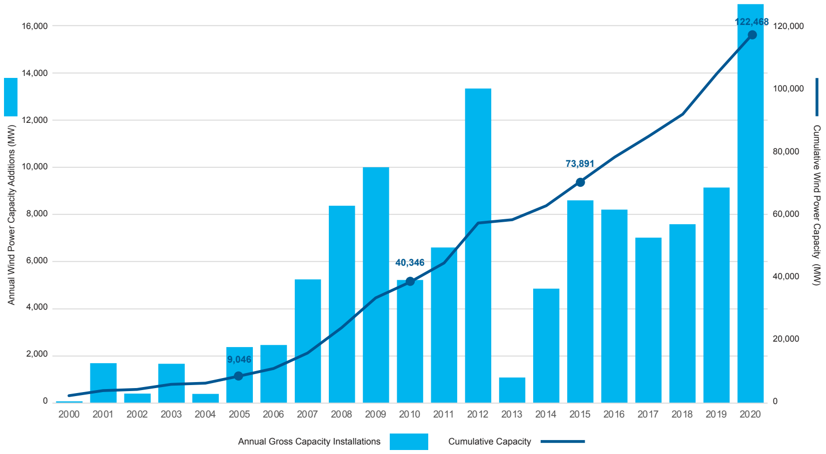


Figure 10.10: The annual additions to and cumulative wind generating capacity in the US over the past 20 years. From *Americal Clean Power Market Report 4Q2020*.

generating capacity be from renewable or low-carbon sources. Most important to US wind development has been the ***Production Tax Credit (PTC):*** the US government has, in most years, paid owners of wind farms \$0.026 per kWh of electricity sold over the first 10 years of operation (\$0.023 more recently). Clearly this drops the net CoE for wind power by roughly two cents per kWh and has allowed wind owners to compete on the market, incentivizing wind production.

But in many years, the US Congress has been unable or unwilling to extend the PTC on a timely basis so that investors and producers can plan for profitable projects. On several occasions the PTC has expired or come within months of expiration before being renewed. In 2013 it was not renewed at all. The precipitous drop in new wind capacity in 2013 is a result of this (in)action. In December 2014, the PTC was extended for projects beginning construction by the end of 2014—two weeks later! These gyrations in the wind turbine market have been very difficult for US wind turbine manufacturers to weather, needless to say. The budget bill passed in Dec 2015 gave the first long-term predictability to the PTC situation in many years, renewing the \$0.024/kWh PTC but setting a gradual phase-out thereafter. The wind PTC for construction in 2021 will be at \$0.018/kWh.

If we use the capital costs of \$1440/kW in our CoE formula with an interest rate of 8%/yr and 30-yr financing, we obtain \$0.036/kWh. This is very close to the costs reported in the top panel of Figure 10.11. The lower panel gives the actual prices at which wind-generated electricity have been sold to utilities by producers. In the windy interior states, wind power is now, incredibly, below \$0.02 per kWh—the difference between this and the COE corresponding to the PTC. In the interior regions, wind is competitive with gas even without the PTC.

Wind power is now a viable source of large-scale renewable energy, but even if the 2020 growth rate of 16 GW of capacity per year is sustained, it will be another decade before wind produces 20% of US electrical needs, and having it produce a majority of *total* US energy needs would take many decades. A much faster rate of installation will need to be attained if we wish to approach carbon neutrality within 20–30 years.

Global wind power usage

Global statistics are from LBNL *WTMR* and the *REN21* report.

Other countries have different wind resources and incentives. Globally, 60 GW of capacity were added in 2019, nearly half in China, to bring total installed wind capacity is 650 GW. If the global capacity factor is 40%, these devices will produce 8.1 EJ of electricity in 2020, out of $\approx$ 97 EJ of electricity produced worldwide—about 8%. China is by a substantial margin the leader in total wind capacity,

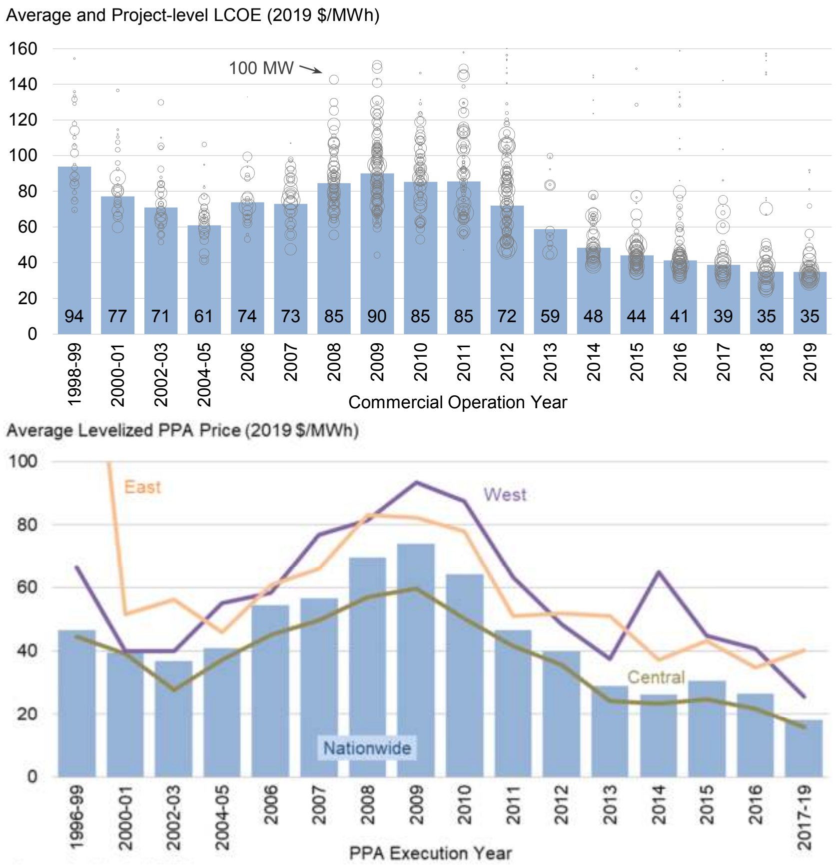


Figure 10.11: The cost of producing wind electricity (levelized cost of electricity, "LCOE") and the prices at which wind electricity is being sold (PPA is "producer price agreements") in different regions of the US over the past 20 years. From the LBNL *WTMR*.

with China and the US together providing more than half of the new installations and cumulative capacity. But there are countries whose *fraction* of total electricity provided by wind are much higher than in the US or China—led by Denmark, obtaining over 40% of electricity from wind power (see Figure 10.12). The rate of wind capacity addition in the world has been steady for the past few years (about 50 GW/yr worldwide), i.e. the growth is linear, not exponential.

In summary, wind power is mature and feasible. The potential resource is large, easily sufficient to reach the 20% of electricity supply that is expected to be manageable without demand-curve issues. The cost of wind power is similar to (currently cheap) natural gas power in the US, and in favorable regions wind power is routinely producing the cheapest new electricity—although of course it is not a dispatchable resource, so other sources are needed. Wind power capacity and production are increasing by about 0.5–1.0% of total electricity per year, a trend that seems likely to continue. But the rate of wind installations will need to accelerate considerably if our goal is to largely displace fossil fuel use within a few decades.

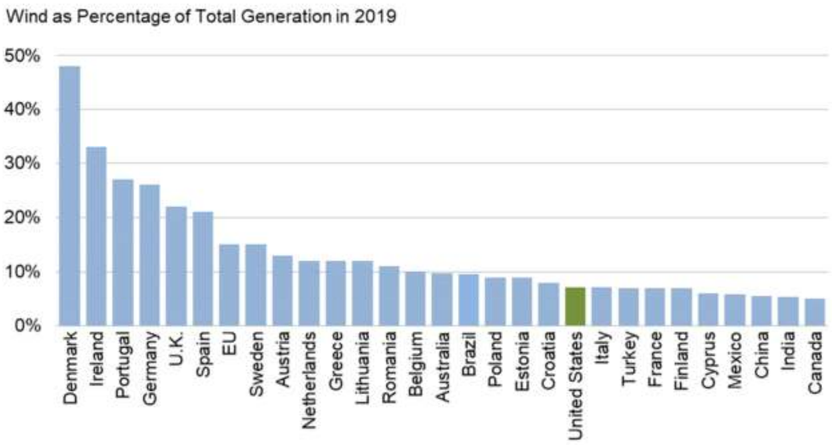


Figure 10.12: Fraction of electricity consumption that is generated by wind in leading countries. From LBNL WTMR.

Practice Problems

- If we wanted to supply all of current electrical demand with wind, plus power most of the transportation fleet with 90%-efficient electric vehicles instead of 25%-efficient ICE's, what average wind power output would we need? What wind *capacity* would we need? How much would it cost to build this? What would the cost per year per person be if we were to issue 25-year bonds at 4% interest rate, then spread the repayment cost among the whole US population?
- Would you vote in favor of sustaining PTC for wind power in the US?

11

Solar Power Basics

Sunlight is the mother of renewable energies as well as the grandmother of all fossil fuels. Wind power, for example, taps the small fraction of sunlight that is converted to kinetic energy of air. It would seem we could capture more power by tapping the photons directly. There are four currently feasible ways to tap solar power on a large scale:

1. Solar light and heat directly provide indoor lighting, space heating, and/or water heating in buildings.

2. ***Concentrated solar power (CSP)*** is converted to heat, and used to run a heat engine and generate electricity.

3. ***Photovoltaic*** devices convert photons directly to electricity, without heat engines.

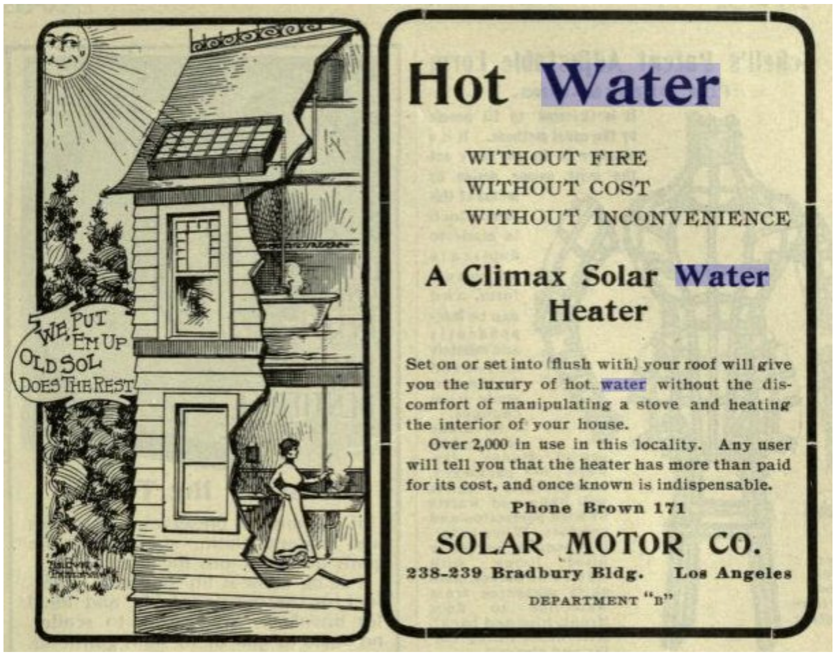
4. Photosynthesis in plants (or potentially in synthetic systems) store the solar energy in ***biofuels.***

Use of solar energy for light and heat is not a new idea (Figure 11.1), but we are capturing it on a much more massive scale than any time before—can we expand much farther? This chapter will describe when, where, and how much solar energy is available on Earth to power any of these four techniques, as well as simple methods for direct use of solar energy (method 1). Methods 2–4 will be covered in succeeding chapters.

Flux and insolation

The Sun has a very stable power output[1]. Recall that the power that passes through a flat window of area A pointed directly at the Sun is the ***solar flux*** of $f_{\odot} = P/A = 1360\,\mathrm{W\,m^{-2}}$ if we place the window

[1] The rise in global mean temperature over the past 150 years is definitely *not* attributable to increased solar output!



Hot Water

WITHOUT FIRE
WITHOUT COST
WITHOUT INCONVENIENCE

A Climax Solar Water Heater

Set on or set into (flush with) your roof will give you the luxury of hot water without the discomfort of manipulating a stove and heating the interior of your house.

Over 2,000 in use in this locality. Any user will tell you that the heater has more than paid for its cost, and once known is indispensable.

Phone Brown 171

SOLAR MOTOR CO.
238-239 Bradbury Bldg. Los Angeles
DEPARTMENT "B"

Figure 11.1: A 1902 ad for rooftop solar water heaters. From "Out west (1902) - Solar water heater advert" by Solar Motor Co., via Wikimedia Commons.

above the atmosphere. Several factors, however, cause a solar collector at the surface to recieve substantially less power than $f_{\odot} A$.

Night: Since sunlight cannot pass through the Earth, sunlight is available to only half the planet's surface at any given moment. In a particular site, the Sun is visible on average for only 12 hours per day. But in winter it's less than 12 hours, more in summer (see below for why).

Clear-sky absorption: Molecules in Earth's atmosphere absorb (or reflect back into space) about 1/4 of the solar energy before it reaches the surface. Therefore, *the peak flux at Earth's surface, which occurs in clear skies, is* $1000\,W\,m^{-2}$. Whenever you purchase a solar-power device, its *capacity,* or rated maximum power output, will be what it produces at this flux, *i.e.* when pointed directly at the Sun in these clear-sky conditions. The *capacity factor* of a solar device is necessarily below 50%—even in a site that never has clouds, we can obtain this maximum power less than half the hours of an average day.

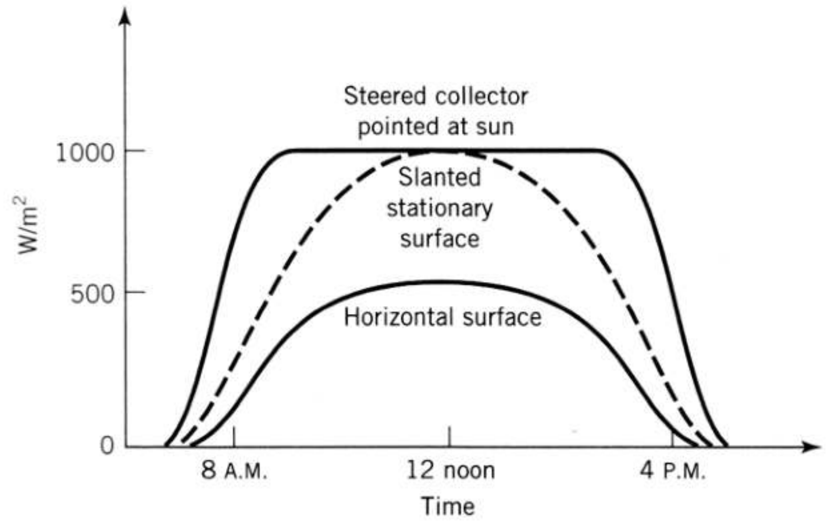


Figure 11.2: The solar power per unit area incident on a solar panel is plotted vs time of day for a clear winter day at temperate latitude. The three curves are for different mounting schemes for the panel. From Figure 4.4 of Ristenen.

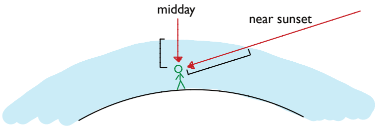


Figure 11.3: When the sun is far from zenith, such as at sunset, its rays have to pass through more air (blue) before reaching us than they do when the sun is high in the sky.

The uppermost curve in Figure 11.2 shows the power that is received (per unit area) by a panel that is kept pointing at the Sun on a clear day. As expected, the power is near $1000\,\mathrm{W\,m^{-2}}$ during the daytime and zero at night. But we see it drops gradually at the start and end of the day. In the mid-day, the sunlight takes the shortest path through the air to reach the panel. As illustrated in Figure 11.3, near sunrise or sunset, the Sun is low in the sky and sunlight passes through much more air before reaching the ground. The air absorbs more of the Sun's energy before it reaches the panel. This is also why you get sunburned more quickly at midday than early morning or evening.

Orientation: if a flux f is reaching the ground, our panel with area A obtains power $P = fA$ from the light if we point the panel directly at the source. But if the source (*e.g.* the Sun) is some angle θ away from the direction the panel is pointed, the panel cannot intercept as much of the light. As illustrated in Figure 11.4, a panel tilted by θ to point at the Sun could intercept all of the same sunlight as ours, but would need a smaller area $A' = A\cos\theta$ to do so. So we see that *a solar collector illuminated from an angle θ by a source with flux f receives a power*

$$P = fA\cos\theta. \tag{11.1}$$

This means, for example, that if we consider an area A of horizontal ground (or a panel laid flat on the ground), it can only collect $1000\,\mathrm{W\,m^{-2}}$ of power when the Sun is located directly overhead. Below we will introduce a little astronomical knowledge to show that the Sun is *never* directly overhead in Philadelphia, nor for that matter anywhere in the continental US. Looking at the lower curve in Figure 11.2 for a horizontal collector, we see that it falls short of 1000 W m^{-2} even at midday when the Sun is at its highest point in the sky. As we get closer to sunrise or sunset, the Sun is lower in the sky, the angle θ gets larger, and $\cos\theta$ decreases, suppressing the sunlight. This is why it's hotter at midday—there is more solar power per square meter of (horizontal!) land area than at the start or end of the day. Using a little astronomy (see below) we can calculate what the angle of the Sun will be at any time in any location on Earth.

Clouds: Finally, the sky is not always clear, and the presence of clouds will of course reduce the light reaching the ground by either absorbing photons or reflecting them back into space. Clouds are also very good ***scatterers*** of light, meaning that they redirect incoming photons. An observer on the ground can see either ***direct*** illumination, coming from the Sun's direction, or ***indirect*** illumination coming from other parts of the sky. If there were no indirect illumination, the sky would look pitch black away from the Sun. The normally blue clear sky means that blue light is scattered by molecules and particles in the air, but clouds greatly increase scattering and produce more indirect light. On a fully cloudy day, the direct sunlight is gone, *i.e.* you cannot see that Sun.

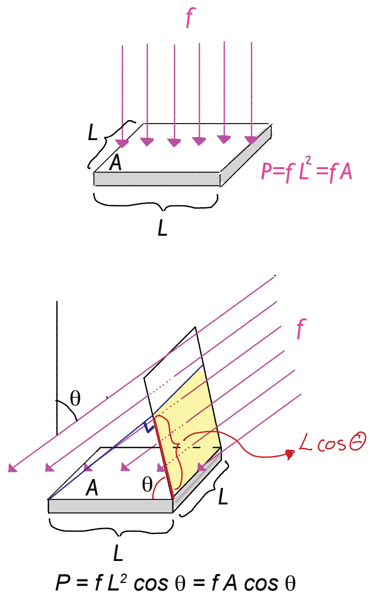
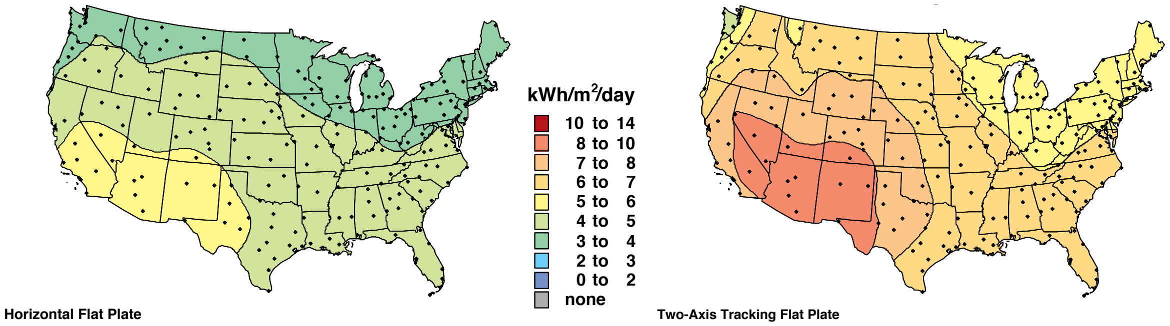


Figure 11.4: Top: a panel of area A pointed directly at the source of light with flux f recieves power $P = fA$. But at bottom, the the panel is pointed some angle θ away from the light source. It only intercepts the light passing through the yellow region, which carries power $P = fA\cos\theta$.

Insolation



Horizontal Flat Plate

Two-Axis Tracking Flat Plate

Figure 11.5: Insolation maps of the continental US, courtesy of the National Renewable Energy Lab.

At any given site, astronomy can tell us when the Sun is up and what orientation it will have will respect to our panel. Weather records tell us what the average impact of clouds will be. From these one can calculate the ***insolation,*** which is the *average solar flux* received at a surface. Figure 11.5 plots the annual average insolation on the continental US. The insolation depends on the orientation of the surface receiving the sunlight. At left is insolation for a horizontal flat plate. This would be the appropriate map to use to calculate the total sunlight incident on the land, since the land area is horizontal. At right is the insolation for a panel that tracks the Sun across the sky, which gets the *most* possible sunlight per square meter of panel, corresponding to the upper curve in Figure 11.2. The insolation maps tell us the average power we can expect to get from a solar collector. If the collector has area A and an efficiency ϵ in converting the incident sunlight into the desired form of output energy, then its *average* power output will be

$$P_{\text{avg}} = \epsilon A \times (\text{insolation}). \tag{11.2}$$

When you purchase your collector, it will be sold on the basis of its capacity, or *peak* power output occurring when pointed directly at the Sun on a clear day:

$$P_{\text{max}} = \epsilon A \times 1000\,\text{W m}^{-2} = \epsilon A \times 1\,\text{kW m}^{-2}. \tag{11.3}$$

The capacity factor f for your solar installation will be the ratio of these. For instance, the insolation map gives a value of $4\,\text{kWh m}^{-2}\,\text{day}^{-1}$ for a horizontal flat plate in Philadelphia. For such an installation we'd expect a capacity factor

$$f = \frac{P_{\text{avg}}}{P_{\text{max}}} = \frac{4\,\text{kWh m}^{-2}\,\text{day}^{-1}}{1\,\text{kW m}^{-2}} \times \frac{1\,\text{day}}{24\,\text{hr}} = \frac{4}{24} = 17\%. \tag{11.4}$$

We could obtain a higher capacity fraction—more electricity from the same panel—by motorizing its mount to track the Sun, yielding a higher insolation of $\approx 6\,\text{kWh m}^{-2}\,\text{day}^{-2}$. But typically it would be cheaper to buy 1.5 times as many flat plates than to buy the hardware and motors to track the Sun to produce the same power.

It is hardly necessary to state that we would get more energy out of solar-collector investment if we installed it in the Southwest. The insolation maps tell us exactly how much more. The higher capacity factor means that solar power is more competitive with other energy sources.

The Sun's path on the sky

Some astronomical knowledge will guide us toward the most cost-effective way to use the Sun's light. The Sun is very far away and its rays are essentially parallel when they get to Earth. At any given moment there is only one point on the spherical Earth where the Sun is located directly overhead at the ***zenith*** of the sky. Recall that a flat horizontal plate gets the full $1\,\text{kW m}^{-2}$ of clear-sky flux only when the Sun is at zenith. It can be helpful to remember that the zenith (straight up) is the opposite of the direction to the center of Earth (straight down). On the ***equinoxes*** that occur near March 21 and September 21 each year, the Sun is located above the equator of Earth, so a person sitting on Earth's equator will see the Sun at

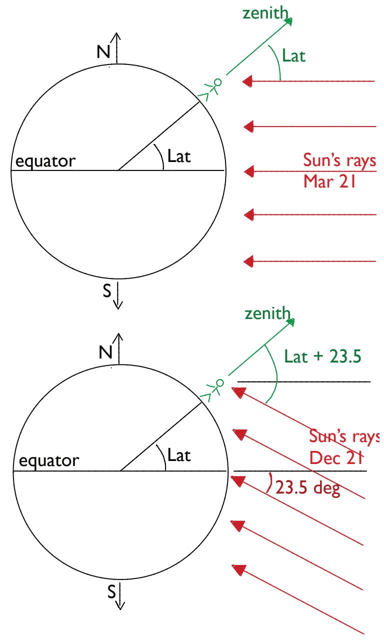


Figure 11.6: Top: On the equinox, the sun's noontime highest point in the sky is given by your latitude. Bottom: On the (northern) winter solstice, the sun's rays approach from 23° south of the equator, and the sun is farther from the zenith at noon for Northern Hemisphere residents.

zenith. Since the Earth is rotating, the Sun reaches this point only at midday. The Sun traces an arc across the sky from due East, 90° from overhead, at 6 AM, to directly overhead at noon, to the due West at 6 PM.

Someone who is not on the equator will have the Sun reach its highest point at noon, but the Sun will *not* reach the zenith. From Figure 11.6 we can see that at at midday, the angle θ between the Sun and the zenith is equal to the ***latitude*** of the observer. Philadelphia has latitude 40° North of the equator. On the equinox, we'll see the Sun rise due E and follow a tilted arc across the sky that misses the zenith by 40°. If we take a solar collector and tilt it 40° from horizontal toward the south, it *will* be pointed directly at the Sun at noon. This is the best position to mount our panel if we do not want to buy motors to move it during the day. The middle curve in Figure 11.2 shows the flux on this panel during a clear winter day. At midday we just match the output of a fully-tracked panel, and we're always doing better than a horizontal panel.

In the Northern hemisphere, the most bang for the buck in residential or commercial solar collectors is typically to permanently mount them tilted South by an angle equal to the latitude. Figure 11.7 is the NREL insolation map for panels mounted this way.

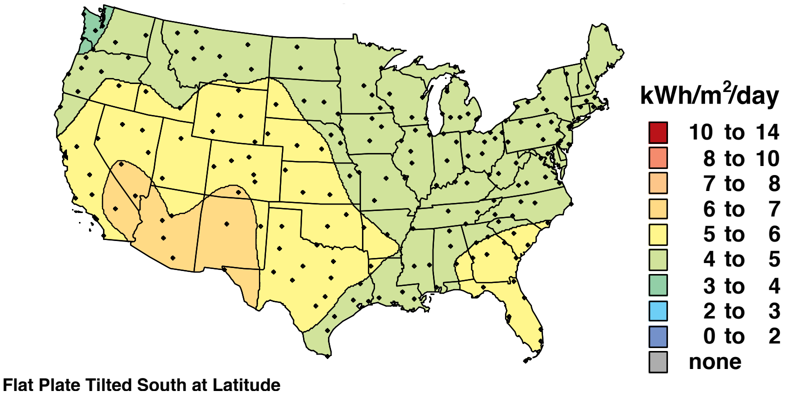


Figure 11.7: Insolation for flat plates tilted south by latitude, again from NREL.

The Sun is not always over Earth's equator, because Earth's spin axis is tilted with respect to its orbit around the Sun. In the northern summer, Earth's northern pole is tipped toward the Sun, reaching 23.5° at the solstice near June 21. Now the Sun is above a point 23.5° north of the equator. Those living at this latitude (the Tropic of Cancer) will see the Sun pass through zenith on the solstice. In Philadelphia, the Sun will get within $40^\circ - 23.5^\circ = 16.5^\circ$ of the zenith. This higher position means more solar power hitting each square meter of ground, making this the warm season. At the same time in the southern hemisphere, the Sun is *farther* from the zenith, and it is winter.

Three months later is the equinox when the Sun is over the equator. It continues drifting south until Dec 21, another solstice, when the Sun is 23.5° *south* of the equator, and those on the Tropic of Capricorn at latitude $-23.5°$ see the Sun pass overhead. Up in Philly, the Sun gets no higher than 63.5° from the zenith. Another way to see how this tilt causes seasons is to look at things from the Sun's point of view: on June 21, the majority of the Earth that the Sun can see is in the northern hemisphere, so:

1. More of the Sun's energy is deposited in the Northern hemisphere, where it's summer as a result.
2. The North (South) Pole is always (never) visible, and sees daylight (nighttime) 24 hours per day. More generally the daylight hours are $>$12 hours in the Northern hemisphere, $<$ 12 in the South.
3. The Sun rises and sets north of due east and west.
4. The Sun reaches higher (lower) in the sky in the northern (southern) hemisphere.

If we now view this from the point of view of someone in the temperate latitudes of the North, as in Figure 11.8, we see that the Sun's arc across the sky is nearly always in the southern half of the sky (except for summer mornings/evenings). The Sun's arc is somewhat higher in the sky in summer, and lower in winter, than the latitude of the site. So tilting your solar collector at the latitude is a pretty good compromise for most effective year-round collection of sunlight. The solar collector will of course produce more power, for more of the day, in summer than in winter, if the seasons are similarly cloudy.

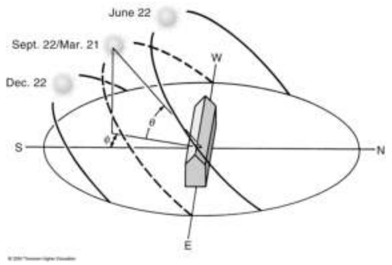


Figure 11.8: Path of the Sun across the sky for a mid-northern-latitude site. From Hinrichs (2006).

Direct use of solar light and heat

In a residential or commercial building, one can save significant energy by designing the structure to make use of solar light and heat, reducing the load on artificial lighting, and space heating and cooling. Figure 11.9 shows some techniques: first, in a Northern home, windows placed on the southern exposure of the building will admit copious sunlight into the home, reducing lighting and heating bills. In summer we might not want the sun's energy entering at midday, since this will increase our air conditioning needs. Placing an overhang above the windows is useful, since these will block the Sun on summer middays when the Sun is high in the sky, but admit winter/evening light. Another good passive strategy is to plant deciduous trees on the southern side of the house.

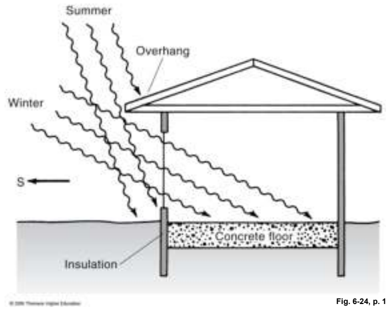


Figure 11.9: Some elements of energy-efficient passive solar design. From Hinrichs (2006).

Modern windows have good "solar gain," meaning that they admit most of the Sun's energy without letting the interior heat escape.

We have seen in Chapter 3 that double-paned windows can reduce heat transfer by conduction and convection. To get solar gain on *radiative* heat transfer, we want to coat the window with a material that is transparent to *visible* light that carries most of the Sun's energy (and that makes the window nice for humans to look out of!), but is opaque to the *infrared* light that is generated by the room-temperature environment inside the house.

We can take a more active approach to harvesting the Sun's heat for our home by putting solar hot-water heaters on the roof. A good solar collector can be $>$ 70% efficient at transferring the incident sunlight into thermal energy of water flowing through pipes on the roof. Elements of good design are illustrated in Figure 11.10: the collectors are painted black to absorb all photons; they are covered in glass to prevent the heat from convecting away; and the glass covers should transmit visible light while blocking the infrared light, creating a miniature greenhouse effect inside.

