Building-Integrated Solar Energy Systems

Robert E. Parkin

CRC Press
Taylor & Francis Group
Building-Integrated Solar Energy Systems
Contents

About the Author

Preface

1 Energy Sources, Energy Uses, and Global Warming
   1.1 History of Fossil Fuels
   1.2 Composition of Fossil Fuels
   1.3 Fossil Fuels, Uses and Reserves
   1.4 Environmental Costs of Fossil Fuels
   1.5 Energy Usage and Income
   1.6 Planet Earth
   1.7 Direct Solar Energy
   1.8 Indirect Solar Energy
      1.8.1 Winds of the World
      1.8.2 Wave Power
      1.8.3 Hydroelectric Power
      1.8.4 The Biomass and Biofuel Potential
   1.9 Other Energy Sources
      1.9.1 Tidal Power
      1.9.2 Geothermal Energy
      1.9.3 Nuclear Power
   1.10 Energy Use in the United States
      1.10.1 Energy in the Industrial Sector
      1.10.2 Transportation in the Modern World
      1.10.3 The Commercial and Residential Sectors
      1.10.4 Food Production
   1.11 Global Warming
      1.11.1 Greenhouse Gases
      1.11.2 The Effects of Greenhouse Gases

2 The Internal Environment of a Residence
2.1 Electrical Use and Its Contribution to Sensible Heat
2.2 Human Needs and Demands in a Residence
2.3 Temperature Levels
2.4 Ventilation Guidelines
   2.4.1 Air Changes per Hour
   2.4.2 Air Velocity
   2.4.3 Relative Humidity
   2.4.4 Air Infiltration
   2.4.5 Air Flow through Exterior Doors
   2.4.6 Respiration and Carbon Dioxide
2.5 Heat Recovery Ventilation
2.6 Humidity Sources in a Residence
2.7 Sick Building Syndrome and VOC Mitigation
2.8 Illumination and Daylighting
2.9 Domestic Water Usage
2.10 Electrical Fixtures and Appliances in a Residence
   2.10.1 ENERGY STAR Appliances
   2.10.2 Electrically Generated Lighting
2.11 Noise Control
2.12 Return on Investment and Service Life

3 Heat Flow from a Residence
3.1 Thermal Terms
3.2 Empirical Nature of Heat Loss Calculations
3.3 Fundamentals of Glazing Losses
3.4 Heat Flow Across an Air Gap
3.5 Thermal Resistance R and Conductance C
3.6 Windows
   3.6.1 Heat Loss Through Windows
   3.6.2 Spacers and Their Influence on Heat Loss
3.7 Heat Losses from a Residence
   3.7.1 Heat Losses through Walls and Roofs
   3.7.2 Heat Losses through Structural Integrated Panels
   3.7.3 Heat Losses from Basements and Foundations

4 Residential Construction Techniques
4.1 Fire, Civilization, and Simple Dwellings
4.2 Construction Techniques for Exterior Walls
   4.2.1 Wooden Walls
4.2.2 The Development of Plywood
4.2.3 Bamboo and Other Grasses
4.2.4 Stud Wall Construction
4.2.5 Structural Integrated Panels

4.3 Masonry Walls
4.3.1 Brick Wall Construction
4.3.2 Poured Concrete Foundations and Walls
4.3.3 Concrete Block Walls
4.3.4 Rock Construction
4.3.5 Earthen Materials
4.3.6 Insulated Concrete Formwork
4.3.7 The AAC Block Wall

4.4 Insulation, Types and Installation Methods
4.4.1 Fiberglass Batts
4.4.2 Cellulose Insulation
4.4.3 Urea Formaldehyde Foam
4.4.4 Foam and Foam Board
4.4.5 Multi-Layer Radiant Barriers
4.4.6 The EIFS Insulation System

4.5 Roofs, Construction and Problems
4.5.1 Roof Construction Techniques
4.5.2 Unwanted Heat and Humidity Flow through Roofs
4.5.3 Roof Problems

4.6 Sealing and Insulating Basements

4.7 Water Damage and Vapor Migration in a Residence

4.8 Air Infiltration
4.8.1 Air Leakage from Conditioned Spaces
4.8.2 Air Leakage around Windows
4.8.3 Air Leakage around Electrical Outlets, Switches, and Lights

4.9 Energy Codes and Their Effect on Construction Methods
4.10 Errors in Energy Codes
4.11 Insulation Guidelines and Their Influence on Construction

5 The Seasons and Solar Angles
5.1 Obliquity and the Seasons
5.2 The Importance of Time
5.3 Longitude
5.4 The Julian Calendar
5.5 Geometry of the Ellipse and Kepler
5.6 Points, Vectors, and Planes
5.7 Rotation of Vectors
5.8 The Obliquity Vector
5.9 Angle of the Sun on the Earth’s Surface
5.10 Angle of the Sun on Solar Collectors
5.11 The Solar Collection Day
5.12 Change in Solar Angle With Time
5.13 Seasonal Angles
5.14 Equation of Time and Sundials
5.15 The Analemma and How to Capture It
5.16 Approximating the EoT

6 Transmission of Light through the Atmosphere
6.1 The Incoming Solar Radiation
6.2 Solar Radiation in the Atmosphere
6.3 Air Mass and Dimming
6.4 Rayleigh Scattering
6.5 Mie Scattering
6.6 Solar Radiation Measures
6.7 Instruments for Measuring Solar Radiation
   6.7.1 Pyranometers
   6.7.2 Other Radiation Measurement Instruments
6.8 Isotropic and Circumsolar Radiation

7 Solar Gain and Solar Attenuation
7.1 Solar Fraction and Solar Gain
7.2 Albedo and Specularity
7.3 Air Temperature, Pressure, and Clouds
7.4 Models for Radiation Dispersion in the Atmosphere
7.5 Solar Radiation Data for Boston
7.6 Clear Sky Radiation Data for U.S. Cities
7.7 Solar Radiation Maps of the World

8 Transmission of Solar Energy through Glazing
8.1 A Brief History of Glass
8.2 Transmissivity of Diathermanous Materials to Light
8.3 The Polarization of Light
8.4 Fresnel’s Equations
8.5 Extinction Coefficient and Absorptivity of Glass
8.6 Transmission, Reflection, and Absorption
8.7 Transmission and Absorption for Normal Radiation
8.8 Daily Solar Gain through Glazing
8.9 Transmissivity through LoE Glass
  8.9.1 Standard LoE Glass
  8.9.2 New Generation LoE Glass

9 Climate and the Siting of a Solar Structure
  9.1 The Köppen Climate Classification
  9.2 Climates Zones by the Number
  9.3 Monthly Averages for Heating and Cooling
  9.4 Other Data Sources for HDD and CDD
  9.5 Southern Exposure
  9.6 Magnetic Declination
  9.7 Visibility Distances
  9.8 Design With Nature
  9.9 Clearances on Flat Sites
    9.9.1 Individual Residences
    9.9.2 Communities of Solar Homes
  9.10 Clearances on Sloping Sites
  9.11 Wind Protection

10 Solar Structures from Early to Modern
  10.1 Double Glazing
  10.2 Historic Use of Solar Heating
  10.3 Solar Heating for Residences in the United States
  10.4 Solar Heating for Residences in the Rest of the World
  10.5 Efficient Building Envelopes
  10.6 The Passivehaus Design

11 Passive Solar Collection
  11.1 Solar Gain on Vertical Collectors with Non-Zero Yaw Angles
  11.2 Solar Gain Through Glass
  11.3 Shading
  11.4 Absorption of Glass — Not All Is Lost
  11.5 Solar Blinds
  11.6 Thermal Storage
    11.6.1 Direct Storage
    11.6.2 Rock Storage
11.6.3 Trombe Walls
11.6.4 Water Storage
11.6.5 Eutectic Salts
11.7 Solar Stills
11.8 Solar Cookers and Ovens
11.9 Solar Water Heaters

12 Non-Concentrating, Active Solar Collectors
12.1 Solar Gain on Sloping Surfaces
12.2 Air Collectors
12.3 Liquid Carrying Flat Plate Solar Collectors
   12.3.1 Trickle Type Collectors
   12.3.2 Closed Loop and Fully Wetted Collectors
12.4 Stratified Water Tanks
12.5 Selective Absorptivity/Emissivity Surfaces
   12.5.1 A Statement of the Problem
   12.5.2 Some Early Results
   12.5.3 Cermat Surfaces
   12.5.4 Selective Coating Films
   12.5.5 Optically Selective Paints and Sprays
12.6 The Bread Box Solar Collector
12.7 Evacuated Tube Collectors
12.8 Heat Pipes

13 Photovoltaic Panels
13.1 Semiconductor Fundamentals
13.2 p-n Junctions
13.3 The First Transistors
13.4 The Field Effect Transistor
13.5 Photovoltaic Cells
13.6 Energy in Photons
13.7 Responses of First through Third Generation PV Cells
13.8 Perovskite Solar Cells
13.9 Building Integrated Photovoltaics
13.10 Temperature Sensitivity
13.11 Concentrating Photovoltaic Arrays

14 Smart Grids, FiTs, and Net Metering
14.1 History of Electrical Utilities
About the Author

Robert E. Parkin received his Ph.D. degree from Imperial College, University of London, in 1965. He has spent most of his professional life in academia. He joined the Electrical Engineering faculty at the University of Massachusetts, Lowell, in 1982, transferred to the Mechanical Engineering Department in 1991, which better represented his research interests in manufacturing theory and robotics, and is presently Professor of Mechanical Engineering. He has held several administrative positions, including President of the Faculty Senate and President of the Massachusetts Society of Professors. He presently serves as Chair of the College of Engineering Personnel Committee, and Chair of the Research and Development Committee of the Faculty Senate. His primary area of research is now solar energy. He is Graduate Coordinator of the Renewable Energy degree programs at UMass Lowell.
The language of science is broad, and needs to be to cover an immense array of subjects: in size from the molecular to the cosmic scale, from the mass of a hydrogen atom to that of the universe; in times from nanoseconds to billions of years; in electrical energy to describe the communication to and from a single cell in a human body to crushing hydrogen atoms into helium to produce electrical power by nuclear fusion; in temperatures from absolute zero to the center of a star.

The metric system was developed at the time of the French Revolution. It established the meter as the unit of length and the kilogram as the unit of mass. Subsequently it added the second as the unit of time and became the MKS system.

Eugene Niemi, a colleague of mine at UMass Lowell, describes the MKS (SI) system as consistently inconsistent. He called it that since it stubbornly clings to 360 degrees in a circle and 24 hours in a day. He also called the English system consistently inconsistent, with its reliance on historically archaic definitions.

What are the anchor units we need to live with? The day is a physical reality, but why describe it as 24 hours? The Earth year is also real at 365.2425 days. \( \pi = 3.141592 \) is also real, as is the exponential \( e = 2.71828 \ldots \), but most other measures are artificial constructs of archaic origin.

There are problems with the MKS system. In particular, as far as the scope of this text is concerned, the units may not be appropriate. The unit of time \( S \) does not work as well as the hour, which even in the decimal-based SI system is 3600 seconds. The electrical power we buy from the utility company is measured in kWh, kilowatt hours. The standard size sheet of plywood in North America and in Europe is 4’ \( \times \) 8’, 4 feet by 8 feet, where the foot length is approximately 0.3048 meters; the Europeans call the same size 2440mm \( \times \) 1220mm (they put the larger dimension first, the opposite to the U.S. convention).

Efforts were made in 1948 to expand the MKS system, and in 1960 the Systeme International d’Unites was established. It now has seven measures, adding the ampere (electrical current), the kelvin (unit of temperature), the mole (unit of substance at the atomic level), and the candela (unit of light intensity). It is now called the SI system. It is decimal based, and larger or smaller units are indicated with a prefix, so cm means centimeter, and km means kilometer.
The language of science is relatively new, and this can lead to contradictions and possible confusion. Luckily, in a confined subject such as solar energy one can focus on the terms needed, and omit other terms that may be indispensable in other branches of science. We will use the tonn (1000 kg), try not to use the long ton (2240 pounds or 1016 kg) or the short ton (2000 pounds), will use the pascal and the bar for pressure and try and avoid the psi (pound per square inch), will use the watt W and will try (unsuccessfully) to avoid the BTU (British Thermal Unit, 0.2931 Wh), and will rarely use the unit of energy the joule.

What is the underlying philosophy in this text? I like diagrams that convey the maximum amount of information in as simple a way as possible. I could not have prepared my manuscript without MATLAB®. Most of the response curves in this book were rendered using MATLAB. About 100 MATLAB files are on a CD included with this book. Many of the exercises for each chapter can be answered with MATLAB.

Virtually all formulae I use are derived from first principles. I favor a vector based notation, in contrast to the direction cosines that most authors in solar engineering use. The advantages of vectors over direction cosines include simpler notations and far simpler diagrams. Perhaps I gravitated to the vector notation because I could not make sense of the diagrams that others use.

Why building integrated? Often solar installations are far too expensive, and the reason is simple: the system is installed on an existing structure that was constructed with no consideration for a future solar system. For example, a number of companies offer flat plate collectors that come with glazing, insulation, a frame, and a support system, and cost about $700/m^2 installed. However, if the roof structure was designed to accept those flat plates, the glazing is the roof, the frames are the rafters, and the insulation is the insulation of the house, and there is no need for an additional support system. Thus, the additional cost above normal construction costs is for the flat plate collector panels and associated piping, far less than the $700/m^2.

As another example, Fraunhoffer, the German non-profit, recently introduced PV panels that glue on top of existing roof shingles, and need no frame or support system. Installed cost is about $1.50/watt, compared to the typical $4.90/watt.

In the 1970s I designed and built a number of solar structures, mostly single family houses. I favored passive solar at the time, and to some extent still do. It is cost effective and architecturally pleasing. The last house I built, in 2010 for my daughter’s family, has semi-passive solar with a central atrium to collect the heat, and a fan/duct system to push the heated air into a rock storage system beneath the floor. It has 3.8 kW of PV on the roof, and is grid connected. It also has geothermal heating and cooling, an experimental system that experts said would not work: In fact, it works better than anticipated.

The chapter development in this book is intended to be logical. The first chapter defines the energy problem we face. Subsequent to the industrial revolution we used an incredible amount
of energy in three sectors — industrial, transportation, and in buildings. It is the energy use in buildings, particularly residences, that is the central theme here.

If the objective is to use solar energy to produce most, or all, of the energy needed in a residence, one must start by defining what is needed for human comfort (Chapter 2), then delve into heat losses from a structure (Chapter 3). Materials used in residential construction and construction techniques themselves are the subject of Chapter 4.

The vector based geometry of the motion of the Earth around the Sun following Kepler’s laws in Chapter 5 is, I believe, an important contribution to the field of solar engineering. Without loss of generality one can assume that the Sun and the Earth lie in the x-y plane, so then the axis of rotation of the Earth is a known vector that I call the obliquity vector.

I do not know how declination tables were constructed. It seems probable that they are based on celestial observation. There are papers that purport to derive declination, but they do not carry it to a conclusion and an actual table. In Chapter 5 a derivation for declination is presented based on the Kepler’s laws, but there is a discrepancy with the established tables that is troubling. Even further, a similar technique to derive “equation of time” tables finds an even bigger discrepancy with published tables. In this latter case, I have presented evidence of errors in this table, realizing that perhaps this is provocative.

The transmission of light through the atmosphere as defined by Planck’s law and attenuating factors such as oxygen, ozone, and water vapor is the subject of Chapter 6. The solar constant is 1366 W/m², which is reduced to about 880 W/m² at the Earth’s surface. The unanswered problem is to explain where the 486 W/m² went. A partial answer is that it has been scattered, of which Rayleigh scattering is the most important constituent.

Chapter 7 attempts to further identify the factors responsible for the loss of 486 W/m². There is an inverse relationship between beam (direct normal) radiation and diffuse radiation. That is, diffuse is low when beam is high, and diffuse increases and beam decreases as clouds and aerosols increase in the atmosphere until a certain point, and then diffuse decreases until fully occluded conditions occur. A problem with most models for solar attenuation is that they do not provide a rationale for their development. Attempts here to correct this deficiency were not successful, but at least the problem is defined and possibly will be of use in future work, either by me or others.

Chapter 8 provides a comprehensive treatment of the transmission of light through glass. It ends with a discussion of LoE glass, sometimes called low-E or low-e glass, including recent developments. There is a major problem with the requirement in the United States that all window glass be LoE, which is the kiss of death for passive solar collection. SHGC (Solar Heat Gain Coefficient) through LoE glass is low. It makes sense to use LoE glass in windows not facing the equator, since solar gain through them is probably unwanted, but glass facing the equator has little, or no, solar gain in summer but maximum solar gain in winter.
Chapter 9 defines climate and uses three measures. The first is the universal Koppen climate classification based on vegetation, and the second is the climate zone classification used in the United States predominantly to advise gardeners as to which plants will survive. These are gross measures, and have limitations in determining energy flow to and from a residence in a particular location. The third measure is the heating and cooling degree days, which is a more satisfactory measure since there are extensive individual sites that measure temperature daily across the populated world. It was found that a satisfactory way to reduce the random changes in temperature was to use monthly averages for each site and to collate the data over a number of years.

The suitability of a site to accept a solar structure should be ascertained before building commences, and this is included in Chapter 9. In particular, can the structure face the Sun without impediments, such as other structures or trees that are in the neighbor’s yard, blocking the Sun and such that they cannot be removed? The angles and dimensions needed are discussed.

Solar structures are considered in Chapter 10. What makes a house heated by the Sun successful is glass. The large sheet of glass is a recent invention. There are very good architectural and energy reasons to employ passive solar radiation, and these are discussed in Chapter 11. The Passivhaus houses in Germany and Austria have been highly successful, even though the climate is not as welcoming for passive solar heating as are vast swaths of the United States.

Active solar systems for residential use is explored in Chapter 12. Photovoltaic cells from first to third generation are treated in Chapter 13, and that chapter ends with a discussion of perovskite cells that could become the savior of our planet by moving us into a hydrogen economy.

The grid and the connection of renewable energy power generators is discussed in Chapter 14. The present day, utility run, power grid is an anachronism. It has barely changed in 100 years, but with the rise of distributed generation from renewables, such as wind and PV, it is evident to the knowledgeable observer that the investor-owned utility company will not survive in its present form.

Architectural requirements for building integrated solar elements are considered in Chapter 15. The energy storage problem is discussed in the last chapter, with the limitations of battery technology given full voice.

Who will use this book? This is a textbook for seniors or graduate students in engineering, physics, or chemistry. It is too technical for students in humanities or social sciences, or even biology. There are exercises for every chapter. The students in these classes must have access to MATLAB; not a problem as MATLAB is a requirement in most engineering environments. It is also intended to be a reference book for working engineers in the solar field, and for solar system designers and installers.
Coding experts will find much to criticize regarding my programming skills. I was never into reverse Polish notation, although it saves keystrokes on a handheld calculator. In MATLAB I can achieve the result I want, although the code is not pretty. My excuse is that when I press the save and run button, the output is produced in a fraction of a second, and the fact that additional electrons had to work harder does not concern me.

Of all subjects covered, the one that gives me the greatest problem is that of heat transfer. I rely on ASHRAE Fundamentals, and tend to doubt formulae with all those numbers — Nussault, Reynolds, Prandl — that are empirically based.

This book is definitely not attuned to the tropics, or to a Mediterranean climate. Perhaps one could say it is applicable to latitudes 30° to 55°, both south as well as north. And this requires a word of apology to the southern hemisphere reader for the northern-centric language that permeates throughout. I do it for simplicity, and no slight is intended.

Finally, we live and work in a connected world where the Internet is an essential part. No longer do we use printed telephone books, but I do confess that I rely on a printed edition of the Oxford English Dictionary. The ease with which key words enable a browser to instantly bring up the information sought makes an extensive bibliography unnecessary. I trust this does not lead the reader to assume that this work is weak in scholarship. I include essential references at the location needed rather than as a list at the end of each chapter, eliminating the discontinuity of flicking through pages to find reference number 73.

### Units

<table>
<thead>
<tr>
<th>quantity</th>
<th>symbol</th>
<th>equivalent unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>second</td>
<td>s</td>
<td></td>
</tr>
<tr>
<td>hour</td>
<td>h</td>
<td></td>
</tr>
<tr>
<td>year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>meter</td>
<td>m</td>
<td>1 m = 39.37” = 3.281’</td>
</tr>
<tr>
<td>centimeter</td>
<td>cm</td>
<td>1 m = 1000 cm</td>
</tr>
<tr>
<td>kilometer</td>
<td>km</td>
<td>1 km = 1000 m</td>
</tr>
</tbody>
</table>
inch 
foot
mile
light year
area
sq. meter
sq. kilometer
hectare
square feet
acre
volume
liter
cubic meter
gallon (U.S.)
Imperial gallon
cubic feet
mass
kilogram
accel due to gravity
force
newton
dyne
pound force
weight
gram
<table>
<thead>
<tr>
<th>Unit</th>
<th>Symbol</th>
<th>Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>kilogram</td>
<td>kg</td>
<td>1 kg = 1000 grams</td>
</tr>
<tr>
<td>tonne</td>
<td>tonn</td>
<td>1 tonn = 1000 kg</td>
</tr>
<tr>
<td>1 pound</td>
<td>lb</td>
<td>1 lb = 0.453592 kg</td>
</tr>
<tr>
<td>long ton</td>
<td>ton</td>
<td>1 ton = 2240 lb = 1016 kg</td>
</tr>
<tr>
<td>short ton</td>
<td></td>
<td>1 short ton = 2000 lb</td>
</tr>
</tbody>
</table>

**Chemical Weight**

<table>
<thead>
<tr>
<th>Unit</th>
<th>Symbol</th>
<th>Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>mole</td>
<td>mol</td>
<td>6.022x10^{23} molecules</td>
</tr>
</tbody>
</table>

**Density**

<table>
<thead>
<tr>
<th>Unit</th>
<th>Symbol</th>
<th>Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>density</td>
<td></td>
<td>$\rho = \frac{\text{kg/m}^3}{\text{lb/ft}^3}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 lb/ft^3 = 16.03 kg/m^3</td>
</tr>
</tbody>
</table>

**Pressure or Elasticity (Modulus of)**

<table>
<thead>
<tr>
<th>Unit</th>
<th>Symbol</th>
<th>Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>pascal</td>
<td>Pa</td>
<td>1 Pa = 1 N/m^2, 50 Pa = 0.00725 psi</td>
</tr>
<tr>
<td>kilopascal</td>
<td>kPa</td>
<td>1 kPa = 1000 Pa = 0.145 psi = 0.29533”Hg”</td>
</tr>
<tr>
<td>pounds per square inch</td>
<td>psi</td>
<td>1 psi = 6895 Pa</td>
</tr>
<tr>
<td>bar</td>
<td>bar</td>
<td>1 bar = 100 kPa = 1.019716 kg/cm^2 = 14.5038 psi</td>
</tr>
<tr>
<td>millibar</td>
<td>mbar</td>
<td>1 mbar = 10^{-3} bar</td>
</tr>
<tr>
<td>atmospheric pressure</td>
<td>atm</td>
<td>1 atm = 1.013 bar = 14.696 psi</td>
</tr>
</tbody>
</table>

**Speed**

<table>
<thead>
<tr>
<th>Unit</th>
<th>Symbol</th>
<th>Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>speed</td>
<td></td>
<td>$\frac{\text{m}}{\text{s}}, \frac{\text{km}}{\text{h}}, \frac{\text{mph}}{\text{h}}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 km/h = 0.6213 mph</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 mph = 1.6095 km/h</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 c = 2.998x10^5 km/s = 1.863x10^5 miles/s</td>
</tr>
</tbody>
</table>

**Angle**

<table>
<thead>
<tr>
<th>Unit</th>
<th>Symbol</th>
<th>Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>angle</td>
<td></td>
<td>$\theta^\circ = \frac{360^\circ}{\pi} \text{ radians}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>360° = 2\pi rad</td>
</tr>
</tbody>
</table>

**Rotational Speed**

<table>
<thead>
<tr>
<th>Unit</th>
<th>Symbol</th>
<th>Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>rotational speed</td>
<td></td>
<td>$\frac{\text{rad}}{\text{s}}$ = 1 rad = 180/\pi degrees</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 rpm = 0.10472 rad/s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>revolutions per minute</strong></td>
<td><strong>rpm</strong></td>
<td></td>
</tr>
<tr>
<td>1 rpm = 0.10472 rad/s</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>power</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>watt</td>
<td>W</td>
</tr>
<tr>
<td>kilowatt</td>
<td>kW</td>
</tr>
<tr>
<td>megawatt</td>
<td>MW</td>
</tr>
<tr>
<td>gigawatt</td>
<td>GW</td>
</tr>
<tr>
<td>terawatt</td>
<td>TW</td>
</tr>
<tr>
<td>horsepower</td>
<td>hp</td>
</tr>
<tr>
<td>1 W = 1 N*m/s = 1 J/s</td>
<td></td>
</tr>
<tr>
<td>1 kW = 10^3 W</td>
<td></td>
</tr>
<tr>
<td>1 MW = 10^6 W</td>
<td></td>
</tr>
<tr>
<td>1 GW = 10^9 W</td>
<td></td>
</tr>
<tr>
<td>1 TW = 10^12 W</td>
<td></td>
</tr>
<tr>
<td>1 hp = 550 ft.lb/s = 746 W</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>energy</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>joule</td>
<td>J</td>
</tr>
<tr>
<td>exajoule</td>
<td>EJ</td>
</tr>
<tr>
<td>calorie</td>
<td>cal</td>
</tr>
<tr>
<td>kilocalorie</td>
<td>kcal</td>
</tr>
<tr>
<td>watt hour</td>
<td>Wh</td>
</tr>
<tr>
<td>kilowatt hour</td>
<td>kWh</td>
</tr>
<tr>
<td>foot pound</td>
<td>ft.lb</td>
</tr>
<tr>
<td>exajoule</td>
<td>EJ</td>
</tr>
<tr>
<td>terawatt year</td>
<td>TWy</td>
</tr>
<tr>
<td>electron volt</td>
<td>eV</td>
</tr>
<tr>
<td>British Thermal Unit</td>
<td>BTU</td>
</tr>
<tr>
<td>quadrillion BTU</td>
<td>quad</td>
</tr>
<tr>
<td>1 J = 1 N.m = W<em>s = 1 kg</em>m^2/s^2</td>
<td></td>
</tr>
<tr>
<td>1 J = 4.187 J = 0.00116 Wh = 3.97 x 10^{-3} BTU</td>
<td></td>
</tr>
<tr>
<td>1 kcal = 1000 cal</td>
<td></td>
</tr>
<tr>
<td>1 Wh = 3.412 BTU = 0.8598 kcal</td>
<td></td>
</tr>
<tr>
<td>1 kWh = 3.412 BTU = 0.8605 x 10^6 cal</td>
<td></td>
</tr>
<tr>
<td>1 ft.lb = 1.355817 J</td>
<td></td>
</tr>
<tr>
<td>1 EJ = 1018 J = 0.9478 quad = 2.77 x 10^{11} kWh</td>
<td></td>
</tr>
<tr>
<td>1 TWy = 8.76 x 10^{12} = 29.89 quad</td>
<td></td>
</tr>
<tr>
<td>1 BTU = 0.2931 Wh = 252 cal</td>
<td></td>
</tr>
<tr>
<td>1 quad = 1015 BTU = 0.03346 TWy = 1.055 EJ</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>energy density</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>joules/gram</td>
<td>J/gram</td>
</tr>
<tr>
<td>kilowatt hours/kilogram</td>
<td>kWh/kg</td>
</tr>
<tr>
<td>watts/meter^2</td>
<td>W/m^2</td>
</tr>
<tr>
<td>1 J/gram = 0.4299 BTU/lb</td>
<td></td>
</tr>
<tr>
<td>1 kWh/kg = 0.4299 BTU/lb</td>
<td></td>
</tr>
<tr>
<td>1 W/m^2 = 0.317 BTU/ft^2*h</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>temperature</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kelvin</td>
<td>°K</td>
</tr>
<tr>
<td>0° K is absolute zero</td>
<td></td>
</tr>
</tbody>
</table>
**celsius** °C 
0° C = 273.15 °K = 320 °F

**Rankin** °R
0° R is absolute zero

**fahrenheit** °F
0° F = 459.67 °R

**specific heat capacity** 1 cal/gram·°C = 1 BTU/lb·°F = 4.187 J/gram·°C

In SI units
thermal conductance is measured in W/°C over 1 m² area for the specified thickness
thermal resistance is measured in °C/W over an area of 1 m² for the specified thickness
thermal conductivity is measured in W/°C over 1 m² area for 1m thickness
thermal resistivity is measured in °C/W over an area of 1 m² for 1m thickness

In Imperial units
thermal conductance is measured in BTU/°F over an area of 1 ft² for the specified thickness
thermal resistance is measured in °F/BTU over an area of 1 ft² for the specified thickness
thermal conductivity is measured in BTU/°F over an area of 1 ft² for 1” thickness
thermal resistivity is measured in °F/BTU over an area of 1 ft² for 1” thickness

\[
C_{si} = 6.933C_{Imp} \quad C_{Imp} = 0.1442C_{si} \quad R_{si} = 0.1442R_{Imp} \quad R_{mp} = 6.933R_{si}
\]

solar constant 1366 W/m² = 433 BTU/ft²·h
solar radiation is measured in W/m²

**SI Units Versus English Units**

SI (Système International) units are the accepted, scientific measurement system. But even within the SI world there is a confusion between force and weight. Newton’s second law of motion is, where is force, is mass, and is acceleration. Mass in SI units is measured in Kg (kilograms), and force is measured in newtons, where on the Earth’s surface 1 Kg is subject to a gravitational force of 9.81 newtons.

On the surface of the Earth the acceleration due to gravity is taken as 9.81 m/s², or 32.174 ft/s², plus or minus 0.5 percent. Earth’s Moon has about 1/6 of the gravitational acceleration as on Earth, so a 1 Kg mass on the Moon is subject to a force of 9.81/6 = 1.635 newtons.

One does not go into a store to purchase 8.81 newtons of cheese. The scale on which cheese is weighed is calibrated in kilograms, not newtons. Although scientifically incorrect, this is the accepted protocol. In English units, the unit of weight is the pound (lb), where 1 Kg = 2.2 lb.
The English unit of force is designated as lb, and the unit of mass is the slug, where 1 slug weighs 32.174 pounds. Whereas the kilogram is a precisely defined unit, the pound is not. The avoirdupois pound contains 16 ounces, but another measure of weight is the troy system, where 1 troy ounce = 0.8229 avoirdupois ounce. The troy system is used for precious metals such as gold. Troy calculates 12 ounces/pound, compared to 16 ounces/pound in avoirdupois.

In order to avoid any confusion, the only mass that will be considered is the kilogram. In discussing weight on Earth both the kilogram and the pound will be used, among other measures.

**Useful Trigonometric Formulae**

\[
\begin{align*}
\cos^2 \theta + \sin^2 \theta &= 1 \\
\cos 2\theta &= 2\cos^2 \theta - 1 = 1 - 2\sin^2 \theta = \cos^2 \theta - \sin^2 \theta \\
\sin^2 \theta &= 2\sin \theta \cos \theta \\
\sin(\alpha \pm \beta) &= \sin \alpha \cos \beta \pm \cos \alpha \sin \beta \\
\cos(\alpha \pm \beta) &= \cos \alpha \sin \beta \sin \alpha \cos \beta
\end{align*}
\]

\[
\begin{align*}
2\sin \alpha \sin \beta &= \cos(\alpha - \beta) - \cos(\alpha + \beta) \\
2\cos \alpha \cos \beta &= \cos(\alpha - \beta) + \cos(\alpha + \beta) \\
2\sin \alpha \cos \beta &= \sin(\alpha + \beta) - \sin(\alpha - \beta) \\
2\cos \alpha \sin \beta &= \sin(\alpha + \beta) + \sin(\alpha - \beta) \\
2\sin(\alpha + \beta)\sin \beta &= \cos \alpha - \cos(\alpha + 2\beta) \\
2\cos(\alpha + \beta)\cos \beta &= \cos \alpha + \cos(\alpha + 2\beta) \\
2\sin(\alpha + \beta)\cos \beta &= \sin \alpha + \sin(\alpha + 2\beta) \\
2\cos(\alpha + \beta)\sin \beta &= \sin(\alpha + 2\beta) - \sin \alpha \\
\end{align*}
\]

\[
\begin{align*}
\sin 2\theta &= 2\sin \theta \cos \theta \\
\cos 2\theta &= \cos^2 \theta - \sin^2 \theta = 1 - 2\sin^2 \theta = 2\cos^2 \theta - 1 \\
\end{align*}
\]

Law of sines: \( a \sin \alpha = b \sin \beta = c \sin \gamma \)

Laws of cosines: \( a^2 = b^2 + c^2 - 2bc \cos \alpha \)

\( a = bc \cos \gamma + c \cos \beta \)
Solar and Astronomical Terms Applied to the Earth

**absorptivity**: the ability of a body to absorb solar or thermal energy, where 1 is a perfect absorber and 0 is a perfect reflector: see [blackbody](#).

**air mass**: a number $M$ to help describe attenuation of solar energy as it passes through the atmosphere, where $M = 1$ on a clear day when the [zenith](#) angle is zero.

**albedo**: the reflective property of a non-luminous object, where 0 is a perfect absorber and 1 is a perfect reflector.

**analemma**: the figure-8 path that the Sun follows at a particular solar hour throughout the year.

**angle of obliquity**: 23.45° for the Earth, the angle between the Earth’s rotational axis and the ecliptic plane.

**anisotropic**: solar radiation that is not equal from all directions, excluding [beam radiation](#).

**antipodal point**: point on direct opposite side of Earth (or any planet).

**aphelion**: the point on the Earth’s orbit when it is furthest from the Sun.

**apogee**: point in orbit of the Moon where it is furthest from the Earth.

**apsis**, pl. **apsides**: point of least or greatest distance between two bodies. For the Earth and Sun, the greatest and least distances are termed the **aphelion** and the **perihelion**, respectively.
aurora: the interaction of charged particles from the Sun with the Earth’s magnetic field near the poles to produce dancing, multicolored curtains of light, capable of moving from horizon to horizon at light speed. It is called aurora borealis, or northern lights, visible at latitudes 65–72 north, and aurora australis near the south pole.

autumnal equinox: the date in the year when the Sun crosses the Earth’s equatorial plane from north to south.

azimuth: angular distance around the horizon with respect to south (astronomers use north).

barycenter: the center of mass of the Earth and the Moon.

beam radiation: the direct solar radiation, excluding diffuse and reflected radiation, in W/m².

blackbody: a perfect absorber of solar or thermal energy.

centripetal acceleration: the acceleration towards the center of motion (as occurs when a mass is swung in a circle at the end of a string).

circumsolar: the area within an angle of 2.5° or 3° of the center of the Sun.

coriolis force: the force that occurs on a system with two or more bodies when at least one body is rotating and the other body is moving with respect to the first body. The rotation of the Earth interacting with wind creates a coriolis force that results in hurricanes rotating counter clockwise in the northern hemisphere and cyclones rotating clockwise in the southern hemisphere.

decidation  δ: the relationship between the obliquity vector and the position of the Earth with respect to the Sun on the ecliptic plane.

diathermanous: having the property of transmitting visible light; glass, clear plastics, or water, which can transmit sunlight, are the most common diathermanous materials.

diffuse radiation: solar radiation that excludes beam radiation.

dimming: the apparent diminution of light caused by scattering and absorption, caused by gases and particles in the atmosphere.

eccentricity: the measure of the extent that the orbit (of the Earth about the Sun, or the Moon about the Earth) departs from circularity, where 0 is a perfect circle and 1 is a parabola.

eclipse: the alignment of the Sun, the Earth, and the Moon such that the Moon casts a shadow on the Earth, or the Earth casts a shadow on the Moon.

ecliptic plane: the plane containing the Sun and the orbit of the Earth, defined by the unit vector P=[001] orthogonal to the surface of the plane.

electromagnetic spectrum: the frequency spectrum of electromagnetic radiation (audio, radio, radar, visible light, x-rays, gamma rays to cosmic rays)

emissivity: the ability of a body to radiate energy, where 1 is a perfect radiator.

ephemeris table: table of values, usually position, for astronomical bodies as a function of time.

equation of time EoT: the deviation of longitude noon from solar noon.

equinox: the two times a year when the declination of the Earth is zero and all points off the Earth receive 12 hours of daylight.

extinction: the apparent dimming of the Sun when low to the horizon.

first point of Aries: place in the sky where the Sun passes the equator at the vernal equinox.
**geodetic:** referring to a feature on the Earth’s surface, such as latitude or longitude.

**global radiation:** the total energy striking a surface in W/m².

**global horizontal radiation:** the total energy striking a horizontal surface.

**gnomon:** the triangular blade that casts a shadow on the horizontal surface of the **sundial**.

**gravitational constant:** \( G = 6.67 \times 10^{-11} \), the universal constant in the formula determining the gravitational attraction between celestial bodies.

**gravity:** the mutual attraction of large bodies such as the Earth, Sun, and Moon.

**greenhouse effect:** the rise in the Earth’s temperature caused by the increase in greenhouse gases such as carbon dioxide due to the burning of fossil fuels.

**GMT:** Greenwich mean time, the clock time (standard, not summer) for Greenwich, England, which is at longitude zero.

**heliocentric:** centered on the Sun.

**heliostat:** an orientable, large, flat mirror that focuses solar energy on a central receiver; usually deployed with a large number of mirrors focused on the top of a tall tower.

**infrared meter:** the meter, usually hand held, that can measure air leakages and cold areas from the inside of a structure.

**infrared radiation:** the wavelengths of light greater than about 0.7μ (microns, or \(10^{-6}\) meters); not visible to the human eye.

**insolation:** solar radiation received on a unit area surface, in W/m².

**isotropic:** equal in all directions, and applied here with respect to solar radiation.

**Julian angle:** angle \( jn360365.25Jn \).

**Julian day:** the day of the year with \( J_1 \) being January 1.

**Kepler’s laws:** The laws that describe the motion of the Earth around the Sun, with the first two laws being:

1. The Earth orbits the Sun in an elliptical orbit with the Sun at one focus,
2. The area of the elliptical orbit swept by the Earth in a given time is constant.

**latitude L:** the line on the surface of the Earth at a fixed angle from the equatorial plane, where \( \theta = 0^0 \) is the equator and \( \theta = \pm 90^0 \) are the poles.

**latitude inclined:** a surface facing south (in northern latitudes) at an angle to the horizontal the same as the latitude.

**light year:** the distance light travels in one year, approximately \( 9.5 \times 10^{12} \) km or \( 5.85 \times 10^{12} \) miles.

**longitude:** an arc on the Earth’s surface cut by a plane intersecting the axis of the Earth.

**meridian:** arc on the Earth’s surface from north to south pole with the same longitude.

Greenwich observatory in England is, by definition, on meridian zero. Boston, Massachusetts is just west of the 71st meridian.

**micron μ:** meters, the standard unit when discussing wavelengths of light.

**Mie scattering:** the scattering of a photon as it passes close by a particle in the atmosphere that has about the same size as the wavelength of the photon.

**millibar:** a measure of atmospheric pressure, 1013 millibars at sea level.
**nutation**: the slight wobble of the axis of the Earth caused by the gravitational effects of the Moon and the Sun.

**obliquity vector**: the unit vector $q = [0.3860, -0.0969, 0.9174]$ defining the Earth’s axis of rotation with respect to the **ecliptic plane** $p$.

**penumbra**: the area of illumination caused by a partial **eclipse**.

**perigee**: point in the orbit of the Moon when it is closest to the Earth.

**perihelion**: the point on the Earth’s orbit when it is closest to the Sun.

**photon**: the theory that describes light as comprised of massless bundles of energy called photons, which have a range of wavelengths described in **Planck’s law**.

**Planck’s Law**: the universal law that defines the spectral flux emitted by a radiating body.

**precession of the equinoxes**: change in orientation of the axis of the Earth by 50.3 seconds/year, so a complete precession cycle takes about 25,765 years.

**prime meridian**: the meridian passing through Greenwich, England, defining longitude zero.

**pyranometer**: a device to measure **global** solar radiation.

**pyroheliometer**: a device to measure beam solar radiation.

**Rayleigh scattering**: the scattering of a photon as it passes close by a particle in the atmosphere that is about 1/10 the size of the wavelength of the photon; the extent of the scattering is a function of the wavelength of the photon to the power minus four, and results in the sky appearing blue.

**refractive index**: the ratio of the speed of light in a substance to the speed of light in a vacuum, a number always greater than 1; it is a measure of the amount a beam of light is bent when leaving one substance and entering another.

**right ascension**: the position of a celestial object, or a point on the surface of the Earth, as seen from the center of the Earth, as measured from the **first point of Aries**.

**shading coefficient**: the fraction of the solar energy passing through a glazing system compared to 3 mm clear float glass: now an archaic term.

**SHGC**: Solar Heat Gain Coefficient, the fractional measure of a window’s ability to block sunlight. SHGC used to be called the solar factor in the U.S.

**sidereal**: relating time or position with respect to the stars.

**sidereal day**: the time for a star to pass a given meridian: 23 hours 56 minutes 4.09 seconds: the Earth rotates 360.98° in 24 hours.

**sidereal month**: the time it takes for the Moon to return to the same apparent position among the stars.

**solar constant**: 1,366 W/m², the amount of solar energy received in space at the Earth’s distance from the Sun by a surface normal to the Sun’s rays.

**solar eclipse**: when the Earth passes into the shadow of the Moon.

**solar fraction**: the fraction of incident solar radiation striking a fixed orientation surface of unit area compared to a unit surface normal to the radiation.

**solar gain**: the clear sky solar gain from solar noon to dawn or dusk: short for half day solar gain; a function of declination and latitude.

**solar noon**: the time when the azimuth of the Sun is zero on a particular day.
**solstice**: the times of the year when the declination is a maximum or a minimum, and the hours of daylight are longest or shortest, respectively.

**specularity**: the reflectivity of a surface in terms of both *emissivity* and coherence (preserving essential features of the incoming energy).

**Stefan-Boltzmann Law**: the law that defines the relationship between the temperature and the energy a blackbody emits.

**steradian**: the solid angle *omega* defined as the area $A$ on the surface of a sphere by a cone with apex at the center of the sphere divided by the radius $r$ of the sphere squared, so $\omega = A/r^2$.

**sundial**: the simplest of clocks, with no moving parts, that uses the shadow cast by the Sun to tell time.

**topocentric**: measuring the Sun’s position with respect to a particular point on the Earth’s surface.

**ultraviolet**: solar energy with wavelengths less than about 0.3 microns that is not visible to the human eye.

**umbra**: the area of total darkness during an eclipse.

**vernal equinox**: the date in the year when the Sun crosses the Earth’s equatorial plane from south to north.

**visible light**: electromagnetic energy visible to the human eye with wavelengths approximately 0.3 to 0.7 micrometers.

**Wein’s displacement law**: the temperature at which the energy emitted from a *blackbody* is a maximum.

**whitening**: the apparent lightening of the sky on the area close to the Sun.

**zenith**: a point directly overhead.

---

**Earth Measures Related to the Sun and the Moon**

Center of Earth to center of Sun average distance: $1.4966899 \times 10^{11} \pm 1.668\% m$. With the average distance from Earth to Sun normalized to unity, the perihelion, the closest distance, is $0.98326622$, and the aphelion, the furthest distance, is $1.016619$.

Center to center of Earth to Moon has average distance $3.84392 \times 10^8 m$, with extrema of $5.564 \times 10^8$ and $4.0670 \times 10^8$.

Radius of Earth, average $6.371 \times 10^6 m$, at equator $6.378 \times 10^6$, through poles $6.3576 \times 10^6 m$.

Radius of Sun $6.955 \times 10^8 m = 4.3 \times 10^5 miles$  
Radius of Moon $1.738 \times 10^3 m$  
Observed radius of Sun on Earth

Masses in Kilograms: Earth $5.97 \times 10^{24}$
Sun $1.99 \times 10^{30}$
Moon $7.36 \times 10^{22}$

The gravitational constant between two masses $m_1$ and $M_2$ at distance $r$ apart is $G = 6.67 \times 10^{-11}$, and comes from the formula $F = G \frac{m_1 m_2}{r^2}$.

Surface temperature of Sun is $6000 \, ^\circ K$

Solar energy on a normal surface on a clear day is about $880 \, W/m^2 = \frac{280}{3600} \, BTU/ft^2 \cdot hr$

Meter — times the distance along a meridian from the pole to the equator passing through Paris. The measure was off by 0.2 mm because the engineers did not take into account the bulging of the Earth at its equator. A competing definition, the length of a pendulum whose half period time is 1 second, lost out.

The English measure, the foot, has no solid foundation like the meter. It is defined as the length of a human left foot from big toe to heel when standing with the weight evenly distributed on both feet.

AU — Astronomical Unit, where $1 \, AU = 1.4859787 \times 10^{11}$ Km, slightly less than the average distance of the Earth from the Sun. It was intended to define the distance from the Sun of a massless body in circular orbit whose time to complete one orbit is 365.242199, the same as the Earth, and is used to define distances within the solar system.

**Basic MATLAB Commands**

**Getting started**

MATLAB executes an m-file, a file such as test.m.
The first lines of an m-file should set the appearance that you want for the output.

`clc` → clears the command window

`clear` → removes items from memory

`format compact` → suppresses excess line feeds, producing a more compact output.

`format short, format long` → sets the length of the numeric output

The use of italics in the lines above is used to identify a section of MATLAB code; there are no italics in the actual code.

The best way to get started in MATLAB is to jump in, create a simple file, and from the pull-down menu Debug press “Save File and Run test.m”, where test.m is the m-file you created.
Don’t worry if it finds a fatal error. The computer will not blow up. Just doing it is the fastest way to learn it.

The output of an m-file is put in a command window. Click on the MATLAB icon on the bottom of the screen and the split screen enables you to click on the command window.

One of the first computers I used during my doctoral research used 80 column hollerith cards, where each card was a line of code. I handed my stack of cards in over a counter, and several hours later my printout would be ready for pickup. I could get in two, possibly three, runs a day. Now, with the modern personal computer, a MATLAB run is semi-instantaneous in most cases. This makes it more convenient to try something out to see if it works, rather than consulting a manual or textbook.

Variables are case sensitive, so Test is different than test. All variables must start with a letter, and can designate a scalar, a vector, a matrix, or a polynomial. The statement in which the variable is used determines its form.

Some words are restricted, such as end, for, exp, and sin, so care must be taken in choosing variable names. When in doubt, put a number at the end of the variable name. end is not a legal variable, but end2 is.

It is good practice to choose variable names that are consistent, at least as far as is practical, with standard notation in engineering mathematics. Use lower case letters for vectors, and upper case for matrices. There is no boldface, subscripts, or superscripts in MATLAB. The i\textsuperscript{th} member of array x is x(i). The first member of array x is x(1), unlike the programming language C, which starts with iterand 0.

Each line of code, even lines with no code in them, are numbered. MATLAB is an interpretive language, meaning it executes lines of code sequentially. If it finds a fatal error it stops, and identifies the line number and the type of error. The line of code x=2 will produce the output x=2 when that line is executed. To suppress this output, place a semicolon at the end as x=2;.

Interpretive languages are slow, compared to compiled languages. For some applications, such as finite element analysis, MATLAB may not be suitable in its standard form. MATLAB can be compiled, but coding in C from scratch may be the best option if the computational intensity of the code is high.

Comments should be inserted to explain what is being entered. The symbol \% makes anything entered after that symbol on that line a comment, which is not executable. Anything on the line before \% is treated as code.

Defaults
\( i \) or \( j \rightarrow \sqrt{-1} \\
2 + 3i \rightarrow \text{complex number, but } 2 + 3i \text{ assumes } i \text{ is a variable, which must have been defined} \\
pi \rightarrow \pi = 3.1416 \text{ with format short, } \pi = 3.141592653589793 \text{ with format long} \\
pi = 3 \rightarrow \text{sets } \pi \text{ as a user defined variable} \\

\text{Arithmetic} \\
MATLAB follows the usual arithmetic protocols. \\
3 * 2 \rightarrow 6, \text{ multiplies the two numbers} \\
3 \wedge 2 \rightarrow 9, \text{ and the symbol } \wedge \text{ raises the first number by the power of the second number} \\
12 / 3 \rightarrow 4, \text{ divides the two numbers} \\
12 / 3 * 2 \rightarrow 8, \text{ it is bad practice to use this. Instead use parentheses as done below.} \\
(12 / 3) * 2 \rightarrow 8, \text{ parentheses always take precedence} \\
\text{real}(x) \rightarrow \text{real part of complex } x \\
\text{imag}(x) \rightarrow \text{imaginary part of complex } x \\
x = y \rightarrow \text{replaces what was in } x \text{ with what is in } y, \text{ regardless of what was orginally in } x \\
\text{sqrt}(x) \rightarrow \text{square root of } x \\

\text{Vectors and Matrices} \\
x = [1 4 3 2] \rightarrow \text{vector [1234], where one or more spaces between the numbers indicate the} \\
\text{next entry in the array.} \\
y = [2 3 1] \\
y(6) = 4 \rightarrow y = [2 3 1 0 0 4], \text{ these two lines of code illustrate dynamic indexing} \\
A = [1 2 2 3; 2 0 1 1; 2 1 4 1] \rightarrow \\
\begin{bmatrix}
1 & 2 & 2 & 3 \\
2 & 0 & 1 & 1 \\
2 & 1 & 4 & 1
\end{bmatrix} \\
[122320112141] \\
The semicolons indicate that what follows starts the next row. \\
A = [1 2 2 3 \\
2 0 1 1 \\
2 1 4 1] \rightarrow \text{this produces the same matrix as before} \\
(\text{Although the numbers of entries in the rows are not required to be the same, it is good practice} \\
\text{to make them the same}) \\
x = [1 4 3 2] \\
y = [2 3 1 4]
\[ A = \begin{bmatrix} 1 & 2 & 2 & 3 \\ 2 & 0 & 1 & 1 \\ 2 & 1 & 4 & 1 \end{bmatrix} \]

\[ x(2) \rightarrow 4, \text{ the second entry of vector } x \]
\[ A(3,2) \rightarrow 1, \text{ the entry in the third row and second column of } A \]
\[ A(2,:) \rightarrow [2 \\ 0 \\ 1 \\ 1], \text{ the second row of } A \]
\[ A(:,3) \rightarrow [2 \\ 1 \\ 4 \\ 1], \text{ the third column of } A \]
\[ A(:,2:3) \rightarrow [2 \\ 0 \\ 1 \\ 1; 2 \\ 1 \\ 4 \\ 1; 2 \\ 1 \\ 4 \\ 1], \text{ the second and third row of } A \]
\[ x+y \rightarrow [3 \\ 7 \\ 4 \\ 6], \]
\[ 2*x+y \rightarrow [4 \\ 11 \\ 7 \\ 8] \]
\[ x*y \rightarrow [2 \\ 12 \\ 3 \\ 8], \text{ multiplies the two arrays, element by element} \]
\[ x./y \rightarrow [0.5000 \\ 1.3333 \\ 3.0000 \\ 0.5000], \text{ divides the two arrays, element by element} \]
\[ x.^3 \rightarrow [1 \\ 64 \\ 27 \\ 8], \text{ raises each element of } x \text{ to third power} \]
\[ x+yj \rightarrow [1+2i \\ 4+3i \\ 3+i \\ 2+4i], \text{ adds real } x \text{ to imaginary } y \]
\[ x*y' \rightarrow [1432] [2314]=25 \quad x'\times y' \rightarrow [1432] [2314]=[[2314812416693124628] \]
\[ x*y \text{ and } x'\times y' \text{ both produce an error (in multiplying two vector/matrices, the number of rows in the first entity must match the number of columns in the second)} \]
\[ sort(x) \rightarrow [1 \\ 2 \\ 3 \\ 4] \rightarrow \text{ sorts the elements of } x \text{ in ascending order} \]
\[ sort(A)=[1011121212243] \]
\[ \rightarrow \text{ sorts columns of } A \text{ in ascending order} \]
\[ A' \rightarrow , [122201214311] \text{ the transpose of } A \]
\[ A(2,5)=6 \rightarrow \text{ changes matrix } A \text{ to } [122302011621410] \]
\[ x=linspace(0,2*pi,30) \rightarrow \text{ creates an array of 30 equally spaced points from 0 to } 2\pi \]

**Control Loops**

MATLAB is rich in control statements, such as \textit{for, if, else, while}, which start a number of lines of code that are always terminated with \textit{end}. These are best illustrated by example.

\[ \text{for } k=1:3 \]
\[ x(k)=2^k; \]
\[ \text{end} \]
x \rightarrow [2 \ 4 \ 8], this is an example of dynamic programming. If the line \(x(5)=3\) is inserted statement and \(x\) between the end then the final line would be \(x=[2 \ 4 \ 8 \ 0 \ 3]\).

\[
\text{for } j=-0.3:0.1:2 \rightarrow \text{begins for loop at -0.3, steps by 0.1 until 2 reached}
\]

\[
x=1;
\]

\[
\text{while } x<10 \ % \text{ continues to execute loop as long as } x < 10
\]

\[
x=2\times x;
\]

\[
\text{end}
\]

\[
x \rightarrow \text{returns } x=8
\]

\[
x=1;
\]

\[
\text{if } x<10 \ % \text{ continues to next line if } x<10
\]

\[
x=2\times x;
\]

\[
\text{end}
\]

\[
x \rightarrow \text{returns } x=8
\]

\[
\text{for } k=1:25
\]

\[
x=k^{1.2};
\]

\[
\text{if } x>2 \ break \ % \text{if } x \text{ is greater than 2, break out of the loop}
\]

\[
\text{end} \ % \text{this end is associated with break}
\]

\[
\text{end}
\]

\[
x \rightarrow \ 2.2974
\]

\[
sind(x) \rightarrow \text{sine of } x \text{ where } x \text{ in degrees: also cosd(x), tand(x)}
\]

\[
sin(x) \rightarrow \text{sine of } x \text{ where } x \text{ in radians: also cos(x), tan(x)}
\]

\[
asin(x) \rightarrow \text{inverse sine where result is in radians}
\]

\[
exp(x) \rightarrow , \text{where } e = \exp(1) = 2.718281828459046, \text{base of natural logarithms}
\]

\[
log(x) \rightarrow \text{natural logarithm (log to base } e)\]

\[
log10(x) \rightarrow \text{logarithm to base 10}
\]

\[
sinh(x) \rightarrow \text{hyperbolic sine}
\]

\[
max(x) \rightarrow \text{the maximum element in vector, also } min, mean, median
\]

\[
max(A) \rightarrow \text{the maximum value in each column of } A
\]

\[
abs(x) \rightarrow \text{absolute value of scalar } x
\]

\[
sort(a) \rightarrow \text{sorts vector in ascending order}
\]

\[
sort(A) \rightarrow \text{sorts each column of matrix } A \text{ in ascending order}
\]

\[
sign(x) \rightarrow \text{returns } -1\text{if } x < 0
\]

\[
0\text{if } x = 0
\]

\[
1\text{if } x > 0
\]

\[
round(x) \rightarrow \text{rounds scalar } x \text{ to nearest integer}
\]

\[
fix(x) \rightarrow \text{rounds to integer nearer zero}
\]
floor(x) → rounds to nearest integer towards -∞
ceil(x) → rounds to nearest integer towards ∞
sum(x) → sums elements of vector
sum(A) → sums columns of columns of A
size(A) → size of matrix A
length(x) → length of vector x
prod(x) → computes the product of all entries in
prod(A) → vector whose entries are the products of the entries in each column
std(x) → standard deviation of the n samples defined by variance σ² = \(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\)
eye(n) → identity matrix of order n
ones(n) → square matrix of order n with unity entries
zeros(n) → square matrix of order n with all entries zero
rand(n) → square matrix of order n whose entries are random numbers in the interval (0,1)

inv(A) → \(A^{-1}\) for \(A_{mn}x_n = y_n\)
A\(\backslash\)y → x (this is more efficient than inv(A)*y)
y/A → row vector x for xA = y
det(A) → \(\det(A)\), the determinant of A
rank(A) → the rank of A
eig(A) → eigenvalues of A

[N,E]=eig(A) → matrix of eigenvectors \(V_{nn}\) and diagonal matrix of eigenvalues \(E_{nn}\), where AV = VE

dot(a,b) → scalar dot product of vectors a and b whose order is the same
cross(a,b) → vector cross product of vectors a and b whose order is the same
length(a) → returns the order of vector a
conv(a,b) → convolves vectors \(a_n\) and \(b_m\) by treating them as polynomials and forming
resulting polynomial of order \(n + m - 1\)
diff(a) → difference vector \([a_2 - a_1 a_3 - a_2 ... a_n - a_{n-1}]\)
diff(a,2) → second difference for vector \(a_n\), resultant has order n-2
diff(A) → column differences for matrix A

Plotting Data

plot(y) → plots y against numbers 1 to i, with chords, where i is the number of entries in y
plot(x,y) → plots points y against x, with chords, where x(i) and y(i) stored
plot(x,y,’+’) → plots points with + at each point, where y(i) stored; also ’o’, ’*’, ’.’
loglog(x,y) → log-log plot of vectors and
semilogx(x,y) → log abscissa, linear ordinate
semilogy(x,y) → linear abscissa, log ordinate,
hold on → holds the current plot so that subsequent plots are superimposed on the original plot
axis([xmin xmax ymin ymax]) → sets limits on x-y plot
grid on → shows grid lines on a plot

Polynomials

\[ p = [1 \ 6 \ 11 \ 6] \rightarrow \text{polynomial } x^3 + 6x^2 + 11x + 6 \]
\[ \text{roots}(p) \rightarrow \text{roots of polynomial } p \text{ as } -1, \ -2, \ -3. \]
\[ \text{poly}(r) \rightarrow \text{forms polynomial } p \text{ from its roots} \]
\[ \text{polyval}(p,2) \rightarrow \text{value of polynomial } p \text{ at } 2 \text{ as } 60 \]
\[ \text{polyfit}(x,y,3) \rightarrow \text{polynomial of order } 3 \text{ for data } i = 1,2,...,n, \text{ where by least squares} \]
\[ \text{polyder}(r) \rightarrow \text{derivative of } r \]
\[ d = \text{conv}(b,c) \rightarrow \text{convolution of polynomials } b \text{ and } c \]
\[ [d,r] = \text{deconv}(b,c) \rightarrow \text{divides } b \text{ by } c \text{ to form } d \text{ with remainder } r \text{ (} r \text{ null if } \text{conv}(c,d)=b) \]
\[ [r,p,k] = \text{residue}(n,d) \rightarrow d(s) = \sum r_i s^i - p_i + k(s), \text{ where } k(s) \text{ is null if } O(n(s)) \leq O(d(s)), \text{ (if a multiple root occurs, then terms } p_i r_i s^i - p_i + r_i + 1(s-p_i)^2 + \ldots \text{ occur}) \]

Transforms

\[ \text{fft}(x) \rightarrow \text{Fourier transform of vector } x \]
\[ \text{fft}(A) \rightarrow \text{Fourier transform of each column of } A \]
\[ \text{ifft}(x) \rightarrow \text{inverse Fourier transform of vector} \]
\[ \text{fourier}(f) \rightarrow \text{forward Fourier transform of symbolic } f \]
\[ \text{invfourier}(f) \rightarrow \text{inverse Fourier transform of symbolic } f \]
\[ \text{laplace}(f) \rightarrow \text{forward Laplace transform of symbolic } f \]
\[ \text{invlaplace}(f) \rightarrow \text{inverse Laplace transform of symbolic } f \]

Logical Operators

if \( x = 3 \) \% the double equals sign means identically equal
\[ y = 6; \text{ end; } \rightarrow \text{replaces } y \text{ with number } 6 \text{ if } x \text{ equals } 3 \]
if \( x = 5 \)
\[ y = 6; \]
else
\[ y = 4; \]
\[ end \rightarrow \text{if } x \text{ is not 5, then set } y = 6, \text{ otherwise set } y = 4 \]
\[ \text{if } x = 2 \& y = 3 \]
\[ z = 4 \]
\[ end \rightarrow \text{if } x=2 \text{ and } y=3 \text{ then } z=4 \]
\[ \text{if } x = 2 \mid y = 3 \]
\[ z = 4 \]
\[ end \rightarrow \text{if } x=2 \text{ or } y=3 \text{ then } z=4 \]
\[ \text{if } x == 2 \]
\[ y = 3 \]
\[ \text{elseif } x == 3 \]
\[ y = 6 \]
\[ end \rightarrow \text{if } x=2 \text{ set } y=3, \text{ otherwise if } x=3 \text{ set } y=6 \]

**Symbolic Expressions and Functions**

A symbolic expression is entered by encasing it in quotation marks: \( \text{sym('}a+b*\text{t}^\text{2}') \) → stores \( a+b*t^2 \) as symbolic function with \( t \) as independant variable

\( \text{symvar('}a*\text{x}+1/(y+3)') \) → finds symbolic independant variable as \( x \)

\( \text{fmin('}g(\text{x}),3,11) \) → returns the minimum value of stored \( g(x) \) in the range \( 3 \leq x \leq 11 \)

\( \text{fplot('}\text{funct],[-2,5]) \) → plots stored \( \text{'funct} \) between limits -2 and 5: eg funct=\text{sin}.

\( f = '2/x^2' \) → \( f=2x^n \), which can be manipulated symbolically

\( \text{diff(sym(f))} \) → the derivative \( \text{- frac}2x^{n+1} \)

\( \text{int(sym(f))} \) → the integral \( \text{- frac}2x \)

\( \text{int(sym(f),0,3)} \) → the definite integral \( \int_0^3 2x^2 \text{dx}=\text{inf} \), where \( \text{inf} \) means \( \infty \)

\( \text{quad('}f',1,3) \) → evaluates definite integral of stored \( f \) between points 1 and 3

\( \text{int('}f') \) → indefinite integral of stored \( f \)

\( \text{diff(int(f))} \) → partial fraction expansion on stored rational polynomial \( f \)

\( g = 'a*\text{x}^2+b*\text{x}+c' \)

\( \text{diff(sym(g))} \) →, \( 2ax + bd \)ifferentiating wrt

\( \text{diff(sym(g),'}a') \) →, \( 2ax + bd \)ifferentiating wrt

\( d = '\text{sin(x)}' \)

\( \text{int(sym(d))} \) → the integral \( -\text{cosx} \)

\( \text{int(sym(d),0,pi)} \) → the definite integral \( \int_0^\pi \text{sinxdx}=2 \)

**Subroutines**
The user can store subroutines. This is particularly useful when it needs to be executed a number of times, since it avoids duplication of the same code. The subroutine starts with the statement \texttt{function out = funcname(x)} where \texttt{funcname} is chosen by the user, \texttt{x} is the control variable, and ends with the statement \texttt{out=yy} where the answer the user wants is in \texttt{yy}.

EXAMPLE: Newton’s method for finding the square root of number \texttt{x} is to use an initial guess \texttt{yy=x/2}, and then successive guesses \texttt{yy=(yy+x/yy)/2} until \texttt{yy/xx}. The m-file called \texttt{Newroot.m} to perform this is \texttt{function out = Newroot(x)}

\begin{verbatim}
y=x;
yy=x/2;
while abs(yy-y)>0.000001
  yy=(yy+x/yy)/2;
y=x/yy;
end
out=yy
\end{verbatim}

The statement \texttt{Newroot(2)} in another m-file will return 1.44213562. Notice that the left-hand side of the equation has the same variable that appears on the right-hand side in the statement \texttt{yy=(yy+x/yy)/2}; this is standard practice in MATLAB.

**Scope of MATLAB**

There is far more to MATLAB than described here. There are a number of books covering MATLAB from specific viewpoints. What is covered here will suffice, barely, for those who wish to follow the MATLAB files used in preparing this manuscript.
Chapter 1

Energy Sources, Energy Uses, and Global Warming

The energy available for use on Earth has two sources — solar and nuclear. The Sun’s energy is generated by nuclear fusion of hydrogen into helium during which some mass $M$ is lost to produce energy $E$, according to Albert Einstein’s famous formula $E = MC^2$ where $C$ is the speed of light in a vacuum. Thus, it could be argued that the only source of energy is nuclear. So how is the solar energy reaching Earth used, stored, or lost? That is the subject of this chapter.