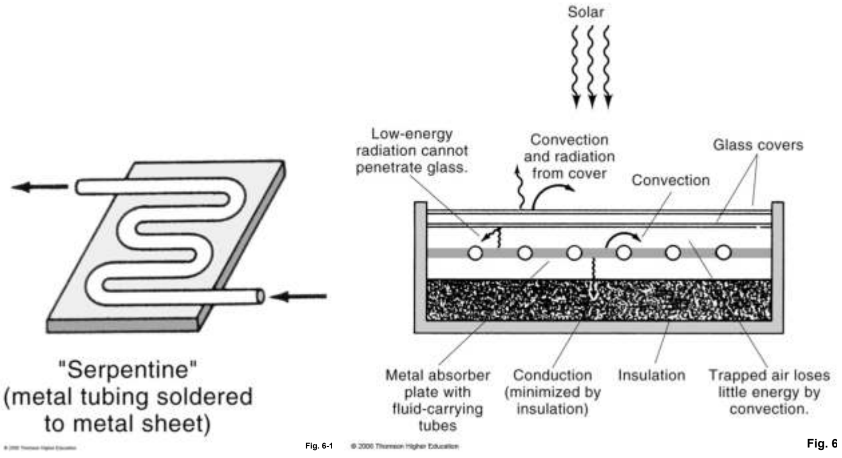


Figure 11.10: Design elements of a rooftop solar hot-water heater. From Hinrichs (2006).

Once we have collected the solar energy in our water pipes, we can use this directly as our hot-water supply (heating it further with a traditional natural-gas or electric water heater, if the Sun has not finished the job). We can also use a heat exchanger to transfer the heat to our hot-water lines (a "two-loop" system), or we can transfer the heat to the interior air for use in space heating. A nice application is to use solar for heating a swimming pool, since swimming season and solar-power seasons coincide. Rooftop solar water heaters are ubiquitous in some sunny locales, such as Israel and Spain.

Practice problems

- You build a rooftop solar water heater that is 70% efficient, and you install it on your south-facing rooftop in Southern California. What capital costs (dollars per square meter of collector) will make the cost of the solar heater less than the cost of natural gas that you save by using it? (Take typical interest rates and retail gas prices, and a 25-year lifetime for your system). What if you have an electric water heater instead of gas?
- What fraction of the area of Arizona would need to be covered with solar collectors operating at 15% efficiency if we wanted them to produce energy equal to the total US consumption rate?
- A tricky question: if we to collect all the solar energy from 1 square km of land in Pennsylvania, we do not need to buy a full 1 square km of solar panels. Why not?
- Compare the energy density (per unit land area) of a 20%-efficient solar panel to a typical wind farm. Which would you consider to be the more environmentally benign use of land to produce the same amount of energy?

12

Solar Electricity

Converting solar power to electricity makes it easy to distribute and use at high efficiency in many applications. There are two methods that have now been successfully implemented with many GW of global installed capacity: *solar thermal* methods, which tranform sunlight to heat and then into electricity using steam turbines; and *photovoltaics (PV)*, which convert photons directly into electricity. PV is in use on a much larger scale and is the cheaper, more promising technology. Both methods, however, are subject to the same issues mentioned for wind power in Chapter 10, namely that lack of dispatchability means that its use for more than about 20% of the electric supply can become inefficient because of mismatch between supply and the demand curves. Peak solar power arrives at midday, when electric demand typically peaks, but the match is not perfect: the early-evening period of high demand is obviously not easily supplied by solar power, and of course there are clouds to consider. Solar thermal power, while currently more expensive than PV, could prove useful because it may offer a cost-effective way to store the solar energy and offer dispatchability.

In Equation 10.12 we showed that a renewable-electricty plant operating near 30% capacity factor, with low O&M costs, must have capital costs in the neighborhood of $2000 per kW capacity in order to generate electricity near $0.05 per kWh. For whatever reason, solar capital cost is usually given in $/W instead of $/kW, but the result is the same: we should *target capital costs of $2 per W electrical capacity for solar electricity to become competitive for grid power without subsidies.*

Solar thermal electricity

We can simply take the concept of a coal-fired power plant and replace the coal boiler as a heat source with solar power—of course also getting rid of all the emissions-control equipment too! Knowing that a steam turbine (or any other heat engine) will be more efficient the higher temperature we can produce at the input, we will want to *concentrate* the Sun's energy onto a smaller area and use it to heat steam to high temperature and pressure. More commonly we will heat some other high-heat-capacity liquid (mineral oil or molten salt), then transfer the heat from this working fluid to the steam. The steam runs a turbine, which runs a generator.

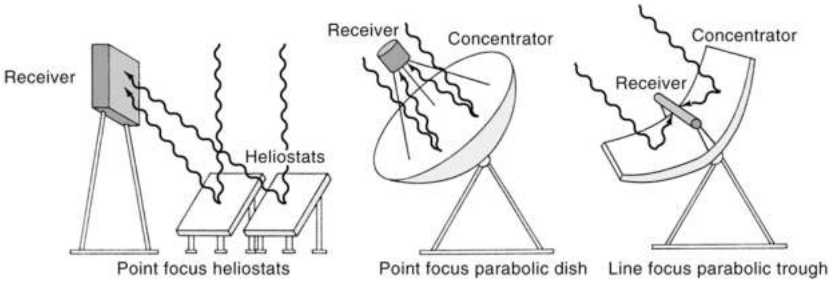


Figure 12.1: Three types of solar concentrators. The left-hand one is also known as "solar power tower." From Hinrichs (2006).

Solar thermal is more commonly called ***concentrated solar power (CSP)*** for this reason. Figure 12.1 shows different ways of concentrating sunlight. The two that have been used for large-scale (>100 MW) installations as of 2019 are the ***solar power tower*** method, whereby a field of flat mirrors ***(heliostats)*** direct sunlight toward a central receiver tower; and ***parabolic trough*** concentrators, which direct sunlight onto a line running down the trough. At this "line focus" we place a pipe filled with circulating oil to absorb and carry the solar energy as heat. Each method has its advantages: troughs are simpler, as they need to rotate only around their long axis to track the sun. You can keep growing a trough solar field continuously by adding more and more troughs. The solar power tower is more complex, as the heliostats need two-axis control to direct the sunlight toward the tower. They need a tall tower in the center for the receiver, so heliostats cannot be put too far away. But because the sunlight is more concentrated, power towers can operate at higher temperatures and efficiencies, closer to 20% than the 15% more typical of parabolic troughs.

CSP installations of both types share several characteristics:

- Obviously only the *direct* sunlight can be successfully reflected to the target. CSP installations therefore are best suited to sites with very few clouds. The majority of CSP capacity to date is in the US Southwest and in Spain, although neither country is currently building new CSP plants. We need insolation maps that do not count indirect sun. Arizona insolation for full tracking concentrators is in excess of 8 kWh $\mathrm{m}^{-2}\,\mathrm{day}^{-1}$. This implies that

Figure 12.2: Top: Heliostat mirrors being delivered to the Ivanpah Solar Electric Generating System in Nevada (Wikimedia Commons). Bottom: aerial view of the full facility nearing its 2014 completion. it has 173,500 heliostats and 392 MW capacity (Brightsource Energy).

CSP plants will be operating at capacity factors up to $\approx 8/24 = 33\%$.

- CSP plants are not backyard operations: the great majority of capacity is in utility-scale operations.
- Steam turbines require cooling water for higher efficiencies. Water is generally scarce in locations with no clouds. This would become an issue for truly large-scale CSP deployment. Air cooling can greatly reduce the amount of water needed, but at some expense in capital costs and efficiency.
- CSP, unlike PV or wind, offers the opportunity for *built-in energy storage.* We can store *heat* much more easily than sunlight or electricity! We can use the sunlight to heat a fluid—typically molten salt—and keep some or all of the hot liquid in an insulated tank for hours or days before using it to generate steam. This greatly improves the practical (and monetary) value of CSP electricity. The plant output can remain steady and predictable even if clouds

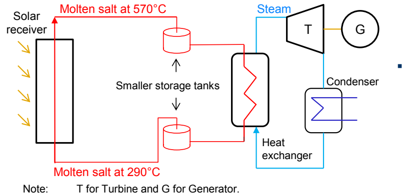


Figure 12.3: Schematic diagram of CSP with molten-salt thermal energy storage. From the SBC Energy Institute *Concentrating Solar Power* factbook (2013).

pass briefly overhead. The electric output can be delayed by a few hours to better match solar energy input to the daily peak in the demand curve. All recent CSP installations include thermal storage, typically for ≈ 10 hours' worth of plant output, for this purpose. Also one can design the steam turbine/generator system for a capacity lower than the peak solar input, storing away the excess heat until later, which leads to lower capital costs and higher capacity factor for the turbine.

Figure 12.4: Left: closeup of a parabolic trough solar collector (source ???). Right: aerial view of the Solana Generating Station in Gila Bend, Arizona, which generates 250 MW of electricity at 36% capacity factor using parabolic trough collectors and molten-salt storage (from Abengoa Solar). The individual troughs at Solana are substantially larger than the one shown at left. Note Interstate 19 at right, for scale.

CSP technology should still be considered to be in developmental phase—there are 6.2 GW of global CSP capacity as of the end of 2019, with 1.7 GW (and the 5 largest installations) in the US and 2.5 GW in Spain. Only 0.6 GW of CSP were commissioned in 2019, these in Israel, China, South Africa, and Kuwait. Capital costs (and *CoE*) for the large US installations completed in ≈ 2014 were several times higher than PV costs. At present, interest in CSP is limited, since PV has proven better in cost and reliability. Substantial CSP price drops could be possible with further development, which could be inspired by the energy-storage options for CSP. In the long term, if the water-suppy issues can be addressed, CSP could become a significant share of electricity supply in sunny regions such as the southwest US, Spain, North Africa, and Chile.

Renewables 2018 Global Status Report

Photovoltaics

Photovoltaic power has been in use for over 60 years, for instance on spacecraft where a power source with no need of fuel, oxygen, or moving parts is extremely valuable. While NASA will pay top dollar for a few kW of solar power, it is a totally different story to be able to deliver many GW at capital costs near our $2/W target. The improvement in PV performance and cost over the past few decades has been remarkable, and PV use is growing rapidly worldwide, though still a small fraction of total electric capacity.

Physics of photovoltaics

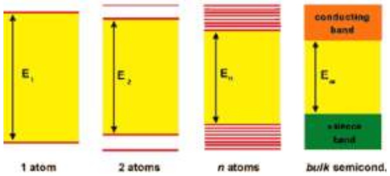


Figure 12.5: Electronic energy bands in a solid. From http://www.nanolytics.de.

To understand how a photovoltaic device can create electricity without a heat engine, we need a simplified picture of how electrons are arranged inside a solid.

Many solids are ***crystals*** in which atoms are arranged into regular arrays. Recall that isolated atoms have "orbitals" in which electrons can exist with a certain amount of energy; only one electron can go into each orbital. The electrons will fill up the lowest available energy levels (the ground state) unless something gives the electrons extra energy. When placed into a crystal, the electrons can be shared among many atoms, and the individual atoms' energy levels turn into *bands* that span a range of energies. The *valence band* is the highest-energy one that is filled with electrons when the solid is at its lowest energy; and the *conduction band* is the next-higher energy band, which will be empty with the solid at its lowest energy.

In an ***insulator*** there is a substantial ***band gap*** between the energies of the valence and conduction bands. The insulator cannot carry electricity at all because the valence band is *full.* Basically all the orbitals are full, and the electrons have no place to go and hence cannot flow—it's like a parking lot that is completely filled with cars.

A photon can give one of the electrons enough energy that it can reach a state in the conduction band. Now the material *can* conduct electricity, because this electron always has empty neighboring states it can move into, carrying its electric charge across the device. In addition, there is a ***hole*** left behind in the valence band. As electrons shift to fill in the hole, the hole can be thought of as moving in the opposite direction, in fact *the hole in the valence band acts just like a positive charge,* a kind of anti-electron! The material will continue to conduct electricity until the electron falls back into the hole, releasing its energy as heat.

In a ***semiconductor*** the energy it takes to cross the band gap is small, typically equal to the energy of a single visible or infrared photon. A semiconductor can be turned into a photovoltaic device as follows: a photon is absorbed to create an electron and a hole that can carry electricity. We arrange for the electron and hole to be pulled to opposite ends of the semiconductor, so that the electron will travel through an electrical circuit before combining with the hole. The electron's energy is then used in the circuit instead of turning into heat!

Photovoltaics are less than 100% efficient for several reasons:

- Some of the photons reflect off the surface of the cell. A good PV cell will look very black, not shiny.

- Some photons don't have enough energy to push an electron

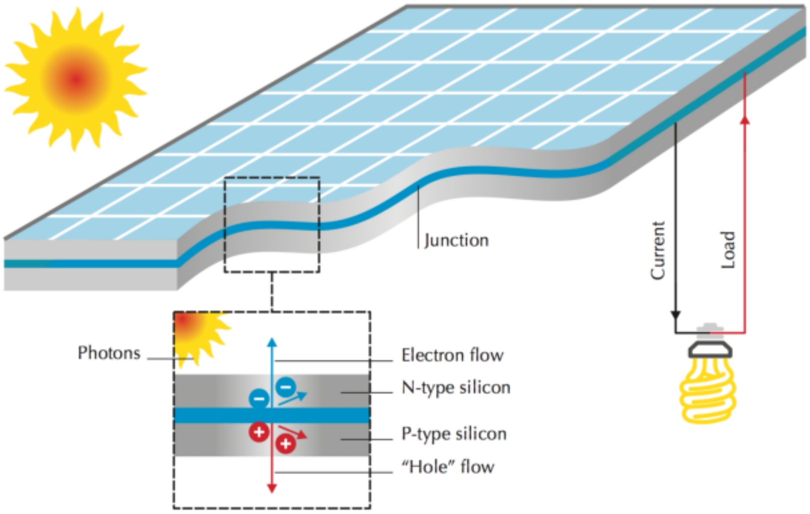


Figure 12.6: Diagram of a photovoltaic cell. From the IEA *Solar Energy Perspectives* (2011).

across the gap. These pass through the device.

- If a photon has *more* energy than the gap, this extra energy is usually wasted.
- If an electron combines with a hole before it leaves the device into the circuit, its energy becomes heat instead of electricity.

There are many different semiconductors that have been made into PV devices. The most common is ***crystalline silicon (c-Si),*** which is the material that virtually all electronic devices are made from. We are now exquisitely talented at manipulating electrons in silicon. The challenge with PV is to take materials that are usually used to make fingernail-sized integrated circuits, and instead make millions of square meters of PV cells—and do it cheaply!

Figure 12.7 shows the world-record efficiencies for various types of PVs advancing year-by-year. Our friend c-Si has a theoretical maximum efficiency of 29%; the filled blue squares show that efficiencies up to 26% have been obtained in the lab. Commercial c-Si cells have efficiencies more like 15–20%, with the average for PV installed in 2019 being about 19%.. Remember that the goal is not to have the highest efficiency, or kW per square meter: the goal is to obtain the most kW per *dollar* spent on the cells. It's not worth spending a lot of money to get slight efficiency gains. On the efficiency chart there are laboratory PV cells at 47%, but these are not cost-competitive. While c-Si cells are most common now, the "thin-film" cells have a substantial market share, and it is possible that other technologies will be cheaper in the future.

LBNL *Tracking the Sun* 2019

PVs are a very attractive energy source:

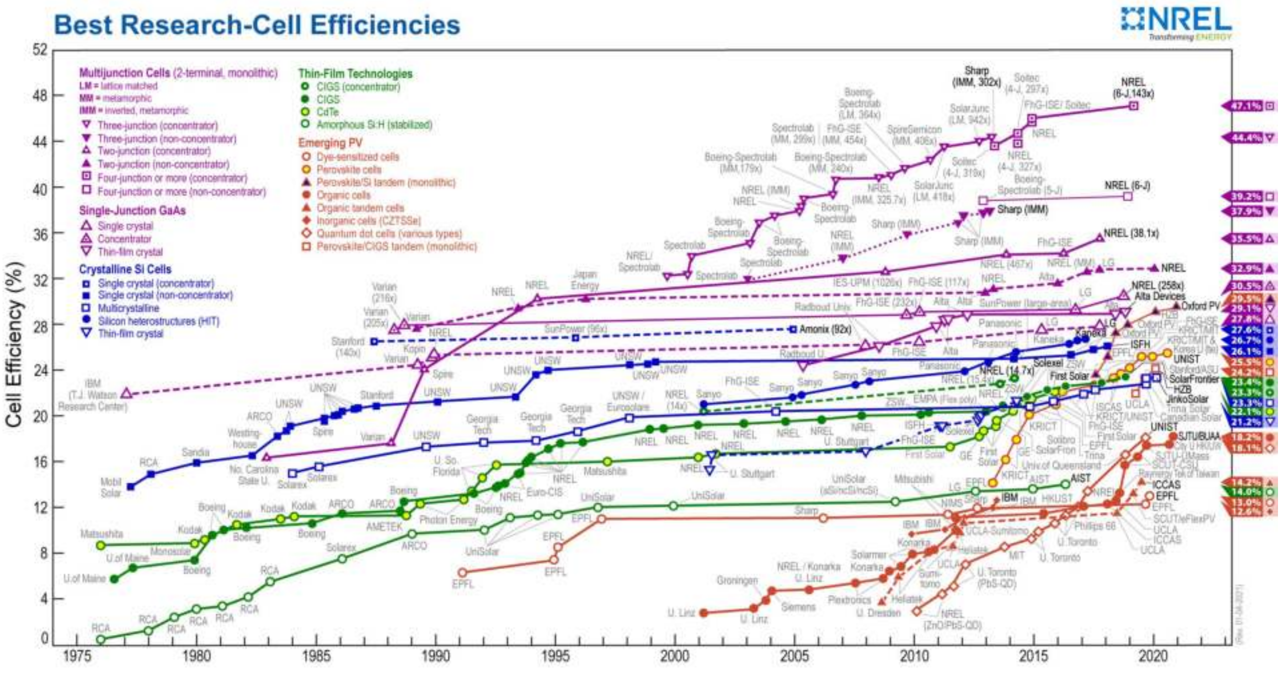


Figure 12.7: NREL's scoreboard of world-record efficiencies for different types of PVs. Commercial devices obtain lower efficiency than research cells.

- They of course have no CO_2 emission in operation and no fuel costs. The manufacturing process releases CO_2, but far less than any fossil fuel would emit to make the energy a PV panel produces over its lifetime.

- They have no moving parts (unless tracking the Sun) and very low O&M costs. You need to wash the dust off as it blocks some of the light.

- They do not require cooling water.

- They produce zero-entropy electricity, which is used at high efficiency for many purposes and is readily transmitted long distances.

- Once installed they are long-lived. Sunlight gradually degrades the power output of most PVs, but they should have 25-year useful lifetimes or longer.

- They can be installed in small batches. One does not need to commit a large amount of capital to generate PV electricity. In fact until ≈2012 most PV capacity in the US was installed in non-utility settings (residences and commercial buildings). The ability to build out in small increments with very little lead time gives businesses desirable flexibility.

- PVs can be installed on rooftops and carports, so there are billions of square meters of appropriate mounting surfaces available "for free."

- The most common kind of PV panel does not require substantial quantities of any rare materials. The primary component (silicon) comes from sand. This is important when we are considering whether we can cover thousands of square km with cells!

Unlike other electricity sources, there are large PV markets for residential and commercial installations, as well as utility-scale PV fields. The former need only beat the *retail* cost of electricity, not the *wholesale,* so they can tolerate higher PV prices. There is also a substantial market for ***off-grid*** PV power, for remote installation or sale in countries with less-developed electrical infrastructure—in these markets, PV decisively beats the cost of alternatives such as diesel generators.

PVs do of course have the drawback of non-dispatchability, but we have seen that up to 20% of our electricity could be produced by them without demand-curve trouble. There are no technical barriers to reaching that level of PV implementation.

PV costs

Costs for PV installations are generally divided into the costs of the ***modules,*** the actual packaged semiconductor devices, and the ***balance-of-system (BoS)*** costs for installation hardware and labor, wiring, and the ***inverters*** that are needed to convert the DC output of the PV cells into the AC electricity that is required for the grid and for most household appliances.

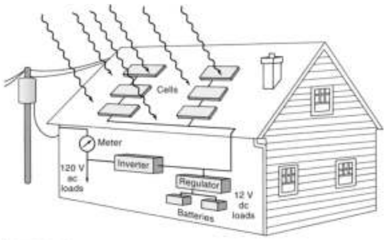


Figure 12.8: A PV system needs not just the solar cells (modules), but also mountings and connections to the AC power system. From Hinrichs (2005).

Until recently, the costs of the PV modules would be about half or more of the total costs. So to reach our goal of grid-competitive PV power, a total system cost of \$2 per Watt of capacity would require module costs of \$1/W. PV module prices have been on a sustained and dramatic fall, from ≈\$100/W in the early 70's, to \$10/W in 1990, and passed \$1/W in 2012. At the end of 2020, PV modules in the US are down to *an insanely cheap \$0.19/W* as illustrated in Figure 12.9. The tremendous drop in PV module prices has been driven by low-cost manufacturing in China, which has been the subject of major trade disagreements between China and the US and Europe, with cycles of overcapacity, bankruptcies, and subsidies in all these places. In Jan 2018, President Trump announced import tariffs on PV cells and modules from China, which have led to US prices being well above average global prices.

Figure 12.10 shows that module costs are now so low that they are a minority constituent of the total cost of PV installations, especially

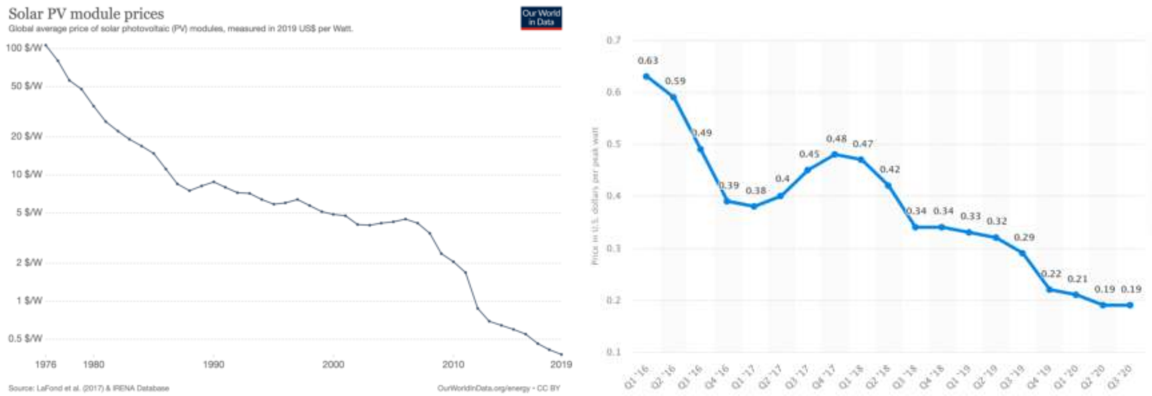


Figure 12.9: Costs of photovoltaic modules have dropped 500× in the past 45 years (left panel), down to an incredible $0.19/W in late 2020 (right panel). From OurWorldInData and Statista, respectively.

for residential installations. In 2019, the median installed cost of residential PV was about $3.70/W, according to the LBNL report, with the more optimistic Solar Energy Industry Association (SEIA) reporting average costs of $2.90/W for residential installations in late 2020. These costs vary widely between countries or across the US, indicating that some installers are significantly less expensive than the averages. Utility-scale installations control the BoS costs much better and gain economies of scale. For the largest US installations in 2019, the average price was near $1.10/W, according to the LBNL report, or $0.91/W at the end of 2020 according the SEIA. This is the price per Watt of peak power emerging from the modules; not all of this electric output is put onto the grid, so the price per actual power-grid Watt of capacity is a bit higher. Figure 12.11 shows that utility-scale installations are obtaining capacity factors of 20–25%, depending of course on regional insolation. We have reached the range where PV electricity should be cost-competitive with all fossil fuels for grid electricity, in regions with favorable insolation!

PV has for several years been the electricity source of choice for off-grid power. PVs are low-maintenance and free of capital costs, and well suited to underdeveloped locales in sunny climates. Of course battery storage must be added if night-time electricity is needed. Accessing the off-grid market has helped PV manufacturers sell enough capacity to develop the experience to lower production costs.

But reaching grid-competitive prices has vastly expanded the possible markets for photovoltaics. Let's look again at the economics. Figure 12.11 shows capacity factors attained in real-life PV installations, which we see are near 25% in the southwest and California. Note also that it is now the case that the added expense of tracking motors on PV arrays is starting to become worthwhile to raise the capacity factor to 30%. Using our CoE formula, we calculate that at

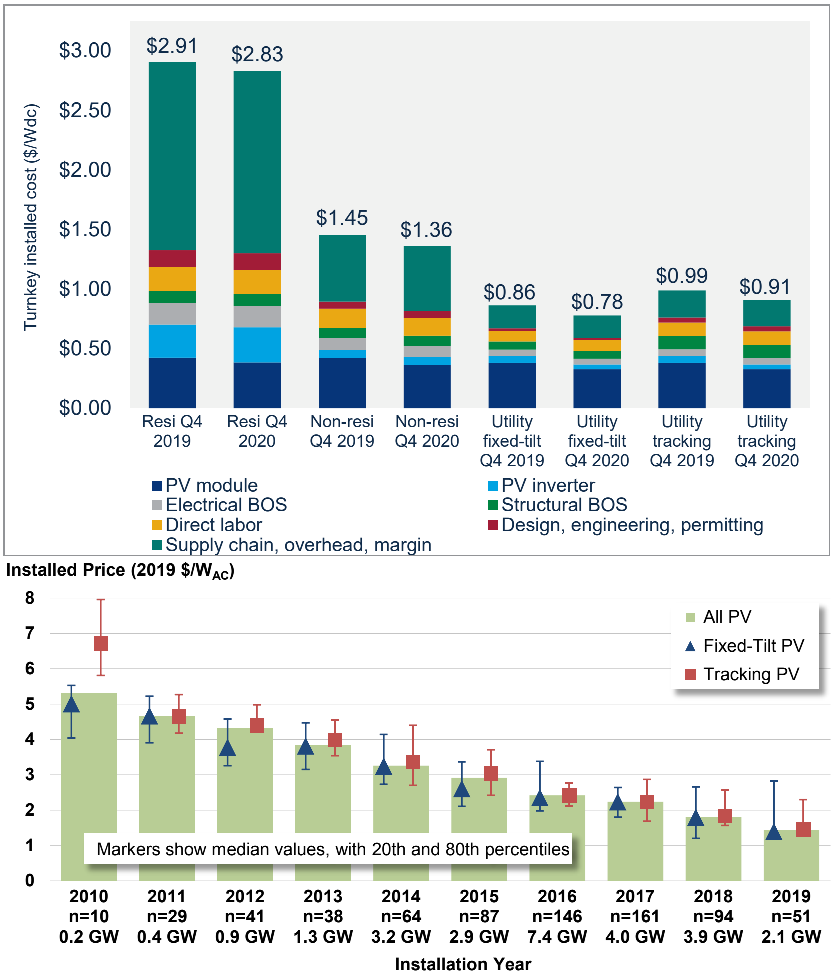


Figure 12.10: Top plot shows that most of the costs of solar installations in the US are in the "soft costs" of installation, not in the modules themselves, especially for residential installations, and that this estimate places *total* costs for utility-scale PV below $1/W. (from the Solar Energy Industry Association Market Insight report). At bottom are prices for US utility-scale PV installations over the past 10 years, now given per Watt of electric grid capacity (from the LBNL report *Utility-Scale Solar*, Bolinger et al.).

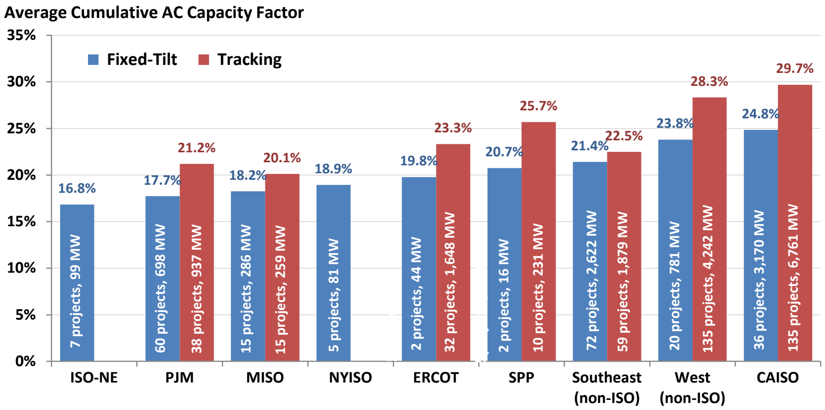


Figure 12.11: Capacity factors attained by US PV installations in 2019 in different regions of the US. "PJM" is the mid-Atlantic region, "ERCOT" is Texas, "CAISO" is California. From LBNL *Utility-Scale Solar* report.

a capital cost of \$1.10/W, $f = 25\%$, 6% interest rate, and 30-year lifetime (and ignoring O&M costs), we find CoE of \$0.036/kWh. Figure 12.12 shows the average CoE for new PV installations in the US over the years, which are indeed now below \$0.04/kWh. The actual prices at which PV electricity is being sold to utilities in the US are lower, in the range \$0.02–0.03/kWh (!), which is even lower than we calculated. The main reason for this is a Federal subsidy, in this case the solar Investment Tax Credit (ITC), which currently rebates 26% of capital costs back to the owners of a solar power installation. Under current law, the ITC will drop to 22% for construction started in 2023, and then 10% (for utility-scale) or zero (for residential) in 2024..

SEIA web site

Like wind power, a further form of preference for solar electricity is created in many states by ***Renewable Portfolio Standards,*** which mandate that a certain fraction of electricity be produced by these means. The net effect is to raise the bid prices for PV power. Subsidies for PV installations have dropped substantially at both Federal and state levels in the US. We see, however, that in favorable climates the prices for PV are now low enough to be competitive with NGCC electricity, and beat coal, even without these subsidies.

Since PV can be readily installed on residential or commercial rooftops, it can be economically viable by competing with *retail* electricity rates, whereas utility-scale PV and CSP must compete on the *wholesale* market. So even at \$3–4/W, PV can produce electricity at \$0.10–\$0.20/kWh in a top-quality solar locale, and compete with typical retail prices. Many states offer further tax incentives, and there are favorable depreciation treatments available for solar investments as well. Residential PV installations are growing rapidly, even outside the sunny southwest. An important policy is ***net metering,*** in which utilities are required to let residential or commercial power meters run backwards if a solar installation is producing more power than the consumer needs. In other words, the utilities are required to pay you as much for the electricity you produce as you pay them for electricity they produce.

Utilities are understandably unhappy about net metering. They still need to have generating capacity sufficient to replace residential solar power if it's cloudy, and they still have to maintain the grid. Net-metered customers are in effect getting to use these services for the power they make without having to pay for them.