![Figure 1.1: Worldwide Energy Use in Exajoules](image)

The SI (Système International) units to describe energy and power are the joule J and the watt W, where 1 joule equals 1 watt second, or $1 \text{ J} = 1 \text{ W} \cdot \text{s}$. These units are small. Residential electric power is measured in kilowatts (kW), and customers of the electric utility company are billed based on the number of kilowatt hours (kWh) they consume. To determine the total
amount of energy used worldwide, we need a much larger measure since numbers with many zeros at the end are difficult to comprehend. The standard is the exjoule (EJ), defined as $10^{18}$ J. World energy use in EJ is shown in 1.1. Approximate equivalents are:

$$1 \text{ EJ} = 2.77 \times 10^{11} \text{ kWh} = 7.6 \times 10^9 \text{ gallons of gasoline (US)} = 9.2 \times 10^{11} \text{ ft}^3 \text{ natural gas} = 3.4 \times 10^7 \text{ tonnes coal} = 0.948 \text{ quads},$$
where 1 quad = $10^{15}$ BTU (British Thermal Units)

Worldwide energy use is now increasing by about 71 EJ per decade. It is difficult to see this energy use decreasing since population growth and increased living standards require more energy. The challenge for the future is to replace energy generated from fossil fuels with renewable energy.

### 1.1 History of Fossil Fuels

Fossil fuels are hydrocarbons and were formed in the Carboniferous Period from about 360 to 286 million years ago, when much of the Earth was forested. Much of the land was swampy with ferns and a great variety of leafy plants as well as huge trees. When such plants die and fall into the swamp, the waterlogged material sinks to the bottom and over time gets covered in sediment, and peat is formed as a wet, spongy substance.

Peat remains a popular domestic fuel in Ireland. It is cut out of the ground in slabs and air dried before burning in fireplaces. It is a dirty fuel, but is free for the taking, so the price to be paid, apart from the cost to the environment, is the labor of harvesting and drying.

Over time peat becomes covered with more and more material, silt, sand, rock, etc., and the pressure on the peat converts it to lignite, a soft coal. Additional compression and time and the lignite becomes bituminous coal. Eventually, it becomes anthracite, a hard coal.

Some coal deposits of coal were formed in the late Cretaceous period of 65 million years ago, but these deposits are considerably smaller than those of the Carboniferous Period. Late forming coal is less compacted and contains more moisture; it is often called “brown coal.”

Fossil fuels remained largely untouched until the Industrial Revolution. Initially, water power provided the motive power for the textile mills in Northern England. The Massachusetts towns of Lowell and Lawrence copied the English model, and other New England towns followed, producing great wealth for the mill owners. But the amount of power to be extracted from rivers is limited, and we have harnessed most of the power from rivers.

The English moved from total reliance on water power to a predominance of coal power. The coal seams of England are, by chance, near the textile manufacturing centers, so soon coal became the fuel of choice. This was not the same in New England, and a variety of factors led
to the collapse of textile manufacturing in New England and its displacement to the South to be closer to cotton.

The thickness of an English coal seam can be as thin as a film, or as thick as 15 m. Seams as thick as 60 m can be found in India and France.

When coal was discovered around 1750, not much was made of it. It was difficult to burn in a fireplace made for wood. Transportation difficulties were dominant. Then blacksmiths found it a longer lasting fuel than wood with more energy per unit volume. Shipments from the principal coal mine near Wilkes Barre, Pennsylvania, was 365 tons in 1820, but increased to 40,000 tons in 1822, and 140,000 tons in 1833. A major industry was born, supported by an increasingly effective and fossil fueled railroad network. However, it was only until 1855 that the use of coal surpassed that of charcoal, made from wood.

![Figure 1.2: Growth in Coal Consumption](image)

The recent use of coal worldwide in short tons is shown in Figure 1.2. The United States Energy Information Administration is the source of this information, and it can be found at [www.indexmundi.com/energy.aspx?product=coal&graph=production](http://www.indexmundi.com/energy.aspx?product=coal&graph=production).

China is by far the world’s largest producer and consumer of coal, accounting for 46% of global coal production and 49% of global coal consumption — almost as much as the rest of the world combined... China has accounted for 69% of the 3.2 billion ton increase in global coal production over the last 10 years [www.eia.gov/todayinenergy/detail.cfm?id=16271. The coal consumption in billions of tons in China versus the rest of the world is shown in Figure 1.3. The source for this data is the U.S. Energy Information Administration, International Energy Statistics.

<table>
<thead>
<tr>
<th>Country</th>
<th>2011</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>3162</td>
<td>3200</td>
</tr>
<tr>
<td>U.S.</td>
<td>932</td>
<td>928</td>
</tr>
<tr>
<td>India</td>
<td>538</td>
<td>540</td>
</tr>
<tr>
<td>Australia</td>
<td>353</td>
<td>355</td>
</tr>
<tr>
<td>South Africa</td>
<td>255</td>
<td>257</td>
</tr>
<tr>
<td>Russia</td>
<td>248</td>
<td>250</td>
</tr>
<tr>
<td>Indonesia</td>
<td>173</td>
<td>175</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>105</td>
<td>107</td>
</tr>
<tr>
<td>Poland</td>
<td>77</td>
<td>78</td>
</tr>
<tr>
<td>Colombia</td>
<td>74</td>
<td>75</td>
</tr>
</tbody>
</table>

The earliest use of oil, as far as we know, was from oil seepages in the Middle East about 3000 BC; the oil was used for ship caulking and road construction, but apparently was not used as fuel. The Chinese discovered natural gas around 200 BC when they drilled a well for brine...
and hit gas instead. They then figured out that the gas could burn, so they used it to dry the brine into salt.

Similarly, in 1819, a well being drilled for brine in Kentucky struck black petroleum; without a use for this “black stuff,” the well was abandoned. In 1829 another well produced a massive flow of oil; the only use to be found was as a liniment. Finally, in 1859 a well in Pennsylvania was bored for the purpose of extracting oil.

The use of petroleum was virtually unknown in the nineteenth century. Before that the principal fuel, particularly in the United States, was wood. The population density was in the northeast, which was heavily forested. The climax trees were not only massive but were slow growth, high density, and high strength. Their timbers were perfect for ship building — the U.S.S. Constitution which played an important role in the Revolutionary War, was termed “Old Ironsides,” since enemy cannon balls bounced off her sides; she is the oldest warship still under commission in the United States.

### 1.2 Composition of Fossil Fuels

Whereas coal is formed from plants, petroleum is formed from marine life. In particular, the basis of petroleum is phytoplankton, algae, and bacteria. The petroleum, oil, or gas, found under “trap rock,” usually shale, was formed in the Cenozoic Period about 50 million years ago.

Coal is low in hydrogen ($\text{H}$) and high in carbon ($\text{C}$). \[ \text{C} + \text{O}_2 \rightarrow \text{heat} + \text{CO}_2. \] Coal combustion adds almost twice as much CO$_2$ to the atmosphere per unit of energy produced as does natural gas. Crude oil falls between these two [Energy Information Administration, “Emissions of Greenhouse Gases in the United States 1985-1990,” DOE/EIA-0573, Washington, DC, September 1993].

The percentage composition of coal is

<table>
<thead>
<tr>
<th>Element</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>carbon</td>
<td>75</td>
</tr>
<tr>
<td>ash</td>
<td>10</td>
</tr>
<tr>
<td>oxygen</td>
<td>8</td>
</tr>
<tr>
<td>hydrogen</td>
<td>5</td>
</tr>
<tr>
<td>nitrogen</td>
<td>1.5</td>
</tr>
<tr>
<td>sulphur</td>
<td>0.5</td>
</tr>
</tbody>
</table>

The percentage composition of crude oil, the liquid extracted from the ground, is

<table>
<thead>
<tr>
<th>Element</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>carbon</td>
<td>83–87</td>
</tr>
<tr>
<td>hydrogen</td>
<td>10–14</td>
</tr>
<tr>
<td>nitrogen</td>
<td>0.1–2</td>
</tr>
<tr>
<td>oxygen</td>
<td>0.1–1.5</td>
</tr>
</tbody>
</table>
Natural gas is mostly methane ($CH_4$). A typical analysis of natural gas, by mole, is

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methane</td>
<td>95.2%</td>
</tr>
<tr>
<td>Ethane</td>
<td>2.5%</td>
</tr>
<tr>
<td>Propane</td>
<td>0.2%</td>
</tr>
<tr>
<td>Butanes/pentanes/hexanes</td>
<td>0.09%</td>
</tr>
<tr>
<td>Nitrogen/carbon dioxide</td>
<td>2%</td>
</tr>
</tbody>
</table>

Estimated coal reserves in million short tons per state in 2009 was

<table>
<thead>
<tr>
<th>State</th>
<th>Reserve (million short tons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Montana</td>
<td>74,770</td>
</tr>
<tr>
<td>Wyoming</td>
<td>38,743</td>
</tr>
<tr>
<td>Illinois</td>
<td>37,913</td>
</tr>
<tr>
<td>West Virginia</td>
<td>17,390</td>
</tr>
<tr>
<td>Kentucky</td>
<td>14,480</td>
</tr>
<tr>
<td>U.S. TOTAL</td>
<td>260,553</td>
</tr>
</tbody>
</table>

Figure 1.4: From Energy Source to Energy Use

1.3 Fossil Fuels, Uses and Reserves

The source of energy and its uses in today’s world is shown in Figure 1.4. The largest energy source is petroleum. Petroleum imports from 1950 to 2010 increased from almost zero in 1950.
to about 50% by 2005 [U.S. Energy Information Administration, Monthly Energy Review, Table 3.1, April 2013]. It is now declining as shown in Figure 1.5 [C.E. Behrens & C. Glover, “U.S. Energy: Overview and Key Statistics,” Congressional research Service, April 2012].

![Figure 1.5: Petroleum Imports to the United States.](image)

Per capita energy use in North America, at the equivalent of approximately 91 MWh annually, is higher than in any other part of the world. Europe is next at about 42 MWh. Asia is at 10 MWh, and South America at 12 MWh. The worldwide average is about 19 MWh annually [timeforchange.org].

![Figure 1.6: Worldwide Energy, Produced and Projected](image)
The energy in quads produced in recent years, and a projection out to 2035, is shown in Figure 1.6. Petroleum, crude oil, is the dominant source. About 50% of petroleum is used as gasoline or diesel oil, about 27% for heating oils, both domestic (diesel oil) and heavy bunker oils. Of the remaining 23%, 11% is used to generate electricity, 6% goes to industrial lubricants, and the remaining 6% is converted into plastics, synthetic rubbers and fibers, fertilizers, paints, detergents, and other products.

![Figure 1.7: Fuel Source for Electrical Generation](image)

Estimated technically recoverable natural gas resources in United States is 2,587 trillion ft$^3$ [Energy Information Administration - Annual Energy Outlook 2010], most of it onshore. This is far higher than the estimate of natural gas resources [National Petroleum Council, “Facing the Hard Truths About Energy,” 2007] of 1,451 trillion ft$^3$. However, a report by Jad Mouawad in the New York Times on June 18, 2009 showed that the estimates of natural gas in shale rocks was 35% higher than previously thought and stood at 2,074 trillion ft$^3$.

We badly underestimated the recoverable reserves of natural gas. In fact, estimates of reserves of natural resources are almost always far too low. Until recently, estimates of recoverable oil were way off. True, most of the “easy to get out” oil has been fully exploited, but new technologies are finding ways to get at shale oil deep beneath the Earth’s surface, and off-shore drilling platforms are becoming more adventurous and so, more dangerous: case in point, Deep Water Horizon, discussed briefly in Section IV.


“Crude oil – proved reserves is the stock of proved reserves of crude oil, in barrels (bbl). Proved reserves are those quantities of petroleum which, by analysis of geological and
engineering data, can be estimated with a high degree of confidence to be commercially recoverable from a given date forward, from known reservoirs and under current economic conditions [The World Factbook, www.cia.gov/library/publications/the-world-factbook/rankorder/2244rank, 2014].”

The countries with the highest proven reserves, in millions barrels of oil, together with an estimate of these reserves [op cit] are given in Table 1.1.

Table 1.1: Oil Reserves

<table>
<thead>
<tr>
<th>country</th>
<th>reserves</th>
</tr>
</thead>
<tbody>
<tr>
<td>Venezuela</td>
<td>297,700</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>268,400</td>
</tr>
<tr>
<td>Canada</td>
<td>173,200</td>
</tr>
<tr>
<td>Iran</td>
<td>157,300</td>
</tr>
<tr>
<td>Iraq</td>
<td>140,300</td>
</tr>
<tr>
<td>Kuwait</td>
<td>104,000</td>
</tr>
<tr>
<td>United Arab Emirates</td>
<td>97,800</td>
</tr>
<tr>
<td>Russia</td>
<td>80,000</td>
</tr>
<tr>
<td>Libya</td>
<td>89,470</td>
</tr>
<tr>
<td>Nigeria</td>
<td>37,140</td>
</tr>
<tr>
<td>United States</td>
<td>30,530</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>30,000</td>
</tr>
</tbody>
</table>

The annual natural gas production in millions of m³ for the top producing countries in the world [http://jthomasranken.com/is-natural-gas-really-the-next-big-thing] are given in Table 1.2.

The uncertainty in determining natural gas reserves is a consequence of the late realization of natural gas as an energy source, particularly in electrical power generation. As can be seen in Figure 1.7, the importance of natural gas as the source of electrical energy only became apparent around 1990. The units are in billions of watt hours. A game changing technology has vastly increased the amount of natural gas that is being produced in the United States — hydrofracking.

Oil and gas carrying shale exists over large tracts in the United States. The greatest deposit of oil shale in the world is in the Green River Formation spanning parts of Colorado, Utah, and Wyoming as shown by the shaded areas in Figure 1.8. It is estimated to contain between 1.2 to 1.8 trillion barrels. Most of this formation over 65,000 km² is under federal land. This location is arid, a major problem since shale oil extraction and processing requires several gallons of water for every gallon of oil produced [Bartis, J.T., LaTourrette, T., Dixon, L., Peterson, D.J., and Cecchini, G., “Oil Shale Development in the United States; Prospects and Policy Issues,” Rand Corporation, MG-414-NETL, 2005].
Initial investigations for development of the Green River shale oil began in 1967, but it required the OPEC oil embargo of 1973–74 to provide real impetus. Several parcels on public land were put up for bid under the Federal Prototype Oil Shale Leasing Program. Traditional deep shaft mining techniques were employed without much success. The most productive was the Unocal plant constructed in 1980 that produced 4.4 million barrels. It produced about 5,900 barrels of oil a day, but was not financially successful, so it was closed in 1991.
Recent studies now project much larger estimates for the oil reserves in the Green River formations. “A stretch of largely vacant federal lands in Utah, Wyoming, and Colorado may hold more recoverable oil than all the rest of the world put together. That is what Anu Mittal, Director of Natural Resources for the General Accounting Office, informed the House Science Subcommittee on Energy and the Environment in her written testimony on May 10, 2012. The Green River Formation — an assemblage of over 1,000 feet of sedimentary rocks that lie beneath parts of Colorado, Utah, and Wyoming — contains the world’s largest deposits of oil shale. USGS [U.S. Geological Survey] estimates that the Green River Formation contains about 3 trillion barrels of oil, and about half of this may be recoverable, depending on available technology and economic conditions [B. Walker, “The Green River Formation: World’s Largest Oil Shale Deposits,” thenewamerican.com].”

“People have squeezed oil from Western shale since the early 20th century, but it has yet to flow in commercial quantities despite untold millions spent on research and failed operations. Now, Enefit and another Utah operator called Red Leaf Resources, using a much different process, seek to bring industrial-scale development to a petroleum resource believed to dwarf Saudi Arabia’s reserves . . . . The 60-foot seam Enefit is targeting will yield 27 gallons of oil per ton of ore, according to CEO Rikki Hrenko. In other words, the company is hoping to extract two-thirds of a barrel from every cubic yard of ore it plans to strip mine on the private portion of its holdings. By the 2020s, Enefit hopes to be shipping 50,000 barrels of sweet crude to Salt Lake City-area refineries each day [B. Maffly, “Green River Formation may dwarf Saudi reserves, but environmentalists are worried, Salt Lake Tribune, Aug 10, 2013].”

A secondary area of shale oil is the Devonian deposit of about 725,000 km² over parts of Tennessee, Kentucky, Indiana, and Ohio. The first commercial ventures tapped the shale for natural gas. Now efforts are underway to develop economical extraction of Devonian oil.

Another large deposit is the Marcellus and adjoining Devonian Shales in the northeast; the location of these two deposits is shown in Figure 1.9. “Named for the exposed outcrop in Marcellus, NY, Marcellus Shale is a large deposit of black shale and is characterized by being very rich in unoxidized carbon. Marcellus Shale covers most of New York State and ranges in depths down to 7,000 feet below the surface and is included in most of the Appalachia. The larger Marcellus shale formation becomes thinner as it moves from east to west and is named for an exposed outcrop in Marcellus, NY. Utica Shale is deposited broadly across the Appalachian Basin and into Ontario, Canada. In New York there is an outcrop along the west and southeast sides of the Adirondack Mountains, and is also exposed along the northern Alleghany Plateau. It ranges in depths to over 9,000 feet in the southern portion of the state. One of the characteristics that is common with black shale is that it contains trace levels of uranium. The concentration of this uranium at the surface, on drilling equipment, and in combination with drilling muds, fracking fluid, and other elements exposed in the process of drilling is the primary cause for concern with exposing this shale. The proper disposal of these cuttings, worker exposure, and the potential contamination from open on-site storage must be adequately
Oil shale is the second largest energy source behind coal. It was formed 40 to 50 million years ago in the same way as peat, in swamps and lake beds, from plant and animal matter. The United States oil shale resource is estimated at more than 2 trillion barrels. The problem is how to economically extract it and how to do this while limiting damage to the environment. Oil shale is a viscous form of petroleum called bitumen that may not flow at the ambient temperature.

It is more difficult and expensive to develop the shale oil resource than conventional oil resources. Shale oil can be mined like coal. This means underground mining or surface (strip) mining. The problem is, unlike coal, which typically is an almost totally combustible
substance, mined shale oil is mostly rock. The result is that large amounts of waste material must be disposed of. The economic cost of this disposal could be largely mitigated if a commercial use for it could be found.

A more promising technology is to drill a well and inject heat to make the oil in the shale flow, a process called retorting; this process does not result in the material waste problem of mining, but it requires large amounts of energy, reducing the net efficiency of the overall process. Another technique is to inject the well with lighter forms of petroleum to effectively dissolve the oil so it can be pumped out. Shell Oil uses electric heaters in the shale, heating the shale to 650 – 700° F over two or three years until the oil is released. The oil then flows into collection wells and is pumped to the surface.

In order to prevent groundwater from flowing into the heated zone, a freeze wall perimeter is erected by drilling wells and injecting cold liquids. This wall also prevents shale oil from leaking from the site.

Historically, the United States is late to the table for exploiting shale oil. About 1839, oil shale was extracted at Autin, France. Scotland started an oil shale industry in 1959 and worked about 20 different beds; by 1881 production reached a million metric tons. Mining ceased in 1962. Canada has developed some of its extensive oil shale reserves starting in the 1980s. Other countries with oil shale deposits include Venezuela, China, and Australia.

### 1.4 Environmental Costs of Fossil Fuels

What is the real cost of petroleum products? Not the cost of gasoline at the pump, but the long term costs, the subsidies, the mitigation costs, in sum the complete cost for the product. It has been estimated that nearly half of our defense expenditures of $700 billion is for ensuring that the 730 million barrels of crude keep flowing. Just this alone raises the cost of a gallon of gas from $2.70/gallon to $6.50/gallon [Steve Christ, “Oil Price Fantasy: The True Cost of Crude,” Wealth Daily, September 21, 2010].

The costs of fossil fuels start with getting it out of the ground, then transporting it to the refining facility, storing it and transporting it to the end user who then burns it. We investigate some of these costs in this section.

The Keystone XL pipeline will be the safest pipeline ever built, according to its owner TransCanada Corporation. And since it is so safe, they say it does not need infrared sensors or fiber-optic cables to detect spills. “There are lots of things engineering-wise that are possible, that the industry doesn’t do,” said Carl Weimer, executive director of Pipeline Safety Trust, a fuel-transportation safety advocacy group in Bellingham, Washington. As pipeline executives say they’re changing their industry’s culture to tolerate zero incidents, companies aren’t spending on technology to catch even pinhole-sized leaks that can turn into bigger problems,
Weimer said . . . Keystone XL is part of an additional 4.7 million barrels a day of new U.S. oil pipeline capacity expected to be built during the next two years, according to the Association of Oil Pipe Lines, a Washington-based industry group. About 19.2 million barrels of crude are transported each day in the U.S. Pipelines spilled an average of 112,569 barrels a year in the U.S. from 2007 to 2012, a 3.5 percent increase from the previous five-year period, according to U.S. Transportation Department figures” [R. Penty and M. Lee, “Keystone XL Pipe Shuns Infrared Sensors to Detect Leaks,” June 2013, http://www.bloomberg.com].

On November 6, 2015, Secretary of State John Kerry announced that construction of the Keystone XL pipeline would not serve the national interest. President Obama agreed with Kerry and announced his opposition to this project.

It is amazing how glib the executives that operate oil tankers, pipelines, and railroads carrying bulk oil or gas can be when minimizing the risks that their enterprises carry. In the years 2000 to 2010, over 130 pipeline incidents occurred in the United States alone. In July 2010 a 40’ section of the Enbridge Energy pipeline in Michigan ruptured and 877,000 U.S. gallons of heavy crude oil escaped into Talmadge Creek. The cost of ‘cleanup’ as of the summer of 2012 has been $765 million, and rising.

“Bad news for Enbridge, TransCanada and all the other companies working to build controversial pipelines across North America: Pipelines spill three times as much oil over comparative distances as rail, the International Energy Agency (IEA) says. The IEA found the risk of a rail spill is six times as high as the risk of a pipeline spill, but pipelines simply spill more when they rupture” [D. Tencer, “Pipeline Spills Release Three Times As Much Oil As Rail Spills,” Huffington Post Canada, May 14, 2013, http://www.huffingtonpost.ca]. The number of incidents reported, the property damage in millions of dollars, the oil spilled in thousands of barrels, and the number of thousand barrels lost that resulted from pipeline spills are given in Table 1.3.

“The number of spills and other accidents from railroad cars carrying crude oil has skyrocketed in recent years, up from one or two a year early in the previous decade to 88 last year. Only four of those were classified as serious by the Pipeline and Hazardous Materials Safety Administration (PHMSA), and none involved injuries. So they didn’t even approach the human tragedy caused by a runaway oil train in Quebec earlier this month . . . . The number of oil-filled tank cars moved by rail jumped from about 10,000 in 2009 to more than 230,000 last year, according to the Association of American Railroads” [M. Soraghan, “Crude mishaps on trains spike as rail carries more oil,” EnergyWire, July 17, 2013, http://www.eenews.net/stories].

On July 6, 2013, a train with 72 tankers, with each tanker containing 30,000 gallons of crude oil, was parked for the night. According to the train engineer, brakes were set on the carriages, but either this was not the case or someone maliciously released these brakes. The train was on a section of track that was relatively steep, so it became an out-of-control runaway. At high
speed the train ran into the small community Lac-Megantic, Quebec, where it crashed. The crude oil exploded, destroying 40 buildings, killing 13 and leaving 37 missing.

Table 1.3: Pipeline or Rail

<table>
<thead>
<tr>
<th>year</th>
<th>number</th>
<th>cost</th>
<th>spilled</th>
<th>lost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>229</td>
<td>$29</td>
<td>117</td>
<td>58</td>
</tr>
<tr>
<td>1994</td>
<td>245</td>
<td>$62</td>
<td>164</td>
<td>114</td>
</tr>
<tr>
<td>1995</td>
<td>188</td>
<td>$33</td>
<td>110</td>
<td>53</td>
</tr>
<tr>
<td>1996</td>
<td>194</td>
<td>$85</td>
<td>160</td>
<td>101</td>
</tr>
<tr>
<td>1997</td>
<td>171</td>
<td>$55</td>
<td>196</td>
<td>103</td>
</tr>
<tr>
<td>1998</td>
<td>153</td>
<td>$63</td>
<td>150</td>
<td>61</td>
</tr>
<tr>
<td>1999</td>
<td>167</td>
<td>$86</td>
<td>167</td>
<td>105</td>
</tr>
<tr>
<td>2000</td>
<td>146</td>
<td>$151</td>
<td>109</td>
<td>57</td>
</tr>
<tr>
<td>2001</td>
<td>130</td>
<td>$25</td>
<td>98</td>
<td>77</td>
</tr>
<tr>
<td>2002</td>
<td>459</td>
<td>$52</td>
<td>97</td>
<td>78</td>
</tr>
<tr>
<td>2003</td>
<td>434</td>
<td>$67</td>
<td>81</td>
<td>51</td>
</tr>
<tr>
<td>2004</td>
<td>377</td>
<td>$166</td>
<td>89</td>
<td>69</td>
</tr>
<tr>
<td>2005</td>
<td>369</td>
<td>$306</td>
<td>138</td>
<td>46</td>
</tr>
<tr>
<td>2006</td>
<td>354</td>
<td>$75</td>
<td>138</td>
<td>54</td>
</tr>
<tr>
<td>2007</td>
<td>332</td>
<td>$60</td>
<td>96</td>
<td>69</td>
</tr>
<tr>
<td>2008</td>
<td>375</td>
<td>$148</td>
<td>102</td>
<td>70</td>
</tr>
<tr>
<td>2009</td>
<td>342</td>
<td>$74</td>
<td>55</td>
<td>32</td>
</tr>
<tr>
<td>2010</td>
<td>350</td>
<td>$1,176</td>
<td>174</td>
<td>123</td>
</tr>
<tr>
<td>2011</td>
<td>346</td>
<td>$252</td>
<td>136</td>
<td>105</td>
</tr>
<tr>
<td>2012</td>
<td>364</td>
<td>$122</td>
<td>53</td>
<td>31</td>
</tr>
</tbody>
</table>

The number of U.S. railroad accidents involving hazardous materials has not increased in recent years, despite the massive increase in crude oil shipments. In 2012 there were 88 railroad spills, of which 23 resulted in a “visible sheen,” the federal standard for a significant oil spill from a train; by contrast, Federal law requires pipeline companies to report spills of 5 gallons or more, a more lax standard. The total volume of railroad spills in 2011 and 2012 was less than 4,000 gallons. North Dakota led the nation with 26% of the total spills — the increase in crude production from the Bakken Shale formation in this state could not be handled by the existing pipelines, so it needed to be shipped by rail.

Unlike oil spills on land that can be contained with earthen dams, oil spills at sea spread, sometimes for hundreds of miles of coastline. Land animals can avoid oil, but sea birds and marine life have no choice. Oil on feathers or fur reduces their insulating properties, and these animals often die of hypothermia. Catastrophic oil spills at sea receive considerable publicity, but illegal cleaning of oil tanks and ejection at sea of oily wastes contribute significantly to environmental damage.
On March 23, 1989, the tanker Exxon Valdez ran aground on Bligh Reef, Prince William Sound, Alaska, spilling 11 million gallons of crude oil. 1,300 miles of coastline were impacted, with 200 miles heavily oiled. It has been estimated that 250,000 birds, 2,800 sea otters, 300 harbor seals, 250 bald eagles, and 22 killer whales perished. The National Transportation Safety Board blamed the master of the vessel, who by some accounts was drunk, some members of the crew, the lack of an effective pilot, and the failure of the Exxon Shipping Company to supervise the master and crew. The vessel was single hulled — the company argued that a double hulled vessel would be an unnecessary expense since there was no chance of the single hull being penetrated.

The BP (British Petroleum) off-shore oil rig Deepwater Horizon blew up on April 20, 2010, causing the death of 11 men and the worst accidental oil spill in history. The investigation that followed led to BP pleading guilty to criminal charges of violations of the Clean Water Act in November, 2012 and paying a fine of $4.5 billion: this on top of payments to injured parties. As reported by ABC News on April 22, 2015, a federal court is deciding the size of the fine for violating the Clean Water Act.

Hydrofracturing, a terms usually shortened to hydrofracking or simply fracking, is a safe and effective method for increasing the yield from low flow water wells. The depths involved are typically 200 to 500’. After the well is dug an inflatable sleeve called a packer is inserted in the well below the casing and in the bedrock. The packer is hydraulically inflated and fresh water under high pressure is pumped through a pipe in the center of the sleeve. Several thousand gallons of water are pumped into the well. Up to 3000 lb/ft\(^2\) of pressure is possible, but a well that will maintain such pressure is unlikely to benefit from fracking. What is desirable is a pressure around 500 lb/ft\(^2\). This pressure in the seams in the bedrock is sufficient to raise the strata above. The movement is imperceptibly at the surface. After the full load of water is pumped into the well, the pressure is released and the water is allowed to jet out like a fire hose. The strata moved apart never regain their old seating since pebbles and gravel act as mini pillars.

Hydrofracking is now being employed to aid in the extraction of natural gas and oil. It is mostly water that is injected in deep wells. Some chemicals, between 0.5 and 2%, are added to the water, and it is claimed that these chemicals are the source of contamination. Water contamination could be caused by these chemicals, or could be caused by the well itself linking ultra-deep strata containing toxic substances with the shallow water-bearing strata.

After hydrofracking, the Marcellus shale field in the eastern United States has been found to be one of the biggest producers of natural gas in the world. The result is that it is estimated that the U.S. production of natural gas will exceed domestic demand by 2019, and then natural gas will be exported [DOE/EIA Annual Energy Outlook 2013: With Projections to 2040, July 25, 2013].
1.5 Energy Usage and Income

The rapid rise in energy use began with the industrial revolution in northern England, then rapidly spread to Europe and the New World. Water power and wood power changed to coal power, then crude oil became a factor, and more recently natural gas. The historic usage of these fuels in EJ (exajoules) is shown in Figure 1.10 [http://ourworldindata.org/data/resources-energy/energy-production-and-changing-energy-sources]. Wood was the dominant fuel at the beginning of the nineteenth century. By the beginning of the twentieth century, coal became dominant. In the middle of the twentieth century, crude oil had caught up with coal, and natural gas started to become a factor.

Shown in Table 1.4 is the relationship between national average energy use per capita (E) in kg of oil and per capita income (I) in 2011 U.S. dollars for some of the nations in the world, including the most populous [http://ourworldindata.org/data/resources-energy/energy-production-and-changing-energy-sources].

There is a direct relationship between energy usage and affluence. A subsistence farmer and the urban poor, particularly in third world countries, consume little energy. The poor farmer in India is likely to use dried dung for his primary fuel. The European and American farmer has powerful tractors running on diesel oil, lives in a house that is heated in winter with natural gas, and thinks nothing of running into town in an SUV to get a six pack of beer.
This data is presented in Figure 1.11. The cluster of six in the bottom left corner, within the green ellipse, represent the six nations at the bottom of 1.4. The red trend line is the function

\[ I = 8.83E - 157.8 \] (1)

Excluded from Figure 1.11 is the datum point for Iceland, It is an outlier in a number of ways. First, its energy use per capita of 17964 kg is over double that of the nation with the second highest energy consumption. Next, very little of Iceland’s energy use is with fossil fuels. Most of its energy comes from geothermal and hydroelectric; 71% of electrical power generation is produced from hydro, the remainder from geo. 90% of households are heated with geothermal energy. The breakdown of where geothermal is utilized is given in 1.5. Finally, all oil is imported, and 95% of this oil is devoted to fishing and transportation.

To determine how good the trend line shown in Figure 1.11 is in linking income to energy we invoke the principle of least squares approximation. Suppose \( f(x) \) is to be approximated by \( g(x) \) and the fit tested at \( n \) points \( x_i \). With a least squares approximation one minimizes

\[ \varepsilon = \sum_{k=1}^{n} (f(x_i) - g(x_i))^2 \] (1.2)
The goodness of fit $G$ determines how close $g(x)$ approximates $f(x)$. The sample average is

$$f^\bar=n\sum_{k=1}^{n}f(x_i) (1.3)$$

and a zero order approximation will be $D=\sum_{k=1}^{n}(f(x_i)-f^\bar)^2$, so the goodness of fit $G$ is given by

$$G=D-\varepsilon D=\sum_{k=1}^{n}(f(x_i)-f^\bar)^2-\varepsilon \sum_{k=1}^{n}(f(x_i)-f^\bar)^2 (1.4)$$
a number between 1, a perfect fit, 0, a useless fit since it is no better than the zero order approximation, and negative numbers indicating that the approximation is worse than useless.

For the specific data in this study, \( f^2=26,213, D=1.0267 \times 10^{10}, \) and \( G = 0.704, \) a fairly good fit, indicating that there is definitely a correlation between income and energy use.

Some interesting facts on energy use in the United States are:
* Energy released by burning a wood match, 1 BTU
* Total energy used annually in United States, about 100 quads
* U.S. energy use from fossil fuels, 86%
* U.S. population relative to world, 5%
* U.S. contribution to global warming, 20%

[“Fast Facts on Energy Use,” Energy Star, as accessed on 7/15/12]

### 1.6 Planet Earth

Life as we know it in the solar system is limited to planet Earth. We are at the optimum distance from the Sun, our own star. The closest planet to the Sun is Mercury, followed by Venus, then Earth and Mars. The solar constants for the five planets closest to the Sun are given in the column labelled SC in Table 1.6. Solar constant is energy striking a surface normal to the incident solar irradiance at that planet’s distance from the Sun [http://rredc.nrel.gov/solar/standards/am0/newam0.html].

The mean distances, the closest distances, and the furthest distances in millions of kilometers are in columns 3 through 5 in Table 1.6. The last column is the product of the mean distance squared and the solar constant, without the multiplying factor of \( 10^{12} \): each number is close to \( 30 \times 10^{12} \) and the units are \( \text{km}^2 \cdot \text{W/m}^2 \).

Given its proximity to the Sun, and so the highest of all solar constants, one would expect Mercury to have the highest average surface temperature. However, that honor goes to Venus, which has a dense atmosphere of carbon dioxide (think global warming) and sulfur dioxide. Mercury spins very slowly about its axis and has no atmosphere, so the side facing the Sun can reach \( 465^\circ \text{C} \), but the dark side drops to \( -184^\circ \text{C} \) [www.universetoday.com/35664/temperature-of-the-planets].
The average temperature on Mars is about $-55^\circ\text{C}$, although it can reach up to $20^\circ\text{C}$ at the equator. Due to its thin atmosphere it cannot sufficiently trap solar energy, and it is much colder than Earth.

The next planet out from the Sun is Jupiter. Unlike the rocky Earth, Jupiter is gaseous. At the top of its clouds the temperature is $-145^\circ\text{C}$, and this temperature increases towards the center of the planet. At the center the temperature is about $35,700^\circ\text{C}$, hotter than the surface of the Sun. Somewhere between the top of the atmosphere and center of the planet is a comfort zone [op cit] which, however, is unlikely to support life.

### 1.7 Direct Solar Energy

Direct solar energy is sunlight. It is the energy from the Sun that reaches planet Earth. It is attenuated as it passes through Earth’s atmosphere before reaching the Earth’s surface. It is the source of all life on Earth. We will consider this energy in detail in Chapters 5 through 8.

### 1.8 Indirect Solar Energy

Indirect solar energy may be “ready to use” almost right away, like wind energy. It may take longer, over several years, for forestry products. When winds blow, waves are created, so waves are a double indirect form of solar energy.

Water power has three principal forms. Hydroelectric is the most developed, and results from rainfall, which is itself the consequence of evaporation due to the Sun. Wind causes ocean waves, which contain enormous amounts of energy, and there are many schemes for capturing that energy but very little has been captured to date.

Tidal power is not the result of solar energy; it is caused by gravitational pulls of the Moon and to a lesser extent the Sun.
1.8.1 Winds of the World

The distribution of the world’s winds is a strong function of latitude. In particular, winds are strong between latitudes 30 and 60, both north and south, resulting in what are called the prevailing westerlies, as shown in Figure 1.12.

![Prevailing Winds of the World](image)

Figure 1.12: Prevailing Winds of the World

The Sun is strongest in the tropics, from the Tropic of Cancer at latitude 23.45° north to the tropic of Capricorn at 23.45° south. Close to the equator, between latitudes ±5° the heated air rises; as far as a navigator can tell there is little or no wind, and these latitudes are known as the doldrums. The heated air moves towards the poles, cooling as it does, and between latitudes 30 and 35°, both north and south, the air falls to Earth and has the same effect as in the doldrums — little wind to propel a sailing ship. Sailing masters called latitudes 30 to 35° the horse latitudes, supposedly (but dubiously) because they would throw their horses and cattle overboard to save on provisions when becalmed for a long period.

The horse latitudes are noted for high barometric pressures with little precipitation: the world’s major deserts, the Sahara and Great Australian, occur in the horse latitudes.

The southern westerlies are somewhat different than the northern westerlies since there is little land below latitude 40 south to attenuate the wind. Thus, latitudes 40 to 49 south are known as the roaring forties. Trade ships in the age of sail welcomed the roaring forties since it speeded their passage and so increased their profits. In some cases the profits from one voyage would pay for the construction cost of the ship itself.

The furious fifties and the screaming sixties have conditions similar to the roaring forties, but there is little financial incentive to ply these waters.
1.8.2 Wave Power

The following statements are taken from “Wave Energy Potential on the U.S. Outer Continental Shelf,” Minerals Management Service, U.S. Department of the Interior, May, 2006: “Incoming solar radiation levels that are on the order of 100 W/m² are transferred into waves with power levels that can exceed 1,000 kW/m of wave crest length. The transfer of solar energy to waves is greatest in areas with the strongest wind currents (primarily between 30° and 60° latitude), near the equator with persistent trade winds, and in high altitudes because of polar storms.

“Waves are also efficient transporters of solar energy. Storm winds generally create irregular and complex waves. In deep water, after the storm winds die down, the storm waves can travel thousands of kilometers in the form of regular smooth waves, or swells, that retain much of the energy of the original storm waves. The energy in swells or waves dissipates after it reaches waters that are less than ~200 m deep. At 20-m water depths, the wave’s energy typically drops to about one-third of the level it had in deep water.

“The total annual average wave energy off the U.S. coastlines (including Alaska and Hawaii), has been estimated [Bedard, R., et al., “Final Summary Report, Project Definition Study, Offshore Wave Power Feasibility Demonstration Project,” EPRI Global WP 009 – US Rev 1, Jan. 14, 2005] at 2,100 TWh. This estimate was made at a specified water depth of 60 m (irrespective of the distance from the shore at which that depth occurs) in order to allow comparisons of wave energies between coastal areas and to eliminate the possible, but unpredictable loss of energy of the wave through its interactions with the sea bottom (scouring) at shallower depths. Typical wave energy in U.S. offshore regions ranges from 2 to 6 Kw/m in the mid-Atlantic, 12 to 22 kW/m in regions such as Hawaii with trade winds, and 36 to 72 kW/m in northwestern U.S. coastal areas near Washington and Oregon.

“Estimates of the worldwide economically recoverable wave energy resource are in the range of 140 to 750 TWh/yr for existing wave-capturing technologies that have become fully mature [European Thematic Network on Wave Energy (ETNWE), “Results from the Work of the European Thematic Network on Wave Energy,” ERK5-CT-1999-20001, 2000–2003, 2003 — available at www.wave-energy.net]. With projected long-term technical improvements, this could be increased by a factor of 2 to 3 [Thorpe, T.W., ‘A Brief Review of Wave Energy,’ ETSU Report R-122, prepared for the United Kingdom Department of Trade and Industry, 1999]. The fraction of the total wave power that is economically recoverable in U.S. offshore regions has not been estimated, but is significant even if only a small fraction of the 2,100 Twh/yr available is captured. (Currently, approximately 11,200 TWh/yr of primary energy is required to meet total U.S. electrical demand.) WEC devices have the greatest potential for applications at islands such as Hawaii because of the combination of the relatively high ratio of available shoreline per unit energy requirement, availability of greater unit wave energies due to trade winds, and the relatively high costs of other local energy sources.”
The wave energy as measured in kW/m is shown in Figure 1.13. This information was extracted from the document, “Options for the Development of Wave Energy in Ireland: A Public Consultation Document, by the Marine Institute and Sustainable Energy Ireland,” November 2002, available online. In turn, the cited document relied on report ETSU-R120 [T.W. Thorpe, “A Brief Review of Wave Energy,” UK department of Trade and Energy, November 1999]. Notice the wave dominance of the western shores. When wave energy is converted to electrical energy, a near shore location is needed, and the map shows the most suitable area of the globe is northern Europe, particularly the west coasts of Ireland and the U.K.

Looking at Europe in greater detail in Figure 1.14, we see the west coast of Ireland, Scotland and Iceland have the best potential for producing electricity from waves [“Wave Energy Utilization in Europe: Current Status and Perspectives,” by the European Thematic Network on Wave Energy].

The book by David Ross, Power from the Waves, Oxford University Press, 1995, provides a good initiation into the study of wave energy.
1.8.3 Hydroelectric Power

The first hydroelectric plant in the world harnessed a small amount of the energy at Niagara Falls in 1879. In 1882 a hydroelectric facility was constructed on the Fox River at Appleton, Wisconsin. About 700GW (1 GW = $10^9$ W) of electrical energy is produced by hydroelectric means worldwide, about 20% of the total electric production. Of this total, the United States produces about 80 GW, about 12% of the nation’s consumption.

Hydroelectric plants are expensive to construct, but cheap to operate. Once the capital cost is recovered, no other method of electrical generation can compete on price. The layout of a typical hydroelectric facility is shown in Figure 1.15. The dam creates a head of water, which is converted to kinetic energy to power the turbine. A control gate regulates the amount of
water flowing through the penstock. The rotational energy produced by the turbine feeds the electrical generator whose output is passed through a transformer and on to the power lines. The energy remaining in the outflow water is minimal.

The power output of a hydroelectric plant is given by $P = \eta \rho \rho g h Q = k h Q$ where $\eta$ is the efficiency of the plant, and $Q$ is the volume of water flow per second, and where $k = \eta \rho g$ is a constant. The quantity $\rho g h$ is the potential energy of the water per unit volume, where $\rho$ is the density of water, $g$ is the acceleration due to gravity, and $h$ is the head of water.

“Hydropower use reached a record 3,427 TWh, or about 16.1% of global electricity consumption, by the end of 2010, continuing the rapid rate of increase experienced between 2003 and 2009 . . . . China was the largest hydropower producer and is expected to continue to lead global hydro use in the coming years. The country produced 721 TWh in 2010, representing around 17% of domestic electricity use. China also had the highest installed hydropower capacity, with 213 GW at the end of 2010. It added more hydro capacity than any other country, 16 GW in 2010, and plans to add 140 GW by 2015 [www.worldwatch.org/use-and-capacity-global-hydropower-increases-0].”

The two largest hydroelectric facilities in the world are the 2008, 22 GW Three Gorges Dam on the Yangtze River in China, and the 14 GW Itaipu Plant in Brazil that first began operations
in 1984. Water flow over a year on the Yangtze River is highly variable, so annual production at Itaipu is about the same as that of Three Gorges. Third is the 2013, 13.9 GW Xiluodu facility on the Jinsha River in China, followed by the 10.3 GW Guri Dam in Venezuela, built in 1978. The fifth largest is the 8.4 GW Tucuruí plant on the Tocantins River in Brazil, the first hydro facility in the Amazon rainforest, built in 1984. Next is the venerable 6.8 GW Grand Coulee Dam on the Columbia River in the northwestern United States, which began operations in 1933 [http://water.usgs.gov/edu/hybiggest.html].

New Zealand, Iceland, and Norway generate most of their electricity from hydro. Twelve other countries generate at least 90% of their electric power from hydro. At the other end of the scale, Africa produces little hydroelectric power, but it has the potential for the greatest expansion [op cit].

We have exploited most of the locations in the world that are suitable for large scale hydroelectric power. Smaller facilities offer future opportunities. Small hydro facilities of 0.1 GW or less have the greatest potential in developing countries, and communities at some distance from commercial centers. “As of 2009, roughly 60 GW of small hydro was installed worldwide, accounting for less than 6% of the hydropower total. Small hydro is likely to expand, especially as populous countries like India continue to pursue rural electrification [op cit].”

Once built, hydroelectric plants are carbon neutral. Contrast this with biofuels that burn a source that has captured carbon in the recent past (a few months to a number of years) and emit an equal amount of carbon in the form of CO₂, or fossil fuels that take a stable substance that has lain inert over geological time and is burned to produce CO₂ and other pollutants.

Hydroelectric power is not ecologically neutral. Silting occurs in the lake created by the dam. The flooding that often occurred in the un-dammed situation brought nutrients downstream that fertilize the fields; the dams prevent this, and a classic case of this is the Aswan High dam project in Egypt. Furthermore, fish such as salmon are often prevented from reaching their spawning grounds.

Hydroelectric power is not considered an energy storage system, although it stores potential energy in the reservoir. However, it is common to classify something as an energy storage scheme if it can receive and emit energy in a closed cycle. An energy storage system is a rechargeable battery. Another energy storage method is pumped hydroelectric storage, a close cousin to hydroelectric.

The 14-year drought in the American West, in particular the Colorado River basin, has resulted in Lake Mead being at 40% of capacity. At 43 m down from its filled capacity, Lake Mead is at the lowest level since the Hoover Dam, originally named the Boulder Canyon Dam, was constructed in the 1930s to create Lake Mead, the nation’s largest reservoir at 640 km², 247 square miles. Its power output is about 2 GW.
It is not rainfall that fills Lake Mead, it is the snowpack in the Rocky Mountains, and the winter of 2014–2015 provided less than half the normal amount [K.Siegler, “As Lake Mead Levels Drop, The West Braces For Bigger Drought Impact,” www.npr.org/2015/04/17/400377057/as-lake-mead-levels-drop-the-west-braces-for-bigger-drought-impact].

About 70% of the Colorado River water and its reservoirs go to growing crops to feed the nation, with the rest to major cities such as Las Vegas, Phoenix, San Diego, and Denver. Nevada is the nation’s driest U.S. state, and to build a city like Las Vegas in the middle of a desert is increasingly looking unwise.

The drought in the American West extends all the way into Canada. In July, 2015, the time of this writing, the rain forest near Seattle is experiencing a wildfire that is out of control.

### 1.8.4 The Biomass and Biofuel Potential

Up to the middle of the nineteenth century, wood provided 90% of the energy used in the United States. Biomass/biofuels now account for 48% of all renewable energy sources, but it is a mature industry with little prospects for expansion. Second is hydropower at 35%, also a mature industry, followed by wind energy at 13%, geothermal at 2.5%, and solar at just over 1%. Of these, solar is the most promising and likely to be the dominant renewable energy producer by 2025.


Energy from biomass is a mature source, some would say too mature. Deforestation is just one adverse effect: With 80% of its population living under the poverty line, Haiti is the poorest country in the Western Hemisphere, and while its principal energy source is wood, less than 1.5% of its original tree cover remains. The result is massive erosion. Hurricane Jeanne killed over 3000 people in 2004, many of them by landslides. By contrast, the Dominican Republic, which occupies the eastern side of the island of Hispaniola with Haiti on the western side, has preserved more of its forests, and environmental degradation is less than that for its neighbor.

Energy production in the United States over the years 1973 to 2009 is shown in Figure 1.16 [op cit]. Over these years renewable energy production has risen from 4.43% to 7.78%. The 2009 percentage energy consumption by category is shown in Figure 1.17 [op cit].

Unfortunately, biogasses and biofuels are more expensive than classic fossil fuels, and there is a limit on the amount of biomass that the planet can produce.
1.9 Other Energy Sources

We have seen that solar energy, direct and indirect, is responsible for most energy sources that we have become reliant on. One can easily make the case that we have become overly reliant on one such source, namely fossil fuels. Non-solar renewable energy sources are considered in this section.

1.9.1 Tidal Power

Few locations worldwide, about 20, are suitable for trapping large tidal ranges behind dams called barrages. The trapped water is held until a sufficient head of water (as the tide goes out) exists, and the potential energy in that trapped water is passed through generators to produce electrical power. The tidal flow must be least 5 m for a location to be considered suitable for this system. The bottom line is that estimated capacity of tidal power is 50 times smaller than that for hydroelectric power worldwide.
The tidal cycle is 12 hours and 25 minutes, or 6 hours and 12.5 minutes between high and low tides. Power can be generated close to high or low tide, not in the middle of the tide. Thus, the most one can expect to generate electricity from tidal power is 10 hours in a 24-hour day. This creates problems for reliable power to the residential or commercial consumer. If surplus power can be generated over the 10-hour active period, it may be possible, depending on the neighboring topography, to pump water to a hilltop reservoir, then convert the potential energy stored into electrical energy during the 14-hour down period.

![Figure 1.17: Energy Consumption in the United States in 2009](image)

The first tidal plant was built on the estuary of La Rance, near St. Malo in Brittany, France, with a tidal range of 8 m. Construction started in 1963 and was completed in 1966 for a cost of $620 million. The barrage is 700 m long and uses 24 turbines to generate up to 240 MW and $6 \times 10^{11}$ watt hours/year at a cost of about 1.8c/Kwh, and has paid back its investment. Silting behind the barrage is a problem, and certain fish stocks have been eliminated while other species flourish.

A small plant located in Russia on the White Sea has a capacity of 0.5 MW. Other small tidal plants include the Dutch sea barrages. Far larger at 20 MW and $5 \times 10^{10}$ Wh/year is the Annapolis tidal power station completed in 1984 in the Bay of Fundy in Canada, which has an average 10 m tidal range.

The places in the world with the highest tidal ranges in meters are shown in Table 1.7 [A.M. Gorlov, “Tidal Energy,” Northeastern University, Academic Press, 2001]
According to the Fundy Ocean Research Center for Energy, the Bay of Fundy has a tidal flow of over 100 billion tonnes of water: This is greater than the combined total of all the freshwater rivers and streams in the world. The bay has the potential to generate several thousand megawatts of power [Bay of Fundy Tidal Energy: a Response to the Strategic Environmental Assessment, Nova Scotia Department of Energy, undated].

Table 1.7: Sites with the Greatest Tidal Ranges

<table>
<thead>
<tr>
<th>country</th>
<th>site</th>
<th>range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>Bay of Fundy</td>
<td>16.2</td>
</tr>
<tr>
<td>England</td>
<td>Severn Estuary</td>
<td>14.5</td>
</tr>
<tr>
<td>France</td>
<td>Port of Ganville</td>
<td>14.7</td>
</tr>
<tr>
<td>France</td>
<td>La Rance</td>
<td>13.5</td>
</tr>
<tr>
<td>Argentina</td>
<td>Puerto Rio Gallegos</td>
<td>13.3</td>
</tr>
<tr>
<td>Russia</td>
<td>Penzhinskaya Guba</td>
<td>13.4</td>
</tr>
<tr>
<td>Russia</td>
<td>Bay of Mezen (White Sea)</td>
<td>10.0</td>
</tr>
</tbody>
</table>

The Severn Estuary has a tidal range that averages 14.5 m, suitable for substantial electricity generation. “It has been estimated that a barrage across the Severn River in western England could supply as much as 12 GW, about 10% of the country’s electricity needs [Ocean Energy Council, www.oceanenergycouncil.com/oceanenergy/tidal-energy].”

A barrage across the Severn would have ecological effects, not all of them bad, but such considerations were not the reason for lack of action. Instead, the cost of such a project has denied its construction. A mile long barrage was first proposed by Thomas Fulljames in 1849, whose purpose was to provide a spacious shipping harbor. Electric generation became the main purpose for a 1925 proposal. In 1933 an official recommendation for construction was shelved due to the impending World War II. After the war further studies predicted:

* The barrage would generate $2.2 \times 10^{12}$ Wh/year and cost £60 million in 1948.
* in 1971 the barrage would cost £500, and in that era of cheap oil was not economically viable.
* A 1981 committee recommended a 10-mile-long barrage that would generate $7.2 \times 10^9$ W.
* In 2006 the consulting firm of Parson Brinkerhoff proposed a barrage that would generate $2.75 \times 10^{12}$ Wh/year with an average output of 313 MW.

The latest design, relying extensively on the experience of the French and Canadian models, would have 216 turbines, each generating 40 MW. Sluices would permit the tide in, and then trap it until a substantial head was created. The tidal range would be reduced by about half, and the Severn bore (a tidal wave front) could be completely eliminated.
The construction would take about 8 years, and the life of the project is estimated at between 120 and 200 years. Electric generation time would be about 8 hours. Cost of the project is estimated between £10bn and £34bn.

There have been several generations of plans for capturing tidal power on the Severn, but nothing has been done. The bottom line is the British can find many excuses for delaying major projects. They delayed constructing major highways, called motorways, until the economic loss for their inaction was too painful to bear. It can be argued that the Chunnel, the tunnel under the English Channel, was only constructed since the British were in partnership with the French.

Scotland is considered by some to have the best tidal and wave power potential in the world. The channels between land masses can intensify both tide and wave action, and the proximity to land eases the logistics for connecting the electricity generated to the consumer.

“The crown estate and Scottish government today unveiled a £4bn project to build 10 wave and tidal power sites around the Orkney islands and the Pentland Firth, with the potential to power up to 750,000 homes. The narrow sea channel has some of the most powerful currents and tidal surges in the world, with speeds up to 16 knots or 19 mph recorded. The area also experiences some of the biggest waves in the UK. OpenHydro, a large underwater turbine resembling a jet engine and bolted to the sea floor, is built by Cantick Head Tidal and will harness the firth’s fierce tides at a 200 MW site south of Orkney” [guardian.co.uk, 16/3/2010].

### 1.9.2 Geothermal Energy

Geothermal is a word derived from the two Greek words geo (earth) and therme (heat). Geothermal energy means extracting heat from the Earth’s surface or subsurface. It can tap into hot springs, or heat water in naturally hot rocks. It can also use a heat pump to extract energy from cool earth or water. Here we describe where heat can be extracted from the ground.

Most volcanic and hot spring activity occurs close to a tectonic plate. The intersection locations of the major tectonic plates of the world is shown in Figure 1.18, and the red circles indicate the places of major geothermal activity. The Pacific plate is limited to the north, east, and west by what is called “the ring of fire,” with major activities along the U.S. west coast, Alaska, the Korean peninsula, Japan, and Indonesia.

A major volcanic eruption occurred at Mount St. Helens in the Cascade Range of Washington State on May 18, 1980. There were plenty of warnings of a pending eruption, but volcanology is a non-exact science. Fifty seven people died. The eruption produced a magnitude 5.1 earthquake, and over $5 \times 10^5$ kg of ash shot 24 km into the sky. Half the mountain was ripped off and it is estimated that material shot sideways at over 480 km/h, destroying all in its path. The volcano remains active, and the caldera is rebuilding.
An undersea eruption can trigger a tsunami, what is often referred to incorrectly as a tidal wave in England. On December 26, 2004, a 9.0-magnitude earthquake in the Indian ocean near the island of Sumatra unleashed the energy of 23,000 Hiroshima-sized atomic bombs and triggered the largest tsunami in recorded history. The USGS estimated that the ocean floor was displaced vertically by about 9 m over an almost 1000 km front. The resulting tsunami killed 150,000 people.

The most lethal characteristic of a tsunami is its innocuous start. The ocean recedes, and all seems still. Then the ocean rises and rises and does not stop, pushing everything in its path. The most lethal part occurs when the water recedes, holding people in its grip. Many of these are never seen again. Some bodies wash onto shore days or weeks later.

On March 11, 2011, an 8.9-magnitude earthquake about 130 km off the northeast shore of Japan caused a massive tsunami. A 9 m wall of water penetrated about 5 km inland. In all, 15,839 were killed and 3,642 remain missing. The most destructive ecological result was to the Fukushima Daiichi nuclear power plant. The plant was flooded and the operators could not cool the reactors because they had located the emergency diesel generators in the basements. Meltdowns at reactors #1 and #3 caused hydrogen explosions that tore the roofs off and radiation escaped, requiring evacuation of the surrounding area. Three other reactors managed to achieve cold shutdown.

The country with the most hot springs in the world is Iceland. It lies on the mid Atlantic Ridge, which is the intersection of two tectonic plates. It has about 800 hot springs with an average temperature of 75° C, enough to heat structures by circulating water at that temperature through radiators. The name of its capital city Reykjavik means “bay of steam,” and hot water is piped to every structure. Geothermal energy heats about 85% of the structures in Iceland, including
greenhouses, thus taming Iceland’s frigid climate by producing tropical fruits. Iceland is famous for its naturally heated mineral baths. It is also famous for having Geysir, a 180’ water column we call a geyser, so we know where that word is derived.

One of the most popular tourist spot in the United States is Yellowstone National Park in the northeast corner of Wyoming. Its principal attraction is its hot springs and more than 300 geysers in a beautiful natural setting. The most famous is Old Faithful, so called for the regular periodicity of eruption: It spews from \(10^4\) to \(3 \times 10^4\) liters of water over 50 m in the air.

### 1.9.3 Nuclear Power

The electrical power generated by nuclear fission in 2012, according to the U.S. Energy Information Administration [www.eia.gov/energyexplained/index.cfm?page=nuclear_power_plants](http://www.eia.gov/energyexplained/index.cfm?page=nuclear_power_plants) is shown in Figure 1.19. The numbers are in Wh times \(10^{12}\). Although the United States is the largest generator of such power, only 19% of its domestic needs are supplied by nuclear means. By way of contrast, France uses nuclear to supply 78% of its needs.

Before the tsunami took out the nuclear facilities in Fukishima, Japan was the third largest generator of electricity by nuclear means, with about \(26 \times 10^{12}\) Wh annually. The shock to the nation has soured it on nuclear power. Similarly, Germany has declared its intention to reduce or eliminate electrical generation by nuclear power.
“The USA is the world’s largest producer of nuclear power, accounting for more than 30% of worldwide nuclear generation of electricity. The country’s 104 nuclear reactors produced 821 billion kWh in 2011, over 19% of total electrical output. There are now 100 units operable and three under construction. Following a 30-year period in which few new reactors were built, it is expected that 4–6 new units may come on line by 2020, the first of those resulting from 16 license applications made since mid-2007 to build 24 new nuclear reactors. However, lower gas prices since 2009 have put the economic viability of some of these projects in doubt.

Government policy changes since the late 1990s have helped pave the way for significant growth in nuclear capacity. Government and industry are working closely on expedited approval for construction and new plant designs [“Nuclear Power in the USA,” updated July 31, 2013, www.world-nuclear.org, as accessed Sept 17, 2013].”

Most nuclear power plants in the United States, because of their large generating capacity, require complicated and expensive pumps to ensure that overheating, and a possible nuclear meltdown, does not occur. “However, if safety is the primary goal, as it is in the United States today, it is much easier to assure that adequate cooling will be available if there is only half as much heat to dissipate. In fact, in a 1,200,000 kW reactor, cooling requires elaborate pumps, while in a 600,000 kW reactor it can be handled by simple gravity flow with natural convection — cool water enters the bottom of the reactor, which heats it, causing it to rise because warm water is lighter. This process sets up a natural circulation driven only by gravity. Unlike pumps which can fail and are driven by electric power which may not always
be available, gravity never stops working. That makes the 600,000-kW reactor inherently safer. This is called ‘passive stability,’ since no active measures by operators or by mechanical or electrical control systems are required. The operator could shut off the electric power and go home without any harm coming to the reactor [University of Pittsburgh, Department of Physics and Astronomy, www.phyast.pitt.edu, as accessed Sept. 17, 2013].”

New designs to semi-passively produce automatic shutdown include a frozen salt plug in the bottom of the reactor vessel. In event of a problem the cooling of that plug ceases, the plug melts, and the nuclear material flows into a large holding tank below.

Operator error is minimized if future plants are all the same design and layout. Everything should be standardized, including the furniture and color scheme, so a trained engineer is instantly familiar with everything when walking into a plant for the first time.

Any mention of nuclear power often invokes strong societal reaction. There are some who claim it is the ultimate carbon-free energy source with an admirable safety record. Others cite Chernobyl and Fukushima, and believe that Three Mile island was a near disaster, and also claim that the waste products of nuclear fission, with their almost infinite half lives, are a ticking time bomb of nuclear radiation. The reality is somewhere between the opposite poles.

In response to the Fukushima Daiichi disaster, the German government closed seven nuclear power stations, and another plant would not reopen after being on shutdown since 2009. Soon thereafter it was announced that all its nuclear facilities would be permanently decommissioned by 2022.

Nuclear fission has two main problems. First, once the chain reaction starts, it is a delicate task to control it so it does not spin out of control, as happened at Chernobyl. An analogy to gasoline is to set fire to the fuel tank, and then control the oxygen it receives so it does not explode. Second, the byproducts are radioactive for many years. For example, plutonium-239 has a half life of 24,000 years; half life is the time for its radioactivity to reduce by half. After 10 to 20 half lives, a radioactive material is no longer considered hazardous.

Contrast this to nuclear fusion, where reducing the hydrogen feed slows the reaction so it can never get out of control. Also, there are no dangerous, radioactive byproducts.

The structure of the thermonuclear bomb, known as the hydrogen bomb, is an atomic bomb, a fission device. Around the atomic bomb is a layer of lithium and an isotope of hydrogen called deuterium. Around this is an outer layer designed to increase the temperature and pressure immediately after the atomic explosion so that nuclear fusion occurs.

The first hydrogen bomb was tested at the Eniwetak Atoll in the Marshall Islands in 1952. It had the power of $10^7$ tons of TNT. Four other nations followed in developing their own hydrogen bomb.
A recent article explains advances in the science and technology of fusion physics [L. Grossman, “A Star is Born,” Time Magazine, November 2, 2015]. Since the World War II era there has been considerable effort and expense to produce a sustainable fusion reaction for peaceful purposes. After billions of dollars expended we know a lot more about the mechanism of the plasma, the ionized cloud of free electrons, protons, and neutrons, that needs to get to an incredibly high temperature and pressure in order for a fusion reaction to occur. The Sun can do it since its mass is large enough for gravity to produce the required pressure, and the continual fusion reaction sustains a temperature of about 17°K million.

The configuration that dominates the fusion landscape is the tokamak, a hollow metallic doughnut surrounded by electromagnetic coils to produce an intense magnetic field. Tokamaks are costly. The International Thermonuclear Experimental Reactor being constructed near Marseilles, France is behind schedule and its cost has risen from $5 billion to $20 billion [op cit]. None of the tokamaks have yet reached a sustainable reaction.

Most money for the research into nuclear fusion has come from governments, much of it to universities. This may be changing. Now, private money is trying new techniques. With such money, Tri Alpha uses two cannons firing plasmas straight at each other to produce massive pressure and temperature between them, into which hydrogen atoms are injected. Who knows if they can reach the break-even point with a fusion reaction, but early data is encouraging [op cit].

The classic joke is that sustainable, nuclear fusion will be achieved in 30 years, and this will be true into the future. However, if that sustainability is reached there will be an avalanche of resources committed, and a limitless energy source may result. Until then, we need to produce more and more renewable energy from the Sun and wind.

1.10 Energy Use in the United States

In the United States the breakdown of energy use in 2009, excluding agriculture, was:

<table>
<thead>
<tr>
<th>Sector</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>industrial</td>
<td>30%</td>
</tr>
<tr>
<td>transportation</td>
<td>29%</td>
</tr>
<tr>
<td>commercial buildings</td>
<td>19%</td>
</tr>
<tr>
<td>residential buildings</td>
<td>22%</td>
</tr>
</tbody>
</table>

In this section we shall examine how energy is used in these sectors.

There is no way that we can or will dismantle the mechanisms of production. It would mean returning to the farms that existed before the industrial revolution. It would result in famine, war, and pestilence. But what can be done to produce a sustainable system that reduces our carbon footprint? The answer is a lot, provided there is a political will to do so.
1.10.1 Energy in the Industrial Sector

Substantial reductions have been made in U.S. industrial energy usage since 1975 [www.need.org/needpdf, as accessed 7/13/12]:

- Petroleum refineries 30%
- Chemicals 47%
- Steel 45%
- Aluminum 58%
- Paper and pulp 42%
- Cement 38%

The Energy Used in Petroleum Production

“The petroleum refining industry in the United States is the largest in the world, providing inputs to virtually any economic sector, including the transport sector and the chemical industry. The industry operates 146 refineries (as of January 2004) around the country, employing over 65,000 employees. The refining industry produces a mix of products with a total value exceeding $151 billion. Refineries spend typically 50% of cash operating costs (i.e., excluding capital costs and depreciation) on energy, making energy a major cost factor and also an important opportunity for cost reduction. Energy use is also a major source of emissions in the refinery industry making energy efficiency improvement an attractive opportunity to reduce emissions and operating costs [http://industrial-energy.lbl.gov/node/79, as accessed on July 25, 2012].”

“The U.S. petroleum refining industry provides 23% of world production and including inputs to virtually all economic sectors, including the transportation and the chemical manufacturing sectors. Significant quantities of energy are used to produce the petroleum products used by consumers and industry. The petroleum refining industry is one of the most energy-intensive manufacturing industries in the U.S. The industry used 3.1 quadrillion BTUs in 2002. The most energy-intensive refining processes include: distillation, hydrotreating, alkylation and reforming [www.epa.gov/region9/waterinfrastructure/oilrefineries.html, as accessed on July 25, 2012].”

The Power Needed in the Chemical Industry

“The US chemical industry has already been ahead of the energy efficiency curve for years, and it has reduced its fuel and power energy consumed per unit of output by 53% since 1974, according to the American Chemistry Council (ACC). . . . ACC president and CEO Cal Dooley noted total annual energy savings of 18.9 trillion BTUs achieved from 56 US projects implemented by several chemical firms last year. The total carbon dioxide emissions reductions achieved by the projects were 3.3m tonnes [www.icis.com/Articles/2009/08/17/9238790, as accessed on July 25, 2012].”
“The chemical industry accounts for 6 percent of energy usage in the United States (Wells, 2008). Approximately half of this energy is contained in hydrocarbon raw materials - primarily from oil and natural gas. The other half is used to transform raw materials into useful chemical products through reaction and purification steps.

“The key difference between the chemical industry and the fuels sector is that, in the production of chemicals, most of the enthalpy of the starting materials is preserved in the final products. In the fuels sector, the enthalpy is completely consumed to generate energy . . . .

“Because of the magnitude of its energy consumption, the chemical industry is motivated to conserve, and U.S. producers have reduced their fuel and power usage per unit output by nearly half since 1974. But opportunities for energy savings go beyond internal consumption. Because more than 96 percent of manufactured goods involve chemistry, the industry can also improve energy usage by consumers through the careful shaping of the product life cycle . . . .

“The chemical industry consumes large amounts of finite energy sources to process raw materials that are in limited supply. The industry has undertaken five major initiatives to improve its energy efficiency: (1) improving existing processes; (2) commercializing new processes; (3) recycling waste; (4) investing in renewable raw materials; and (5) creating products that enable energy savings. Innovation in all of these areas is an absolute necessity for long-term sustainability [www.nae.edu/Publications/Bridge/EnergyEfficiency14874, as accessed on July 25, 2012].”

Recycling and Energy Savings in the Steel and Aluminum Industries

“As our country ‘breaks through’ to new, green energy supplies, the North American steel industry must ‘breakthrough’ to new production methods that take advantage of these clean energy sources. The North American steel industry has continually reduced its need for energy, thus minimizing our footprint on the environment. Since 1990, energy intensities to make one ton of steel have been reduced by 27 percent. Because of these advances the steelmaking processes we use today are approaching the limits defined by the laws of physics. To make further reductions in energy use and CO₂ emissions, new processes are required. The CO₂ Breakthrough Program is an international research project striving to meet this challenge . . . .

“Manufacturing steel by today’s steelmaking process produces CO₂ as a by-product.” CO₂ is one of the major Greenhouse Gas identified as contributing to climate change. On average 1.2 tons of CO₂ was emitted in 2009 for every ton of steel produced in the United States. Today, the American steel industry operates with the lowest average energy consumption per ton of steel produced. Because of the close relationship between energy use and GHG emissions, the industry aggregate CO₂ emissions per ton of steel shipped were reduced by approximately 33 percent since 1990.
“Because of industry’s voluntary investments in R&D and resulting new technology, US
steelmaking processes are highly optimized, and efforts will be made continue to achieve
incremental improvements. However, in order to make major reductions in future energy/ CO₂
reductions, new methods of making steel will require completely fresh and innovative thinking.

“The North American steel industry has been actively investing in research and development
into new transformational processes for making steel that will dramatically reduce or eliminate
CO₂ emissions. This R&D is called the AISI CO₂ Breakthrough Program

“The U.S. Geological Survey (USGS) reports that bauxite is the only raw material used on a
commercial scale in the United States in the production of alumina and aluminum (NAICS
3313). As a general rule, four tons of dried bauxite is required to produce two tons of alumina,
which in turn provides one ton of primary aluminum metal (NAICS 331312). As reported in
USGS Mineral Commodity Summaries 2006, in 2005:
* Nearly all of the bauxite consumed in this country was imported; more than 90 percent was
converted to alumina at domestic refineries located in Louisiana and Texas.
* Of the total alumina used domestically, about 90 percent went to primary aluminum smelters.
* Six companies operated 15 primary aluminum smelters at about two-thirds of rated or
engineered capacity; another four smelters were idle. All modern primary aluminum smelting
plants employ the ‘Hall-Heroult’ process to reduce alumina to aluminum through electrolysis .

“The industry-wide average energy consumption per kilogram of aluminum production has
generally declined in recent years through a number of factors: (1) the closure of older, more
energy-intensive ‘Soderberg’ smelters in the Pacific Northwest; and (2) the implementation of
best management energy efficiency practices, including (a) improvements in the molten cryolite
chemical bath composition; (b) improved training of cell operators and monitoring to reduce
anode effects (AE); (c) use of improved, computerized cell control systems and other process
controls to prevent AE; and (d) installation of alumina point feed systems. As is the case with
other capital-intensive industries, replacing older equipment/processes with state-of-the-art
equipment/processes holds potential for energy efficiency improvement. In 2000, typical
energy consumption achieved by operating smelters was between 13 kWh/kg of Al for state-of
the-art facilities (e.g., point feed pre-bake) to 20 kWh/kg of Al for older Soderberg smelters
(many of which were located in the Pacific Northwest and have now been shut down).

“Aluminum recycling also has an impact on sector energy use, as production from recycled
aluminum requires only five percent of the energy required for primary ore production.
Recycling one kilogram of aluminum can save up to 14 kWh of electricity
[www.epa.gov/sectors/pdf/energy/ch3-1.pdf, as accessed on July 25, 2012].”

The Paper and Pulp Industry
“In 2002 the U.S. Paper Industry produced 99.5 million tons of pulp and paper products while consuming 2,361 trillion BTUs. It should be noted that since 2002, the Pulp and Paper Industry has reduced its energy consumption, primarily through the use of waste energy streams, i.e. capturing the energy in waste heat streams, both air and liquid, as well as installing energy saving devices such as variable speed motors and more efficient lighting. [www1.eere.energy.gov/manufacturing/industries_technologies/forest, as accessed on July 24 2012].”

“The forest products industry consumed almost 3.3 quads of energy in 1998. This represents about 14% of domestic manufacturing energy use, making the forest products industry as a whole the third largest industrial consumer of energy, behind only petroleum and chemicals. Within the forest products industry, the pulp and paper industry uses the vast majority of the energy, 2.75 quads, while the wood products industry uses only 0.51 quad. The industry has made good use of wood residues and byproducts (black liquor) and self-generates a large portion of its own energy needs. According to the 1998 Manufacturing Energy Consumption Survey (MECS), the wood products industry produced almost 50% of its energy requirement by burning wood residues from tree harvesting and sawmill operations. In 1998, the pulp and paper industry self-generated about 46% of its energy needs, and the industry itself estimates this figure to be closer to 55% in 1997. The pulp and paper industry spent $7.6 billion for energy in 1998, roughly 3.0% of the value of shipments in that year, and remains the manufacturing sector’s fourth largest consumer of fossil fuels[www.eia.gov/emeu/meecs/iab98/forest/energy_use, as accessed on July 24, 2012].”

Energy in Cement Production

The world cement industry uses more than 300TWh of electricity, while Germany uses about 3TWh. Cement production consumes 150 Kwh/ metric ton.

“Just as cement is the binding -?- glue in concrete production, the U.S. cement industry is the building block of the nation’s construction industry. Cement manufacturing accounts for 1 to 2 percent of U.S. industrial energy use, but more than 5 percent of the nation’s industrial carbon dioxide (CO2) emissions.

“The most energy-intensive step in modern cement manufacturing is the calcination reaction, which requires extremely high temperatures — up to 3,000 degrees Fahrenheit (1,700 degrees Celsius) — in order to transform limestone (calcium carbonate) into lime (calcium oxide), a necessary component of cement. CO2 emissions result from the combustion of fuel used to reach these high temperatures, but they also are produced as a by-product of this calcination reaction. Overall, the resulting emissions are disproportionately large when compared to those produced by other industries. With few viable alternatives currently available and the worldwide demand for cement increasing, investment in energy- and CO2-reducing technologies and processes in the cement industry represents one key opportunity to help the
United States attain its energy and climate goals [www.reliableplant.com/Read/23445, as accessed July 23, 2012].”

S Peddanna, in “Energy Audit to Evolve Energy Conservation Measures,” ERCOM Consulting Engineers Pvt. Ltd, India, wrote: “Overall energy efficiency of cement production in a modern plant is quite low. About 60% of total electrical energy is used in grinding of the raw materials, coal and clinker. The energy efficiency of grinding systems employed in cement plants is estimated as only about 20%. It means that about 80% of grinding energy i.e. about 48% of total energy of the plant is wasted as heat, sound and vibration [www.energymanagertraining.com/Journal/24032006, as accessed July 23, 2012].”

**Government Regulation and the Reaction from Industry**

The common reaction of the U.S. industrial sector to regulation is to protest, typically saying it will make their products too expensive and/or noncompetitive. A case in point was and is the objection to pollution controls and tighter standards on power station smoke stacks. The cost of compliance was projected to be orders of magnitude greater than the actual cost.

Another case was the objection of automotive manufacturers to the requirement on the installation of air bags in cars — far too expensive. Yet within a few years they all touted the great safety feature that they, voluntarily, and in the spirit of public service, had installed to save lifes.

Two of the most important and successful U.S. government regulations of all time are the CWA (Clean Water Act) and the CAA (Clean Air Act). It is now difficult to see these acts being implemented in the highly partisan political environment at the time of this writing.

“The Clean Water Act (CWA) establishes the basic structure for regulating discharges of pollutants into the waters of the United States and regulating quality standards for surface waters. The basis of the CWA was enacted in 1948 and was called the Federal Water Pollution Control Act, but it was significantly reorganized and expanded in 1972.

“Under the CWA, EPA has implemented pollution control programs such as setting wastewater standards for industry. We have also set water quality standards for all contaminants in surface waters.

“The CWA made it unlawful to discharge any pollutant from a point source into navigable waters, unless a permit was obtained. EPA’s National Pollutant Discharge Elimination System (NPDES) permit program controls discharges. Point sources are discrete conveyances such as pipes or man-made ditches. Individual homes that are connected to a municipal system, use a septic system, or do not have a surface discharge do not need an NPDES permit; however, industrial, municipal, and other facilities must obtain permits if their discharges go directly to surface waters [www.epa.gov/lawsregs/laws/cwa.html, as accessed July 24, 2012].”
“The Clean Air Act (CAA) is the comprehensive federal law that regulates air emissions from stationary and mobile sources. Among other things, this law authorizes EPA to establish National Ambient Air Quality Standards (NAAQS) to protect public health and public welfare and to regulate emissions of hazardous air pollutants.

“One of the goals of the Act was to set and achieve NAAQS in every state by 1975 in order to address the public health and welfare risks posed by certain widespread air pollutants. The setting of these pollutant standards was coupled with directing the states to develop state implementation plans (SIPs), applicable to appropriate industrial sources in the state, in order to achieve these standards. The Act was amended in 1977 and 1990 primarily to set new goals (dates) for achieving attainment of NAAQS since many areas of the country had failed to meet the deadlines.

“Section 112 of the Clean Air Act addresses emissions of hazardous air pollutants. Prior to 1990, CAA established a risk-based program under which only a few standards were developed. The 1990 Clean Air Act Amendments revised Section 112 to first require issuance of technology-based standards for major sources and certain area sources. ‘Major sources’ are defined as a stationary source or group of stationary sources that emit or have the potential to emit 10 tons per year or more of a hazardous air pollutant or 25 tons per year or more of a combination of hazardous air pollutants. An ‘area source’ is any stationary source that is not a major source [op cit].”

1.10.2 Transportation in the Modern World

Energy efficiencies can be wrung out of transportation. The passenger car, often traveling with a single occupant, is very wasteful. Too often, that single occupant travels many miles a day to office work in an oversized pickup truck that needs a gallon of gas for every 14 miles traveled. The government had to act, since the market had no will to. “First enacted by Congress in 1975, the purpose of CAFE (Corporate Average Fuel Economy) is to reduce energy consumption by increasing the fuel economy of cars and light trucks. NHTSA (National Highway Traffic Safety Administration) administers the CAFE program, and the EPA (Environmental Protection Agency) provides the fuel economy data. NHTSA sets fuel economy standards for cars and light trucks sold in the U.S. while EPA calculates the average fuel economy for each manufacturer. This site contains an immense amount of information about the CAFE program including a CAFE overview, rulemaking actions, fleet characteristics data, compliance activities, summaries of manufacturers’ fuel economy performances since 1978, and related studies [www.nhtsa.gov/fuel-economy as accessed on July 20, 2012].”

“It’s official. After a long and contentious battle, the government and automakers have settled on the new 2025 Corporate Average Fuel Economy (CAFE) regulations that will begin taking effect in 2017. At a ceremony today in Washington, D.C., President Obama announced that the new CAFE standards for vehicle fleets will be 54.5 mpg by 2025. The increase piggybacks
Obama’s 2009 mandate for a CAFE average of 35.5 by 2016 and is the largest mandatory fuel economy increase in history. The standard is just shy of the 56.2 mpg average that the Obama administration was considering just a month ago (though not nearly as ambitious as the 62 mpg target the government floated at one time).

“The ambitious new standards have encountered strong opposition from automakers, who suggest that the rules will mean large increases in cars’ sticker prices. But, as part of the announcement (where the CEOs of the Detroit big three and several foreign automakers were in attendance), Obama said that consumers would save an average of $8000 per vehicle in reduced fuel costs once the regulations are in full effect in 2025.

“Make no mistake about it: The new regulations are hugely important. They will save consumers boatloads of money they would’ve spent on gas, drastically reduce American’s fuel consumption and carbon footprint and change the way cars are made. But, they present a major challenge to automakers, who must determine what technologies or combination of technologies will allow average fleet fuel economy to climb so high. They’ve got a lot of work to do [www.popularmechanics.com/cars/news/fuel-economy, as accessed on July 20, 2012].”

So how are the automobile manufacturers going to achieve these mileage figures? With smaller cars, all electric and hybrid electric vehicles.

In contrast to the automobile, the bicycle is the most efficient means of transportation for a single person. In some cities, particularly in Asian countries, the bicycle is the preferred means of transportation. Some European cities are bicycle friendly, particularly in the low countries. Amsterdam is a notable example of this.

For long-distance land transportation it is difficult to match the efficiency of the locomotive. A current advertisement on radio claims the rails can transport a ton of freight an average of 480 miles on a single gallon of fuel. “America’s passenger rail is a global joke, but our freight rail carries over 40% of our intercity cargo. Trains carry much less of Europe’s freight, which is why trucks clog Europe’s highways. And America’s rail shipping rates are the world’s lowest, reducing the cost of doing business in the U.S.; they’ve fallen 45% in real dollars since the industry was deregulated three decades ago . . . . I love railroads because they’ve got all the right enemies. The welfare queens of Big Ag, who whine about government interference while relying on government handouts, want congress to lower the prices railroads charge to ship their subsidized grain and ethanol. The similarly retrograde King Coal, which resists regulations that could stop it from poisoning our air and water, also clamors to reregulate freight rates [Michael Grunwald, ‘Back on tracks,’ Time magazine, July 9, 2012].”

So why don’t the railroads carry more than 40% of intercity freight? The answer is partially political — hidden subsidies make long distance trucking economically competitive. The flexibility of the truck compared to a rail car for short hauls is difficult to match. However, when the hidden subsidies and the pollution coats of road transportation are factored in, it is evident that the country needs more goods to go by rail and less to go by road.
One other factor that reduces transportation costs is the computer. Now, an increasing number of people work from their homes, eliminating the energy used in a daily commute.

Table 1.8: Energy Use by Category in the United States

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>natural gas</td>
<td>68.6</td>
<td>60.5</td>
<td>64.6</td>
<td>57.0</td>
<td>57.9</td>
<td>53.4</td>
<td>51.7</td>
<td>52.5</td>
</tr>
<tr>
<td>electricity</td>
<td>31.3</td>
<td>30.1</td>
<td>29.7</td>
<td>28.9</td>
<td>28.8</td>
<td>30.5</td>
<td>32.2</td>
<td>35.2</td>
</tr>
<tr>
<td>oil/kerosine</td>
<td>22.1</td>
<td>19.0</td>
<td>15.9</td>
<td>13.6</td>
<td>14.7</td>
<td>13.5</td>
<td>11.1</td>
<td>10.7</td>
</tr>
<tr>
<td>petroleum gas</td>
<td>4.0</td>
<td>4.4</td>
<td>3.7</td>
<td>3.5</td>
<td>3.6</td>
<td>3.5</td>
<td>3.0</td>
<td>3.6</td>
</tr>
<tr>
<td>total</td>
<td>126</td>
<td>114</td>
<td>114</td>
<td>193</td>
<td>105</td>
<td>101</td>
<td>98</td>
<td>102</td>
</tr>
</tbody>
</table>

### 1.10.3 The Commercial and Residential Sectors

Commercial buildings have rarely been designed for energy efficiency. Too often the architectural statement seals the deal, and the mechanical functioning is an afterthought. When function follows form, inefficiency results. Happily, this line of thinking is diminishing, and form is following function. More commercial buildings are being designed and built to LEED standards: see Section XI of Chapter 3. The Bullitt Center in Seattle is a 50,000 ft² office complex that uses photovoltaic panels on the roof, a system of geothermal wells for space heating, composting toilets to reduce water usage, windows programmed to automatically open and close for temperature regulation, and daylighting to reduce the need for electrically generated lighting [Tim Newcomb, “Silver Bullitt: A Seattle Office Goes Ultra-Green,” Time magazine, July 2, 2012]. The result is that the energy reduction compared to prior norms of 78% and a LEED platinum rating.

Average annual residential energy use in million BTUs for the years 1980 through 1987 [www.allcountries.org/uscensus/949, as accessed 7/16/12], based on the U.S. Energy Administration, broken down into categories, is given in Table 1.8. One can assume that the oil/kerosine and petroleum gas usage goes to space heating. Most of the natural gas usage also goes to space and water heating, although a small amount will go to cooking.

### 1.10.4 Food Production

Food production consumes (no pun intended) about 10% of the total energy used in the United States [sustainabletable]. Inefficiencies account for a considerable chunk of this 10%. Livestock are raised on one place, and the grain they eat in another. The distances between them preclude the use of the manure produced, so the manure is largely wasted while the grain needs synthetic fertilizers produced from atmospheric nitrogen and natural gas. “They have too much of it, and not enough acreage, if any, where to spread it. This leads to accumulation of manure and other related problems, such as stench, high concentrations of minerals in the soil.
and eventually in the waterways and drinking water reserves [http://hfgfoodfuturist.com/2011/03/12/the-fertilizer-of-the-future/ as accessed on 7/18/12].”

About 23% of the energy to put food on our tables is used in packaging and processing. Another 32% is for domestic refrigeration and cooking [sustainabletable].

“Among the many challenges that the agriculture of the future faces, soil fertility ranks high on the list of priorities. Originally, most farms were mixed. They had land to grow crops and they had animals for milk, eggs and meat. Markets were mostly local, and food was consumed in the villages and towns near the farms. Food waste was fed to farm animals; the manure produced was mixed with straw and returned to the fields where the crops had been grown. Over time, farming has evolved. Agriculture has become much larger scale, global and specialized. This evolution has been driven by the use of oil, mechanization, and by the development of mineral fertilizers.

“That model, which has been greatly based on cheap energy and resources, needs to be looked at critically as the economic environment changes. Energy is no longer cheap and, like oil, the resources used for the production of fertilizers have been depleted. New solutions are required to be able to produce optimally.

“The production of nitrogen fertilizers requires a lot of energy. According to estimates, it uses 5% of the world’s natural gas production, and half the fossil fuels used in agriculture. Because nitrogen is quite mobile when dissolved, as this happens when it rains, a large amount of these high-energy-consumption compounds are lost. An estimated 50% of the nitrogen spread on crops leaches through the soil. It ends up in the water system [http://hfgfoodfuturist.com/2011/03/12/the-fertilizer-of-the-future/ as accessed on Jul. 18, 2012].”

Ethanol is added to gasoline as an oxygenator, which allows for complete combustion and so reduces tailpipe emissions. In the United States, ethanol is produced from corn, and in the year 2009, 25% of the corn crop was diverted from food and into ethanol. It is interesting that in the summer of 2012, when this section was being written, the corn belt areas of the United States were experiencing a drought of historic proportions, and food prices were expected to shoot up. Perhaps we should reverse our policy of imposing stiff tariffs on sugar-cane-based ethanol from Brazil.

1.11 Global Warming

There is no doubt that global warming caused by human activity is occurring, and occurring at an ever-increasing rate. Scientists agree that the dominant cause is greenhouse gases in the atmosphere.
1.11.1 Greenhouse Gases

The gaseous composition of the atmosphere is

\[ \text{N}_2 78.08\% \quad \text{O}_2 20.95\% \quad \text{Ar} 0.934\% \]

This leaves 0.036\% for trace gases, whose principle components are

\[ \begin{align*}
\text{CO}_2 & \quad 3.7 \times 10^{-2}\% \\
\text{CH}_4 & \quad 1.7 \times 10^{-4}\% \\
\text{N}_2\text{O} & \quad 3 \times 10^{-5}\% \\
\text{O}_3 & \quad 4 \times 10^6\%
\end{align*} \]

The relative effectiveness of greenhouse gases with respect to CO$_2$, so CO$_2$ is unity, is

\[ \begin{align*}
\text{H}_2\text{O} & \quad 0.1 \\
\text{CH}_4 & \quad 30 \\
\text{N}_2\text{O} & \quad 160 \\
\text{O}_3 & \quad 2,000 \\
\text{CFC} & \quad 22,000
\end{align*} \]

where CFC stands for chlorofluorcarbons. Water in the atmosphere is highly variable. It is low over the poles and high over the oceans in the tropics. We will use the lower and upper limits of 0.1\% and 4\%.

Multiplying the effectiveness by the amount of these gases in the atmosphere and normalizing, so again CO$_2$ is unity, produces numbers for the real effect on the greenhouse gases as

\[ \begin{align*}
\text{CO}_2 & \quad 1 \\
\text{H}_2\text{O} & \quad 0.01-0.4 \\
\text{CH}_4 & \quad 5.1 \times 10^{-3} \\
\text{N}_2\text{O} & \quad 4.8 \times 10^{-3} \\
\text{O}_3 & \quad 8 \times 10^{-3}
\end{align*} \]

So the dominant greenhouse gas that is causing global warming is CO$_2$, and as can be seen for recent years in Figure 1.20 this gas is rapidly increasing, and the cause is the burning of fossil fuels. The data for this figure was collected at Mauna Loa Observatory, Hawaii [Scripts Institution of Oceanography, NOAA Earth System Research Laboratory].

The big global emitters of CO$_2$ are shown in Figure 1.21. The ordinate is the number of giga tonnes of CO$_2$ [www.globalcarbonproject.org/carbonbudget/14/GCP_budget_2012]. The increase in Chinese emissions is frightening.
$O_3$ (ozone) is a greenhouse gas, and as such has a deleterious effect, but it also has the benefit of absorbing harmful ultraviolet radiation. It was found that large holes in the ozone layer over Antarctica was due to CFCs, the coolant used in air conditioners and refrigerators. CFC use has since been dramatically reduced following the 1987 Montreal accord. Unfortunately, CFCs have a lifetime of from 60 to 300 years in the atmosphere, so it could be a considerable time before these ozone holes are naturally repaired.

![Figure 1.20: Atmospheric Carbon Dioxide](www.electronicbo.com)

Methane has increased by over 150% since 1760. Major causes include industrial pollution and flatulence by farm animals. Nitrous oxide is increasing by 0.2% to 0.3% per year, primarily due to human activity [www.physicalgeography.net/fundamentals].

Does human respiration contribute to $CO_2$ in the atmosphere? The online consensus opinion is that humans are part of a natural cycle in which plants extract $CO_2$ from the atmosphere and give out $O_2$, humans eat the plants for their nutrition, breathe in $O_2$, and exhale $CO_2$. Without human interference, the plants would die and natural decay would release their $CO_2$ to the atmosphere. Thus, humans just add another step in a natural process. Unfortunately, this is simplistic thinking.

When humans were hunter/gatherers the just-presented scenario was basically correct. In that epoch, the number of humans on the planet was minuscule compared to the present. There are now about 7.3 billion people in the world, and few if any of us are hunter/gatherers. Many of us, a large majority in advanced nations, live in cities and buy our food, and that food is increasingly grown in a mono-culture environment with advanced machinery, pesticides, and fertilizers made from fossil fuels. That food needs to be shipped to the end users using fossil-fuel-based transportation; many are concerned about the number of food miles to get our food to market. Thus, just the growing of our food is no longer a natural process.
A number that appears approximately the average amount of CO₂ exhaled by a human is 0.9 Kg/day, so worldwide the total CO₂ entering the atmosphere from human respiration is $0.9 \times 365 \times 7.3 \times 10^9 = 2.4 \times 10^{12}$ Kg/year. The amount of CO₂ produced by combustion of fossil fuels, some of it in food production and transportation, is estimated at about $35 \times 10^{12}$ Kg, so our respiration directly adds almost 7% of CO₂ to the atmosphere. Yes, a significant fraction of the 7% is part of a growth/harvest/eat cycle, but it is nowhere near all of it.

The problem is numbers. A little boy peeing into a stream is not a problem, but a village using the stream that runs through it as an open sewer is a problem.

### 1.11.2 The Effects of Greenhouse Gases

Given below as bullet points is an edited version of “Global Warming Fast Facts,” *National Geographic News*, updated June 14, 2007:

* Average temperatures have climbed 1.4 degrees Fahrenheit (0.8 degree Celsius) around the world since 1880, much of this in recent decades, according to NASA’s Goddard Institute for Space Studies.
* The rate of warming is increasing. The twentieth century’s last two decades were the hottest in 400 years and possibly the warmest for several millennia, according to a number of climate
studies. And the United Nations’ IPCC (Intergovernmental Panel on Climate Change), 2007, reports that 11 of the past 12 years are among the dozen warmest since 1850.

* Arctic ice is rapidly disappearing, and the region may have its first completely ice-free summer by 2040 or earlier. Polar bears and indigenous cultures are already suffering from the sea-ice loss.

* Glaciers and mountain snows are rapidly melting — for example, Montana’s Glacier National Park now has only 27 glaciers, versus 150 in 1910.

* Coral reefs, which are highly sensitive to small changes in water temperature, suffered the worst bleaching — or die-off in response to stress — ever recorded in 1998, with some areas seeing bleach rates of 70 percent.

* An upsurge in the amount of extreme weather events, such as hurricanes, droughts, and wildfires is also attributed in part to climate change by some experts.

* Humans are pouring carbon dioxide into the atmosphere much faster than plants and oceans can absorb it.

* Sea level could rise between 7 and 23 inches (18 to 59 centimeters) by century’s end. Some hundred million people live within 3 feet (1 meter) of mean sea level.

A valuable measure of global temperature is the land/ocean temperature anomaly over time as shown in Figure 1.22 [source: National Geographic News, updated June 14, 2007]. This measure averages temperature changes in air and water. Air temperature is measured at a height of 2 meters for 2000 or more meteorological stations throughout the world. Water surface temperatures are taken using satellite data.

Things have not improved since the 2007 IPCC report:
* “The Pine Island Glacier ice shelf in western Antarctica is in the process of ‘calving’ a massive iceberg that, when it breaks away, will have an area of about 350 square miles (900 square kilometers) — 15 times the size of Manhattan [www.huffingtonpost.com/2012/02/23, as accessed March 5, 2013].”

* Hurricane Sandy, the second most costliest storm in U.S. history, struck New York and New Jersey in October, 2013 with devastating impact. Now it seems that 100-year storm events are occurring with periodicity less than 10 years.

* Wildfires in the American west in 2012 were quantitatively worse than prior events: looking at the number of wildfires in California for the decades beginning 1920, as shown in Figure 1.23, one sees that it is on an exponential rise. Even worse, for the years 2010, 2011, and 2012 alone the number of wildfires in California was 69.

* January 2013 temperatures in Australia broke all records, requiring the color maps of temperature to add the color bright purple to the high end, and wildfires about this time consumed 1.2 million acres.

Figure 1.23: Wildfires in California
In the summer of 2015, the time of this writing, the biggest wildfires in the American West were in Oregon and Washington State. In these states there are two factors at work. The first is global warming and the drying out of vegetation, so it burns with extreme intensity. The second is the building of houses in areas not previously settled, and the lack of knowledge of these new residents about the need to clear dead brush from around their houses.


There are always a few who have bizarre beliefs, and unfortunate that a few of these are in positions of influence. But the fact is that at least 97% of scientists firmly believe in anthropomorphic-based global warming and the potential for catastrophic physical change of our planet.

Of course, Senator Inhofe has his own self-interest in opposing global warming initiatives. He represents the State of Oklahoma, a major producer of gas, oil, and coal. “Inhofe has led the battle in the Senate to block cap-and-trade legislation after it had passed the House. Cap-and-trade is supposed to be a market-based plan to reduce pollution, and in this case CO2, with the goal of halting or slowing global warming. He has had quiet support in the Senate, but he was the one willing to be hammered by the media and his Senate colleagues for not toeing the line and agreeing that this is settled science and necessary for the preservation of earth as we know it [Roger Aronoff, op cit].”

![Figure 1.24: Political Sign](http://www.electronicbo.com)

As reported by the satirical “The Onion” on Sept 6, 2000, demonstrators on the steps of the Kansas Capitol carrying signs such as “I don’t accept fundamental tenets of science and I vote,” and the eliminate entropy sign shown here, pressured lawmakers to repeal the Second Law of Thermodynamics.
“The second law of thermodynamics, a fundamental scientific principle stating that entropy increases over time as organized forms decay into greater states of randomness, has come under fire from conservative Christian groups, who are demanding that the law be repealed. ‘What do these scientists want us teaching our children? That the universe will continue to expand until it reaches eventual heat death?’ asked Christian Coalition president Ralph Reed, speaking at a rally protesting a recent Kansas Board of Education decision upholding the law. ‘That’s hardly an optimistic view of a world the Lord created for mankind. The American people are sending a strong message here: We don’t like the implications of this law, and we will not rest until it has been reversed in the courts.’

“The controversial law of nature, which asserts that matter continually breaks down as disorder increases and heat is lost, has long been decried by Christian fundamentalists as running counter to their religion’s doctrine of Divine grace and eternal salvation. ‘I wouldn’t want my child growing up in a world headed for total heat death and dissolution into a vacuum,’ said Kansas state senator Will Blanchard (R-Hutchinson). ‘No decent parent would want that.’ Calling the second law of thermodynamics ‘a deeply disturbing scientific principle that threatens our children’s understanding of God’s universe as a benevolent and loving place,’ Blanchard is spearheading a nationwide grassroots campaign to have the law removed from high-school physics textbooks. The plan has already met with significant support in the state legislatures of Kansas, Oklahoma, Missouri, Tennessee, Georgia, and Mississippi [The Onion, September 6, 2000].”

I am sorry Onion, the second law of thermodynamics is NOT controversial. It is fact, just as Ohm’s law is fact in describing the relationship between voltage across and current through a resistor. These and many other laws of physics should never be subject to legislative action. Whenever politically and/or religiously motivated individuals attempt to impose their beliefs on others it ultimately ends in farce and embarrassment. Case in point, the intelligent design trial in Dover, Pennsylvania in 2004. “In one of the biggest courtroom clashes between faith and evolution since the 1925 Scopes Monkey Trial, a federal judge barred a Pennsylvania public school district Tuesday from teaching ‘intelligent design’ in biology class, saying the concept is creationism in disguise. U.S. District Judge John E. Jones delivered a stinging attack on the Dover Area School Board, saying its first-in-the-nation decision in October 2004 to insert intelligent design into the science curriculum violates the constitutional separation of church and state. The ruling was a major setback to the intelligent design movement, which is also waging battles in Georgia and Kansas. Intelligent design holds that living organisms are so complex that they must have been created by some kind of higher force. Jones decried the ‘breathtaking inanity’ of the Dover policy and accused several board members of lying to conceal their true motive, which he said was to promote religion. A six-week trial over the issue yielded “overwhelming evidence” establishing that intelligent design “is a religious view, a mere re-labeling of creationism, and not a scientific theory,” said Jones, a Republican and a churchgoer appointed to the federal bench three years ago [Associated Press, as reported by NBC News, December 20, 2005].”
Chapter 2

The Internal Environment of a Residence

The modern residence is very different from those of hundreds of years ago. It has evolved under the influence of societal changes, particularly due to the industrial revolution. Much of the world’s population exists in an urban setting, far removed from the agrarian norm that existed from the beginnings of civilization until the age of the steam and internal combustion engines, universal electrification, telephones, paved roads, aeroplanes, radio and television, not to mention computers and cell phones.

The amount of energy we consume, compared to the medieval farmer, is enormous. One person commuting to work in an SUV can consume as much energy in a day as the farmer did in a year; of course, the farmer had oxen to pull the plough rather than a tractor. The ox lives off the land while the tractor needs gasoline or diesel fuel.

The standard unit of agrarian area is the hectare, which is 100 m by 100 m, or 10,000 m\(^2\). The English unit of agrarian area is the acre, 43,560 ft\(^2\). How did this unit become established? It was defined as the amount of land that an ox could plough in a day. Aren’t you pleased with such an exact system of measurement? I am sure the archaic English units of length, such as the rod, the chain, and the furlong, followed the same exactitude.

Few cars were air conditioned in the 1950s; now it is rare to see a car without it. Even our houses are energy hogs, contributing mightily to global warming. Virtually everyone in the United States lives in a residence with air conditioning. But times are changing, and now there is a rising awareness of the human impact on the environment. Solar energy is becoming an ever-increasing tool to reduce the carbon footprint of our houses. This chapter defines the need. Future chapters indicate economical solar solutions to our energy needs.

2.1 Electrical Use and Its Contribution to Sensible Heat

The breakdown of electrical use in a U.S. residence is as follows:
heating/cooling 45%
hot water 11%
washers/dryers 10%
lighting 11%
refrigeration 6%
electronics 2%
miscellaneous 15%
electric toothbrushes 0%

The last item is a dig at some environmentalists who, during the 1970s gas crises, deplored the use of electric toothbrushes: These use very little energy, and recall at the time almost all domestic lighting used incandescent bulbs. Please, focus on the important and dismiss the trivial.

The percentages of domestic energy use are changing as appliances become more efficient: see Section X. Also, the awareness of the impact on energy of the behavior of the occupants is increasing. It is recognized that energy efficiencies in residences is the low-hanging fruit. Alternative energy resources, such as passive and active solar thermal, photovoltaics, wind turbines, and geothermal are becoming a small but rapidly increasing part of the equation.

Electrical energy used in a residence reduces winter heating requirements. Most appliances are used intermittently, maybe once a day or less. Electric lighting is the exception, and may be used during all the waking hours of the occupants. We will see in Section 10.2 that incandescent lighting is extremely wasteful. For reading purposes a 70-watt incandescent bulb may be adequate provided the light is less than a meter from the page. If used for 8 hours a day it burns 0.56 kWh/day, or 204.4 kWh/year. At 16c/kWh this costs the homeowner $32.70/year, just for providing lighting for a single person in a fixed position. To illuminate a larger area, such as a kitchen, with incandescent bulbs could easily be consuming 600 watts, and with 8 hours of use per day this costs $280/year.

Typical yearly energy use in kWh for various electric appliances is as follows [trimutilities@aol.com]:

- television 200
- microwave oven 200
- dishwasher 600
- washing machine 900
- clothes dryer 850
- refrigerator 1150
- freezer 800
- electric stove 700
- computer 550

Assuming a uniform use of these appliances throughout the year, the average daily electric use by these is almost 16.3 kWh/day, which contributes directly to heating the residence.

The most common way of operating a clothes dryer is to vent the hot, humid air to the outside. This is wasteful during the heating season, not only for the heat lost but by losing the needed water vapor. A simple vent bypass can be installed to direct the air flow either in or out. The
outlet to the interior of the house should be screened to prevent lint and dust from being recirculated to the interior; women’s hosiery makes a satisfactory and inexpensive filter.

Other appliances not listed above include:

* coffee maker, using about 1 kWh to make a single pot of coffee,
* clothes iron 1–2 kWh,
* hair dryer, about the same,
* toaster, 0.8–1.4 kWh,
* vacuum cleaner, 1–1.5 kWh.

Overall, we can assume that appliances contribute about 20 kWh of sensible heat every winter day.

A typical American house with about 180 m² of floor space requires about 12 kW of heating when the interior temperature is 21°C and the external temperature is about -1°C, the average winter temperature in New England. Over a day this amounts to 288 kWh. Thus, the sensible heat provided by electrical appliances, at 16.3 kWh, provides about 5.7% of the winter heat load, or raises the interior temperature by about 1.3°C.

Most lighting use will occur during occupancy. Lighting used for security purposes can be motion activated — something that typically scares off an intruder while illuminating one’s path at night — a good idea in some neighborhoods.

Most of the energy consumed by heating water and operating washers is wasted. Modern dishwashers use about 15 liters of water. Hand-washing dishes uses about ten times this amount, meaning that from an energy reduction standpoint, machine dishwashers make the best sense. However, there are some caveats. Dishwashers use the same amount of energy for an empty load as for a full load, so always run it full, even if it means holding dirty dishes for a day or more; dishwashers can clean even baked-on food scraps. Also, scrape larger food particles from the dishes before loading them into the dishwasher; do not hand rinse, which typically uses enormous quantities of hot water. Air dry the dishes, and avoid using machine drying: in winter, the air drying provides needed water vapor to the interior air.

The U.S. DOE 2003 rule requires dishwashers to use less than 2.17 kWh per load. After January 2012, dishwashers must use 9% less electricity and 27% less water. However, compared to European standards, the United States is sorely lagging.

Notably, the performance of the dishwashers manufactured by the German company Bosch far exceeds those of their American counterparts.

Top-loaded clothes washers require about 150 L of water per load. The tub within which the clothes are cleaned is vertically oriented, so the level of water in the tub is high. The tub in a front loaded washing machine is horizontally oriented, so as the tub, rotates the clothes tumble to the bottom side of the tub which is covered by water, and this requires less water, from 45 to
95 L per load. After the water has done its task, it is discharged into the sewer or equivalent, and the sensible heat held by it is lost.

Clothes dryers use substantial amounts of energy, either electricity or gas. A clothes dryer has a horizontal-axis rotating cylinder with internal ridges so that as the cylinder rotates, the clothes are tumbled upwards and then drop to the bottom. A small electric motor drives the cylinder. Heated air passes through perforations in the cylinder to dry the clothes. The air is heated by electricity or gas. The American standard for efficiency is at least 1.4 kg of clothes for every kWh of electricity, or 1.2 pounds for every equivalent kWh of gas. The energy used depends on the size of the load, and all modern dryers have sensors that terminate the drying process when the humidity inside the cylinder is below a level that indicates that the clothes are dry. Unlike clothes or dish washers, small loads can be run through a clothes dryer efficiently. There is no ENERGY STAR rating for clothes dryers.

2.2 Human Needs and Demands in a Residence

Residences are designed for the human needs of shelter, security, comfort, and convenience. The temperature and humidity ranges must be controlled. Ventilation and fenestration are required. Natural lighting requires artificial augmentation, particularly in the evening hours.

For a human body to feel comfortable a number of physical parameters must lie within a defined range. These include:

* temperature, which includes radiated energy from surrounding surfaces,
* ventilation,
* humidity,
* air motion (drafts or pleasing wafts),
* control of dust,
* odor (unwanted bathroom or aromatic flowers),
* noise (grating machinery or delightful Mozart),
* lighting, artificial and natural,
* aesthetics.

All of these items are to some extent subjective, some more than others. In this chapter we determine limits on most of these, with objective limits set whenever possible.

2.3 Temperature Levels

What is required to produce the sensation of comfort with respect to the temperature? According to ISO 7730, “That condition of mind which expresses satisfaction with the thermal
environment.” The feeling of comfort is not an unchanging function of temperature. After moderate exercise the body heat generated may make the room seem too hot. After relaxing for some time the body at rest produces minimal heat, and the room may feel chilly. However, there are acceptable limits that can be set, and we will attempt to do so here.

The core body temperature is 37° C. Skin temperature is a little lower, but at 34° C sensors in the skin send signals to the brain that the body is getting cold.

A fully clothed body in a 5°C environment can be more comfortable than a naked body in a stiff 20°C breeze. The heat loss from the body must be contained within certain limits to maintain comfort. Clothing helps in this regard. At 22°C to 25°C a body can be lightly clad and be comfortable. A person at rest can be quite comfortable if warmly clad when the temperature is 18°C.

Several years ago it appeared impossible to breed polar bears in captivity. Zoo keepers attempted to simulate the ice caves of the polar bears with concrete caves, with no success. Then someone studied the natural internal environment of a polar bear lair and found it warm, hot by human standards from reflected body heat. A duplication of this environment in zoos has resulted in a successful breeding program.

Humans dissipate heat as a function of age, body size, and sex. If it is assumed that average sizes of 70 kg for men and 58 kg for women (these figures were representative in 1950 but are no longer applicable due to our increased heights and substantially increased weights), then the heat loss in watts of a human at rest as a function of age and sex is shown in Figure ?? [C.C.W. Voss, *A Synopsis of Physiology*, John Wright, London, 1954].

If the person is other than at rest, sitting in an easy chair, the following modifications must be made:

* when asleep, deduct 20–30 W,
* engaged in light activities, add 10–20 W,
* engaged in moderate activity (moving furniture), add 30–40 W.

Assuming a family of four with two 35 year old adults and two children ages 5 and 10, typical human energy produced to heat the house is:

**Man:** 14 h/day occupancy, 8 sleeping, 4 sitting, 2 in light activity, produces $8 \times 110 + 4 \times 135 + 2 \times 150 = 1820$ Wh/day.

**Woman:** 19 h/day occupancy, 8 sleeping, 4 sitting, 6 light activity, 1 strenuous activity, producing $8 \times 80 + 4 \times 100 + 6 \times 115 + 1 \times 135 = 1865$ Wh/day.

**10 year old:** 17 h/day occupancy, 10 sleeping, 1 sitting, 4 light activity, 2 strenuous activity, producing $10 \times 95 + 1 \times 115 + 4 \times 125 + 2 \times 145 = 1855$ Wh/day.
5 year old: 18 h/day occupancy, 12 sleeping, 1 sitting, 3 light activity, 2 strenuous activity, producing $12 \times 65 + 1 \times 85 + 3 \times 95 + 2 \times 115 = 1380$ Wh/day.

Thus, the complete family contributes 6.92 kWh/day due to body heat loss to produce sensible heat for the house, as well as latent heat in the form of water vapor from respiration and perspiration. Add this 6.92 kWh to the 16.3 kWh and the 23.2 kWh is 8.3% of the heating needs of a 12 kW house when the temperature differential $\Delta T$ is $22^\circ$C. The heating percentage caused by human occupancy as a function $\Delta T$ is shown in Figure 2.2.

2.4 Ventilation Guidelines

We need air to survive, and we want more from air than just survival. We like breathing fresh air, but what does that mean? We know that air containing volatile organic compounds may be damaging to human health; see Section VII of this chapter.
Breathing in draws air into the lungs where oxygen is extracted. Breathing out exhales carbon dioxide (CO₂). CO₂ levels of outdoor air rarely, if ever, reach discomfort levels. Now, in ultr-tight housing, CO₂ and other airborne contaminants can reach dangerous levels without mechanical ventilation; see Section V of this chapter.

European and American air standards concentrate on air changes per hour: see the next section. European standards also consider air velocity; see Section 4.2. In addition:

“ Adequate means of ventilation shall be provided for people in buildings. This shall be achieved by limiting the moisture content of the air within the building so that it does not contribute to condensation and mold growth, and limiting the concentration of harmful pollutants in the air within the building.

“ Ventilation is the supply of fresh outside air and the removal of stale indoor air to or from spaces in a building. It normally comprises a combination of purpose-provided ventilation and air infiltration. The purpose-provided ventilation may be provided by natural or mechanical means.

“ Air Infiltration is the uncontrollable air exchange between the inside and outside of a building through a wide range of air leakage paths in the building structure. Purpose-provided ventilation is the controllable air exchange between the inside and outside of a building by means of a range of natural and/or mechanical devices.

“ Air permeability: The average volume of air in cubic metres per hour that passes through one square metre of the building envelope when subject to an internal to external pressure difference of 50 Pascals when measured in accordance with the method defined in IS EN 13829:2000


2.4.1 Air Changes per Hour

ASHRAE 62 guidelines specify a minimum of 0.35 air changes/hour, but no less than 25.5 m³/h, 15 cfm, per person. A maximum of 0.6 air changes/hour is recommended to limit energy costs. These guidelines will permit adequate air for respiration purposes, but may not prevent unwanted odors. Some such odors should be controlled at the point of source if at all possible. Odor from a bowel movement should be vented to the outside at time of creation. Cooking odors should be similarly vented. Such vents have low cfm ratings, so they provide a small
load on the domestic energy usage. Smelly socks and sweaty sports clothing should be kept in a sealed container until laundered.

These are the numbers for air: the density of air at 0°C is 1.293 kg/m³, the specific heat capacity of air in the temperature range 0°C → 40°C is 0.24 kcal/kg·°C, and 1 Wh ≡ 0.8598 kcal, so if the air used is $V$ m³/h and the temperature differential is $\Delta T$ °C, then the power needed to heat that air is

$$V \times \Delta T \times 1.293 \times 0.8598 = V \times \Delta T \times 2731 \text{ kW}. \quad (2.1)$$

For a family of four people, assuming an average volume of fresh air needed for respiration of 90 m³/hr, the heat content to raise the temperature of that air by 31°C, from the freezing point of water to body temperature, is $90 \times 312731 = 1.022$ kW. If this respiration air was heated by electricity at 16c/kWh, the monthly cost would be $0.16 \times 1.022 \times 24 \times 30 = \$117.69$.

European standards typically specify a minimum in m³/h depending on the number of rooms. For example, Brussels, the capital city of the European Union, requires 75 m³/h for a living room, 25 m³/h for a bedroom, 50 m³/h for a bathroom with WC, but limited to 75 m³/h: for a whole building, the rate is 3.6 m³/h per m² of floor area. See “Indoor Air Quality, Thermal Comfort and Daylight: Analysis of Residential Building Regulations in Eight EU Member States [www.scpclearinghouse.org/upload/publication_and_tool/file/414.pdf].”

Consider a medium sized house with footprint $8 \times 14$ m on two levels with ceiling heights 2.7 m. The volume of that house is 604.8 m³, and with the ASHRAE recommended minimum of 0.35 air changes per hour, then the volume of fresh air needed is 211.68 m³/hr. To raise that air temperature by 21°C, freezing point of water to room temperature, requires $211.7 \times 212731 = 1.628$ kW.

### 2.4.2 Air Velocity

A number of European cities/countries specify minimum air velocity rates in a dwelling, and usually specify a higher rate in summer than in winter. For example, Brussels specifies a rate of 0.21 m/s, meters per second, in winter, and 0.24 m/s in summer. Poland specifies 0.2 m/s in winter and 0.4 m/s in summer. England and Wales specify 0.15 m/s, both summer and winter, but Germany and France have no requirements on air velocity.

To specify air velocity regulations implies that mechanical ventilation is required. Additional regulations determine the location of the registers, typically high in a wall so that air is not blowing directly at the occupants.

Further, Germany has a non-binding standard called DIN 1946-6 that suggests the use of mechanical ventilation for buildings in which the air volume needed for moisture control is
greater than the air infiltration rate. As in the United States, air infiltration is determined by a blower test; see Section 7.1 of Chapter 4.

2.4.3 Relative Humidity

The relative humidity of a New England house should not be maintained at a level much higher than 30% in winter; otherwise, condensation can occur on the inside surface of the windows. The limit of relative humidity in the house is a function of outdoor temperature, with room temperature about 21°C, as can be seen in Figure 2.3 for single and double glazing [W.J. McGuiness, B. Stein, C.M. Gay and C. de van Fawcett, Mechanical and Electrical Equipment for Buildings, John Wiley, New York, 1964].

The relationship between air temperature and its water content in g/m³ at 100% humidity is shown in Figure 2.4. As can be seen, at low temperatures the capacity of air to hold water vapor is low, but a substantial amount of water can be contained at high temperatures. Air transfer through the skin of a house takes low temperature (and so low water content) air and heats it to room temperature. 0°C air can contain a maximum of 4.83 grams/m³ of water vapor, and if it does the relative humidity is 100%. At 21°C air can contain up to 18.9 grams/m³ of water.
For our typical house with volume 604.8 m³, and 30% relative humidity at 21°C, the amount of water in the air is $604.8 \times 18.9 \times 0.3/1000 = 3.43$ kg. Assuming 0.35 air changes per hour, then $3.43 \times 0.35 = 1.20$ kg/h of water needs to be added the air.

The data shown in Table 2.1 was extracted from Table 2.4. Using MATLAB to fit a polynomial through this data produces a very close fit, and the polynomial of order three is

$$p=3.59\times10^{-4}t^3+8.80\times10^{-3}t^2+0.229t+5.16,$$

(2.2)

where $t$ is the temperature in °C. Choosing air at temperature 30°C and relative humidity 30%, the dew point temperature when the relative humidity is 100% is 8.994°C.

### 2.4.4 Air Infiltration

Air infiltration in winter reduces the relative humidity in the structure. This is countered by sources of humidity within the structure.
Here we use the term infiltration to mean leakage of the skin of the house through the siding, around doors and windows, and through the roof or basement. The air flow caused by ingress and egress though exterior doors is discussed in the following section.

Older structures are commonly so leaky that they far exceed the minimum air changes needed. Tight modern houses may have too little infiltration and air changes must be augmented by opening windows (a bad idea in cold climates) or by heat recovery ventilation as discussed in Section V.

Lack of insulation or air leakage can be detected with an infrared meter. These are widely available and are inexpensive: The non-contact type can be pointed towards a small area to determine its temperature on an exterior wall or window. Although the temperature detected will be lower than the interior ambient temperature in winter, a large differential indicates a heat loss problem.

Care must be taken when using extractor fans or wood stoves that rely on internal combustion air. These create a negative pressure with respect to atmospheric, and this can have detrimental effects. If the negative pressure is 3 pascals or more, back-drafting can occur — flue or chimney gases may be sucked into the structure with potential danger from carbon monoxide. Carbon monoxide is odorless and colorless and absorbs more readily in the blood stream than oxygen. Asphyxiation can occur while the victim is unaware of a problem.

According to the National Fuel Gas Code of the United States, a vented gas appliance using interior combustion air requires a minimum of 50 ft³ of open space for every 1,000 BTU/hr of rated input, or 20 BTU/ft³·h; this converts to $20 \times 0.2931/35.3144 = 0.166$ W/m³. If the combustion air is ducted in from the outside, the 1,000 BTU/hr requirement can be increased to between 2,000 and 4,000 BTU/hr; this converts to $0.332 \rightarrow 0.664$ W/m³.

Most modern furnaces, boilers, and water heaters use external combustion air: There is no good reason to have one of these that relies on internal air for its combustion. The two compelling reasons for using external combustion air are energy efficiency and air quality needed for human respiration.

Ventilation requirements for non-occupied areas or lightly occupied areas may be minimal. However, basements often need humidity control, and in some regions naturally occurring radon gas can leak into the basement with detrimental consequences to human health. Attic ventilation may be employed to cool the structure underneath in summer, or to reduce dew point precipitation in winter.

2.4.5  Air Flow through Exterior Doors
It is difficult to determine the amount of air flow that occurs when an exterior door is opened and then closed to permit someone to pass through. It depends on a number of factors, not least the time it takes for the transition. An elderly person, possibly with disabilities, will probably take much more time than a young adult. Children are liable to be careless in closing doors promptly. It is for these reasons that commercial installations often employ revolving doors that substantially reduce air flow.

The revolving door was invented by Theophilus Van Kannel; see patent #641563 dated January 16, 1900. He termed it a revolving storm door. The revolving door is used on many commercial or institutional buildings for energy efficiency. A study conducted by MIT students in 2006 concluded that the standard swinging door permits eight times more air to pass through per operation. Further, a person can be entering at the same time as a person is exiting with the revolving door with no more energy loss than when a single person is operating it. The problem is the reluctance of many people to use a rotating door, and the difficulty (impossibility) presented by such a door to a disabled person.

In the commercial/institutional buildings that install rotating doors, the air pressure is increased slightly by 2–5 pascals. This means that the energy losses/gains in such buildings could have different characteristics than that of a residence. However, it appears certain that the swing of door will have many times the air flowing the rotating door. The factor of 8:1 will be assumed here.

The configuration of a revolving door is shown in Figure 2.5. Its diameter is $D$, where typically this varies from 2m to 2.7m. With a 2.1m diameter door, opening $A = 1.4$ and wall length $B = 1.6$ m. If the height of the door is 2.1m then the volume of air in a segment is $1.052\pi 4 \times 2.1 = 1.82$ m$^3$.

Raising 1.82 m$^3$ of air at 0°C to body temperature 31°C requires $1.82 \times 1.293 \times 0.24 \times 31 \times 0.86/1000 = 0.015$ kW; the numbers have been rounded to avoid the appearance of higher accuracy than is present. Multiplying this number by eight by replacing the rotating door with a swing door produces 0.12 kW. The heat loss does not seem excessive, but this loss is for one transition from interior to exterior or vice versa. With ten transitions a day, and this number is small when children are in the house, the result is 1.2 kW of heat loss.
A wintertime energy need occurs when the air in the house is too dry, requiring the use of a humidifier. However, in a tight house, particularly one requiring a heat recovery ventilator, the internal sources of water vapor such as showers and cooking could raise the humidity to a comfortable level.

Direct entries with no interlock are not a good idea. An interlock system has two doors, with only one open at any time under normal operation. Two configurations of interlock entryways are shown in Figure 2.5. The closet doors should be fairly tight so that there is minimal air flow around them when one of the two doors is opened. Assuming that these closet doors are tight and the area to the left of the closet wall is the exterior, then the closet itself can act as an additional insulation for the house, with the minor penalty that clothes in that closet will be cold in winter.

The entry on the right will reduce air flow because the two doors are not in line, but suffers in that it will not permit large items, such as sofas, to pass through. This is not a problem provided another entry on the same level provides a clear path in.

For the entryways shown, and assuming tight closet doors and a minimal ceiling height of 2.1 m, the volume of air in the entryway is \(1 \times 1.2 \times 2.1 = 2.52\) m\(^3\). Compare this to \(1.82\) m\(^3\) for the rotating door and its 8:1 advantage over a single swinging door and one can see that the double door system for a residence is most desirable.
2.4.6 Respiration and Carbon Dioxide

Does human respiration contribute to \( CO_2 \) in the atmosphere? The online consensus opinion is that humans are part of a natural cycle in which plants extract \( CO_2 \) from the atmosphere and give out \( O_2 \), humans eat the plants for their nutrition, breathe in \( O_2 \), and exhale \( CO_2 \). Without human interference, the plants would die and natural decay would release their \( CO_2 \) to the atmosphere. Thus, humans just add another step in a natural process. Unfortunately, this is simplistic thinking.

When humans were hunter/gatherers, the just-presented scenario was basically correct. In that epoch, the number of humans on planet Earth were minuscule compared to the present. There are now about 7.3 billion people in the world, and few if any of us are hunter/gatherers. Many of us, a large majority in advanced nations, live in cities and buy our food, and that food is increasingly grown in a mono-culture environment with advanced machinery, pesticides, and fertilizers made from fossil fuels. That food needs to be shipped to the end users using fossil-fuel-based transportation; many are concerned about the number of food miles to get our food to market. Thus, just the growing of our food is no longer a natural process.

A number that approximates the average amount of \( CO_2 \) exhaled by a human is 0.9 Kg/day, so worldwide the total \( CO_2 \) entering the atmosphere from human respiration is \( 0.9 \times 365 \times 7.3 = 2.4 \times 10^{12} \) Kg/year. The amount of \( CO_2 \) produced by combustion of fossil fuels, some of it in food production and transportation, is estimated at about \( 35 \times 10^{12} \) Kg, so our respiration...
directly adds almost 7% of CO₂ to the atmosphere. Yes, a significant fraction of the 7% is part of a growth/harvest/eat cycle, but it is nowhere near all of it.

\[ \text{directly adds almost 7\% of CO}_2 \text{ to the atmosphere. Yes, a significant fraction of the 7\% is part of a growth/harvest/eat cycle, but it is nowhere near all of it.} \]

### 2.5 Heat Recovery Ventilation

Residences in Germany typically use HRVs (Heat Recovery Ventilators) to swap stale interior air with fresh exterior air. At the same time they can save most of the energy in the warm interior air by heating the incoming cold winter air. Why do so few residences in the United States employ the same strategy? There are two answers to this question. First, residences in the U.S. are usually much larger than their German counterparts, so the required number of air changes per hour is reduced — see Section 4.1 in this chapter. Second, residences are constructed to looser standards than those in Germany, so air infiltration is much higher.

Building code standards in the United States and other countries are becoming more stringent so as to reduce energy use, a trend that will continue. We see this in the Massachusetts Stretch Code.

![Figure 2.7: Parallel and Counter Flowing Air](image)

In an HRV, the cold incoming air does not mix with the warmer outgoing air, so any pollutant in the outgoing air is expelled from the structure. The flow of air in the exchanger can be parallel, in the same direction, or can be counter flowing. Both are illustrated in Figure 2.7. On the left the flow is in the same direction, and the incoming air can, at best, pick up half the heat of the outgoing air. However, with counter flowing air it is theoretically possible for the incoming air to reach the interior temperature inside the HRV — the efficiency of the heat transfer would then be 100%. In practice, efficiencies reaching 90% are possible.
There are numerous strategies for achieving heat transfer in the HRV. One is the shell-and-tube type as shown in Figure 2.8. Here the incoming air is carried in a central tube, which is surrounded by an airtight shell. The outgoing air passes through the shell in such a way that as it moves from one end of the shell to the other, it rotates around the central tube.

Another air-to-air heat exchanger is the plate type as shown in Figure 2.9. Stale air is in the air gaps between the first and third pair of plates, etc., while fresh air is in the air gaps between the second and fourth pair, etc. The stale and fresh air do not intermingle. Stale, warm interior air enters one side of the plates and gives up most of its heat to the plates, emerging as cool stale air to the outside. Fresh, cool air enters another side of the plates, picking up heat from the adjacent plates and entering the structure as warm fresh air.
Let us evaluate the energy savings. In Section 4.1 of this chapter we discussed a medium sized house needing 211.68 m$^3$/hr of fresh air. To raise that air temperature by 21$^0$C, from freezing point of water to room temperature, required about 580 W. With an HRV at 90% efficiency, a high but realistic number, the net energy needed for this fresh air is $0.1 \times 580 = 58$ W.

However, this does not tell the whole story. As posted online by Alex Wilson on February 13, 2014 under the title “Our Top–Efficiency Heat–Recovery Ventilator [www.greenbuildingadvisor.com/blogs/dept/energy−solutions],”: “Another measure of efficiency is how much air is moved per unit of electricity consumed. Here we can look at the cfm of air flow per watt of electricity consumption. With this metric, the Zehnder ComfoAir really shines, achieving a remarkable 2.58 to 3.25 cfm/W (depending on the fan speed). The Energy Star criteria for HRVs to be listed as EnergyStar is 1.0 cfm/W, and most good HRVs have air-delivery efficiencies only in the 1.0 to 1.5 cfm/W range. I was able to find only a few others with cfm/W values exceeding 2.0.” Thus, if our house needed 211.68 m$^3$/h, the electrical energy to power the EnergyStar HRV is $211.68 \times 35.3/60 = 124.53$ W; this plus the wattage needed to heat the air adds up to 182.54 W.

One can classify the electrical costs of using HRVs as a parasitic cost. Other parasitic costs occur with active solar thermal systems such as flat plate collectors that use electricity for the pumps involved. These parasitic costs in electrical use produce sensible heat which, in a heating environment, provide a benefit. However, it would be better to minimize parasitic costs, just as it is desirable to reduce air infiltration.

In a cold winter climate, the relative humidity inside a residence, as discussed earlier in this chapter, may be too low, since the incoming air is dry compared to the internal relative humidity. The HRV does not help in this regard. As posted on January 22, 2010 by Martin Holladay, greenbuildingadvisor.com, “An ERV does everything that an HRV does. In addition, an ERV allows some of the moisture in the more humid air stream (usually the stale air in winter and the fresh air in summer) to be transferred to the air stream which is dryer. This transfer of moisture — called enthalpy transfer — occurs with very little mixing of the two air streams. The cross contamination rate for one well-regarded ERV, the UltimateAir RecoupAerator, is 9.6%.”

ERVs may be of no help in the hot and humid summers in the American South. They may help lower the interior temperature, but may not be effective in lowering the humidity. In such a case, if cooling is not required, a dehumidifier may be needed.

### 2.6 Humidity Sources in a Residence

Cooking adds water to the air — 0.15kg for breakfast, 0.07kg for lunch, and 0.5kg for dinner are typical figures. Dishwashing adds 0.09, 0.07, and 0.3 lb after these three meals. Showering
adds about 0.2kg/shower. Clothes washing can add a lot of water — over 2kg per load. House plants can add about 0.5kg/plant per day. The largest source of humidity is an unvented clothes dryer — 12kg/load — so this condition is unacceptable unless the humidity is too low. Thus, a typical family will add about 12kg/day, 0.51kg/h, of water per day to the air in the structure [R.M.E. Diamant, The Internal Environment of Dwellings, Hutchinson Educational Ltd., London, 1971].

Add to this the water vapor added to the air by human respiration. The typical family of four breathe in 90m$^3$/h of air at 30% relative humidity, while breathing out air at close to 100% relative humidity. Air at 31°C at 100% relative humidity contains 32 g/m$^3$ of water, while air at 21°C at 30% relative humidity contains 18.9×0.3 = 5.67 grams/m$^3$ of water. Thus, the family adds 90 × (32 − 5.67)/1000 = 2.37 kg/h water to the air. Thus, total interior water vapor sources add 0.51 + 2.37 = 2.88 kg/h of water to the air. Compare this to the capacity of our typical house to contain water in the air at 30% with relative humidity at 3.43 kg.

Since only 1.2 kg/h of water vapor is required to maintain 30% humidity, while the occupants through respiration and human activities produce 2.88 kg/h, this leads to one of the following scenarios:

* The humidity rises, resulting in condensation on the double glazed windows at low outside temperatures, or
* Infiltration/ventilation must be substantially increased with associated heat loss, or
* A dehumidifier is needed to reduce the humidity.

### 2.7 Sick Building Syndrome and VOC Mitigation

Volatile Organic Compounds (VOC), not formaldehyde, are emitted by humans and many building materials. Humans exude $H_2O$, $(CO_2)$, aldehydes, esters, and alcohols. TVOC (total VOC) levels as measured in μg/m$^3$, which do not distinguish by type of pollutant, are

<200 comfort range200–3,000 multifactorial rangeTVOC -> 3,000–25,000 discomfort range>25,000 toxic range

A more sophisticated approach divides the pollutants into classes with maximum designated levels that should not be exceeded, and the numbers added to get the TVOC, which must not exceed 300 μg/m$^3$. The pollution limits by type are:

alkanes 100
aromatics 50
terpenes 30
halocarbons 30
esters 20
carbonyls (excluding formaldehyde) 20
“others” 50
$CO_2$ levels should be below 1,000 ppm. $CO$ (carbon monoxide) is produced by incomplete combustion, and is extremely dangerous, particularly since we cannot see it or smell it; levels as little as 5 ppm can have deleterious consequences, and if is at 100 ppm the building should be immediately evacuated.

1 olf is the rate of odor generation from one person with a normal standard of sanitation and a moderate to low level of physical activity. Decipol is the perceived level of air pollution in a space with a pollution source of 1 olf ventilated by 10 L/s of fresh air. Thus, if the area is vented at 1 L/s, then with 1 olf pollution the result is 10 decipols [R. Edwards, *Handbook of Domestic Ventilation*, Elsevier, Oxford, 2005].

$$Q = M \varepsilon v (C_i - C_o)$$

where $Q$ L/s is the ventilation rate, $M$ $\mu$g/s is the pollution generation rate, and $C_i$ and $C_o$ $\mu$g/s are the permitted maximum levels of indoor and outdoor air pollution respectively.

*Figure 2.10* shows the percentage dissatisfaction, a subjective measure, as a function of the ventilation rate as measured in L/s times the number of occupants. *Figure 2.11* shows the percentage dissatisfaction as a function of the decipols [Report #33, “Guidelines for Ventilation Requirements in Buildings,” European Collaborative Action: Indoor Air Quality & its Impact on Man, 1992].

The feeling of malaise that some experience in certain buildings is termed Sick Building Syndrome (SBS). Its causes are difficult to pin down. Some believe that the predominant cause is psychological rather than physical. Since the prevalence appears on the increase, and coincides with the increased energy efficiency of buildings, mainly due to reduced air infiltration, the blame is most often placed on inadequate ventilation. Symptoms of SBS
include eye, nose, and throat irritation, skin rash, headaches, and nausea. Mold, caused by high humidity, can cause a number of ailments.