Current PV capacity

Solar electricity, particularly PV, is growing at a phenomenal rate in a percentage sense. The EIA *Monthly Energy Report* shows that PV power *generation* in the US has increased an average of 25% per year

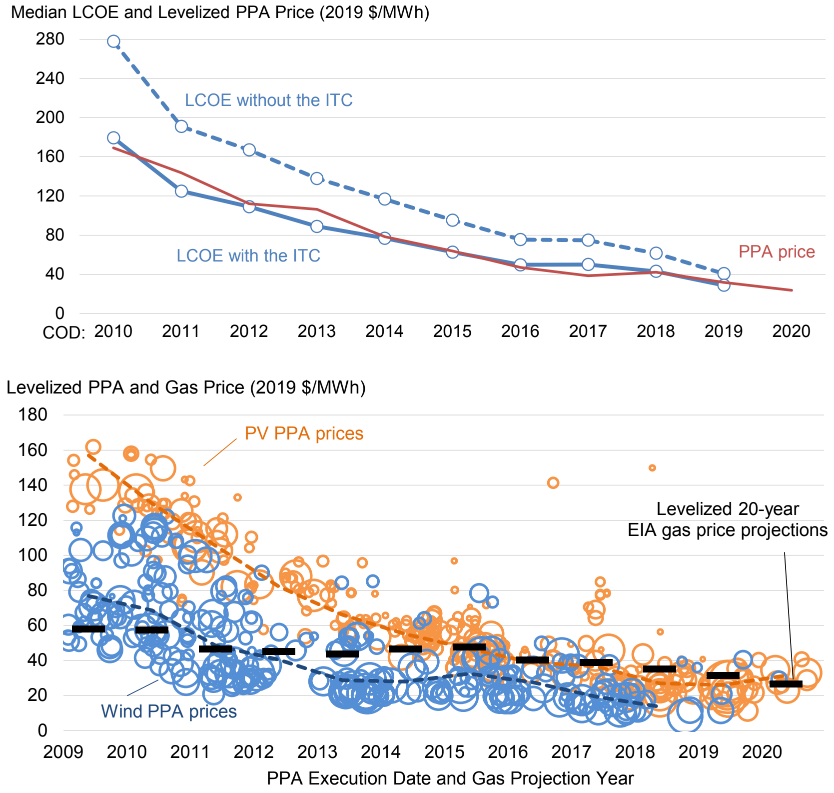


Figure 12.12: Top: The average cost of electricity (LCOE) for utility-scale solar PV installations in the US is shown vs time, both before any tax incentives are included (blue dashed) and after application of the federal investment tax credit (ITC, solid blue). In red is the average price being paid for electricity in producer price agreements (PPAs) between utilities and solar producers in the US. Not surprisingly, the prices being paid for PV electricity closely track the after-tax costs of production. Bottom: PPA's signed for wind and solar power in the US. The size of each circle represents the capacity of the facility being contracted. The heavy dashed white line is the cost of natural gas power—both renewables are now cheaper (after consideration of tax subsidies). From the LBNL *Utility Scale Solar* update in 2020.

for the last 4 years, to 0.48 EJ (133 billion kWh) in 2020, which was about 3.3% of total electricity generation. More than a third of this PV energy came from distributed sources. This is still well below the 8.5% of electricity that came from wind power that year, but PV power is increasing more quickly, in a percentage sense.

Even in the pandemic year of 2020, a record 19.2 GW of PV capacity were installed in the US, which comprised 43% of all new electric capacity. For the second straight year, more PV capacity was added to the US grid than any other kind of power. At the end of 2020, there were 97 GW of PV capacity, about 8.8% of all US generating capacity.

From SEIA *Solar Market Insight 2021*.

In summary, PV produced 3.3% of US electricity in 2020, so is not yet a dominant contributor and is still well behind wind and hydro power even among renewables. The 0.5 EJ of PV electricity produced in 2020 was only 0.5% of the total US primary energy budget. It is quite possible, however, that PV's rapid growth rate will continue as the price of PV power has dropped spectacularly and is now probably the cheapest source of new utility-scale power. In fact 23% of power generation in California is solar—high enough that occasional ***curtailments*** occur, where more PV power is being generated than

Though PV produces electricity, which for many purposes does the same work as 3× as much fossil-fuel energy, or ≈ 1.5 EJ.

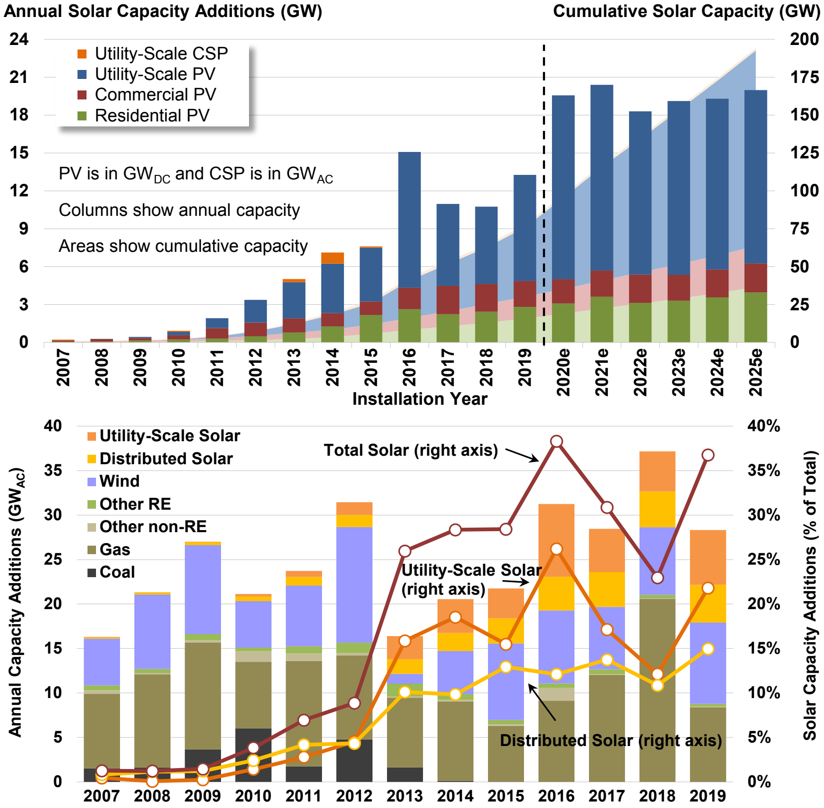


Figure 12.13: Top: total solar capacity additions and cumulative capacity in the US through 2020 (with estimates for future—the 2020 estimate was accurate). Bottom: annual additions of photovoltaic power capacity and other sources on the US grid through 2019. From LBNL's *Utility-Scale Solar* report.

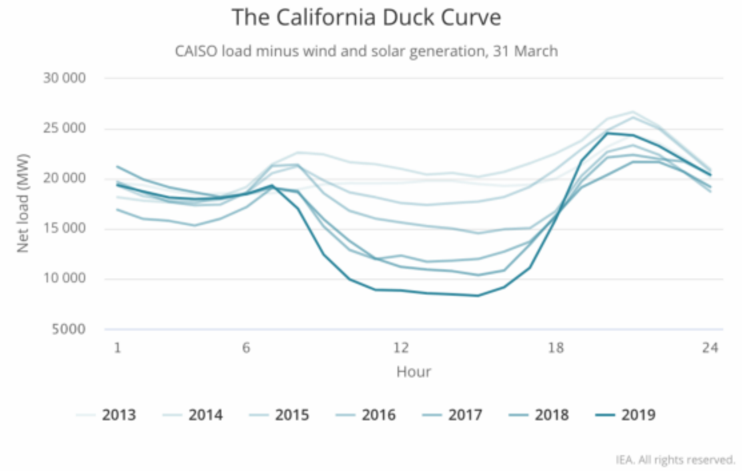


Figure 12.14: The average daily demand curves for California *non-renewable* electricity over several recent years show a growing mid-day depression as renewable capacity supplies an increasing fraction of daytime electricity.

the utilities can use, so some of this power goes unused, cutting into PV capacity factors. It also makes life difficult for California electric utilities, since utility and residential PV produce nearly all needed electricity on sunny afternoons but other sources must be turned on to cover the late-afternoon/evening demand hump. Figure 12.14 shows that California's demand for non-renewable electricity now *drops* on a typical afternoon as solar power peaks, and typically has a maximum in the evening. This has required special attention from utility operators, but has not yet broken the grid. It is true, however, that a few percent of wind and solar generation is curtailed. Nonetheless, solar power is so economical in California that in 2019 it became the first state to *require* installation of rooftop PV on new-home construction.

In 2020, about 10% of US electricity came from PV+wind, and this is likely to approach 20% within a few years. Demand management and storage will become increasingly important to continued expansion; and the rate at which we add PV and wind capacity to our grid will need to increase quite substantially if we are to make the US nearly carbon-neutral in the 2030's. That goal will require vastly increasing our electrical capacity so we can displace fossil fuel use in transportation and heating.

Global PV

The REN21 *Renewables 2020 Global Status Report* estimates that in 2019, 115 GW of PV capacity were installed worldwide, for a total of 627 GW (Figure 12.15). As in the US, the annual installations are increasing about 25% per year. China has the most installed capacity with the US second, but PV power is progressing in other countries too, such as India.

The US is not the most advanced country in solar electricity in terms of percentage of power coming from solar. Japan, Italy, and Germany each produced 6–8% of all their electricity from PV in 2018. The REN21 report notes that rooftop PV electricity is cheaper than retail electricity in multiple countries, including Australia, Brazil, Denmark, Germany, and Italy. As of April 2017, it is reported that 21% of suitable rooftops in Australia have PV systems installed.

Renewable summary

The rapidly dropping costs and increasing production capacity of PV, combined with its ease of installation and financing, are likely to make PV an increasing part of world electricity production in the near future, perhaps reaching several percent, even in the absence of subsidies meant to internalize the cost of fossil-fuel CO_2 emis-

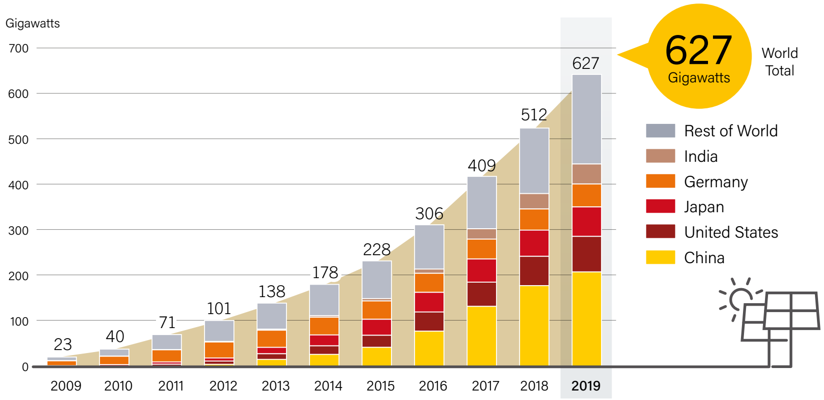


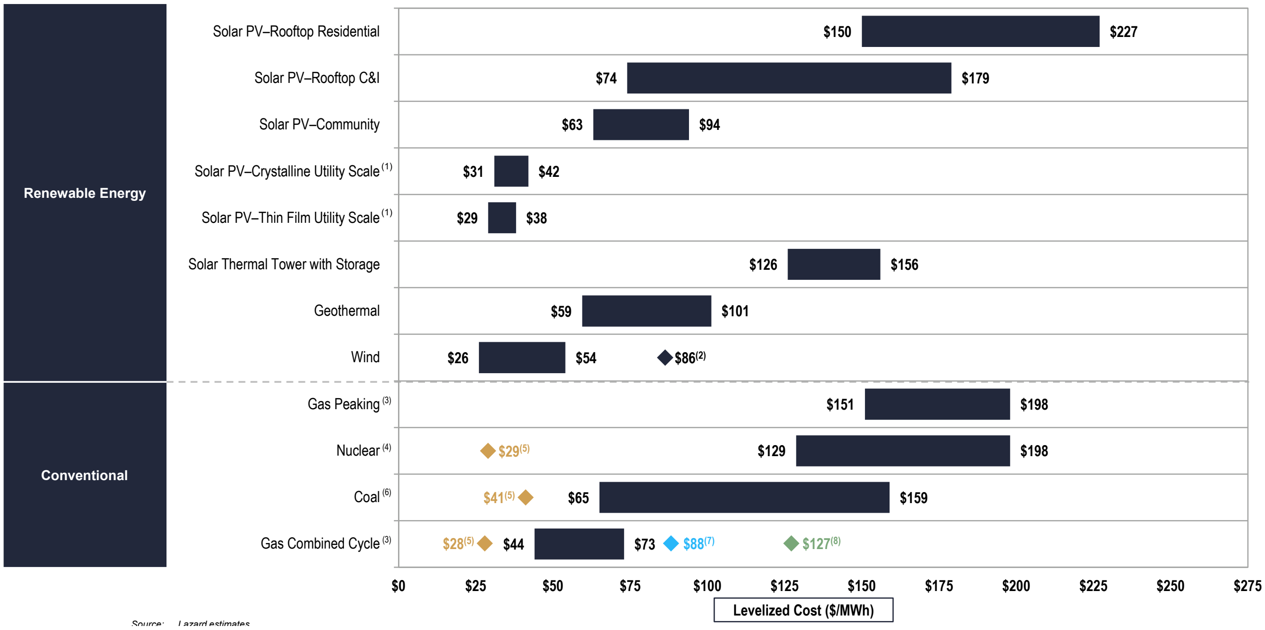
Figure 12.15: Global PV capacity over the past 11 years, showing the cumulative total, and several countries' shares of the total capacity. From REN21 *Renewables 2020 Global Status Report*.

sion. The technology is clearly capable of supplanting fossil fuels entirely, at prices below those of current fossil-fuel energy in many locales. Figure 12.16 shows that current estimates of the LCOE of wind and solar are below that of gas, especially after application of tax subsidies, and this is confirmed by the real prices being offered for electricity contracts. There is substantially more wind capacity than PV at the moment, but more PV is being installed as PV prices drop more rapidly. Internationally, we should expect that PV electricity will play a very large role in supplying growing electrical needs in developing countries, particular in equatorial regions, given the simplicity of buying, installing, and maintaining PV panels.

Future growth of PV and wind will depend on improving grid capacity and, importantly, the progress on viable forms of grid-scale energy storage, as well as on whether manufacturing capacity and international trade practices can keep up with rapidly growing demand.

Practice problems

- How big a tank (what diameter and height) of molten salt would be needed to store enough heat to generate a day's output from a 1 GW solar thermal power plant? Critical values here: the specific heat of molten salt is about 1.6 kJ/kgK; its density is 1.7 kg/m^3; the salt is heated from 280° C to 560° C; and the steam generator is 40% efficient.
- If PV power generation grows by 25% per year, when will it generate 20% of all US electricity? 50%?
- If we wanted to replace nearly *all* fossil-fuel use in the US with PV, i.e. by switching our transportation sector to electric vehicles and



.

space heating from natural gas to heat pumps, how many EJ of electricity would we need to produce per year? How long would that take to build if we keep growing at 30% per year? How much would the PV installations cost at $1.50 per W of capacity?

Figure 12.16: Estimates of the levelized cost of electricity (LCOE) for 2020 construction of different power sources in the US, from Lazard-Freres. The plot shows values *before* government subsidies; current subsidies lower the utility-scale PV costs to the $26–42/MWh range, and wind to $9–43/MWh range, well below NGCC power. The yellow diamonds are the marginal costs of generating electricity from plants that are *already built*.

13
Biofuels

PHOTOSYNTHESIS PROVIDES THE ENERGY for virtually all life on Earth, some tiny fraction of which has been stored away as our fossil fuel supply. Before the coal era, burning of organic materials (wood, dung) was the largest source of human-harnessed energy, and it remains so in some underdeveloped societies. Is it possible for us to scale up some form of agriculture to provide chemical fuels that produce a substantial fraction of our energy supply? Or perhaps reproduce photosynthesis in a synthetic reactor instead of living cells?

Many countries, including the US, make significant energy by burning biomass on or near the site of growth (e.g. paper mills generate energy for their processing by burning wood waste). We will concentrate here on ***biofuels,*** whereby the energy in biomass is converted into an engine-ready chemical fuel. Biofuels are already in large-scale production in two countries. In 2106–2019, the US produced about 1.25 EJ/yr of ethanol via fermentation of corn,[1] which is roughly 4% of the 30 EJ energy used in transportation and 1.2% of *total* primary energy. Brazil produced 1.0 EJ of biofuels in 2019, mostly ethanol from fermentation of sugarcane, roughly 8% of their *total* energy consumption.

Corn ethanol production was down 20% in 2020, due at least in part to lower fuel demand during the pandemic.

In principle, biofuels are sustainable in that all the carbon emitted into the atmosphere by their combustion is carbon that was removed months earlier through photosynthesis. Likewise, biofuels can enhance energy security by displacing petroleum, the least secure source of primary energy, in the transportation sector. In practice, however, we need to think carefully about the sustainability of biofuels:

- Energy is used during the cultivation of biofuel feedstocks and the processing into fuel. Do the biofuels deliver significantly more energy and displace more usage than it requires to produce them?

[1] EIA *Monthly Energy Report* Mar 2021, Table 10.3.

Is the *net* emission of CO_2 negative after considering the fossil-fuel inputs?

- Growing feedstock for biofuels requires repurposing land. The change in land use, particularly if it means clearing forest, often causes emission of CO_2 that must be considered in the balance.

- Many biofuel feedstocks are also elements of the global food system. Will it be sustainable, or even moral, to divert part of the food supply into the energy supply?

- Many crops require water. Is there sufficient sustainable water supply to produce enough biofuels to power the transportation sector?

Here we will examine the current ethanol regimes for fossil-fuel use and CO_2 emission on a "life-cycle" basis, and then ask if there are potential better biofuel options on the horizon. But first we will estimate the ultimate size of the biofuel resource. I highly recommend reading MacKay's biofuel discussions (in the two "Solar" chapters); they are specific to the United Kingdom, whereas we will have a US bias, as usual.

Efficiencies and resource size

There are two different biochemical photosynthesis pathways, called $C3$ (used by algae and most grasses and grains) and the more recently evolved $C4$ pathway used by more dicots. Zhu, Long, & Ort (2008) report that the theoretical maximum efficiency for these two families is to convert 4.6% and 6.0% of incident sunlight into chemical energy of biomass, respectively.[2] They cite the highest efficiencies ever observed in real C3 and C4 plants as 2.4% and 3.7%, and most major crops are closer to 1% efficiency. Furthermore this is the gross energy in the biomass, which we might extract by simply burning all of the biomass. But if we want to produce liquid fuels from biomass for use in transportation, we need to convert the complex carbohydrates of raw biomass into practical engine fuels, and so far we are able to extract only a fraction of the biomass energy in this way.

The 1% or lower efficiency of biofuels makes it clear they will require more land area to produce a given amount of energy than CSP or PV power—plus the fact the electricity produced by CSP or PV will do most tasks with 3–4× less energy than chemical fuels. There are however two very good reasons to nonetheless pursue large-scale biofuels: first, biofuels could be significantly *cheaper* per Joule of output than PV even if they require more land. Second, liquid

[2] Zhu, Long, & Ort, *Current Opinion in Biotechnology,* **19,** 153–159 (2008). This is for current atmospheric temperatures and CO_2 levels, as efficiency can rise at higher CO_2 concetrations.

chemical fuels have much higher energy storage per kg than batteries that store electricity, and this makes them much more practical for transportation—and indeed *essential* for air transport (see Chapter 14).

If we optimistically take 1% efficiency of conversion of solar energy into useful fuel energy for biofuels, then the total output energy available from the average 4.5 $\mathrm{kWh/m^2/day}$ insolation across the full $7.7 \times 10^{12}\ \mathrm{m}^2$ of land in the continental US, we obtain a maximum energy supply of 450 EJ/yr. Only 17% of this land is currently under cultivation, so if we were to use all of this land we could produce perhaps 80 EJ/yr of biofuels, similar to annual yearly energy consumption. At 1% efficiency it would take roughly 1/3 of currently cultivated land to produce the 30 EJ of fuels used annually for transportation. So for biofuels to be the future of transportation, or take an even larger role, we will need some combination of:

- Very high efficiency in converting biomass into biofuels.
- Self-sustaining biofuel conversion processes (small fossil-fuel inputs).
- Converting 10's of percent of the country's land into agricultural use, and biomass crops that can make use of land not currently considered appropriate for agriculture.
- Shifting a substantial fraction of current farmland from food production to energy production, and find the food elsewhere.

Corn ethanol

At present, commercial-scale biofuel production occurs using the ancient technology of ***yeast.*** These organisms are capable of metabolizing *sugar* into *ethanol.* This is how we've been making beer, wine, and liquor for millenia. Only the part of the plant's chemical energy stored as sugar can be metabolized: in sugarcane, a sugar-rich crop, a bit less than half of the chemical energy content is in sugar. The majority of biomass chemical energy is stored in the matrix of *lignin* and *cellulose* that make up the plant cell walls and support the plants. These more complex molecules cannot be digested by yeast. They form the "dietary fiber" which humans (or other vertebrates) cannot digest.[3] This resistance to decay helps make wood such a good construction material! The more straw-like cellulosic residue from sugarcane cultivation is called *bagasse.*

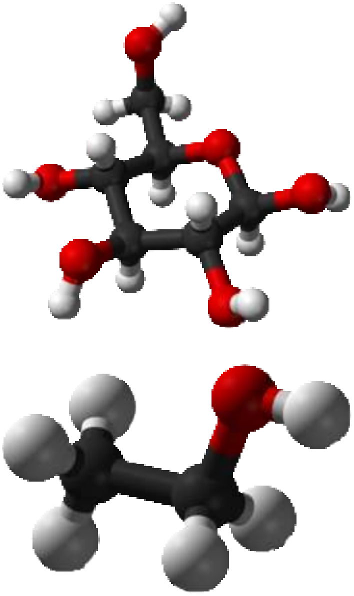


Figure 13.1: Molecular models of glucose (at top), $C_6H_{12}O_6$, which yeast consume, and ethanol (at bottom), C_2H_5OH, which they excrete. Black is carbon, grey is hydrogen, red is oxygen. From Wikipedia.

Sugarcane cannot grow in most of the US, but another sugar-rich crop is corn, which is planted on nearly 1/3 of all cultivated land in the US. The plant produces sugar in the kernels; the remainder of

[3] Horses, goats, even termites that eat cellulosic foods host symbiotic bacteria that break the cellulose into digestible sugars.

cellulosic plant material (stalks, etc.) is called *corn stover.* We have become masters of corn-growing and in the US have undertaken fermentation of corn to ethanol via yeast on a massive scale: in 2019, 1.25 EJ of corn ethanol was produced and blended with gasoline for commercial sale. By comparison, hydro, wind, and solar produced 1.04, 1.21, and 0.48 EJ of electricity in 2020, so in raw Joules, corn ethanol is the single largest source of "renewable" energy in the US for many of the past 20 years. But note that hydro, wind, and solar produce *electricity* Joules, which for most applications would be considered more valuable than an ethanol Joule that is fed into a heat engine.

Ethanol is almost as good as gasoline as a motor-vehicle fuel. Its energy released per kg of fuel is lower than gasoline's because the oxygen atom adds weight, but not combustion energy, to the fuel molecule. The lower heating value (LHV) of ethanol is 27 MJ/kg vs 43.5 MJ/kg for gasoline. Since the densities are similar, this means that *a gallon (or liter) of ethanol contains only 2/3 the energy of a gallon (or liter) of gasoline.* So we should be careful not to compare the price per gallon of the two fuels—use price per MJ! Ethanol can be blended into gasoline, e.g. by producing E15 which is 15% ethanol, 85% gasoline. Automobile engines designed for gasoline use can burn E15 with no modification required. With some modification, engines can run on 85% ethanol (E85) that is available at some Midwestern filling stations. In the US these are labelled "Flex-Fuel" vehicles.

Pure ethanol is harder to handle than gasoline or E15. Water dissolves in ethanol and degrades the combustion properties. This, along with ethanol's corrosive properties, make it difficult to transport ethanol in pipelines. Ethanol is current transported by rail and truck. And of course the input biomass must be trucked to processing sites.

Now let us examine the efficiency of corn ethanol production. Figure 13.2 shows the remarkable and continuing increase in corn yields that has been achieved in the US over time. In an average-weather year, yields are now above 170 bushels per acre (using the archaic but standard units). A bushel of corn has a mass of about 25 kg, and an acre is about $4000\,\mathrm{m}^2$. From Figures 13.3, we see that an astounding 3.6×10^{11} m^2 of corn were planted in 2020, or *5% of the entire land area of the contiguous US!* These acres produced 3.6×10^{11} kg of corn, or nearly 1100 kg per person per year. Did you eat your share? Probably not; very little of this corn is consumed by people, and in fact most corn planted is a variety inedible to humans. The great majority of the corn goes to feeding animals (which we then eat) and to producing ethanol. An area roughly the size of Illinois is currently devoted to growing corn for fuel ethanol!

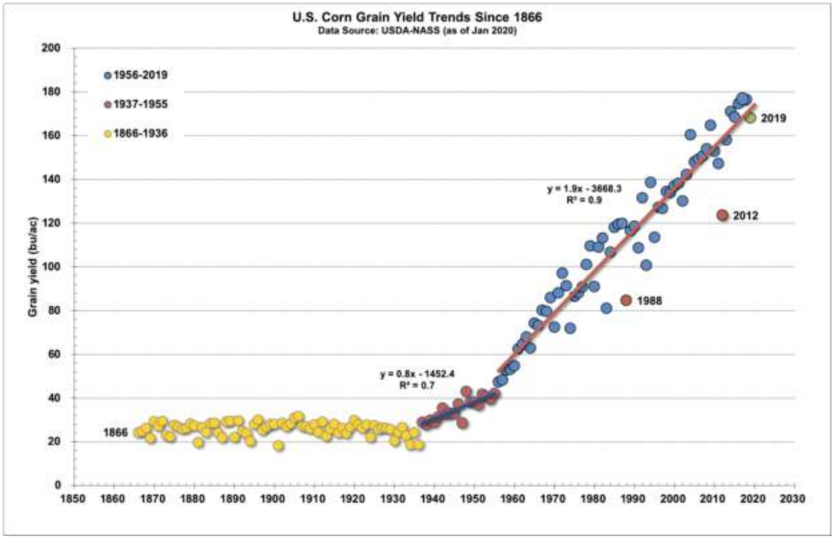


Figure 13.2: History of corn yields in the US, from Purdue Univ..

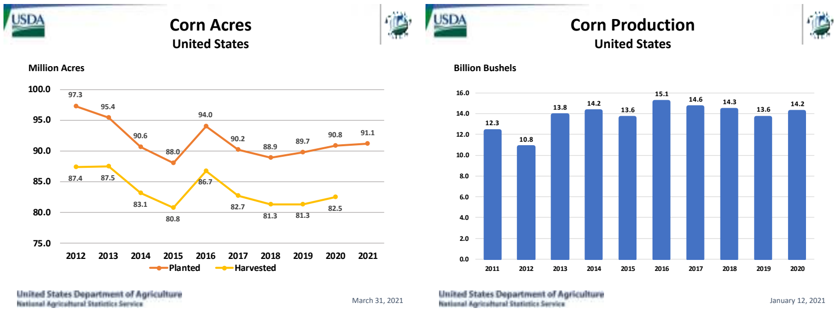


Figure 13.3: The total US acreage planted in corn each year (left) and the total harvest (right). From the US Department of Agriculture's National Agricultural Statistics Service.

One bushel of corn, when fermented, yields about 2.6 gallons of ethanol with an HHV of 87 MJ per gallon. This means that ethanol power produced per unit area of land is

$$\frac{P}{A} = \frac{170\,\text{bu}}{\text{acre yr}} \times \frac{2.6\,\text{gal}}{1\,\text{bu}} \times \frac{87\,\text{MJ}}{1\,\text{gal}} \times \frac{1\,\text{acre}}{4000\,\text{m}^2} \times \frac{1\,\text{yr}}{3.1 \times 10^7\,\text{s}} \qquad (13.1)$$

$$= 0.31\,\text{W m}^{-2}. \qquad (13.2)$$

With the average insolation of corn fields being about $4.5\,\text{kWh m}^{-2}\,\text{day}^{-1} = 190\,\text{W m}^{-2}$, this makes the efficiency of corn ethanol production

$$\epsilon = \frac{\text{ethanol power out}}{\text{solar energy in}} = \frac{0.31\,\text{W m}^{-2}}{190\,\text{W m}^{-2}} = 0.16\%. \qquad (13.3)$$

This is 6× lower than the 1% efficiency we used to estimate the total potential of biofuels in the US. At the real-life efficiency, we would produce only 12 EJ of corn ethanol per year even if devoting *all* currently cultivated land to the effort. So we cannot hope to replace our oil with corn ethanol, although we could replace 10% or so by devoting a large fraction of our arable land to it (and indeed we do use about 12% of our arable land for ethanol production now).

To make matters worse, the above values are the *gross* energy output of corn ethanol production. We need to account for the energy used in producing the corn, transporting corn and ethanol, running the fermentation process, and particularly producing the fertilizers that are needed to generate high yields. Over the past 15 years, several authors have calculated the net energy and CO_2 impact of corn ethanol production. The most favorable results are from the Department of Agriculture's calculation (Shaphouri et al. 2010), who estimate:

- 16 MJ of energy used to grow the corn that makes 1 gallon of ethanol.
- 46 MJ used in transportation and fermentation, for a total of 62 MJ of energy input.
- 89 MJ of energy output in the HHV of the ethanol.
- 21 MJ of energy are obtained from ***coproducts*** of production, which are byproducts like distiller's grain that can be burned for heat in the process or otherwise displace the use of a fossil fuel.

So a net fossil-fuel input of $16 + 46 - 21 = 41$ EJ is needed to produce the 89 MJ of ethanol energy. By this estimation, the net energy value of the corn ethanol is only half as big (per gallon) as the gross value above, or an $\approx 2:1$ gain of energy production per fossil-fuel input. The more pessimistic calculations suggest that corn ethanol has a *negative* net energy production, meaning that each gallon we produce actually *increases* fossil-fuel energy consumption by more than just burning gasoline in the car instead of ethanol! Most studies end up somewhere in between, with a positive energy return of about 1.2–1.3:1 from corn ethanol. A 2018 USDA study reports that corn ethanol production produces 61% of the effective GHG emissions of the energy-equivalent amount of gasoline. While boosters will tout the "net positive energy production" from corn ethanol, the fact remains that it's likely that more than half the energy we are getting from ethanol is really reprocessed fossil-fuel energy, and corn ethanol is at best 40% "renewable" in a greenhouse-gas sense. If ethanol production is getting 50–80% or more of its fuel energy from fossil fuels, then clearly it is not a zero-net-carbon form of energy. We have to ask how much lower the net carbon emission of corn ethanol is compared to simply using gasoline. This will depend on many factors, particularly what source of energy the farming and distilling is using—is it coal or natural gas, for example? Figure 13.4 from the US Congressional Budget Office (CBO) collects estimates of the net CO_2 emission reduction from a switch from gasoline to corn ethanol. The most favorable estimates are that corn ethanol use eliminates 1/3 of CO_2 emission; other reliable estimates suggest that corn ethanol use actually *increases* CO_2 emission, particularly if we account for the diversion of land from other uses.

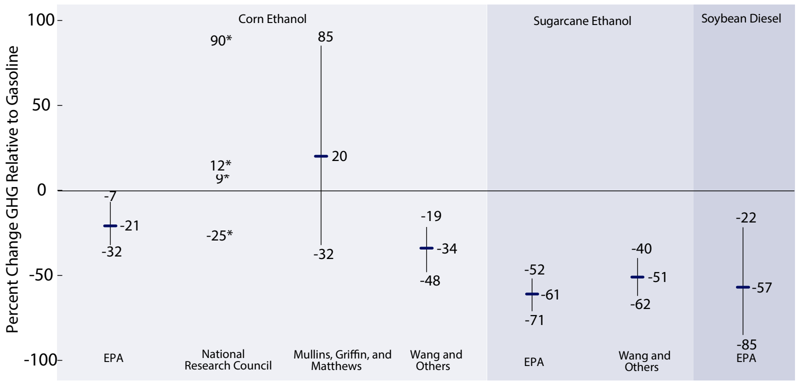


Figure 13.4: Various estimates of the net CO_2 emissions from production and use of corn ethanol (left), sugarcane ethanol (middle), and soybean biodiesel (right) are collected here by the Congressional Budget Office's report on *The Renewable Fuel Standard: Issues for 2014 and Beyond.* A value of 0% means that the fuel's net CO_2 emission is the same as using gasoline. A value of -100% would mean zero net CO_2 emission.