The attributes of a building producing SBS include:

* forced ventilation/air conditioning,
* carpets and drapes,
* lack of operable windows.

## 2.8 Illumination and Daylighting

Illuminance or light intensity is quantified as lux (SI units) or foot-candles (Imperial units). A lux is the light intensity from a source of one candlepower that is incident on a surface of unit area normal to the source at a distance of unity from the source. A foot-candle is light intensity from a source of one candlepower that is incident on a surface of one square foot normal to the source at a distance of one foot from the source. 1 foot-candle = 10.752 lux. Light energy is measured in lumens, where 1 lux = 1 lumen/m², and 1 foot-candle = 1 lumen/ft².

Outdoor light levels in lux [www.EngineeringToolbox.com, as accessed 1/3/15] are approximately:

<table>
<thead>
<tr>
<th>Condition</th>
<th>Illuminance (lux)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sunlight</td>
<td>100,000</td>
</tr>
<tr>
<td>daylight</td>
<td>10,000</td>
</tr>
<tr>
<td>overcast day</td>
<td>1000</td>
</tr>
<tr>
<td>twilight</td>
<td>10</td>
</tr>
<tr>
<td>full Moon</td>
<td>0.1</td>
</tr>
<tr>
<td>starlight</td>
<td>0.001</td>
</tr>
</tbody>
</table>

The level of illumination on a surface in full sunlight is 10 times that for an outside surface exposed to the light but out of the Sun, and 100 times that of a similar surface on an overcast day. The ability of the human eye to adjust to an enormous range of illuminations is only matched by a similar range for the human ear. On a clear night with a full Moon, the illumination level is a million times lower than that in full Sun: moonlight is not adequate for reading, but it is good enough for navigating oneself along a path.

For the reading of papers and books, an illumination level of 300–400 lux is recommended. This is too bright for a large open space, so a library may want to have directional table lamps that can shine on the reading matter but permit lower levels elsewhere. Worktables need about 500 lux. There is some difficulty in providing the correct level of lighting while minimizing glare on a computer monitor.

The subject of daylighting has been given considerable attention in solar circles. Free natural lighting is obviously desirable over carbon-hogging incandescent lighting, so it should be used whenever possible. Traditionally, daylighting means the use of windows and skylights, as
shown in Figure 2.13. As we will see in Chapter 11, passive solar requires south-facing vertical glass.

A potential problem with windows and skylights is the glare that can occur with direct sunlight; this can be alleviated with light-diffusive blinds. Another downside may be the unwanted heat in summer, and the heat losses in winter: Low-e glass can reduce these effects, but should not be used for passive solar glazing on the south-facing vertical glass. As always, the best solution is a compromise, and the architectural layout and interior surfaces are important factors. Dark interior walls are not an option, except as limited decor elements. South-facing passive solar elements during the predominance of the daylight hours provide heat in winter and minimal solar gain in summer: see Chapter 11.

Until 1986, one could argue that the subject of daylighting was the purview of architects and interior designers. What happened in 1986 was the issuance of a patent on a TDD (Tubular Daylighting Device) to the Australian company Solatube International. The result was a technically superior way of directing natural lighting to locations in buildings that would otherwise be in the dark. The sketch of a TDD is given in Figure 2.12. The device brings light from a pitch roof via a flexible cylinder into a room below. A transparent hemispherical dome collects the light that bounces off the interior reflective surface of the flexible cylinder. It enters the room below after passing through a diffuser to spread the light and reduce the glare. The flexibility of the reflective cylinder makes the device capable of adapting to a broad range of roof pitches.
“It is remarkable to see the impact we have made worldwide in just two decades,” says Solatube International CEO and founder David Rillie. “By conservative estimates, we have sold enough Solatube Daylighting Systems to offset 68,590,000 pounds in carbon emissions annually. That equates to taking 5,715 cars off the road for a full year. Imagine if everyone used daylighting.”

Suppose the efficiency of light transmission of a TDD is 10% and the area of light collection is 0.1m². Then, under full Sun at 880 W/m², the light energy transmitted into the structure is 8.8W. An ideal light source delivers 683 lm/W (lumens/watt) as discussed in Section 10.2 of this chapter, and we will see in Section I of Chapter 6 that about 48% of solar energy is in the visible range, so the TDD delivers $8.8 \times 683 \times 0.48 = 2885$ lm.

A relatively open interior architecture permits light from south-facing windows to penetrate deep into the house. Insulating blinds at nighttime substantially reduce heat losses. East and west windows are usually the most problematic, but specialty glass can reduce solar gain and heat loss.
The cross-sectional sketch of a bermed house built on a south-facing slope is shown in Figure 2.13. The sunlight may not penetrate deep into the structure. One solution is to use electrically generated lighting; see Section 4.2. The other solution is to allow light to penetrate the roof through a skylight. Skylights are problematic, at least in regions subject to snow and ice, which can make them subject to leakage. Another problem is the unwanted solar gain in summer and the substantial heat loss in winter. The bottom line on skylights is to use them sparingly and only when other options are less attractive.

2.9 Domestic Water Usage

Domestic water usage worldwide is incredibly uneven, with the ratio of the country with the highest use to the lowest per person being 270:1. Here are a few numbers as measured in cubic meters of domestic water use per person per year for the average of years 1987–2003 [www.worldmapper.org, as accessed on December 1, 2012]:

<table>
<thead>
<tr>
<th>Country</th>
<th>Use (m³/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>487</td>
</tr>
<tr>
<td>Canada</td>
<td>259</td>
</tr>
<tr>
<td>United States</td>
<td>209</td>
</tr>
<tr>
<td>Cambodia</td>
<td>1.8</td>
</tr>
</tbody>
</table>

Table 2.2: Profligate Water Usage

<table>
<thead>
<tr>
<th>Fixture</th>
<th>Time or Usage</th>
<th>Liters</th>
<th>Total Daily</th>
</tr>
</thead>
<tbody>
<tr>
<td>showerhead</td>
<td>8 minutes per</td>
<td>15/minute</td>
<td>15×32 = 480</td>
</tr>
<tr>
<td>toilets</td>
<td>5 times per</td>
<td>13/flush</td>
<td>13×20 = 260</td>
</tr>
<tr>
<td>faucets</td>
<td>5 minutes per</td>
<td>10/minute</td>
<td>10×20 = 200</td>
</tr>
<tr>
<td>kitchen</td>
<td>15 minutes total</td>
<td>10/minute</td>
<td>10×15 = 150</td>
</tr>
<tr>
<td>dishwasher</td>
<td>1 use</td>
<td>33/use</td>
<td>33</td>
</tr>
<tr>
<td>top-loader</td>
<td>1 use</td>
<td>170 liters per</td>
<td>170</td>
</tr>
</tbody>
</table>
The average domestic consumption for a home in Canada is approximately 350 to 400 liters per person per day (300 l/d for indoor use, and 100 l/d for outdoor use). Canadians use considerably more water than most other nations, using water that flows through the summer from mountain glaciers and snowpack, but this resource is diminishing. To sustain our access to water, it will be important to conserve water — now and into the future.

Domestic water is used in the bathroom (toilets, showers, and faucets), the kitchen (dishwashing and food preparation), as well as for laundry. With little or no concern for profligate use, the water use for a family of four people could easily be as shown in Figure 2.2. The following assumptions are made. Each person uses the shower once a day for 8 minutes per, flushes the toilet 5 times a day, and operates a bathroom faucet for 5 minutes/day. The conventional dishwasher is used once a day, as is the top-loaded clothes washer.

Thus, a typical family of four can easily use 1293 liters of water a day inside the residence according to these calculations, or 341.6 U.S. gallons/day. These numbers are at the high end of domestic water usage — compare this to the 209 gallons/day estimated by worldmapper.

There is another villain that needs to be identified as contributing to water usage — leaks often account for 10% of the total. A slow-dripping faucet can leak a several liters a day. A poor toilet sealer can leak many times that amount. Vigilance on the part of the residents is required.

In 1992 a set of federal regulations set upper limits on flow rates from shower heads and faucets, and also reduced the number of gallons/flush of the toilets. It required a shower head to have a flow rate of less than 2.5 gallons/minute at a water pressure of 80 psi, or 2.2 gallons/minute at 60 psi. Before this time, shower heads could draw water at 5.5 gallons/minute. Shower heads are of the aerating kind, or the laminar flow kind. The aerating kind mixes air with the water to produce a misty spray; in many locations this is the preferred kind, but presents problems in humid climates. The laminar flow kind sends out individual streams, and this produces less steam than the aerating kind.

<table>
<thead>
<tr>
<th>Fixture</th>
<th>Time or Usage</th>
<th>Liters</th>
<th>Total Daily</th>
</tr>
</thead>
<tbody>
<tr>
<td>showerhead</td>
<td>5 minutes per</td>
<td>9.5/minute</td>
<td>9.5×20 = 190</td>
</tr>
<tr>
<td>toilets</td>
<td>3 times per</td>
<td>6/flush</td>
<td>6×12 = 72</td>
</tr>
<tr>
<td>faucets</td>
<td>5 minutes per</td>
<td>2/minute</td>
<td>2×20 = 40</td>
</tr>
<tr>
<td>kitchen</td>
<td>15 minutes total</td>
<td>8/minute</td>
<td>8×15 = 120</td>
</tr>
<tr>
<td>dishwasher</td>
<td>1 use</td>
<td>11/use</td>
<td>11</td>
</tr>
<tr>
<td>front-loader</td>
<td>1 use</td>
<td>55 liters per</td>
<td>55</td>
</tr>
</tbody>
</table>

Faucets have an aerator that performs two functions: It reduces the flow rate out of the faucet, and it substantially reduces splashing when the stream hits an object, such as a dish or human hand. New kitchen faucets have a flow rate of 2.2 gallons/minute, while bathroom faucets are
reduced from 1.5 gallons/minute to 0.5. Aerators do not last forever, but are inexpensive and simple to replace.

The 1992 federal regulations also require that a standard U.S. toilet uses 1.6 gallons/flush. This is a substantial reduction over the 1960s-era toilets that used 3.5 gallons/flush. The early 1.6 gallon models were poorly designed, most likely designed in haste to satisfy the federal regulation, and had problems handling solid waste, sometimes requiring multiple flushes to clear. The more recent models are much better.

It is interesting to note that the bathroom is the single highest source of water consumption in the home, using, according to 2.2 over 70 percent of the domestic water. Logically, the greatest opportunities for water reduction in the home are the shower-heads and toilets. Further, the numbers given in Table 2.2 may be excessive, such as 8 minutes for a shower at 15 liters/minute; the time taken and the flow rate can be substantially reduced, as is discussed next.

Low-flow fixtures include dual flush toilets in which one button provides a 3 liter flush for urine, and a second button provides a 6 liter flush for solids. For families with on-site sewage disposal, known as septic systems, excessive use of water could be a major concern, so the mantra with respect to toilets is:

“If it’s yellow, let it mellow.
If it’s brown, flush it down.”

Showerheads and faucets can be designed to reduce flow without a noticeable reduction of performance; and new appliances like dishwashers and front-loading clothes washers have significantly lower water usage. Using low-flow fixtures can reduce water consumption in the home by over 50%, from 1293 l/d (interior use) to 600 l/d. This does not include outdoor use for landscaping.

Thus, with modern showerheads, faucets, toilets, and dish and clothes washers, for the modern house whose residents are conscious of water use, the water use for the family of four could be as shown in Table 2.3.

<table>
<thead>
<tr>
<th>Table 2.4: How Water Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>use</td>
</tr>
<tr>
<td>showers</td>
</tr>
<tr>
<td>faucets</td>
</tr>
<tr>
<td>kitchen</td>
</tr>
<tr>
<td>dishwasher</td>
</tr>
<tr>
<td>clothes</td>
</tr>
</tbody>
</table>
This family of four uses 488 liters of water a day inside the residence, which is about 38% of the usage of the earlier family with older appliances and faucets.

Under the assumption that water used in faucets and the clothes washer is 50% hot water, while all the water used in showering and the dishwasher is hot water, then the total amount of hot water in liters used by two families of four, one with high usage (old appliances and faucets) and the other with low water usage (new appliances and low-flow faucets), as discussed in Chapter 2 is given in Table 2.4. The high usage family uses 773 liters of hot water a day, while the low usage family uses just 324.5. In gallons these numbers become 204 and 86, and in pounds they become 1693 and 714, respectively.

Supposing the water temperature entering the residence is 9.5°C and the hot water temperature is 42°C, then the temperature rise is 32.5°C. The daily energy needed for these two families is 29 kWh and 12 kWh, respectively; these are year-round numbers. According to the U.S. Energy Information Administration about 18% of the energy needs for an American residence is used to heat water [www.eia.gov/consumption/residential/data/2009/c&e].

After usage, the hot water is disposed of down the drain. Thus, the energy put into producing hot water is lost. Contrast this to the energy in respiration air, 90% of which is captured and returned to the residence via a heat recovery ventilator.

There has been some talk recently in recovering energy from what we send into our municipal sewers using a system similar to that of a ground-loop geothermal system. In such a system, a fluid-carrying pipe is laid horizontally in the ground at a depth such that seasonal temperature changes are minimal. The challenge with such a system is to ensure that there is adequate heat flow between the ground and the pipe.

I built a ground-loop geothermal system in the year 2010 in Massachusetts. The pipe was set 2 meters below the surface where the native temperature is only 9°C. A heat pump with a coefficient of performance about 3 raises this temperature to a useable temperature for heating purposes; the heat pump is reversible, so the system can provide air conditioning in summer.

A city or municipality does not have a problem in assuring a satisfactory heat transfer from sewage. Sewage at about 17°C will enter a tank in which a heat exchanger is set. The sewage leaving the tank will be at a lower temperature. The fluid inside the heat exchanger will pass through a heat pump to raise its temperature moderately for radiant floor heating in municipal buildings in winter, or high enough to generate electricity.

The energy needed for domestic hot water in the United States is larger, by far, than that of less wealthy countries, or countries with a limitation on piped-in water. Changing water usage by legislation or mandate is difficult, if not impossible.

The location of the hot water tank with respect to the kitchen and the bathrooms is of some importance. It should be located to minimize the length of plumbing runs to the point of use.
The standard size of pipe for such runs in North America is $3/4"$, which has an I/D (Internal Diameter) of 0.745”. 1’ of such pipe contains 0.0227 gallons, so a 50’ run contains 1.135 gallons. If 2 or more bathrooms must be serviced, the required feed pipe size could be 1” with I/D of 0.995” and hold 0.0404 gallons per foot; a 50’ run of such pipe holds 2.02 gallons. Thus, to draw water until it is warm enough to wash one’s hands wastes 1.135 gallons or 2.02 gallons, depending on pipe size, before the washing itself commences.

The prices for low-flow fixtures such as toilets, showerheads, and faucets are competitive with those of standard fixtures. Performance is also comparable for quality fixtures (though there are still myths being circulated that toilets must be double flushed and showerheads can’t remove soap). There is no reason not to buy quality, low-flow fixtures for a residence.

Stresses between nations towards the end of the twentieth century often centered around oil — oil needed for transportation, heat, to fuel industrial demand, etc. Now, in the early part of the twenty-first century, more talk is concerning potable water, or, more accurately, lack of it.

At the time of this writing, the southwestern states of the United States are experiencing a record drought. California, in particular, is running out of potable water. Drilling more wells is causing permant damage. “In some places, water tables have dropped 50 feet or more in just a few years. With less underground water to buoy it, the land surface is sinking as much as a foot a year in spots, causing roads to buckle and bridges to crack. Shallow wells have run dry, depriving several poor communities of water. Scientists say some of the underground water-storing formations so critical to California’s future, typically, saturated layers of sand or clay are being permanently damaged by the excess pumping, and will never again store as much water as farmers are pulling out [A.C. Revkin, ‘California’s Wasteful Water Habits Run Up Against a Dry Future — and Past,’ New York Times, April 5, 2015].”

2.10 Electrical Fixtures and Appliances in a Residence

There is a considerable amount of erroneous information regarding energy use and energy efficiency, and environmental impact. When it was first introduced, tubular fluorescent lighting was considered so efficient that it was not worth turning it off; particularly, it was said that most energy was needed at startup. Consider a 120-volt, 40-W fluorescent fixture taking a second to start up. If the energy it needs is the same as one hour’s usage then the startup current would be $40 \times \frac{120}{3600} = 1200$ amps, obviously ludicrous.

Almost always ignored is the standby electricity in residential appliances and electronics, estimated to consume up to 10% of the total electric usage in the structure, as per the U.S. Department of Energy.

Mercury in fluorescent and compact fluorescent lighting elements is of concern, so some would continue to use inefficient incandescent lighting. However, the amount of mercury in
fluorescent lighting is far less than that produced in the power station that feeds that incandescent bulb.

2.10.1 ENERGY STAR Appliances

As a response to the need for energy conservation, in the early 1990s the U.S. government created the ENERGY STAR® rating for domestic appliances. Initially it was used to rate computers and printers. It now includes all common appliances, such as refrigerators, stoves, washing machines, furnaces, televisions, and water heaters.

Consumers can now see the savings in energy costs of a new appliance over their old existing one. When the payback period for a new appliance is less than ten years, there is an incentive to replace. When the payback period is twenty or more years, there is little incentive. The payback period of the 3.8 kW photovoltaic system I installed in 2011 on the house I built was about 4 years, well worth installing. Another consideration is the expected life of an appliance, which has increased to close to 20 years on average.

To illustrate how much lower in energy usage new appliances are than the old ones, these figures provided by the Natural Resources Defense Council are instructive [www.nrdc.org/air/energy/fappl.asp, as accessed on July 15, 2016]. “New refrigerators consume 75% less energy than those produced in the late 1970s. A family replacing a 1980 vintage refrigerator with one that meets today’s energy standards will save more than $100 a year in utility costs. Go one step further and buy an ENERGY STAR-qualified model, and your new refrigerator will save you an additional 15% or more by employing better insulation, more efficient compressors and more precise temperature control and defrost mechanisms . . . . Refrigerators with freezers on top use 10 to 15% less energy than a side-by-side model of equivalent size.”

ENERGYSTAR®-rated dishwashers and front-loading clothes washers are also becoming more competitive with conventional appliances, but there is still a significant price gap. As with other environmentally friendly technologies, the ‘payback’ is dependent upon use — the less you conserve, the faster these technologies financially pay back.
The $350 difference between a conventional and an ENERGYSTAR® dishwasher will have a simple payback over approximately 18 years, at one use per day. This is based on a reduction of water use and energy required to heat the water to 60°C. This financial payback exceeds the design life-cycle of the dishwasher, but has a lower environmental impact. As such, an energy efficient dishwasher is recommended.

“The implementation of energy standards for appliances by the U.S. Department of Energy established new benchmarks for energy and water efficiency. A significant proportion of the energy savings for today’s automatic dishwashers comes from the reduction in hot water use. Because energy is used to heat water, less water use by a dishwasher also means reduced energy use. In 1978, 83% of a dishwasher’s energy use went to heating water, with 10% used for washing and 7% for drying. By 1994, only 56% of the energy used by the dishwasher was to heat water. A significant reduction in water usage resulted from designing more efficient wash systems that incorporate direct water delivery and improved soil-handling systems. The average water use per dishwasher cycle decreased from a range of 11–15 gallons per normal cycle in 1978 to 6–10 gallons per normal cycle in 2000 [www.sdahq.org/dishwash/understanding_automatic_dishwashing.html, as accessed December 3, 2012].”

Most major retailers in the United States try and sell the product on an extended warranty. In almost all cases, this is a bad deal for the consumer, but a major source of revenue for the store and its sales assistant, who typically gets a commission on this un-needed warranty. The original manufacturer provides a warranty that is good for failures and breakdowns for a year or so. Appliances and electronics are subject to ‘infant mortalities,’ failures that occur early. The failure rate of an appliance can be illustrated with what is called the bathtub curve, shown in Figure 2.14. After an initial steep period of failure, what is called in electronic circles the burn-in period, the appliance has a low failure rate until the failure period commences. The use of the term bathtub is evident from the shape.

The ENERGY STAR® label on an appliance indicates that it is more efficient than the
minimum efficiency standard set by the U.S. government. Some state governments or utility companies offer rebates on ENERGY STAR® models, to encourage energy conservation. Since all new appliances must have an ENERGY STAR® rating, the label itself is not an indication of the most efficient model. The consumer must shop carefully, and consider the need before looking at the models. For example, oversized appliances waste energy and cost more.

Electric water heaters have far more expensive running costs compared to natural gas models. The cost of natural gas in the United States has fallen in recent years, caused by the rapid development of new sources using hydrofracking, see Section III of Chapter 1.

2.10.2 Electrically Generated Lighting

Incandescent lighting is a major energy hog. Very little of the electrical energy is converted to visible light: most of it is dissipated as heat. In effect it is really an electric heater that gives off a small amount of light as a byproduct. Incandescent lighting is banned in Germany due to its inefficiency: its output efficiency is about 15 lm/W (lumens/watt).

Typical light outputs for such bulbs are:

\[
\begin{align*}
40 \text{ watt} & - 450 \text{ lumens} \\
60 \text{ watt} & - 800 \text{ lumens} \\
100 \text{ watt} & - 1600 \text{ lumens}
\end{align*}
\]

Fluorescent bulbs are much more efficient. A 15-watt T8 bulb produces over 800 lumens for 1/4 the energy use of the incandescent bulb, and it has a considerably longer life.

Early fluorescent bulbs had a color problem — they were too white and so too cold-looking. Their color temperature was about 5000°K. A cool white bulb has the color temperature of 4200°K, while a warm white bulb has the color temperature of 3000°K. LED lighting has the same temperature problem as the early fluorescents, but this problem will soon be overcome.

The LED is compact and durable, and will be the central part of the light sources of the future. Unfortunately, at time of this writing, few LED light fittings are available.

Visible light is wavelength dependant, and so color dependant, from violet at 0.3μ to red at 0.7μ. The eye is most sensitive to green light at 0.55μ: With the same total energy levels, illumination with predominant energy at about 0.5μ appears more intense than light that has an energy spectrum which is flat across the visible range.
It is commonly assumed that the maximum illumination that can possibly be obtained is 683 lm/W (lumens per watt); this is an ideal and unattainable number. By way of comparison, here are the outputs in lm/W of common light sources [www.mge.com/home/appliances/lighting, as accessed on September 21, 2012]:

- incandescent 10-17
- halogen 12-22
- white LED 20-254
- mercury vapor 25-60
- linear fluorescent 30-110
- compact fluorescent 40-70
- high pressure sodium 50-140
- metal halide 70-115

That is, the efficiency of an incandescent light bulb is about 2%, of a compact fluorescent about 8%, high pressure sodium bulb about 15%, and white LED at 254 lumens/watt is 37%. These numbers are changing, and changing for the better, largely due to governmental prodding.

The Energy Independence and Security Act was signed into law by President George W. Bush in 2007, requiring a 25% increase in efficiency for traditional light bulbs. Manufacturers devoted considerable effort to meeting these standards. Advanced Lighting Technologies has developed an incandescent light bulb, the Vybrant 2X, that produces 32 lm/W or double the efficiency of a standard incandescent bulb, with versions in the pipeline that are projected to produce 45 lm/W, or triple the efficiency of today’s standard. Compact fluorescent bulbs are
reaching the market with double the efficiency of earlier versions. The most significant player in the lighting market is the LED, with plummeting prices and soaring efficiencies.

The LED is a semiconductor diode, whose symbol is shown in Figure 2.15, in which electrical current in the forward-biased PN junction produces electroluminescence. The voltage/current characteristic response of the LED is shown in Figure 2.16.

The first LED was developed at General Electric Company in 1962 by Nick Holonyak, Jr., earning him the Lemelson-MIT Prize in 2004. The early LEDs were expensive. It took until 1968 for the Monsanto Company to use gallium arsenide phosphide (GaAsP). Hewlett Packard introduced these LEDs into hand-held calculators the same year. By the 1970s, Fairchild Optoelectronics was producing LEDs for under 5 cents.

Early LEDs produced red light, unacceptable for most lighting purposes. Now the color of LED-produced light can be tailored into almost any wave band, where white (neutral) light is the most desirable.

By 2002, white LEDs were at about 20 lm/W, and by the next year were at 65 lm/W. In 2006 Nichia Corporation was producing white LEDs at about 150 lm/w. Cree, Inc., a North Carolina-based company, announced on April 12, 2012 that they had developed a white LED that produced 231 lm/W; This is over four times the efficiency of the T8 bulb. The color temperature of this LED is 4408°K — compare this to the surface of the Sun at about 6000°K, so this LED will appear warm white to the eye. It is inevitable that the efficiency of LED lighting will continue to improve until a limit is reached based on fundamental physics. Cree announced in February 2013 that it had developed a white LED at 276 lm/W, or 40% of the theoretical maximum possible. One can anticipate the efficiency of LEDs to continue to rise to over 400 lm/W. At the time of this writing it is doubtful that they will ever reach 500 lm/W.

The cost of LED lighting is high, due in part to the complicated circuitry needed, but its lifespan is 25,000 hours, almost 3 years of 24/7 use, better than fluorescent, and far better than incandescent. However, a sensible solution would be to have one point containing the circuitry for every LED in the structure, and radiate out from that with low voltage wiring that connects to the individual LEDs with no need for individual circuitry.

2.11 Noise Control

Noise can be reduced in a residence using passive noise control or active noise control, or both. It is rare that active noise control can be economically employed to reduce noise, internally or externally generated, in a residence. It can be used in a commercial setting when the noise source is well defined and predictable.
The noise of a machine in a factory can be mitigated by producing the same noise but in antiphase. Near the machine a microphone sends the noise to a signal processor that identifies the dominant frequencies and sends the same dominant frequencies in their same relative magnitude, but in anti-phase, through an amplifier to a speaker: this is shown in Figure 2.17. As shown in Figure 2.18, if the noise has a dominant natural frequency that is identified and subtracted from the noise, this leaves a much reduced noise level. Here the signal processor identified the dominant frequency but calculated its magnitude at 95% of the true magnitude, and calculated the phase with an error of 6 degrees.

![Figure 2.17: System for Noise Reduction](image)

For residential applications, good thermal insulation is often the best defense against unwanted external noise.

### 2.12 Return on Investment and Service Life

In this chapter a number of items have been considered that can be classified as a capital expense — appliances, doors, heat recovery ventilators, and noise control equipment. Solar collectors of all kinds are a major expense. The questions asked by the consumer are:
Some costs are unavoidable. Building codes are being upgraded in most places in the world, but particularly so in northern Europe. For example, new codes regarding window U-factors in Europe and North America will be difficult, meaning expensive, to satisfy. The least-expensive choice may fail to satisfy its assigned task from the beginning, or may break down in short order.

It is evident that in order to reduce our carbon footprint, the cost of fossil fuels will continue to rise over the long term. More governments are willing to add the environmental costs of fossil fuels to the cost the consumer pays.

The service life of a residential structure should be 100-plus years, meaning it should be considered to its owner to have an unlimited life. This may be different in some commercial structures such as hotels in popular vacation locations, where replacement is preferable to renovation. Major integrated features of a structure, such as windows, exterior doors, and roofs, should have a service life of 25-plus years. Major domestic appliances such a refrigerators, or dish or clothes washers, should have a service life of 15-plus years. Water heaters appear to fail more frequently, and, at least in the United States, a failure after 7 years is common. Small appliances, such as coffee makers and hand mixers, have a shorter life. Possibly the item with the shortest service life is the newspaper — one day.

PV arrays are a major cost, and most have a declared service life of 25 years. The owner should study the cost/benefit equation. The cost is evident. When I had a turn-key PV system costing $27,000 installed on the house built in 2010, I received a 30% federal tax credit, a 17% state tax credit, and since the array is grid connected, $0.16 for every kilowatt hour generated via reduction in the electric utility bills from the power company. On top of this are the SRECs (State Renewable energy Credits): One credit is issued for every 1000 kWh generated, which at the current time is about $430 per SREC. The 3.8 kWh system has been averaging 420 kWh/month, so over a year it produces 5 SRECs worth $2150 and 420×12×0.16 = $806 (in reduced electric bills) from the utility company. The out-of-pocket expense was 27,000×(1-0.3-0.17) = $14,310. This cost is being reduced by 2150 + 806 = $2956/year, so in less than 5 years the system is paid for. This is a simplistic analysis since it does not cover the the cost of borrowing to cover the original capital cost, and it does not anticipate the increasing cost of utility-produced electricity, but it does clearly indicate that the system was money well spent.
Chapter 3

Heat Flow from a Residence

In Chapter, the temperature, humidity, ventilation, and lighting requirements of a residence to satisfy the needs of people were considered. This chapter concentrates on the heat loss from a residence so as to control the temperature requirements of its occupants. Further, since the heating needs for climate zones 4 through 8 (see Section II of Chapter 9) in winter are considerable, but such needs are minimal for climate zones 1 and 2, it is the colder zones that will be addressed here.

The approach taken in this chapter to heat loss follows the traditional approach to be found in the American Society of Heating, Refrigerating, and Air-Conditioning Engineers Handbook of Fundamentals. This handbook is periodically updated in two versions: SI and IP. IP stands for inch/pound, the Imperial system. Henceforth we term this book, either SI or IP, as ASHRAE Fundamentals.

We leave consideration of the recent advances in LoE (low emmissivity) glass until Chapter 8, and its substantial advantages in a passive solar house, termed passivhaus in Germany, to Section VI of Chapter 10. Germany has an energy rating system for windows that makes passivhaus technology attractive. Other countries, mostly European, are way ahead of the United States in energy standards. This will be discussed in Sections VIII and IX of Chapter 4.

In Chapters 5 and 8 we derive equations from first principles since the underlying physics is well established and reasonably simple. Here, this is not the case, so we are forced on rely on empirical data.

3.1 Thermal Terms

There is some confusion regarding the difference between conductivity and conductance, between resistivity and resistance, and what units we should use. For example, the conductivity of a material is measured in watts per unit thickness over unit area for a defined temperature difference between opposite faces, so the assumption is that the material has a
uniform thickness. Thus, conductivity $\kappa$ can be described as $\kappa \text{ W/m·K/m}^2$, where the symbol K, commonly used to indicate degrees Kelvin, means the temperature difference in degrees celsius $\Delta^\circ$C. You are now scratching your head, since this does not make sense.

The only solution is to use Fourier’s law

$$q = \kappa \cdot A \cdot \Delta T \cdot \ell, (3.1)$$

where $q$ is the heat flow in watts (W), $\kappa$ is the thermal conductivity whose units are to be determined, $A$ is the area in square meters, $\Delta T$ is the temperature difference in degrees celsius, and $\ell$ is the thickness of the material in meters. The assumptions in using this equation are that the material is planar, rather than curved in any way, so that the opposite faces are parallel, and that heat flow is perpendicular to the surfaces: We can ignore fringe effects. In this chapter we are determining heat flow though the skin (walls, roofs, and foundations) of a structure, so these assumptions are fully satisfactory.

Rewriting Fourier’s law so that $\kappa$ is explicit gives $\kappa = q \cdot \ell A \cdot \Delta T$, where the units on the right-hand side are $W \cdot m/m^2 \cdot \kappa = W/m \cdot K$. Thus, the dimensions of conductivity $\kappa$ are $W/m \cdot K$, the material is 1 m thick, and with the assumption that the area over which the heat flow occurs is 1 $m^2$.

Conductance defines the heat flow through a material per unit area per °C and for a specific thickness, so we divide the conductivity by the thickness to get the conductance, and conductance has units $W/\kappa \cdot m^2$.

Resistivity is the reciprocal of conductivity, so it has the units of $m \cdot K/W$ per unit area. Resistance is found by multiplying the resistivity by the thickness of the material in meters. When calculating the heat losses through a composite of different materials, such as occurs in the wall of a house, resistances add up, so working with resistances rather than conductances is common. To find the overall conductance of a composite material, the easiest procedure is to add the individual resistances to get the overall resistance, and then take its reciprocal.

As an example, consider a material $\ell = 0.1$ m thick and $A = 2$ $m^2$ in area that has thermal conductivity of $\kappa = 3$ W/m·K. If the temperature differential across the material is $\Delta T = 5$ °C, then the heat flow across it is $\kappa \cdot A \cdot \Delta T \cdot \ell = 3 \times 2 \times 50.1 = 300$ W.

The United States is the only major country to cling to the old Imperial units. The United Kingdom had mostly converted to the SI system by 1980, although they still refer to distances on the highway in miles as well as kilometers, and the English pint of beer endures. In English units, thermal conductivity is given in BTU/h·"·F, or BTU/h·’·F, or BTU/h·yard·F, or BTU/h·rod·F (I am joking of course).

The BTU (British Thermal Unit) is defined as the amount of heat required to raise the temperature of 1 pound weight of water at temperature 39°F, when it is at its highest density, by
one degree Fahrenheit. The jouls and the BTU define energy, whereas the watt defines power, a flux, so to convert a BTU to a flux we need to use BTU/h.

In the United States, the unit of thickness in determining resistance from resistivity is the inch, but the unit of area is the square foot. This is inconsistent but has the saving grace that the physical sizes feel more appropriate. The thickness of a wood-framed wall in the United States is 3.5” if the studs are 2x4s, or 5.5” if the studs are 2x6s, and the fiberglass insulation is available in 3.5” thickness or 5.5” thickness to accommodate the stud sizes; fiberglass is also available in other thicknesses, particularly larger thicknesses for roof insulation.

The only satisfactory way of converting from SI to Imperial units is to perform a dimensional analysis using Fourier’s law as

\[ 1 \text{W/m-K} = q \cdot l \cdot A \cdot \Delta T(W \cdot \text{mm}^2 \cdot \text{K}) = 3.412 \times 39.3710.764 \times 9/5(\text{BTU/h \cdot \text{inch}^2 \cdot \text{ft}}'F0) = 6.933 \text{ BTU/h \cdot \text{inch}'F0} \]

The problem with a strict application of SI units is that thickness measured in meters is not ideal when considering the wall of a residence. Inches are better, but should not be mixed with SI units.

Dimensional lumber in the United States refers to wood rough-sawn to a rectangular shape and then planed with eased edges to exact dimensions, so the thickness is 1.5”, and the width is 3.5” for a 2x4 to 11.25” for a 2x12. The sidewall of a framed house typically uses 2x4s or 2x6s as its structural members, so the overall thickness with 1/2” wallboard on the inside and 1/2” plywood sheathing on the outside will be 4.5” or 6.5” before the siding is applied.

The CLS (Canadian Lumber Standard), the European Scant standard, and the UK standard all specify dimensions in millimeters. Unlike the U.S. designation, which quotes the small number first, often the reverse is true in Europe. Typical rough-sawn CLS sizes are 75mm x 50mm, which finishes to about 63mm x 38mm (2 1/2” x 1/2”), and 100mm x 50mm, which finishes to about 89mm x 38mm (3 1/2” x 1 1/2”). The finished European Scant is usually a little larger than the CLS.

What Americans call lumber, people in the UK call limber. Some rough-sawn limber sizes in millimeters in the UK are 25x59, 38x87, 38x100, 47x75, 47x100, 50x75, 50x100, 75x75, 75x100, and 75x125. The species called whitewood is actually spruce. As in Canada, the planed limber is considerably smaller than the original rough sawn piece.

There is a void between dimensions in millimeters and those in meters. The number that seems to work best is the archaic inch. What could work better than the SI meter is to replace it, as far as thermal calculations are concerned, with the centimeter.

This means that typical plywood thicknesses would be 1.25cm or 1.6cm, and typical framing limber would be 9cm or more in width. However, it appears that the millimeter holds sway,
Scientific correctness insists that one use SI units, but as far as practicality in considering heat transfer through the wall of a house is concerned, this presents some difficulty. The meter is a large and awkward measure of wall thickness. The centimeter would be better, but unfortunately Europeans, including the British, insist on using millimeters (mm) when discussing building materials.

Since 75\% of the lumber cut and processed goes to its giant neighbor to the south, Canada uses Imperial sizes for construction materials. Europe is a mixed bag. A sheet of plywood in England is 1220 × 2440 in millimeters (mm). Converting this to inches results in 48.03” × 96.06”, essentially the same as the standard U.S. size of 4’ × 8’. Plywood sizes are not standard in the rest of Europe. In Finland, a major producer of plywood, sheet sizes are 1200, 1220, or 1250 mm wide, with lengths of 2400, 2440, 2500, 3000, and 3600 mm.

A wood cavity wall is sheathed in plywood with the longer dimension being horizontal, and to save on cuts the studs supporting the plywood need to be spaced evenly. This means that the studs are either 16” on center or 24” on center when the plywood sheet is 8’ (2440 mm) long. In Norway the standard stud spacing is 600 mm, 23.62”, so the plywood is 2400 mm long. In Australia the stud spacing is either 450 mm or 600 mm.

### 3.2 Empirical Nature of Heat Loss Calculations

Most of the technical information presented in this text is based on the underlying physics, but the study of the mechanisms of heat transfer is far more difficult to analyze using physics. We know there are three mechanisms at work — conduction, convention, and radiation. We know that convection requires a fluid, in our case almost always air, so a vacuum between two surfaces prohibits heat flow by convection. But soon our analytic knowledge is exhausted and we are forced to rely on experimental observations made under controlled circumstances.

Heat loss through vertical, double glazing would seem to be an easy problem to solve. Unfortunately, it is not. With the outside temperature much lower than the internal temperature, there is heat flow from the interior to the inside surface of the inside glass sheet, and this heat flow is controlled by more than the temperature differential: Internal humidity matters, as does interior air flow. As heat passes into the glass, a natural flow of air bathes the glass in a downwards direction. If the window is directly above a baseboard radiator, then that air flow is reduced or possibly reversed. Air flows upwards against the outside glass surface, and this can be considerably disrupted by wind against the glass — windy conditions always increase winter heat losses. This is only the beginning of our problems. Heat loss near the window frame is greater than at the center of the window, so the heat loss per unit area is a function of the size and shape of the window.
The standard method for determining heat losses from a structure, at least in the United States, is to consult the bible for heat flow calculations — the ASHRAE Fundamentals. Some may find this unsatisfactory since virtually all the data regarding heat flow in that reference are empirically based: that is, based on measurements rather than derived from the physics. However, the mechanisms of heat loss are quite complicated and based on many factors, so the empirical data are more compelling than theoretical calculations.

The need for theoretical calculations arises when one considers more exotic systems, such as the heat pipes discussed in Section 8 of Chapter 12, since there are few ASHRAE empirical studies to rely on. Further, the salient variables in heat pipe calculations vary widely with the design, so one is forced to rely on physics-based studies or in-house experimentation.

3.3 Fundamentals of Glazing Losses

Air flow on the inside and outside of double glazing, as well as flow between the glass panels, can be described as shown in Figure 3.1. The heat that arrives at the inner layer of glass is dominated by convection and radiation; conduction is insignificant. The radiation is from bodies at about 290°K that have a peak radiation energy at wavelengths of about 10 \( \mu \) wavelength of light is measured in micrometers, \( 10^{-6} \) m, termed microns and given the symbol \( \mu \).

Glass is opaque to radiation energy with wavelengths longer than 2.7\( \mu \); see Section II of Chapter 8. Between the glass panels there is convection and conduction. Outside the outer glass panel the predominant mechanism of heat loss is convection, and this is a function of temperature differential between the room and the environment, and the wind shear across the glass.

The temperature changes from interior to exterior for single glazing are shown in Figure 3.3. The two surfaces of the glass, inside and outside, provide resistance to heat flow. The interior surface resistance is \( R_i \) and the external surface resistance is \( R_e \). The total surface resistance is \( R_s = R_i + R_e \cdot R_e < R_i \) due to wind. The standard wind speeds assumed are 12 km/h and 24 km/h. 24 km/h is the same as 6.7 m/s and 24/1.609 = 14.9 mph, which is commonly taken as 15 mph. 12 km/h is the same as 7.5 mph.

For vertical walls, the surface resistance to the interior is rated as \( R_i = 0.68 \text{ ft}^2\cdot\text{h} \cdot \text{F}/\text{BTU} \), and the surface resistivities of the outer skin are rated as \( R_e = 0.25 \) and \( R_e = 0.17 \text{ ft}^2\cdot\text{h} \cdot \text{F}/\text{BTU} \) for 7.5 and 15 mile/hour winds, respectively. Thus, the total surface resistance for the vertical surfaces is \( R_s = 0.68 + 0.25 = 0.93 \) and \( R_s = 0.68 + 0.17 = 0.85 \text{ ft}^2\cdot\text{h} \cdot \text{F}/\text{BTU} \) for exterior winds of 7.5 and 15 mile/hour, respectively.
The heat loss $U$ for a homogeneous vertical panel, such as glass, of conductance $k$ is given by $U=1Rs+1/k$ ft$^2$·h·F/BTU. If glass was a perfect conductor so $k \to \infty$, $U=1Rs=10.93=1.075$.

Glass has conductivity $\kappa = 5.9$ for a 1” thickness, so a single glazed 3/16” window has $U=10.93+0.1875/5.9=1.040$. Thus, the insulating effect of single glazing lies not in the thermal resistance of the glass but in its surface resistance. This can be seen in Figure 3.3, where there is a negligible temperature difference between the inside and the outside of the glass.

The conversion from Imperial to SI units is simple. Surface resistance $R_i$ BTU/ft$^2$·h·F translates to $R_i = 0.68/6.933 = 0.098$ m$^2$·K/W. Why do we not use SI units first, or even exclusively? The problem is the inconstance in the numbers given in SI units, and this includes the venerable ASHRAE Fundamentals, SI Edition. Also, there are few reliable
sources, and we have already talked about the awkwardness of the meter applied to building elements.

![Figure 3.3: Temperature Gradients with Single Glazing](image)

The heat loss through double glazing is more difficult to calculate due to the air gap. As seen in Figure 3.4 there is a significant temperature differential between the glass surfaces facing each other. The classical way of determining the heat transfer is given in the next section, but even this is not satisfactory.

### 3.4 Heat Flow Across an Air Gap

If the air gap between two vertical sheets of glass is small, there is no convection, so the heat losses are by conduction and radiation. The controlling parameter to determine if there is convective heat loss is the Nusselt number $N_u$: if $N_u \leq 1$ there is no convection. If $N_u \geq 1$ the conduction and convection heat loss is $N_u$ times the conduction heat loss [Y.A. Cengel, *Heat Transfer: A Practical Approach*, McGraw Hill, New York, 1998].
Consider a situation where the inner and outer glass temperatures are $T_i \, ^\circ C$ and $T_o \, ^\circ C$, respectively, $H$ is the height of the glass panels, and $D$ is the width of the air gap. Then

$$Nu = 0.197Ra^{1/4}(H/D)^{-1/9}(3.2)$$

where the Rayleigh number $Ra$ is given by

$$Ra = g(T_i + T_o)D^3Pr^2v^2(Ta+273)(3.3)$$

where $g = 9.81 \, \text{m/s}^2$ is the acceleration due to gravity, $Pr$ is the Prandl number, and $v$ is the kinematic viscosity in m$^2$/s [op. cit.]. It is commonly taken that if $Ra \leq 2000$ then $Nu = 1$. Further, the minimum value of $Nu$ that can be used is unity. The conductive/convective heat loss across the air gap per unit width is given by

$$Q = kNuH(Ti - T_o)D. (3.4)$$

$k$ is the thermal conductivity of air in W/m·K.
Suppose $T_0 = 0 \, ^\circ C$, then the air gap to produce $N_u = 1$ as a function of $T_i$ is shown in Figure 3.5. Notice that as the temperature differential between the panes of glass rises, the air gap reduces to avoid convection. Also shown in Figure 3.5 is the Rayleigh number.

Maintaining $N_u = 1$ to avoid convection losses, the conductive and radiative heat losses across the air gap as a function of temperature differential are shown in Figure 3.6. As can be seen, the radiative heat loss is about double the conductive heat loss. Further, the radiative heat loss is proportional to the temperature differential.

It is rare to use the analysis just presented to determine the heat loss through glass. Instead, the empirically based tables and formulae given in the ASHRAE Fundamentals are most often consulted. It is well known that the radiative air gap resistance depends strongly on the emittance of the surfaces — if one or both surfaces have a low emissivity, the radiative heat loss will be negligible. The radiative and conductive components of heat flow are independent of the tilt angle of the surfaces. The conduction and convection air gap resistances depend on the width of the gap — they increase with the gap until it reaches about 1.1” (3 cm), and thereafter are fairly constant. There is more discussion of air gap resistance in the next section.
With an air gap between two surfaces of emissivity $\varepsilon_1$ and $\varepsilon_2$, the air gap has an effective emissivity of

$$\varepsilon_a = \frac{1}{1/\varepsilon_1 + 1/\varepsilon_2 - 1}. \quad (3.5)$$

Consider the following cases:

1. $\varepsilon_1 = \varepsilon_2 = 0.9$, then $\varepsilon_a = \frac{1}{1/0.9 + 1/0.9 - 1} = \frac{1}{1.2222} = 0.819$.

2. $\varepsilon_1 = \varepsilon_2 = 0.05$, then $\varepsilon_a = \frac{1}{1/0.05 + 1/0.05 - 1} = 0.0256$.

3. $\varepsilon_1 = 0.9$ and $\varepsilon_2 = 0.05$, then $\varepsilon_a = \frac{1}{1/0.9 + 1/0.05 - 1} = \frac{1}{1.111 + 20 - 1} = \frac{1}{0.0497}$.

Table 3.1: Air Gap Resistances

<table>
<thead>
<tr>
<th>emissivity</th>
<th>horizontal</th>
<th>up</th>
<th>down</th>
<th>45° up</th>
<th>45° down</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.026</td>
<td>2.64</td>
<td>1.84</td>
<td>8.94</td>
<td>2.08</td>
<td>0.85</td>
</tr>
<tr>
<td>0.82</td>
<td>0.92</td>
<td>0.80</td>
<td>1.23</td>
<td>3.57</td>
<td>1.02</td>
</tr>
</tbody>
</table>

The air space resistance in h·F/BTU for one square foot for air gaps of 3/4” (19 mm) for the indicated heat flow is shown in Table 3.1 [Table 3, Chapter 25, 2001 ASHRAE Fundamentals].

![Figure 3.7: Air Gap Resistance as a Function of Effective Emissivity](image)

The definitive works on the thermal resistance of air spaces can be found in ASHRAE Fundamentals. However, this cites, as do others, the work of Robinson and Powlitch [H.E. Robinson and F.J. Powlitch, “The Thermal Insulating Value of Air Spaces,” USA Division of Housing Research, HHFA, Research Project ME-12 - HR32, April, 1954].
The air gap resistance \([\text{op cit}] R\) in \(\text{m}^2 \cdot \text{K}/\text{W}\) for a 25 mm 1” unvented air space at mean temperature of 25°C is shown in Figure 3.7 as a function of \(\epsilon_a\), the effective emittance of the two parallel surfaces. Notice that the resistance to heat flow upwards in the air gap between two horizontal surfaces is much lower than that for heat flow downwards. Further, the thermal resistance between two parallel vertical surfaces, termed “wall” in the diagram, is a little lower than that of horizontal surfaces with heat flow upwards for low \(a\), but is essentially the same for \(a > 0.5\).

![Figure 3.8: Resistance as a Function of Air Gap](http://fricker.net.au)

Figure 3.8 shows the thermal resistance \(R\) for air flow up, down, and horizontally as a function of the air gap in millimeters [http://fricker.net.au](http://fricker.net.au), as accessed on October 1, 2012, which refers to Fricker’s published paper, [J.M. Fricker, “Computational Analysis of Reflective Air Spaces,” AIRAH Journal, October 1997]. The air gap emissivity is 0.03, the mean temperature is 27°C, and the temperature differential is 4°C. \(R\) continues to increase for heat flow downwards as the air gap increases. There is little increase in \(R\) for air gaps greater than 30 mm for heat flow upwards, where \(R \approx 0.5 \text{ m}^2 \cdot \text{K}/\text{W}, \text{ or } 2.8383 \text{ ft}^2 \cdot \text{h} \cdot \text{F}/\text{BTU}\).

The thermal resistance for the wall peaks at about 0.76 \(\text{m}^2 \cdot \text{C}/\text{W} \approx 4.316 \text{ ft}^2 \cdot \text{hr} \cdot \text{F}/\text{BTU}\) for an air gap of 30 mm (about 1.2”), and declines a little after that. This air gap is much larger than is normal for double glazing [“Thermal Insulation Performance of Reflective Material Layers in Well Insulated Timber Frame Structures,” Sivert Uvsløkk, Senior Scientist, and Heidi Arnesen, Research Scientist, both at SINTEF Bygforsk, [http://web.byv.kth.se/bphys/copenhagen/pdf/118–2.pdf](http://web.byv.kth.se/bphys/copenhagen/pdf/118–2.pdf), as accessed on January 14, 2012]

The thermal conductivity in BTU/ft\(^2\)-h·F for air gap in inches and with surface emissivities of \(\epsilon = 0.9\) is shown in Figure 3.9. Notice that for air gaps of 7.5 mm (0.3”), or larger the dominant mechanism of heat transfer is radiation.
Changing the emissivity of one face of the air gap to 0.05 results as shown in Figure 3.10. Here the radiative heat loss is minimal, the conduction heat loss is dominant for air gaps less than 5 cm (2”), and the convection is dominant for air gaps over 5 cm.

The thermal resistance for walls, roofs, and floors as a function of air gap with the emissivity of one face of the air gap at 0.05 results as shown in Figure 3.11. Also shown is the resistance for fiberglass insulation as a function of thickness. For air gaps (thickness of fiberglass) greater than 3” it is evident that fiberglass is superior, and fiberglass is competitive for lower gaps (thicknesses).

![Figure 3.9: Heat Losses across an Air Gap with Emissivity 0.9](image1)

![Figure 3.10: Heat Losses across an Air Gap with Emissivity 0.05](image2)
Thermal conductivity is an inherent attribute of the homogeneous material being considered, be it wooden studs or fiberglass insulation. Thermal conduction typically refers to the heat flow in watts across a composite of a number of different materials per unit area and per °C temperature differential. How do we get to thermal conduction $C$ from thermal conductivity ($c$)? We go through thermal resistance ($R$).

We will term thermal conductivity in SI units as $c_{si}$, and in Imperial units as $c_{imp}$, where $1\ c_{si} = 6.974\ c_{imp}$. Thermal resistivity is $r=1c_{si}$, and so $1\ r_{si}=r_{imp}6.974$. In Europe it is common to use thermal conductivity, whereas in the United States, resistivity is more commonly used.

Interior air gaps in walls and roofs can have good insulating properties. The resistance of an interior air gap $R_a$ is a function of orientation, width of the air gap, and the emissivity of the materials bordering the air gap. Glass, wood, masonry, non-metallic paints, and most building materials have high emissivity (about 0.9, where the maximum emissivity possible is 1.0). Bright polished metal sheets, such as aluminum foil, have emissivities about 0.05. The emissivities of common construction materials along with the thermal resistivities are given in tables later in this section.
Consider a composite wall shown in Figure 3.12, composed of three layers of materials with resistances $R_1$, $R_2$, $R_3$ and interior and exterior skin/air resistances of $R_i$ and $R_e$, respectively. The overall resistance is given by

$$R = R_i + \ell_1 r_1 + \ell_2 r_2 + \ell_3 r_3 + R_e = R_i + R_1 + R_2 + R_3 + R_e \tag{3.6}$$

where $\ell_1$, $\ell_2$, $\ell_3$ are the respective thicknesses of the materials.

There are typically no air gaps in a wall, or the air gaps are so small as to have no effect on the heat transfer characteristics. The exception to this is a wall with multilayers of reflective material whose individual emissivity is low; see Section 3.5 of Chapter 4.

Notice that heat flow is unidirectional through a wall, so the calculations are one-dimensional. When considering total heat loss from a wall with different composites, we calculate the resistance $R_j$ and surface area $A_j$ of each composite, then the heat loss for the complete wall is

$$H = \Delta T \sum_j = 1 n A_j R_j \tag{3.7}$$

where $\Delta T$ is the temperature differential from interior to exterior. Following standard procedure, we consider only steady-state heat losses. We check the dimensions in Equation 3.5 in both SI and Imperial units:

$$\text{oC} \sum_i = 1 \text{nm}^2 \text{m}^2 \text{oC}/\text{W} = \text{W} \text{oF} \sum_i = 1 \text{ft}^2 \text{ft}^2 \text{oF}/\text{BTU} = \text{BTU}/\text{h} \tag{3.8}$$

One composite could be a cavity wall filled with insulation. Another could be the framing members, called studs. Another could be a window. Typically, we ignore any cross-coupling effects, which are usually insignificant, as discussed earlier. The exception is the decrease in resistivity of windows close to the frame; this is discussed at the end of this section.
Heat losses are sometimes calculated by using electrical resistance models. Thermal resistance in ohms is equivalent to resistance in $m^2 \cdot K/W$. Temperature in °C is equivalent to voltage. Situations may occur in which heat transfer calculations involve composites that do interact and permit heat to flow cross-wise as well as linearly; in this case the electrical resistance model is valuable. However, this is not the case in determining heat transfer from a structure, so the electrical resistance model is unnecessary.

<table>
<thead>
<tr>
<th>$m^2 \cdot °C/W$</th>
<th>up</th>
<th>horiz</th>
<th>down</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_i$</td>
<td>0.1</td>
<td>0.13</td>
<td>0.17</td>
</tr>
<tr>
<td>$R_e$</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
</tbody>
</table>

The surface resistance in $m^2 \cdot K/W$, as extracted from Table 1 of EN ISO6946:2007, “Building Components and Building Elements – Thermal Resistance and Thermal Transmittance,” is shown in Table 3.2. Notice that external surface resistances $R_e$ are the same: this is due to an assumed non-zero wind speed. Also notice that $R_i$ is maximum for downward heat flow, indicating that downward heat flow is less than horizontal heat flow, which makes perfect sense. However, as seen in Section IX of Chapter 4, the IECC code considers downward heat flow through a floor to be greater than horizontal heat flow through a wall, which makes no sense at all.
The thermal resistivities $r$ in h·F/BTU for one inch thickness and over one square foot of area for masonry materials are shown in Table 3.3. Also shown are the emissivities of these materials. In some cases a range is given.

In Table 3.3 one sees concrete (vermiculite aggregate) can have the resistivity of 1.6, which is above the resistivity of softwoods. However this resistivity is for a mix of 1/7 of portland cement to vermiculite, and this mix at 100 psi is far below the typical strength of regular concrete at 2000 psi or higher; in particular, it should not be used as the foundation of a structure. Further inhibitors to using it are availability and cost.

AAC blocks are also included in Table 3.3. AAC (autoclaved aerated concrete) blocks are the preferred material for exterior residential walls in parts of Europe and the Middle East, as well as Asia, Australia, and South America. More will be said about AAC blocks in Section 2.27 of Chapter 4.

<table>
<thead>
<tr>
<th>masonry materials</th>
<th>$r$</th>
<th>$\varepsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td>concrete (sand &amp; gravel aggregate)</td>
<td>0.08</td>
<td>0.92</td>
</tr>
<tr>
<td>concrete (cinder aggregate)</td>
<td>0.2</td>
<td>0.92</td>
</tr>
<tr>
<td>concrete (vermiculite aggregate)</td>
<td>0.4–1.6</td>
<td>0.94</td>
</tr>
<tr>
<td>common brickwork</td>
<td>0.2</td>
<td>0.93</td>
</tr>
<tr>
<td>cement mortar</td>
<td>0.2</td>
<td>0.87</td>
</tr>
<tr>
<td>hollow concrete block (sand &amp; gravel aggregate)</td>
<td>0.14</td>
<td>0.9</td>
</tr>
<tr>
<td>4” hollow concrete block (cinder aggregate)</td>
<td>0.3</td>
<td>0.93</td>
</tr>
<tr>
<td>8” hollow concrete block (cinder aggregate)</td>
<td>0.22</td>
<td>0.95</td>
</tr>
<tr>
<td>12” hollow concrete block (cinder aggregate)</td>
<td>0.16</td>
<td>0.97</td>
</tr>
<tr>
<td>AAC blocks</td>
<td>1.25</td>
<td></td>
</tr>
<tr>
<td>stone</td>
<td>0.8</td>
<td>0.93</td>
</tr>
<tr>
<td>asphalt</td>
<td>0.2</td>
<td>0.93</td>
</tr>
<tr>
<td>slate</td>
<td>0.1</td>
<td>0.67–0.80</td>
</tr>
<tr>
<td>dry sand</td>
<td>0.6–1</td>
<td>0.75</td>
</tr>
<tr>
<td>moist sand</td>
<td>0.07–0.6</td>
<td>0.75</td>
</tr>
<tr>
<td>saturated sand</td>
<td>0.04–0.07</td>
<td>0.75</td>
</tr>
<tr>
<td>dry clay</td>
<td>1</td>
<td>0.91</td>
</tr>
<tr>
<td>moist clay</td>
<td>0.08</td>
<td>0.91</td>
</tr>
<tr>
<td>saturated clay</td>
<td>0.06</td>
<td>0.91</td>
</tr>
<tr>
<td>dry earth</td>
<td>1.4</td>
<td>0.9</td>
</tr>
<tr>
<td>damp loose earth</td>
<td>0.27</td>
<td>0.94–0.96</td>
</tr>
<tr>
<td>damp tight-packed earth</td>
<td>0.15</td>
<td>0.96</td>
</tr>
</tbody>
</table>
Table 3.4: Thermal Resistivities and Emissivities for Interior Materials

<table>
<thead>
<tr>
<th>interior materials</th>
<th>r</th>
<th>$\varepsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td>exterior glue plywood, softwoods</td>
<td>1.25</td>
<td>0.87</td>
</tr>
<tr>
<td>hardwoods</td>
<td>0.9</td>
<td>0.87</td>
</tr>
<tr>
<td>sawdust</td>
<td>2.8</td>
<td>0.75</td>
</tr>
<tr>
<td>wood fiberboard</td>
<td>1.8</td>
<td>0.85</td>
</tr>
<tr>
<td>hardwood fiberboard</td>
<td>0.7</td>
<td>0.85</td>
</tr>
<tr>
<td>corkboard</td>
<td>3.3</td>
<td>0.75</td>
</tr>
<tr>
<td>standard fiberglass batt</td>
<td>3.3</td>
<td>0.75</td>
</tr>
<tr>
<td>closed cell polyurethane board or foam</td>
<td>6</td>
<td>0.97</td>
</tr>
<tr>
<td>expanded polystyrene</td>
<td>3.8</td>
<td>0.95</td>
</tr>
<tr>
<td>extruded polystyrene</td>
<td>5.0</td>
<td>0.96</td>
</tr>
<tr>
<td>loose fill rockwool</td>
<td>2.8</td>
<td>0.4</td>
</tr>
<tr>
<td>loose fill cellulose</td>
<td>3.7</td>
<td>0.34</td>
</tr>
<tr>
<td>acoustic tile</td>
<td>2.4</td>
<td>0.62–0.63</td>
</tr>
<tr>
<td>gypsum board, plaster board, wallboard</td>
<td>0.85</td>
<td>0.9</td>
</tr>
<tr>
<td>gypsum lath &amp; plaster (sand aggregate)</td>
<td>0.18</td>
<td>0.9</td>
</tr>
<tr>
<td>expanded vermiculite</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>iron</td>
<td>0.02</td>
<td>0.16–0.69</td>
</tr>
<tr>
<td>steel</td>
<td>0.0032</td>
<td>0.28–0.59</td>
</tr>
<tr>
<td>aluminum, polished</td>
<td>0.0006</td>
<td>0.05</td>
</tr>
<tr>
<td>copper, polished</td>
<td>0.00036</td>
<td>0.05</td>
</tr>
<tr>
<td>brass</td>
<td>0.0013</td>
<td>0.22–0.61</td>
</tr>
</tbody>
</table>

The resistivities and emissivities of interior materials are shown in Table 3.4. The insulating materials fiberglass batts and foamboard are of most interest here.

The resistivities and emissivities of assorted materials not included in Tables 3.3 and 3.4 are shown in Table 3.5.

Some materials have a standard thickness based on their use, and these are listed in Table 3.6 along with their fixed resistance. For example, vinyl siding has a resistance of 0.61, with no thickness specified.

### 3.6 Windows

When discussing heat flow through a window, the conductance is given the symbol $U$. It has the same units $W/m^2\cdot K$ or BTU/ft$^2\cdot h\cdot F$ as before. Recall, $1 \text{ W/m}^2\cdot \text{K} = 6.933 \text{ BTU/ft}^2\cdot \text{h}\cdot \text{F}$. 
### Table 3.5: Thermal Resistivities and Emissivities for Miscellaneous Materials

<table>
<thead>
<tr>
<th>miscellaneous materials</th>
<th>r</th>
<th>( \varepsilon )</th>
</tr>
</thead>
<tbody>
<tr>
<td>glass</td>
<td>0.17</td>
<td>0.9</td>
</tr>
<tr>
<td>acrylic, clear</td>
<td>0.77</td>
<td>0.94</td>
</tr>
<tr>
<td>water</td>
<td>0.25</td>
<td>0.9</td>
</tr>
<tr>
<td>ice at 32°F</td>
<td>0.67</td>
<td>0.97</td>
</tr>
<tr>
<td>snow</td>
<td>0.59–2.9</td>
<td>0.8</td>
</tr>
<tr>
<td>heavy linen drapes (estimated value)</td>
<td>0.33</td>
<td>0.88</td>
</tr>
<tr>
<td>carpets with rubber pad</td>
<td>1.23</td>
<td>0.85</td>
</tr>
<tr>
<td>carpets with felt pad</td>
<td>2.1</td>
<td>0.85</td>
</tr>
<tr>
<td>felt building paper</td>
<td>0.12</td>
<td>0.87</td>
</tr>
<tr>
<td>plastic film, 6 mil</td>
<td>0</td>
<td>0.84–0.95</td>
</tr>
<tr>
<td>black paint</td>
<td>0</td>
<td>0.9</td>
</tr>
</tbody>
</table>

### Table 3.6: Thermal Resistances and Emissivities for Siding and Roofing Materials

<table>
<thead>
<tr>
<th>siding &amp; roofing materials</th>
<th>R</th>
<th>( \varepsilon )</th>
</tr>
</thead>
<tbody>
<tr>
<td>asbestos shingles</td>
<td>0.21</td>
<td>0.88</td>
</tr>
<tr>
<td>asphalt roll roofing</td>
<td>0.15</td>
<td>0.86</td>
</tr>
<tr>
<td>asphalt shingles</td>
<td>0.44</td>
<td>0.91</td>
</tr>
<tr>
<td>built-up roofing, 3/8” thick</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>pine, laid siding</td>
<td>0.78</td>
<td>0.9</td>
</tr>
<tr>
<td>stucco</td>
<td>0.2</td>
<td>0.89</td>
</tr>
<tr>
<td>vinyl/aluminum siding</td>
<td>0.61</td>
<td>0.91–0.93</td>
</tr>
<tr>
<td>cedar/pine bevel siding</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>wood shingles</td>
<td>0.94</td>
<td>0.9</td>
</tr>
<tr>
<td>3/4” barnboard siding</td>
<td>0.94</td>
<td>0.87</td>
</tr>
</tbody>
</table>

### 3.6.1 Heat Loss Through Windows

The heat loss through a window is commonly determined from the conductance of that window, and the authority for this, at least in the United States, is ASHRAE Fundamentals.

The *U*-factor for double or triple glazing is a function of the air gap(s) between glass sheets, and the emissivity of the glass. Losses are reduced if the spacing between the glass is filled with argon or krypton instead of air. *U* in BTU/ft\(^2\)·h·F are shown in Figure 3.13 for double glazing and low and high emissivity surfaces as a function of the air gap in inches [ASHRAE Fundamentals, 1993, p 27.2, Figure 1]. For an air gap of 1/4” and emissivity 0.84 the *U*-factor is 0.57 BTU/ft\(^2\)·h·F, so the resistance is 1.75 ft\(^2\)·h·F/BTU. For an air gap of 1/2” the *U*-factor is 0.48 BTU/ft\(^2\)·h·F, so the resistance is 2.08 ft\(^2\)·h·F/BTU.
From Table 3 of Section 25.4 of *ASHRAE Fundamentals*: For a vertical 1/2” air gap, so the heat flow is horizontal, the effective emissivity of the air gap being 0.82, the following air gap resistances $R$ ft$^2$·h·F/BTU as a function of mean temperature $T_m$ and temperature differential $\Delta T$ are:

\[
\begin{align*}
T_m &= 50^\circ F, \quad \Delta T = 30^\circ F, \quad R = 0.90; \\
T_m &= 50^\circ F, \quad \Delta T = 10^\circ F, \quad R = 0.91; \\
T_m &= 0^\circ F, \quad \Delta T = 100^\circ F, \quad R = 1.15
\end{align*}
\]

Although the dominant factor in air gap resistance is the effective emissivity of the surfaces facing the gap, the next most important factor is the width of the air gap. Other factors include the mean temperature, the temperature differential, and the proximity to the metal spacer along the perimeter of the glass panes. Also to be considered is the solar heating of the glass — see Section IV of *Chapter 11* for a discussion on this. In total, the heat loss through glass is a complex problem, subject to many factors that cannot be determined in advance. Thus, approximations and inexact rules-of-thumb need to be invoked.
The total surface resistance for vertical glazing with a 15-mph wind will be \( R_s = 0.68 + 0.17 = 0.85 \text{ ft}^2 \cdot \text{h} \cdot \text{F}/\text{BTU}. \) The resistances \( R \) for various glazing situations, assuming a 15-mph wind and all panels vertical, are calculated as follows:

**Single Glazing:**
- 3/16” glass, \( R = 0.85 + 0.1875 \times 0.17 = 0.88 \)
- 1/4” acrylic, \( R = 0.85 + 0.25 \times 0.77 = 1.04 \)
- 1/2” acrylic, \( R = 0.85 + 0.5 \times 0.77 = 1.23 \)

**Double Glazing:**
- (3/4” overall) 3/16” glass, 1/4” air gap, \( R = 0.85 + 0.625 + 2 \times 0.1875 \times 0.17 = 1.54 \)
- 1/4” acrylic, 1/4” air gap, \( R = 0.85 + 0.625 + 2 \times 0.25 \times 0.77 = 1.86 \)
- 3/16” glass, 1/2” air gap, \( R = 0.85 + 0.81 + 2 \times 0.1875 \times 0.17 = 1.72 \)
1/4” acrylic, 1/2” air gap, \( R = 0.85 + 0.81 + 2 \times 0.25 \times 0.77 = 2.04 \)
3/16” glass, 3/4” air gap, \( R = 0.85 + 0.92 + 2 \times 0.1875 \times 0.17 = 1.83 \)

**Triple Glazing:**

(1 3/8” overall)
3/16” glass, 1/4” air gaps, \( R = 0.85 + 2 \times 0.625 + 3 \times 0.1876 \times 0.17 = 2.20 \)
3/16” glass, 1/2” air gaps, \( R = 0.85 + 2 \times 0.81 + 3 \times 0.1875 \times 0.17 = 2.56 \)
3/16” glass, 3/4” air gaps, \( R = 0.85 + 2 \times 0.92 + 3 \times 0.1875 \times 0.17 = 2.79 \)

### 3.6.2 Spacers and Their Influence on Heat Loss

The development of double glazing is discussed in Section I of Chapter 10. Two sheets of glass are separated by spacer bars at their perimeter to create an air gap between the sheets. The standard spacer was a rectangular section aluminum pipe containing a desiccant. Heat flow is considerably higher close to the spacers than in the center of the glass.

The standard spacer in double or triple glazed windows is aluminum, whose thermal conductivity is one of the highest. Replacing it with material having lower thermal conductivity is called using WES (warm edge spacers). The improvement may be small, such as using stainless steel, but is it still called SES [glassmagazine.com/article/commercial/defining-warm-edge-commercial-market].

The R-factors given for double and triple glazing are considered “center-of-glass” figures. The edge-of-glass numbers will be considerably lower. The edge-of-glass effect is from the outside of the spacers to about 6 cm (2.5”) into the glass. Typically, the frame covers the spacers, but does not cover the 6 cm of the edge-of-glass effect.

The old European standard DIN 4108:1981 determined the thermal conductivity of a window as

\[
K_w = A_g U_g + A_f U_f + A_g + A_f(3.9)
\]

where \( A_g \) is the visible area of the glass in \( m^2 \), \( U_g \) is the center-of-glass conductance in \( W/m^2 \cdot K \), \( A_f \) is the visible area of the frame in \( m^2 \), and \( U_f \) is the conductance of the frame in \( W/m^2 \cdot K \). This does not take into account the thermal bridging caused by the spacer(s) in the window.

The new European standard EN ISO 10077 takes into account the spacer, and defines the window U-value \( U_w \) by

\[
U_w = A_g U_g + A_f U_f + l g \psi g A_g + A_f(3.10)
\]
where \( l_g \) is the perimeter of the glass in m, \( \psi_g \) is the linear thermal transmittance due to the combined effect of glazing, spacer, and frame in W/m·K.

Table 3.7: Default Values for \( \psi_g \)

<table>
<thead>
<tr>
<th>frame</th>
<th>metal/nonE</th>
<th>metal/lowE</th>
<th>WES/nonE</th>
<th>WES/lowE</th>
</tr>
</thead>
<tbody>
<tr>
<td>wood/plastic</td>
<td>0.06</td>
<td>0.08</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>TB metal</td>
<td>0.08</td>
<td>0.11</td>
<td>0.06</td>
<td>0.08</td>
</tr>
<tr>
<td>NTM metal</td>
<td>0.02</td>
<td>0.05</td>
<td>0.01</td>
<td>0.04</td>
</tr>
</tbody>
</table>

EN ISO 10077 gives default values for \( \psi_g \) as shown in Table 3.7. In the first column the designations TB and NTB stand for thermal break and non-thermal break, respectively. It is difficult to calculate the actual value for \( \psi_g \), so it is common to rely on tables such as that given here.

An example produced by Eckelt Glass [www.glassolutions.at/en/downloads/produkte/warm_surfaces_warm_edges.pdf] is as follows. Consider a window \( 1.23 \times 1.48 = 1.820 \text{ m}^2 \) and whose frame with \( U_f = 1.4 \text{ W/m}^2\cdot\text{K} \) is 0.11 m deep, so the exposed glass area is \( 1.01 \times 1.26 = 1.273 \text{ m}^2 \) and the glass perimeter is \( 2(1.01 + 1.26) = 4.54 \text{ m} \). Assume \( \psi_g = 0.08 \), a typical number; then according to the old formula of Equation 3.9 and the new formula of Equation 3.10,

\[
K_w = 1.273 \times 1.1 + 0.55 \times 1.41.27 + 0.55 = 1.192 \tag{3.11}
\]

and

\[
U_w = 1.273 \times 1.1 + 0.55 \times 1.4 + 4.54 \times 0.081.27 + 0.55 = 1.392 \tag{3.12}
\]

Assuming that these calculations are typical, then the spacer causes substantial thermal bridging leading to considerable heat loss. The strategy should be to reduce the total perimeter length of the windows while having an equivalent total area of window. This can be done by having larger windows but less of them, and also by choosing windows whose aspect ratio is close to square.

### 3.7 Heat Losses from a Residence

We calculate a composite wall resistance as the sum of the resistances of the individual components. For the situation illustrated in 3.12, \( R = R_i + R_1 + R_2 + R_3 + R_e \). The overall
thermal conductance is the reciprocal of this overall resistance, a single number. Calculating
the overall thermal conductance from the individual conductances involves

\[ C = \frac{1}{1C_1 + 1C_2 + 1C_3 + 1Ce} \] (3.13)

This could add unnecessary confusion and would add a lot more work.

Suppose we have a wall with one section of conductance \( C_1 \) (resistance \( R_1 = 1/C_1 \)) over area \( A_1 \) and another section with conductance \( C_2 \) (resistance \( R_2 = 1/C_2 \)) with area \( A_2 \). Then the heat loss through the complete wall is

\[ H = (A_1C_1 + A_2C_2)\Delta T \] (3.14)

where \( \Delta T \) is the temperature difference from the interior to the environment.

The IECC code and other codes consider a resistance number such as R20 for a stud wall to mean the composite resistance \( R_c \) of the wall, insulated cavity, and stud, so the composite conductance is \( C_c = 0.05 \).

The composite conductance of a wall with two sections of individual conductances \( C_1 \) and \( C_2 \) with areas \( A_1 \) and \( A_2 \), respectively, is

\[ C_c = A_1C_1 + A_2C_2A_1 + A_2 \] (3.15)

Take the reciprocal of this composite conductance to find the composite resistance.

The lessons we can draw from this section are:
* use resistivities, not conductivities.
* calculate resistance per building section, from inside to outside, then convert to conductances.
* calculate heat loss from a wall as the sum of the products of the individual conductances times each area, finally multiplied by the temperature differential.
* if the effective resistance of a wall is needed, calculate the effective conductance and take its reciprocal.

### 3.7.1 Heat Losses through Walls and Roofs

Consider a 2x6 stud wall as shown in Figure 4.2 with 5.5” of fiberglass at R3.3, 1/2” of gypsum wallboard at r0.85 on the inside, and 1/2” plywood at r1.25 dressed with vinyl siding at R0.61 on the outside. The interior surface resistance of \( R_{si} = 0.68 \) will be adopted for the rest of this chapter. Assuming a 7.5-mph wind speed, it is reasonable to use the value of \( R_{se} = \)
0.25 for the exterior surface resistance as specified in *ASHRAE Fundamentals*. With the resistivity of the fiberglass at R3.3, the wall resistance through the fiberglass is

![Figure 3.14: 2x6 Stud Wall](image)

\[ R_I = 0.68 + 0.5 \times 0.85 + 5.5 \times 3.3 + 0.5 \times 1.25 + 0.61 + 0.25 = 20.74 \text{ (3.16)} \]

and the resistance through the studs is

\[ R_S = 0.68 + 0.5 \times 0.85 + 5.5 \times 1.25 + 0.5 \times 1.25 + 0.61 + 0.25 = 9.47 \text{ (3.17)} \]

The standard wall has a double sill on top of the plywood, and a wood floor 0.75” thick is laid up to the wall, so the height of the sills exposed to the interior is 3−0.75 = 2.25”. Similarly, with a double plate on the top of the wall and 0.5” gypsum board on the ceiling, the exposed plate thickness is 2.5”. The standard spacing of the studs on a 2x6 stud wall is 24”, but the average spacing is less than this due to the stud-work around windows and exterior doors, and multiple stud column to support a major bearing beam. An average spacing of 20” is reasonable. If the 2x6 stud wall is 96” in height, 4.75” passes through the sills and plates, then

\[ (96-4.75)96(20-1.5)20=0.879 \text{ (3.18)} \]

is the fraction of the wall that is insulated, leaving 0.121 to pass through the studs, so

\[ C_c=0.87920.74+0.1219.47=0.0552 \text{ (3.19)} \]

and the wall does not pass code.
The IECC building code and other codes may require a component, wall or ceiling, to satisfy one of two measures. For example, the requirement R20 or R13+R5 means either $R_c = 20$, or the cavity wall is to have $R_c = 13$ through the cavity wall insulation and then be wrapped, either inside or outside, with continuous insulation whose resistance is at least R5. We have already established how to handle the R20 requirement. If the wall is wrapped with foam insulation with resistance R5, then the resistances through the insulated cavity and through the studs are

$$RI = 0.68 + 0.5 \times 0.85 + 5.5 \times 3.3 + 5 + 0.5 \times 1.25 + 0.61 + 0.25 = 25.74 \text{(3.20)}$$

and

$$RS = 0.68 + 0.5 \times 0.85 + 5.5 \times 1.25 + 5 + 0.5 \times 1.25 + 0.61 + 0.25 = 14.47 \text{(3.21)}$$

respectively. With the same area ratios as before,

$$Cc = 0.87925.74 + 0.12114.47 = 0.0429 \text{(3.22)}$$

and now the wall passes code.

![Diagram of 2x4 Stud Wall](image)

If the wall with the R value 13+5 has 2x4 studs at 16” on center as shown in Figure 3.15, but the average spacing is closer to 13.33”, through the fiberglass with R3.3,

$$RS = 0.68 + 0.5 \times 0.85 + 3.5 \times 3.3 + 0.5 \times 1.25 + 0.61 + 0.25 = 14.14 \text{(3.23)}$$
and through the studs,

\[ R_S = 0.68 + 0.5 \times 0.85 + 3.5 \times 1.25 + 0.5 \times 1.25 + 0.61 + 0.25 = 6.97. \quad (3.24) \]

The fraction of the wall through the insulation is

\[ \frac{(96-4.75)96(13.33-1.5)13.33}{13.33} = 0.843, \quad (3.25) \]

leaving 0.157 to pass through the studs, so

\[ C_{eff} = 0.84314.14 + 0.1576.97 = 0.0822. \quad (3.26) \]

Replacing the 2x6 studs with 2x8 studs, which are 7.25” wide, then the wall resistance through the insulation is

\[ R_I = 0.68 + 0.5 \times 0.85 + 7.25 \times 3.3 + 0.5 \times 1.25 + 0.61 + 0.25 = 26.52 \quad (3.27) \]

and the wall resistance through a stud is

\[ R_S = 0.68 + 0.5 \times 0.85 + 7.25 \times 1.25 + 0.5 \times 1.25 + 0.61 + 0.25 = 11.66. \quad (3.28) \]

The nominal spacing of the studs will be 24”, and assuming average spacing is 20”, so the fraction of the wall with insulation is 0.879, then

\[ C_{eff} = 0.87926.52 + 0.12111.66 = 0.0442. \quad (3.29) \]

and the wall passes code.

As stated previously, evolving building codes continue to require higher component resistances, making it increasingly difficult to satisfy them in a framed cavity wall with fiberglass insulation. Expanded closed cell foam, which is sprayed through a nozzle directly into the wall, is increasingly popular, since its resistivity is 6. A problem with foam is the inflexibility this presents for future electrical chases in the exterior wall, but the answer to this is to strap the sidewall after insulating, and add a polyethylene vapor barrier inside as shown in Figure 3.16: the overall resistance of this wall is analyzed in this section.

The resistances for typical wall and roof sections are discussed in this section. It is assumed that the surface area exposed to the outside is very large compared to the fringe areas — the areas at the sides and ends abutting other areas; thus, fringe effects can be ignored.

Repeating the last example with polyurethane foam insulation with resistivity of 6 BTU/h·F, the resistance through the foam is

\[ R = 0.68 + 0.5 \times 0.85 + 3.5 \times 6 + 0.5 \times 1.25 + 0.61 + 0.25 = 23.59. \]
Repeating the last example, but with 2x6 studs, produces resistance through the fiberglass insulation of $R = 0.68 + 0.5 \times 0.85 + 5.5 \times 3.2 + 0.5 \times 1.25 + 0.61 + 0.25 = 20.19$, and through the stud of $R = 0.68 + 0.5 \times 0.85 + 5.5 \times 1.25 + 0.5 \times 1.25 + 0.61 + 0.25 = 9.46$.

Replacing the fiberglass insulation in the 2x6 stud frame wall with polyurethane foam insulation produces resistance $R = 0.68 + 0.5 \times 0.85 + 5.5 \times 6 + 0.5 \times 1.25 + 0.61 + 0.25 = 35.59$.

Next, repeating the last example, but with 2x8 studs, produces resistance through the fiberglass insulation of $R = 0.68 + 0.5 \times 0.85 + 7.25 \times 3.2 + 0.5 \times 1.25 + 0.61 + 0.25 = 25.79$, and through the stud of $R = 0.68 + 0.5 \times 0.85 + 7.25 \times 1.25 + 0.5 \times 1.25 + 0.61 + 0.25 = 11.65$.

Replacing the fiberglass insulation in the 2x8 stud frame wall with polyurethane foam insulation produces resistance $R = 0.68 + 0.5 \times 0.85 + 7.25 \times 6 + 0.5 \times 1.25 + 0.61 + 0.25 = 46.09$.

2x6 sidewalls are architecturally acceptable: The visible depth of the frame of a window or the style of a door looks fine. 2x8 sidewalls make these visible depths look like caverns — perhaps getting used to them will make them acceptable with time. There is no way a 2x10 sidewall can be made acceptable. However, the depth of a rafter is not evident after construction is completed, so 2x12s or 2x14s set 24” o/c are fine.

A typical roof section using 2x10 rafters is shown in Figure 3.17. Assume the interior surface resistance $r_{si} = 0.62$. With fiberglass insulation the resistance through the insulation is $R = 0.62 + 0.5 \times 0.85 + 9.25 \times 3.2 + 0.5 \times 1.25 + 0.61 + 0.25 = 32.12$. 

![Figure 3.16: Sidewall with Strapping](image)
Figure 3.17: Typical Roof Section

and the resistance through a rafter is \( R = 0.62 + 0.5 \times 0.85 + 9.25 \times 1.25 + 0.625 \times 1.25 + 0.44 + 0.25 = 14.08 \).

Replacing the fiberglass with polyurethane foam produces resistance through the foam of \( R = 0.62 + 0.5 \times 0.85 + 9.25 \times 6 + 0.5 \times 1.25 + 0.61 + 0.25 = 58.02 \).

Replacing the 2x10 rafters with 2x12, and with fiberglass insulation, the resistance through the insulation is \( R = 0.62 + 0.5 \times 0.85 + 11.25 \times 3.2 + 0.5 \times 1.25 + 0.61 + 0.25 = 38.52 \), and the resistance through a rafter is \( R = 0.62 + 0.5 \times 0.85 + 11.25 \times 1.25 + 0.625 \times 1.25 + 0.44 + 0.25 = 16.58 \).

Replacing the fiberglass with polyurethane foam produces resistance through the foam of \( R = 0.62 + 0.5 \times 0.85 + 11.25 \times 6 + 0.5 \times 1.25 + 0.61 + 0.25 = 70.02 \).

### 3.7.2 Heat Losses through Structural Integrated Panels

A modern alternative to the traditional stud wall is the SIP (structural integrated panel) with a foam core sandwiched between structural facings. The facings are OSB (oriented strand board), a modern version of particle board but with superior glues. The older particle board would decompose when soaked, but this does not happen to OSB.

SIP panels can be quite large, requiring a crane to be used during construction. When properly prepared, the time for constructing the frame can be considerably less than for a stick built frame. The advantage of a SIP building is tighter construction, and so lower air infiltration, and higher R values.
The R values for SIP walls depend on their thickness as follows: R16 for 4.5”, R24 for 6.5”, R32 for 8.25”, R40 for 19.25”, R48 for 12.25”. Notice that these thicknesses are the same as for a 2x4, 2x6, 2x8, 2x10, and 2x12 stud wall with 1/2” wallboard on the inside and 1/2” plywood on the outside. Assuming the resistivity of the OSB is the same as plywood, then the resistance contribution of the two layers of OSB is 1.25, so the resistivity of the foam is about 4.2. This is a little better than open-cell, air filled polyurethane foam at 3.7, but considerably less than gas-filled closed cell polyurethane foam at 6.

Wiring chases are built into the SIP panels at assigned points. This does not afford the flexibility of a stick built structure, where, even if expanding foam insulation is used, electrical wiring in an exterior wall is always installed before the foam is blown in. The possibility of future wiring needs can be addressed by installing conduits at the same time as the rough wiring. There is typically no need to have plumbing in an exterior wall, and the SIP panels are only constructed on exterior walls where insulation is needed.

### 3.7.3 Heat Losses from Basements and Foundations

Shown in Figure 3.18 is a sectional elevation of a poured concrete basement wall, sitting on a concrete footing, and supporting the stud wall of the structure above. Also shown are arrowed lines indicating heat flow from the interior to the exterior. The greatest winter-time loss per unit area from the basement wall is in the aboveground area exposed to the elements. Notice that the heat flow lines below ground are drawn as arcs of a circle. The heat: heat flow from a basement loss per unit area is an inverse function of the length of the arc.

<table>
<thead>
<tr>
<th>depth</th>
<th>path length</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>0→1</td>
<td>0.68</td>
<td>2.44</td>
</tr>
<tr>
<td>1→2</td>
<td>2.27</td>
<td>4.50</td>
</tr>
<tr>
<td>2→3</td>
<td>3.88</td>
<td>6.45</td>
</tr>
<tr>
<td>3→4</td>
<td>5.52</td>
<td>8.40</td>
</tr>
<tr>
<td>4→5</td>
<td>7.05</td>
<td>10.42</td>
</tr>
<tr>
<td>5→6</td>
<td>8.65</td>
<td>12.66</td>
</tr>
<tr>
<td>6→7</td>
<td>10.28</td>
<td>14.49</td>
</tr>
</tbody>
</table>

The thermal path lengths, in feet, and resistances in h·F/BTU of a typical soil based on depth below the surface are given in Table 3.8 [J.K. Latta and G.G. Boileau, “Heat Losses From House Basements,” Canadian Building, vol XIX #10, October 1969]. As can be seen, soil depth $x$ is a strong factor in determining heat losses.

The thermal resistance follows the linear law $x$
where $x$ is measured in feet.
The thermal conductivity of concrete is 12 BTU/ft²·F per inch, and the thermal conductivity of earth can vary from 0.735 BTU/ft²·F to 6.55 BTU/ft²·F per inch depending on moisture content and compactness. Further, the ground shown in Figure 3.18 slopes away from the foundation, but the corresponding diagram in the Latta/Boileau article shows the ground is flat. The empirical nature of such heat losses, and the uncertainties that result, are evident from this simple analysis.

Adding insulation to a basement wall will significantly reduce heat losses. For insulation with a conductivity of 0.24 BTU/ft²·F, the overall conductivity through the insulation and the concrete wall is as shown in Figure 3.19 as a function of depth below ground, with no insulation and 1”, 2”, and 3” of insulation.
Next consider the heat loss from a 10” thick basement wall of concrete; In North America the standard thickness of a poured concrete foundation wall is 10”. Assuming the same surface resistances as for glass and a 7.5-mph wind, the total resistance is $R = 0.93 + 10 \times 0.08 = 1.73$. This will produce an unacceptable heat loss in winter, so to reduce this we use 2” of closed cell foam board on the outside as shown in Figure 3.20. The resistance becomes $R = 0.93 + 10 \times 0.08 + 2 \times 6 = 13.73$, an acceptable value.

If the concrete was buried underground with an average berm of 24” as shown in Figure 3.21, the resistance would be $R = 0.93 + 10 \times 0.08 + 24 \times r = 1.73 + 24 \times r$, where $r$ is the resistivity of the earth that ranges from 1.4 for dry earth to 0.15 for damp tightly packed earth. Assuming an intermediate figure of $R = 0.4$ that approximates damp loose earth, the resistance becomes $R = 1.73 + 24 \times 0.4 = 11.33$.

Adding 2” of urethane foam to the prior example, as shown in Figure 3.21, gives $R = 1.73 + 2 \times 6 + 24 \times 0.4 = 23.33$. 
Chapter 4

Residential Construction Techniques

Architecture reflects the climate it is applied in. An adobe house looks appropriate and works well in Santa Fe, New Mexico, but would be totally out of place in Illinois or Sweden. A wood frame, garrison colonial looks fine in New England, but would be ridiculous in southern Florida or Sicily.

Housing developments in the United States and Britain that followed the end of World War II discarded some of the lessons learned by earlier generations. Orientation to the Sun was rejected in favor of more housing units per unit area. We should know better.

In a mediterranean climate the Sun is rejected. That is, windows are designed to take in light, but not the Sun’s heat. Exterior walls are painted white to reflect rather than absorb solar energy. White walls absorb much less solar energy than dark walls, while re-radiated infrared energy is a function of the emissivity of the surface, and this emissivity is largely unaffected by the color of that surface.

Urban streets in southern Italy are designed to exclude direct sunlight. It goes so far as to consider a suntan undesirable. When turning a corner in a city so that the Sun is on them, people cross the road to walk in shade. Contrast this to northern Europe and North America, where tanning oneself is common, either by lying on a beach or going to a tanning salon.

In a predominantly cooling climate one is concerned with reducing heat gain, including solar heat gain. The Sun is to be excluded, and air flow around the residence is encouraged.

Most of Europe and North America have predominantly heating climates, and capturing solar energy rather than rejecting it is the objective. Solar heat gain is welcomed in winter, and the predominance of this is through glass, either passively or to solar thermal collectors, or indirectly with photovoltaic electric generation. Air flow through the skin of the residence is not welcomed since it contributes to heat loss. Thus, it is important to know how to keep heat in with effective building elements, and that is the subject of this chapter.
Thus, the structure itself should be designed for its climate. Foundations must be insulated in heating climates, whereas in cooling climates the earth can help cool the foundation walls and so cool the house. Interior courtyards, shielded from the Sun, are common in cooling climates. Conservatories to welcome the Sun are common in heating climates.

The major heat loss from a structure per unit area is through windows and exterior doors. Further, air infiltration is a major source of heat loss in the United States, but much less so in northern Europe with tighter building standards. The air leakages are highest with exterior doors and operable windows.

This chapter discusses building materials and techniques, including some that are relatively recent advances. Houses made of factory manufactured panels, such as SIPs discussed in Section 2.14, are now important building elements. AAC block walls have become an important part of the construction industry in Europe. Insulation materials are changing from fiberglass batts to closed cell polyurethane foam. Spurred by the need for increased energy efficiency in buildings, expect additional advances in the technology of buildings into the future.

Towards the end of this chapter a number of national and international agencies with an alphabet soup of acronyms are cited, most of which are listed here for convenience:

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASHRAE</td>
<td>American Society of Heating, Refrigerating and Air Conditioning Engineers</td>
</tr>
<tr>
<td>ASTM</td>
<td>American Society for Testing Materials</td>
</tr>
<tr>
<td>DOE</td>
<td>U.S. Department of Energy</td>
</tr>
<tr>
<td>EIA</td>
<td>U.S. Energy Information Administration</td>
</tr>
<tr>
<td>GBPN</td>
<td>Global Buildings Performance Network</td>
</tr>
<tr>
<td>IEA</td>
<td>International Energy Agency</td>
</tr>
<tr>
<td>IECC</td>
<td>International Energy Conservation Code</td>
</tr>
<tr>
<td>IESNA</td>
<td>Illuminating Engineering Society of North America</td>
</tr>
<tr>
<td>ICER</td>
<td>International Confederation of Energy Regulators</td>
</tr>
<tr>
<td>IRC</td>
<td>International Residential Code</td>
</tr>
<tr>
<td>ISO</td>
<td>International Organization for Standardization</td>
</tr>
<tr>
<td>NIST</td>
<td>National Institute of Standards and Technology</td>
</tr>
<tr>
<td>NFPA</td>
<td>National Fire Protection Association</td>
</tr>
<tr>
<td>NREL</td>
<td>National Renewable Energy Laboratory</td>
</tr>
<tr>
<td>NOAA</td>
<td>National Oceanic and Atmospheric Administration</td>
</tr>
</tbody>
</table>
4.1 Fire, Civilization, and Simple Dwellings

“New research lends credence to the theory that humans’ early ancestors were using fire long before Homo sapiens stepped on the scene 200,000 years ago, according to a study published this month in the *Journal of Human Evolution*. Scientists have found evidence of controlled fire use inside an ancient cave in Israel that was home to several lineages of prehistoric hominins, the predecessors of modern humans. Artifacts discovered at the site, located in the Nahal Me’arot Nature Reserve in northern Israel, suggest that fire use became routine among hominins around 350,000 years ago [P. Ross, “When Did Man Discover Fire? Ancestors Of Modern Humans Used Fire 350,000 Years Ago, New Study Suggests,” www.ibtimes.com/when-did-man-discover-fire-ancestors-modern-humans-used-fire-350000-years-ago-new].”

Fire provides warmth in winter, but its greatest advantage was in the cooking of meat. Cooked meat is far easier to digest than raw meat. In contrast, meat-eating animals spend most of their time in the wild seeking prey, then eating and digesting it. One unwelcome consequence of eating cooked food is that we have no need for our appendix, with the result that it sometimes atrophies, becomes toxic, and must be removed surgically.

The hunter/gatherer lifestyle was essentially nomadic — they followed the herds, and knew where and when certain fruits would ripen.

The nomadic Indians of North America constructed portable dwelling units — wigwams, a frame of flexible tree branches arranged in a circle on the ground and tied together at the top. The frame was then covered with bark or animal skins. Typically it was no bigger than 5 meters in diameter at the base, and could house more than one family [www.ducksters.com/history/native_american_homes].

The yurt as a dwelling unit dates back several thousand years to the nomads in the steppes of Central Asia, in particular Mongolia and Turkestan. It is a circular structure with vertical side walls and a conical roof. The frame is made from saplings covered with felted mats made from the wool of the sheep that were an essential part of the nomads’ survival. The poles in the Turkic yurt are bent to form the side wall and the roof, but in the Mongolian yurts the sidewall and roof poles are separate, so the roof poles are straight. These structures are not crude. They often have considerable ornamentation. They provided warmth and shelter in winter when additional mats were added. In the hot summer, sections could be raised so air could flow throughout.

Yurts have become popular in the United States. A number of companies manufacture a modern version of the yurt, with steel cables and architectural fabrics [www.yurtinfo.org/yurtstory-the-history-of-yurts-ancient-and-modern].
The end of the ice age, about 8,000 BC, was followed by the evolution of humans from hunter/gatherers to farmers. The enabling mechanism was fire, reducing digestion of meat from an almost full time chore, creating free time to try new things such as farming. Farming started in the Middle East, and then entered and spread in Europe by about 4,000 BC.

No longer nomadic, these farmers had every incentive to develop substantial dwellings for comfort and security. The first houses were mud brick: mud formed into blocks and air dried [www.localhistories.org/houses]. Roofs were made from leaves, reeds, or grasses.

The Celts built small round houses with a central support pole. Benches were built around the walls, and at night they were used as beds. Roofs were thatched; see Section 2.12.

English peasants in the Middle Ages lived in simple huts with wooden frames on which wooden slats were woven and plastered with mud reinforced with animal hair. The floor was mud, usually covered with straw. By the early seventeenth century, houses, even for the poor, were built of stone or fired bricks. Soon most houses had fireplaces and chimneys. Glass, once a luxury, became commonplace, and operable windows were the standard, first as hinged casement windows and then as sliding sash windows. Houses for the poor were small, with one, two, or three rooms [www.localhistories.org/homes].

Farming substantially increased the food supply over that of the hunter/gatherers. This led to a great increase in population, and the spreading of that population into previously sparsely populated areas. The population of the world over recorded history is given in Figure 4.1; the numbers are in millions. Notice the logarithmic scale of the ordinate. There are two surges, the first a little after 5000 BC, coinciding with the predominance of farming over hunting/gathering. The second was due to the industrial revolution and the enormous increase in productivity that resulted.
The industrial revolution, which began in northern England in the eighteenth century, drew young people from their farms to newly formed cities. Before this revolution, most of the population of the world toiled on the land. Now, in Europe and North America, few work the land, and how did this become possible when in the seventeenth century, with over 90% working the farms, we could hardly feed ourselves? The answer is mechanization — the farm tractor, the combine harvester, and other equipment, and a transportation infrastructure of roads, railroads, and canals.

Shanty towns spring up around major cities, and are difficult or impossible for a government to regulate. The slums in Brazil are called favelas. Until recent times favelas were constructed in the far suburbs — as far as the city dwellers were concerned, it was out of sight, out of mind. This changed in the 1970s when a massive rural exodus occurred and people were drawn to the cities. Now about 6% of the population of Brazil, over 11 million people, live in favelas.

There are slums in every country in the world. None are desirable places to live, but they are a place to live. Some are better than others. Some have municipal services, of which the most important are clean water and sewage disposal.

A shanty can be built in a day or two, and the authorities can do little to stop this. Roofs are often constructed out of corrugated metal or plastic. Thus, the roof is light and transportable and is often the most valuable part of the structure. If evicted, they take the roof with them.

“The poor you will always have with you” is a quote from the Bible. There are poor people in every country in the world. By necessity, the poor are good scavengers. They gather materials
that others ignore or discard.

The most pernicious of all dwellings for the poor is the high rise. The Pruitt Igoe housing project was constructed in Saint Louis, Missouri in 1956, with 233 buildings, each with 11 storeys. It was demolished in the mid 1970s. Almost immediately after its construction it become crime and blight ridden. High-rise apartments and condos work well for the middle class and the wealthy, so why does it fail the poor? The answer is in the aspirations that poor people have for a dwelling unit, at least in the United States. These aspirations include privacy, so that when the front door is closed the outside world is excluded.

4.2 Construction Techniques for Exterior Walls

4.2.1 Wooden Walls

Wood is a universally popular construction element for furniture and for buildings. Old growth wood harvested in New England was denser and much stronger than second growth wood. Eighteenth and nineteenth century furniture is desirable not only for its antique value but also for the quality of the wood. Quarter-sawn oak with distinct medullary rays is highly prized. Modern oak furniture is made from trees whose growth rate was much faster than that from a native forest.

Wood is an unsatisfactory material for a dwelling if it is in direct contact with the ground, even if is treated: Wood rot occurs in damp and wet wood, not dry wood. Treated fence posts are acceptable, even though their service life is 10–15 years. For a structure, more durability is required. Wood on top of a masonry foundation can have an unlimited life time.

Whole logs can be used to construct a cordwood house. Introduced to the United States by Scandinavian immigrants, a single room structure about 3 m wide and 4 to 6 m long, the log cabin proved to be durable. Preparing the logs involves removing the bark and sorting the logs by size. Since logs taper, the log chosen on a partially constructed wall should be oriented so its smaller diameter is above the larger diameter of the log below. In the four walls there needs to be at least two openings, one for a door and the rest for windows. Few if any nails are used in constructing the walls since the logs on the two sides making a corner are notched and overlapped. The foundation can be as simple as a pier at each of the four corners, with the result that the floor often remains earthen. Better still is enough piers to support floor joists, on top of which the logs rest. Better again is a full-perimeter foundation of stone or concrete. If corner piers are used and the floor is earthen, and a small masonry chimney can be built on the dirt.

Now, log cabins can be very sophisticated, with two or more storeys and many rooms. Rather than use natural logs, with their uneven sizes and round profiles, some manufacturers offer
precision-sawn timbers that neatly fit against each other with a tongue and groove type of joint, but this loses some of the charm of a natural log cabin.

The gaps between adjacent logs are filled with wood chips and mud. The sidewalls are heavy, but the roof typically is light, with sawn lumber or corrugated metal. A desirable element of the roof is large overhangs so that the sidewalls are somewhat protected from the elements. More primitive log cabins would not have had operable windows. The high infiltration rate makes natural ventilation more than adequate for human respiration, but with the undesirable side effect of making heating in colder climates a problem.

Wood for construction or cooking fires is a limited resource. For example, there is concern for the denuding of sections of the rain forest in Brazil. The rain forest is possibly the most diverse natural environment in the world, but that diversity is threatened by illegal logging and clearcutting to make fields to raise cattle.

The island of Hispaniola is shared by Haiti and the Dominican Republic. The Haitian side of the island has essentially no trees left, whereas on the Dominican side the forests are reasonably healthy. Erosion is endemic on the Haitian side, but not on the Dominican side. It can be argued that Haiti is a failed society, like a number of others.

In Europe the population density is such that wood is no longer a dominant resource for fires, but it remains an important component in residential construction, particularly for roofs. A house will last for generations, a cooking fire for a few hours. A wooden rafter could provide fuel for a cooking fire, and will be gone in a week. Wood will remain a valuable resource, but not by providing heat from combustion.

There are other parts of the world where people have essentially wiped out forests and woodlands, and this is most severe in third world countries. Unfortunately, the alternative is a fossil fuel.

Referring back to Figure 4.1, we conclude that if we had the same world population as we had in the middle ages, we could burn fossil fuels at the same rate per capita as we now do with no damage to the planet.

### 4.2.2 The Development of Plywood

Until very recently all wooden walls were constructed from dimensional lumber (2 × 4’ though 2 × 12’) or larger lumber beams and columns in a post and beam house. The frame of a house constructed in the early part of the twentieth century was sheathed with individual boards. Since the 1940s the sheathing of choice has been plywood.
The technique of gluing thin layers of wood together dates back to the pharaohs. Through the ages, laminated wood was used for furniture and other household items. Although some patents were awarded in the nineteenth century, little was done with plywood until 1905, when, as part of an exhibit for the World Fair, the Portland Manufacturing Company used paint brushes to glue three thin layers of softwoods together. The product created enough interest for the company to develop an automatic glue spreader. Early sales were predominantly for door panels. The problem with these panels was that the glues were not water resistant, and the fragmented group of small manufacturers in the plywood business at that time did not have the resources to develop a more durable product [www.apawood.org/apas-history].

This changed in 1933 when the Douglas Fir Plywood Association was formed. Durable glues were developed in 1934 by James Nevin, a chemist working for the Harbor Plywood Corporation in Aberdeen, Washington [op cit]. Plywood with designations such as CDX, where the X means suitable for outside use, soon appeared to become the standard product for sheathing a wood frame wall or covering the wooden rafters. The association changed its name to APA (American Plywood Association) in 1964.

The APA stamp is on most plywood sold in the United States. A cheaper product was introduced in the 1970s, called mill certified plywood. It had fewer laminations and was subject to delamination and buckling when wetted. A leaky roof sheathed with mill certified plywood would need to be removed down to the rafters and replaced. Another feature of this defective product was that it was not dimensionally true — a side could be as much as 1/2” off the standard 4’ or 8’. A number of builders who thought they were saving money learned to rue the day they chose mill certified.

Particle board became popular in the 1970s, but even though its glues were water resistant the product lost structural integrity when soaked in water. Like the mill certified product, using particle board in a location that could be wetted was asking for trouble.

OSB (oriented strand board) was introduced in the 1980s, a big improvement over particle board. 4’×8’ OSB panels are inexpensive, less than half the cost of corresponding plywood panels.

Factory-made modular homes constructed out of dimensional lumber, and delivered to the site on which a foundation was ready and waiting, became an alternative to the stick-built house. Usually the selling point for this type of construction was more about the architecture than the quality, although these homes were quality products. Now the factory-made wooden wall of choice is the SIP; see Section 2.14. Both sides of a SIP panel are OSB.

4.2.3 Bamboo and Other Grasses
In some parts of the world, wood may not be available at an affordable cost, and the climate would not permit adobe or earthen walls, so what is the alternative? The answer may be bamboo. Bamboo is not a wood: it is in the grass family. There are over a thousand varieties of bamboo. In compression it is stronger than wood, and in tension it has about half the strength of steel.

The best places to grow bamboo are warm, damp regions with relatively high humidity, where it can grow at a rate exceeding one meter per day. Southeast Asia is considered the best place for bamboo to grow, but it thrives in many other locations. The American Southeast was once covered with bamboo, called “cane break.” Cane break lost out to food crops.

The stems of the bamboo, known as culms, emerge from the ground at full size and grow to essentially full height before branching out; this growth occurs in one season, and the culm can be as tall as 30 meters and have a diameter up to 0.2 meters. In the second season the culm hardens, and further hardens in the third season. The timber bamboo is then harvested. If left alone, a fungal growth appears on the culm in the second through fifth season, which finally kills the bamboo in its fifth through eighth season.

In the United States, bamboo is a popular rival to hardwood flooring: it is durable, has a pleasing natural pattern, and is affordable. However, rarely if ever do you see bamboo used in framing a U.S. house, although this is common in Asia and South America.

What is interesting is the fact that it seems that natural materials are more durable than manufactured products in the right circumstances. In Section VII of Chapter 9 we discuss Ian McHarg’s analysis of the durability of dune grasses over concrete bulwarks to survive coastal storms. A similar argument can be made of bamboo structures over more conventional ones.

“A 7.5 earthquake in Limón, Costa Rica, in April 1991 destroyed homes built with concrete and rebar, but all 20 of the more-flexible bamboo houses at the earthquake’s epicenter remained standing. When three typhoons swept into the Cook Islands in 2005, one producing winds of 173 mph, they devoured everything in their path — everything, that is, except a group of bamboo houses on the beach [E. Best, “Bamboo Houses to the Rescue,” www.psmag.com/environment/bamboo-houses-to-the-rescue-16347, July 2010].”

Field grass and swamp reeds are widely used in third world construction, mostly on side walls. When dry they are reasonably durable. For the poor of the world in earthquake-prone regions, they are preferable to heavy masonry materials.

There are many uses for grasses other than bamboo. Fields of ordinary grass can be harvested as hay, which is fed to cattle. Grass can also be composted into silage, which is fed to cattle in the winter months when fields are unproductive. Similarly, a combine harvester can separate the seeds of oats, wheat, or barley and automatically create hay bales from the green stalks. Hay can also be harvested from legumes such as alfalfa and clover. Hay bales can be lined up and stacked to create a wall; leaving out a bale when building that wall creates a window area.
A use for grass that survives into modern times is the thatch roof. Durable and functional, it keeps the interior warm in winter and cool in summer. It is also architecturally attractive. Early dwellings had walls that could not carry a heavy load, and thatch was the lightest of all roofing materials available at that time.

Thatch is dried straw, grasses, or reeds. To make a thatch roof, these are gathered into bundles, and laid over wooden supports, and held in place with wooden pins. An upper layer is laid on top, and along the ridge a reinforcing layer is added.

The first thatch dwellings were all thatch: these were located in north central Europe. The thatch dwelling was built like a tent, with a ridge pole that supported the top of the roof while the bottom was on the ground. An open fireplace provided warmth and light in winter. It also sent out creosote to preserve the structure — of course, it did serious damage to the lungs of the occupants, particularly children.

4.2.4 Stud Wall Construction

In North America, wood intended for construction purposes is rough sawn, dried to a specific water content, and then planed into a precise cross-sectional dimension so it can be used as studs (vertically erected structural members in a sidewall), or horizontal beams as floor joists, or sloping roof rafters. Boards have a thickness of 3/4”, and can be used for a wooden ceiling, or, if a hardwood such as oak or maple, a floor on top of plywood that is supported by joists. Europe does not follow this system, and final dimensions vary.

The plan view of a standard frame wall is shown in Figure 4.2. The studs, not shown, are typically spaced 24” on center. Since the thickness of the fiberglass insulation is 5 1/2” it implies that the studs are 2×6, while their actual size is 1 1/2” × 5 1/2”. One standard thickness in the United States for fiberglass batts is 5 1/2”.
The standard spacing of studs in the United States is 16” or 24” on centers, so full 4’×8’ plywood sheets can be used without cutting: 16×6 = 96, and 24×4 = 96. Recall, a standard size plywood sheet in Europe is 4’×8’, the same as in the United States, so the logical spacing of studs to avoid waste, and to save time by reducing the number of cuts to be made, will be in Imperial units.

The most common framing method is the box frame, sometimes called the platform frame. Joists of 2×10 or 2×12 are laid on top of a foundation or a framed wall as shown in Figure 4.3. Notice the small amount of overlap over the central support beam shown in red in the diagram. This beam is typically two or three timbers the same size as the joists and nailed together. It is
supported with posts in the area below. If this area is designed to be open it could require the use of a major supporting beam, such as a glulam or a steel I-beam.

Rather than overlap the joists, they can be held on either side of the bearing beam with joist hangers. A joist hanger is a U-shaped metal support whose internal width accommodates the width of the joist. It has flanges with perforations to nail it to the support, and other perforations on the sides of the U to nail into the joist.

Between the joists and the concrete foundation sit the sills, usually two 2×6s with staggered joints and bolted to the foundation; anchor bolts are set in the concrete after the concrete is poured and before it sets. In the United States the sills are now required to be pressure treated to inhibit rot. The ends of the joists are set back 1.5” from the outside of the foundation so that a box frame, the same size as the joists, can be nailed to the ends of the joists.

With the joists in place, plywood is laid. A good framing carpenter will have positioned the joists so the number of plywood cuts is minimized, as is wastage of time and material. The result is a solid deck from which to build the next level. The gable ends are constructed next for a two-storey house with a pitched roof. The wall is constructed flat on the deck, with metal straps holding the shoe so it does not slide out when the wall is lifted.

Another popular framing method is post and beam. Here the posts are few and the beams are major structural elements. The structure can be erected and roofed before most of the interior walls are erected. The posts and the beams are designed to be exposed, resulting in a rustic, barn-like appearance that many people like.

A different framing technique, unique to the United States, is balloon framing. Here the studs for the sidewall are two stories high. This is not a good method of construction since it presents difficulties in setting the intermediate floor. Further, some codes prohibit balloon framing. Even more damming, it is now virtually impossible to get the long studs needed.

Joists are omitted at the stairwell location. Those joists adjacent to the stairwell are typically doubled up. In the design of a house, it is the stairwell that is the most difficult to design and locate so that there is an easy flow to and from it.

Stick-built construction will allow greater air infiltration than any of the wall techniques discussed in the next sections. Care must be exercised by the framing crew. Gaps between plywood sheets should be taped on the outside or filled with expanding foam. In the United States a product called Great Stuff comes in a hand-held spray can with a thin wand that can slip into small spaces.

<table>
<thead>
<tr>
<th>Table 4.1: SIP Resistances</th>
</tr>
</thead>
<tbody>
<tr>
<td>foam</td>
</tr>
<tr>
<td>expanded polystyrene</td>
</tr>
<tr>
<td>extruded polystyrene</td>
</tr>
</tbody>
</table>
The standard framing method for windows and exterior doors is to provide a rough opening that is a little larger than the unit itself. When the unit is in place the gap around the window can be filled, a little at a time, with Great Stuff: The foam expands when it comes out of the wand, and can exert considerable pressure on the door or window frame, so it is prudent to fill the cavity in stages. The adhesion of the foam to almost any surface means that so-called cold joints do not occur.

Contrast this to adhesion with poured concrete. Fresh concrete will not adhere to concrete poured two or more hours before: if there is any delay in pouring a fresh batch of concrete against concrete beginning to set, then the old concrete is mechanically agitated to mix with the new and so create a strong bond.

A significant air infiltration rate also indicates that, under certain circumstances, water vapor can migrate through a fiberglass-insulated wall. When the temperature outside is significantly lower than that inside, dew point could be reached in the wall, resulting in water wetting the fiberglass, which could drag it down, leaving a void in the top of the wall, and possibly dripping down to the sills. Even when great care is exercised to avoid these problems, a simple and inexpensive way to ensure that it does not happen is to staple 6-mil polyethylene sheeting to the face of the interior studs after the rough plumbing and electric are completed and the wall is insulated.

### 4.2.5 Structural Integrated Panels

The frame of a house, including sheathing and insulation, can be delivered to the building site as SIPs (Structural Integrated Panels). The size of each panel is very specific, so considerable care must be taken to ensure that the framing elements on the foundation (sills, joists etc.) be the precise size. The tops of the SIPS are designed to accept the plates that hold adjacent panels together. The SIP facing panels are made of OSB discussed in Section 2.11.

SIPs use a modular format. Panels are 4’ or 8’ wide, and the thicknesses, face to face, are 4.5, 6.5, 8.25, 10.25, or 12.25”. Since the OSB panels are 1/2” thick, this means that connecting dimensional lumber can be used to connect panels together; the thickness of the foam insulation is 3.5, 5.5, 7.25, 9.25, or 11.25”.

The R values in ft²·h·°F/BTU of a SIP as a function of its thickness and the foam used is given in Table 4.1. The 4.5” panels will not satisfy code, and neither will the 6.5” panel with expanded polystyrene when one takes into account the thermal bridging of the 2×6 connecting pieces.

Plumbing should never be run though walls exposed to the elements in climates subject to winter freezing; this mandate does not apply to sill cocks, which can always be constructed with interior shutoffs. This is not the case with electrical wiring. U.S. electrical codes require
that there be no point on a wall more than 2 m from an outlet where the wall space is useable. Useable space does not include doors. Thus, SIP panels have factory-installed electrical conduits.

4.3 Masonry Walls

4.3.1 Brick Wall Construction

A brick is a block of clay/sand/lime/concrete, mixed together with water and sometimes straw as a binder and then fire-hardened in a kiln to create a very durable product, or simply air dried. The use of air-dried bricks dates back to before 7,500 BC [www.ask.com/wiki/Brick], and they are still used in more arid parts of the world, including in the adobe houses of the American Southwest. Fire-hardened bricks date back to about 2,900 BC, and these have an almost unlimited life. Early bricks were mixed by hand (and foot by driving oxen over the paste to properly mix it) and shaped by hand. Now bricks are manufactured to a precise shape and size.

A standard brick is 3.625×2.25×8, the width, height, and length in inches, so three courses with the mortar in between are 8”, and these bricks when laid occupy 4 x 2.5 x 8” = 80 cubic inches. A standard brick in India is 9x9x19, where the dimensions are centimeters. The firing process makes these bricks expensive, and often too expensive for some in third world countries.

The phrase “making bricks without straw” is meant to say that an impossible task is about to be undertaken. It dates back to Exodus 5 in the Bible. In those ancient times straw was added to the wet mud before it was cast into bricks and air dried. This was done to speed up the drying process. Egypt was a land without trees, so firing the brick was out of the question.

Fifty years ago the majority of houses built in England were brick. There was little else. Now there are many more choices, not just in England but everywhere. We divide these choices into two categories, wooden and masonry.

About 70% of the residences built in England have brick on the outside. Brick is impervious to the wet English climate. Until the second half of the twentieth century, the sidewall was two separate brick walls with an air space between, to act, not very efficiently, as an insulator. Now it is more common for the inside wall to be concrete block, which is plastered and painted. The air gap between the walls is replaced with an insulating material, which preferably also acts as a vapor barrier.

Brick veneer over a wood frame is an alternative in the northern tier in the United States. The brick is 3/4” thick and is pinned to the underlying frame by laying metal straps on the mortar
between layers of bricks as they are laid, then bending the straps vertically against the frame and nailing them to the studs of the frame. It is impossible to tell that the brick is only a veneer.

By volume, concrete is far cheaper than brick. A yard (meaning a cubic yard) of 3000 lb concrete in 2014 in New England cost $110, with a six-yard minimum when delivered by a specially designed truck with a hopper that can hold 10 to 12 yards. A cubic yard is 27 ft$^3$ or $4.6656 \times 10^4$ cubic inches, the same volume as taken by 583 bricks, or 0.76 m$^3$. Concrete can be poured into concrete forms or poured to form a slab at a great rate, whereas bricks must be laid individually. However, a red brick facade is far more attractive than concrete.

The problem with brick and concrete is its high thermal conductivity. A poured concrete foundation subject to low winter temperatures needs to be insulated. In U.S. the usual method is to construct a stud wall on the inside of the foundation and fill the cavity wall created with insulation. Now, insulated foam board can be placed on the outside of the concrete foundation, but this is not fully satisfactory since such board is easily damaged. The answer may be the use of ICF (Insulated Concrete Formwork), as discussed in Section 2.26, as outside protection for the foamboard.

4.3.2 Poured Concrete Foundations and Walls

Concrete is comprised of portland cement, sand, and small stone as aggregate; a common mix is one part portland cement, three parts sand, and five parts stone. Concrete is the most important material in the world for the construction of foundations of buildings.

Early cements were concoctions of lime and clay, and were typically used as mortar between rocks or bricks. A major step towards modern day portland cement was developed by Louis Vicat in 1817. He graduated from the École Polytechnique, then obtained a higher degree from France’s oldest civil engineering school, the École Nationale des Ponts et Chaussées (the National School of Bridges and Roads). He was commissioned to build a bridge over the Dordogne River in 1812. He experimented with a mixture of limestone and clay, which he burned while controlling the temperature involved. The mixture would set solid after immersion in water. The bridge was completed in 1822 and was the world’s first bridge to use cement.

Louis Vicat declined to file for a patent, for altruistic reasons. So, Joseph Aspidin, a bricklayer from Leeds, England, improved on the work of Vicat and obtained British patent #5022 in 1824. He called it portland since it resembled a rock found in Portland, a district in southern England.

However, the use of lime/clay cements as mortar between bricks or rocks dominated construction up to the twentieth century. Now, the dominant masonry construction material, by far, is ready-mix concrete delivered by truck to the building site and poured inside wooden
forms held together by tie rods so the thickness of the resulting concrete wall is at an assigned thickness; in America this is commonly 10” for a single-family residence. Alternative portland cement products include manufactured concrete beams, usually with prestressed steel cables set inside the beams, or concrete blocks.

We saw in Chapter 1 that making portland cement is energy intensive. The stone aggregate is typically 3/4” nominal, which means there are stones somewhat smaller and somewhat larger, but not as small as pea stone (so called because it is the size of a pea) or as large as 1.5”. These quantities can be varied, where a higher amount of portland cement makes the concrete stronger. The compressive strength of concrete is measured in lb/in$^2$, where the minimum acceptable for construction purposes is 2500 lb/in$^2$.

Another factor in the final strength of the concrete is the amount of water added to the dry mix; builders must use only enough water so the mix holds together — too dry it crumbles apart, too wet it slumps and is weakened.

Once water is added to the mix of portland cement, sand, and stone, hydrolization commences, and there is a limited window of time in which to use the concrete before it sets. This time is measured in at most a few hours, where the higher the ambient temperature the faster the set. Hydrolization itself causes internal heating, which is good in colder weather since the concrete cannot be allowed to freeze before it is set. A common practice in cold weather, when constructing a foundation by pouring concrete into wooden forms, is to cover the top of the forms with straw.

Concrete should not be poured when the ambient temperature is way below freezing. It is possible to accelerate the hydrolization process and internal heating by adding calcium chloride to the mix before the concrete truck leaves the plant, but this practice is often prohibited by code, and is always prohibited for large commercial jobs.

Salt is destructive to concrete. Sand taken from the seashore should never be used in making concrete, not only because of the salt content but because the sand has been rolled by wave action to round the individual grains; sand used in making concrete and masonry cement is always specified as sharp, often as one of the products coming from a stone crushing operation.

### 4.3.3 Concrete Block Walls

A two-hole hollow concrete block is 7.625×7.625×15.625, so with 0.375” mortar the block spans 8×8×16”. Concrete block is considerably less expensive than brick in material and labor, but it is not pretty. An end concrete block is shown in Figure 4.4. The plane end is used at the beginning of a wall. It has two rectangular voids, termed “core” in the figure, to make the block lighter and easier to handle without compromising its strength. The stretcher end, drawn without some of the detail needed, is intended to be joined to its neighbor with a mortar joint.
4.3.4 Rock Construction

Rock is possibly the most durable of the masonry materials. It can be used in an unaltered state, such as field stone. In colonial times in New England the farmers cut down native forests for fuel, building materials, and to make fields to grow their crops. There were no stones in these newly created fields. At least, there were no stones on the surface, but after a few years stones appeared. The leaf mulch when the area was forested prevented frost from penetrating the ground, but cleared fields had no such frost protection. The freezing/thawing cycles forced the buried stones to pop up. These stones interfered with the farmer’s ploughs, much to their annoyance. They built stone walls around their fields, not because they looked nice, but because it got the stones out of the way.

Old New England houses had field stone foundations. When properly constructed they are durable. Some extant early houses, dating back to the mid seventeenth century, in Salem, Massachusetts, have field stone foundations and a central chimney structure servicing several fireplaces; the Witch House, dated 1642, is an outstanding example. I live in a house that was constructed around the middle of the eighteenth century, and its foundation is of field stone. Most houses in New England followed this practice for the better part of another century.

Soon after the U.S. Civil War, the foundations of Boston’s major buildings were made of cut stone, usually granite dug from Cape Ann north of Boston. Cut stone permits each individual stone to butt up tight to its neighbors, making for a stronger and more stable foundation than is possible with field stone. Now, the largest of these granite quarries, Halibut Cove in Rockport, Massachusetts, is a state park; large cut stones abound in this park, left behind presumably since they did not meet construction specifications.

Availability and affordability are everything. Industrialization in England led to the railroads, and affordable, durable slate from Wales could be shipped almost anywhere in England, so slate roofs became the preferred roofing material. The life of a slate roof is determined by the
metal pins holding the slate in place. A natural slate roof is common in England, less so in
north America, except on churches.

Old slate, harvested from existing roofs, is the most desirable and thus expensive. Some
unscrupulous contractors in England would offer to replace the roof on a slate-roof house free
of charge, claiming that the roof needed repair, taking the valuable old slate and replacing it
with a cheaper product.

In the hill country of northern Italy it is common to find houses with stone roofs. This stone is
not slate but sometimes 10 cm or more thick. It is an example of using what is available rather
than picking the best material. In regions that are geologically active, this is a dangerous
practice: In such areas houses should use light materials that will reduce the chance of a
collapse crushing the occupants. Parts of Japan are subject to extreme earthquakes, and it is
common to see people living in houses with paper sidewalls and light roofs.

Nonflammable roofing materials should be required in all areas subject to wildfires, such as
Australia and the Western United States. These materials include metal, slate, and ceramic.
Ceramic roofs are very popular in Mediterranean climates and southern England, but are less
popular in colder climates. Ceramics do not absorb water to the extent that freezing is a
problem, so they should have wider appeal. They are more expensive than most other roofing
materials, but have a long service life, often over 100 years.

4.3.5 Earthen Materials

The use of unfired earthen materials for foundations is common in third world countries, as is
the use of unfired bricks. There are lots of choices in the use of earthen materials. In the
American Southwest, adobe is popular and very attractive. Adobes are mud bricks dried by the
Sun, where in this case the mud is predominantly clay and sand. The adobe bricks are then
stacked to create walls on a foundation of concrete or stone with cement or mud for mortar.
Such material would not survive in rainy climates, such as England. Even in a semi-arid
climate care must be taken, such as having large overhangs of the roof to protect the adobe
from the beating rain, and the use of gutters to prevent splashing.

Rammed earth is just what it sounds like. Simple wooden forms are filled with earth that is not
predominantly clay to create a wall from 18” to 24” thick. Again, such a wall is not durable in
rainy climates.

A variant on the rammed earth wall is the earthbag wall. Here, earth-filled bags of burlap or
polyethylene are stacked and pressed down with barbed wire as a binder. Since sunlight
deteriorates the polyethylene it must be covered, often with plaster.
4.3.6 **Insulated Concrete Formwork**

A fairly new technique, made possible because of recent advances in expanded polystyrene board, is ICF (insulated concrete formwork). Set on a solid concrete foundation, the hollow polystyrene blocks, about 8×8×16” long, and reinforced to prevent the 16” sides from separating during construction, are laid to the full height of a wall. The reinforcement is with ties or straps made of metal or plastic. The metal reinforcement has the unfortunate result of causing thermal bridging, which can result in a substantial increase in thermal conductivity. The blocks are designed to interlock, an important detail for when they are concrete filled.

When the wall of polystyrene block is completed and braced, the hollow interior of the blocks is filled with pumped concrete. The concrete must be fresh, not beginning to set, and the concrete must be directed deep in the wall with a flexible hose so that there are no voids in the concrete.

When the concrete is set, another ICF wall can be constructed on top, or a wooden joist system can be laid, or roof trusses set. Openings for windows and doors are framed in during the laying of the blocks.

The main advantages of a stress intensity factor house are low thermal conductivity, a high thermal mass to even out external temperature changes, and good sound-deadening properties. Also, the cost is marginally more than a brick and block construction.

4.3.7 **The AAC Block Wall**

Autoclaved aerated concrete was invented by Swedish architect Johan Axel Ericksson in the 1920s as a replacement for wood, which at the time was scarce. It is manufactured from powdered silica, which is mixed with water to a slurry to which limestone powder, portland cement, and a small amount of aluminum powder are added. The aluminum produces an instantaneous chemical reaction in which hydrogen gas is produced, causing small bubbles in the mixture, which rises like bread dough. It is left to partially set, then wire-cut into blocks of exact dimensions before entering a 400°F pressurized oven. What emerges is a block that has 1/5 the weight of concrete yet can withstand a compressive pressure of 1100 lb/inch$^2$ [www.concreteconstruction.net/concrete-construction/building-with-aac.aspx as accessed on, December 18, 2014]. A 8”x8”x24” block weighs only 35 pounds. Usually, the wall is dressed inside and outside with a mortar stucco.

The labor cost of constructing AAC walls is low. Individual blocks are connected with a thin layer of mortar, about 2.5 mm thick. Beams laid across the top are affixed to the top blocks. The resultant walls are fire, wind, and water resistant, and have good sound insulation properties. AAC blocks, like concrete blocks, come in modular-friendly sizes. The three most
popular sizes are 4×8×24”, 8×8×24”, and 12×8×24” (100×200×608mm, 200×200×608mm, and 300×200×608mm), where the first dimension is width, the second is height, and the third is length; the dimensions given include the mortar layer. Rather than using individual AAC blocks, AAC panels are also available up to 20’ tall. They come 24” across, consistent with the modular format. Panels make for rapid construction; a small crane will be needed to set these panels.

The thermal properties are good, but not great. The r-factor per inch is about the same as that of softwoods at 1.25 h·°F/BTU, compared to polystyrene board at 3.8, and fiberglass at 3.3, but it is better than concrete at 0.08. This makes their use in latitudes 40 or greater problematic, since they do not satisfy building codes for thermal conductivity. Exterior treatments can be used to satisfy these codes, and the best choice could be the EIFS system; see Section 3.6.

AAC walls have a high thermal mass, which makes them very desirable in climates with large daily temperature swings. By the time the heating effect of the Sun passes through the wall and enters the interior the outside temperature may be lower than the interior, temperature. This works well in Mediterranean climates, and here these panels will satisfy building codes for insulation that could not be satisfied in northern Europe.

Another advantage of AAC walls is that they muffle sound very effectively. This can be important in some locations, such as near a highway, or a manufacturing plant, or a nightclub.

### 4.4 Insulation, Types and Installation Methods

Older buildings were heavy buildings. Most were masonry, but even the wooden buildings were heavy. With industrialization came the availability of dimensional lumber with standard sizes, and with more sophisticated building techniques the wooden buildings became much lighter. Most of these had cavity walls that leaked like a sieve. Central heating became a standard feature in early twentieth century America, at least amongst the northern tier states. The fuel used was wood or coal, and the cost of fuel was low, so who cared about the heat losses!

Convenience became a critical factor in the 1950s, and the heating fuels of choice became oil and natural gas. Then the gas crises with rapidly escalating prices of fuels in the 1970s changed everything. The domestic economies of the advanced world woke up to the need for energy efficiency. There was a need to fill the cavity walls with an effective insulator.

The traditional insulating material was whatever one could get one’s hands on, such as crumpled newspaper, dried corn husks, dried grass, and shells of nuts. These were low in R value and high in variability, but it was the only show in town. This changed by the 1940s.
In the late 1980s my wife and I bought a property, mentioned in Section 2.24, conveniently close to my university. The house was a mess but the property was well worth the asking price. I evaluated the value of the house at negative $5,000, and in retrospect I was accurate. We ended up replacing every window and every inch of plumbing and electric. The ceiling of the kitchen was sloping and the vertical separation between the back burners of the stove and the ceiling was about 1 m. In taking down the ceiling I found the insulation was crumpled newspapers, just 1 m above an open flame. Please be assured that the fire hazard no longer exists.

4.4.1 Fiberglass Batts

What was the insulating material that replaced newspapers? It is what we now call fiberglass batts, and was discovered by accident in 1932 when a researcher called Dale Kleist was trying to create a seal between glass blocks. A high pressure jet of air turned a stream of molten glass into fine glass fibers. He saw a potential product, and continued to work on it. He turned to steam rather than air, and the result was a method for producing large quantities of these glass fibers [www.protectall.com/artfiber.aspx]. The name Fiberglass was trademarked in 1938, the same year that Dale Kleist and his supervisor at Owens-Illinois, Jack Thomas, were awarded patent number 2121802.

Corning Glass was also experimenting with glass fibers in the 1930s, and in 1938 they joined Owens Illinios to form the Owens-Corning Fiberglas company, creating a virtual monopoly. In 1949 an antitrust ruling prevented either parent company from controlling Owens-Corning, so in 1952 it went public as an independent entity [www.encyclopedia.com/topic/Owens-Illinois_Inc.aspx].

Craft-faced fiberglass should be used for walls, ceilings, and between rafters. It should not be used to insulate an attic floor when the only access is from the top. When craft-faced fiberglass batts are installed the craft face must be positioned on the heated side of the structure. The rule is to prevent moisture from entering into the insulated cavity, but, if it does, let it out. The craft face is a vapor barrier, of sorts. After the fiberglass is installed an inexpensive but effective vapor barrier is 6 mil polyethylene sheet, which should cover the complete surface and lap over to the adjoining surfaces. For example, with a wall the plastic should extend a little way onto the ceiling and over the floor to prevent corner and edge leaks.

Fiberglass is inert and inorganic, so it has an unlimited life. It will not deteriorate when wetted, unlike cellulose, and will not sag, particularly when craft faced. However, to ensure no gap in the insulation will occur it is prudent to cut the length of the batt a little longer than the height of the cavity and cram it in.

The problem with the fiberglass products presently on the market is the variability or their stated thermal resistivity values. A recent survey (May 15, 2015) at the Home Depot and
Lowe’s of fiberglass insulation revealed the following resistances and associated thicknesses: R13 3.5”, R15 3.5”, R19 6.25 (made to compress to 5.5” in a 2×6 studded cavity wall), R21 5.5”, R30 7.5”, R30 9”, and R30 9.5”. The lowest, average, and maximum resistivities are 3.0, 3.7, and 4.3. This is a disturbingly wide range, and may indicate puffery rather than performance.

What is the effective resistance of the fiberglass R19 6.25, designed to compress to 5.5”? If it has resistance R19 it has resistivity of 3.45, whereas uncompressed it has resistivity of about 3.