So does corn ethanol succeed in increasing energy security and sustainability?

- All agree that it *does* reduce the amount of petroleum needed. The fossil-fuel inputs to corn ethanol production can be primarily from coal and natural gas, hence shifting the transportation fleet away from petroleum dependence.

- Corn ethanol is at best a minor assist in reducing total fossil-fuel usage. Typical estimates are that each Joule of ethanol produced displaced only ≈ 0.3 J of fossil-fuel energy.

- Corn ethanol is at best a minor assist in reducing total CO_2 emission. It is likely somewhere between a 40% reduction and an *increase* in net GHG compared to gasoline.

- It is *not* a viable fuel for replacement of the bulk of petroleum in transportation. Even using the gross efficiency of 0.29% rather than a lower net efficiency, there is not enough farmland in the US to support the transportation fleet.

Figure 13.5 shows that despite corn ethanol's sustainability benefits being minor at best, its use increased substantially over the 2000–2010, and the fraction of US transportation energy that comes from biofuels has since held fairly steady at $\approx 5\%$.

If the sustainability benefits of corn ethanol are minor at best, why are we producing so much? Perhaps the price is lower than gasoline? This is not the case. Corn prices are highly volatile—the Midwestern weather matters a lot here. And the wholesale cost of buying ethanol has usually been the same or higher than buying the energy-equivalent amount of gasoline. So how can ethanol be getting sold when it costs more than gasoline? There are several reasons:

- At the retail level, E85 prices track gasoline prices, but typically so that the price per *gallon* is nearly the same. Consumers apparently do not realize that ethanol has only 2/3 as much energy as gasoline and therefore they should only pay 2/3 as much.

- More importantly, the US government offered a tax credit of $0.51 for every gallon of ethanol that a fuel blender put into gasoline. This lowered the effective cost of ethanol directly. This tax credit expired at the end of 2011, after being the most expensive Federal alternative-energy subsidy for many years.

- Even more importantly, since 2005 the US government *mandates the inclusion of a certain quantity of ethanol into gasoline!* This is called the ***Renewable Fuel Standard (RFS).***

We note that the RFS and the blender's credit are not the only subsidies given to corn ethanol. Corn growers are paid several billion dollars per year in direct payments from the US government. To protect the US ethanol market, an import duty of $0.54/gallon was placed on Brazilian sugarcane ethanol. This also expired at the end of 2011. Hence through the 2000's, corn ethanol was the target of the largest alternative-energy subsidies and mandates even though it is contributing very little to sustainability and (at that time) was probably increasing net CO_2 output. This painful irony is often attributed to the fact that Iowa is the site of one of the first two presidential primary elections—it was considered mandatory for candidates to travel to Iowa and pledge their fealty to corn ethanol, and most presidential candidates still do this.

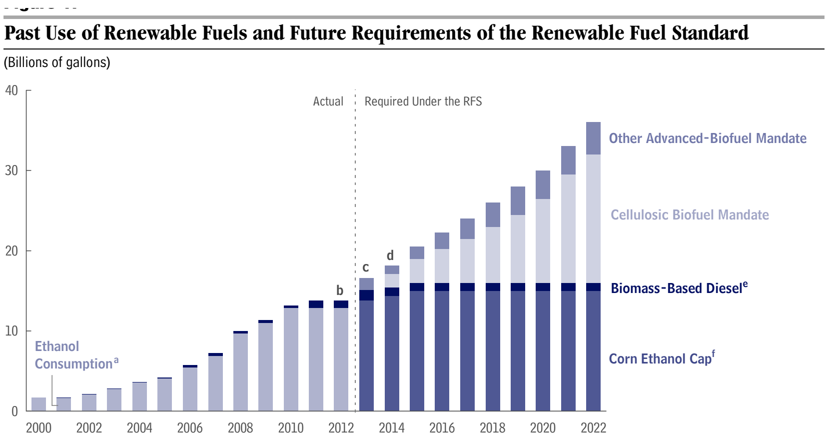


Figure 13.5: The amount of ethanol required to be blended into gasoline according to the Renewable Fuel Standards legislation.

As it became more widely recognized that corn ethanol is not delivering environmental benefit, and the recession hit, and many lawmakers were strongly focused on reducing Federal spending, the ethanol subsidies became vulnerable and were canceled. The RFS mandates remained, however, after their renewal in 2007, but they did cap the contribution of corn ethanol (at the now-current level) and it was mandated that future growth in ethanol production come from more sustainable sources than corn fermentation, as illustrated in Figure 13.5.

Better biofuel possibilities

We can improve substantially on yeast-based corn fermentation as a source of biofuels. Brazil produces a large fraction of its transportation fuel by fermentation of cane sugar. The sugar (and ethanol) output per unit of input energy is much higher, resulting in larger net reductions of fossil fuel input and CO_2 output relative to gasoline. [In the continental US, sugarcane grows well only in Florida.] Figure 13.4 shows estimates that cane ethanol probably yields a net reduction in CO_2 emission of 50% or more relative to gasoline. This climate-change advantage is lost, however, if we get the land for sugarcane agriculture by clearing rain forest.

The higher energy yield per unit area, and lower processing energy/costs of cane ethanol make it cheaper to produce than corn ethanol. But as noted earlier, the US had a substantial import duty on Brazilian ethanol in order to protect the domestic producers.

The major anticipated breakthrough in biofuels is the development of a commercially viable process for ***cellulosic ethanol,*** i.e. some kind of bacterium or enzyme that can metabolize the tougher cellulose or lignin molecules into a usable fuel, or into sugar that can be fermented with yeast into ethanol. This would open the door to utilization of a much larger fraction of the chemical energy produced by the plant, for instance by using corn stover or cane bagasse in addition to the sugar.

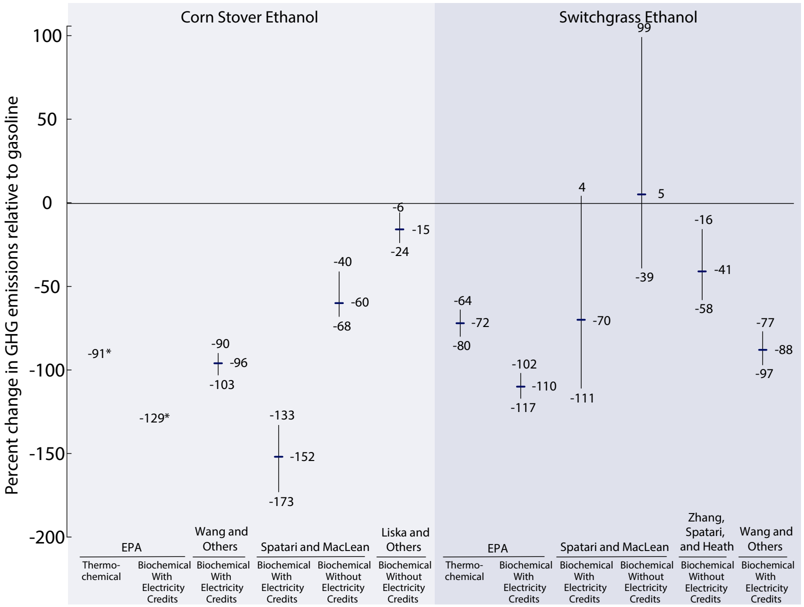


Figure 13.6: CBO collection of estimates of greenhouse gas emission reductions from switching gasoline usage to cellulosic ethanol production.

Perhaps more attractive is the production of cellulosic ethanol from crops that are faster-growing and more robust than corn, e.g. requiring less water and fertilizer to thrive. One such proposed energy crop is switchgrass. Accessing more of the photosynthetic output, and requiring less energy input, would raise the net energy per unit area and also drop the net CO_2 reduction. Figure 13.6 gives estimates of the latter for corn stover and switchgrass cellulosic ethanol, which we see are often close to 100% (meaning zero net carbon emissions) or sometimes even *more* than 100%. This seems impossible but it's not: the production of fuel ethanol can have coproducts that displace additional fossil fuels, and/or some of the biomass can be sequestered to remove carbon permanently from the atmosphere.

Cellulosic ethanol is promising as a potential route to more sustainable liquid fuels, but unfortunately no economically competitive pathway has been developed. When the RFS legislation was last updated in 2007, numerous venture-capital-funded companies were claiming to have commercially viable processes almost ready to go. As can be seen from Figure 13.5, the RFS mandates production of large and increasing quantities of cellulosic biofuel as of 2013, doubling ethanol production by 2022. However the EPA has had to rescind the RFS mandates to date because there were no large-scale plants in existence. In 2014, the first few commercial cellulosic ethanol plants opened in the US, producing < 100 million gallons of ethanol per year (< 0.009 EJ). The production capacity for cellulosic ethanol is far below even the first steps on the cellulosic ethanol mandate ladder; the EPA RFS requirements for 2020 are for only 5% of the amount of cellulosic ethanol originally written into the RFS law in 2007. At this point the prospects for significant production of cellulosic ethanol to the US or world energy budget are dim.

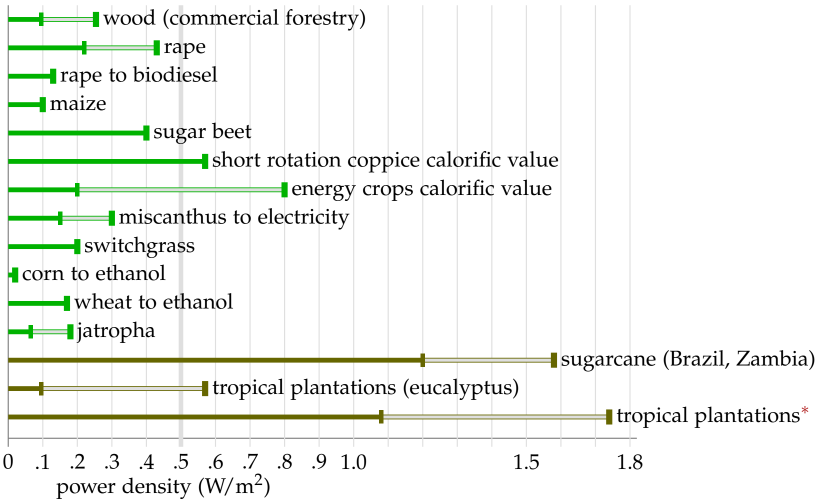


Figure 13.7: MacKay's estimates of the net energy yield of different biofuel crops in the UK (green) or tropical climates (brown).

Let's be optimistic and assume that an economically viable process for converting cellulosic biomass into fuels is developed. What would the prospects then be for generating biofuels to replace current petroleum use? Figure 13.7 illustrates MacKay's pessimism on this account by showing estimates he has collected for net energy production per unit land area. He considers not only ethanol production, but also fast-growing oil-seed crops (rapeseed, jatropha) that can be refined to generate fatty carbohydrates suitable for use in engines (biodiesel). He finds all of them short of 1 W m^{-2} in temperate climates (the UK) and potentially slightly above 1 W m^{-2} in tropical climates. Note that these represent net conversion efficiencies of $\approx 0.5\%$

or less, a factor two less than we took as our upper limit. For the UK, he estimates biofuels as capable of being at best a minor contributor to the energy budget. In the sunnier, more sparsely populated US, we might be able to consider replacing a substantial fraction, though not all, petroleum usage with biofuels, again requiring us to commit a substantial fraction of all land area to growing feedstock. We could, for example, easily imagine generating enough biofuels to fuel air transportation, where electricity from other renewables is not a feasible replacement.

Another element would be cost. A 2011 National Academy of Sciences panel estimated of the costs of feedstocks for cellulosic ethanol and concludes that cellulosic ethanol would become competitive with gasoline only when oil prices to rise to $\approx$ \$200 per barrel (or with carbon taxes or ethanol subsidies that make up the difference with a lower price). This supports my guess that biofuels will find use primarily in countries with very favorable land/sun/water endowments.

In the next chapter we will see that switching the entire transportation sector to renewable electricity is a very difficult job—air transport, for example, will probably still require liquid fuels. Biofuels could be a route to a lower-carbon (if not zero-carbon) way to maintain demanding transportation applications. If we get to the point where air travel is our dominant source of CO_2 emissions, then I believe we can claim victory over greenhouse gasses, and could tolerate the not-really-renewable aspects of biofuel usage to supply this need.

While biofuels do not offer a currently forseeable route to supplying $> 10\%$ of total energy demand, we should keep in mind that biotechnology may surprise us. Selective breeding and genetic engineering have been extremely successive at raising yields of commodity crops like corn, and we should expect that yield gains will occur for potential energy crops, maybe giving us a 1%-efficient crop with minimal water needs. There is an intensive research effort worldwide in breeding or engineering new photosynthetic organisms (plants or algae) that can generate liquid fuels more directly. And ultimately it may be possible to create bio-inspired photosynthetic processes that operate outside living cells. This may allow biofuels to make a larger impact in the more distant future.

But for now, the story of corn ethanol should serve as a cautionary tale. The RFS was a centerpiece of the Bush energy policy, sold as a way to increase US energy independence and cut carbon emissions at the same time. In fact it has probably increased carbon emissions, led to increased food prices, and the conversion of large amounts of land to pointless corn and sugarcane production. It was not hard to predict that this would be the case if anyone looked in any detail at the quantitative story, but the political appeal of a subsidy of many billions of dollars per year to midwestern farmers while falsely claiming to help the environment proved too great to resist. While the good news is that the direct ethanol payments were phased out several years ago, the bad news is that the ethanol RFS remains in place.

14
Transportation and Energy Storage

TRANSPORTATION IS THE MOST DIFFICULT challenge for decarbonization of the energy supply because it requires a portable energy supply. And it relies on petroleum, currently the most expensive of the fossil fuels and the one with least secure supplies for the US. Can we find a new primary source for the 30 EJ per year we use for transportation? Can we reduce the 28% of our energy use and 36% of CO_2 emission that comes from transportation?

In the US in 2018, according to EIA MER.

The transportation system's energy use can be factored into three elements:

1. *The amount of energy it takes to run the vehicles:* how many km need the vehicles travel, how many Joules at the driveshaft (i.e. mechanical energy) does it take to move the vehicle 1 km, and at what efficiency does the engine of the vehicle convert its fuel supply into driveshaft energy?
2. *The storage medium for the energy:* in what form is the energy kept aboard the vehicle?
3. *The primary source of energy:* where are we getting the Joules from in the first place? How efficiently is the primary energy converted into the storage? For the now-standard gasoline-powered internal combustion engine (ICE), the primary source is petroleum, basically the same as the storage medium except for the refining step. But if the storage medium is a battery, then the electricity used to charge it can be produced from a variety of fossil or non-fossil fuels.

Figure 14.1 is from the *Transportation Energy Data Book (TEDB)* issued by the US Department of Energy's Oak Ridge National Lab, a publication we will use frequently in this chapter. The graph shows the fraction of transportation energy used by different types of vehicles, and the forms of fuel that are used by each type. We will focus mostly on the "light vehicles"—cars, motorcycles, plus the "light trucks" such as vans, pickups, and sport-utility vehicles (SUVs) used for personal transport—that account for the majority of transportation energy use.

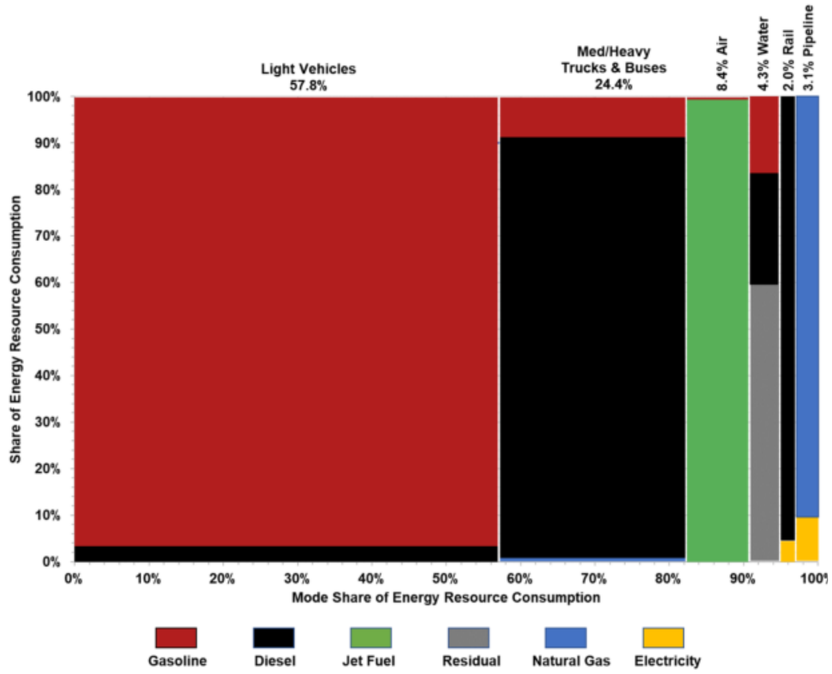


Figure 14.1: Division of transportation energy use and fuel type among types of vehicles in 2018. Figure 2.6 from the 2021 *TEDB*

Energy use of vehicles

Let's break down the primary energy use into four factors:

$$\frac{\text{primary energy use}}{\text{year}} = \frac{\text{vehicle-miles}}{\text{year}} \times \frac{\text{driveshaft energy}}{\text{vehicle-mile}} \times \frac{\text{stored energy}}{\text{driveshaft energy}} \times \frac{\text{primary energy}}{\text{stored energy}} \tag{14.1}$$

We recognize the last two terms as the (inverses of the) efficiencies of the onboard engine and the primary/storage technologies, respectively, which we will examine later.

We are also of course interested in CO_2 emission per year. To get this, we would take the above primary energy use per year and multiply by the $CO_2$2 emission per unit energy for our source of primary energy. Remember that this is *not* necessarily zero for an electric vehicle, since the primary energy could be coal or natural gas.

Vehicle usage

The first term is simply how many cars are driven how far each year. Clearly if we can reduce the number and/or distance of vehicle trips taken, we reduce energy consumption and CO_2 emission. Unfortunately the number of vehicle-miles driven is literally cast into concrete: aside from a handful of city cores with viable public transportation networks (Chicago, NY, Boston, DC, SF, Philadelphia, ??), the entire country is constructed under the assumption that every trip will be made by private automobile. Vast numbers of people in the US live in low-density suburbs where work and shopping are well beyond walking distance and public transportation coverage is sparse. Transportation is yet another area where high-density city living is much more favorable from an energy and environmental perspective.

					Energy intensities		
	Number of vehicles (thousands)	Vehicle-miles (millions)	Passenger-miles (millions)	Load factor (persons/ vehicle)	(Btu per vehicle-mile)	(Btu per passenger-mile)	Energy use (trillion Btu)
Cars	111,242.1	1,419,571	2,186,139	1.5	4,374	2,840	6,209.2
Personal trucks	3,678.4	1,235,245	2,248,145	1.8	5,966	3,278	7,369.4
Motorcycles	8,666.2	20,076	24,091	1.2	2,843	2,369	57.1
Demand response[b]	70.1	1,702	1,821	1.1	15,687	14,660	26.7
Buses	[b]	[b]	[b]	[b]	[b]	[b]	217.9
Transit	72.3	2,543	19,559	7.7	35,075	4,560	89.2
Intercity[d]	[b]	[b]	[b]	[b]	[b]	[b]	37.3
School[d]	708.8	[b]	[b]	[b]	[b]	[b]	91.4
Air	[b]	[b]	[b]	[b]	[b]	[b]	1,965.4
Certificated route[e]	[b]	6,092	722,935	118.7	277,822	2,341	1,692.6
General aviation	215.4	[b]	[b]	[b]	[b]	[b]	272.8
Recreational boats	12,568.5	[b]	[b]	[b]	[b]	[b]	212.1
Rail	20.7	1,475	38,449	26.1	31,639	1,214	46.7
Intercity (Amtrak)	0.4	273	6,363	23.3	36,708	1,535	10.0[f]
Transit	13.0	826	19,452	23.5	19,871	844	16.4
Commuter	7.2	377	12,634	33.6	53,794	1,603	20.3

Figure 14.2: Statistics on usage and efficiency of different forms of transportation in the US as of 2018, from the 2021 *TEDB* Table 2.13. One important caveat is that the energy usage from electricity counts only the electricity, not the primary energy usage—this matters primarily for the rail categories.

Table 14.2 from the *TEDB* highlights several important facts about US vehicle usage as of 2018. From this table and others in the *TEDB*, we learn that:

- There are about 260 million cars and trucks in use the US, *roughly 0.84 vehicles per person, or 1.2 per licensed driver!* The US leads the world in vehicles per person, with Canada and western Europe at 0.6–0.65, and China and India at 0.04 and 0.17, respectively, in 2018.
- US vehicle ownership has stablized over the past ≈ 20 years after growing continuously since the invention of the automobile.
- Only 8.5% of US households do *not* own a vehicle. 22% have 3 or

more.

- The number of *vehicle-miles* grew steadily over time until stabilizing at about 3 *trillion* from 2004–2014. In 2018 it was 8% higher. slightly since its 2007 peak. This is roughly *10,000 vehicle-miles per person per year,* over 90% of which are in light vehicles.
- The average trip to work is 12 miles (20 km), and journeys to work account for 30% of all miles driven. The average driver takes 3 trips per day, with an average length under 10 miles (16 km) each.
- 76% of workers drive to work alone. 5% take public transport, and only 3.4% walk or bike. These percentage have changed very little in the past 20 years (but the number of people working at home has gone from 3% in 1990 to 4.9% in 2018.)
- About 16% of total expenditures by US households is for transportation (only 1.3% for public transport). This fraction has held fairly steady for 30 years, and holds across the income range. This is more than for food (14%) or health care (8%); only housing (33%) is higher than transportation.
- 3.4% of household expenditures are for buying gasoline in 2018, when gas prices were fairly low.
- Fuel costs represent only 13% of the total cost per mile of owning and operating a car in 2019. The highest this has been since 1985 is 18% in 2012–2013, before the oil fracking era and gas price collapse.
- The average car (light truck) sold in 2019 obtained 29.9 (22.3) miles per gallon, for an average of 25.5 miles per gallon. These numbers began to rise in 2007 after having been essentially constant since 1985.
- The average vehicle on the road is 12 years old, so we should even if every new vehicle sold were electric, it would take a decade for most cars on the road to be electric.

Travel by private car is, therefore, a huge part of US life and its economy, and therefore energy use and CO_2 emissions. What are the prospects for reductions in vehicle-miles? The real-estate and development industries are even more powerful than oil companies in US politics, especially at the local level, and for 50 years after WWII the emphasis was on expanding into lower-density housing and total reliance on automobiles. While the past 10–20 years have seen a move back toward high-density living, and we have seen car ownership and usage stabilize, we cannot very quickly "unbuild" the last century's development. Funding for increased capacity of public transport in the US continues to be miniscule in comparison to that for the automobile infrastructure. So I do not expect dramatic near-term changes in the first factor of our transportation energy equation in the US. The long-term trend will depend on whether there is any change in the current political hostility toward public transport at the Federal level of US government. Probably the only way to move the needle would be substantial increases in the cost of oil, as we have already seen this have some effect on consumer decisions on housing density and vehicle purchases. In the near term, we should not be looking for substantial reductions in vehicle-miles as a route to drastically lower oil consumption and CO_2 emission. And to make long-term changes, we need to design and build cities, towns, and suburbs to be better accessed by public transit.

How much power and energy does a vehicle need?

At first glance it would seem like it should take *zero* energy to transport a vehicle. At the start and end of the trip, the car is at rest and there is zero kinetic energy. For any round trip, the height at the start and end are the same, so both the KE and the gravitational PE of the car are unchanged, suggesting that no net energy was added to the car.

The first answer is that we know that energy was put into the car along the road since the KE and height did change along the trip. This means the engine did need to produce power for the driveshaft. There are four primary needs for driveshaft power while driving.

First, there is ***acceleration,*** which means change of speed. Suppose the car speeds up from speed v_1 to v_2 during a time Δt. The increase ΔE in KE requires a power

$$P_a = \frac{\Delta E}{\Delta t} = \frac{\frac{1}{2}mv_2^2 - \frac{1}{2}mv_1^2}{\Delta t} \tag{14.2}$$

$$= m \times \frac{v_1 + v_2}{2} \times \frac{v_2 - v_1}{\Delta t} \tag{14.3}$$

$$= mva. \tag{14.4}$$

The acceleration $a = (v_2 - v_1)/\Delta t$ is the rate at which we are increasing speed (or decreasing, if negative). Consider for example a car with mass $m = 1500\,\text{kg}$ which we want to accelerate from a side-road speed of 25 miles per hour ($11\,\text{m s}^{-1}$) to 65 mph ($29\,\text{m s}^{-1}$) within 15 seconds to merge onto the highway. The power needed is

For calculus afficionados, $a = dv/dt$.

$$P_a = 1500\,\text{kg} \times 20\,\text{m s}^{-1} \times \frac{(29 - 11)\,\text{m s}^{-1}}{15\,\text{s}} = 36\,\text{kW}. \tag{14.5}$$

If your car's engine cannot deliver 36 kW of power to the driveshaft, you will not be able to accelerate to highway speed fast enough, and you will be unhappy and unsafe. In the US, engine power is given in *horsepower,* $1\,\text{hp} = 746\,\text{W}$, and for this example you want 48 hp or more.

Second we need power for ***climbing.*** The power needed to supply the gravitational potential energy to get from height h_1 to height h_2 while traversing distance d in time Δt is

$$P_{\text{climb}} = \frac{\Delta E}{\Delta t} = mg \times \frac{d}{\Delta t} \times \frac{h_2 - h_1}{d} \tag{14.6}$$

$$= mgv \times (\text{grade}). \tag{14.7}$$

Here the *grade* of the slope is the height gain (or loss, if negative) per distance driven. As a numerical example, consider driving your 1500 kg vehicle up the Grapevine Hill on Interstate 5 just north of Los Angeles. If you want your car to maintain 65 mph ($29\,\text{m s}^{-1}$) while ascending this 6% grade, your engine will need to give the driveshaft this much power:

For trigonometry fans, this is just the sine of the angle of the road to horizontal.

The steepest points on the US Interstate highway system have 8% grade.

$$P_{\text{climb}} = 1500\,\text{kg} \times 9.8\,\text{m s}^{-2} \times 29\,\text{m s}^{-1} \times 0.06 = 25.6\,\text{kW} = 34.2\,\text{hp}. \tag{14.8}$$

If you want to *accelerate* while driving uphill, you need to add this power to the P_a from (14.4).

Third, we need power to overcome ***air resistance*** or ***drag.*** The car must push air out of the way to move forward. In the process it acts like a reverse wind turbine in that it *gives* kinetic energy to air that was nearly still. The power formula therefore resembles our wind turbine power formula in Equation 10.5:

$$P_{\text{drag}} = \frac{1}{2} C_D A \rho_{\text{air}} v^3. \tag{14.9}$$

In this equation, A is the *frontal area* of the vehicle: if you look at it from the front, how many square meters of area does it take up? The density of air $\rho_{\text{air}} = 1.226\,\text{kg m}^{-3}$ reappears. The ***coefficient of drag*** C_D describes how aerodynamically smooth the shape of the vehicle is, i.e. how much does it disturb the air as it passes? A block-shaped car has $C_D = 1$, but it is theoretically possible to attain $C_D = 0.03$ with a teardrop-shaped vehicle. Efficient cars (*e.g.* Toyota Prius, Nissan Leaf, Chevy Volt) attain C_D values of 0.25–0.30. Let's say you're driving your new Lexus Q50 sedan, with $C_D = 0.26$ and $A \approx 2.5\,\text{m}^2$, at our chosen highway speed. The power needed to overcome drag is

The product $C_D A$ is called the *drag area.*

$$P_{\text{drag}} = \frac{1}{2} \times 0.26 \times 2.5\,\text{m}^2 \times 1.223\,\text{kg m}^{-2} \times \left(29\,\text{m s}^{-1}\right)^3 = 9.7\,\text{kW} = 13\,\text{hp}. \tag{14.10}$$

Notice that the drag power has v^3 on it, so it grows quickly with speed.

The final main power sink for vehicles is ***rolling resistance.*** The primary source of rolling resistance is the continual squeezing and un-squeezing of the rubber in the tires as they roll along the ground. The power being sapped by rolling resistance is often written

$$P_{\mathrm{roll}} = C_r m g v, \tag{14.11}$$

where C_r is the ***coefficient of rolling resistance.*** Tires that are stiffer, inflated to higher pressure, have smaller contact area on the ground, and less distance between the wheel rim and ground, have lower C_r, but these also are poorer at absorbing small road bumps and holding traction on the road, so there is a tradeoff to be made. There are smaller contributions from friction in the wheel bearings. A good automobile tire has rolling resistance below 0.01. Going back to our standard highway-travel example, the power required for a car with such a rolling resistance is

Bicyclists will know that fat squishy mountain-bike tires make it harder to pedal than thin high-pressure racing-bike tires, but it would be suicidal to take racing-bike tires down a bumpy trail.

$$P_{\mathrm{roll}} = C_r m g v = 0.01 \times 1500\,\mathrm{kg} \times 9.8\,\mathrm{m\ s^{-2}} \times 29\,\mathrm{m\ s^{-1}} = 4.3\,\mathrm{kW} = 5.7\,\mathrm{hp}. \tag{14.12}$$

The total power that must be delivered to the driveshaft by the engine at any given moment is the sum of $P_a + P_{\mathrm{climb}} + P_{\mathrm{drag}} + P_{\mathrm{roll}}$ in the current circumstances. The car designer must make sure the engine has a *maximum power* (its capacity) that is big enough to handle the peak total of these 4 expected during its lifetime—typically trying to accelerate uphill at highway speeds, when you're getting hit with all 4. Our example calculations demonstrate these important patterns:

- Drag and rolling resistance powers are lower than the acceleration and hill-climbing power demands. But you *always* have the drag and rolling resistance, whereas you only *sometimes* are climbing a hill or accelerating. So most of the time, when the car is either cruising on a level highway or moving through city streets, the power needed from the engine is several times lower than its peak power capacity.

- Acceleration and climbing power can be *negative* if you are decelerating or going downhill! What this means of course is that you can get back the KE or gravitational PE stored in your car. Typically we *do not* because we slow down using our brakes, which put the car's KE into heat. But a car that has ***regenerative braking*** can save this energy to use later, making the net power from acceleration and climbing equal zero (as we initially guessed).

- Power for drag and rolling resistance are, on the other hand, *lost immediately* from the car, turning into heat either in the air, or in the tires. There is no way to get these back, hence they set the floor on how much energy it takes to move the car.

- The drag has v^3 in it, but rolling resistance has v. So as speed increases, the drag will become dominant to rolling resistance. In our examples, at highway speeds, this is the case. At low speeds, however, rolling resistance uses more power. All of the power types have v in them, so the faster you drive, the more power you need from your engine.

Driving faster takes more Joules per second, but also gets us home faster. We're interested in the energy used per unit *distance* travelled. We use the symbol F for energy per distance.

Because this is also equal to the force, in physics lingo.

$$F = \frac{E}{d} = \frac{E}{t} \times \frac{t}{d} = P \times \frac{1}{v}. \tag{14.13}$$

Notice that we are using the energy per unit distance, whereas cars are usually compared with miles per gallon (in the US), which is effectively distance per unit energy, the inverse of our F. We can now take our 4 types of transportation energy needs and convert them into energy needed per distance:

$$\begin{aligned} F_a &= ma & &\text{(acceleration)} \\ F_{\text{climb}} &= mg \times (\text{grade}) & &\text{(climbing)} \\ F_{\text{drag}} &= \frac{1}{2} C_D A \rho_{\text{air}} v^2 & &\text{(air resistance)} \\ F_{\text{roll}} &= C_r mg & &\text{(rolling resistance)} \end{aligned} \tag{14.14}$$

The strategies for reducing the energy needed per vehicle mile are clear:

- Reduce the *mass* and *size* of the vehicle, since all energy needs grow with m or with A.

- Use regenerative braking if possible, otherwise avoid hills and unnecessary acceleration/deceleration.

- Drive more slowly, and build as low C_D shape as possible, to reduce energy lost to air drag. This is the only energy-per-mile term that depends on the speed of your trip!

Rolling resistance actually increases slightly with speed, but it's small enough for us to ignore.

- Use tires with low C_r to minimize rolling resistance.

You might contrast the designs of the Toyota Prius vs the Hummer H2 (no longer manufactured) as examples of vehicles whose purpose is to minimize and maximize consumption, respectively, and see how they fill out Equations (14.15).

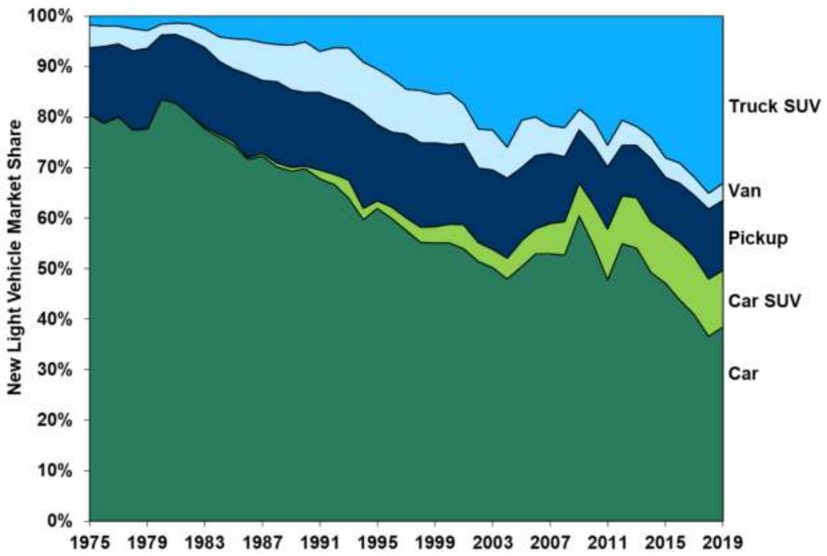


Figure 14.3: Fraction of light vehicles sold in different size/type categories each year since 1975. From *TEDB* Figure 4.4.

Since it would be difficult to make C_D and C_r much lower than they are now, it is clear that the main way we could reduce the energy required per vehicle-mile would be to reduce the size and mass of the vehicles, and implement regenerative braking as widely as possible. Unfortunately the trend over the past 40 years has been for US cars to get larger. Figure 14.3 shows the steady trend from 1975–2019 for US consumers to move away from "cars" to larger pickups, vans, and SUVs. This was partly a way to avoid the fuel economy standards enacted in the 1970's, which set weaker standards for "light trucks" than for cars. Manufacturers reclassified their larger, higher-profit-margin mass-market passenger vehicles as "light trucks" so they could meet weaker fuel economy standards. Thus was born the SUV era. From 2003–2012 the SUV fraction did not change much, but with the return of consistently low gas prices in the fracking era, the car fraction has continued to drop. In fact US auto makers have almost completely abandoned the market for actual cars. Figures 14.4 illustrate the way in which automotive technology advances have been channeled as oil prices have varied since 1975. Compare this to Figure 7.11 and note the faithful response of vehicle mileage to oil prices in the past 40 years!

Internal combustion engines

The on-board stored energy needed for a trip is the crankshaft energy divided by the engine efficiency. Essentially all cars use one (or both!) of two kinds of engines: internal combustion engines (ICE), or elec-

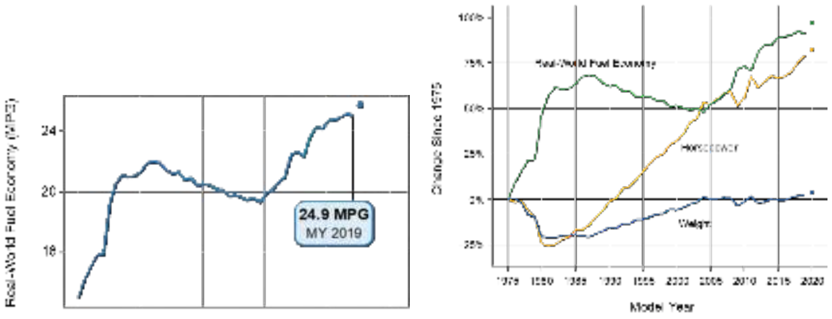


Figure 14.4: At left is the average gas mileage of the US light vehicle fleet vs year for 1975–2019. At right are plotted the changes in mean vehicle mileage, weight, and horsepower. After the oil crises, car sizes and engine powers dropped quickly and resulted in rapid fuel economy increases. Through 1985–2005, engine technology was improving, but oil was cheap and the fuel economy dropped as heavier and more powerful cars were sold. Since 2005, weight has stabilized and technological improvements have yielded higher mileage, and engine power continues to rise. From the Environmental Protection Agency *2020 EPA Automotive Trends Report*

tric motors. Before we compare them, it's good to know how an ICE works. Figure 14.5 illustrates the 4-stroke Otto-cycle (gasoline) engine that powers the great majority of light vehicles. The ***piston*** travels inside a ***cylinder*** with an airtight seal made by ***rings*** around the piston. The ***connecting rods*** cause the the ***crankshaft*** (which points out of the page) to rotate as the piston goes up and down. The cycle starts when the rotating crankshaft pulls the cylinder downward in the ***intake*** stroke. A mixture of gasoline vapor and air is sucked into the piston through the open intake valve at top right. Next comes the upward ***compression*** stroke, with the valves closed, which compresses the flammable mixture into smaller volume, raising its pressure and temperature. When the piston reaches the top of the compression stroke, the ***spark plug*** ignites the fuel-air mixture, which releases heat as it burns inside the cylinder (hence "internal combustion") and greatly increases the temperature and pressure in the cylinder. The piston gets pushed down very hard on the ensuing ***power stroke.*** Next the left-hand valve opens, and the upward piston motion expels the combustion products in the ***exhaust stroke.***

The ***compression ratio*** is how much smaller the cylinder volume is with the piston at top than at bottom. A bigger compression ratio means higher pressure and temperature at the start of the power stroke, which means higher efficiency (as per Carnot). A problem, however, is that if the compression ratio is too high, the fuel-air ratio will ignite without the spark plug, before the compression stroke begins, pushing the driveshaft the wrong direction. This is called *knocking* and it's very bad. An engine running on standard gasoline cannot have a compression ratio above about 10:1.

The ***Diesel*** cycle differs slightly in that pure air is induced during the intake cycle. The fuel is injected into the cylinder at the start of the power cycle so premature combustion cannot occur. Much higher compression ratios (up to 22:1) are possible. The compressed air is hot enough to ignite the fuel—no spark plug is needed. Diesel

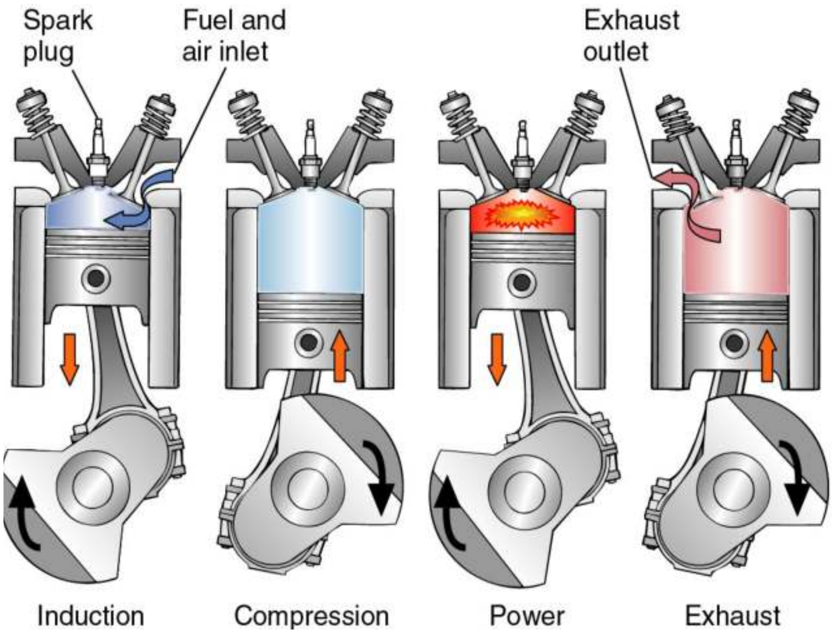


Figure 14.5: The 4-stroke Otto cycle engine. From Prof. Nick Collings.

engines are hence more efficient, and more robust to variation in the composition of the fuel.

ICE's deliver $> 30\%$ efficiency, up to 40% for diesels, but in practice the typical efficiency is more like 20–25%. Figure 14.6 shows an efficiency plot for an example ICE. This engine obtains its optimum efficiency of 34% at a power output of 25 kW (near its maximum) and a certain engine speed. If we only need 5 kW of power for cruising then the engine will have lower efficiency. The more *maximum* power we pack into a car's ICE—so that it has zippy acceleration—the less efficiency it will have at the much lower power levels encountered in *normal* operation.

Also a more powerful engine generally means more weight.

Electric vehicles and hybrids

Vehicles can obtain their driveshaft power from electric motors instead of ICEs. We know that electric motors are highly efficient, typically 90–95%. Thus *a car requires 3–4× less energy input to an electric motor than it would require in fuel input to an ICE.* Electric motors offer other advantages over ICE's as the propulsion for cars:

- Electric propulsion enables *regenerative braking.* A motor can convert electric energy to mechanical energy on the driveshaft. But we can also use the motor as a generator, to convert the kinetic energy of the car into electric energy. This will slow down (brake) the vehicle. If this electric energy is stored in a battery, it can then

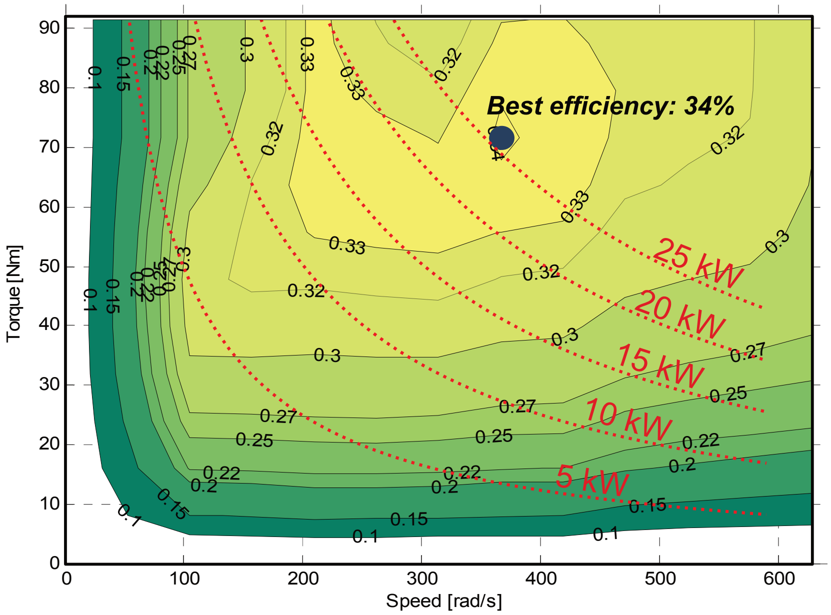


Figure 14.6: ICE efficiency diagram. The contours/colors give the efficiency of the engine, which depends on the engine's speed (x axis) and the power we want from it (red dashed lines). Its highest efficiency of 34% is obtained only when providing 25 kW of power. At lower power and speed the efficiency can drop quickly. Adapted from Roy et al (2015).

be re-used to accelerate the vehicle later. Thus *a vehicle with electric propulsion can recover most of the energy that goes into acceleration and hill-climbing.* This is a particularly large benefit in city driving, where one is frequently stopping and speeding back up.

- Electric motors have high efficiency at all speeds and power levels. They do not have the ICE's difficulty in providing both high peak power, and high efficiency at normal operating power. In fact they usually don't need a transmission, and are generally much simpler, more reliable, and lower-maintainence than ICE's.
- Electric motors do not need to ***idle***, that is keep running while the car is stopped.
- Electric motors emit no CO_2 and no other pollutants. Of course such pollutants might be emitted at the power plant where the electricity was produced, but the primary power could be nuclear, wind, solar, or hydro with no emissions, decarbonizing the transport sector.

We will return later to the issue of how to provide the electric energy; we just note now that storing energy in a battery requires much more weight and expense than storing energy in gasoline. As a consequence, electric vehicles have a much shorter ***range*** that they can drive before running out of stored energy. These issues are summarized in Table 14.1. The Nissan Leaf, Chevy Bolt, and the Tesla vehicles are the best-known electric vehicles (EVs). They obtain all of their energy from the power grid. When the batteries are drained, the car is stuck.

Characteristic	ICE	Electric	Hybrid	Plug-in Hybrid
Typical efficiency	20–25%	90–95%	30% ?	90–95%
Regenerative braking?	No	Yes	Yes	Yes
Efficiency @ low power?	Poor	Good	Good	Good
Primary energy source	Gasoline	Grid	Gasoline	Grid
Range	Long	Short	Long	Long
Main pollution site	Onboard	Power plant	Onboard	Power plant

Table 14.1: Characteristics of different vehicle engine systems.

Surprisingly it can be highly cost-effective and efficient to have *both* an electric and ICE engine in the same car! This is called a ***hybrid*** electric vehicle (HEV). Figure 14.7 illustrates two possible ways of doing this. At left is the ***series*** hybrid design: an ICE engine is used strictly to run a generator. The electricity stored is stored in a battery, and then is withdrawn from the battery to run an electric motor which powers the driveshaft. This seems like a waste, but in fact can be very efficient, because it can offer the best of both the ICE and electric worlds:

- Regenerative braking will work, saving all of the acceleration and climbing energy.
- The ICE can run at its most efficient power output *all the time* and be considerably smaller than needed in an ICE-only vehicle. It does not need to produce the peak power: the electric motor can provide more power than the ICE's capacity for short periods by discharging the battery faster than the ICE is charging it. The battery needs only to be large enough to store the excess energy needed in these short periods. The ICE needs only the maximum power level that will be sustained for long periods. This makes the ICE smaller, lighter, and more efficient.
- We get the robust high-power performance characteristic of electric motors.
- **The battery is not the energy storage medium nor the primary energy source.** All the energy that enters the car is in the form of gasoline, the battery is just a temporary holding tank for energy. Hence we get the cheap, high-capacity gasoline tank instead of heavy battery storage, and avoid the range problem!

At right we see the ***parallel hybrid,*** thus called because energy can flow from both the ICE and electric motor to the driveshaft at the

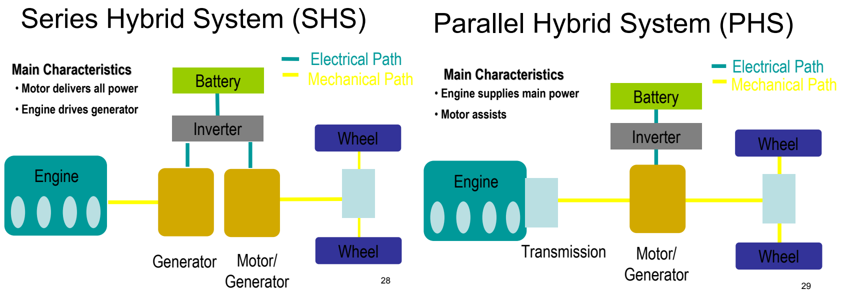


Figure 14.7: Schematic diagrams of series and parallel hybrid ICE-electric power systems. From Prof. Nick Collings.

same time, instead of having to flow serially from ICE to electric. In this setup, both engines are connected to the driveshaft. When peak power is needed, the ICE's power is added to electric power being extracted from the battery to run the car. At times of lower power demand, the ICE power is partly running the car and partly recharging the battery. At very low power needs, the ICE is shut off and the car runs on stored battery electricity. The advantages are the same as for a series hybrid, there are just different engineering tradeoffs here.

The Toyota Prius uses a more complicated hybrid system than either of these diagrams, involving two motor-generators.

Let me reiterate that the basic hybrid *does not run on electricity from the grid.* All energy that enters the car is in the form of gasoline. It is just a way to make more efficient use of this gasoline, because the ICE runs at higher efficiency, and we can use regenerative braking. This more than makes up for the extra complexity and weight of having two motors (plus the ICE can be lighter as its peak power output is lower). The Toyota Prius was the first such vehicle to achieve sales volume comparable to ICE models. The Prius's high gasoline mileage is attributable in part to its hybrid power train, but also to careful attention to vehicle mass, $C_D A$, and C_r.

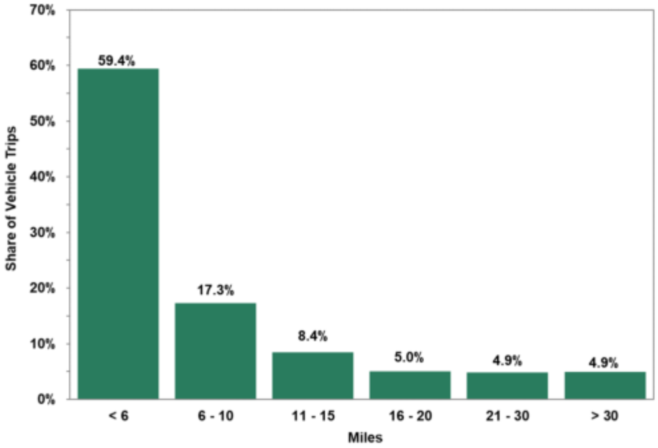


Figure 14.8: The distribution of lengths of US automobile trips. From the *TEDB* Figure 9.3.

Finally there is the ***plug-in hybrid*** vehicle (PHEV). These are just like the hybrids discussed, except that the batteries are enlarged to store enough power to run in all-electric mode for some modest range. Then we add a plug so that these batteries can be recharged with grid electricity. As long as the trip is shorter than the battery range, the car behaves like an EV, and it obtains all its energy from the grid. But for longer trips, the ICE turns on and it continues operation like an HEV. Thus a tank of gasoline (or biofuel) gives you 100's of km of range for your longer trips, plus the rapid refueling we want on these trips. Yet the battery capacity does not need to be large enough to have its cost and weight dominate the vehicle's budget. In 2020 there were 40–50 different PHEV models available for purchase.

Figure 14.8 suggests that $> 90\%$ of people's car trips could be completed without ever using fossil fuels, for a PHEV with just 20-mile battery capacity. But they would not get stuck if they decided to take a longer trip.

Energy storage and primary sources

Except for electrified rail, all modes of transportation must carry their energy with them. We must store enough energy onboard to get us to the next refueling station. Let's make a rough estimate of the amount of energy we would want to store onboard a small passenger vehicle for a highway trip. In the best case, we have perfect regenerative braking and have no net energy expense for acceleration and hill climbing. We need to budget energy for drag and rolling resistance. Let's make an estimate for the 2019 Chevy Bolt, an all-electric compact car that costs about \$37,000 (minus Federal subsidy up to \$7500 available). The Bolt's ***curb weight*** is about 1650 kg; this does not count the passengers, so let's add 150 kg for two adults, $m = 1800\,\mathrm{kg}$. The Bolt has $C_D = 0.31$ and $C_D A \approx 0.9\,\mathrm{m}^2$. If we assume a rolling resistance of $C_r = 0.01$, the driveshaft energy needed at highway speed $v = 29\,\mathrm{m\,s}^{-1}$ is

This is not a great drag coefficient. The Tesla Model 3 claims $C_D = 0.24$.

$$F = F_{\mathrm{drag}} + F_{\mathrm{roll}} = \frac{1}{2} C_D A \rho_{\mathrm{air}} v^2 + C_r m g \tag{14.15}$$

$$= \frac{0.9\,\mathrm{m}^2}{2} \times 1.23\,\mathrm{kg\,m}^{-3} \times \left(29\,\mathrm{m\,s}^{-1}\right)^2 + 1800\,\mathrm{kg} \times 9.8\,\mathrm{m\,s}^{-2} \times 0.01 \tag{14.16}$$

$$= 641\,\mathrm{J\,m}^{-1} = 0.64\,\mathrm{MJ\,km}^{-1}. \tag{14.17}$$

If we guess an efficiency of 90% for the electric motor, we need to store 0.71 MJ of energy for each km of ***range*** that we want the Bolt to have between recharges. The Bolt's EPA highway efficiency rating is $0.31\,\mathrm{kWh\,mile}^{-1} = 0.68\,\mathrm{MJ\,km}^{-1}$ for highway energy use—very close to our estimate, considering that we were optimistic about perfect regenerative braking, we ignored energy needed to run the car's accessories, and our constant-speed calculation is simpler than the EPA's test drive protocol.

What is the Bolt's range? Chevy claims that the Bolt has a 60 kWh (216 MJ) battery, so the above highway efficiency we would expect a range of $216\,\mathrm{MJ}/0.71\,\mathrm{MJ\ km}^{-1} = 304\,\mathrm{km} = 189\,\mathrm{mi}$. In fact the EPA rates the Bolt at 248 mile range (I don't understand the source of the difference). Getting a consumer-priced EV over 200 mile range is a signficant achievement; in 2020, 7 car models, plus all the Tesla models, offer this. The Tesla Model 3 and Model Y are now by far the best-selling EVs. A typical ICE car has ≈500 km range on a single tank of gasoline. For an EV like the Bolt to attain this and fully alleviate "range anxiety" for customers, we see that it would need to store something like 350 MJ or about 100 kWh.

The best-selling electric vehicle is the Tesla Model 3—95,000 were sold in the US in 2020 and 137,000 in China. The "Standard Range Plus" 2021 model has a 54 kWh (195 MJ) battery and costs $38,500, with an EPA-estimated range of 263 miles. The "Long Range AWD" model comes with an 82 kWh (295 MJ) battery, and an EPA-estimated range of 353 miles, and a cost of $47,490. The EPA also estimates that these vehicles consume about 25 kWh per 100 miles of travel, averaging over city and highway usage. At face value, the extra $9,000 for 28 kWh of additional battery works out to a cost of $320/kWh for battery storage, although we cannot assume that this price differential accurately reflects battery costs. It's also of interest the the long-range battery has mass of 478 kg, which is about 1/4 of the total vehicle weight.

I know, the arithmetic of the EPA estimates does not work out to the right battery capacity. I don't know what special version of multiplication is involved in EPA regulations.

Info from `https://evspecifications.com`.

So we see that we should store $\approx$ 300 MJ on board an EV, and preferably more for a larger car and larger range. ICE vehicles need 3–4× more energy stored because the engine efficiency is 20–25% instead of the $>$ 90% we expect for EVs. Recalling that 1 gallon of gasoline holds 130 MJ of energy, the $\approx$ 40 miles per gallon we expect to get on the highway from an efficient small car converts to 1.9 MJ per km, which is indeed 3× more than the Bolt. But this higher energy need is no problem for an ICE car, because gasoline is an incredibly compact way to store energy: just 8 gallons of gas holds enough energy for a 500 km range. No one has ever complained that the gas tank on their car costs too much or takes up too much space!

They only complain that it costs too much to fill up.

Our vehicle's energy-storage mechanism must be adequate on several criteria, which we list below. We are also extremely interested in energy storage for grid electricity, which would allow renewable solar and wind power to better match fluctuating electric demand. We'll note below which aspects of energy storage are important for

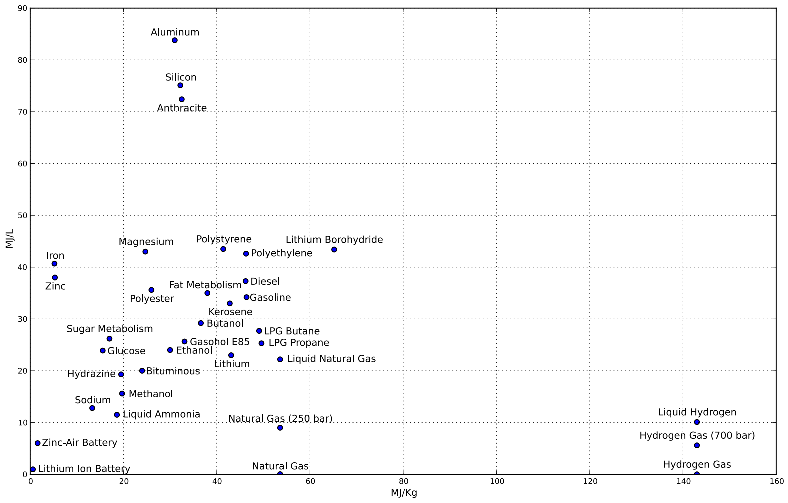


Figure 14.9: Energy density of many chemical energy storage substances. These values do not include the weight or volume of the storage containers. Gasoline is in the center-left of the diagram. Batteries are at the bottom left corner, the worst place! From Wikipedia.

this application too.

- **Energy density per unit mass:** The storage "tank" should be a small fraction of the mass of the car. A heavy storage device means the car's mass is higher, which means it needs more energy to move, which means we need a bigger storage device.... You can see this is a bad cycle to be in. Figure 14.9 gives the energy density of many chemical storage options. For grid storage, mass and volume are not important.

- **Energy density per unit volume:** The storage device cannot be too large either, else the car needs to get large, making it have higher drag as well as being unpalatable to the market. For aircraft, low mass and volume are hugely important—airplanes already have a good fraction of their mass in fuel, and they simply could not fly if the mass of the fuel were much larger than for jet fuel.

- **Efficiency:** We need the storage device to release almost all the energy we put into it.

- **Discharge rate (power):** When we're accelerating or climbing a hill, we need to get the energy out of storage fast enough to provide the peak power needed by the vehicle. Grid storage typically has less need for short high-power bursts over seconds or minutes.

- **Recharge rate:** We do not want to spend too much time recharging the storage device before we can get back on the road. We are accustomed to spending only a few minutes at the gas station to refill after 5–6 hours of driving. Not so important for grid storage.

- **Supply chain:** How feasible/expensive will it be to build and supply a network of recharge stations across the country?
- **Primary source:** What source(s) of primary energy can be used to recharge the storage, and at what cost in dollars, CO_2 emission, and energy security?
- **Pollutants:** What undesirable exhaust products are released from the vehicle?
- **Safety:** Will the stored energy be released catastrophically when the vehicle crashes? For grid storage, we don't want to have explosions at locations storing many hundreds of GJ.
- **Lifetime:** Can the storage be refilled the hundreds or thousands of times that will be required to last the lifetime of the vehicle?
- **Cost:** Of course. Buying the storage should not ruin the economics of owning the vehicle.

That's a lot to think about! We'll concentrate on the issues of energy density and cost. Table 14.2 summarizes the state of the art for the three energy storage methods predominantly considered for our transportation fleet: gasoline, batteries, and hydrogen. All three of these are storing chemical energy, but have very different characteristics.

Storage	Efficiency	MJ/kg	MJ/liter	$/MJ	Emissions	Fill time
Gasoline	≈25%	47	35	near 0	onboard	minutes
Li-ion battery	90%	0.9 ?	1 ?	30–50 ?	plant	hours
Hydrogen, compressed 700 bar	50% ??	5	3	??	plant	minutes

Table 14.2: Characteristics of three transportation energy storage technologies. A typical efficiency of conversion of the stored energy into driveshaft energy is given assuming ICE for gasoline, electric motor for battery, and fuel-cell electric for hydrogen. Many values are evolving with time and/or highly uncertain. We note whether primary emissions occur onboard the vehicle vs at a centralized power plant.

Hydrocarbons

Gasoline and other petroleum products like diesel and jet fuel are truly fantastic energy-storage media. The energy density is so large that it is simply not an issue for ground transport, even though the heat engines that use the fuel are inefficient so we need to store triple the driveshaft energy. Distributing all this energy throughout the transportation network, and avoiding catastrophic release (explosions and fires) is not necessarily easier than for other fuels, but the infrastructure is *already in place,* which makes it cheap (in the near term).

Of course petroleum derivatives have the major drawbacks of poor long-term security and high CO_2 emissions. These could in principle be overcome with a switch to liquid biofuels. While ethanol has only 2/3 the energy density of gasoline, this is still more than adequate for use in transportation.

Another gasoline alternative is the use of natural gas in vehicles. While NG has even more energy per kg than gasoline, it is of course a gas instead of a liquid and has far lower energy per liter. Feasible NG-fueled vehicles must compress the fuel by storing it in high-pressure tanks—this yields lower energy per liter than gasoline, but high enough to store energy for shorter-range usage. NG use would cut CO_2 emissions below gasoline's, though still far from zero-carbon. NG is in use for city buses, which travel limited range, can always be refueled at specialized depots, have plenty of room for the NG tanks. But since global NG reserves are not larger than petroleum's, and NG is not a sustainable fuel, I will not further consider the possible of building a large NG vehicle infrastructure.

Hydrogen

Way out on the right-hand side of FigureJ 14.9 sits the very tempting target of hydrogen (H_2), with an energy density at 143 MJ per kg, more than 3 times that of gasoline. This is not surprising, since hydrogen is the lightest element and should pack the biggest chemical-energy punch per kg. Another benefit of hydrogen as a vehicle fuel is that its combustion yields just H_2O, no CO_2 at all. We could consider using hydrogen as a fuel for ICE's, but its true benefit would come from the use of ***fuel cells.*** A fuel cell is basically a battery that we continuously feed new chemical energy, instead of "charging" it by putting electricity in to create that chemical energy. It accomplishes the same purposes of diverting the path electrons in a chemical reaction through a circuit.

Fuel cells are in principle a superior way of using chemical storage in a vehicle compared to burning the fuel and using the released energy to run a heat engine. A fuel cell will not be subject to the Carnot efficiency limit and will not require high temperatures to work well. It avoids the pollutants that are formed by high-temperature combustion. By producing electricity, we can use electric motors for propulsion, which means that regenerative braking is easily incorporated. These properties were attractive enough that the (W.) Bush administration made development of hydrogen-powered cars the centerpiece of Federal research funding for energy-efficient vehicles, calling it the *Freedom Car* initiative. There are however, some formidable obstacles to decarbonizing the transportation fleet with hydrogen.

The Bush administration was very fond of naming things after freedom.