There do not appear to be any recent independent investigations that have tested the resistances of fiberglass batts. An early experimental study of the actual thermal resistance values of fiberglass insulation batts showed that the actual R values for R11 batts were consistently below stated value, but within 10% of the nominal value. However, the R value of one manufacturer’s R19 batts tested more than 10% below the stated value [R.P. Tye, A.O. Desjalais, D.W. Yarborough, and D.L. McElroy, “An Experimental Study of the Thermal Resistance Values (R-values) of Low-Density Mineral-Fiber Building Insulation Batts Commercially Available in 1977,” report ORNL/TM-7266, Oak Ridge National Lab, April 1980]. Therefore, taking an average value for the stated resistances and subtracting 10% results in R3.3, the number assumed in Table 3.4 of Chapter 3.

### 4.4.2 Cellulose Insulation

Cellulose insulation is the closest competitor to fiberglass batts. It is made from at least 80% recycled paper by volume. With cellulose insulation, are we coming full circle?

Cellulose is blown into place. This can be done by drilling a small hole into each cavity and filling it with cellulose insulation. It can be used where it would be difficult to use fiberglass — namely in cavity walls and other spaces that would be difficult to open up. Older homes, often lacking any insulation, can be greatly improved thermally at modest cost and little fuss. However, cellulose should never be used when the cavity is fully accessible.

There are a number of things not to like about cellulose. Paper is flammable and rapidly deteriorates, so a number of chemicals are added, including borax, boric acid, sulfuric acid, ammonium sulphate, formaldehyde, fungicides, insecticides, and glues. “The New England Journal of Medicine published a review on health effects of cellulose insulation which reported that ‘the fire retarding chemicals used to recycle cellulose insulation could be potentially carcinogenic’ [Deyanda Flint, www.ehow.com/about_6173833_health-risks-cellulose-insulation].” Additionally, there are respiratory problems associated with cellulose, so care must be exercised to clean up after installation and to properly fill the holes.
4.4.3 Urea Formaldehyde Foam

Urea formaldehyde foam is installed through small holes into an otherwise inaccessible cavity wall, just like cellulose, but unlike cellulose it hardens and does not sag. It has a thermal resistance of R. Its high flowability allow it to completely fill a cavity wall, but it has a shrinkage problem. Typical shrinkage after installation is about 0.5%, but can be several times that amount, compromising its insulation value.

URF became popular in the 1970s when energy prices sky rocketed. It seemed to be the ideal insulation for older, existing houses. However, in 1982 the United States and Canada banned it in schools and residences since it off-gassed formaldehyde, which is a known carcinogen. The owner of a house with urea formaldehyde insulation was required to remove it at their expense. This removal was not easy, and it was not cheap. It required the interior walls to be removed, the foam to be removed, and the cavity neutralized with a solution of sodium bisulphite. Further, the work had to be done by a certified environmental contractor. Interestingly, the builder who installed the urea formaldehyde foam in the first place walked away, leaving the problem to the homeowner. The U.S. Court of Appeals overturned the ban in 1983, but the industry never recovered [www.carsondunlop.com/resources/articles/ureaformaldehyde-foam-insulation].

It should be pointed out that formaldehyde occurs naturally in fruit and vegetables. It is also present in a number of building products, especially fiberboard and hardwood plywood. It is present in softwood plywood used in sheathing and subfloors, but at about 10% of the level in hardwood plywoods.

The typical formaldehyde level in a house is around 0.03 ppm. The off-gassing from building products decreases with time such that formaldehyde levels in a five-year-old house are of little concern. The use of polyethylene sheeting covering studs and rafters substantially reduces interior formaldehyde levels.

4.4.4 Foam and Foam Board

Expanding polyurethane foam is an effective way of insulating a cavity space. It comes in open cell with r3.4 h·°F/BTU, or closed cell with r6. Open cell permits air and water to pass through, whereas closed cell does not. An effective vapor barrier should be installed on the heated side of the wall if the open cell polyurethane is used.

Foamular®, an extruded polystyrene board insulation, is a registered trademark of Owens Corning Company, with a compressive strength 49, 60, or 80 psi. It is a faced extruded polystyrene board suitable for foundation insulation. The section of the foam board above ground and for at least 0.5 m below grade should be shielded from physical damage with
inorganic solid board; one such product is 1/2” thick Permabase Backerboard, and comes in sheets 3’×5’.

4.4.5 Multi-Layer Radiant Barriers

The thermal resistivity of the insulating materials considered so far runs from r3.2 h·°F/BTU per inch for fiberglass to r6 for polyurethane board. To obtain a well-insulated sidewall, a fairly thick layer of such material is needed. There is an alternative that has been employed since the 1950s by NASA. It is radiant heat barrier, defined by the DOE (Department of Energy) as a product having emittance of 0.1 or less and reflectance of 0.9 or more.

Multi-layer radiant barriers are an ideal insulator for satellites in space, as well as for shielding humans against damaging extraterrestrial ultraviolet rays. However, as an insulator they do not work as well on Earth. In space, heat transfers though the multi-layer barrier by radiation only, whereas all three mechanisms (conduction, convection, and radiation) occur on Earth.

The emissivity of polished aluminum foil is about 0.05, while its thermal conductivity is 1416, one of the highest of all common materials. Multiple layers of aluminum foil, constructed so as to maximize the number of layers per unit thickness while minimizing the contact area of adjacent sheets, can produce an astonishingly low thermal conductivity. Not only are these so-called radiant barriers thin, they also weigh little.

An alternate to multiple layers of aluminum foil is a sandwich of polyethylene foam between aluminum foil. They are employed extensively in refrigerators, automobiles, and protective clothing for firefighters.

Prodex AD5 has aluminum foil on the outside, polyethylene backing next, and a core of closed-cell polyethylene foam. It has an effective thermal resistance of 15.67 h·°F/BTU, equivalent to 4” of fiberglass wool, yet is only 0.2”, 5 mm, thick. It comes in rolls, 16”x175’, 24”x175’, 48”x175’, and 72”x100’. A substantial contributor to the resistance is the air gap necessary to be maintained on both sides of the product. For example, with two layers of AD3 and no air gap between them, and air gaps on either side, the total resistance is 21.1: with an effective air gap between them the resistance would be 31.4. Prodex AD10 is 10 mm thick and has thermal resistance of R21.36.

Similar products are available from different manufacturers. To guard against moisture penetration in the wall, 6 mil polyethylene sheeting should be installed against the studs before the wallboard is affixed.

A number of companies, such as Reflectix and Radiantguard, offer double bubble radiant barrier insulations. The outer layers are thin aluminum foil with high reflectivity followed by a
polyethylene layer, to give the overall construction structural integrity. The inner two layers are polyethylene bubbles, using the trapped air inside each bubble as insulation. A cross section of this type of product is shown in Figure 4.5. When used as part, or all, of the sidewall insulation, the material should be tabbed over and stapled as shown in Figure 4.6 so the material is about 3/4” from the interior wall. This creates a dead air space with the low emissivity of the radiant barrier minimizing the radiant heat flow from the interior to the exterior: as was seen in Chapter 3, the radiant heat flow across an air gap lined with high emissivity surfaces is dominant, considerably larger than the convection plus conduction heat losses. Staples should be applied every 10 cm. The radiant barrier material provides resistance of about R13 when applied in this manner.

![Figure 4.5: Double Bubble Radiant Barrier](image1)

![Figure 4.6: Affixing Radiant Barrier with Staples](image2)

The most common application we find for radiant barriers is in the sidewall. However, manufacturers of such products discover other applications that may produce more sales. One of these applications is in radiant under-floor heating, particularly beneath the subfloor, to direct radiant energy upwards into the living space above. It can also be used under a poured-in-place concrete slab. In unfinished basements and in crawl spaces the material can be stapled across the floor joists and provide a resistance of about R17. Another application that we do not consider here is to reflect summer heat to keep it from penetrating the roof and entering the
living space below. Since we recommend the cool roof system described in Section 3.2, there is no need for a radiant barrier in this location.

A well-established practice when insulating to keep heat in the structure is to put a vapor barrier on the heated side, but never on the unheated side. When insulation is designed to keep heat out, humidity levels must be investigated to determine on which side, if any, the vapor barrier should be placed. Any vapor that gets into the insulation proper due to flaws in the barrier may travel through that insulation without inhibition. Thus, when reducing heat flow into an attic space, a radiant heat barrier should be stapled under the attic floor joists, not on top of them.

Breathable radiant barrier sheets are also available, and these should be used when the only practical place for these products is on the unheated side of a wall, typically under the siding.

### 4.4.6 The EIFS Insulation System

EIFS (Exterior Insulation and Finish Systems) is an exterior, non-load bearing wall system that is insulating and decorative. It was first marketed in Europe in the 1950s as a way to insulate older masonry buildings and to smarten their appearance. The EIFS wall has a number of layers, the outer layer for appearance, another layer for insulation, and a water-resistant layer.

The insulation value for the EIFS depends on the thickness of its foam board. Double its thickness and the insulation value is almost doubled.

EIFS walls were introduced to the United States in 1969, and soon became reasonably popular, but their image suffered in North Carolina when flashing caused water infiltration and a number of lawsuits. Other problems occurred in New Zealand and British Columbia. The manufacturers cite improper construction techniques when EIFS was installed over wood frame walls. EIFS may be the answer for facelifting older and drab masonry buildings, but there are other and better methods for insulating wood frame structures.

### 4.5 Roofs, Construction and Problems

It is the roof of a residence that receives the most solar energy and everything that nature can throw at it — rain, snow, hail, and wind. Attic areas are subject to greater extremes than the living environment below them. We consider some of these problems in this section.

#### 4.5.1 Roof Construction Techniques
Unlike the sidewall of a house, there are fewer options for the roof. The most common technique is to use wooden rafters, overlaid with plywood. As with a wood frame sidewall, the spacing of the rafters will be 16” or 24” on centers to make best use of the 4’×8’ plywood sheets.

There is a wide selection of roof surfaces, and sometimes the best to use is dictated by local climate, aesthetic preference, or conformability. Ceramic roofs are very common in southern England where winter temperatures are moderate. They are also common in the American Southwest. Less common today in southern England, but considered highly desirable, is the thatch roof — discussed in Section 2.12 of this chapter. Thatching is a dying industry. There are few thatchers left to ply their trade.

It is claimed that SIP panels are rigid enough to be used in place of rafters. This seems a poor way to use these panels.

4.5.2 Unwanted Heat and Humidity Flow through Roofs

The summer Sun can significantly heat the area below a roof. To avoid this it is desirable, perhaps essential, to employ a cool roof strategy that permits air flow beneath the roofing plywood, as shown in Figure 4.7. The facia board covers the bottom portion of the rafter tails, so a manufactured metal drip edge, with slots to permit air flow while protecting against water penetration, should be used; one such preformed drip edge is Hick’s Starter Vent. At the ridge the plywood is cut back a few inches so air flows from under the plywood though a preformed metal ridge vent.

Figure 4.7: Preferred Method for Roof Insulation
In order to ensure air flow beneath the plywood, simple and inexpensive rafter vents should be used. There does not appear to be a commonly accepted term for these devices, which go under proprietary names such as Duravent, or Vent Chute. The rafter vent is 22” wide, as is Duravent, or 14” wide, as is Vent Chute. Made of plastic, they fit between the rafters, which are 24” O/C, or 16” O/C and are stapled in place. The fiberglass or foam insulation is installed to fill the cavity between the rafters, but the air gap between the rafter vent and the plywood remains. Thus, a clear path for the air to enter the starter vent to travel above the rafter vent, and emerge at the ridge vent, produces a cool roof system in summer and winter.

An additional benefit of the cool roof system is the prevention of ice dams in winter. With an unvented roof, heat leaking from the house melts snow or ice on the roof. As the water trickling down the roof gets to the overhang, the heating effect from the house to the roof surface is eliminated, or substantially reduced, and ice dams can be formed. The result is a backup of water under the roof shingles, particularly in low-pitched roofs, and leakage into the house can occur. Home owners that do not have a cool roof system have had to resort to electrical heater cables, which are clipped to the roof shingles in a zig-zag pattern that runs from the drip edge to some distance above where ice dams could form. When activated, the cables provide a pathway for water flow.

Radiant heat barrier sheets or polethylene sheets can be stapled under the rafters to reduce heat loss by inhibiting air infiltration. The choice to be made is whether to have a cathedral ceiling below, or to lay horizontal joists to create an attic space. If the choice is to have an attic, the joists can be insulated and a vapor barrier at the bottom edge of the joists is recommended. Craft-faced fiberglass insulation can be used with the craft face on the underside. A more convenient method could be to finish the ceiling below while using a radiant heat barrier, then laying unfaced fiberglass batts or blown-in cellulose insulation before laying a plywood floor in the attic: The finished ceiling provides support for the fiberglass or cellulose.

Folding stairs are most commonly used to access the attic. Such stairs are difficult to insulate, but the major problem is that it allows the warm humid air to enter the cooler attic. Unless some venting of the attic is provided, dew point can be reached and damaging condensation can occur. Fortunately, humidity reduction is easier to achieve in an attic than temperature reduction. Small vents at the gable ends take care of the humidity problem. Hip roofs create problems, and such roofs should not be insulated and should have ridge venting; instead, the attic floor insulation should be beefed up to provide all the heat loss control needed.
4.5.3 Roof Problems

We are sometimes dismissive or forgetful regarding considerations of roof construction, or even the roof surface itself. At the time this passage is being composed, wildfires in Washington State are blazing out of control. A wildfire can be a self-feeding engine, generating its own wind, which whips up burning undergrowth and small branches that can cross fire barriers such as roads and rivers. When a wildfire reaches a house, the survival of the house most likely depends on its roof, since the burning embers, known as firebrands, land on the roof, and if the roof is combustible the house is gone.

Unfortunately, many houses in regions susceptible to wildfires have cedar shake roofs. Some communities ban such roofs and require noncombustible roofs made of ceramic, clay, or metal. Certainly for new construction, non-combustible roofs should be mandated for all susceptible areas. Roofs in the United States are rated as Class A, B, or C; wood shake roofs are not rated. Further, gutters may not be a good idea in wildfire-prone areas since they can accumulate debris that a firebrand can ignite.

The Insurance Institute for Business and Home Safety has a check list:

* Does your roof have features or details that can increase the chance that it will catch on fire and the fire enter your attic?
* Are your gutters full of vegetative debris (pine needles, leaves, and twigs) or are gutter covers allowing debris to accumulate on the roof?
* Does your roof have dormers or other features such as those found in a split level home where vertical walls intersect with the roof, or where the roof of one level overhangs the roof of a lower level?
* Does debris such as pine needles accumulate at the vertical wall-to-roof intersections?
* Is the exterior siding at these locations combustible? Wood and vinyl are common types of combustible siding. If so, does the siding come within 4 inches of the roof?
* Are your attic or crawlspace vents protected with a fine screen or some other method for reducing the chances that embers will enter your attic or crawlspace?
* Do you have a deck, patio, or porch with flammable material attached to your home or business?
* Are your windows and glass doors protected? Do you have wood structures, including wood gates, attached to your home or business that could provide a path for wildfire flames to come into direct contact with your building?
* Have you evaluated the surrounding terrain and vegetation and taken action to reduce the intensity of a wildfire approaching your building?

Roofs can collapse under snow load, or fail under wind load. 10” of dry snow has the same weight as 1” of water, or 5.2 lb/ft². A typical live load rating in the northern United States is 40 lb/ft², equivalent to 77” of dry snow. Following the brutal winter of 2010/2011 when a number of roofs collapsed in Massachusetts, the snow load capability of roofs was increased. The intense snowstorms in January through March 2015 will probably result in a further increase. It
should be pointed out that failures of conventional roofs on residences due to snow were rare, and most failures occurred on large commercial buildings such as big-box stores; the worst performing roofs were metal salt-shed roofs, typically flimsily constructed and with shallow pitch and considerable height above ground level.

The snow load rating of a roof is a function of the pitch of the roof: the shallower the pitch, the higher the snow load requirement. However, the reverse is true when assessing wind loads. The result is that the load rating of a roof, snow and wind, can be considered independent of the roof pitch.

4.6 Sealing and Insulating Basements

The traditional waterproofing technique in the United States for concrete foundations is tar. It is not a satisfactory substance since it is organic and breaks down over time. It is easily pierced. A portland cement-based product that works much better is a product called U-Seal.

Now, with tightened energy codes requiring insulation of basements, the choice is using closed cell foamboard on the outside, or studwork on the inside, creating a cavity wall that can be insulated. The problem with foamboard is that it is easily damaged and needs to be protected, as discussed in Section 3.4.

A problem that seems to be prevalent in England is rising damp. A masonry wall can wick water up the wall by capillary action, or so conventional wisdom asserts. The result is an unsightly “high tide” mark on plaster walls in the house that can rise to 0.8 m. This can be prevented by installing a barrier to water at the base of the wall, and to provide adequate drainage around the foundation.

Now there is a considerable body of evidence that there is no such thing as rising damp, and a high tide mark on the wall is due to trapped water inside the structure. The solution is to reduce the internal humidity, and the problem goes away.

4.7 Water Damage and Vapor Migration in a Residence

Water in the fabric of a building can cause significant and possibly irreparable damage. It can eat away at wood without any indication of its insidiousness, until failure occurs. Possibly the first sign is a window that fails to open or a door that jams as the building settles. The type of damage that occurs depends on the climate and the materials and architecture used. Besides the undetected water damage, here are obvious problems:
* Flooding or structural damage caused by hurricanes.
Amongst some lax standards, a number of states and/or municipalities in the United States have permitted building in floodplains. One example is building in one of the many floodplain areas of the Mississippi River. Another is building on a sandy barrier island, such as occurred on the New Jersey Shore; this is described by Ian McHarg in _Design With Nature_, John Wiley, New York, 1997: more about this in Section VIII of Chapter 9.

One of the effects of global warming is the melting of polar ice caps, as well as glaciers and snow caps on mountains, with the result that sea levels are rising. Also, as sea temperatures rise some experts warn that storm events are going to occur more frequently and with more severity. Some 100-year storm events are now occurring every five to ten years. Flood plains need to be reassessed and tighter restrictions placed on construction.

In Chapter 2 there was a discussion of the ability of air to hold moisture as a function of its temperature: the higher the temperature, the more moisture. In climates where winter temperature differentials from interior to exterior can be 25°C or more, and with the interior humidity level for comfort at about 30%, the vapor pressure is higher on the inside than the outside. Thus, water vapor migration is from the inside to the outside.

Moisture vapor in walls becomes a problem when the relative humidity reaches 100% and dew point is reached. In a wood frame structure, the sills can be wetted, even saturated, and rot will result. To avoid this condition, a vapor barrier should be placed on the inside of the studs, underneath the wallboard or wood paneling or other interior material. Cognizant of this, some building codes in the United States require pressure treated, rot-proof sills.

Even with rot-proof sills, it is important to prevent water droplets from forming in the walls. A simple and inexpensive solution is to cover the interior studs with a vapor barrier such as heavy duty polyethylene sheeting. This sheeting would rot over a season if exposed to sunlight, but inside the wall it lasts forever. Never should the vapor barrier be placed on the cold side of the wall. For that reason, boosting the R-factor of a cavity wall with foam insulation on the outside could be a problem, particularly if fiberglass or cellulose insulation is used in the cavity wall, and even with an interior vapor barrier.

Under no circumstance should the vapor barrier be placed on the outside of the studs in colder climates. This will trap water already in the wall. Foil-faced foam insulation nailed to the sheathing will trap water, and should be avoided. Siding should be breathable. Tar paper under the siding is to be avoided in favor of breathable house wraps such as Tyvec or Typar, which stop water from passing through into the wall structure if the siding leaks, but let water vapor pass easily from the interior to the exterior.
Table 4.2: Vapor Pressure for Saturated Air

<table>
<thead>
<tr>
<th>°F</th>
<th>VP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.038</td>
</tr>
<tr>
<td>10</td>
<td>0.063</td>
</tr>
<tr>
<td>20</td>
<td>0.103</td>
</tr>
<tr>
<td>30</td>
<td>0.165</td>
</tr>
<tr>
<td>40</td>
<td>0.248</td>
</tr>
<tr>
<td>50</td>
<td>0.363</td>
</tr>
<tr>
<td>60</td>
<td>0.522</td>
</tr>
<tr>
<td>70</td>
<td>0.739</td>
</tr>
<tr>
<td>80</td>
<td>1.032</td>
</tr>
<tr>
<td>90</td>
<td>1.422</td>
</tr>
</tbody>
</table>

A new timber frame house has a lot of water in the wood. Green softwoods contain over 2.5 lb of water per board foot; a board foot is a board of thickness 3/4”, width 11 1/4”, and length 12”. When kiln dried to the normal 19%, water content of a board foot drops to 1.95 lb. A typical 2,000-square-foot house contains about 16,000 board feet of lumber. The water content in the wood drops to about 10%, so 9% of the weight of lumber is in water that needs to migrate out of the house, about $16,000 \times 0.09 \times 1.95 = 2808$ pounds of water, or $2808/8.35 = 336$ U.S. gallons.

The rate of moisture transmission through a material, regardless of thickness, is termed its permeability, measured in perm/inch. Multiply the permeability by the thickness of the material in inches and one has its permeance, measured in perms. A perm is 1.0 when one grain (1/7000 of a pound) of water vapor passes through one square foot of the material in one hour when the vapor pressure differential from the warm side to the cold side is 1” of mercury (1”Hg).

The vapor pressure $V_P$ for saturated air in “Hg is given in Table 4.2 as a function of temperature in °F. To find the vapor pressure at a relative humidity of X%, one multiplies the saturated vapor pressure by X/100.

Supposing a residence is at 70°F and a relative humidity of 30% with the outside ambient temperature at 0°F and relative humidity of 40%, then the differential vapor pressure is $0.3 \times 0.7392 - 0.4 \times 0.0377 = 0.2067$ “Hg. For a structure with 3,500 ft² of area (walls, roof, basement) through which vapor can pass through at an average perm rating of 0.1, then the water vapor passing from interior to exterior is $3500 \times 0.2067 \times 0.1 = 72.34$ grains/hour. For the month of January, with 744 hours, this amounts to 53819 grains, or 7.69 pounds of water.

ASTM (American Society for Testing Materials) defines a vapor barrier as one with a value of 1.0 or less in perms. The vapor barrier of choice to nail inside the studs is 6-mil polyethylene sheeting, with 0.06 perms. A thinner polyethylene sheet could be used, but the cost savings are negligible, and the thinner sheet is more susceptible to tearing. Aluminum foils could be used,
with 0.05 perms. Less effective is craft facing on fiberglass batts, with 0.40 perms. Vapor retardant paints, with 0.45 perms, should be used in existing structures.

<table>
<thead>
<tr>
<th>Table 4.3: Relative Humidity Chart</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>18</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>22</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>26</td>
</tr>
<tr>
<td>28</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>32</td>
</tr>
<tr>
<td>34</td>
</tr>
</tbody>
</table>

It appears likely that there will be a marriage between ISO (International Organization for Standardization) and ASTM. It makes sense that these two important organizations join to speak with one voice.

Relative humidity can be determined using a wet/dry bulb thermometer. In fact, this device has two identical thermometers. One has the bulb wrapped in wetted muslin, the other has its bulb exposed to the air. Allow enough time, ten or more minutes, for the evaporative cooling from the muslin to reach steady state. The web bulb thermometer will be lower than the dry bulb thermometer. Read the dry bulb temperature and the temperature difference between the two. Using Table 4.3, where the left hand column is the dry bulb temperature, and the top row is the temperature differential, find the position in the chart where the dry bulb row intersects the differential column, and that is the relative humidity. Temperatures are in °C. For example, at a temperature of 20°C and differential 5°C, the relative humidity is 59%.

### 4.8 Air Infiltration

Air infiltration is almost always considered a bad thing, since it is uncontrollable, and in winter can cause significant heat losses. Ventilation is considered a good thing, since it is needed for human respiration and is controllable. The two may have the same effect, but the needs of the occupants of a house for fresh air are usually, until with tight and ultra-tight houses
in the modern era, far less than the amount of air infiltration. Now, the ultra-tight houses require heat recovery ventilation, as discussed in Section V of Chapter 2.

Winter air leakage brings cold, low-water-content air into a structure, while losing warm, relatively high-water-content air to the environment. Summer air leakage brings warm, high-water-content air into a structure, while losing cooler air to the environment. In either case condensation can occur in the structure, resulting in mold or rot.

Air infiltration is a function of the temperature differential from conditioned spaces to the exterior, or to unconditioned spaces. It also increases with wind speed. Differential pressure across exterior walls in a building can be caused by the stack effect: the more stories in a structure, the greater this pressure, and in cold weather a differential pressure of 4 Pa/storey is typical.

Air infiltration can also be caused by the mechanical equipment in the building. Here are typical air needs of the following appliances, which can be as much as the numbers given in cfm.

- traditional open brick or stone fireplace ∞
- furnaces with internal combustion air 150
- water heaters 100
- clothes dryers 200
- range hoods 2000
- bathroom fans 100

The first item is rated as infinite loss because the use of the open fireplace often cools the house rather than heats it due to the enormous air flow up the chimney. It was not always so. In the days before central heating, houses were often at not much more than ambient temperature, so if the outside temperature was -5°C the inside may have been 7°C. People bundled up. The point is that an open fireplace can raise the ambient temperature by 12°C but not by 20°C.

Furnaces should use external combustion air and a counter-flowing heat exchanger with the combustion air. Even with this heat exchanger, the overall efficiency is often little more than 50%. A typical distribution of energy to yield 100 W is as shown in Figure 4.9: this figure and the numbers in it were adopted from “Better Duct Systems for Home Heating and Cooling,” U.S. Department of Energy (DOE). Furnace loss is 43 W, and a combination of conduction losses to non-conditioned spaces, air leakages, again to non-conditioned spaces, and system interactions, account for 54 W. This means that 197 W of fuel is needed to produce that 100 W of useful heat.

Gas- or oil-fired water heaters should also use external combustion air and a counter-flowing heat exchanger.

The U.S. Department of Energy DOE states that energy wasted from leaky residential ducts alone is equivalent to the energy burned by 13 million cars a year.
4.8.1 Air Leakage from Conditioned Spaces

A conditioned space is heated in winter and possibly cooled in summer: it is the complete living area of a house, including kitchen, dining area, bedrooms, and bathrooms. It is surrounded by exterior walls and walls to unconditioned spaces such as attics, unheated basements, and attached garages.

Water at 40 lb/inch$^2$ is running in a 3/4” pipe when it transitions to a 1/2” pipe. The water speed in the 1/2” pipe is 2.25 (the ratio of pipe areas) times water speed in the 3/4” pipe, and the pressure in the 1/2” pipe drops. The reason is conservation of mechanical energy. The kinetic energy increases in the 1/2” section, so its pressure must drop to conserve energy. This is called the Venturi effect. So what happens when wind flows over a low-pitched roof? The pressure on the roof is less than the pressure under the roof and any air leakage through the roof increases with increased wind speed.

The pressure of a 26 km/h, 16-mph wind is around 7 KPa, or 1 lb/ft$^2$. However, it is unlikely that a house will experience this differential pressure, but it is common on a high-rise building.

Wind against a house produces an increased pressure on the windward side, and decreased pressure on the leeward side. Even without air leakage, wind increases heat loss from a house in winter. Recall, in Chapter 3 we saw that heat loss through glass increases with increased wind speed, even without air leakage.

Air at 20°C or 68°F has a density of 1.206 kg/m$^3$ or 0.0753 lb/ft$^3$. It has a specific heat capacity of 1.005 kJ/kg-°C or 0.240 BTU/lb-°F. For a house considered a rectanguloid 40’ long by 24’ deep by 16’ tall, the volume is 15,360 ft$^3$, so the mass of air is $15360 \times 0.0753 = $
1157 lb with heat capacity of $1157 \times 0.24 = 278$ BTU/°F. For an ambient temperature of 30°F and an interior temperature of 70°F, and an air change rate of $x$ per hour, the heat loss from the structure is $278 \times (70 - 30)x = 278 \times 40x = 11103x$ BTU/hour.

Figures for air exchange rates and the heat loss from the house described above for various categories of construction are:

<table>
<thead>
<tr>
<th>Rate (hour$^{-1}$)</th>
<th>Description</th>
<th>Heat Loss (kW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25-0.35</td>
<td>tight construction</td>
<td>0.81-1.14</td>
</tr>
<tr>
<td>1.75</td>
<td>typical new house</td>
<td>5.7</td>
</tr>
<tr>
<td>2.5</td>
<td>older or poorly sealed house</td>
<td>8.1</td>
</tr>
</tbody>
</table>

The heat loss at 2.5 air changes per hour is about $1/3$ of the total heat loss from the house — a number that is unacceptable. The problem with a leaky house is how to make it less leaky. Air leakage in new construction is fairly easy to control, with careful attention to gaps that permit air flow from interior to exterior, and with vapor barriers on the inside of the bare studs before the finished interior walls are erected. A common myth is that fiberglass insulation in the cavity walls controls infiltration — not so. Polyethylene sheeting is far more effective than the craft face on the fiberglass. However, if fiberglass insulation is used in a cavity wall it should be craft faced, since it is a vapor barrier if somewhat imperfect, but more importantly it provides structural support to the fiberglass to prevent sagging.

The standard test to determine air leakage from a building is the blower door test. In this test the building is pressurized above ambient to 50 Pa, approximately 0.007 psi. The air necessary to maintain this pressure is measured as $\chi$ cfm, referred to as $\chi$ cfm50. The volume of the building is calculated as $\chi$ cubic feet. $60 \chi/V$ is air flow over one hour per cubic foot of interior space, the number of air changes per hour for the entire building when pressurized to 50 Pa, and is known as ACH50.

The ACH50 number is substantially greater than will occur in practice. One commonly used rule of thumb is to divide the ACH50 number by 20 to obtain an estimate for the number of air changes per hour, a number based on the analysis of J. Kronvall and A. Persily [M.H. Sherman, “Estimation of Infiltration from Leakage and Climate Indicators,” *Energy and Buildings*, 1987].

Many factors control the air leakages from a building besides the construction itself. These include the amount of exposure — a house totally exposed on the top of a hill will have substantially more air leakage than a similarly constructed house surrounded by trees and in a valley; see Section XII of Chapter 9. A procedure was developed in the 1980s at Lawrence Berkeley Laboratory (LBL) that determines infiltration based on a table, now known as the LBL table [M. Sherman and D. Dickerhoff, “Air-Tightness of U.S. Dwellings,” report LBR-35700, Lawrence Berkeley Laboratory, University of California, Berkeley, California]. This procedure has been subsequently simplified into an $n$-factor table by George Tsongas in *Home Energy Magazine*, which is given here as Table 4.4.
The first column of Table 4.4 gives the zone in the United States Zone 3 is:

* the Southeast, encompassing the states Louisiana, Tennessee, Kentucky, and Virginia and all states to the east of these, if any, and excluding the eastern half of Alabama and a small part of the western side of Georgia.

Table 4.4: Air Infiltration from a Structure

<table>
<thead>
<tr>
<th>zone</th>
<th>well-shielded</th>
<th>normal</th>
<th>exposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18.6</td>
<td>16.7</td>
<td>14.9</td>
</tr>
<tr>
<td></td>
<td>15.5</td>
<td>14.0</td>
<td>12.4</td>
</tr>
<tr>
<td></td>
<td>14.0</td>
<td>12.6</td>
<td>11.2</td>
</tr>
<tr>
<td>2</td>
<td>22.2</td>
<td>20.0</td>
<td>17.8</td>
</tr>
<tr>
<td></td>
<td>18.5</td>
<td>16.7</td>
<td>14.8</td>
</tr>
<tr>
<td></td>
<td>16.7</td>
<td>15.0</td>
<td>13.3</td>
</tr>
<tr>
<td>3</td>
<td>25.8</td>
<td>23.2</td>
<td>20.6</td>
</tr>
<tr>
<td></td>
<td>21.5</td>
<td>19.4</td>
<td>17.2</td>
</tr>
<tr>
<td></td>
<td>19.4</td>
<td>17.4</td>
<td>15.5</td>
</tr>
<tr>
<td>4</td>
<td>29.4</td>
<td>26.5</td>
<td>23.5</td>
</tr>
<tr>
<td></td>
<td>24.5</td>
<td>22.1</td>
<td>19.6</td>
</tr>
<tr>
<td></td>
<td>22.1</td>
<td>19.8</td>
<td>17.6</td>
</tr>
</tbody>
</table>

The second column defines the type of exposure of the structure. The numbers 1, 1.5, 2, and 3 heading the last four columns give the number of stories in the structure. So based on zone, exposure, and number of stories, the buildings n-factor can be determined. Thus, simply divide the ACH50 number by n to get an estimate of the air changes per hour.

### 4.8.2 Air Leakage around Windows

Air leakage is greatest around windows and doors. The ASHRAE/IESNA Standard 90.1 sets maximum air leakage around windows at 0.4 cfm per square foot of fenestration area, and around swinging doors to 1.0 cfm per square foot of each swinging door.

ASHAE (The American Society of Heating and Air-Conditioning Engineers) was founded in 1894, and ASRE (The American Society of Refrigerating Engineers) was founded in 1904. The two organizations merged in 1959 to become ASHRAE (The American Society of Heating, Air-Conditioning and Refrigerating Engineers), whose stated mission is “To advance the arts and sciences of heating, ventilation, air conditioning and refrigeration to serve humanity and promote a sustainable world.”
IESNA is the Illuminating Engineering Society of North America, a non-profit learned society that was founded in New York City with the mission to improve the lighted environment. The year was 1906, only 24 years after Thomas Edison’s Pearl Street power station provided street lighting to the city; see Section I of Chapter 14.

A commonly accepted rule of thumb is that air infiltration around an operable window is 10 times that around fixed glazing.

### 4.8.3 Air Leakage around Electrical Outlets, Switches, and Lights

The sequence of events after the frame of a structure is erected is rough plumbing first, electrical wiring next, communications third, followed by insulation in the exterior cavity walls and the ceilings to unconditioned spaces and the roof, and then installation of the 6-mil polyethylene vapor barrier. Typically, inspections followed by building card signoffs are required after completion of the rough plumbing, the rough electrical, and the insulation.

Plumbing in exterior walls in regions subject to freezing conditions is to be avoided; the exception to this is plumbing to exterior faucets, and there should be an easily accessible shutoff in a conditioned space.

The electrical outlets are uncovered using a utility knife cutting an X in the polyethylene from opposite corners of the box, and the plastic is pressed around the box to create a seal.

A typical electrical code requirement is that there should not be any position along useable wall space further than 6’ from an electrical outlet. Electrical outlets and switches along interior walls are of no concern as far as air leakage is concerned, but those on exterior walls can be a serious problem. Older metal out boxes have holes in them not associated with the cutouts for wiring. Modern plastic boxes have fewer such holes or no holes at all. Assuming fiberglass insulation is to be used, to reduce air leakage in and around an electrical outlet or switch box, the following steps can be taken:
* Caulk all places where wiring enters a box.
* After the wall board is installed, caulk any gaps between the box and the board.
* Install foam socket and switch plate sealers under each cover plate.

These steps are unnecessary if foam insulation is used to fill the cavity wall.

The same steps should be taken around the boxes for surface-mounted ceiling lights when an unconditioned space is above them. Recessed lighting presents greater difficulties, and such items should only be purchased after much thought. In particular, look carefully at the fixture to see if it is suitable for use with unconditioned space above.
7.4 Leakages in Air Ducts

Duct leakage must be less than 10–30% of the heat load of the structure, with leakage less than 4 ft³/minute per 100 ft² of floor area. Air handler leakage must be less than 2% of the air flow.

A plumbing leak is immediately recognized and corrected. An air leak is often (usually) not observed. Thus, without rigorous enforcement and strong codes, requirements on duct leakages were widely ignored. Now, rough-in testing an/or post-construction testing is universally required.

The IRC (International Residential Code) is used by some states, but it has a much larger focus than energy. The NFPA (National Fire Protection Association) has commercial and residential energy codes based on ASHRAE Standards 90.1 and 90.2, respectively. Some states (e.g., Florida and California) have developed and adopted their own energy codes.

Supply side duct leaks in unconditioned spaces depressurize the conditioned spaces, while return side leaks do the opposite. Either type of leak can cause severe energy loss. Leaks inside a conditioned space is far less detrimental to energy efficiency.

Figure 4.10: Air Pressure Changes due to Furnace
Figure 4.10 shows a simple furnace and duct system. The furnace is located in the basement area, which is not heated, so the furnace should be insulated, as well as the ducts in the basement. The supply side has positive air pressure, so a leak will have air leaving the duct. The return side has negative pressure, so a leak will have air entering the duct. Most of the conditioned space has neutral pressure. In a good duct system, duct runs are as short and straight as possible. Ducts should not run in exterior wall cavities. As much as possible, the furnace and associated ducts should be located in conditioned spaces.

Hardware store duct tape is not approved for duct sealing — it has a limited life.

The 2009 IECC code requires air leakage to be less than 0.08 ft\(^3\)/ft\(^2\) of conditioned floor area, and the 2012 IECC code reduces this to 0.04 ft\(^3\)/ft\(^2\). All register boots, the connection piece between the duct and the register, must be sealed at a test pressure 25 Pa.

IECC (The International Energy Code Council) was founded in 1998 as the successor organization to IEC (The International Code Council), which was founded 20 years earlier. The vision and mission of these organizations is “to protect the health, safety, and welfare of people by creating better buildings and safer communities,” and “to provide the highest quality codes, standards, products, and services for all concerned with the safety and performance of the built environment.”

In 1994, International Code Council was founded as an association of BOCA (Building Officials and Code Administrators International), ICBO (The International Conference of Building Officials), and SBCCI (The Southern Building Code Congress International, with the mission to develop a single, comprehensive national building code.

A standard test for duct leakages uses a blower door and a pressure pan. The blower door reduces the house pressure by 50 Pa, and the pressure pan is placed over the registers, one at a time, to measure the pressure differential from the house to the duct. The pressure pan has a gasket that seals the pan to the register. Pressure differentials from 1 Pa to 45 Pa are typical, where 1 Pa indicates a tight duct, and 45 Pa indicates a terrible duct. However, since the network of ducts are connected as a single system, determining where the leaks occur may be difficult. Poor workmanship should be identified early so the leaks can be corrected at the time of installation.

There is a considerable amount of guess work in the blower door/pressure pan test, so the duct blower test is preferred. Here, all supplies and returns are sealed with plastic sheeting and tape, and the air handler is turned off. A duct blower is attached to a duct near the furnace and pressurizes the ducts to 25 Pa; 25 Pa is the typical supply side operating pressure. The system is then rated in cfm25, where <50 cfm25 is a tight system, and >500 cfm25 is a very leaky system.

A fogger can be used with the duct blower. With the blower running, the fog nozzle is pointed at the fan blades of the blower, and the locations where the fog comes out are noted.
4.9 Energy Codes and Their Effect on Construction Methods

The world is waking up to the problems of global warming, and in doing so recognizes the value of energy conservation. Buildings consume enormous amounts of energy, and recent changes in building codes are intended to substantially reduce their energy use.

According to the U.S. EIA (Energy Information Administration), in 2014, 41% of total U.S. energy consumption was consumed in residential and commercial buildings, or about 40 quads [http://www.eia.gov/tools/faqs/faq.cfm?id=86&t=1].

“Primary energy consumption in the residential sector totaled 20.99 quads in 2009, equal to 54% of consumption in the buildings sector and 22% of total primary energy consumption in the U.S. [http://buildingsdatabook.eren.doe.gov/ChapterIntro1.aspx].”

“Buildings are responsible for 40% of energy consumption and 36% of CO2 emissions in the EU. While new buildings generally need less than three to five liters of heating oil per square meter per year, older buildings consume about 25 liters on average. Some buildings even require up to 60 liters. Currently, about 35% of the EU’s buildings are over 50 years old. By improving the energy efficiency of buildings, we could reduce total EU energy consumption by 5% to 6% and lower CO2 emissions by about 5% [https://ec.europa.eu/energy/en/topics/energy-efficiency/buildings].”

“Almost 60% of the world’s electricity is consumed in residential and commercial buildings. At the national level, energy use in buildings typically accounts for 20–40% of individual country total final energy use, with the world average being around 30%. Per capita final energy use in buildings in a cold or temperate climate in an affluent country, such as the United States and Canada, can be 5–10 times higher than in warm, low-income regions, such as Africa or Latin America [www.iiasa.ac.at/web/home/research/Flagship-Projects/Global-Energy-Assessment/GEA_Chapter10_buildings_lowres].”

ASHRAE is a leading authority on heat and moisture transfer, and air infiltration in a structure. Its Handbook of Fundamentals is considered by many to be the best source of technical information. Its mission statement is: “To advance the arts and sciences of heating, ventilating, air conditioning and refrigerating to serve humanity and promote a sustainable world.” Its vision statement is: “ASHRAE will be the global leader, the foremost source of technical and educational information, and the primary provider of opportunity for professional growth in the arts and sciences of heating, ventilating, air conditioning and refrigerating.”

ASHRAE Standard 90.1 2010 is 24% to 30% more efficient than 90.1 2004. Expect future versions to continue to improve energy efficiency.
The IECC (International Energy Conservation Code) is contained in a number of documents and is continually being updated. A recent one is “Residential Requirements of the 2009 International Energy Conservation Code; U.S. Department of Energy Building Energy Codes Program.” In it are the minimum resistance requirements in h·°F/BTU for 1 ft$^2$ according to the U.S. climate zones: see Section II of Chapter 9. These values are given in Table 4.5.

The designation 4M in Table 4.5 means maritime region. A small part of the northeastern United States has a maritime climate — part of Long Island and Cape Cod.

An updated set of insulation requirements, version IECC 2012, is given in Table 4.6. The differences between the 2009 and 2012 codes are wall insulation requirements increased in zones 6, 7, and 8, and increased floor insulation for zones 7 and 8.

### Table 4.5: The IECC 2009 Residential Code

<table>
<thead>
<tr>
<th>zone</th>
<th>wall</th>
<th>ceiling</th>
<th>floor</th>
<th>basement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13</td>
<td>30</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>30</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>30</td>
<td>19</td>
<td>5/13</td>
</tr>
<tr>
<td>4</td>
<td>13</td>
<td>38</td>
<td>19</td>
<td>10/13</td>
</tr>
<tr>
<td>4M+5</td>
<td>20 or 13+5</td>
<td>38</td>
<td>30</td>
<td>10/13</td>
</tr>
<tr>
<td>6</td>
<td>20 or 13+5</td>
<td>49</td>
<td>30</td>
<td>15/19</td>
</tr>
<tr>
<td>7&amp;8</td>
<td>21</td>
<td>49</td>
<td>30</td>
<td>15/19</td>
</tr>
</tbody>
</table>

### Table 4.6: The IECC 2012 Residential Code

<table>
<thead>
<tr>
<th>zone</th>
<th>wall</th>
<th>ceiling</th>
<th>floor</th>
<th>basement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13</td>
<td>30</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>38</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>20 or 13+5</td>
<td>38</td>
<td>19</td>
<td>5/13</td>
</tr>
<tr>
<td>4</td>
<td>20 or 13+5</td>
<td>49</td>
<td>19</td>
<td>10/13</td>
</tr>
<tr>
<td>4M+5</td>
<td>20 or 13+5</td>
<td>49</td>
<td>30</td>
<td>15/19</td>
</tr>
<tr>
<td>6</td>
<td>20+5 or 13+10</td>
<td>49</td>
<td>30</td>
<td>15/19</td>
</tr>
<tr>
<td>7&amp;8</td>
<td>20+5 or 13+10</td>
<td>49</td>
<td>38</td>
<td>15/19</td>
</tr>
</tbody>
</table>

Some interpretation of these numbers is needed. The numbers for the sidewall are for a stud-framed structure with cavity insulation only. For zones 3 and above there are numbers for the standard stud-framed with additional continuous foam-type insulation or insulated siding. For zones 3 through 5/4M the insulation should be at least 13+5, meaning R13 for the cavity wall plus R5 for the continuous insulation. For zones 6, 7, and 8 these numbers are 13+10.

The evolving building codes will continue increasing the R factors of building components, making the use of fiberglass insulation in a cavity wall problematic. 2x4 or 2x6 stud walls are acceptable, but 2x8 walls may create an appearance problem for the frames around windows.
and doors. It appears likely that foam-filled walls will become standard in stud walls in the future.

Fiberglass insulation has an assumed resistivity of r3.3, so a cavity wall with R20 would need close to 6” of fiberglass. This is an awkward size, bigger than achieved with a 2×6 stud. It would be achieved in a 2x4 stud wall with polyurethane foam at r6. A ceiling at R49 would require over 15” of fiberglass; a 2×12 joist is 11.25” deep, and larger joists are difficult to find and are expensive. An R49 ceiling insulated with polyurethane foam would require 2×10 joists, which are 9.25” deep.

There is a modification for the floor insulation: an alternative to the stated R value is “insulation sufficient to fill the framing cavity,” but there is a major problem with this. To construct a wood-frame floor, the joists are put in place, then plywood is laid on top, possibly with polyethylene sheeting above the joists but below the plywood. After the house is framed, the rough plumbing comes next, followed by the rough electric. Only then is any insulation installed, but how can this be done with fiberglass batts or blown-in cellulose?

<table>
<thead>
<tr>
<th>country</th>
<th>wall</th>
<th>roof</th>
<th>floor</th>
<th>window</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>0.28</td>
<td>0.2</td>
<td>0.28</td>
<td>1.3</td>
</tr>
<tr>
<td>England/Wales</td>
<td>0.3</td>
<td>0.2</td>
<td>0.25</td>
<td>2</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.18</td>
<td>0.13</td>
<td>0.15</td>
<td>1.3</td>
</tr>
<tr>
<td>Holland</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Cellulose works well on attic floors after the ceiling below is in place and before any boards or plywood are laid in the attic, but the only way to do this on a lower floor is to put a ceiling on the unconditioned space below the floor, and then drill holes in the plywood above the joists to facilitate the blowing-in process. One cannot use craft-faced fiberglass batts in the floor since the vapor barrier craft will be on the unheated side, and this will cause moisture buildup in the fiberglass. The most viable method is expanding foam.

In the IECC 2012 is the requirement that new wood-burning fireplaces shall have gasketed doors and outdoor combustion air. This makes the location of a wood stove likely to be on an outside wall. The code also limits air infiltration to ACH50 ≤ 5 in climate zones 1–2, and ACH50 ≤ 3 in climate zones 3–8. Recall at ACH50 the air infiltration is about 20 times natural infiltration, so natural infiltration in climate zones 3-8 must be less than 0.15 air changes an hour. This will be very difficult to achieve.

Also in the IECC 2012 is a requirement that the air leakage around a window, sliding glass doors, or skylight is less than 0.3 cfm/ft², and around a swinging doors is less than 0.5 cfm/ft².
“The Global Buildings Performance Network (GBPN) is a globally organized and regionally focused organization whose mission is to provide policy expertise and technical assistance to advance building energy performance and realize sustainable built environments for all. . . . The GBPN was founded in 2010 with the mandate to advance knowledge and expertise globally on building energy performance and the structure to achieve it [www.gbpn.org/databases-tools/bc-detail-pages/germany].”

Extracted from the GBPN document are the maximum conductance numbers shown in 4.7 SI units of W/m²·K for a number of European countries. To compare these numbers $C_{si}$ to the Imperial resistance $R_{am}$, we use

$$R_{am} = 6.933C_{si}, (4.1)$$

so the wall resistance in Germany is $R_{am} = 24.8$, in England/Wales is $R_{am} = 23.1$, and for Sweden is $R_{am} = 38.5$. These numbers are fairly similar to the Imperial ones. Holland is the outlier with $R_{am} = 17.3$, as it is with all the measures. For roofs, Germany is $R_{am} = 34.7$, England/Wales is the same, and Sweden is $R_{am} = 53.3$.

We have seen the United States is wedded to thermal resistance, whereas Europe is wedded to thermal conductance, at least as far as walls and ceilings are concerned. As far as windows are concerned, both use thermal conductance, where

$$C_{am} = 0.1442C_{si}, (4.2)$$

IECC 2012 requires the maximum conductance for windows in climate zone 2 to be $C_{am} = 0.4$, in zones 3–4 $C_{am} = 0.35$, in zones 4M–8 $C_{am} = 0.32$. The corresponding numbers are Germany/Sweden $C_{si} = 1.3$ and England/Wales $C_{si} = 2.0$

### 4.10 Errors in Energy Codes

In recent years the energy codes of buildings have been substantially tightened. The vast majority of these changes make sense and are beneficial, but some errors have crept in. One such error is in the requirement that floors have a substantially higher resistance to heat loss than walls.

One column in Table 4.5 gives the required resistance values for a floor as R19 in climate zones 3 and 4, and R30 in climate zone 4M though 8. These are more stringent than for a wall and make no sense. Heat rises, not falls. Further, the interior surface resistance for heat flow downwards as given by Table 3.2 of Chapter 3 is 0.17 m²·K/W, but the resistance sideways is 0.13 and upwards is 0.10. Although the reduction in heat loss due to the interior surface resistance is small, it is indicative of the relative overall heat flow.
It is interesting that the 2009 *ASHRAE Fundamentals* makes no mention of heat loss through floors. It does cover the temperature of a floor for comfort, and discusses heat loss from a slab, but nary a thing on heat loss through a floor with unconditioned space below.

Air infiltration through a floor with unconditioned space below is a concern, particularly if that unconditioned space is a garage; garage doors are notoriously leaky. The solution to this problem is to lay polyethylene sheeting on top of the joists before the plywood is laid.

U.S. climate zones 4–8 are predominantly heating areas, unlike zones 1–2. In zones 1–2 a house built on or near wetlands may be built on stilts, so there is uncontrolled flow of air beneath the floor. Here, in a predominantly cooling environment, floor insulation is needed for cooling purposes. No one would think of building a house, designed for year-round occupancy, on stilts in Minnesota or North Dakota. Building codes prohibit it. So why specify floor insulation in zones 4M–7 at R30 when wall insulation is only R20?

### Table 4.8: Floor Resistances in the U.S.

<table>
<thead>
<tr>
<th>zone</th>
<th>IECC</th>
<th>NY</th>
<th>NC</th>
<th>CO</th>
<th>MT</th>
<th>IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

### Table 4.9: Fenestration Numbers

<table>
<thead>
<tr>
<th>zone</th>
<th>2009 U</th>
<th>2013 U</th>
<th>SHGC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>≤6.8</td>
<td>≤2.8</td>
<td>≤0.3</td>
</tr>
<tr>
<td>2</td>
<td>≤3.7</td>
<td>≤2.3</td>
<td>≤0.3</td>
</tr>
<tr>
<td>3</td>
<td>≤2.8</td>
<td>≤2.0</td>
<td>≤0.3</td>
</tr>
<tr>
<td>4</td>
<td>≤2.0</td>
<td>≤2.0</td>
<td>-</td>
</tr>
<tr>
<td>5-8</td>
<td>≤2.0</td>
<td>≤1.8</td>
<td>-</td>
</tr>
</tbody>
</table>

If the IECC was the only authority with ridiculous requirements for wood frame floor insulation, one could cite other authorities, but these “other authorities” follow the same path as IECC “as can be seen in” Table 4.8. The U.S. states are New York (NY), North Carolina (NC), Colorado (CO), Montana (MT) and Iowa (IO). These were not selected to make the point, but were the only ones found. No other state of authority exists, as far as could be determined, to omit wood frame floor insulation requirements.

The GBPN document echoes the error made in the IECC 2009 and IECC 2012 by imposing stringent requirements on the maximum $C_{si}$ for floors [www.gbpn.org/databases-tools/bc-detail-pages/germany].

At last, an article critical of the IECC mandates on fenestration has appeared. The REL (Residential Energy Laboratory) points out the errors in IECC 2009 and IECC 2012
Energy codes that reduce energy use are good, but codes that do the opposite and are possibly designed for the advantage of window manufacturers are bad.

“The IECC, which might be mostly driven by the U.S. Dept. of Energy (DOE) and window manufacturers, has come up with the code requirements for fenestration ... for IECC 2009 and IECC 2012 [op cit].” The U-factor in W/m²·K and $S_{HGC}$, the solar heat gain coefficient according to the IECC 2009 and 2012 fenestration codes, are shown in Table 4.9.

<table>
<thead>
<tr>
<th></th>
<th>U-factor</th>
<th>SHGC</th>
</tr>
</thead>
<tbody>
<tr>
<td>low SHGC</td>
<td>1.9</td>
<td>0.30</td>
</tr>
<tr>
<td>high SHGC</td>
<td>2.1</td>
<td>0.53</td>
</tr>
</tbody>
</table>

“Some typical window thermal properties for high and low solar gain double-pane, low-E windows were taken from the DOE LBNL (Lawrence Berkeley National Laboratory) computer model RESFEN (RESidential FENestration), and were labeled as typical for wood or vinyl-frame windows [op cit].” These numbers, given in Table 4.10, are for air filled units, and show the windows with high SHGC do not satisfy the U-factor of 2.0 required in the IECC code. Argon filled units can satisfy the new code, but they are expensive and over the service life of the unit the argon often leaks out, negating their benefit.

Notice that the low solar gain windows meet the specification for northern U.S. latitudes of having U-factors less than 2.0 W/m²·K, but the high solar gain windows do not meet the 2009 or 2012 standards.

“The high solar gain windows let in 77% more solar energy, and conduct only 9% more energy out of the house, but are forbidden by code to be used in northern latitudes where they might be very helpful in providing passive solar heating for homes [op cit].” There is an onerous appeals procedure to allow the use of high SHGC windows if the owner/builder can show offsetting savings elsewhere, but what is needed is a single entity to prove the error of the code and have it modified.

In their concluding remarks, REL opined, “... in some cases, the IECC code requirements for windows result in the use of low solar gain windows that waste potential energy savings from passive solar heating. These codes should be changed to allow maximum U-values of 2.3 W/m²·K in climate zones 4 through 8 so that air-filled, double-pane, high solar gain windows can be used in these zones.”

Unfortunately, the REL has errred in saying that “the use of high solar gain windows would result in more solar energy entering the home, but an increase in heat losses from the home. During the winter, the extra solar energy would be an advantage, but during the summer, it
would be a disadvantage.” South facing glass allows little if any solar energy to pass through in summer, particularly if shading is used; see Section III of Chapter 11. There is almost no solar energy above the wavelength of 2.7μ, and window glass is essentially opaque at wavelengths above 2.7μ; see Section II of Chapter 8. It is possible that long wavelength radiation will be reflected more efficiently from a low-E surface than from clear glass, but this is a minor factor in the energy balance of a residence. More importantly, a passive solar structure needs a high SHGC, as recognized by the Passivhaus standards.

### 4.11 Insulation Guidelines and Their Influence on Construction

Let us consider what these numbers translate to as far as the construction is concerned for zones 3 and colder. In Section IX we calculated the resistance of a 2x4 stud wall with fiberglass insulation as $R = 14.27 \text{ ft}^2\cdot\text{hr}\cdot\text{°F}/\text{BTU}$, so this does not satisfy code. 2x4 stud wall with foam has $R = 22.35 \text{ ft}^2\cdot\text{hr}\cdot\text{°F}/\text{BTU}$. A 2x6 stud wall with fiberglass insulation has $R = 22.49 \text{ ft}^2\cdot\text{hr}\cdot\text{°F}/\text{BTU}$, so this also satisfies code.

The burning question is whether to use foam or fiberglass in the sidewall. There are advantages and disadvantages to each. The big plus for foam is increased R values. The two minuses are cost and inflexibility — once the foam is installed, it is virtually impossible to feed electrical romex cable through it. The interior wallboard must be removed, and this is typically done when the feed is vertical; this is almost always less disruptive than removing the siding and sheathing. An alternative is to plan for future electrical expansion by setting electrical conduit in the walls. Further, the differential between foam and fiberglass is diminished by the substantial losses through the windows.

Closed cell foam insulation prohibits water vapor migration, but there remains the exposed studs that can easily wick water. Again, an interior vapor barrier is recommended. However, consideration should be given to ventilation requirements when constructing a tight house, particularly if the interior volume is relatively small and the occupancy (number of residents) is relatively large.

The numbers for floor loss are only appropriate for structures that permit air flow beneath the floor. These include temporary structures, or structures on stilts, such as in tidal areas or areas with a high water table. However, an unheated garage with heated living space above can result in substantial heat loss. The first thing is to use a vapor barrier to substantially reduce or eliminate air infiltration, then insulate with fiberglass or foam; the vapor barrier should be on the upper heated side of the insulation, and the easiest way to do this is to lay it over the joists and under the plywood subfloor. Heat flow downwards, in the absence of air flow, is far less than heat flow upwards, so a little insulation goes a long way.
ASHRAE Fundamentals gives the zone of floor temperatures for comfort, with the upper limit being $28^\circ C \equiv 82^\circ F$. The lower limit depends on the floor material as follows:

- textiles $21^\circ C$
- pine floor $22.5^\circ C$
- oak floor $24.5^\circ C$
- linoleum $24^\circ C$
- concrete $26^\circ C$

A comparison of the minimum R numbers for windows and sidewalls in zone 5/4M are 3 and 20 $ft^2 \cdot hr \cdot ^\circ F/\text{BTU}$, respectively. The guideline for the maximum percentage of fenestration to sidewall in zone 5/4M is 35%. There are also codes that dictate a minimum percentage, where 15% is a typical figure. At an area of 15% the window losses versus frame sidewall losses for an area of 100 ft $153 = 5$ and $8520 = 4.25 \text{ BTU/hr} \cdot ^\circ F$, respectively.
Chapter 5

The Seasons and Solar Angles

This chapter begins with a discussion of the seasons, which leads into a discussion of time and the definition of longitude noon. We will follow the standard practice of relying on the Julian calendar to define a day of the year.

Longitude noon is defined with respect to longitude and Greenwich; longitude zero, known as the prime meridian, passes through the Royal Observatory in Greenwich, England. The orientation of the axis of rotation of the Earth, called the obliquity vector, is derived, and forms the basis for the position of the Sun as a function of time of day, Julian day, and declination.

The Earth rotates around the Sun in what is called the ecliptic plane. Most of the planets move in essentially the same the ecliptic plane, suggesting that the solar system was formed from a disc of material. Mercury’s orbit is 7° off the ecliptic plane, and Pluto is 17° off.

Is Pluto a planet? At the 2006 International Space Union Conference it was decided that in order to be a planet it must orbit the Sun, be spherical, and be the biggest body in its orbit so that it can knock space rocks out of its orbit. Pluto fails that last requirement, so it was reclassified as a “dwarf planet.” The negative response to this demotion was especially intense with school children.

It is well known that sundials do not read the same time as a clock, a difference known as EoT, the “Equation of Time.” Here, an analysis reveals an intimate bond between the declination and EoT. This analysis may be valuable to some researchers, but the deviation from standard tables is potentially troublesome.

In the study of solar azimuth and elevation, declination angle and EoT are typically made available in tabular form. What is the basis for these charts? Is it celestial observation, or is it calculation?

The work of Johannes Kepler is fundamental to an understanding of the motion of the Earth around the Sun, and this leads to a numerical method for calculating declination. It also leads
to a method for calculating the EoT, but the difference between the numbers resulting and the commonly accepted EoT table is disturbing.

The astronomical terms used in this chapter can be found in the frontpiece materials of this book. It would be well to review them before proceeding further into this chapter.

5.1 Obliquity and the Seasons

There are seasons of the year due to the angle of obliquity: 23.45 degrees. The result is that the circle on the Earth with the Sun directly overhead varies from the Tropic of Capricorn at latitude 23.45 degrees south to the Tropic of Cancer at latitude 23.45 degrees north. The angle presented by the Earth to the Sun is called the angle of declination (δ).

The situation on December 21, the date of winter solstice, when δ = −23.45, is shown in Figure 5.1. The Earth rotates over a day while the sunbeams remain constant. There are several features to notice on this December 21 day:
* The locus of point on the Earth’s surface directly facing the Sun at solar noon over the complete day is the Tropic of Capricorn.
* The Sun’s rays just touch the edge of the arctic circle, 66.55°N.
* The Sun’s rays cover the complete antarctic circle, 66.55°S.
The situation is reversed at the time of the summer solstice, June 21, when $\delta = 23.45$. Here, the Sun’s rays are over the Tropic of Cancer and the northern latitudes, north of the Tropic of Cancer, are in the middle of their summer.

The region between the Tropic of Capricorn and the Tropic of Cancer is termed the tropics, and this region is considered to have no seasons. The locus of the point on the Earth with the Sun vertically overhead is a latitude between the tropics.

At times of equinox, March 22 and September 22, when $\delta = 0$, the equator is directly facing the Sun, as shown in Figure 5.2. At these dates every point in the surface of the Earth (idealized to remove mountains and valleys) receives 12 hours of sunlight a day. Between the Fall equinox and the Spring equinox, the days of sunlight are shorter north of the equator and longer south of the equator.

5.2 The Importance of Time

When humans evolved from hunter/gatherers to farmers, the need to gauge the seasons became vitally important. In the Nile valley of Egypt, the ability to predict the yearly flooding of fields,
bringing nutrients to the crops, was a valuable asset. The priests had the education, rudimentary though it was, and more importantly the free time (free that is from the daily drudgery of the rest of the population) to observe and record. They became a powerful force in that society [H.W. Smith, Man and His Gods, Little Brown and Company, Boston, 1952]. This is now an e-book that can be accessed online at www.positiveatheism.org/hist/homer1a, together with a forward by Albert Einstein].

![Water Clock Diagram]

**Figure 5.3: The Water Clock**

The first clock with moving parts was the water clock. About –270 BC the Greek Ktesibios invented a float regulator, a necessary part for the water clock. The function of this regulator was to keep the water level in a tank at a constant depth. This constant depth yielded a constant flow of water through a tube at the bottom of the tank, which filled a second tank at a constant rate. The level of water in the second tank thus depended on time elapsed. The regulator of Ktesibios used a float to control the inflow of water through a valve; as the level of water fell the valve opened and replenished the reservoir. This float regulator performed the same function as the ball and cock in a modern flush toilet. These clocks were not accurate, but it
took until the middle ages before pendulum-driven mechanical clocks with intricate mechanisms were invented. These early pendulum clocks were not constructed to tight tolerances.

The industrial revolution produced low-cost, high-accuracy timepieces — see the next section of this chapter on the clock that solved the navigational problem of determining longitude.

Greenwich is at longitude 0° (by definition), and the world’s time is often quoted with respect to Greenwich Standard Time or Greenwich Mean Time (GMT). In an effort to make time seem more universal and less national, Coordinated Universal Time (CUT) was proposed to replace GMT. This was countered by the French Temps Universal Cordonné (TUC). A compromise was struck to use the acronym UTC, which since 1972 has been the standard for defining time zones worldwide.

Most time zones are integral deviations from UTC. For example, the United States has five time zones, Eastern Standard Time, 5 hours after UTC, or −5, Central Standard Time −6, Mountain Standard Time, −7, Pacific Standard Time, −8, and Alaska Standard Time, −9. A country ahead of UTC such as Turkey is +2.

An important, but non-physical, entity is the International Date Line. It travels from the north to the south pole near (and sometimes on) longitude 180. It passes through the Bering Straight between Russia and Alaska, due south on longitude 180 to the equator, then meanders eastwards and for a short distance northwards before moving westwards closer to longitude 180, then south once more between American Samoa to its east and Tonga to its west before skirting New Zealand to its west and arriving at the south pole.

On the positive side of the international date line, such as in the United States, UTC times are never greater than 12, and on the negative side of the international date line, such as in Turkey, UTC times are never less than −12.

5.3 Longitude

Astrophysicists have produced highly accurate tables and numbers to assist in navigation. Using hand-held instruments to make celestial measurements, and with the assistance of charts and tables, mariners could determine with satisfactory accuracy the latitude of the vessel. However, it is interesting that the solution to the problem of determining longitude was, until relatively recently, intractable. In 1714 the English Parliament offered an award, worth about $12 million today, for the first person to produce a device or method to crack the problem; in particular, to determine longitude within half a degree, equivalent to 60 nautical miles. It took an unschooled craftsman called John Harrison, 1693–1776, to provide the solution — an accurate determination of time, within 3 seconds per day, which adds up to 2 minutes over a 40-day journey. He delivered his clock called H1 to the Royal Society, and it had its first
official trial in 1736, which it passed [D. Sobel, *Longitude*, Walker & Co., New York, 1995]. He built a smaller and more robust version, called H2, in 1741. It took him until 1773 to claim his financial reward, denied him for so long by class prejudice, but he never officially won the competition.