- **There is no such thing as a hydrogen well!** Hydrogen is not a primary source of energy. Hydrogen is used in many industrial

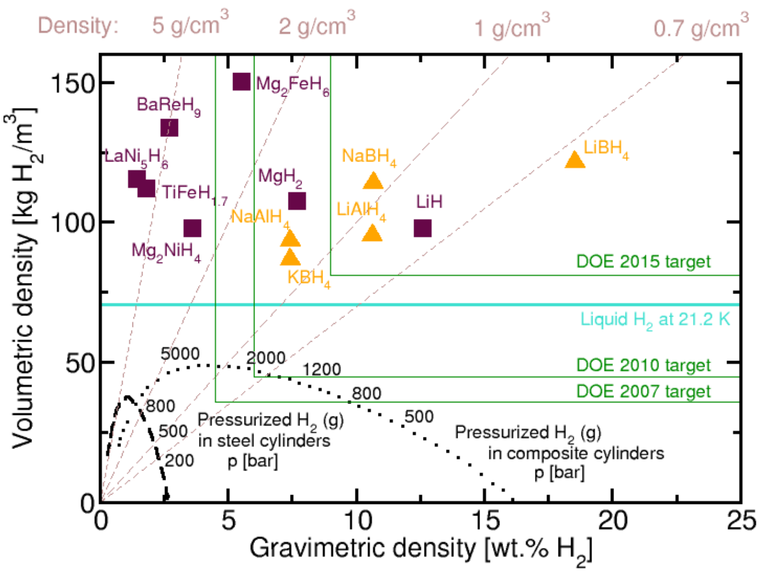


Figure 14.10: Energy density of hydrogen in various possible forms of storage. The x axis gives the fraction of the total weight that is actually hydrogen; multiply this by the 143 MJ/kg energy content of hydrogen to get the net energy density of the storage. Likewise multiply the y axis by this value to get MJ per cubic meter. An indication of the slow progress is that the Department of Energy goals for 2020 were the same as the "2007" targets in this old plot from Wikipedia!

processes already, but most is produced by "reforming" natural gas. Running cars with NG-derived hydrogen would just be a less-efficient way of using NG to fuel cars, so this is not a path to sustainability. Hydrogen is of course plentiful on Earth in the form of water, which is of course "already-burned" hydrogen. Every Joule of energy that we extract from burning the hydrogen is a Joule we must have earlier used to separate the H and O in water! One method of separating water is ***electrolysis***, by use of electricity. But there are inherent inefficiencies, as we need to (1) use some primary energy source to make electricity, which involves some efficiency factor; then (2) run a electrolysis process to make hydrogen at 70–80% efficiency; and (3) use the hydrogen in a $\approx$ 60%-efficient fuel cell to make electricity again. Hydrogen cars would be a net environmental loss without a very cheap and abundant renewable source of primary energy available to produce the hydrogen. But perhaps PV or wind electricity could do this job—we'd just have to accept that only half (80% × 60%) of the electric energy produced by the PV would make it out of the fuel cell, even if the hydrogen storage method required no energy.

- Hydrogen **distribution** would require a new infrastructure to be built, apart from our existing gasoline and electric distribution systems. This should be feasible given that we already have natural-gas pipelines threading the nation, but it's clearly a large undertaking.

- Hydrogen is a *low-density gas* and we need to compress it somehow to obtain a feasible MJ-per-liter performance. Figure 14.10 shows the resultant energy densities for different methods, and we see that *every feasible mode of hydrogen storage has much more mass in the container than in the hydrogen!* The simplest storage for hydrogen, and the one being used in the vehicles being offered for sale/lease in 2021, is a simple tank of gas compressed to 700× atmospheric pressure. The tank to withstand this pressure outweighs the H_2 stored within it. More exotic schemes for H storage involves having the hydrogen bond to other elements and from which it can be released by relatively mild heating. These "hydride" structures would be solid and have higher density. One question is how much of the energy would be lost in the binding/releasing cycle.

- The *fuel cell* needs to be made efficient, reliable, and cheap enough for vehicle use. Fuel cells have been used in spacecraft for decades, but those are specialized and expensive. The most efficient fuel cells need very high temperatures for operation, not easily managed in a vehicle. Fuel cells can work not just for hydrogen, but also natural gas, or ethanol. If practical fuel cells could be developed, it is better to use them to produce electricity fed to motors to run a vehicle than it is to burn them in an ICE.

When Steven Chu, a Nobel Laureate in Physics, become Secretary of Energy in 2008, he famously referred to these four major unsolved problems with a hydrogen transportation system by stating "If you need four miracles, that's unlikely. Saints only need three miracles." The Freedom Car program was felt to be emphasizing a technology that would not be commercially viable for decades, and funding emphasis was shifted to batteries and other renewable-energy technologies that were closer to reality.

Hydrogen research did not stop however, in 2021 there were three fuel-cell cars available in the US: the Honda Clarity (midsize), Hyundai Nexo (SUV), and Toyota Mirai (subcompact, $52,000), with EPA-estimated ranges of 360, 380, and 402 miles, respectively. But only about 1000 were sold in the US in 2020 and according to the *TEDB* there are ≈ 60 hydrogen filling stations in the US.

The Mirai is measured by the EPA to obtain 66 highway miles per kg of hydrogen, which means 1.36 MJ/km for a subcompact. This is about 2× the energy use of the Chevy Bolt, which suggests the Mirai's fuel cell is near 50% efficient, in good agreement with data that I have seen.

The best current technology for hydrogen storage appears to be high-pressure tanks, which are still just 4% hydrogen by weight when filled (1.4 kWh/kg), and projected to cost $500 per kg H_2, or $15/kWh, if produced in high volume. Both of these are better than Li-ion batteries, even if we allow for the need to store twice as much energy in hydrogen as in electricity because of fuel-cell inefficiencies.

Meanwhile the two other problems remain: there is not yet a sustainable, high-efficiency mode of hydrogen production, and there are only 63 hydrogen filling stations in 2020. There is much work to be done before we could consider hydrogen vehicles as a sustainable alternative to petroleum, but we could think of hydrogen produced using photovoltaic electricity as a long-term possible replacement for fossil fuels in transportation scenarios requiring energy densities better than batteries, but not as high as liquid fuels.

Batteries

Batteries are devices that can convert stored chemical energy into electricity. In most chemical reactions, electrons shift from being resident on one atom to another; a battery arranges for the electron to make this shift via a circuit, so we can use the energy instead of having it turn into heat first. Rechargeable batteries can reverse the chemical reaction by putting electric current into the battery. There are many different chemical reactions that can be harnessed to construct a working battery, but it is a significant challenge to build a battery that is efficient, high-density, does not degrade over repeated discharge cycles, works safely over a wide range of temperatures, and can be produced cheaply in huge volumes.

Electric-motor vehicles have many advantages over ICE's: higher efficiency; easy implementation of regenerative braking; high power delivery at any speed; and zero emissions of CO_2 or other pollutants from the vehicles themselves, which is important in reducing urban air pollution. We already have an electricity-distribution infrastructure in place, though we would need to greatly increase grid capacity if we switched to an electric vehicle fleet, and modify service stations to handle high-power electricity instead of gasoline. Most importantly, the electricity can be generated from many non-petroleum primary energy sources, greatly increasing energy security. And, *if we generate electricity from renewable sources,* we decarbonize transportation.

But *batteries are the weak point of an electric vehicle fleet.* As seen in Table 14.2, current batteries have $\approx 35\times$ less energy density than gasoline. Even considering that we only need to store 1/3 as much energy in batteries than in gasoline, this leaves us with storage using

$\approx 12\times$ as much weight and volume as a gasoline vehicle with the same range.

Significant progress has been made in battery technology in the past 20 years, especially with the advent of lithium-ion (Li-ion) batteries as opposed to the lead-acid batteries that were available in earlier eras. Looking at the periodic table, we can see that Li is the third-lightest element whereas lead (Pb) is one of the heaviest, so we can shuffle electrons amid much less mass now—although Li-ion batteries also typically require one atom of manganese and cobalt for each lithium atom, and cobalt is a material of limited supply. Small, lightweight cellphones and laptops are made possible by Li-ion batteries, and it is no coincidence that viable all-electric cars came on the market only as Li-ion batteries became available in bulk. In 2019, the Nobel Prize in Chemistry was awarded to three of the top contributors to the development of Li-ion batteries—John Goodenough, Stanley Whittingham and Akira Yoshino.

Prices

Battery prices still remain the biggest sticking point for electric vehicles. A price of $100 per kWh of battery capacity has been considered as a "tipping point" level at which EV's can begin to seriously compete with ICE's in the consumer market. At $100/kW, and an energy density of 0.9 MJ/kg, then the 350 MJ (97 kWh) battery necessary to give the Bolt a 500 km (300 mile) range would cost $10,000 and have mass 390 kg. The battery is about 40% of the cost of a typical ICE subcompact car, and raises its mass by 25% too. On the other hand, electric motors are considerably less complex than ICE's and require much less maintenance. And of course one will save money on fuel. The average vehicle in the US is driven 12,000 miles per year. If your subcompact attains 40 miles per gallon, and gas costs $3 per gallon, you're paying $900/yr for gasoline. I leave it as an exercise to estimate the cost of electricity to run the car—but $100/kW clearly puts us in a ballpark where an EV's life-cycle costs are close to an ICE vehicle's.

Exact prices that car manufacturers are paying for battery packs are proprietary, but some information and estimates are available. A BloombergNEF battery-price survey from December 2020 reports that average Li-ion battery packs for EV's averaged $126/kWh in 2020, with some individual sales dropping below the $100/kWh mark for the first time. They note that just 10 years ago, the price for Li-ion batteries was $1100/kWh! They also estimate that "by 2023 average pack prices will be $101/kWh. It is at around this price point that automakers should be able to produce and sell mass market EVs at the same price (and with the same margin) as comparable internal combustion vehicles in some markets. This assumes no subsidies are available. . . .”

Further price reductions for Li-ion batteries seem likely, but not another factor of 10. The BloombergNEF report forecasts $58/kWh by 2030. A study from the National Renewable Energy Laboratory (NREL) of costs for utility-scale Li-ion storage also suggests 2030 costs that are 20–60% lower than today.

Adoption of EVs

It certainly seems that we are entering the era of price-competitive EV’s. Current EV sales remain a small fraction of vehicle sales in the US: about 2.5% as of early 2021 (Figure 14.11), counting plug-in hybrids. But this fraction is growing. Because cars last a decade or more, it will take a long time for most vehicles *on the road* to be electric even after most of the vehicles *being sold* are electric.

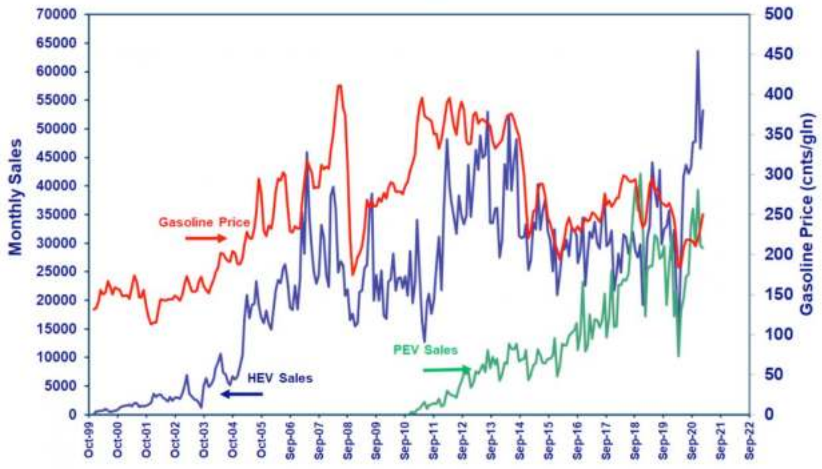


Figure 14.11: Monthly electric vehicle sales in the US, from Argonne National Labs. Here “PEV” combines battery and plug-in hybrid vehicles, and “HEV” are gas-only hybrids. For reference, total US monthly vehicle sales average around 1.4 million.

Several very large automakers have announced their intentions to switch entirely to electric-vehicle manufacture. Notably, General Motors, the largest US automaker, pledged in January 2021 to phase out ICE’s by 2035. Some European manufacturers have more aggressive plans.

A difficulty with EV’s is recharge time: if you are using a typical 20 A, 120 V household circuit (2.4 kW), it will take 40 hours to recharge a 100 kWh battery! Practical EV recharging requires installation of higher-voltage, higher-current charging stations to move more power. This isn’t a problem for daily commuting—one can recharge at night—but most people taking long-distance trips will frown upon having to stop for a 30-minute recharge every few hundred miles. And if you don't have your own garage, how do you recharge overnight? One possible solution is to make battery packs standardized so that "refueling" becomes a quick battery swap operation: you drop your discharged battery at the "gas" station, and a fully charged one is plugged into your car. The station recharges the battery and sells it later to someone else. This of course means that we need > 1 battery per vehicle on the road, raising the cost. This idea has not yet caught on.

Plug-in hybrids are a good option for this reason, since refueling is fast for the (relatively rare) longer-range trips. In the past there have been more *non-plugin* hybrids (HEV's) sold than plug-in vehicles, but it's important to remember that HEV's *get all their energy from gasoline,* not the grid. They are just a more efficient way to use gasoline energy, they do not facilitate a switch to non-petroleum primary fuels. An HEV is no more "green" than an ICE vehicle with the same mpg rating! They came onto the market first because they do not require as much battery capacity as a PHEV—they only need the capacity to provide the energy for high-power periods of acceleration and hill climbing, so the associated weight and cost are lower.

A more complete solution is to have high-power chargers be as readily available as current gasoline filling stations. Governments of course have a potentially big role to play in transitioning the vehicle fleet to electricity. There are Federal tax credits of up to $7500 available for EV/PHEV purchase to help the market and spur development of economies of scale, and many states add other subsidies—including the priviledge of driving in special express lanes on some crowded highways. But we know that the EV's themselves are only part of the infrastructure needed to make EV's dominant: the EV manufacturers know this, and Tesla in particular has been aggressive in making rapid-charging stations available to facilitate use of their cars. The Biden administration's pending US infrastructure bill reportedly includes funds to push for installation of 500,000 EV charging stations across the US by 2030. And of course if most of our vehicles are electric, we're going to need to build more generating and grid capacity.

Other countries are ahead of the US in the transition to EV's. Over 1 million EV's per year are being sold in China. In Norway, plug-in vehicles are already a majority of sales, and they're above 15% in several other Nordic countries.

Grid Storage

What are the prospects for substantial improvements to the energy density and cost of batteries? Cheap batteries matter not only for the transportation sector—we know that economical utility-scale energy storage that would enable nondispatchable renewables to provide most of our electricity (and energy density is not really an issue here). Unfortunately, advances in chemistry do not occur at an exponential pace like miniaturization of electronic circuits has.

One interesting idea is that *electric cars themselves* could help shape the load curve for electricity, because the cars have batteries that can charge at midday or other high-availability periods and use the electricity later.

Since there have been steady improvements in cost per kWh of Li-ion batteries—which is what we usually expect as production volume rises—we are just beginning to see grid battery storage emerge at GW scales, often in association with photovoltaic installations, but even in standalone configurations. The economics of various forms of grid storage are examined in a report from Lazard. They assume present-day capital costs in a range $160–300 per kWh of storage—somewhat higher than we see above for raw Li-ion battery costs, but we can expect there to be additional costs. From this they derive that a grid-storage system that is charged and discharged once per day would end up costing about $0.10–0.22 per kWh of electricity delivered to the grid (above the cost of purchasing the electricity to charge the batteries). This is not yet cheap enough to be viable for most utilities, but in locations with abundant solar/wind power and expensive power, it is starting to see use. For example Tesla installed a 129 MWh battery system in a low-density part of South Australia. The world's largest storage system is now in operation at the Moss Landing facility in California, holding 1.2 GWh of energy. California utility regulators have required their operators to build a substantial amount of storage.

Battery improvement is an area of frenzied academic and industrial research. There are many possible combinations of chemicals that can form a battery, but many fewer that can yield an inexpensive, practical storage medium. Figure 14.12 gives estimates for densities of some other battery types under study. There are emerging technologies with double the MJ per kg of Li-ion. But an important thing to remember is that *stationary batteries do not require low weight,* and can draw from a wider set of possibilities. The development of a new chemistry could lead to substantial drops in the cost of storage. One interesting possibility for grid storage, for example, is a zinc-air battery. "Flow batteries" can hold a reservoir of electrolyte instead

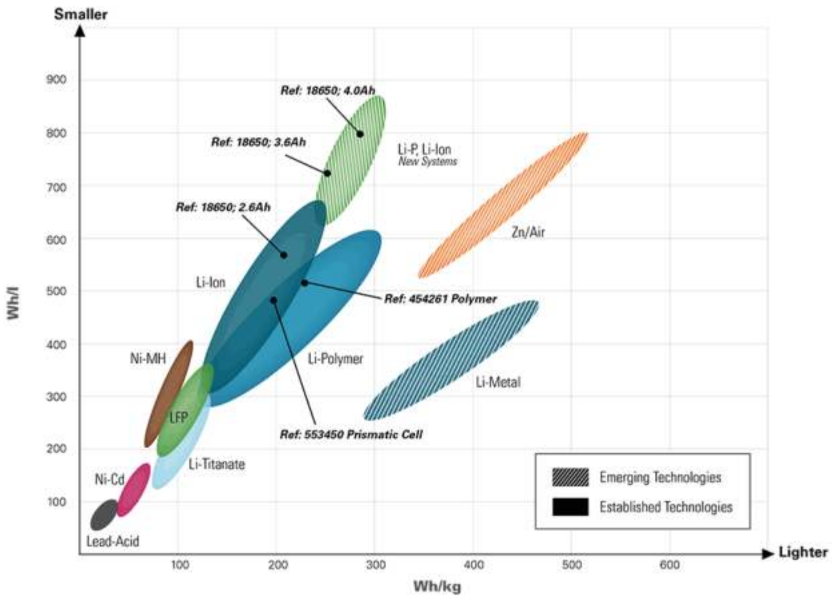


Figure 14.12: Energy densities of different battery chemistries. From ICCN Energy (2011).

of just a limited amount sealed inside. Several companies claim to be developing zinc-air batteries that can store energy well for below $100/kWh.

Practice questions

1. How far do you have to travel at highway speeds before the energy you've lost to drag is equal to the kinetic energy of the vehicle?

2. If you have a 240 V charger, how many amps of current would you need to put into the battery of your all-electric vehicle to charge it in 1 hour or less? What if you want to recharge a plug-in hybrid overnight? Make your own estimates of the energy stored in the batteries of your EV and modest-range plug-in.

3. Using the values from the examples, at what speed range is the rolling resistance higher than the drag?

4. What will happen if your car's engine does not have enough power to supply P_{climb} on a highway upgrade?

5. At current prices, how much would it cost to buy the Li-ion batteries necessary to give most passenger vehicles in the US an all-electric range of 500 km? If we did this, how many hours' worth of average US electricity generation would be stored in those batteries?

6. Using Figure 14.2, make some arguments about whether the advent of Uber and Lyft makes transportation more or less sustainable in the long run.

15

Nuclear Energy and Fusion

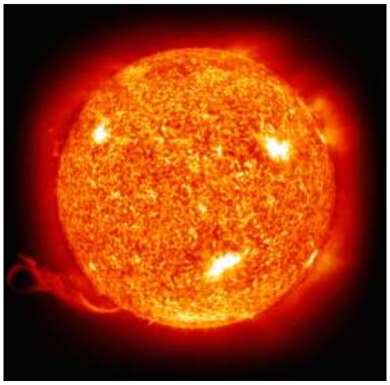


Figure 15.1: The Sun, as viewed in ultraviolet light by the *SOHO* spacecraft. Reprocessed by Ian O'Neill.

THE SUN PROVIDES 99.9% OF THE ENERGY reaching Earth's surface. Since Conservation of Energy should hold in the whole Universe, not just on Earth, we should ask *where is the Sun getting the energy to emit all these photons?* And of course, the next question is whether we can use that energy source directly to run our civilization. The source of the Sun's energy has been a mystery for millennia; it was only determined in the 1930's.

The Sun's power source

The first step in figuring out the Sun's power source is to estimate the ***luminosity*** of the Sun, which is the total power $P_\odot$ contained in the stream of photons leaving its surface.[1] Above Earth's atmosphere, the solar flux is $f_\odot = 1370\,\mathrm{W\,m^{-2}}$. Imagine building a solar collector that completely envelopes the Sun to collect *all* the photons that it emits. If this collector is a sphere around the Sun, at Earth's distance[2] $R = 1.5 \times 10^{11}$ m, then the total power collected will be

$$P_\odot = f_\odot A = f_\odot \times 4\pi R^2 = 1370\,\mathrm{W\,m^{-2}} \times 4\pi \left(1.5 \times 10^{11}\,\mathrm{m}\right)^2 = 3.9 \times 10^{26}\,\mathrm{W!!} \tag{15.1}$$

This is an incredibly large number. To give some sense of scale, the Sun produces an energy in one *second* that is equivalent to the explosion of 10 billion hydrogen bombs. Human civilization will soon be using 600 EJ per year; the Sun produces this energy in 0.0000015 seconds!

It clearly takes a lot of fuel to provide this energy. For most of human history, the most potent known fuels had energy densities in the neighborhood of coal, at $3 \times 10^7\,\mathrm{J\,kg^{-1}}$. The mass of the Sun is known to be $M_\odot = 2 \times 10^{30}$ kg, so if the Sun were a lump of coal, its total energy reserve would be $E_\odot = 6 \times 10^{37}$ J. The *lifetime* T of a

[1] The Sun also emits energy in the form of *neutrinos*, which we won't count since they pass right through Earth!

[2] How do we know the distance to the Sun or the mass of the Sun? Take Astro 001 to find out.

coal-powered sun at its current power output would therefore be

$$P_\odot = \frac{E_\odot}{T} \quad \Rightarrow \quad T = \frac{E_\odot}{P_\odot} = \frac{6 \times 10^{37}\,\mathrm{J}}{3.9 \times 10^{26}\,\mathrm{J\,s^{-1}}} = 1.5 \times 10^{11}\,\mathrm{s} = 4800\,\text{years}. \tag{15.2}$$

This is a comforting number if you agree with many literal interpretations of the Bible that the Earth and Sun were created roughly 5000 years ago, and if you think the Sun is about to go out. But it is inconsistent with many lines of biological, geological, and physical evidence that the Earth is billions of years old, and that presumably the Sun has been around without being hugely larger, brighter, or fainter during Earth's lifetime. We thus conclude that the Sun cannot be feasibly fueled by coal burning (where would the oxygen come from anyway?) or any other chemical process.

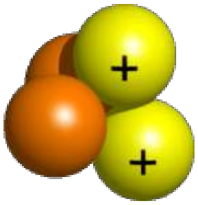


Figure 15.2: A helium nucleus has two positively-charged protons and two zero-charge neutrons. From here.

By examining the spectral lines in sunlight, we can determine the composition of the outer layers of the Sun to be 74% hydrogen, 25% helium, with ≈1% composed of any heavier elements. It is therefore wise to look carefully at hydrogen and helium as potential energy sources. Hydrogen is a single electron orbiting a single proton. The nucleus of helium (Figure 15.2) reveals a conundrum: it contains two protons in close proximity. Putting these two protons close together creates a huge electric potential energy—where did that come from? And the protons should repel each other, instantly disintegrating the nucleus. For helium and even larger nuclei to exist, there must be a ***nuclear force*** that is *stronger than the electric force when protons are extremely close together.* We therefore expect there to be ***nuclear (potential) energy*** released when building a helium nucleus from protons, and we expect the size of this nuclear energy release to be greater than the electromagnetic energy that is involved in chemical processes.

It was proposed, therefore, that the Sun produces energy by the conversion of protons (i.e. hydrogen nuclei) into helium (or larger) nuclei, releasing this nuclear energy in the process. This ***nuclear fusion*** process should be, however, very difficult to initiate—for which we should be thankful, because otherwise all of the nuclei in our bodies would have long ago combined in a blast of nuclear energy. The reason is that protons *do* repel each other at distances larger than the size of a nucleus. We need to give the protons very high kinetic energy before they collide so that they can overcome the electric repulsion to get close enough for nuclear attraction to take over, then release the nuclear energy. Temperatures above 10 million K are needed for hydrogen nuclei to attain the KE required for fusion. Below we will see how the Sun generates fusion conditions, then examine the possibilities for doing this on Earth.

Mass-energy equivalence

Figure 15.3 describes the steps that might be used to fuse hydrogen into helium in the Sun. First, two protons collide and stick together—but in the process, one of them turns into a neutron, losing its positive charge by emitting a ***positron.*** The positron is the evil twin ***antiparticle*** of the electron: identical except with opposite charge. A neutral exotic particle called a ***neutrino*** is emitted at the same time. The positron will quickly annihilate with a nearby electron, releasing energy as photons. The neutron-proton pair is ***deuterium,*** (^2H) an *isotope* of hydrogen having having 1 neutron instead of the normal zero.[3] The next step in hydrogen fusion is for the deuterium nucleus to collide with another proton to make a ^3He nucleus, releasing more nuclear energy. The last step is the collision of two ^3He nuclei, to make a "normal" ^4He nucleus and leave behind two protons.

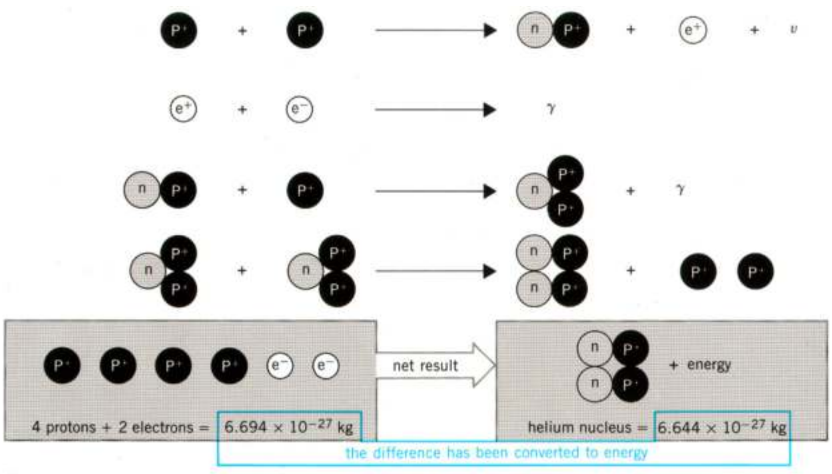


Figure 15.3: The steps of fusion of hydrogen to helium in the Sun. From Zeilek (??).

The net effect is that four protons and two electrons are converted into one helium nucleus—plus the energy, in the form of photons (and neutrinos) that we want. But as the bottom line of the figure shows, the mass of the output helium does *not* equal the mass of the proton and electron inputs added together! We find that

$$\frac{\text{mass missing}}{\text{mass input}} = \frac{0.050 \times 10^{-27}\,\text{kg}}{6.694 \times 10^{-27}\,\text{kg}} = 0.007 = 0.7\%. \qquad (15.3)$$

Just as energy is conserved—can be neither created nor destroyed, just changed in form—we have implicitly assumed that also *mass* is conserved, and therefore the fusion process should not work because this 0.7% of the mass cannot just disappear.

The answer to this puzzle is found in Albert Einstein's 1905 theory of special relativity. He proposed that *mass and energy are not seper-*

---

[3] The number to the upper left of an isotope/nucleus label is the sum of the neutron and proton count. The number of protons determines the chemical properties, and hence the element symbol.

ately conserved: they can be traded for each other, and are essentially equivalent. The "exchange rate" between mass and energy is the famous equation

$$E = mc^2 \tag{15.4}$$

where $c = 3 \times 10^8\,\mathrm{m\ s^{-1}}$ is the speed of light. Einstein suggests that the 0.7% of the mass that disappears has in fact been *converted into energy.* The amount of energy released by the fusion of 1 kg of hydrogen should therefore be the energy equivalent of the 0.007 kg of mass that disappeared:

$$E = mc^2 = 0.007\,\mathrm{kg} \times \left(3 \times 10^8\,\mathrm{m\ s^{-1}}\right)^2 = 6.3 \times 10^{14}\,\mathrm{J}. \tag{15.5}$$

This is an enormous amount of energy, because c is a big number (which we have squared). It is roughly the amount of energy released by a small H-bomb—indeed an H-bomb *is* the fusion of a few kg of hydrogen.

What Einstein tells us is that *any* time that energy is released from a system, we should see its mass go down slightly also. This is not specific to nuclear energy release. Consider for example a lithium-ion battery having a mass of 1 kg and holding 0.8 MJ of energy. After we discharge the battery, it should (and does) have less mass then when fully charged. The change in mass is

$$\Delta m = \frac{E}{c^2} = \frac{8 \times 10^5\,\mathrm{J}}{(3 \times 10^8\,\mathrm{m\ s^{-1}})^2} = 9 \times 10^{-12}\,\mathrm{kg}. \tag{15.6}$$

This mass loss—a few trillionths of the original mass—is so tiny that we would not notice it. Similarly for other chemical reactions. But nuclear energy is so much more powerful, we can notice the mass difference when it is released.

Another illustration of the difference between nuclear and chemical energy: if we took our 1 kg of hydrogen and *burned* it to make H_2O, Figure 14.9 says we would get back 143 MJ, or 1.4×10^7 J. The energy from *fusing* the same hydrogen to make helium, in Equation (15.5), is 40 millions times larger!

The lifetime of the Sun, if it starts as 75% hydrogen, would be

$$T = \frac{E_\odot}{P_\odot} = \frac{0.75 \times 2 \times 10^{30}\,\mathrm{kg} \times 6.3 \times 10^{14}\,\mathrm{J\ kg^{-1}}}{3.9 \times 10^{26}\,\mathrm{J\ s^{-1}}} = 2.4 \times 10^{18}\,\mathrm{s} = 75\,\text{billion years}. \tag{15.7}$$

Fusion power can easily power the Sun for the ≈5 billion years since we believe Earth was created, and indeed many more!

Fusion requires a combination of high temperature (so protons collide at high KE) and high density (so that collisions between protons are frequent, giving many attempts at fusion). The enormous mass of the Sun means that *gravity* naturally generates these conditions at the

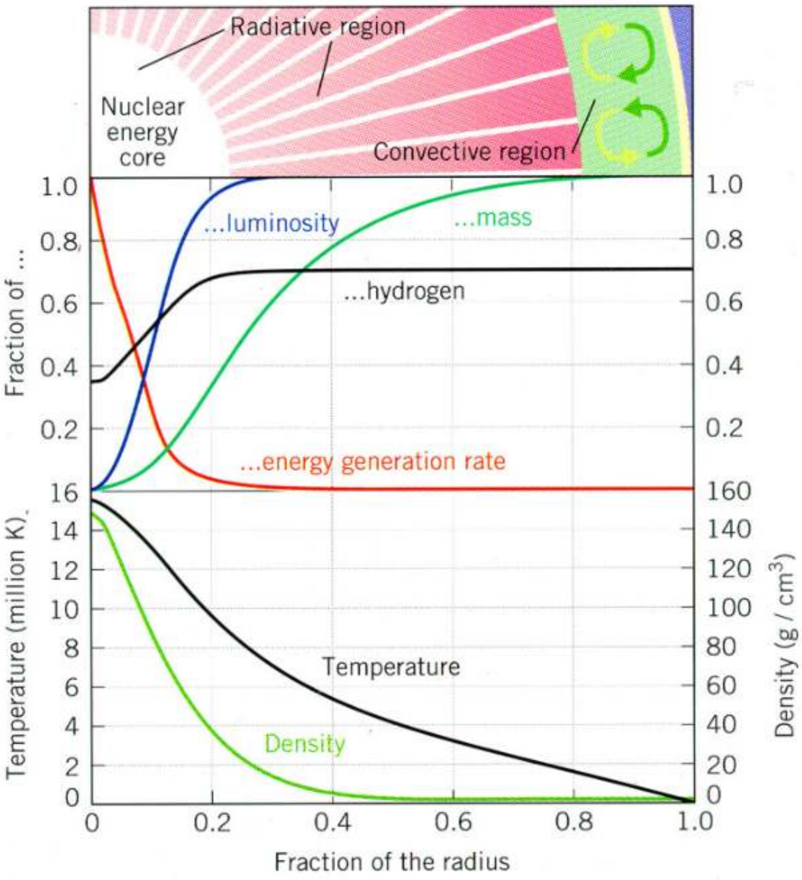


Figure 15.4: What is happening inside the Sun, as a function of distance from the center (radius 0) to the surface (radius 1). The blue "luminosity" curve shows how much of the Sun's power is produced *interior* to the given radius: nearly all fusion occurs in the central 20%. At bottom, the black and green curves show the temperature and density rising to the center, making the conditions amenable to hydrogen fusion. From ??

Sun's core: as we descend from the Sun's surface, we feel the weight of all the hydrogen/helium gas above us. This weight compresses the gas, raising its temperature, pressure, and density. At the center of the Sun, the temperature is over 15 million K and the gas is $10\times$ the density of lead! The center of the Sun is hence a natural, highly prodigious fusion machine. To produce the measured $P_\odot$ requires the fusion of 6×10^{11} kg of hydrogen (600 million tons) of hydrogen every *second*.

The black line in the middle panel of Figure 15.4 shows that in the central fusion zone of the Sun, half of the hydrogen as already been converted to helium. In another 5 billion years, the hydrogen will run out and the Sun will begin to die. So solar energy is sustainable for "only" another 5 billion years, then any beings left on Earth will need to find another planet to live on.

Controlled nuclear fusion

Hydrogen fusion would be the ideal energy source on Earth. The input fuel, hydrogen, is enormously abundant: the oceans are filled with H_2O.[4] We would need only a few thousand tons of hydrogen per year to power all of civilization, so the fuel source would be inexhaustible, for practical purposes. The waste product of fusion is helium—an inert, harmless gas. There is no CO_2 output, of course.

Thoughts of harnessing fusion power on Earth began very soon after we realized that it powers the Sun. The first fusion reactions created by humans were in hydrogen bombs (1952). The high temperatures needed for fusion were created by an atomic bomb (using fission, described in the next Chapter). The sudden release of all the energy in ≈1 kg of hydrogen is highly destructive, and we clearly need a method for controlled release of fusion energy.

The principal difficulty of controlled fusion is: how do you contain a gas at temperatures of $> 10^7$ K? Obviously any material would vaporize. To make matters worse, we cannot produce the high densities of gas present at the Sun's core, so our fusion reactor needs temperatures above 10^8 K to succeed. Earthly fusion reactors (and bombs) use ***DT*** fusion reaction, in which a deuterium (^2H) and a tritium (^3H) nucleus are fused. This can occur at lower densities than the proton-proton fusion pathway in the Sun. These hydrogen isotopes are rare, but there is still plenty in the oceans to fuel a fusion society for many years before we'd need to switch to the more difficult fusion chains.

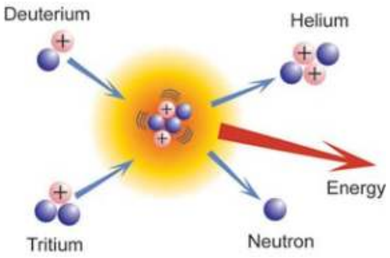


Figure 15.5: The DT fusion reaction. From ??.

[4] The chemical energy needed to separate the chemical bonds in water is millions of times less than the fusion energy we would get out of the resultant fusion process.

Magnetic confinement

The most-explored strategy for controlled fusion is to use magnetic fields to confine the very high-temperature hydrogen gas (called a ***plasma*** because the atoms will be ionized). As illustrated in Figure 15.6, electrons and protons will tend to spiral around magnetic field lines. A ***tokamak*** creates a doughnut-shaped circuit of magnetic field; if the plasma follows the magnetic field, it can be kept away from the container walls.

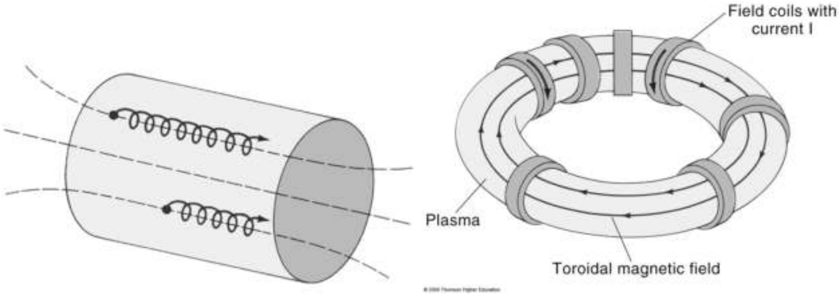


Figure 15.6: *Left:* Charged particles tend to spiral around magnetic field lines. *Right:* A tokamak creates a doughnut-shaped magnetic "racetrack" for hot plasma, confining the superheated plasma and keeping it away from the walls. From Hinrichs.

Tokamak research has been underway since the 1950's. Experimental tokamaks in the US and elsewhere have succeeded in producing nuclear fusion in plasmas above 100 million K. With DT fusion, they have reached ***break-even*** conditions, when the power being produced by fusion momentarily exceeds the power being put into the machine to heat and confine the plasma. After 50 years of research, however, we are still not close to a commercially viable fusion reactor. The next step will probably be taken by the *ITER* reactor, under construction in France by a consortium of most of the industrialized countries of the world (including the US). ITER's goal is to produce 500 MW of fusion power (10 times the input energy) for 1000 seconds at a time. The project is far behind schedule and over budget: it will cost at least $25B. The current schedule is to produce the first plasma in 2025 and first experiments with "burning" DT fuel not until 2035. ITER will not produce electricity from the fusion power that it generates. Further generations of experiments will be needed to develop the techology for harnessing the fusion power; and of course to bring the costs of fusion power into any viable range.

Magnetically confined fusion is clearly difficult to pull off: we can do it, but after 60 years of work, we are clearly still at least a few decades away from having a practical power source.

Inertial confinement

An H-bomb works because the hydrogen is heated to enormous temperatures so quickly that it fuses before it has time to spread out. ***Inertial confinement*** attempts to replicate this on a non-destructive scale. An inertial confinement reactor would drop a stream of millimeter-sized frozen hydrogen pellets into a reactor chamber. Instead of an atomic bomb, the hydrogen is heated by blasting it with the most energetic lasers ever constructed. A football-field-sized array of lasers is focused onto this tiny pellet, with exquisite synchronization necessary to get all the energy into the pellet before it disintegrates. The largest inertial-confinement experiment to date is the *National Ignition Facility (NIF)* at the Lawrence Livermore National Laboratory in Cal-

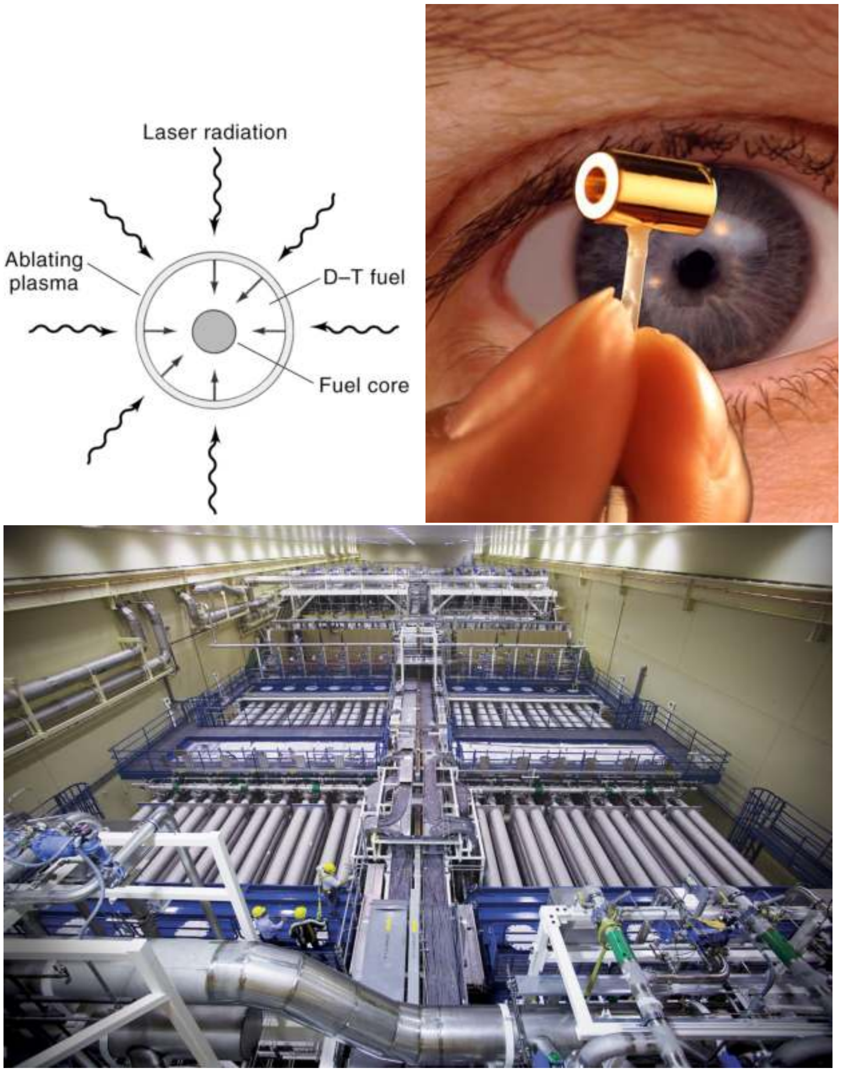


Figure 15.7: *Left:* Illustration of inertial-confinement fusion (Hinrichs). *Right:* This small cylinder holds the frozen hydrogen pellets that are used to fuel the *National Ignition Facility* inertial-confinement experiment (LLNL). *Bottom:* The hall holding half of the NIF lasers that will be focussed onto the small pellet of hydrogen (LLNL).

ifornia.[5] The NIF, like ITER, is far over budget and behind schedule. Unlike ITER, it is now in operation. It was designed to feed several 100 MJ of electricity into its lasers to generate a pulse of 1.8 MJ of photons that hit a 3 mm pellet of hydrogen in a few-nanosecond pulse, and ideally get more than 1.8 MJ of fusion energy released. This "scientific break-even" is of course not the same as having an operating power plant that yields net energy. But NIF was not able to reach scientific break-even after several years of operation, yielding far less energy than originally projected, probably due to instabilities in the heated fuel. As of 2021, some fraction of the NIF's time is still being used on fusion energy, although most of its experiments are to help in the design of nuclear weapons. The fusion energy released during each "shot" has improved but is still far from scientific break-even.

In conclusion, fusion is clearly the energy source of choice, but is not likely to make a significant impact for at least a few decades. At current rates it's not clear we would ever reach the goal of a practical energy source. Funding for fusion research has been inconsistent, and of course is a miniscule fraction of global energy spending. A better-funded research program might get us there sooner.

Practice problems

- How many kg of water per year would be needed to fuel a fusion reactor that produced all the energy needed in the US?
- When matter comes into contact with antimatter, they annihilate each other: *all* of the mass is converted into pure energy (e.g. photons). Antimatter would therefore be the ultimate vehicle fuel, giving the highest possible energy per kg. If you could store antimatter in your car, what mass of antimatter would you need to run your car for its entire lifetime? Assume a typical lifetime and energy needs for a current vehicle.
- Would it be safe to stand next to an operating fusion reactor?

[5] Livermore is where they design nuclear weapons. NIF funding is partly (mostly?) justified by experiments that help understand weapons design.

16

Nuclear Fission

THE MASS OF A NUCLEUS is an indicator of its nuclear potential energy, since $E = mc^2$ tells us that mass and energy are equivalent. Because the mass of ^{4}He is only 0.993 times the mass of 4 protons (H), we know that hydrogen fusion will release 0.007 of the input mass in the form of energy. Are there other ways that we can release nuclear energy by rearranging the ***nucleons*** (protons and neutrons) of atomic nuclei? Figure 16.1 plots the mass (per nucleon) of the most common isotope of all the naturally-occurring elements. Helium is, as expected, well below hydrogen on this chart, and in fact the mass per nucleon continues to decrease as the nuclei get larger, up to iron. Stars can produce energy by fusing hydrogen into helium and then into larger elements.[1]

Nuclei larger than ^{56}Fe, (26 protons, 30 neutrons), generally have *more* mass per nucleon as they grow larger. This means that we could potentially release nuclear energy by breaking a single nucleus into *smaller* pieces. For example the most massive naturally-occurring element, uranium, could potentially release about 0.1% of its mass by converting to elements like iron or lead. Accessing this nuclear energy turns out to be much easier than hydrogen fusion, and will be the topic of this Chapter.

Radioactivity and its hazards

There are hundreds of known nucleii (a.k.a. isotopes), each with its own mass and implied nuclear potential energy. For many of them, a lower nuclear potential energy can be attained by rearranging the protons and neutrons. Such nuclei can undergo ***spontaneous decay,*** in which they'll rearrange themselves and release the nuclear energy. This spontaneous energy release is called ***radioactivity,*** which was discovered by Becquerel in 1896—though it took work over the following decade to realize that this spontaneous release of energy was

[1] This is in fact where all the atoms heavier than helium were created, including the ones in your body.

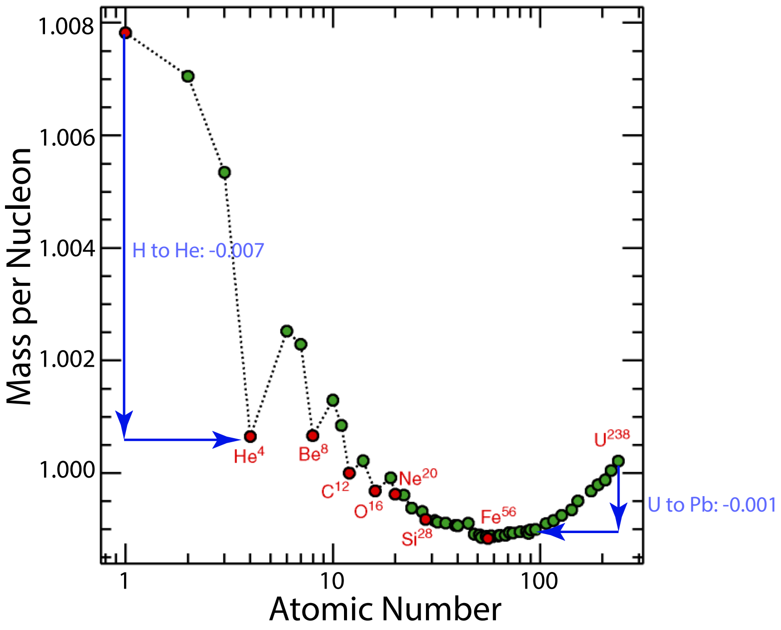


Figure 16.1: The mass per nucleon of the most stable isotope of each naturally occurring element. Fusing hydrogen to helium released 0.7% of the mass as energy; decay of uranium into lead (and other mid-sized elements) releases about 0.1% of mass as energy too.

due to nuclear decay. The energy can be released in several forms, which we now understand as follows:

- In ***alpha decay,*** the nucleus spits out a helium nucleus (2 protons plus 2 neutrons, also known as an α particle), thus reducing its atomic number by 2. The resulting "α rays" are easily blocked, e.g. by a sheet of paper.

- In ***beta decay***, one of the neutrons converts to a proton, and the energy is carried away by an electron (a.k.a. β particle or ray). The atomic number *increases* by 1. β-rays are more penetrating than α's, making it through a sheet of paper, but they can be blocked by a few mm of metal. Sometimes a proton can turn into a neutron, with a positron carrying away the energy instead of an electron.

- In ***gamma decay,*** the nucleus undergoes internal rearrangement, releasing a very high energy photon (γ ray) and keeping the same proton/neutron count. γ rays are highly penetrating: it takes several cm of lead to stop them, or greater thicknesses of steel, concrete, or water.

- A *neutron* can be expelled from the nucleus. The early researchers did not know about this, since the neutron has no electric charge and does not leave a mark on photographic film. Neutrons are highly penetrating, like γ rays.

- In ***fission,*** the nucleus breaks into two smaller nuclei (but bigger than helium), sometimes spewing out one or more neutrons at the same time.

An isotope that does not undergo spontaneous decay is called ***stable.*** Each unstable nucleus has a ***half-life*** $t_{1/2}$ that describes how quickly it will decay. We can never predict exactly when an individual nucleus will decay. But we do know that *on average,* half of them will decay in time $t_{1/2}$. Of the surviving nuclei, another half will decay in the next half-life, leaving 1/4 of the original nuclei. After time $3t_{1/2}$, 1/8 of the nuclei are left, on average. The general formula is

$$(\#\text{ nuclei left at time } t) = 2^{-t/t_{1/2}} \times (\#\text{ nuclei at time } t = 0)\,. \qquad (16.1)$$

Some isotopes have half-lives of microseconds or less, some take billions of years to decay. An isotope might be capable of decaying in more than one of the above fashions. Again we can never guess exactly which decay path it will take, but we can know the average ***branching ratios*** giving the probability of each type of decay.

Often the "daughter" product of a radioactive decay is itself a radioactive isotope. An sequence of successive decays may occur before a stable isotope is reached, with some steps very slow and others rapid. Figure 16.3 shows how ^{238}U, with $t_{1/2} = 4.5$ billion years, will go through 14 decay steps, taking anywhere from milliseconds to hundreds of thousands of years, before ending up as stable lead. Since isotope can have multiple decay paths, there can be multiple different decay chains and endpoints for a given isotope.

Radioactivity can be very useful. It can be a power source—a chunk of radioactive material can generate heat (and radiation) for decades, since the nuclear energy within is so potent. The *New Horizons* spacecraft that passed Pluto in 2015, and the *Curiousity* rover on Mars are both powered by the heat emitted by a few kg of the α-emitting isotope ^{238}Pu, with 87-year half-life.[2]

Radioactive dating allows us to determine the *ages* of many substances. The basic idea is to determine how much of an isotope is present compared to its daughter products, and then using Equation 16.1 we can estimate the time since the sample was created. Radioactive dating of Earth and Moon rocks and meteorites tells us that the solar system (including the Sun) formed about 4.6 billion years ago. The radioactive isotope ^{14}C, with $t_{1/2} = 5730$ yr, can be used to determine the age of archeological finds.

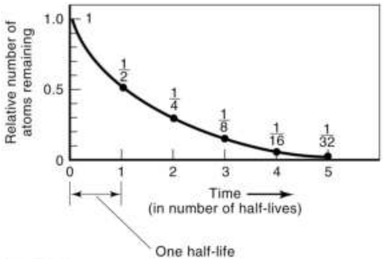
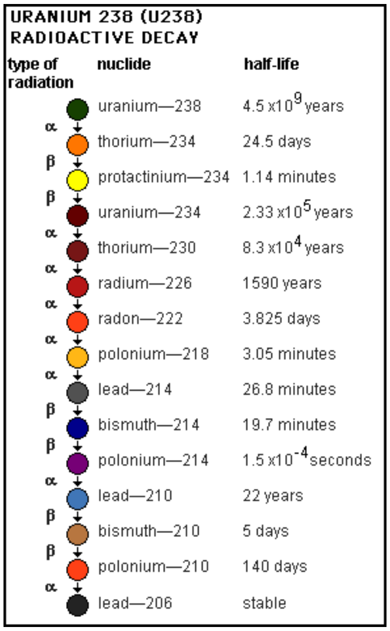


Figure 16.2: Graph of radioactive decay, as in Equation (16.1). From Hinrichs Figure 13-10.



URANIUM 238 (U238)
RADIOACTIVE DECAY

type of radiation	nuclide	half-life
	uranium—238	4.5×10^{9} years
α	thorium—234	24.5 days
β	protactinium—234	1.14 minutes
β	uranium—234	2.33×10^{5} years
α	thorium—230	8.3×10^{4} years
α	radium—226	1590 years
α	radon—222	3.825 days
α	polonium—218	3.05 minutes
α	lead—214	26.8 minutes
β	bismuth—214	19.7 minutes
β	polonium—214	1.5×10^{-4} seconds
α	lead—210	22 years
β	bismuth—210	5 days
β	polonium—210	140 days
α	lead—206	stable

Figure 16.3: The dominant decay chain for ^{238}U, showing the half-life of each isotope along the way. From ???

Health effects of radiation

Radioactivity is not all good, however. The particles emitted by decaying nuclei have very high energy, and are capable of breaking

---

[2] Photovoltaic panels get too weak to run these spacecraft as they get farther from the Sun.

many molecular bonds before they are stopped. In other words nuclear decay produces *ionizing radiation,* with the same health hazards as the high-energy photons described in Chapter 5. We know, therefore, that exposure to radioactive decays will increase the chances of cancerous mutations in cells, and higher exposures to decaying isotopes will cause ***acute radiation poisoning*** whereby tissue function is disrupted by the destruction of cells.

We need to limit, therefore, our exposure to emissions from radioactive decays. The potency of a given source of radioactivity is measured in ***Becquerels (Bq):*** a 1 Bq source has 1 radioactive decay per second.[3] Note that sources with very *long* half-lives (millions or billions of years) will have very low potency, because in any given second, very few of the nuclei will decay. We also don't worry very much about isotopes with very *short* half-lives (seconds or less), because by the time any of these could get out of a laboratory, nearly all nuclei would have decayed already. The isotopes we need to look out for are those with half-lives of days to thousands of years.

Next we need to consider how much of the radioactive emissions (ionizing radiation) makes it into our tissue. Exposure to radiation is measured in ***Grays (Gy):*** a 1 Gy exposure means that 1 Joule of energy has been deposited in each kg of tissue. A related quantity is the ***Sievert (Sv),*** which adjusts the exposure by some factors related to how damaging different particles (alphas, gammas, neutrons, etc.) are to tissue.[4] An exposure of 1 Gy will generally cause significant acute damage, and doses of 4 Gy or higher are usually fatal. Doses below 0.1 Sv do not appear to cause detectable tissue damage.

The dose you recieve from a radioactive source depends not just on its Bq output, but also of course on your distance. And very importantly, on the type of radiation it emits and the ***shielding*** between you and the source. As noted above, virtually any barrier will block α radiation, and β's are easy to block as well. It is the neutron and γ emitters, which require meters of concrete or water to stop, that are hardest to shield. But α and β emitters are still a problem if you ingest or inhale them, since of course any radiation from them is absorbed in your body. A primary example is the radon isotope ^{222}Rn, which is part of the decay chain of the uranium and thorium that occur naturally in rocks. This dense gas emerges from the ground and collects in poorly ventilated basements. It is an alpha emitter with $t_{1/2} = 3.8$ days. The US EPA estimates 21,000 lung cancers per year are caused by decay of inhaled radon.

While the results of acute radiation damage are obvious and easily measured, it is much more difficult to ascertain the rate at which ionizing radiation exposure causes cancer. Long-term studies of survivors of the Hiroshima and Nagasaki atomic bombs measure their

[3] The Curie (Ci) is an older unit of radioactivity.

[4] Older units for exposure are the *rad* and the *rem*, which are $100\times$ smaller than the Gy and Sv, respectively.

death rates from cancer as a function of the radiation exposure they were estimated to have when the bombs drop. This and other evidence suggest that each Sv of exposure raises the incidence of cancer by about 5%.[5] This varies with age of the victim, organs exposed, etc. What remains very controversial is whether the ***linear no-threshold (LNT)*** model,

$$(\text{excess cancer rate}) \approx 0.05 \times (\text{Sv exposure}) \qquad (16.2)$$

is valid for very low exposures, in the mSv range. This would be our first guess, if we expect each mutation caused by ionizing radiation to have the same chance of causing cancer. On the other hand, we are exposed to ***natural background radiation*** constantly: radon emitted from the ground, cosmic radiation raining down from the sky, and other naturally occurring radioactive elements.[6] On average, we recieve a few mSv of dose per year from natural sources, and many sites (with more radioactive geology, or at high altitude with less atmospheric shielding of cosmic rays) are substantially higher. Medical x-rays typically add a few more mSv per year in the US. As the human race (and other species) have done pretty well for themselves over eons with this level of exposure, we might guess that our bodies can repair low-level radiation damage. At the least, it is clear that we should not be concerned with radiation sources that are smaller than the natural background of a few mSv per year. US regulations limit occupational exposure to <50 mSv per year.

It is very difficult to measure the cancer effects of low-level radiation, because (a) a cancer death might be many years after the radiation exposure that caused it; (b) people die of cancer all the time, and we can't tell which ones are caused by manmade radiation vs all other possible causes. So the LNT model is not well tested. Nonetheless the LNT model remains the standard means for estimating cancer deaths from radiation, since it's simple, it's the most conservative choice, and nobody has good evidence for an alternative either. Just keep in mind that estimates of cancer deaths from radiation exposure are just that—estimates—and they may be erring significantly on the high side.

Radioactivity is also a boon to health care: the penetrating power of radiation enables imaging of the insides of our bodies through various techiques like positron emission tomography (PET) scans. And the cell-killing abilities of ionizing radiation are frequently deployed against malignant tumors.

[5] International Commission on Radiological Protection Publication 103 (2007).

[6] Bananas, high in potassium, are also high in β-emitting ${}^{40}K$! You get a dose of $\approx 10^{-7}$ Sv from eating a banana.

Chain reactions

Spontaneous radioactive decay can power a small spacecraft but not a GW power plant. To harness nuclear power on a large scale, we need to be able to induce nuclei to break apart "on demand." Figure 16.4 illustrates that ^{235}U is a ***fissile*** nucleus: we can induce fission by bombarding the nucleus with a neutron. When ^{235}U fissions, it also releases neutrons—an average of 2.5 of them. If 1 or more of these neutrons are, on average, absorbed by other ^{235}U nuclei, we have a ***chain reaction.*** We can have the number of fissioning nuclei double every microsecond (or less). This is a classic example of exponential growth: in a very short time, a huge amount of energy can be released. These are the conditions that are produced in ***atomic bombs.*** When ^{235}U fissions, roughly 0.1% of its mass is converted to energy.

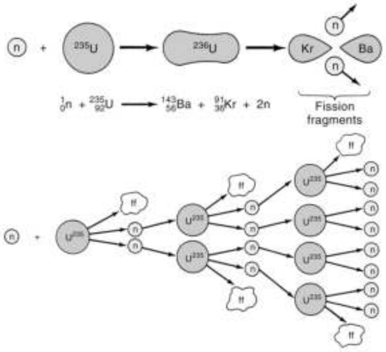


Figure 16.4: *Top:* stimulating fission in ^{235}U by adding a neutron. *Bottom:* A chain reaction for fissile nucleii. From Hinrichs, Figs 14.3 and 14.5

There are three fissile isotopes: ^{235}U, which makes up 0.72% of natural uranium (most of the rest is non-fissile ^{238}U); ^{239}Pu, which is not found in nature, but can be produced by bombarding ^{238}U with neutrons; and ^{233}U, which can be produced by bombarding the naturally occurring element thorium with neutrons. One of the atomic bombs dropped on Japan used ^{235}U and the other was a plutonium bomb.

To build a successful nuclear weapon (or reactor), we need to insure that >1 neutron from each fission is absorbed by another fissile nucleus. Neutrons are highly penetrating, so if the chunk of fuel is small, most of the neutrons will escape before being captured. There is a ***critical mass*** of fissile material necessary to sustain a chain reaction. This critical mass depends upon the purity of the fuel, i.e. what fraction of the fuel is made up of the fissile nucleus. Absorbing neutrons in non-fissile nuclei will impede the chain reaction. For pure ^{235}U, the critical mass is about 50 kg, a sphere of just 17 cm diameter. It can be lowered to just 15 kg if we can surround it with a *neutron reflector* material. The critical mass of a plutonium weapon is even lower—about 5 kg with a neutron reflector.

Anti-proliferation

The hardest part of building a uranium bomb is ***enriching*** the uranium from its natural 0.7% of the fissile ^{235}U to the 90% level in weapons-grade *highly enriched uranium (HEU)* that is needed for a bomb. It is very difficult to separate two different isotopes of the same element, since they behave the same in every chemical reaction. Once you have HEU in hand, the bomb is relatively easy to make, perhaps within the abilities of non-state (terrorist) organizations: you build two sub-critical mass chunks, and at the desired moment slam them together to exceed critical mass. To prevent the proliferation of nuclear weapons, we want to keep very close track of any HEU that is produced, and get very concerned when nations (like Iran) develop the technology for isotopic enrichment.

Plutonium is created in "breeder" reactors where ^{238}U can be bombarded by neutrons to create the Pu. Plutonium can be separated from the uranium using chemistry, which is much easier than the isotopic separation that is necessary to create a uranium bomb. But for various reasons it is much more difficult to build a Pu bomb than a U bomb (for one thing, Pu is highly toxic and more radioactive, so will tend to kill the people working on it). It is widely believed that building a Pu bomb requires the resources of a nation and is beyond terrorist capability—"non-state actors" would need to steal some HEU and fashion a bomb from it. North Korea is believed to have extracted plutonium from its nuclear reactors and tested plutonium bombs. It is this activity that negotiations aim to halt.

A ***thermonuclear*** weapon (H-bomb) begins with a *primary* explosion of a uranium or plutonium fission core. The radiation from the primary is used to heat hydrogen in a *secondary* device to the point of fusion. The radiation must do its fusion job before the blast wave from the primary obliterates the secondary. Nations known to have exploded H-bombs are the US, UK, France, Russia, and China (and possibly, but not likely; Israel, India, and Pakistan). North Korea claimed to have exploded two H-bombs in 2017, though it is not clear that these were real fusion bombs rather than fission.

Practical nuclear power

To harness a fission chain reaction for useful power (instead of bombs), we need a steady power output which means that exactly one neutron from each fission must be absorbed by another fissile nucleus. Less than one and the reaction fizzles out; more than one and it grows out of control. To maintain this delicate balance we must have some kind of *negative feedback* in the reactor: for example, arrange for higher power output to create higher temperatures which then cause to the fuel to expand slightly and lower the neutron capture rate.

A practical nuclear reactor design must pay close attention to managing the *heat* and the *neutrons*. It will include these elements:

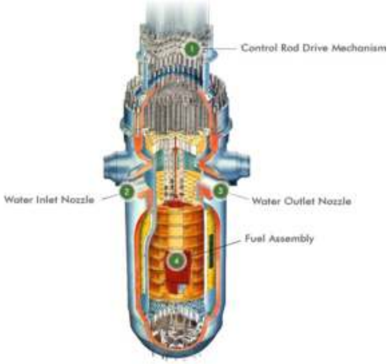


Figure 16.5: Cutaway of the reactor vessel for the Daya Bay (China) pressurized-water reactor. Control rods extend down into the fuel assembly. The vessel is 13 m tall, 4.4 m in diameter, with walls 20 cm thick. From HK Nuclear.