The original purpose of ephemerides tables was to assist mariners in the determination of longitude; the singular of ephemerides is ephemeris, and it is now common to use the word ephemeris as both singular and plural. The first ephemeris tables were published by Connaissance de Temps in France in 1679. The Royal Observatory at Greenwich, started the *The Nautical Almanac* and *Astronomical Ephemeris*, now referred to as the *Nautical Almanac*, in 1767, and it continues to this day. Since Britannia ruled the waves in the seventeenth century, Britain defined longitude zero as passing through the Royal Observatory at Greenwich: This is known as the prime meridian. In 1852 the U.S. Naval Observatory began publishing *American Ephemeris and Nautical Almanac*, with its dominant meridian passing through Washington, D.C. The British and American publications combined in 1981 under the name, *The Astronomical Almanac*, using the Greenwich prime meridian.

### 5.4 The Julian Calendar

The Sun and the Earth’s orbit lie in a plane, termed the ecliptic plane. Without loss of generality we can assume that this is the x-y plane $p=[001]$, and the Sun lies on the positive x-axis close to the origin $[000]$. The angle of obliquity $23.45^\circ$ for the Earth is defined as the angle between the Earth’s rotational axis $q$ and $p$; we term $q$ the obliquity vector, which is calculated in Section VI. This angle changes slowly over time, by $50.3^\prime$ per year, a phenomenon called the precession of the equinoxes. A complete cycle takes 25,765 years. As far as our human experience is concerned, $q$ is fixed.

The Julian calendar marks the days of the year after January 1, known as Julian day $J_1$, so February 1 would be $J_{32}$. The Earth’s perihelion, when it is closest to the Sun, is commonly taken to occur on $J_3$ (this value may be suspect, as the analysis of Section VIII indicates), and the winter solstice occurs on about Julian day $-11.3$ or $J_{355}$.

The orbit of the Earth around the Sun is shown in Figure 5.4. The Earth is shown at four significant times of the year, the solstices and equinoxes, plus at $J_3$. The eccentricity is exaggerated. Each depiction of the Earth shows an arrow on the split circle indicating the daily rotation about its axis, and an arrow facing the Sun indicating solar noon. The perihelion occurs on the positive x-axis, and the aphelion occurs on the negative x-axis.

Also shown in Figure 5.4 are the days between solstices and equinoxes. This indicates that the angular velocity of the center of the Earth changes, and is higher on the section of the orbit
nearer the winter solstice than at the section nearer the summer solstice. This, we will see, is a consequence of adherence to Kepler’s second law; see the next section.

Figure 5.4: Julian Days on the Earth’s Orbit

Table 5.1: First of the Month Julian Days

<table>
<thead>
<tr>
<th>Month</th>
<th>Julian Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 1</td>
<td>J_1</td>
</tr>
<tr>
<td>February 1</td>
<td>J_32</td>
</tr>
<tr>
<td>March 1</td>
<td>J_60</td>
</tr>
<tr>
<td>April 1</td>
<td>J_91</td>
</tr>
<tr>
<td>May 1</td>
<td>J_121</td>
</tr>
<tr>
<td>June 1</td>
<td>J_152</td>
</tr>
<tr>
<td>July 1</td>
<td>J_182</td>
</tr>
<tr>
<td>August 1</td>
<td>J_213</td>
</tr>
<tr>
<td>September 1</td>
<td>J_244</td>
</tr>
<tr>
<td>October 1</td>
<td>J_274</td>
</tr>
<tr>
<td>November 1</td>
<td>J_305</td>
</tr>
<tr>
<td>December 1</td>
<td>J_335</td>
</tr>
<tr>
<td>January 1</td>
<td>J_366</td>
</tr>
</tbody>
</table>
The number of days from summer solstice to winter solstice is 184, and from winter solstice to summer solstice 182. The number of days from vernal to autumnal equinoxes is 186, but from autumnal to vernal is 180, a considerable difference indicating increased velocity along the orbit near the perihelion, a phenomenon classified by Kepler’s second law, discussed in the next section.

Since there are no fractional Julian days, it is evident that rounding errors in the Julian calendar are as much as half a day. The first-day-of-the-month Julian days is given in Table 5.1 for a non-leap year. The significant dates (solstices, equinoxes, perihelion, and aphelion) are given in Table 5.2, where tbd means “to be determined.” We will refer to $\text{j}_n=360.365.24\text{J}_n=0.9857\text{J}_n$ as the Julian angle, so at winter solstice $J_n = 355, j_n = 0.9857 \times 355 = 349.9 - 10.09^0$; the average Earth year is 365.2421897.

There is another built-in error when using Julian days’ numbers caused by leap years. Further, it is as likely that the autumnal equinox will occur on September 23, $J_{267}$, as on $J_{266}$. The other significant dates could be a day off (March 21, June 20, December 22), but the most likely dates are the ones given.

A leap year occurs whenever the year is divisible by 4 (no fraction), except when that year is divisible by 100, and the exception to this modification is when the year is divisible by 400. So, the year 2000 is a leap year but 1900 is not. This sets the average year to be 365+14=1100+1400=365.2425 days long, close to the average Earth year.

### 5.5 Geometry of the Ellipse and Kepler

The Earth’s orbit is elliptical, so its motion with respect to the Sun is governed by the first two of Kepler’s laws:
1. The Sun is at a focus of the ellipse, and
2. The area swept by the Earth to the focus over a given time is the same regardless of position on the orbit.
Kepler’s second law is illustrated in Figure 5.5, where the arcs on the orbits represent equal time intervals, and area $A_1 = A_2$.

Consider the ellipse shown in Figure 5.6 with major axis $a$ and minor axis $b$, so

$$x^2a^2+y^2b^2=1.\quad (5.1)$$
The two focal points are shown, and \( a^2 - b^2 = e^2, r_1 + r_2 = 2a \). \( e \) is the distance from the center of the ellipse to a focal point. Eccentricity of an ellipse is defined as \( e = e/a \). The normalized distances of the Earth from the Sun range from its perihelion \( \beta \) (its closest distance) of 0.9832622, which occurs about \( J_3 \), to its aphelion \( \alpha \) (its furthest distance) of 1.016619, which occurs about \( J_{187} \) [The American Ephemeris and Nautical Almanac, U.S. Gov. Printing Office, 1977]. These numbers produce \( e = \alpha - \beta = 0.01669985, a = \alpha + \beta = 0.99996205 \) and \( e = ea = 0.01667841.016619 = 0.01670048 \), so \( b^2 = 0.9996452 \) and \( b = 0.9998225925 \).

When the Earth is at \( y = 0 \), so the Earth is on the major axis \( x \) of its elliptical orbit, it is at its perihelion or its aphelion.

In the triangle shown in Figure 5.6 with sides \( r_1, r_2 = 2a - r_1, \) and \( 2e \), if \( \psi \) is known then \( r_1 \) can be determined from the law of cosines:

\[
 r_2 = r_1 + 4e^2 - 4r_1 \cos \psi, \quad \text{so} \quad (2a - r_1) = r_1 + 4e^2 - 4r_1 \cos \psi = r_1 + 4e^2 - 4r_1 \cos \phi, \quad \text{giving} \quad e^2 - a^2 + ar_1 + r_1e \cos \phi = 0, \quad \text{and}
\]

\[
 r_1 = a^2 - e^2a + e \cos \phi = b^2a + e \cos \phi \quad (5.2)
\]

The orbit angle of the Earth’s position with respect to the Sun is \( \phi \), where \( \phi = 0 \) at perihelion, and \( \phi = 180 \) at aphelion. The orbit angle at vernal equinox will be less than 90° since the angular velocities of the Earth for the period of time between perihelion and vernal equinox is greater than the angular velocities for the period between the equinoxes.

The Earth year is 365 days, 5 hours, 48 minutes and 46 seconds long, or 365.242199 days. If the complete ellipse is broken down into incremental segments \( \Delta \phi_i = \phi_{i+1} - \phi_i, \) \( i = 0, 1, \ldots 365 \) of equal area \( A \) that covers the range \( 0 \leq \phi \leq 360(1 - 0.242199365.242199) = 359.761277 \); the year
is 365 days but the orbit in a non-leap year is less than a full orbit to account for the partial day of 0.242199.

The area of an ellipse is $\pi ab$, so the area covered in one day is

$$ A_i = \pi ab 365.242199. \quad (5.3) $$

For the geometry shown in Figure 5.7,

$$ A_i \pi 365 \Delta \phi i r_i r_i + 1 \quad (5.4) $$

where $r_i = b^2 a + e \cos \phi_i$ and $r_i + 1 = b^2 a + e \cos \phi_i + 1$, so $\pi ab 365.242199 \pi 365 b^3 \Delta \phi i a + e \cos \phi i (a + e \cos \phi_i)$ (a + e cos $\phi_i + 1$), or $1 = 1.000663559 b^3 \Delta \phi i a + e \cos \phi_i (a + e \cos \phi_i + 1)$. The term $(a + e \cos \phi_i + 1)$ can be written explicitly as $a + e \cos \phi_i + 1 = 1.000663559 b^3 \Delta \phi i a + e \cos \phi_i$, so

$$ \cos \phi_i + 1 = f a + e \cos \phi_i \Delta \phi_i - ae \quad (5.5) $$

where $f = 1.000663559 b a e$. Expanding the left-hand side using the Taylor series produces

$$ \cos \phi_i + 1 = \cos(\phi_i + \Delta \phi_i) = \cos \phi_i \cos \Delta \phi_i - \sin \phi_i \sin \Delta \phi_i = \cos \phi_i (1 - \Delta \phi_i 22! + \Delta \phi_i 44! + \ldots) - \sin \phi_i (\Delta \phi_i - \Delta \phi_i 33! + \Delta \phi_i 55! + \ldots) \cos \phi_i - \sin \phi_i \Delta \phi_i, $$

so \( \cos \phi_i - \sin \phi_i \Delta \phi_i f a + e \cos \phi_i \Delta \phi_i - a e \) and

$$ \Delta \phi i \cos \phi_i + a e \sin \phi_i f a + e \cos \phi_i. \quad (5.6) $$

Using MATLAB to perform the calculations for the complete year produced $\phi_{365} = 359.7532^\circ$, an insignificant error of $0.00808^\circ = 29''$; the first order approximation will be used in subsequent calculations.

### 5.6 Points, Vectors, and Planes

A point is $u = [xyz]$, and a vector is $v = [efg]$, so they can be confused with each other. To avoid this we can use what is known as the homogeneous representation $U = [xyz1]$ and $V = [efg0]$ for a point and a vector, respectively. The lower case is used for the non-homogeneous and upper case for the homogeneous representation.

A point has a location but no size. A vector has length and direction but no location. We can add vectors, and the result is a vector. We can add a vector to a point, and the result is a point. It makes no sense to add two points, resulting in the fourth term being 2, but we can find the average of two points as a point.
The most common definition for a line is in terms of two points and a scalar parameter. \( \mathbf{u}(\alpha) = \mathbf{u}_1 + \alpha (\mathbf{u}_2 - \mathbf{u}_1) \) is a line as well as a parametric point in terms of parameter \( \alpha \); when \( \alpha = 0 \) the parametric point is at \( \mathbf{u}_1 \), and when \( \alpha = 1 \) the parametric point is at \( \mathbf{u}_2 \).

EXAMPLE: The line joining the points \( u_1 = [124] \) and \( u_2 = [2-11] \) is \( \mathbf{u}(\alpha) = u_1 + \alpha (u_2 - u_1) = [124] + \alpha [2-11] - [124] = [124] + \alpha [1-3-3] = [1+\alpha 2-3\alpha 4-3\alpha] \)

Given two vectors \( \mathbf{v}_1 = [e1f1g1] \) and \( \mathbf{v}_2 = [e2f2g2] \), the scalar or dot product is \( \mathbf{v}_1 \cdot \mathbf{v}_2 = [e1f1g1] \cdot [e2f2g2] = e1e2+f1f2+g1g2 \) and the vector or cross product is \( \mathbf{v}_1 \times \mathbf{v}_2 = [e1f1g1] \times [e2f2g2] = [f1g2-g1f2g1e2-e1g2e1f2-f1e2] \).

Consider the four vectors \( \mathbf{v}_i, i = 1,\ldots, 4 \) then

\[
\begin{align*}
\mathbf{v}_1 \cdot \mathbf{v}_1 &= |\mathbf{v}_1|^2, \\
\mathbf{v}_1 \times \mathbf{v}_1 &= 0, \\
\mathbf{v}_1 \cdot \mathbf{v}_2 &= \mathbf{v}_2 \cdot \mathbf{v}_1 = |\mathbf{v}_1||\mathbf{v}_2|\cos \theta, \\
|\mathbf{v}_1 \times \mathbf{v}_2| &= |\mathbf{v}_1||\mathbf{v}_2| \sin \theta, \\
\mathbf{v}_1 \times \mathbf{v}_2 &= \mathbf{v}_2 \times \mathbf{v}_1, \\
(\mathbf{v}_1 + \mathbf{v}_2) \cdot \mathbf{v}_3 &= \mathbf{v}_1 \cdot \mathbf{v}_3 + \mathbf{v}_2 \cdot \mathbf{v}_3, \\
(\mathbf{v}_1 + \mathbf{v}_2) \times \mathbf{v}_3 &= \mathbf{v}_1 \times \mathbf{v}_3 + \mathbf{v}_2 \times \mathbf{v}_3, \\
\mathbf{v}_1 \cdot (\mathbf{v}_2 \times \mathbf{v}_3) &= \mathbf{v}_1 \times (\mathbf{v}_2 \times \mathbf{v}_3), \\
(\mathbf{v}_1 \cdot \mathbf{v}_3) \mathbf{v}_2 - (\mathbf{v}_1 \cdot \mathbf{v}_2) \mathbf{v}_3 &= (\mathbf{v}_1 \times \mathbf{v}_2) \times \mathbf{v}_3 - (\mathbf{v}_1 \times \mathbf{v}_3) \mathbf{v}_2 - (\mathbf{v}_2 \times \mathbf{v}_3) \mathbf{v}_1 
\end{align*}
\]

where \( \theta \) is the angle between vectors \( \mathbf{v}_1 \) and \( \mathbf{v}_2 \).

Vector or cross multiplication has precedence over scalar or dot multiplication, so \( \mathbf{v}_1 \times \mathbf{v}_2 \cdot \mathbf{v}_3 \equiv (\mathbf{v}_1 \times \mathbf{v}_2) \cdot \mathbf{v}_3 \). The construction \( \mathbf{v}_1 \times \mathbf{v}_2 \times \mathbf{v}_3 \) cannot be interpreted, so it is incorrect.

The cross product of two vectors \( \mathbf{p} = \mathbf{v}_1 \times \mathbf{v}_2 \) is said to define the plane that contains them, and this resultant vector is orthogonal to the original two vectors.

A plane is a sheet of infinite extent in two dimensions and so is impossible to draw, except on edge, when it looks like a line as shown in Figure 5.8, where \( \mathbf{p} = [abc] \) partially defines a plane — it gives its orientation but not its position. To determine its position one uses the
homogeneous representation $\mathbf{P} = [a\ b\ c\ d]$. Point $\mathbf{U}$ lies on plane $\mathbf{P}$ if $\mathbf{P}\mathbf{U} = 0$. Thus point $[2311]$ is on plane $[2\ 3\ 3\ 2]$ since $[2-332][2311]=0$.

The distance of a plane $\mathbf{P} = [a\ b\ c\ d]$ from the origin can be determined by noting that vector $\mathbf{p}=[\mathbf{a}\mathbf{b}\mathbf{c}]$ from the origin intersects $\mathbf{P}$ at the minimum distance from the origin, so $[\mathbf{a}\mathbf{b}\mathbf{c}\mathbf{d}]$ 
$[\alpha\alpha\beta\alpha\gamma]=\alpha(a2+b2+c2)+d=0,,\ and$

$$\alpha=-da2+b2+c2(5.7)$$

and the distance from the origin is $\alpha a2+b2+c2=da2+b2+c2$.

If $\mathbf{p}=[abc]$, is a unit vector, so $\mathbf{p}\cdot\mathbf{p}=[abc][abc]=a2+b2+c2=1$, then the distance of $\mathbf{P}$ from the origin is $d$, and we refer to the plane as being normalized.

The successive cross products of $\mathbf{v}$, which lies in $\mathbf{P}$ as shown in Figure 5.9, are instructive. $\mathbf{p} \times \mathbf{v}$ lies in $\mathbf{P}$ and is orthogonal to $\mathbf{v}$; We use the right hand rule – curl the fingers from the first to the second vector over the smaller angle and the thumb points in the direction of the resultant vector. $\mathbf{p}\times(\mathbf{p}\times\mathbf{v}) = -\mathbf{v}, \mathbf{p}\times(\mathbf{p}\times(\mathbf{p}\times\mathbf{v})) = \mathbf{p} \times (-\mathbf{v}) = -\mathbf{p} \times \mathbf{v}$, and $\mathbf{p} \times (\mathbf{p} \times (\mathbf{p} \times (\mathbf{p} \times \mathbf{v}))) = \mathbf{p} \times (-\mathbf{p} \times \mathbf{v}) = \mathbf{v}$.

EXAMPLE: Rotate vector $\mathbf{v}1=[100]$ successively by $90^\circ$ in the $x-y$ plane.

$v2=\mathbf{p}\times\mathbf{v}1[001] \times [100]=[010], v3=\mathbf{p}\times\mathbf{v}2=[001] \times [010]=[-100]=v1, v4=\mathbf{p}\times\mathbf{v}3=[001] \times [-100]=[0-10]=-v2, v5=\mathbf{p}\times\mathbf{v}4=[001] \times [0-11]=[100]=v1.$

A plane can be defined by three points provided they are not colinear. If the three points are $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$, then $[\mathbf{abc}]=(\mathbf{u}_2-\mathbf{u}_1)\times(\mathbf{u}_3-\mathbf{u}_1)=\mathbf{u}_2\times\mathbf{u}_31$. $[a\ b\ c\ d] \ \mathbf{U}_1 = 0$ to find $d$.

EXAMPLE: $\mathbf{u}1=[124], \mathbf{u}2=[2-11], \text{and } \mathbf{u}3=[033]$, so $[\mathbf{abc}]=\mathbf{u}2\times\mathbf{u}31=[2-1-1-21-4] \times [0-13-23-4]=[1-3-3]\times[-11-1]=[64-2] \text{ and } [64-2d][1341]=6+8-8+d=0$, so
A plane can also be defined by a point \(U\) on the plane and vector \(p=[abc]\) orthogonal to the plane. It can also be defined by a line and a point, which is not on the line, both of which lie in the plane.

EXAMPLE: Find the plane containing line \(u(\alpha)=[1+\alpha 2\alpha 3]\) and point \(u_1=[2020]\).

When \(\alpha = 0\) the point on the line is \(u(0)=[103]\). A vector lying on the line is \(u(1)-u(0)=[110]\). Another vector lying on the plane is \(u(0)-u(1)=[1-20-23-0]=[1-23]\). The plane is the cross product of these two vectors in the plane, so \(p=[abc]=[120] \times [1-23]=[6-30]\). Finally, since \(u_1\) lies on \(P\), \( [6-30d][2200]=0\), so \(d=-6\).

### 5.7 Rotation of Vectors

Suppose vector \(v_1\) lies in normalized plane \(P\), so \(p \cdot v_1 = 0\). \(v_1\) is to be rotated in plane \(P\) by angle \(\theta\) to produce vector \(v_2\) as shown in Figure 5.10; \(v_1\) and \(v_2\) are shown as hollow arrows. Dropping a normal from the end of \(v_2\) to \(v_1\) gives vectors \(\alpha = v_1 \cos \theta\) and \(\beta = p \times \alpha \tan \theta\); the latter is true since it is in the required direction and has the correct length. Therefore \(v_2 = \alpha + \beta = v_1 \cos \theta + p \times \alpha \tan \theta\), so

\[
v_2 = v_1 \cos \theta + p \times v_1 \sin \theta \tag{5.8}
\]

If \(\theta = 0^\circ\), then \(v_2 = v_1\), and if \(\theta = 90^\circ\), then \(v_2 = p \times v_1\), as expected.
\[ v_2 \cos \theta - p \times v_2 \sin \theta = (v_1 \cos \theta + p \times v_1 \sin \theta) \cos \theta - p \times (v_1 \cos \theta + p \times v_1 \sin \theta) \sin \theta = v_1 \cos 2\theta + p \times v_1 \sin \theta \cos \theta - p \times v_1 \sin \theta \cos \theta + v_1 \sin 2\theta = v_1 \]

so we return to the original vector, as expected.

We can consider \( p \) as a plane or as an axis. That is, if in Equation 5.8, vector \( p \) is considered an axis, then \( v_1 \) must be orthogonal to axis \( p \).

The major axes are the x-axis [100], the y-axis [010] and the z-axis [001]. The x-axis is also the y-z plane, the y-axis the x-z plane, and the z-axis the x-y plane.

We can rotate \( v_1 \) about \( p \) even if these two vectors are not orthogonal. We decompose \( v_1 \) into \( v_1 = v_a + v_b \) such that \( p \cdot v_a = 0 \) and \( v_b = \alpha p \), where \( \alpha \) is a scalar. Taking the dot product of this decomposition with \( p \) produces \( p \cdot v_1 = p \cdot v_a + p \cdot v_b = 0 + p \cdot (\alpha p) = \alpha \), so \( v_a = v_1 - v_b = v_1 - \alpha p = v_1 - (p \cdot v_1)p \).

Rotating \( v_1 = v_a + v_b \) by \( \theta \) in plane \( p \), where \( v_a \) lies in \( p \), changes Equation 5.8 to

\[
v_2 = v_a \cos \theta + p \times v_a \sin \theta + v_b = (v_1 - (p \cdot v_1)p) \cos \theta + p \times (v_1 - (p \cdot v_1)p) \sin \theta + (p \cdot v_1)p (5.9)\]

![Figure 5.11: Vector Decomposition](image)

Example: Decompose \( v = [211] \) into the sum of two vectors \( v_a \) and \( v_b \) where \( v_a \) lies in plane \( 12[110] \) and \( v_b = \alpha p \)

\[
\alpha = p \cdot v = 12(2+1+0) = 32, \text{ so } v_b = \alpha p = [3/23/20] \text{ and } v_a = v - \alpha p = [211] - [3/23/20] = [1/2-1/21]. \text{ By way of verification, } v = v_a + v_b = [1/2-1/21] + [3/23/20] = [211] \text{ as required.}

Any vector can be translated by Equation 5.8 to any other vector of the same length. Vector \( [x1y1z1] \) can become \( [x2y2z2] \), where \( x1 + y1 + z1 = x2 + y2 + z2 \), using Equation 5.8 by rotating in plane

\[
p = x1 + y1 + z1 \cdot [x1y1z1] \times [x2y2z2] = x1 + y1 + z1 \cdot [y1z2 - z1y2z1x2 - x1z2x1y2 - y1x2] (5.10)\]

www.electronicbo.com
and the angle of rotation is found from two of the three equations in \( \mathbf{v}_1 \cos \theta + \mathbf{p} \times \mathbf{v}_1 \sin \theta = \mathbf{v}_2 \).

**EXAMPLE:** Determine the plane of rotation and rotation angle to transform \([110]\) into \([002]\).

\[ \mathbf{p} = 12[110] \times [002] = 12[2\,\tilde{2}0], \] and using this plane of rotation in **Equation 5.8** gives \([110] \cos \theta + [1/2\,\tilde{1}/20] \times [110] \sin \theta = [110] \cos \theta + [002] \sin \theta = [002] \) so \( \theta = 90^\circ \).

Suppose point \( \mathbf{u}_1 \) is to be rotated about axis \( \mathbf{u}_a + \alpha(\mathbf{u}_b - \mathbf{u}_a) \). An axis is a line, so it has location as well as direction, unlike a vector, which has direction only. The plane of rotation is \( \mathbf{p} = \mathbf{u}_a - \mathbf{u}_b | \mathbf{u}_a - \mathbf{u}_b | \). The distance \( D \) of point \( \mathbf{u}_1 \) from the axis is given by

\[
D^2 = (x_1 - x_a - \alpha(x_b - x_a))^2 + (y_1 - y_a - \alpha(y_b - y_a))^2 + (z_1 - z_a - \alpha(z_b - z_a))^2
\]

and for this distance to be a minimum, \( \partial \alpha D^2 = 0 \), which determines \( \alpha \), and so the closest point on the axis as \( \mathbf{u}_2 \).

\[
\mathbf{p} \cdot (\mathbf{u}_1 - \mathbf{u}_2) = \mathbf{p} \cdot \mathbf{u}_{12} = 0,
\]
so the rotation of point \( \mathbf{u}_1 \) by angle \( \theta \) about the plane \( \mathbf{p} \) to produce point \( \mathbf{u}_3 \) can use **Equation 5.8** as

\[
\mathbf{u}_3 = \mathbf{u}_2 + \mathbf{u}_{12} \cos \theta + \mathbf{p} \times \mathbf{u}_{12} \sin \theta. \tag{5.11}
\]

**EXAMPLE:** Rotate point \([221]\) about axis \([0\alpha\alpha]\) by angle \( \theta \).

Choosing \( \alpha = 0 \) and \( \alpha = 1 \) to determine \( \mathbf{p} = [011] = 12[011] \).

![Figure 5.12: Vector at a Winter Solstice](image)

\[
D^2 = 22 + (2 - \alpha)^2 + (1 - \alpha)^2 = 2\alpha^2 - 6\alpha + 9, \partial \alpha D^2 = 4\alpha - 6 = 0 \] at max or min, so \( \alpha = 32 \) producing axial point \([03/23/2]\), and from **Equation 5.11** \( \mathbf{u}_3 = [03/23/2] + [22-3/21-3/2] \cos \theta \\
+ 12[011] \times [12-3/21-3/2] \sin \theta = [03/23/2] + [21/2-1/2] \cos \theta + 12[-11-1/2] \sin \theta. \]

## 5.8 The Obliquity Vector
Winter solstice occurs \( \sigma \) days before perihelion, which is equivalent to an angle of \( \sigma^- = -360^\circ 365.2422 \) and can be represented by unit vector \( \mathbf{v}_1 = [\cos \sigma^- \sin \sigma^- 0] \) lying in ecliptic plane \( \mathbf{p} \), as shown in Figure 5.12; the plane \( \mathbf{p} \) in which \( x, y, \) and \( \mathbf{v}_1 \) lie is shown encircled.

Consider the plane \( \mathbf{v}_1 \times \mathbf{p} \) in which vectors \( \mathbf{p}, \mathbf{q}, \) and \( \mathbf{v}_1 \) lie as shown encircled in Figure 5.13. If we rotate \( \mathbf{p} \) in this plane by the negative of the angle of obliquity 23.45° we produce unit vector \( \mathbf{q} = \mathbf{p} \cos 23.45 - (\mathbf{v}_1 \times \mathbf{p}) \times \mathbf{p} \sin 23.45 = \mathbf{p} \cos 23.45 + \mathbf{v}_1 \sin 23.45. \) We call this the obliquity vector, which is the axis of rotation of the Earth. Substituting in the known values for \( \mathbf{p} \) and \( \mathbf{v}_1 \) produces

\[
q = [\cos \sigma^- \sin 23.45 - \sin \sigma^- \sin 23.45 \cos 23.45] = [0.39795 \cos \sigma^- - 0.39795 \sin \sigma^- 0.91741].
\]

(5.12)

We are being somewhat coy by leaving \( \sigma^- \) as a variable although we identified it as a number when we discussed the Julian calendar. The reason for this will soon be apparent.

Declination is given by the angle between the vector \( [\cos \phi i \sin \phi i 0] \) from the Sun to the Earth and the obliquity vector \( [0.39795 \cos \sigma^- - 0.39795 \sin \sigma^- 0.91741] \) as

\[
\delta = \sin -1([0.39795 \cos \sigma^- 0.39795 \sin \sigma^- 0.91741] [\cos \phi i \sin \phi i 0]) = \sin -1(0.39795 \cos \sigma^- \cos \phi i - 0.39795 \sin \sigma^- \sin \phi i)(5.13)
\]

With \( \phi = 0 \) and \( \sigma^- = 14.09 \), the declination as given by Equation 5.13 is \(-22.705\), which according to the declination table occurs at \( J_5 \), a substantial deviation of two days from the expected day. Although the declination difference is small close to \( J_3 \), it will be substantial at the equinoxes, when the deviation of 2.035 days produces a deviation in declination of \( 0.484^\circ \). Where is the error? Either the analysis here is wrong, or the tables are wrong, or both are wrong.
The commonly accepted value for $\sigma$ is $\sigma = 14.3$, so $\sigma^- = -14.095$, giving

$$q = [0.3860 - 0.09690.9174].(5.14)$$

If the angle of obliquity was zero, the obliquity vector would be the ecliptic plane $p = [001]$, a unit vector. The vector dot product of two unit vectors is the cosine of the angle $\theta$ between them, so the angle between vectors $q$ and $p$ is

$$\theta = \cos^{-1}(q \cdot p) = \cos^{-1}[0.3860 - 0.09690.9174] \cdot [001] = \cos^{-1}0.9174 = 23.45,(5.15)$$

which is the angle of obliquity, as expected.

Knowing the ecliptic plane and the obliquity vector is not enough to determine the declination for a given day of the year. If the orbit of the Earth around the Sun was circular, then the orbital speed of the Earth would be constant and the Sun would be at the origin of the ecliptic plane, where $x=0$ and $y=0$. Then we would know with precision the exact angle of the Earth with respect to the Sun at every day of the year. Instead, Earth’s elliptical orbit produces a constantly changing position with the result that there does not appear to be an analytic method of determining the declination. The declination, as plotted from standard declination tables, is shown for the year in Figure 5.14.

Declination Calculations

To determine the elevation and azimuth of the Sun as observed from a point on the Earth’s surface, one needs to know the declination. The easy way to determine this is to consult the declination table. In this section we will discuss some simple models for declination, some approximations to the table, and a numerical model based on Kepler’s laws.

The seeming accuracy in the declination table is misleading, since at the equinoxes the daily change in declination is $24'$, while the leap year corrections induce a half a day error every four years, equivalent to $12'$. That is, a determination of the declination angle with an error of the order of $12'$ is acceptable.

The orientation of the axis of the Earth changes by $50.3$ seconds/year: This is known as the progression of the equinoxes. The result is that the perihelion point on the Earth’s orbit increases from about $J_3$ today, so that after about 12,500 years it will be close to the summer solstice.

An accurate resource and the one judged to be the most acceptable is Spencer’s Fourier series approximation [J.W. Spencer, “Fourier Series Representation of the Position of the Sun,” Search, vol 2 #5, 1971] to the commonly accepted declination tables:

$$\delta s = 0.3964 - 22.9133 \cos \theta + 4.0254 \sin \theta - 0.3872 \cos 2\theta + 0.0520 \sin 2\theta - 0.155 \cos 3\theta + 0.0848 \sin 3\theta,$$
where $\phi=360362.25(J_n-1)$. At the winter solstice, $J_n = -11.25$, at the equinox add 91.25 to $J_n$, and at the summer solstice add 182.5 to $J_n$. Spencer’s approximation is shown in 5.14 and covers a single cycle over the complete year.

If the orbit of the Earth around the Sun was circular, then the orbital speed of the Earth would be constant and the declination would be $\delta_c = 23.45 \sin \left(\frac{360}{365} J_n - 80\right) \equiv -23.45 \cos \left(\frac{360}{365} J_n + 10\right)$. (5.17)

An approximation that is similar is Equation (3.7) in Dan Chiras’ book [D. Chiras, Power from the Sun: A Practical Guide to Solar Electricity, New Society Publishing, Philadelphia]. Also found at www.powerfromthesun.net/Book/chapter03/chapter03, and is

$$\delta_p = \sin -10.39795 \cos (0.98563(J_n-173)). (5.18)$$
Figure 5.15: Differences Between Various Models and Spencer’s Approximation

Table 5.3: Difference Between Kepler and Spencer

<table>
<thead>
<tr>
<th>Julian</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>46</td>
<td>+0.308</td>
</tr>
<tr>
<td>138</td>
<td>-0.092</td>
</tr>
<tr>
<td>217</td>
<td>+0.133</td>
</tr>
<tr>
<td>316</td>
<td>-0.201</td>
</tr>
</tbody>
</table>

Using the numerical analysis summarized by Equation 5.13 produces the declination values that differ from Spencer’s approximation, as shown by the response $\delta_s - \delta_K$ in Figure 5.15; the difference between these is shown at its maximum in Table 5.3 with the corresponding Julian day. Also shown in Table 5.15 are the two responses $\delta_s - \delta_c$ and $\delta_s - \delta_p$.

$\delta_s$ has significant second and third order components. These are not present in $\delta_c$ or $\delta_p$. More significantly, the Kepler-based analysis does not have these either. There is no doubt that $\delta_s$ accurately models the existing declination table. How was the declination table created? By celestial observation, or by calculation?

All the responses show two cycles per year. This is similar to the EoT discussed in Section XVI of this chapter. The Kepler approximation is the closest to Spencer, but which of these
two is more accurate? Also notice that at the two solstices, all the approximations give the same value of zero for the declinations as does Spencer.

5.9 Angle of the Sun on the Earth’s Surface

Consider a point at latitude $L$ at solar noon as the origin of Cartesian coordinates $x$ and $z$, as shown in Figure 5.16: The $z$ axis is vertical, and the $x$ axis is the horizon due south. Further, suppose $v_1$ is the vector in the direction of the axis of the Earth with declination $-23.45 \leq \delta \leq 23.45$ and lying on the x-z plane, so

$$v_1 = [-\cos L \sin L]. (5.19)$$

Also shown in Figure 5.16 is $v_2$, the vector in the direction of the Sun at solar noon also lying in the x-z plane, so

$$v_2 = [\sin (L-\delta) \cos (L-\delta)]. (5.20)$$

The angle in degrees between $v_1$ and $v_2$ is given by

$$\cos \theta = v_1 \cdot v_2 = [-\cos L \sin L] [\sin (L-\delta) \cos (L-\delta)] = -\cos L \sin (L-\delta) + \sin L \cos (L - \delta) = \sin \delta = \cos (90 - \delta)$$

as expected.
If the time away from solar noon is given in hours by $\tau$, then the solar hour angle is $\phi = 15\tau$ degrees (this follows, since in a 24-hour day the Earth revolves about $360^\circ$, and $360/24 = 15$), so the vector in the direction of the Sun can be found using Equation 5.21 as

$$v_s=[-\cos \, L0 \sin \, L] \sin \, \delta(1-\cos \, 15\tau)+[\sin \, (L-\delta)0 \cos \, (L-\delta)] \cos \, 15\tau +[-\cos \, L0 \sin \, L] \times [\sin \, (L-\delta)0 \cos \, (L-\delta)] \sin \, 15\tau[-\cos \, L \sin \, \delta(1-\cos \, 15\tau)+\sin \, (L-\delta) \cos \, 15\tau(\sin \, L \sin \, (L-\delta)+\cos \, L \cos \, (L-\delta) \sin \, 15\tau) \sin \, L \sin \, \delta(1-\cos \, 15\tau)+\cos \, (L-\delta) \cos \, 15\tau]=-\cos \, L \sin \, \delta(1-\cos \, 15\tau)+(\sin \, L \cos \, \delta - \cos \, L \sin \, \delta) \cos \, 15\tau \cos \, \delta \sin \, 15\tau(\sin \, L \sin \, \delta(1-\cos \, 15\tau)+(\cos \, L \cos \, \delta+\sin \, L \sin \, \delta) \cos \, 15\tau)]$$

so finally

$$v_s=[v_s x v_s y v_s z]=-\cos \, L \sin \, \delta+\sin \, L \cos \, \delta \cos \, 15\tau \cos \, \delta \sin \, 15\tau \sin \, L \sin \, \delta+\cos \, L \cos \, \delta \cos \, 15\tau]. (5.21)$$

As partial verification, we set $\tau=0$ so Equation 5.21 becomes

$$v_s=[-\cos \, L \sin \, \delta+\sin \, L \cos \, \delta \sin \, 0 \sin \, L \sin \, \delta+\cos \, L \cos \, \delta]=[\sin \, (L-\delta)0 \cos \, (L-\delta)]=v_2,$$

which is correct. Further, rotation of a unit vector produces a unit vector, and we verify this with

$$|v_s|^2=v_s x v_s y v_s z=-\cos \, L \sin \, \delta+\sin \, L \cos \, \delta \cos \, 15\tau^2+\cos \, 2\delta \sin ^2 \tau+\sin \, L \sin \, \delta+\cos \, L \cos \, \delta \cos \, 15\tau+\sin \, L \sin \, \delta+\cos \, L \cos \, \delta \cos \, 15\tau+\cos \, 2\delta \sin \, 15\tau+\sin \, 2\, L \sin \, 2\delta+\cos \, 2\, L \cos \, 2\delta \cos \, 215\tau-2 \cos \, L \sin \, L \cos \, \delta \sin \, 15\tau +\cos \, 2\delta \sin \, 215\tau+\sin \, 2\, L \sin \, 2\delta+\cos \, 2\, L \cos \, 2\delta \cos \, 215\tau+2 \cos \, L \sin \, L \cos \, \delta \sin \, \delta \cos \, 15\tau+\sin \, 2\, \delta+\cos \, 2\, \delta \cos \, 2 \, 15\tau+\cos \, 2\, \delta \sin ^2 \, 15\tau=1.$$
At solar noon, when the angle of the Sun is maximum for that day, \( \tau = 0 \) and Equation 5.22 becomes

\[
\cos \psi = \sin L \sin \delta + \cos L \cos \delta = \cos (L - \delta), (5.26)
\]

so \( \psi = L - \delta \) and \( \alpha = 90 - L - \delta \).

Plotting \( v_{sy} \) versus \( v_{sx} \), or \( y \) against the x-axis for latitude 42.25° (that of Boston, Massachusetts), for the solstices and equinox, as well as the midway points between solstice and equinox, for all mornings when the Sun’s elevation is above the horizon, produces Figure 5.18. Naturally, the afternoon equivalent of Figure 5.18 is its mirror image.

The declination angle used in producing Figure 5.18 is Spencer’s approximation, Equation 5.16, assuming a 365.25 day year. The days of the year chosen were:

<table>
<thead>
<tr>
<th>( J_n )</th>
<th>( \delta )</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>-11.25</td>
<td>23.42</td>
<td>December 22, winter solstice</td>
</tr>
<tr>
<td>45.66 - 11.25 = 34.41</td>
<td>( \delta = -16.37 )</td>
<td></td>
</tr>
<tr>
<td>91.3125 - 11.25 = 80.0625, ( \delta = 0.31 )</td>
<td>March 22, spring equinox</td>
<td></td>
</tr>
<tr>
<td>136.97 - 11.25 = 125.72, ( \delta = 16.46 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>182.625 - 11.25 = 171.375, ( \delta = 23.45 )</td>
<td>June 21, summer solstice</td>
<td></td>
</tr>
</tbody>
</table>

The declination angle should be zero at the equinox, but instead is \( \delta = 0.31 \) here. The discrepancy is minor, caused partially by the Spencer approximation, and partly by the effects of a leap year every fourth year.
5.10 Angle of the Sun on Solar Collectors

The unit vector that points at the Sun, with azimuth $\beta$ and elevation $\alpha$, is $v_s$. If $\beta = 0$, then $v_s = [\cos \alpha 0 \sin \alpha]$. If instead $\alpha = 0$, then $v_s = [\cos \beta \sin \beta 0]$. It is a simple logical step to deduce, for any $\alpha$ and $\beta$, that

$$v_s = [\cos \beta \cos \alpha \sin \beta \cos \alpha \sin \alpha]. \quad (5.27)$$

Similarly, consider a solar collector at the origin whose angle to the vertical is $\phi$ and whose yaw angle, the angle with respect to the x-axis, due south in the northern hemisphere, is $\gamma$. Suppose the unit vector normal to the collector is $v_c$. If $\phi = 0$ then $v_c = [\cos \gamma \sin \gamma 0]$, and if $\gamma = 0$, then $v_c = [\cos \phi 0 \sin \phi]$, so for any $\phi$ and $\gamma$,
\[ v_c = [\cos \gamma \cos \phi \sin \gamma \cos \phi \sin \phi]. (5.28) \]

The angle \( \theta \) between unit vectors \( v_s \) and \( v_c \) determines the solar collection, which is a maximum when \( \theta = 0 \). This angle is given by

\[
\cos \theta = [\cos \beta \cos \alpha \sin \beta \cos \alpha \sin \alpha] [\cos \gamma \cos \phi \sin \gamma \cos \phi \sin \phi] = \cos \beta \cos \alpha \cos \gamma \cos \phi + \sin \beta \cos \alpha \sin \gamma \cos \phi + \sin \alpha \sin \phi = (\cos \alpha \cos \phi)(\cos \beta \cos \gamma + \sin \beta \sin \gamma) + \sin \alpha \sin \phi
\]

so

\[
\cos \theta = \cos \alpha \cos \phi \cos (\beta - \gamma) + \sin \alpha \sin \phi. (5.29)
\]

If \( \gamma = 0 \), so the yaw angle of the solar collectors is zero,

\[
\cos \theta = \cos \beta \cos \alpha \cos \phi + \sin \alpha \sin \phi, (5.30)
\]

which is the commonly accepted formula. Further, if these south-facing collectors are vertical, so \( \phi = 0 \), then

\[
\cos \theta = \cos \beta \cos \alpha. (5.31)
\]

The angle \( \theta \) of the Sun on south-facing vertical glazing as a function of the solar time angle is shown in Figure 5.19 for latitudes 40° and the three most important times of the year: the winter and summer solstices and the equinox. Also shown in Figure 5.19 are the azimuth angles \( \beta \). Notice that for high solar time angles the azimuth angle \( \beta \) effectively becomes the same as the angle on the collector \( \theta \). Further, the minimum angle of the Sun on the collectors at summer solstice is 73.45°, so even without architectural shading (see Chapter 11) the unwanted heating effect of the summer Sun is minimal.
Changing the angle of the solar collectors from vertical to $30^\circ$ from vertical produces the angles shown in Figure 5.20.

Finally, with the angle of the solar collectors from vertical to $60^\circ$ from vertical produces the angles shown in Figure 5.21.

### 5.11 The Solar Collection Day

Regardless of the orientation of a solar collector, a necessary condition for positive solar collection is for $\alpha \geq 0$. From Equation 5.23 this means $\sin L \sin \delta + \cos L \cos \delta \cos 15\tau = \alpha \geq 0$. When $\alpha = 0$,

$$
\cos 15\tau = -\sin L \sin \delta \cos L \cos \delta = -\tan L \tan \delta, \text{(5.32)}
$$
Figure 5.20: Solar Angles on 30 Collectors

Figure 5.21: Solar Angles on 60 Collectors
so at equinox ($\delta = 0$) $\cos 15\tau = 0$ and $\tau = 6$ as expected, at winter solstice ($\tau = -23.45$) $\cos 15\tau = 0.4338 \tan L$, and at summer solstice ($\delta = 23.45$) $\cos 15\tau = -0.4338 \tan L$. 

The solar elevation will be zero at $\tau = 6$ hours when the latitude is zero and/or we are at equinox. When the solar elevation is zero at the equator the time from solar noon is always $\tau = 6$, regardless of declination. When the solar elevation is zero at equinox the time from solar noon is always $\tau = 6$, regardless of latitude. When $L \neq 0$ and $\delta \neq 0$, the time for $\alpha = 0$ is never $\tau = 6$.

![Figure 5.22: The Relationship between tau and del when alpha = 0](image)

The relationship between time from solar noon to $\alpha = 0$ is shown in 5.22 for a range of latitudes as a function of declination.

There is a balancing symmetry for positive declinations. That is, for declination $\delta$ and latitude $L$, the time for the solar elevation to be zero is $\tau + 6$, where $\tau$ is the time for the solar elevation to be zero if the declination was $-\delta$.

When the solar elevation is zero the azimuth, as given by equation 5.25, becomes

$$\beta = \sin^{-1}(\cos \delta \sin 15\tau)$$

where $\tau$ is determined by equation 5.32. The values of azimuth for $\alpha = 0$ is given in 5.23 for negative declinations. Again, there is a balancing symmetry for positive declinations: The
azimuth for positive declination $\delta$ at latitude $L$ when $\alpha = 0$ is $180 - \beta$, where $\beta$ is the azimuth for declination $-\delta$ at the same latitude and solar elevation zero.

We now consider specific collector orientations. For horizontal collectors we know that the collection day ends when $\alpha = 0$ and the time $\tau$ can be found from equation 5.32.

![Figure 5.23: Azimuth as alpha = 0 as a Function of Declination](image)

For vertical collectors facing south the collection day ends when one of these two conditions happens first: $\alpha = 0$, or $\beta = 90^0$. We can deduce that which one happens first is a function of declination. If $\delta \leq 0$ then $\alpha = 0$ happens first. When $\delta \geq 0$ then $\beta = 90^0$ happens first. At equinox both occur at the same time. On the winter side of equinox the collection day ends with $\beta < 90^0$, whereas on the summer side it ends with $\alpha > 0$.

For passive solar collection through glass the only acceptable angle for the glass is vertical facing the equator. Vertical glass through which occupants of the structure can view the outside works better than sloping glass. Although sloping glass, with an angle possibly at $\phi = 90 - L/2$ to vertical, can increase the winter solar gain, it will have the unfortunate effect of substantially increasing the unwanted solar gain in summer.

In the southern hemisphere the collection day ends with the opposite sign for declination.
Care must be taken in determining that either $\alpha$ has reached zero or $\beta$ has reached $90^\circ$. The problem is that we have a formula for $\sin \beta$, and when $\beta = 91^\circ$ we have the same value as for $\beta = 89^\circ$. The solution to the problem taken here is to use small time steps in the MATLAB program so that in the sequence $\sin \beta_1, \ldots, \sin \beta_n$, $\sin \beta_{n+1}$ if the last value $\sin \beta_{n+1}$ is less than $\sin \beta_n$, then we have passed $\beta = 90^\circ$.

Another important collector orientation is latitude inclined, meaning that the angle of the collector to vertical is given by $\phi = 90 - L$. As discussed in Chapter 7 this results in enhanced year round collection. Solar collection for a declination on the summer side of equinox is the same, regardless of latitude, and the collection day ends at $\tau = 6, \theta = 90^\circ$ and $\alpha > 0$. The solar collection day for a declination on the winter side of equinox ends when $\alpha = 0$ and $\tau < 6$.

### 5.12 Change in Solar Angle With Time

At equinox, $\delta = 0$, and equation 5.24 becomes $\tan \beta = \tan 15\tau/\sin L$ : when $L = 0$ the azimuth angle changes from $-90^\circ$ to $+90^\circ$ in the time span from just before solar noon to just after solar noon.

At equinox the world receives 12 hours of sunlight in a day, regardless of latitude, and at the equator the Sun is directly overhead and changes in longitude by $15^\circ$ in an hour.

Consider the unit vector

$$vn=[-\cos L \sin \delta + \sin L \cos \delta 0 \sin L \sin \delta + \cos L \cos \delta](5.34)$$

which points at the Sun at solar noon with latitude $L$ and declination $\delta$. The angle $\omega$ between $v_n$ and the direction of the Sun at times other than solar noon, given by equation 5.21, is

$$\cos \omega=[-\cos L \sin \delta + \sin L \cos \delta \cos 15\tau \cos \delta \sin 15\tau \sin L \sin \delta + \cos L \cos \delta \cos 15\tau][-\cos L \sin \delta + \sin L \cos \delta 0 \sin L \sin \delta + \cos L \cos \delta] = (-\cos L \sin \delta + \sin L \cos \delta \cos 15\tau)(-\cos L \sin \delta + \sin L \cos \delta) + (\sin L \sin \delta + \cos L \cos \delta \cos 15\tau)(\sin L \sin \delta + \cos L \cos \delta) = (\cos 2L + \sin 2L) \sin 2\delta + (\sin 2L \cos 2\delta + \cos 2L \cos 2\delta) \cos 15\tau$$

$$\cos \omega = \sin 2\delta + \cos 2\delta \cos 15\tau(5.35)$$

This result is surprising since

* $\omega$ is independent of the latitude, and
* for a given positive or negative declination $\omega$ is the same.

The change in angle when $\delta = 0$ is $15\tau$, as expected: the angle of the Sun changes by $15^\circ$ per hour at equinox. However, another surprise is that at the solstices the change in angle is only
slightly less than 15° per hour as shown in 5.24.

Figure 5.24: Angle of the Sun with Respect to Solar Noon

Figure 5.25: Rate of Change of Azimuth

The rate of change of azimuth as a function of latitude and declination is given in 5.25. This rate of change increases with decreasing latitude. Its maximum rate of change is infinite, and this occurs in the tropics when the Sun is directly overhead. Also, the azimuth changes most rapidly at summer solstice and the slowest at winter solstice.
5.13 Seasonal Angles

For Boston, which is at latitude 42.3° north, the relationship between the azimuth and elevation angles at the solstices and equinox is shown in 5.26.

Calculating the time when the Sun’s elevation is at 5°, and then determining the azimuth at this time for winter solstice produces the relationships shown in 5.27. These are useful relationships to have when determining solar collector gains and siting requirements.

5.14 Equation of Time and Sundials

Longitude is defined with respect to Greenwich, England. Greenwich is at longitude 0°, and the world’s time is often quoted with respect to GMT (Greenwich Mean Time). Time can be defined in a specific location with respect to its longitude; we call it longitude time. At Greenwich an accurate clock reads the same as longitude time. At Philadelphia, Pennsylvania, the clocks will show the same time as longitude time, five hours after GMT since it is close to longitude 75°, and 75 = 5 × 15 where the rotation of the Earth is 15° per hour.

Boston, Massachusetts, time is termed Eastern Standard Time, five hours after GMT. However, in Boston longitude noon is \((75−71.1)6015=15.6\) minutes before noon, clock time, since its
longitude is 71.1°

The daily rotation of the Earth takes less than 24 hours. In a 365 day year the Earth rotates 366 times - the extra rotation is caused by the once-a-year orbit of the Sun. This phenomenon was illustrated in 5.4. As the Earth, shown with an arrowed arc to indicate rotation about its axis, orbits the Sun, the orientation of the Earth with respect to the Sun at solar noon is indicated by the arrow pointed approximately (more about this later) at the Sun. Notice at J₈₀ vernal equinox, the Earth has completed 79 rotations plus almost a quarter since J₁. At J₁₇₃, solstice, it has completed 172 rotations plus almost a half. Thus, when it reaches J₃₆₆ ≡ J₁ it has rotated 365+1 times. Since there are 365 days in the year described above, but the Earth rotates 366 times, then the day is short of 24 hours by 1366×241×601=3.934426 minutes. That is, the Earth makes one complete rotation about q every 23 hours, 56 minutes and 4 seconds, or every 23.9344 hours. When estimating in Julian days one must use 24 hour days.

![Figure 5.27: Azimuth and Time when Elevation at 5](image)

There is only one timepiece that has no moving parts - it is the sundial. Sundials have been employed since ancient times to tell time and date. A sundial is a horizontal platform on which a right angle, triangular structure called a gnomon is mounted. The gnomon is mounted in along north-south line with the vertical edge to the north as shown in 5.28. The north-south line is determined by accurate manual orientation at longitude noon. On the flat platform are scribed hour lines; as shown in 5.28 the lines shown are in hours from longitude noon. The top sloping edge of the gnomon is called the style, and part way down the style is a notch or similar feature
termed the nodus, so when the time is away from noon the nodus is visible on the shadow of the gnomon on the platform.

![Figure 5.28: The Classic Sundial](image)

Shown in 5.28 is the shadow cast by the nodus at summer and winter solstices and at equinox. At winter solstice the declination is -23.45°, so the shadow of the nodus is an indication of date. Additional lines on the platform can determine approximate date of the year, with the caveat that two dates of the year have the same declination (except the solstices).

As described, this is the classic sundial. Many sundials are made as works of art and ornamentation. They are attractive to look at, and provide a solid satisfaction that regardless of electrical power, or failure to wind the clock or change its batteries, the sundial provides its valuable service. There are other configurations of sundial, the most notable being wall mounted.

Sundials do not show the same time as a clock. A clock runs at exactly the same rate, but the time shown by a sundial varies according to the ‘equation of time’, the EoT; the phrase using the word equation is awkward and came into existence in Eighteenth Century when equation meant correction.
Longitude noon is not the same as solar noon - the time of day when the elevation of the Sun is at its maximum. The difference between longitude noon and solar noon is the EoT. The maximum deviation is up to 16 minutes 23 seconds fast on November 3 and up to 14 minutes 20 seconds slow on February 12 [A. Carlson, “The Equation of Time,” which can be found at www.sundials.co/equation]. There are two factors at work here. First, the elliptical orbit of the Earth around the Sun means it does not travel at constant speed. Second, the axis of rotation of the Earth with respect to the ecliptic plane. The first factor completes one sinusoidal cycle in a year, the second completes two such cycles. More about this in Section XVI.

Some discussions on the equation of time discuss two mechanisms that interact, obliquity and unequal motion along the Earth’s elliptical orbit. That is, the two mechanisms produce separate graphs which are then brought together to get the desired result [Greenwich Observatory Leaflet #13, “The Equation of Time”, 1993].

The Lamm approximation is accurate to within seconds to the EoT table [L.O.Lamm, “A New Expression for the Equation of Time”, Solar Energy, vol 26, p465, 1981]. What is not clear how the EoT table was produced. There does not seem to be a mathematical derivation, which suggests celestial observation as the source.

5.15 The Analemma and How to Capture It

If one calculates solar elevation against azimuth at longitude noon throughout the year, the pattern followed is that of the number 8; this is called the analemma, and is shown for latitudes 30 in 5.29, 40 in 5.30 and 50 in 5.31. Notice that the shape of the analemma changes little with latitude. Also shown in these analemma diagrams are locations at key Julian days on the analemmas.
Figure 5.29: Analemma at Latitude 30
Notice that the maximum and minimum elevations at longitude noon are $90 \pm 23.45 - L$. Further, if longitude noon was the same as solar noon, in other words the equation of time did not exist, then the analemma would be a vertical line in the sky.

At times other than longitude noon the analemma is no longer essentially symmetric about a vertical axis. The analemmas 3 hours from longitude noon at latitude 30 the analemma is shown in 5.32, at latitude 40 in 5.33, and latitude 50 in 5.34.

It is extremely difficult for an individual without specialized equipment to observe the analemma. However, an indirect method using the shadow cast by the Sun is a simple way to demonstrate the phenomenon. Place a rod in patch of level, open ground that is exposed to the noon Sun throughout the year. Use a spirit level to make sure the rod is vertical. My choice would be a 1/2” rebar, used in reinforcing poured concrete foundations. It should be 5’ or 6’ long and pounded into the ground by about 2’ so 3’ to 4’ is exposed. At noon, preferably longitude noon, drive a small stake into the ground at the end of the shadow. This should be done approximately every other week.
Figure 5.31: Analemma at Latitude 50
Figure 5.32: Analemma at Latitude 50 and Three Hours from Noon

Figure 5.33: Th Shadow Analemma
Rather than pounding these stakes into the ground, a better arrangement would be to set a sheet of plywood by the rod. Recall, in latitudes north of the equator and north of the tropics, the shadows will always be on the north side of the rod. The plywood should be levelled and affixed so that it does not move throughout the year. At noon the shadow tip can be recorded with a small nail. The advantage of this method is at the completion of a year, the plywood can be picked up and the shape of the shadow analemma traced.
For latitude 40 the shadow analemma is as shown in 5.35, where the axes are x pointing north, and y points due west. The height of the rod creating the shadow was unity, so notice that the shadow length, which is maximum at winter solstice, as about 2. Thus, a 4’ rod above ground will require the full length of 8’ for the sheet of plywood. The geometry for the shadow analemma is given in 5.36. The representation of the Sun is at the left and casts a shadow of the rod of length ell of length ell of length ell tan α. The right-hand diagram is a plan view showing azimuth angle β and the shadow length ell tan α, which decomposes into x=ell cos β tan α and y=ell sin β tan α.

An even simpler way to capture the shadow analemma is through a south facing window. The window does not need to be facing due south, but should be within about 20° of due south. The window should extend close to the interior floor, such as would occur with a glass door; for accuracy reasons it is preferable to use the fixed glass panel rather than the sliding panel. Place a target on the glass about 18” above the floor and accurately measure this distance to within 1/16”. Place a white foam board on the floor touching the frame of the glass or the wall beneath the glass. Carefully mark the location selected since the board will be stored after each recording of the shadow created by the target at longitude noon. The accuracy needed in positioning this board should be within 1/16”, and the accuracy in timing should be closer than one minute. On a clear day every week or two place the board on the floor in its assigned position and mark the shadow of the target.

If a glass panel extending down close to the floor is not available, a window can be used if a table can be placed beneath it, and marks made on the wall so the board can be accurately positioned on the table. Also, for accuracy reasons, a hard floor of wood or ceramic is preferable to a carpeted floor.

Continue for a full year. Next, trace the shape of the shadow analemma, and using a straight edge, draw the major axis. The angle of the major axis with respect to the normal of the frame/wall to which the board was placed determines the deviation of the angle of the glass to facing due south.

Next, mark horizontal distances in inches from the glass along the major axis; we can term this the x-axis. Then, using a framing square or similar tool, draw grid lines at right angles to the
major axis; the distance from the major axis gives us the y coordinate. This enables us to produce a table of x-y coordinates. Using MATLAB, these x-y coordinates can be translated into numbers for the solar elevation and azimuth, thus producing the true analemma.

5.16 Approximating the EoT

The EoT is shown as a solid line in Figure 5.37. Notice it has twice the periodic frequency as the rotation of the Earth around the Sun. Also shown in Figure 5.37 is the approximation

\[ \text{EoT} = 9.87 \sin 2B - 7.53 \cos B - 1.5 \sin B, \quad (5.36) \]

where \( B = 360(J_n - 81)/365 \) [www.susdesign.com/popups/sunangle/eot].

The formula

\[ \text{EoT} = 9.87 \sin 2B - 7.67 \sin (B+78.5), \quad (5.37) \]

with the same \( B \) as above produces essentially the same response [www.ips.gov.au/Category/Educational/The%20Sun%20and%20Solar%20Activity/General%.

The difference between the table and \( E_A \) is shown in Figure 5.38. The maximum differences between the two is \(-1.42\), which occurs at \( J_n = 268 \), and \(+1.50\), which occurs at \( J_n = 342 \). More disturbing is the shape of graphical differences. The Fourier approximation should produce a smooth, cyclic response, but the difference is choppy, indicative of errors in the EoT table.
Other Fourier-type approximations to the EoT table include

\[ \text{EoT} = 0.258 \cos x - 7.4161 \sin x - 3.648 \cos 2x - 9.228 \sin 2x, \quad (5.38) \]

where \( x = 360(J_n - 1)/265.242 \), and produces a slightly larger error than the previous ones [www.powerfromthesun.net/Book/chapter03/chapter03.html#Equation of Time].
Chapter 6

Transmission of Light through the Atmosphere

Chapter 2 defines what is needed for human comfort in a residence. Chapter 3 provides fundamental information on heat losses from a residence. Chapter 5 discusses solar angles on surfaces such as solar collectors. The first half of this chapter considers the nature of electromagnetic radiation and the spectoral energy in sunlight. All these subjects present material that is well travelled and without major gaps that require conjecture and possibly controversy to cross.

The attenuation of solar energy in the atmosphere is highly variable. In order to determine how much energy can be harvested, we need mathematical models. A number of models exist, but there seems to be a paucity of models that are data based. An extensive amount of solar data comes from U.S. National Laboratories. There may be a problem with this data, however, since much of it is modeled. Modeled data can validate a defective model.

SERI (Solar Energy Research Institute) became NREL (the National Renewable Energy Laboratory) in 1991. NREL has provided a considerable amount of data for every major city in the United States. One set of data gives the three measures global, beam (direct-normal), and diffuse for solar radiation for every hour of the year for the years 1960 to 2010. We will see that beam and diffuse provide what we need to determine the solar energy striking a collector.

The energy we are interested in here is solar energy in the waveband 0.3 $\mu$ to 2.7 $\mu$, where the dimension $\mu$ means micrometers, $10^{-6}$ meters. Glass, clear plastics, and water are transparent over most of this waveband, as will be detailed in Chapter 8.

Most solar collectors have clear glass or plastic covers, and how sunlight penetrates, reflects, and is absorbed by what is called a diathermanous material is the subject of the next chapter.

6.1 The Incoming Solar Radiation
All bodies with temperatures above absolute zero radiate electromagnetic energy. Frequency $f$ is measured in cycles per second, more commonly known as Hz (Hertz), and wavelength $\lambda$ is in meters. $f$ and $\mu$ are connected by the law

$$f\lambda = c$$

where $c = 3 \times 10^8$ m/s is the speed of light. As far as light is concerned, the unit of length we use is the micron, $10^{-6}$ meters, given the symbol $\mu$.

The full electromagnetic spectrum is shown in Figure 6.1, where the numbers indicate frequency in Hz. At the low end of the spectrum is the audio range, from about 40 Hz through 1,000 Hz. Next is the amplitude modulated (am) radio band, followed by the frequency modulated (fm) radio band, Radar occupies the range $5 \times 10^9$ through $4 \times 10^{10}$ Hz. Visible light appears red in the range $4 \times 10^{14}$ to $4.8 \times 10^{14}$ Hz with wavelength about 0.7 $\mu$, and violet is in the range $6.7 \times 10^{14}$ to $7.9 \times 10^{14}$ Hz with wavelength of about 0.3 $\mu$. The colors of the spectrum, in increasing order of wavelength are violet, indigo, blue, green, yellow, orange, and red.

A black body, one with emissivity unity at all wavelengths of interest, emits radiation according the spectral energy law

$$W^- = c_1\lambda^5(e^{c_2/\lambda T} - 1) = 3.7403 \times 10^8 \lambda^5(e^{14384/\lambda T} - 1)$$

(6.1)

in W/m²·μ, where $\lambda$ is the wavelength in microns (micro-meters), $T$°K is the absolute temperature of the body, $c_1 = 3.7403 \times 10^8$ W·$\eta^4$/m², and $c_2 = 14384 \eta$°K: equation 6.1 is known as Planck’s law.

The absolute temperature of the Sun’s surface is about 6000° K, so the energy emitted according to Planck’s law is

$$W^- = 3.7403 \times 10^8 \lambda^5(e^{2.39733/\lambda T} - 1).$$

(6.2)

Some suggest the slightly lower temperature of 5800° K, but we will stick to the higher figure here.
The spectoral energy versus wavelength diagram as given by Planck’s law for a body at 6,000° K is shown in Figure 6.2.

The human body is at about 310° K, so the energy per unit area it radiates is \((6000/310)^4=140332\) times less than that of the Sun. A semi-logarithmic plot of the energy emitted by the Sun and by a black body at the same temperature as a human, normalized, is shown in Figure 6.3. Notice that there is almost no overlap, and that the crossing point of the two responses is at approximately 2.7 \(\mu\). This is convenient, since we will establish that the effective cutoff point for radiation to pass through window glass is 2.7 \(\mu\).

When the Sun’s energy strikes a body on the surface of the Earth, such as earth, rock, or vegetation, the body warms up. The warming continues until the energy radiated (according to Planck’s law and Stefan-Boltzmann’s equation), convected, and conducted from the body’s surface balances the solar energy received. If a selective filter such as glass, that is transparent
to the short waves but opaque to long waves, is placed between the Sun and the energy absorbing body, then we have a mechanism for energy entrapment. This situation is shown in Figure 6.4: the Sun’s energy passes through the glass and is absorbed by the oval shaped body. The heated body re-radiates the energy like a mini-sun (hence the smaller image of a Sun).

![Figure 6.4: Energy through Glass and Back again](image)

To find the peak $\lambda_{\text{max}}$ of the spectral energy density we differentiate Planck’s law and equate to zero to give

$$\lambda_{\text{max}} T = 2897, (6.3)$$

which is known as Wien’s displacement law. For the Sun, the peak energy density occurs at $2897/6000 = 0.4828 \, \mu$, commonly approximated by $0.5\mu$, which is the center of the band of light visible to humans. Also, light at $0.5\mu$ corresponds to the color green, meaning that green plants can maximize energy capture.

It is not clear how Planck’s law was differentiated, so here we use MATLAB to differentiate Planck’s law, producing

$$d\lambda \bar{W} = e^{-1 - 2.39733/\lambda} \times 10^8 \lambda^7 (8.9667 - 8.7015\lambda), (6.4)$$

which equals zero at maximum or minimum, producing $\lambda_{\text{max}} = 8.9667/18.7015 = 0.47946$, a small deviation from Wien’s law.