- *Fuel:* critical mass of some fissile isotope. Most reactors use ***low enriched uranium (LEU)*** which has been enriched to 3–4% ^{235}U (the rest is ^{238}U). The lower ^{235}U concentration allows the reactor to be bigger (and easier to cool), and also the LEU cannot be used to create a bomb. As the fuel sits in the reactor's neutron bath, ^{239}Pu is produced, which then contributes to the chain reaction. Many other radioactive elements are produced as part of the fissile elements' decay chains. Furthermore many stable nuclei in the reactor will absorb neutrons and become radioactive. So a good fraction of the reactor's power production comes not from fissions, but from spontaneous decays of these "activated" isotopes. It is also possible to fuel a reactor with *thorium,* a more abundant element than uranium, which is not itself fissile, but produces the fissile isotope ^{233}U when under neutron bombardment. So a thorium reactor breeds its own fuel.

- A ***moderator*** is a material that slows down the high-energy neutrons that are emitted during fission. This makes them more likely to be absorbed by other nuclei, allowing us to reach critical mass with less-concentrated fuel. If the moderator is removed, the chain reaction will stop. Light elements make good moderators: most reactors use water as a moderator; graphite (carbon) works too.

- ***Control rods*** are made of some material that is very neutron-absorbent. Lowering these into the reactor core slows (or stops) the chain reaction. Reactors will have a safety mechanism whereby control rods automatically fall completely into the reactor, halting the chain reaction, whenever power is lost or other malfunctions occur. But this does not instantly stop the production of energy in a reactor core, because as noted above, much of the energy is coming from spontaneous decay of activated isotopes. The control rods cannot stop this. The reactor will still generate potentially damaging amounts of heat until the elements with the shortest half-lives decay away. When the tidal wave hit the Fukushima reactors, their control rods shut down the chain reactions as designed. But the continued heat release from spontaneous decays led to damage of the reactor cores and buildings.

- A ***coolant*** is necessary to remove heat from the reactor core. First, because we need to remove the heat to keep everything from melting. Second, because the whole point of building the reactor is to generate energy we can use, so the coolant carries this energy out, where it is used to make high-pressure steam that runs a turbine and generator as in combustion-based power plants (Figure 16.6). The coolant is most often water, but can be helium gas, or molten salts or sodium in other designs. In a ***loss of coolant accident (LOCA),*** the circulation of coolant through the reactor is interrupted and its internal temperature can rise, potentially to the point of ***meltdown*** of the fuel assemblies in the core. This renders the reactor inoperable; and at worst, the heat melts or breaks the *reactor vessel* so that highly radioactive material will leak out of it. US reactors are all built with the reactor vessel inside a thick concrete secondary containment building.

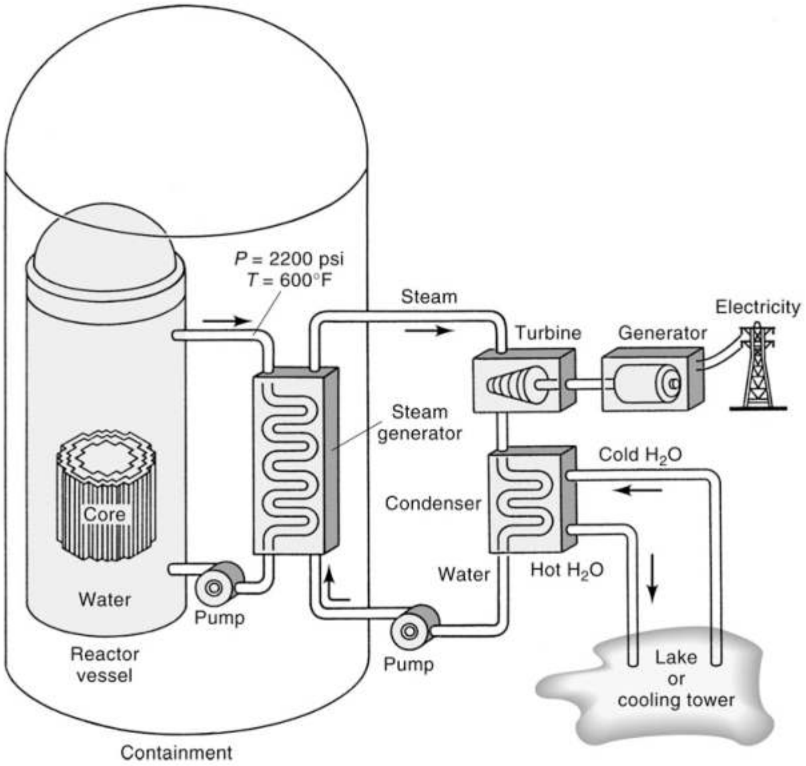


Figure 16.6: Schematic of a pressurized water reactor. From Hinrichs, Fig 14.11.

Nearly all commercial power generating reactors are boiling water reactors (BWRs) or pressurized water reactors (PWRs), where water is used as the coolant and also serves as the moderator. These designs are descendants of the original nuclear reactors designed to power nuclear-missile-carrying submarines for the US Navy in the 1950's.

Nuclear reactor design is a very tricky business. The enormous neutron flux inside the reactor (which would instantly kill anyone exposed to it) means that once in operation, you can really never go in to repair the core itself—yet you need to be able to pull out and replace the fuel assemblies as their ^{235}U is exhausted. The reactor vessel and piping need to withstand high pressures, but the neutrons are constantly *transmuting* atoms into other elements, changing the chemical composition of the components and potentially weakening them. One has to be careful in choosing materials for reactors. One needs to be very careful that the chain reaction always has enough negative feedback to stay in control. And of course care must be taken to prevent escape of the highly radioactive materials present in the reactor core.

As a consequence, reactor designs and construction are scrutinized very carefully, making them quite expensive and slow to build. The thermodynamic efficiency of nuclear reactors tends to be closer to 30% than to the higher values for modern coal or gas plants, because the engineers do not want to push the materials to as high temperatures and pressures. The steam temperature for nuclear reactors is typically $\approx 300^\circ$ C.

Hazards

A nuclear reactor in normal operation is extremely benign to the environment—much more so than a coal power plant. The only significant emission is waste heat (just as for any heat-engine-based power plant), and of course there is no CO_2 emission from nuclear plant operation. Any radiation escaping the containment building in normal operation is well below the natural background, as the reactor vessel is heavily shielded, and coolant that passes through the reactor circulates in a closed loop. In the US, no radiation-related deaths are attributable to the operation of commercial nuclear power plants. A handful of employee fatalities from industrial accidents have occurred, but the rate of these per EJ of energy produced is far below those from coal and oil extraction.

One hazard we do *not* need to worry about is that the reactors will explode as an atomic bomb. The fissile nuclei are not sufficiently concentrated for this to occur. The greater danger is from potential release of radiation when things go awry in a reactor. Radiation releases from commerical power plants are rare—in the past fifty years, I find reference to only about 5 incidents of radiation release outside the reactor building.[7] Three of these are well known and have generated substantial fear and reaction in the community. The first was the 1979 event at one of the Three Mile Island reactors in Harrisburg, Pennsylvania, just 100 miles west (upwind) of Philadelphia. Quoting from the Nuclear Regulatory Commision's fact sheet on this accident, "A combination of equipment malfunctions, design-related problems and worker errors led to TMI-2's partial meltdown and very small off-site releases of radioactivity." The reactor was rendered non-functional (the other TMI reactor is still in operation), but the radiation released was too small to have significant health effects. The additional radiation exposure to a bystander at the site boundary would be equivalent to roughly one year's worth of natural background radiation. For two or three days, however, there was significant fear of a more severe event: after loss of coolant, the tem-

[7] A table is here.

peratures in the reactor became high enough for a chemical reaction in which water reacts with zirconium in the fuel assembly to release hydrogen gas. If the hydrogen mixes with oxygen it can explode, potentially rupturing the reaction vessel. In the TMI accident, the hydrogen bubble in the reactor vessel was deprived of oxygen and did not explode.

Figure 16.7: The Chernobyl nuclear reactor, after the explosion and before it was encased in concrete. From ???

The second accident, and by far the worst nuclear accident in history, was the 1986 explosion of the Chernobyl reactor in the Ukraine. This was an example of how to do everything wrong, and a demonstration of shocking disregard for public welfare by the government. The Chernobyl reactor was an unusual design, graphite-moderated and water-cooled. It was designed to be refueled while still in operation—economical, but intrinsically unsafe, because loss of the water coolant left the graphite moderator in place, so the reaction rates could actually speed up with loss of coolant. Furthermore the reactor was built *without a containment building* (too expensive!). The reactor operators were experimenting with the reactor, unknowingly putting it into a condition with a *positive temperature coefficient,* in which the normal negative-feedback regulation of the reaction rate is defeated, and runaway reactions can (and did) result. Several safety systems were shut off, in violation of operating rules. The reactor ran out of control and exploded, spewing highly radioactive material all over the vicinity and into an atmospheric plume that spread across Europe. Fifty-four people died from acute radiation poisoning or other short-term consequences of the accident.

Through sometimes heroic efforts, the reactor was entombed in concrete. A region of $\approx$ 1000 square miles around the site remains largely evacuated. The World Health Organization reported in 2006 that approximately 6,000 thyroid-cancer cases in Belarus, Ukraine, and Russia are thought to be predominantly attributable to drinking milk with high levels of radioactive iodine, with 99% of these being successfully treated. For the 600,000 evacuees and others most affected by the accident, the LNT model predicts an additional 4000 deaths from cancer over their lifetimes. For the 6 million people in other contaminated areas, the LNT prediction is 5000 additional cancer deaths. It is expected to be impossible to confirm these predictions, because millions of cancer deaths will occur in the affected areas even without Chernobyl, and one cannot determine which cancers were caused by Chernobyl's radiation. And indeed the WHO reports that there is no statistically reliable evidence of additional cancers from Chernobyl accident beyond the thyroid cancers.

The third major reactor accident occurred in 2011 at the Fukushima Daiichi power station near Sendai, Japan following a massive earthquake and subsequent tidal wave that devestated many cities along the Japanese coast. When the earthquake hit, the Unit 1, 2, and 3 reactors were in operation, Unit 4 had its fuel removed (for refueling), and Units 5 and 6 were in cold shutdown conditions. As per design, the control rods automatically shut down the chain reaction in the operating Units 1–3. Diesel generators were used to keep cooling water circulating, however, because spontaneous decays continue after the chain reaction stops. Immediately after shutdown, this produces 6% of the full core heat output, dropping to 1.5% after 1 hour, 0.4% after a day, and 0.2% after a week.[8] A 15-meter-high tsunami arrived 50 minutes later, overtopping the 10-meter-tall seawall. This disabled the diesel backup generators; battery power was knocked out or expired the next day, shutting down the cooling water. For these reactors, the decay heat is still sufficient to cause damage in the absence of active cooling. Hydrogen explosions occurred in the secondary containment buildings of Units 1 and 3, and even in the defueled Unit 4 from hydrogen flowing through ductwork. The explosions, radioactivity, lack of electrical power at the site, and tidal-wave damage to surrounding infrastructure meant it took months to bring the situation under control.[9]

Figure 16.8: Units 1–4 of the Fukushima Daiichi Power Station before (left) and after (right) the tsunami and hydrogen explosions.

Meanwhile, in the reactor cores, heat built up and caused significant melting of much or all of the fuel rods in all three active units. In Unit 1 the fuel melted through the pressure vessel to solidify on the concrete floors of the containment vessel. The hydrogen explosions and venting released large amounts of radioactive material from the building. Furthermore the water being pumped or sprayed in for emergency cooling carried radioactive materials away, and much is believed to have leaked away to the sea.

The total amount of radiation released is estimated to be large,

---

[8] `http://www.anl.gov/sites/anl.gov/files/spent_fuel_nutt.pdf`

[9] A detailed account can be found here.

but about 10× below Chernobyl's. Units 1–4 are unusable and will need to be decommissioned. No deaths or injuries are directly attributable to radiation exposure at the plant (of course many thousands of people, including plant employees, lost their lives in the earthquake and tsunami). The radiation levels at the plant boundaries are below natural background, but a substantial zone remains evacuated. The WHO estimated in 2013 using LNT models that the most vulnerable residents (infants) of the two most-affected locations in Fukushima prefecture would have lifetime cancer risks elevated by 7% of the baseline cancer rates. Beyea, Lyman, & von Hippel[10] estimate roughly 1000 future mortalities from the Fukushima radiation release. This will again prove unmeasurable amongst the much larger numbers of cancers in the population.

What we can conclude about nuclear safety from this history of accidents? First, that accidents have happened and will probably continue to occur, especially if regulatory structures are weak. Second, that there are substantial improvements that could be made to reactor accident safety (even if we assume that boneheaded behavior like Chernobyl will not recur). Both TMI and Fukushima arose from failure of active cooling systems that were needed to keep the core from melting after shutdown. Newer designs aim to be able to remove the decay heat with *passive cooling,* a mechanism that would function without any power at all, removing the need for enormously complex backup systems that drive up costs.

Even if we still have a TMI or Fukushima-scale accident every 30 years, we should ask whether the damage from them is worse than the damage incurred from fossil-fuel alternatives. Will there be more cancers from Fukushima than from coal pollution and coal miners' lung diseases? Is the land that had to be evacuated around Chernobyl more than has been despoiled by surface mining or fracking? Will the effects of the nuclear accident ever come anywhere close to matching the other economic and human casualties of the tsunami? And most important, are these localized pollution problems—or potentially more severe future accidents—worse than the global climate change that we will face if we rely only on fossil fuels?

Economics

The first grid-connected nuclear power plant opened in the Soviet Union in 1954, and the first demonstration in the US was in Shippingport, Pennsylvania starting in 1957. Commercial-scale nuclear power plants came online starting in 1960. As of 2019 there are 94 commercial nuclear reactors in operation in the US, with a total capacity of 97 GW, i.e. an average reactor has 1 GW capacity. They

[10] *Energy Environ. Sci.*, 2013, **6**, 1042

operate with an average capacity factor over 90%, meaning that they are very well-run and reliable—in this respect the industry has made great strides: capacity factors were $< 60\%$ in 1985. While nuclear reactors are only 10% of total US generating capacity, they produce 20% of electricity. The oldest operating reactor was completed in 1969, and several have now been retired.

Nuclear power grew from 1% of total US electricity in 1970 to 20% in 1990, but has changed little since then. The newest operating nuclear plant in the US opened in 2016 in Tennessee; other than this, the newest opened in 1996. All began construction before 1977 (even the one finished in 2016! Construction was suspended 1988–2007). Why did the US stop building new reactors? Part of the reason might be because of public fear after the TMI accident. But the the most important reason is that *nuclear electricity is very expensive.* The increasing complexity of designs, and very frequent cost overruns and design changes, made the cost of electricity for nuclear plants outstrip coal's. Furthermore it was taking 20 years from the first planning of a nuclear plant through design, licensing, and construction into operation. This is too slow for utilities to be able to respond to demand. Hence there has been a drought of > 30 years in US reactor construction. US nuclear electricity production will drop in coming years as older plants need to be retired.

In the mid-2000's there was talk of a "nuclear renaissance" in the US and the first new license applications were filed in decades under a streamlined approval procedure. Only one pair of reactors (Vogtle, in in Georgia) remains under construction. Others were abandoned in mid-construction. Falling natural gas prices made nuclear plants less competitive, even though there are investment tax credits, loan gaurantees, and other subsidies in place for nuclear power in the US. Furthermore the possibility of any taxes or limits on carbon emission began to seem even more remote. The 2×1.2 GW Vogtle plants in Georgia are now expected to cost \$27 billion, or ≈\$11,000 per kW of capacity.

Can nuclear energy be cost-competitive with coal or natural gas? An examination of COE's by Lazard (Nov 2020) places the low end of the cost of new nuclear electricity in the range \$0.13–0.20/kWh, which is well above the \$0.031–0.042, \$0.026-0.054, and \$0.044-0.073/kWh for unsubsidized PV, wind, and NGCC electricity respectively, and also higher than the \$0.065–0.159/kWh for new coal plants. Quite simply, it cannot compete on costs and does not enter into any US utility company's current planning. It is true that nuclear power is the only carbon-free power source which has a reliable steady power output and can provide a large fraction of our energy needs. But we are reaching the point where PV+battery storage is cheaper than nuclear plants as well.

The US EIA is far more optimistic about the costs of new nuclear plants, estimate CoE of about $0.07/kWh. But this strains credulity given the recent history of costs in the US.

One thing to note from the Lazard report is that the CoE from an *existing* nuclear plant is very low—$0.029/kWh. Once the capital costs of a nuclear plant are already sunk, it is one of the most cost-effective sources of electricity. We can expect the current nuclear power plants to continue contributing carbon-free energy until such time as they require expensive upgrades or maintenance.

There are many proposals for nuclear reactor designs that could offer great improvements in safety, cost, reliability, anti-proliferation, and ease of construction over the present BWR/PWR reactors. For instance it should be possible to build a ***molten salt reactor*** in which:

- The fuel is ^{233}U, which can be bred from naturally abundant thorium, separated chemically (instead of requiring isotopic enrichment), but which is harder to turn into bomb material.
- The fuel is dissolved into molten salt which is circulated through the reactor. Because the fuel is in liquid form, one can continuously chemically purify the fuel to avoid the buildup of long-lived radionuclides (like plutonium) that are the biggest waste-disposal hazards.
- The reactor is designed to expand when it warms up, which means that it becomes naturally stable without control rods.
- A meltable salt plug at the bottom of the reactor will melt if the reactor overheats, allowing the liquid salt fuel to spill out into a large shallow "pond" where it can cool safely with air—no pumps or electricity required.
- Because there is no fuel assembly to melt, the reactor can operate at higher temperatures and generate electricity more efficiently.

However only small demonstration reactors of this type have been built. Undoubtedly there would be much engineering development necessary to build a commercial-scale reactor of this or other fundamentally new designs. No utility will undertake the building of a nuclear plant that is expected to cost even more than current ones, so this development will only take place with primarily government funding, and would certainly cost tens of billions of dollars. This is still a small sum, however, compared to energy spending. Even if someone were willing to pay for this development, it would probably take a decade and then another decade or more before large-scale deployment would occur. And of course the costs are not known.

Worldwide nuclear power

Nuclear power generation is not as stagnant worldwide as it is in the US, although the US remains the world's largest producer of nuclear electricity. As of April 2021, the World Nuclear Association reports 443 operating reactors and 54 under construction (17 of them in China, 5 in India). Worldwide nuclear electricity production in 2019 was about 9.6 EJ, 10.1% of world electricity generated according to the above source. Some countries have chosen to produce a much larger fraction of their electric needs from nuclear energy: 72% of France's electricity supply is nuclear (in 2018). Japan and Germany were each $\approx$ 30% nuclear but then shut down nearly all and half of their nuclear capacity, respectively, in response to the Fukushima accident. This is of course a giant step backwards for CO_2 emission. In 2019, Japan is getting only 7.5% of its electricity from nuclear energy, and Germany 12%.

Nuclear fuel cycles and waste disposal

A primary concern over nuclear energy is the ***fuel cycle:*** where does uranium come from and is there enough to sustain a significant expansion of nuclear power? Where will the nuclear waste products go?

Fuel supply

Nuclear energy is incredibly dense. Consider a 1 GW nuclear reactor operating at 25% efficiency and near-100% capacity factor. Its annual energy needs are for $4\,\mathrm{GW} \times 1\,\mathrm{yr} = 1.3 \times 10^{17}\,\mathrm{J}$ of input energy. According to Einstein, this is equivalent to just $m = E/c^2 = 1.4$ kg of mass. Since 0.1% of mass is converted to energy in ^{235}U fission, we'd need 1400 kg of ^{235}U. If the fuel is enriched to 3% ^{235}U, this means the power plant must put about 45 tons of LEU into the reactor each year.[11]

Recall that the LEU requires significant isotopic enrichment from natural uranium's 0.7% level of ^{235}U. This is expensive (and only the government is allowed to do it!). Nonetheless the fuel costs are a very minor part of the cost of electricity for nuclear plants. This tells us that uranium is not yet scarce (even though its price fluctuates rapidly). Known high-quality deposits of uranium are likely sufficient for at least a decade or so of production at current rates.

Even if natural uranium were to become scarce, *a nuclear reactor can produce more fissile material than is put into it!* As mentioned earlier, when (abundant) ^{238}U is bombarded with neutrons, the fissile isotope ^{239}Pu is produced. We can also produce fissile ^{233}U from

[11] A coal power plant of the same capacity uses 45 tons of coal every few minutes!

abundant thorium. When fuel rods are removed from reactors, they are depleted in ^{235}U but we can ***reprocess*** the fuel to extract the fissile plutonium for use in the next generation of fuel rods. Simple chemical processes can do this, no isotopic enrichment is needed. The US has elected *not* to reprocess commercial reactor fuel, because (a) new uranium is sufficiently abundant for now, and (b) we want to minimize the amount of plutonium in circulation to lessen the possibility of diversion for weapons use.

One can design a ***breeder reactor*** especially for producing new fissile material. With breeder reactors, we could produce enough fissile material from non-fissile uranium and thorium to power many generations' worth of nuclear power plants. The US shut down its breeder-reactor program in 1983. Most, but not all, of the other countries that have attempted breeder reactors have shut them down as well. But we know that running out of uranium is not really a barrier to a sustainable nuclear energy supply.

Waste disposal

When fuel rods are inside the reactor, they accumulate many radioactive isotopes in the decay chains of U and Pu as well as isotopes produced by neutron transmutation. The fuel rods are intensely radioactive; while their activity levels drop rapidly, they and hazardous to humans for thousands of years. Figure 16.9 shows that the activity level drops continuously with time. Recall that the very short half-life isotopes decay away quickly and are not a problem; and the isotopes with million-year half-lives decay too slowly to be hazardous. The long-term hazards are the ***actinide*** isotopes that have half-lives of hundreds or thousands of years. These must be kept isolated from human contact for many times longer than human civilization has existed!

There is currently no long-term storage facility for high-level radioactive waste in the US. The Yucca Mountain site in Nevada has been developed for this, but so far local opposition has kept it from being used. While the high-level waste is solid, and you might think you could just seal the material into a cave somewhere, there is a worry that if water flows through the site, the containers might rust through over the eons and the waste could contaminate the water supply in a large region. Some claim the Yucca Mountain site may eventually have water flow. In any case Yucca Mountain's capacity will in fact be too small for high-level wastes that accumulate within another decade or so. One court decided that the facility must be certified as safe for up to a million years, as opposed to the 10,000 time period that DoE was designing and certifying. It is quite bizarre to think of trying to cer-

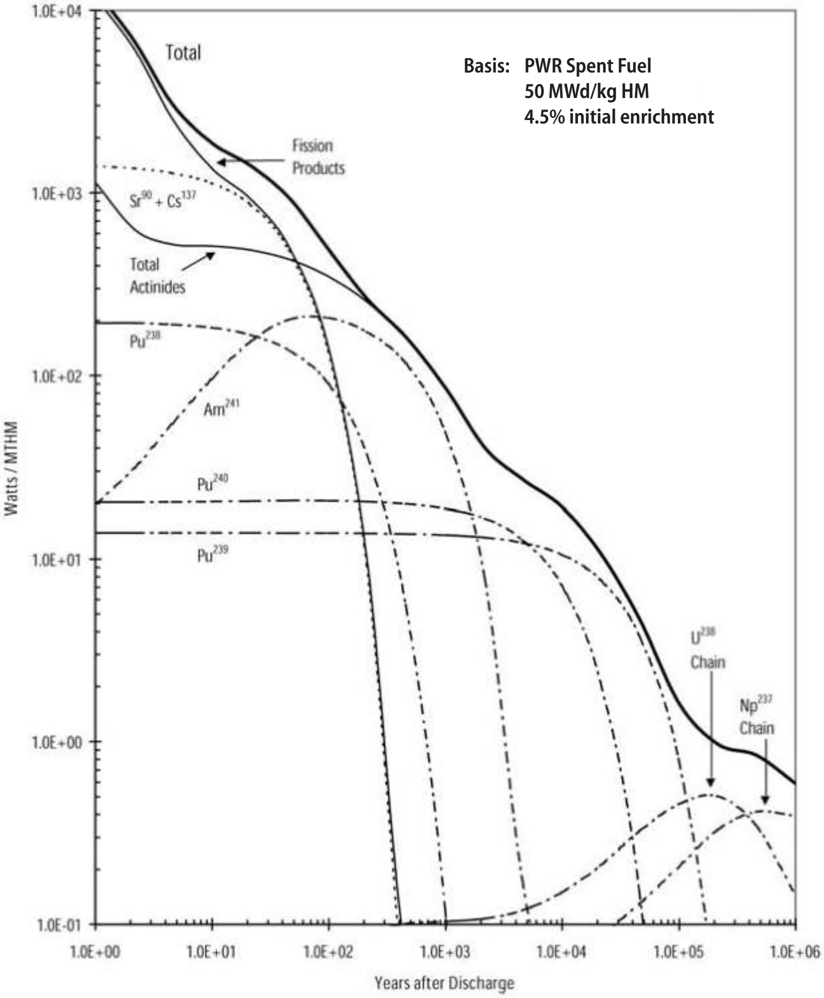


Figure 16.9: The radioactivity of spent fuel rods as a function of years after they are removed from the reactor. The y axis is in Watts of radioactivity power per metric ton. From the MIT *Future of Nuclear Power* report.

tify the quality of something on a time scale longer than the presence of the human species on the planet, never mind hundreds of times longer than the lifetime of any responsible form of government in human society. Particularly in the case where global warming, a near certainty in this century, has no legal constraints at all in this country, it seems odd to demand that the nuclear industry come up with a million-year solution.

Meanwhile each nuclear plant is generating something like 35 tons of spent fuel rods every year. These are all being held in "temporary" storage casks on the sites of the reactors. When they are first removed from the reactor, they are kept in pools of water adjacent in the reactor buildings, because the circulating water is necessary for cooling until the activity level drops (Figure 16.10). After a few months or years, the fuel rods are "cool" enough to store in casks without circulating coolant.

The waste disposal issue is very real. It is, however, possible to avoid the issue altogether. It is only a small fraction of the spent fuel, the actinides, that will cause the long-term radioactivity of the waste. Ironically we can solve this problem by removing the actinides and putting them *back into a nuclear reactor,* where the intense neutron flux would transmute these into shorter-lived isotopes—we can "burn up" the most hazardous substances. This would require investment in fuel reprocessing and in specialized reactors for waste disposal, but again it is possible to remove this barrier to sustainable nuclear fission.

Summary

Nuclear fission is a proven zero-carbon method for providing electricity, and the only one that can provide base load. But it is essentially stalled in the US because nuclear cannot compete with natural gas at current prices. With PV prices dropping, there also seems a good chance that PV+storage would also be a cheaper, lower-risk, faster-building alternative source for carbon-free electricity compared to nuclear.

Transitioning the bulk of our base-load electricity or total-energy needs to nuclear power for the long term will require the economic and political capital to overcome the significant hurdles of accidents, fuel supply, and waste disposal. All three of these, as well as the high cost of nuclear reactors, could be significantly helped by development of better reactor designs than the current ones that are basically the descendants of a 1950's design for nuclear submarines. We will need eventually to move away from our current "once-through" nuclear fuel cycle, to one in which spent fuel is reprocessed to generate new

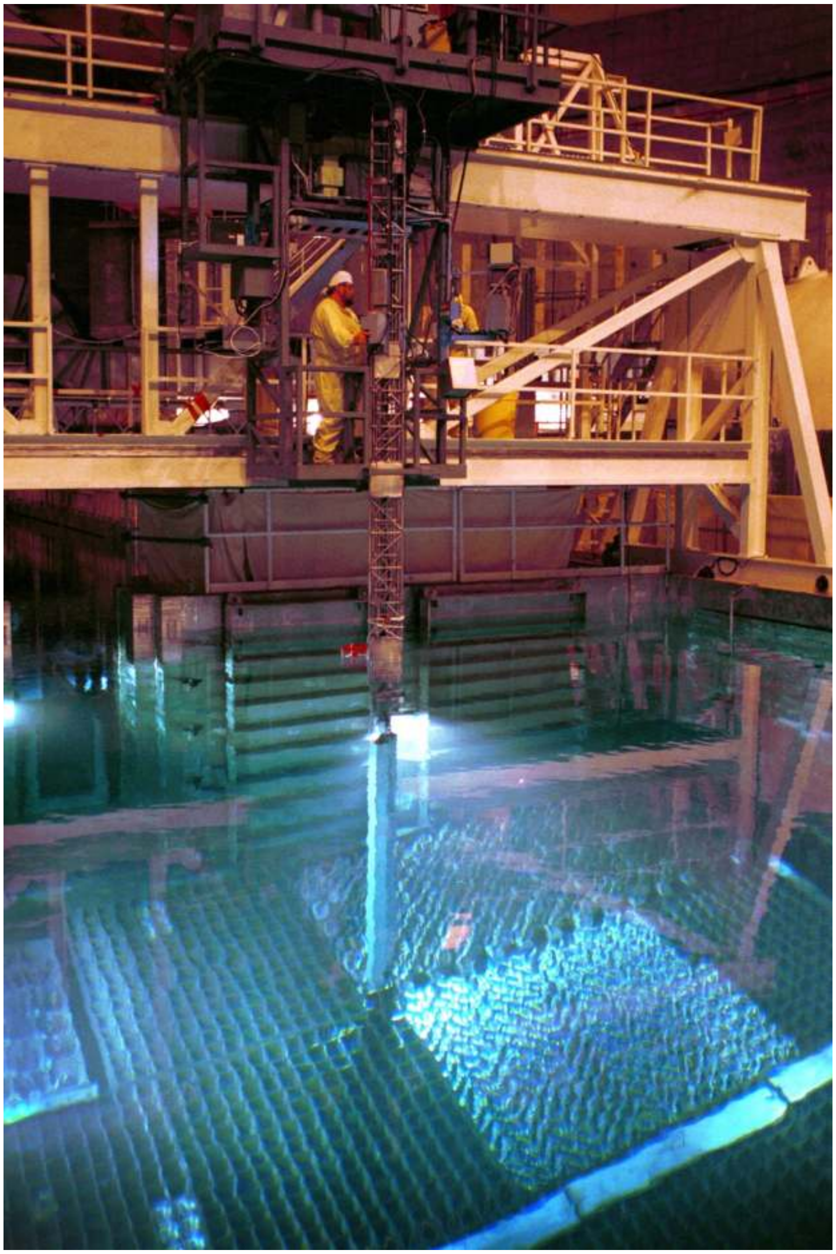


Figure 16.10: A pool containing spent nuclear fuel rods. The beautiful blue glow is produced by radiation moving through the water. From ???

fissile material and to destroy long-lived waste products. This will not be cheap or easy, and it will need to be done carefully to avoid accidents and diversion of nuclear materials to weapons, but it can be a reliable and sustainable source of much more than the 20% of electricity it currently provides. Many scientists and environmentalists believe that nuclear power is the only feasible means to replace our fossil-fuel-based energy economy in this century, or at least an essential component of any feasible plan. Wind and solar are equally sustainable as nuclear and much cheaper and easier now, but would need large-scale storage technology in order to avoid the need for nuclear energy in a decarbonized future.

Practice problems

- Calculate the cost of nuclear electricity from a plant that costs \$6000 per kW of capacity, runs at 90% capacity factor, and is financed for 25 years at 8% interest. Assume that fuel costs and O&M are minimal.
- Uranium oxide (U_3O_8) is currently (2018) selling for about \$65/kg (almost of this mass is in the uranium). Calculate how many kg of uranium oxide must be purchased to supply the LEU for a 1 GW nuclear reactor for a year. Assume the reactor power plant operates at 30% efficiency. Then calculate the contribution of the raw uranium costs to the cost of electricity from the plant.