Planck’s law for a variety of temperatures, from the freezing point of water at 273 °C to the surface of the Sun at 6000 °K, is shown in Figure 6.5. Also shown is the Wein law curve connecting the peak of each response.
Integrating the spectral energy density with respect to \( \lambda \) produces the energy radiated up to wavelength \( \lambda \) as

\[
E(\lambda) = \int_0^\lambda W \, d\lambda (6.5)
\]


\[
E(\infty) = \int_0^\infty W \, d\lambda = 5.679 \times 10^{-8} T^4 (6.6)
\]

and this is known as the *Stefan-Boltzmann* law. For a body with emissivity \( \varepsilon \) the total energy emitted is \( \dot{W} = 5.679 \times 10^{-8} \times 6000^4 \) W/m\(^2\), where \( 0 \leq \varepsilon \leq 1 \) is a dimensionless constant; \( \varepsilon = 0 \) for a totally reflective surface (a perfect mirror) and \( \varepsilon = 1 \) for a black body. The emissivity of the Sun is assumed to be unity, so the energy radiated from the Sun is \( E(\infty) = 5.679 \times 10^{-8} \times 6000^4 = 7.3600 \times 10^7 \) W/m\(^2\). The radius of the Sun is 430,000 miles or \( 6.920205 \times 10^8 \) m, so its
surface area is $4\pi r^2 = 6.0179 \times 10^{18} \text{ m}^2$, and the total energy radiated from the Sun is $7.360 \times 10^7 \times 6.0179 \times 10^{18} = 4.429 \times 10^{26} \text{ W}$.

Table 6.1: Total Solar Energy up to $\lambda$

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$E(\lambda)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>3.946</td>
</tr>
<tr>
<td>0.5</td>
<td>27.341</td>
</tr>
<tr>
<td>0.7</td>
<td>51.618</td>
</tr>
<tr>
<td>1</td>
<td>73.792</td>
</tr>
<tr>
<td>1.5</td>
<td>89.006</td>
</tr>
<tr>
<td>2</td>
<td>94.509</td>
</tr>
<tr>
<td>2.7</td>
<td>97.463</td>
</tr>
</tbody>
</table>

Table 6.2: Solar Energy in Each Waveband

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ultraviolet</td>
<td>$0 \rightarrow 0.3$</td>
<td>3.946%</td>
</tr>
<tr>
<td>visible</td>
<td>$0.3 \rightarrow 0.7$</td>
<td>47.672%</td>
</tr>
<tr>
<td>near infrared</td>
<td>$0.7 \rightarrow 2.7$</td>
<td>45.845%</td>
</tr>
<tr>
<td>far infrared</td>
<td>$2.7+$</td>
<td>2.537%</td>
</tr>
</tbody>
</table>

Integrating Planck’s law in MATLAB with $T = 6000$ and a step length of $h = 0.01$ using the second order correct trapezoidal rule $E\lambda_{n+1} = E\lambda_n + h(W\overline{n} + W\overline{n+1})$, so the error is a function of $h^3$, Table 6.1 is produced; the second column for $E(\lambda)$ should be multiplied by $10^7$. As $\lambda \rightarrow \infty$, the numeric integration gives $E(\infty) = 7.354 \times 10^7 \text{ W/m}^2$, slightly lower than the experimentally based Stefan-Boltzmann law.

The solar amount in the three wavebands are given in Table 6.2. The reason for defining the upper limit of the infrared region as $2.7\mu$ is based on the transmission characteristics of common glass, which has a relatively smooth waveband between 0.3 and $2.7\mu$ and a rapid falloff for wavelengths above $2.7\mu$. The percentage total energy for a body at 6000°K as a function of wavelength is shown in Figure 6.6.
6.2 Solar Radiation in the Atmosphere

As the solar energy radiates outward from the Sun there is some attenuation, so the solar energy reaching a point at the Earth’s distance from the Sun is reduced due to Planck’s Law. Selective filtering occurs as the solar energy enters the Earth’s atmosphere, primarily due to water vapor (a highly variable quantity) and oxygen; secondary filtering is caused by ozone, carbon dioxide, and particulate matter. As a result, the solar energy reaching the Earth’s surface is as shown in Figure 6.7. This is clear-sky data, and is an approximation.

All the energy of the Sun crosses the surface of an imaginary sphere that encloses the Sun. The Earth follows a slightly elliptic orbit, but with an average distance of $1.49 \times 10^8$ km, the radius of the Sun is $6.955 \times 10^5$ km, so the energy reaching a surface at the distance of the Earth from the Sun is $7.36 \times 107 \times (rd)^2 = 7.36 \times 107 \times (6.955 \times 1051.4967 \times 108)^2 = 1,519 \text{ W/m}^2$. 

![Figure 6.6: Percentage Solar Energy up to $\lambda$](image)
However, recent experiments in space have determined that the average solar irradiance in space near the Earth is 1366 W/m$^2$, a number known as the solar constant. The figure of 1366 W/m$^2$ is further confirmed by NASA, whose [website](http://aom.giss.nasa.gov/solar.html) presents a long-term date line of solar energy, as shown in Figure 6.8.

The reason for the reduction in energy from the Sun’s surface is that it is not a black body with emissivity $\varepsilon = 1$. Instead, it has emissivity $\varepsilon(\lambda)$, slightly less than unity for some wavelengths.
Attempting to obtain accurate numbers for useable solar energy at the Earth’s surface is an inexact exercise. Measurements can be made on specific days and averages compiled. However, the most frustrating aspect is the numbers that can be measured but are not explained. For example, the solar constant is 1366 W/m², and the measured normal incident energy in Boston on a clear day close to midday is about 880 W/m², so the percentage of extraterrestrial solar energy available on a normal surface is 64.4%. Is the solar energy that does not reach the Earth’s surface lost? What are the mechanisms of these losses, and is some energy, converted from direct beam energy, recoverable?

With the average radius of the Earth $r$, its surface area is $4\pi r^2$. The area subtended by the Earth to the solar radiation is $\pi r^2$, so the average energy received by the Earth is the solar constant times the ratio of the subtended area divided by the surface area, that is, $1366\pi r^2 / 4\pi r^2 = 1366 / 4 = 341.5$ W/m².

A commonly accepted figure of solar energy on a clear day at sea level, close to solar noon and striking a normal surface, is $880$ W/m² $\approx 279.0$ BTU/ft².hr; an exact conversion factor is $1$ W/m² $= 0.316998$ BTU/ft².hr. This is remarkably close to the series of observations taken on clear days at the Blue Hills Observatory in Milton, Massachusetts, showing $883$ W/m² striking a surface normal to the incident energy [I.F. Hand, U.S. Weather Bureau, Milton, Mass, 1950]; these figures are stated to be the same throughout the year. However, as will be seen later, the elliptical orbit of the Earth leads to a variation of normal radiation on the Earth’s surface of about $60$ W/m².
A considerable energy loss occurs in the ultraviolet region due to Rayleigh scattering (discussed in Section IV of this chapter), for wavelengths smaller than 0.3 \( \mu \): This is fortunate, since the energy in this frequency range is deadly to plants and animal life. Even with this filtering, the violet energy is primarily responsible for skin cancers in humans. However, some ultraviolet radiation is useful to humans, since such radiation can convert fats in the skin to vitamin D.

A typical breakdown of what happens to solar radiation as it passes through the Earth’s atmosphere is as follows [N. Mason, P. Hughes, et al, *Introduction to Environmental Physics: Planet Earth, Life and Climate*, Taylor & Francis, London, 2001]:

* 31% reflected back into space, with 23% from cloud tops and 8% from the ground; this 8% is part of the energy available for solar collectors.
* 21% selectively absorbed by constituents, predominantly water, in the atmosphere
* 25% scattered (Rayleigh and Mie, as discussed in Section IV of this chapter) by the atmosphere.
* 28% reaches the ground directly.

These add up to 96%, but rather than object, one should recognize how speculative these figures are, and how much they can vary. For example, more than half of the scattered radiation reaches the Earth’s surface and can be captured as diffuse radiation (see the definitions in Section VI). As discussed previously, 64.4% of the solar energy in space reaches the Earth’s surface on a clear day close to solar noon; most of this is direct and only a small amount is diffuse. Yet on an overcast day, direct solar radiation reaching the surface may be zero, while diffuse radiation on the surface is significant.

The most important number as far as passive solar heating is concerned is the amount of solar energy striking vertical glazing facing the equator. Thus, for this application, the data shown in Figure 6.9 is more important than the numbers given by Mason and Hughes. Of the 16%
scattered by the atmosphere, approximately 8% reaches the ground on a clear day near solar noon, and can be captured as diffuse radiation. In the diagram, the arrow angled downwards indicates solar energy that can be captured, while the arrow pointed upwards indicates energy lost in space. A total of about 20% of the incoming radiation is absorbed by the atmosphere on a clear day, and so is lost for solar collection purposes; it is this absorption that keeps the average temperature of the Earth comfortable. This leaves 64% to strike the glass directly. Again, these numbers are for clear days. Compare the 64% to the measured amount divided by the solar constant 880136610064%.

The diffuse component only makes up about 10% of the total irradiance on a clear day near midday; a minimum value for diffuse radiation is about 5%. Near sunrise/sunset on a clear day this number is about 20%. A high number for diffuse radiation is 70%, with a maximum number of about 85%.

The NREL site [http://rredc.nrel.gov] has data for every major city in the United States, and gives hourly numbers for solar radiation, terrestrial and extraterrestrial. The average number they give for extraterrestrial radiation on a surface normal to the Sun is 1366 W/m² during daylight hours at all locations. This site also gives the extraterrestrial radiation on a horizontal surface, and the three categories of terrestrial radiation.

![Figure 6.10: Variation of Extraterrestrial Insulation as a Function of Julian Day](image)

The total solar energy striking the interior surface of an inscribing hollow sphere of radius \( R \) around the Sun is constant, and the surface area of that sphere is \( 4\pi R^2 \). The normalized distances of the Earth from the Sun vary from its perihelion of 0.9832622 to its aphelion of 1.016619. Taking the inverse of these numbers and squaring produces 0.967573 and 1.034335, respectively. Multiplying by the solar constant 1366 W/m² means that the actual solar constant varies from 1322 to 1413, a change 91 W/m², or ±3.34%. This solar radiation can be modeled with
where \( j_n = 360365.25 \) J\( n \) is the Julian angle expressed in degrees, with \( J_n \) being the Julian day; recall from Chapter 5 that the perihelion occurs at about J3, and its equivalent Julian angle is 3360365.25 = 2.96. A graphical presentation of Equation 6.7 is given in Figure 6.10 with ‘*’ indicating the extraterrestrial energy striking a normal surface on the first day of each month in the year 1990 at Boston, Massachusetts.

An 11.2 year solar cycle causes a variation in the solar constant of about 1 W/m\(^2\), but the elliptical orbit of the Earth is a much more significant factor.

### 6.3 Air Mass and Dimming

The attenuation through the atmosphere is dependant on the angle of the Sun to the zenith angle. As that angle increases, the length of travel of a photon through the atmosphere increases.

Air mass is defined as the relative length of travel of a photon through the atmosphere with respect to the thickness of the atmosphere \( h \). That is, if the zenith angle is zero the air mass is unity. If the zenith angle is \( \psi \), then the length of travel through the atmosphere is \( h / \cos \psi \). Thus, the air mass is \( 1 / \cos \psi \). In this analysis, as shown in Figure 6.11, it appears that we are advocates of the (now defunct) British Flat Earth Society.

A more accurate model that takes into account the curvature of the Earth is shown in Figure 6.12, where \( r \) is the radius of Earth. Applying Pythagoras’ theorem on the right angle triangle AOB produces 

\[
(r + h)^2 = (r + \ell \cos \psi)^2 + \ell^2 \sin^2 \psi,
\]

so

\[
r^2 + 2rh + h^2 = r^2 + 2r\ell \cos \psi + \ell^2 \cos^2 \psi + \ell^2 \sin^2 \psi,
\]

yielding

\[
\ell^2 + 2r \cos \psi \ell - 2rh - h^2 = 0.
\]

Solving the quadratic produces

\[
\ell = -r \cos \psi + r^2 \cos 2\psi + 2rh + h^2,
\]

where the positive sign is needed. The air mass is then

\[
M = \ell h = (r^2h^2 \cos 2\psi + 2rh + 1)^{1/2} - rh \cos \psi. \quad (6.8)
\]
Both models for the air mass assume that the atmosphere is homogeneous, so its density is constant.

With $r = 6371$ km and $h = 9$ km, the evaluation of Equation 12.3 involves the small difference between two large numbers, a numerically fraught situation, and care must be exercised.

When $\psi = 0$, $M = 1$. When $\psi = 90$, $M = (2rh+1)/2 = 37.64$. At solar noon, $\psi = L - \delta$, and air mass is given by

$$M_0=(r^22\cos(L-\delta)+2rh+1)/2-r\cos(L-\delta).$$

In both models the assumption is that the atmosphere is homogeneous, so its density is constant.

The air masses are calculated, both flat Earth and the more accurate measure taking into account the curvature of the Earth, in Table 6.3 as a function of $\psi$; the flat Earth result will be considered no further.

$1/M = h/l$ as a function of zenith angle is shown graphically in Figure 6.13. The reason for plotting $1/M$ rather than $M$ is the shape of the responses. We will see this in the next two figures.
Aerosols in the atmosphere increase air mass. The increase in sulphurous aerosols in the 1970s was determined to reduce the amount of sunlight reaching the ground by reflection in the atmosphere, a phenomenon known as “dimming.” One effect was increased river flow by up to 25% since evaporation decreased. Less evaporation means less rainfall, the source of most drinking water. Clean air legislation in Europe and the Clean Air Act in the United States have reversed this trend [“Solar dimming caused by air pollution increases river–flows,” October 5, 2014, http://phys.org/news/2014-10-solar-dimming-air-pollution-river-flows].
Terms such as dimming and the light extinction coefficient are often used, but such terms are imprecise. What is needed are definitive figures.

In the next chapter we will attempt to produce a model for insolation on the Earth’s surface for clear sky conditions, and then take away from these numbers as clouds increase in the atmosphere. The sky is clearest over the polar regions. It is also clearer with higher elevations. As one gets closer to the tropics, water vapor in the atmosphere increases, and we will attempt to model this phenomenon.

Aerosols are caused by dust, pollen, and incomplete combustion. There are places where solar energy reaching the surface is significantly attenuated by aerosols due to anthropogenic pollution; Beijing in recent years is one such place. However, in most locations the influence of aerosols is minor, and we will not attempt to model such attenuation.

![Figure 6.14: M at Winter Solstice](image)

The solar zenith angle as given by Equation 21 of Chapter 5 is

\[
\cos \psi = \sin \phi \sin \delta + \cos \phi \cos \delta \cos 15\tau \text{.} (6.10)
\]

The air mass with this zenith angle as a function of the time from solar noon and over a range of latitudes is shown at winter solstice for \( M \) in Figure 6.14, and for \( 1/M \) in Figure 6.15. The maximum possible air mass of 37.64 is shown in Figure 6.14.

Similar responses at equinox and at summer solstice are shown in Figures 6.16 through 6.18.

The attenuation of solar radiation is not directly proportional to the air mass. In the next section a model for attenuation due to Rayleigh scattering is presented that is a function of air mass.
How air mass attenuates otherwise remains to be determined.

6.4 Rayleigh Scattering

All solar energy in space is beam radiation. That is, there is negligible diffuse radiation. When that energy enters the Earth’s atmosphere some may be scattered, some may be absorbed, and the what reaches the Earth’s surface is beam radiation and largely isotropic diffuse radiation.

Sunlight is composed of bundles of energy termed photons that radiate outwards from the Sun in essentially straight lines. As a photon enter the Earth’s atmosphere it may come near or
strike a component of the atmosphere, which results in the trajectory of the photon radically changing, a phenomenon called scattering, or it may give up its energy to heat that component, or it may pass by without change.

When an oscillating electric field created by an electromagnetic wave interacts with an atom, electrons are moved back and forth, creating an oscillating dipole. An oscillating dipole radiates at the same frequency as the incident radiation. When a photon emanating from the Sun enters the Earth’s atmosphere and creates an oscillating dipole, it is scattered.

When the dispersion of photons occurs by particles that have a diameter less than approximately 1/10 the wavelength of the radiation, Rayleigh scattering occurs; in 1871 Lord Rayleigh published a paper describing this phenomenon. The deflection of an individual photon is a function of its wavelength to the power of minus four, so blue photons at 0.3μ are deflected $(0.70 / 0.3)^4=9.38$ times more than red photons at 0.7μ. Oxygen molecules are the dominant cause of Rayleigh scattering, and oxygen content of the air barely changes.

Rayleigh scattering results in the sky appearing blue on a clear day since an observer sees only scattered light, and the blue photons are those scattered the most. However, as the Sun sinks below the horizon and the blue light fades, we often see red and orange colors, especially if there are small particles in the atmosphere. After the Mount St. Helens eruption in Washington State on May 18, 1980, sunsets downwind were spectacularly red.

Given the strong scattering at short wavelengths, why does the sky not appear violet? There are two answers:
* The human eye is less sensitive to violet light, and
* There is much less energy in the violet wavelengths than in the blue wavelengths.
When a photon passes by a particle with a diameter on the order of the wavelength $\lambda$ of the photon, but larger than $\lambda/3$, Mie scattering occurs: See the next section.

For particles whose diameter is much greater than $\lambda$ the resulting Tyndall scattering produces diffuse radiation; this is not a significant factor.

When particles in the atmosphere are large, then their interaction with any photon, red through blue, occurs equally, and is responsible for the white appearance of clouds. Water droplets in clouds have a diameter on the order of $20\mu$. If the scattering due to water droplets favored the long wavelengths of the visible spectrum, clouds would appear red.
The Martian sky appears red in satellite images, which seems to violate the principles underlying Rayleigh scattering. However, dust storms on Mars can throw up iron-rich red soil, and this makes it the “red planet.”

Mars has lost most of its atmosphere, and is destined to lose what is left. Its problem is the lack of a magnetic field such as Earth has, which channels the “solar wind” towards the poles. The solar wind can knock components of the atmosphere free from Mars’ gravitational pull, and so are lost to space [K. Chang, “Solar Storms Strip Air From Mars, NASA Says,” New York Times, November 6, 2015].

In space there is no color, only black and white. The Sun and the stars appear white, and the rest of space is black. There is no scattering in space. The exception is a planet illuminated by the star around which it is orbiting.

The attenuation $1 - \tau$ caused by Rayleigh scattering can be described mathematically [M. Iqbal, An introduction to Solar Radiation, Academic Press, New York, 1983] as

$$\tau_r = e^{-0.008735M\lambda - 4.08}$$ (6.11)

where $\tau_r$ is the transmissivity, a number less than unity, $M$ is the air mass, and $\lambda$ is the wavelength.

For $M = 1$, if we integrate Planck’s law and Equation 6.11 over the range $0 \leq \lambda \leq 2.7$, the results are responses shown in Figure 6.19. The final values for Planck and Rayleigh are 1331.2 W/m$^2$ and 180.2 W/m$^2$, respectively. Similarly, for $M = 2$ the responses are as shown in Figure 6.20: Rayleigh scattering amounts to 278.7 W/m$^2$.

Figure 6.21 shows that Rayleigh scattering is dominant in the short waves only; here $M = 1$. There is a similar result for $M = 2$ shown in Figure 6.22.
Figure 6.19: Rayleigh Scattering with $M = 1$

Figure 6.20: Rayleigh Scattering with $M = 2$
Scattering can follow a particular pattern. In Rayleigh scattering, the pattern can be described by

\[ I = I_0 k \lambda^4 (1 + \cos^2 \phi) \quad (6.12) \]

where \( I_0 \) is the insolation of a beam of light entering the Earth’s atmosphere, and \( \phi \) is the scattering angle. As can be seen from the plot of \( 1 + \cos^2 \phi \) shown in Figure 6.23, the major axis of the scattering is the direction of the beam of light, and the back scattering is an exact duplicate of the forward scattering. Thus, the maximum percentage of Rayleigh scattered energy available on the Earth’s surface is 50%.

When a photon is Rayleigh scattered it retains all its energy and its particular wavelength. When absorbed it releases all its energy in the form of heat. The body heated by this photon re-radiates energy at a significantly lower temperature than the body that issued the photon — the Sun. Rayleigh and Mie cause elastic scattering, meaning that individual photons retain all their energy after scattering. Contrast this to the non-elastic Raman scattering in which energy in an individual photon can change [http://hyperphysics.phy-astr.gsu.edu/hbase/atmos/blusky].
Suppose beam of light traveling along the z-axis in Cartesian space vibrates in the x-y plane. If this vibration is the same in all directions in the x-y plane, the light is said to be unpolarized. Sunlight is unpolarized as it enters Earth’s atmosphere. When the vibration favors a particular direction, the light is said to be polarized. When Rayleigh scattering occurs, a photon is not only displaced from its original trajectory, it also becomes polarized.

The first person to study the scattering of light experimentally was Tyndall. In Rayleigh scattering the deflected photon retains its wavelength; this is termed coherent scattering. Contrast this to incoherent scattering in which the photon is absorbed and reradiated at a different wavelength.
6.5 Mie Scattering

Mie scattering, named after German physicist Gustav Mie (1868–1957), occurs when a photon passes by a particle with a diameter on the order of the wavelength $\lambda$ of the photon, but larger than $\lambda/3$. Typically these particles are water vapor, dust, or industrial air pollution. For particles $\lambda$ the resulting Tyndall scattering produces diffuse radiation; this is not a significant factor.

Mie scattering is not a strong function of wavelength, so it can be termed white light scattering: this becomes apparent when the sky color when looking close to the Sun is almost white.

Mie scattering occurs when a photon with wavelength $\lambda$ passes close to a particle of diameter $d$, where $\lambda/3 \leq \delta \leq \lambda$. Such particles are dynamic in the atmosphere, and their density varies widely. Mie scattering is the reason why wet, coastal skies are whiter than dry, mountainous skies.

Mie scattering favors the direction of the beam of light, and is strongly influenced by the density of the scattering particles. For low concentrations, as in the bottom pattern shown in Figure 6.24, Mie scattering looks much like Rayleigh scattering, but for high concentrations (as shown in the top pattern) the forward lobe is strongly favored.

Scattering has two principal benefits for life on the Earth. The first, caused by Rayleigh scattering but not Mie scattering, is the significant reduction in ultra-violet radiation and the subsequent reduction in cases of skin cancer; in space, solar radiation directly on the skin would result in a high risk of skin cancer. The second is the diminution of shadows. Without scattering, both Rayleigh and Mie, the difference in light intensity from objects partially illuminated by sunlight and the rest in shadow would make the shadowed areas appear far darker, even totally black.

6.6 Solar Radiation Measures
Here are the commonly accepted measures for solar radiation:

**Normal Radiation** is the energy striking a surface normal to the direct energy beam.

**Beam Radiation**, also known as direct normal radiation, is the amount of solar radiation from the direction of the Sun, so it excludes diffuse radiation.

**Circumsolar Radiation** is the amount of solar radiation coming from a disc centered on the Sun disk and having a radius of about 3 degrees.

**Diffuse Radiation** measures scattered and reflected energy striking a horizontal surface, and specifically excludes direct radiation from the Sun. Diffuse radiation is measured over the complete sky dome. That is, the entire sky from horizon to zenith in all directions.

**Global Horizontal Radiation**, often called just **global** for short, is the total solar energy striking a horizontal surface, so it includes beam radiation and diffuse radiation. It also includes ground-reflected radiation, which is usually insignificant and can be ignored.

**Global-Vertical South Radiation** is the total energy striking a vertical, south-facing surface. (Also defined for east, west, and north vertical surfaces.) **Global Latitude inclined** measures the total energy striking a south-facing (in northern latitudes) surface inclined at an angle to the horizontal by the latitude $L$.

Since global horizontal radiation $G$ contains the beam $B$ and diffuse $D$ radiations, the simple formula

$$G = B \cos \Psi + D(6.13)$$

can be used, where $\psi$ is the zenith angle of the Sun.

The standard measure for solar radiation is $W/m^2$. An alternative, now rarely used, is kilojoules per minute per meter squared, kJ/min·m$^2$, where 1 Kj/min·m$^2 = 16.663$ W/m$^2$. In this measure the solar constant is 81 kJ/min·m$^2$. Another measure, also rarely used when considering solar energy, is the Langley, named after the solar energy researcher at the Smithsonian Institution, Samuel P. Langley (1834–1906), where one Langley equals one calorie per square centimeter, cal/cm$^2$.

The term beam radiation implies it comes from a point source, which it does not. The average distance of the Earth from the Sun is $1.49660 \times 10^{11}$ meters, and the radius of the Sun is $6.955 \times 10^8$ meters, so the angle of the Sun as perceived from the Earth as calculated in Figure 6.25 is $\phi_2 = \pm \sin^{-1} rD = \pm \sin^{-1} 16.955 \times 1081.49669 \times 1011 = \pm 0.26626$. 

![Figure 6.25: The Angle Subtended by the Sun](image)
When the sky is clear enough so that the Sun casts clear shadows, the diffuse radiation can be as low as 70 W/m² under clear, dry conditions, or as much as 280 W/m² for high humidity conditions.

The use of the term insolation to describe solar radiation is losing favor, mainly because the word can be confused with the word insulation.

Latitude inclined surfaces are the accepted norm for year round photovoltaic collection. The same is not true for solar thermal using flat plate collectors, or concentrating collectors in residential use. The heat created to be used as space heating is needed in winter. Such collectors may also heat the domestic hot water supply, so some collection is needed in summer. At latitude 40 north, the south-facing roof in which flat plate collectors are mounted should be inclined at an angle greater than 40° to the horizontal. An angle of 60° is probably close to optimal.

6.7 Instruments for Measuring Solar Radiation

6.7.1 Pyranometers

A pyranometer (from the Greek word for fire) measures solar radiation at the Earth’s surface. The most important feature of a pyranometer is a grouping of thermocouples, called the thermopile, that heats up when light falls on it and generates a small voltage. The voltage produced depends on the intensity and wavelength of the incident radiation.

A thermocouple is comprised of two dissimilar metal wires joined at what is known as the hot end. The open end of the wires is called the cold end. When the hot end is heated a small voltage can be detected across wires at the cold end. This voltage is small in millivolts. In pyranometers the hot end is painted black to help absorb the solar energy, and the cold ends are painted white to reflect the solar energy.

The pyranometer is commonly set to measure radiation received on a horizontal surface, so the pyranometer should be mounted horizontally with an unobstructed view of the full sky in all directions. On a clear day, the instrument has a maximum response at solar noon.

A pyranometer can separate the direct and diffuse solar radiation. First the pyranometer must be mounted on a rotatable baseplate whose axis of rotation passes through the thermopile. A simple scaffold holds a disc about 1m from the sensor, as shown in Figure 6.26, such that the distance of the disc from the scaffold is adjustable in the horizontal plane. The base plate is rotated so that the shadow of the disc and the horizontal bar holding it is in line with the thermopile. This bar is then extended or shortened so the shadow of the disc is directly over the thermopile. Orientation is then not a problem since the response of the sensor is measured in
seconds. The size cast by the disc should be more than that subtended by the Sun, $0.533^\circ$, and less than $5^\circ$.

![Figure 6.26: Mechanism to Measure Diffuse Radiation](image)

The pyranometer can be tilted to directly face the Sun and so record the normal incident radiation, a much better reading to have for the purposes of designing a solar heating system. Since the response time is so short, this can be done manually using the simple mechanism shown in Figure 6.27. Here a slender rod is mounted on the same surface as the pyranometer and normal to the surface. The base plate needs to be orientable over two orthogonal axes. The instrument is correctly aligned when the shadow cast by the Sun is coincident with the base of the rod.

To measure only beam radiation, a cylinder should be centered between the sensor and the Sun with the mechanism shown in Figure 6.27. The interior of the cylinder should be non-reflective.

One of the best instruments, and one of the most expensive, is the Eppley PSP (Precision Spectral Pyranometer), whose response ranges from $0.285\mu$ to $2.8\mu$. 
Its specifications include:
* Sensitivity approximately 9 microvolts per W/m².
* Temperature dependence is 1% over -20 to +40°C
* Linearity is plus or minus 0.5% from 0 to 2800 W/m²
* Response time is 1 second
* Cosine response: 1% for 0 to 70 degrees from zenith; 3% from 70 to 80 degrees.
While thermopile-based pyranometers such as an Eppley PSP have fairly uniform spectral response over the solar spectrum, inexpensive solar-cell-based pyranometers, such as the LI-COR, have a constrained spectral response: They respond to wavelengths above 0.4 μ, and have maximum sensitivity at 0.9 μ, but then the sensitivity drops precipitously as shown in Figure 6.28. The LI-COR does not directly measure the total solar energy. Instead, the voltage output is the integral of the product of the response times the spectral intensity with respect to wavelength.

It is possible, but rare, for the measured irradiance to exceed the solar constant. It requires one or more additional light sources to be present: The most common phenomenon is for a tall thunderstorm cloud or body of water reflecting sunlight onto the sensor.

### 6.7.2 Other Radiation Measurement Instruments

A pyrheliometer with an aperture angle of about 5° measures the circumsolar radiation as well as the beam radiation. This is achieved with a collimation tube, a cylinder that fits above the detector with a non-reflecting interior so only rays within that 5° angle can reach the sensor. A quartz window permits only radiation in the solar range, 0.3μ to 3μ, to pass through. The instrument can be fitted with a Sun tracker, or the simple scheme shown in Figure 6.27 can be employed.

To measure beam radiation only would require a pyrheliometer with an aperture angle of 0.53252° and a highly accurate Sun Tracker, conditions very difficult to obtain. So we live with existing instruments, knowing that the numbers recorded over-estimate the beam radiation.

Beam radiation can be measured indirectly with a shadow ring pyranometer, such as the Kipp-Zonen 121B [F. Bason, “Diffuse Solar Irradiation and Atmospheric Turbidity,” available online], which rotates slowly. When the shadow ring blocks the beam radiation it records the diffuse radiation \(I_d\), otherwise it records the global radiation \(I_g\). The beam radiation is \(I_b = I_g - I_d\).
A simplified sketch of the time response of the shadow ring pyranometer is shown in Figure 6.29. The shadow ring should be wide enough so that the thermopile can record the minimum reading when the beam radiation is occluded. A small correction factor is used since the shadow ring shades a complete band of diffuse radiation.

Pyrogeometers detect infrared radiation by shielding the thermopile sensor with a solar blind filter. The shield can be chosen to pass a specific waveband.

The albedo of a body is a number that defines the reflective properties of a non-luminous body, where 0 is a perfect absorber and 1 a perfect reflector. Albedo can also refer to broad band irradiance of daylight. Albedo is discussed in depth in the next chapter.

A pyranometer can be used to determine the outdoor albedo by taking two measurements, the first the zenith reading, and the second the nadir reading, by measuring the reflected illumination from a test surface. The ratio of the two determines the albedo. Care must be taken in taking the nadir reading — setting the pyranometer too close to the test surface will make it cast a shadow on it, but set it too high and the instrument will include reflection from other than the test surface.

An albedometer consists of two pyranometers, with the upper one measuring the global solar radiation and the lower one the solar radiation reflected from the surface below, which is shielded from the direct solar radiation. The ratio of the reflected radiation to the global radiation is the albedo, a number between 0 (no albedo) and 1 (perfect reflection, no absorption at the surface). By means of comparison, the albedo of grass is about 0.15, and of fresh snow is 0.8.

As defined by WMO (World Meteorological Organization), a sunshine duration sensor measures the time beam radiation exceeds 120 W/m². WMO was established in 1950 as an agency of the United Nations devoted to consideration of the interaction of the Earth’s atmosphere with climate, the oceans, and water resources.
There are a number of specialized instruments for measuring solar radiation. The radiometer measures the direct and diffuse solar radiation by deleting the reflected solar radiation. The quantum sensor measures the number of photons striking a surface, also known as the photosynthetic photon flux density.

### 6.8 Isotropic and Circumsolar Radiation

We constructed a polar plot of the Rayleigh scattering pattern $1 + \cos^2 \phi$ in Figure 6.23. If we now split the 1 from the $\cos^2 \phi$, the resulting polar plot is shown in Figure 6.30. The unity circle produces isotropic diffuse radiation on the Earth’s surface. One of the two egg-shaped responses produces radiation that reaches the ground with the maximum value in the same direction as the beam radiation. The other egg-shaped response is reflected back away from the ground.

Integrating the two parts of the Rayleigh scattering pattern produces

$$
\int_0^{2\pi} 1 \, d\phi = \int_0^{2\pi} \cos^2 \phi \, d\phi = \int_0^{2\pi} \phi + 14 \sin 2\phi \, d\phi = \pi, \quad (6.14)
$$

so the first term contains $2/3$ of the scattered energy and is isotropic, and the second term contains the remaining $1/3$ of the scattered energy in two lobes. One of the lobes produces energy that could point towards the Earth’s surface. The energy in this lobe over a range of $\pm f$ is shown in Figure 6.31 and given in Figure 6.4. Notice that over the range $\pm 30^\circ$ the energy is 0.478, or almost 61% of the total energy in this lobe.
If we consider the potentially useful energy in this forward lobe to be in the angle range $\pm 90^\circ$, then the total of this energy is $1/12$ of the Rayleigh scattered total.

On a clear day with beam intensity 880 w/m$^2$ and with air mass of unity, the total scattered energy as illustrated in Figure 6.19 is about 180 w/m$^2$, and of this scattering adds 30 w/m$^2$ to the directed energy, and 60 w/m$^2$ to the diffuse radiation; the Rayleigh scattered component of the directed energy is just 3% of the initial beam energy.

This is not the whole story. The energy in the back lobe is 30 w/m$^2$ with axis pointed at the Sun, so the energy re-scattered from this lobe is $30 \times 0.2 = 6$ w/m$^2$, of which 4 w/m$^2$ is added to the diffuse radiation.
We cannot consider the remaining 2 W/m² to be in two lobes with the same angular pattern as the lobes after the initial scattering. The energy in the original incoming beam appears 0.2662° wide to an observer, but the back scattered lobe is far less coherent. If anything, the $\cos^2 \phi$ component of the re-scattered light adds to the diffuse component, but it is so small that the diffuse radiation can still be considered isotropic.

Increasing the air mass to two, the total scattered energy as illustrated in Figure 6.20 is about 280 W/m², with 47 W/m² added to the circumsolar energy, and 93 W/m² added to the diffuse radiation. Re-scattering adds $47 \times 0.2 = 9.4$ W/m² to the diffuse radiation.

Surveying the extent of diffuse radiation at different locations and on clear days, we see that diffuse radiation is about 10% of the direct radiation. We have seen that the diffuse component of Rayleigh scattering that reaches the Earth’s surface is also about 10% of the direct radiation. The inescapable conclusion is that when diffuse radiation is at a minimum it is caused almost exclusively by Rayleigh scattering. This suggests that one could build a model for attenuation of solar radiation by the atmosphere that starts with clear day numbers, and then adds modifying factors until the model is complete.

Is circumsolar radiation with the beam radiation removed ever significant? We refer to a study performed at UC Berkeley [D. Grether, D. Evens, A., Hunt and M. Wahlig, “The Effect of Circumsolar Radiation on the Accuracy of Pyrheliometer Measurements of the Direct Solar Radiation,” Lawrence Berkeley Laboratory, UC Berkeley, presented at the annual meeting of the American Section of the International Solar Energy Society, Philadelphia, May 1981]. They defined one solar radius $\theta_r = 0.25°$, close to the true figure of 0.26626°. The radial scan of the telescope taking the measurements was $\theta_s = 3°$, about the same as the field of view of a typical pyrheliometer. They then defined the direct solar radiation $S$ and the circumsolar radiation $C$ as follows:

$$S = \int \theta r I(\theta) d\Omega = \int \theta \theta s I(\theta) d\Omega \quad (6.15)$$
where $I(\theta)$ is the direct incident radiation measured at angle $\theta$, and $\Omega$ is the solid angle.

A solid angle is defined in steradians $\omega$, where $\omega = A/r^2$ and $A$ is the area projected onto the surface of a sphere by a cone with apex at the center of the sphere of radius $r$. The surface area of a sphere of radius $r$ is $4\pi r^2$, so if $A = 4\pi r^2$ then $\omega = 4\pi$.

$C + S$ approximates the normal incident radiation $I(0)$ measured by the pyrheliometer, so $CC+S$ approximates the overestimate fraction of the direct solar radiation.

Two cases were considered in the study [op cit]. The first was a clear sky case in the Mojave desert area of California with $I(0) = 932$ w/m$^2$, which had $R < 1\%$. The second case at Atlanta, Georgia, had hazy skies with $I(0) = 623$ w/m$^2$ and $R > 20\%$, or contributing over 125 w/m$^2$ to the total measured solar radiation over the $\pm 3^0$ scan. Thus, with clear sky conditions, the circumsolar radiation outside the $\pm 0.25^\circ$ focus on the Sun is negligible, but with partially cloudy conditions this radiation is significant.

It is unfortunate that, other than the paper cited, there are few experimental studies that use telescopes to measure solar radiation, both beam and diffuse, for a narrow angle that concentrates on angles within $3^\circ$ of the center of the Sun.

How much does the circumsolar radiation add to the direct normal measurements of a pyrheliometer? Quite a bit. “Due to forward scattering of direct sunlight in the atmosphere, the circumsolar region closely surrounding the solar disk looks very bright. The radiation coming from this region, the circumsolar radiation, is in large part included in common direct normal irradiance (DNI) measurements, but only partially intercepted by the receivers of focusing collectors. This effect has to be considered in the performance analysis of concentrating collectors in order to avoid overestimation of the intercepted irradiance. At times, the overestimation reaches more than 10% for highly concentrating systems even for sky conditions with relevant DNI above 200W/m$^2$. The amount of circumsolar radiation varies strongly with time, location and sky conditions. However, no representative sunshape measurements exist for locations that are now of particular interest for concentrating solar power (CSP) or concentrating photovoltaics (CPV) [S. Wilbert, B. Reinhardt, J. DeVore, M. Röger, R. Pitz-Paal, C. Gueymard, and R. Buras, “Measurement of Solar Radiance Profiles With the Sun and Aureole Measurement System,” Solar 2002, Reno, Nevada].”

The circumsolar radiation will be captured by all non-concentrating solar collectors that can collect beam radiation. It will also be captured by low-concentration solar collectors, such as the parabolic trough concentrators with concentration factors less than five.

years of measured data show diffuse to be 38% of total radiation on horizontal surface. On high mountains, as for example Mount Evans, Colo., at an elevation of 14,259 feet, the diffuse radiation from cloudless skies amounts to only 4 to 5 percent of the total solar and sky radiation received on a horizontal surface at noon in midsummer.... Values of diffuse radiation during cloudless conditions are much smaller at high elevations than at sea level or at comparatively low altitudes. Cloudiness is the predominant factor in causing daily variations in the ratio of the diffuse to total radiation; the cloudiness itself is dependent on factors such as time of year, time of day, local geography, nearness to large bodies of water, and the synoptic weather situation.”

The two studies cited above indicate that one can glean a lot of information from experimental data, in particular under clear sky conditions. Cloudiness reduces the beam radiation, even eliminates it completely, and the diffuse component rises from a low base value to a peak before dropping off to almost zero under total overcast conditions. Both studies hint that local conditions determine how increasing cloudiness affects diffuse radiation.

Ground reflection produces horizon brightening. It is most pronounced under clear sky conditions but declines rapidly as cloudiness increases [J.A. Duffie and W.S. Beckman, Solar Engineering of Thermal Processes, Fourth Edition, John Wiley, Hoboken, New Jersey, 2013]. Horizon brightening will always be a small percentage of the energy reaching solar collectors.
Chapter 7

Solar Gain and Solar Attenuation

We start by defining the cosine of the angle of the Sun on a fixed-orientation planar surface, using the formula derived in Chapter 5, as the solar fraction. We also define the integral of this solar fraction from solar noon until the collection day ends as the half day solar gain, or solar gain for short.

Graphs of the solar fraction for the two solstices and equinox are presented for a range of latitudes, together with the solar gains. These numbers represent the ideal, with no reductions for air mass or any other atmospheric condition that almost certainly will reduce the incoming solar radiation. Neither does it take into account glazing losses or other losses that are significant in all solar collectors. However, the responses and numbers generated provide a definitive ceiling to assess potential solar energy as a function of latitude. Also, the seasonal solar gains vary widely depending on latitude, providing insights into architectural considerations such as shading the summer Sun.

The previous chapter presented the nature of solar energy. We saw that the solar constant averages 1366 W/m\(^2\), but terrestrial solar energy is typically less than 1000 W/m\(^2\) on a clear day, and this is for a clear day at elevations such as Denver, Colorado. At sea level, such as Boston, Massachusetts, this figure is closer to 880 W/m\(^2\), or 64% of the solar constant. The model used to describe Rayleigh scattering established that near noon when the air mass is close to unity it scatters 180 W/m\(^2\). Of the 64%, 48% is beam radiation and 16% is scattered or diffuse radiation. This leaves 1366 − 880 − 180 = 306 W/m\(^2\) unaccounted for. Attempts will be made to explain what happened to this radiation.

An analysis of the nature of the solar radiation in space is definitive, but the attenuation in the Earth’s atmosphere is highly variable. There are many models for this attenuation, but not much on the methodology that led to their development.

7.1 Solar Fraction and Solar Gain
In Section XI of Chapter 5 we discussed the solar collection day, which ends when the solar elevation $\alpha$ drops below zero: This is a necessary but not sufficient condition. The necessary and sufficient conditions are:

$$\alpha^* > 0 \text{ and } \cos \theta = \cos \alpha \cos \phi \cos (\beta - \gamma) + \sin \alpha \sin \phi \leq 90^\circ$$

The second condition is Equation 25 of Chapter 5, where $\theta$ is the angle of the Sun on the collector, $\beta$ is the solar azimuth, $\phi$ is the angle of the collector with respect to vertical, and $\gamma$ is the yaw angle of the collector with respect to due south in northern latitudes or due north in southern latitudes.

The solar fraction is the amount of solar energy striking a unit surface compared to a unit surface normal to the incoming solar energy on a clear day. It is a useful base to work from, while acknowledging that it is the highest amount of the incoming solar radiation that can possibly be collected. There is no accounting for the reduction in solar energy due to air mass, cloud cover, or pollutants.

Assuming constant normal radiation throughout the day, how much solar radiation is incident on the surface of a collector? We do this by calculating the half day solar gain $G$ from noon by numerical integration. The integration formula chosen here is the trapezoidal rule

$$G = hI_s\Sigmai=1n−1 (\cos \theta_i + \cos \theta_{i+1})(7.1)$$

where $h$ is the time step, $\theta_1 = 0,$ and $\theta_n = 90$: the imprecise $\theta_n = 90$ means that the integration ends when $\cos \theta_{i+1}$ increases rather than decreases. It can be argued that when $\theta$ increases could be at the beginning of the step close to $\theta_i$ or near the end and close to $q_{i+1}$. A correction could be made, but is unnecessary if the step length is reasonably small, such as 0.25 hours. Further, close to $\theta = 90$ the solar collection is minimal.

The trapezoidal rule is second order correct, meaning that the error is a function of $h^3$. It is also unconditionally stable, a desirable attribute in most circumstances, but probably not essential to have in this situation [A. Ralston, A first Course in Numerical Analysis, MaGraw Hill, New York, 1965].

In the rest of this section we investigate the solar energy on surfaces at latitudes from 30 to 60 with zero yaw angles. The abscissae $\tau$ in the diagrams that follow is the time in hours from solar noon and until the solar collection day ends.

The three essential times of the year are winter solstice, equinox, and summer solstice, marked $W$, $E$, and $S$ in the diagrams that follow. Also shown are the half day solar gains, defined as the integral of the solar fraction over the time from solar noon until the solar collection day ends. Since the yaw angle is zero, the morning collection is the mirror image of the afternoon collection.
Figures 7.1 through 7.4 are when $\phi = 0$, so the collection surfaces are vertical, for latitudes 30, 40, 50, and 60. Following the analysis in Chapter 5, we know that for all declinations $\delta < 0$ the collection day always ends when $\alpha = 0$ and $\beta < 90$. We know that at equinox, the collection day ends at $\tau = 6$, $\alpha = 0$, and $\beta = 90$ for all latitudes. We also know that for all $\delta > 0$ the collection day ends when $\beta = 90$. 

Figure 7.1: Solar Fractions at L=30

Figure 7.2: Solar Fractions at L=40
Shown in 12.2 is the solar fraction for latitude 30 at the three essential times. The difference between the half day solar gains at W of 3.23 and at S of 0.21 is striking, even though no shading was considered in the calculations; see Section VII of this chapter. The sum of the three gains is 5.55. The collection day ends at $\tau = 5.03$ at W, and $\tau = 2.76$ at S.

The responses at latitude 40 are shown in Figure 7.2. Here the gains at W and S are 3.49 and 0.73, respectively, so the summer gain is considerably greater than occurs for latitude 30 and indicates the need for some measure of shading. The collection day ends at $\tau = 4.58$ at W, and $\tau = 3.93$ at S.

![Figure 7.3: Solar Fractions at L=50](image)

The responses at latitude 50 are shown in Figure 7.3. Here the solar gains at W and S are 3.30 and 1.33, respectively, so the summer gain is considerable and indicates the need for a considerable amount of shading. The collection day ends at $\tau = 3.93$ at W, and $\tau = 4.58$ at S.

The responses at latitude 60 are shown in Figure 7.4. Here the solar gains at W and S are 2.55 and 1.93, respectively, but the angles involved make shading difficult. The winter collection day ends at $\pm 2.76$ hours when $\beta = \pm 37.3$. The collection day ends at $\tau = 2.76$ at W, and $\tau = 5.04$ at S. Utilizing solar energy at latitude 60 or greater is difficult.

The next six diagrams cover cases when $\phi \neq 0$. The first, Figure 7.5, is for $L = 30$ and $\phi = 60$, so the collector surface is latitude inclined. Notice that the responses for W and S coincide. This is to be expected when the collector is latitude inclined, the inclination considered close to ideal for maximum year-round solar collection. However, the W response ceases at $\tau = 5.03$ when $\alpha = 0$ with a half day solar gain of 3.29. The gain at E is a little higher at 3.82. At S the collection day ends when $\alpha = 11.4$ and $\theta = 90$. 
We will see that for latitude-inclined collectors and all latitudes:

* the collection day ends at $\tau = 6$ for all $\delta \geq 0$,
* the summer gain is 3.50, and
* the equinox gain is 3.82.

Figure 7.5: $\phi=60$ at L=30

Figure 7.6 shows the solar fractions with tilt of 30 at L=30, and shows the winter gain is the highest and summer gain the lowest. The total gain, the sum of the solar gains for W, E, and S, is 9.18, lower than the sum when the collector surface is latitude inclined at 10.72. At W the collection day ends when $\alpha = 0$ and $\beta = 62.6$, the same azimuth as occurs at the end of the collection day in Figure 7.5, as expected.
At $L = 40$ and latitude inclined, the winter collection day ends at $\tau = 4.58$ with a gain of 3.39, a small reduction over the summer gain of 3.50. With $\phi = 30$ at $L = 40$, the winter gain rises to 3.69, but the summer gain drops to 2.52.

At $L = 50$ and latitude inclined, the winter solar collection day ends at $\tau = 3.93$ with a gain of 3.00. The summer gain is 3.50 and the total gain is 10.32. At $L = 60$ and latitude inclined, the winter solar collection day ends at $\tau = 2.76$ with a gain of 2.76. The summer gain is 3.50 and the total gain is 9.63.

The sum of the gains at the three essential times of the year are given in Table 7.1 as a function of $L$ and $\phi$. For $\phi = 0$ the gain increases with $L$, whereas the opposite is true for $\phi = 90 - L$. 
Further, the numbers for latitude inclined are considerably greater than for vertical collectors, especially for smaller latitudes.

![Figure 7.8: $\phi=30$ at $L=40$](image)

![Figure 7.9: $\phi=40$ at $L=50$](image)

### 7.2 Albedo and Specularity

Solar radiation is comprised of photons, massless bundles of electromagnetic energy with wavelengths predominantly in the waveband $0.3\mu$ to $2.7\mu$; this waveband holds 95% of the total solar energy, with 2.5% at shorter wavelengths and 2.5% at longer wavelengths. What happens to a photon depends on the albedo of the surface it strikes.
Albedo defines the reflective property of a non-luminous object, a number between zero and unity, where zero is a perfect absorber, and unity is a perfect reflector. The albedos of common terrestrial materials are given in Figure 7.11 [www.eoearth.org/view/article/149954]. Fresh snow has the highest albedo, reflecting 95% of the incident solar radiation. At the other end of the scale is the coniferous forest that reflects only 5–15% of the solar energy, and dark wet soil reflects 5%. The albedo of the ideal solar collector is zero. The designation $\alpha$ in “high $\alpha$” and “low $\alpha$” refers to the solar elevation angle.

![Figure 7.10: $\phi=30$ at $L=60$](image)

<table>
<thead>
<tr>
<th>$L$</th>
<th>$\phi=0$</th>
<th>$\phi=90-L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>5.55</td>
<td>10.72</td>
</tr>
<tr>
<td>40</td>
<td>6.68</td>
<td>10.59</td>
</tr>
<tr>
<td>50</td>
<td>7.56</td>
<td>10.32</td>
</tr>
<tr>
<td>60</td>
<td>7.80</td>
<td>9.63</td>
</tr>
</tbody>
</table>

As shown in Figure 7.12, the average albedo of the land areas of the Earth is between 0.2 and 0.3, while the average albedo of the oceans is about 0.08. 92% of the solar energy penetrates the water to a depth of about 2 meters, warming the water and causing evaporation, which increases cloud cover and produces rain. There are no albedo numbers shown for arctic and near arctic regions: These areas have high albedo in winter and much of the year due to snow and ice cover.

Clouds reflect about 50% of the incoming solar radiation, with thick clouds reflecting up to 90% and thin clouds as little as 30%. The Earth’s atmosphere serves as a giant greenhouse that traps enough solar energy, both direct and re-radiated from the surface, to raise the Earth’s average temperature by about 30°C, from –18°C to 14°C. Thus, this greenhouse effect is
beneficial, unlike the effect of burning fossil fuels that increase greenhouse gases and cause global warming, which all but a few climate deniers believe is detrimental.

Some of the factors affecting albedo include
color — light colors have high albedo
Albedo is considered to be isotropic, meaning that the light it reflects is equal in all directions. Contrast this to specularity, a term used in computer graphics, where the angle of reflection $\theta$ equals the angle of incidence as shown in Figure 7.13; the reflection is anisotropic. The albedo reflection is shown as a semicircle.

Matte surfaces have low specularity. Surfaces with high specularity include mirrors and polished metals.

Reflections from specular surfaces have the same color as the incident radiation; if it is from the Sun, the reflection is white. In contrast, reflections from an albedo surface have the same color(s) as that surface. Specular reflection obeys Fresnel’s equation. Chapter 8 discusses the application of Fresnel’s equation to the transmission of light through glass. Here, all we need to know is that light striking a body with a specular surface will reflect some of the light as a function of the refractive index of that body. All bodies have a refractive index, even vegetation. The refractive index of air is unity.

The energy balance for light striking the surface of some body, where the incoming energy is normalized to unity, can be written as

$$1 = R + T + A = R_s + R_d + T + A, \quad (7.2)$$

where $R$ is total reflectance, $R_s$ is specular reflectance, $R_d$ is diffuse reflectance, $T$ is transmittance, and $A$ is the energy absorbed by the body. The photons absorbed heat the body, but the photons reflected or transmitted survive and have no heating effect on the body. If $R_d$ is high we say the body has a high albedo.

Some say that the photons do not bond with the specular surface and so bounce off without modification, hence they do not reflect the surface color. On the other hand, the photons bond with the low albedo surface and adopt the surface color. Some refer to a photon bond with an albedo surface, compared to a geometric interaction of a photon with a specular surface.
Albedo is sometimes called body reflection, whereas specularity is called surface reflection. In a sense this is saying that the photons striking the albedo material have an intimate contact, but the photons striking the specular surface are immediately reflected coherently. The sum of body plus surface reflections produces the intensity of the image.

Most of Earth’s land areas are covered with vegetation, so determining the interaction of solar radiation on leaves is of interest. Studies have shown that leaves absorb most of the energy in visible light, but reflect a significant amount of the infrared radiation [D. Baldocchi, “Solar Radiation Transfer Through Vegetation, Part 2, Lecture 9,” Department of Environmental Science, Policy and Management, UC Berkeley, September 2012]. [S. Jacquemond and S.L. Uctin, “Modeling Leaf Optical Properties,” www.photobiology.info/Jacq.Uctin].

An empirical model to describe specular reflection is attributed to Phong Bui-Tuong [“Illumination for Computer Generated Pictures,” Communications of ACM, vol 18 #6, p 3110317, June 1975]. For a surface with high specularity, he observed that the region of observed optical intensity was small while the intensity itself was high, but for a surface with low specularity, that the region of high intensity was large but its intensity was low. His simple formula for reflected intensity is

\[
R_s = \cos^n \alpha, (7.3)
\]

where \( \alpha \) is the angular deviation from the reflected angle \( \theta \) that would occur for ideal specularity, and \( n \) is a measure of the specularity of the surface, where a matte surface has small \( n \) while a shiny surface has \( n \geq 200 \). Plotting Equation 7.3 for a range of values of \( n \) produces Figure 7.14.

As it stands, Equation 7.3 is not useful, and Phong combined it with basic information about the illumination to produce
where the dots in the formula indicate vector dot products, \( \cos^n \alpha = v_r \cdot v_s \), \( I_a \) is the intensity of the ambient light, \( I_i \) is the intensity of the beam radiation, \( K_a \) is the ambient reflection coefficient, \( K_d \) is the diffuse reflectivity coefficient, and \( K_s \) is the specular reflection coefficient. All coefficients have values that lie between zero and unity. All vectors are unit vectors and lie as shown in Figure 7.15. Without loss of generality, we assume the incident beam radiation strikes the surface at the origin, and all vectors meet at the origin.

The Phong model is a valuable tool, but has its limitations. Determination of the insolation intensities and the coefficients is neither easy or precise, particularly when one tries to factor in the wavelength dependence of each component. However, when it comes to assessing solar energy on the Earth’s surface, nothing is precise, as will be seen in the discussion later in this chapter of the models for insolation dispersion in the atmosphere.

7.3 Air Temperature, Pressure, and Clouds

Most people on this planet do not experience oxygen deprivation, since they live close to sea level. Even at one mile above sea level, in Denver, Colorado for example, there is no discomfort. However, a visitor traveling from Denver to California can experience a loss of breath after a minor exertion on a high pass in the Rockies. Here we look into oxygen content in the atmosphere with a view to identifying what is constant and what is variable in scattering.

The density \( D \) kg/m\(^3\) of dry air is given by the formula

\[
D = PR \times T (7.5)
\]

where \( P \) is the pressure in Pascals, \( R \) is the specific gas constant for dry air 287.05 J/kg\( \cdot ^\circ \)K, and \( T \) is the temperature in degrees Kelvin. At sea level, \( P = 101.3 \) Pascals, but this is often
rounded to 100 Pascals. Both $P$ and $T$ must be established in order to use this formula.

A sketch of the temperature in the atmosphere in °K as a function of height in the atmosphere in km is shown in Figure 7.16. There are three inversion in temperature, the lowest between the troposphere and the stratosphere, the next between the stratosphere and the mesosphere, and the highest between the mesosphere and the thermosphere.

The lowest level of the atmosphere is called the troposphere. The height in this layer varies from about 8 km at the poles to 16 km at the equator. It contains about 75% of the mass of the complete atmosphere, and is responsible for most of the weather events.

![Figure 7.16: Temperature as a Function of Height in the Atmosphere](image-url)}
The temperature drops with altitude in the troposphere, but this reverses in the next layer of the atmosphere, the stratosphere, which extends up to about 50 km. This layer is the ozone layer, and ozone absorbs solar radiation in the ultraviolet band. At the top of this band the temperature is about 0°C.

The temperature in the mesosphere, the region between about 50 km and 80 km, drops to about –100°C. The last significant layer of the atmosphere is the thermosphere, which extends from about 80 km, and here the temperature can reach above 1000°C due to absorption of the incoming radiation by oxygen molecules. However, the mass in this layer is minuscule compared to the levels below, so the attenuation on the 1366 W/m² of solar radiation entering the Earth’s atmosphere is considered negligible.

Plotting data from “The Engineering Toolbox for U.S. Standard Atmosphere Air Properties,” the temperature as a function of height in the atmosphere in kilometers is given by the somewhat jerky blue line in Figure 7.17. The fourth order polynomial

\[ T = 3.18470 \times 10^{-5}H^4 - 0.00665312H^3 + 0.441804H^2 - 10.262864H + 293.22600 \] (7.6)

is shown as the red response in Figure 7.17, and is a good fit to the data.
Next, the logarithm of the density $D$ of air as a function of height in kilometers in the atmosphere, from data in “The Engineering Toolbox for U.S. Standard Atmosphere Air Properties,” is shown as the blue response in Figure 7.18. Also shown is the linear approximation as the red line, a very good fit, which is given by

$$\log_{10} D = -0.061439 H + 1.1601746$$

(7.7)
where 2.30103 = \log_{e}x = \log_{10}x. We use this approximation for D to calculate the cumulative density of the atmosphere for every kilometer of altitude up to 80. The results are shown in Figure 7.19. Also shown in Figure 7.19 are the cumulative densities up to 5 Km and 10 Km, 62.71 and 86.49, respectively. The total up to 80 Km is 109.61, so the percentages of our atmosphere up to 5 and 10 Km are 57.2% and 78.9%, respectively.

The primary attenuation factor on terrestrial solar energy is cloud cover. Pollution in the atmosphere occurs more often near the equator and near major population densities, and is not a factor close to the poles. Typically, pollution is a minor factor in solar attenuation, and will be omitted from consideration here.

The names given to clouds can indicate their height in the atmosphere. The highest clouds around 10 km are called cirrus, and these are wispy and composed of ice crystals. Cirrus clouds refract light and can create a rainbow, and when they warm up they thicken and sink to become cirrostratus, or even lower altostratus, or stratus, or the lowest of all: nimbostratus clouds. Nimbostratus clouds are typically at a height of about 2 km, stratocumulus between 3 and 4 km. Cumulus clouds can span from ground level as mist to 10 km high, and are the primary rain makers. Stratus clouds are thin and may cause drizzle but never heavy rain.

A reasonable assumption is that clouds are limited to 10 km in height. Another assumption is that Rayleigh scattering is directly proportional to air density and distance travelled. Another reasonable assumption is that there is a lower limit of water aerosols that permits peak solar radiation reaching the Earth’s surface and so produces peak clear day scattering, and an upper limit of water aerosols that produce the lowest scattering because more photons are absorbed before they can be scattered.

The question that needs to be answered is: How much of the scattered radiation makes it through the cloud cover to reach the Earth’s surface?

As discussed previously, a photon is never attenuated. It can be scattered with its energy fully intact, or it can be absorbed such that it gives up all its energy to heat the entity that absorbed it, and that entity is far cooler than the source of that photon: the Sun, with a surface temperature of close to 6000°K.

The lightening of the sky in proximity to the Sun is due predominantly to scattering, Rayleigh and Mei. So we can consider that all scattering (Rayleigh, Mei, Tyndall, etc.) produces an anisotropic (unequal energy at different angles) component to diffuse radiation. The term $1 + \cos^2 \psi$ in the Rayleigh scattering pattern has maximum value of 2 at $\psi = 0$ and minimum value of 1 at $\psi = 90^\circ$. As a rough approximation, we assume that Rayleigh scattering is isotropic, so half this scattered energy is added to the diffuse radiation.
7.4 Models for Radiation Dispersion in the Atmosphere

In this section the following symbols are used:

- $I_0$ extraterrestrial solar radiation
- $I$ the solar energy reaching the Earth’s surface

What factors reduce the solar radiation reaching the Earth’s surface from that on a clear, dry day? The most dominant factor is cloud cover — water vapor in the atmosphere. We refer to opacity in the atmosphere as turbidity, where a perfectly clean, dry sky has a turbidity of 0, and a perfectly opaque sky has a turbidity of 1.

Aerosols include any small particle whose diameter $d$ in microns is delimited by $0.002 \leq d \leq 100$ and that tends to stay in the air. Aerosols include smoke, dust, soot, sea salt, and pollen. Water molecules and water vapor are not considered aerosols.

“A large portion of the microscopic particles floating in the air originate from incomplete combustion of coal and oil and from dust storms . . . . One way to gauge an aerosol’s ability to stay aloft is to determine its oxidation rate. Because oxidized aerosols absorb moisture and subsequently form clouds and fall as rain, the faster an aerosol oxidizes, the less time it spends in the atmosphere, and the less impact it has on the climate.” [“Composition and Reactions of Atmospheric Aerosol Particles,” Lawrence Berkeley National Laboratory, www.als.lbl.gov/index.php/contact/284]

Some models that describe the attenuation of solar energy in the atmosphere use something measurable as the starting point. The Lambert-Beer law uses distance traveled, a function of zenith angle. The Linke turbidity model uses air mass. Other models have an actual atmosphere in comparison to a dry, clean atmosphere.

The Lambert-Beer law is

$$I = I_0 e^{-kx} \quad (7.9)$$

where $x$ is the distance traveled through the atmosphere, and $k$ is called the extinction coefficient. Replacing distance travelled $x$ with air mass $M$ produces

$$I = I_0 e^{-kM} \quad (7.10)$$

In 1922, Linke [Linke, F. 1922. Transmissions–Koeffizient und Trübungsfaktor. Beitr. Phys. fr. Atmos., 10:91–103] proposed dividing $k$ into the product $k = \delta_c T_L$, where $\delta_c$ is the optical thickness of an ideal water and aerosol clear atmosphere. Linke determined $\delta_c$ for dry, clean, mountain air and proposed the formula
\[ \delta c = 0.128 - 0.054 \log M. \quad (7.11) \]

\( T_L \) then assesses the number of dry, clean atmospheres that would produce the required attenuation, a number always greater than 1: it is known as the Linke turbidity coefficient. Typical values for \( T_L \) are:

- very clean cold air 2
- moist warm or stagnating air 4–6
- clean warm air 3
- polluted air >6

Angstrom’s turbidity formula for the attenuation caused by all aerosols is

\[ k \alpha \lambda = \beta \lambda - \alpha. \quad (7.12) \]

where \( \beta \) defines sky conditions, which range from 0 (clean), 0.1 (clear), 0.2 (turbid), to 0.4 (very turbid). “A good average value for most natural atmospheres is \( \alpha = 1.3 \pm 0.5 \). The parameters \( \beta \) and \( \alpha \) can be determined simultaneously with a dual-wavelength photometer by measuring aerosol attenuation at two wavelengths where molecular absorption is absent or is minimal. The two wavelengths usually chosen are 0.38 and 0.5 \( \mu \).” [M. Iqbal, An Introduction to Solar Radiation, Academic Press, New York, 1983].

A criticism of these models is that they do not seem to be based on actual data. Here we will attempt, with very limited success, to build a data-based clear sky model. We could use the solar constant \( I_0 = 1366 \) W/m\(^2\) as the starting base, and attempt to build a model by determining the attenuation factors under clear sky conditions. Rayleigh scattering amounts to 180 W/m\(^2\) when the air mass \( M = 1 \), and 279 W/m\(^2\) when \( M = 2 \), as discussed in Chapter 6. To a first approximation we can assume that Rayleigh scattering is isotropic, as was discussed in Chapter 6, so about half of this scattered energy reaches the ground. With no other attenuating factor, this means that about \( 1366 - 180/2 = 1276 \) or \( 1366 - 279/2 = 1220 \) W/m\(^2\), depending on air mass, reaches the Earth’s surface. This obviously is not happening, since about 880 W/m\(^2\) reaches the Earth’s surface on a clear day.

Other researchers have extended, or modified, or proposed alternatives to the work of Linke and Angstrom. Some of these are listed in the work of Ineichen and Perez [P. Ineichen and R. Perez, “A New Airmass Independent Formulation for the Linke Turbidity Coefficient,” Solar Energy, vol 73 #3, pp. 151–157, 2002]. However, these models were proposed before actual numbers for solar energy at the Earth’s surface were available. NREL has the solar radiation data for direct, global, and diffuse radiation for every hour of every day and every year from 1961 through 2010. For the remainder of this chapter we will rely on this actual data to build simple models.
7.5 Solar Radiation Data for Boston

For the month of December 1990 at Boston, Massachusetts, the hourly averages for direct normal, diffuse, and global radiation are shown in Figure 7.22; the cone of aperture for the direct was 5.7°. The monthly average for the hour from noon to 1 pm was 282.52 W/m². The hourly beam maximum was 827 W/m², so for this month the average direct solar radiation was 34% of the maximum. On days of low beam radiation, the diffuse solar radiation is high. Here are some direct/diffuse numbers close to noon: 1/124, 1/186, 1/175, 1/187.

![Figure 7.20: Hourly Radiation for January 1990](image)

For the month of February 1990 at Boston, the hourly averages for beam, diffuse, and global radiation are shown in Figure 7.21. The monthly average for the hour from noon to 1 pm was 452.57 W/m². The hourly beam maximum was 940 W/m², so for this month the average beam radiation was 48% of the maximum.
For the month of January 1990 at Boston, the hourly averages for beam, diffuse, and global radiation are shown in Figure 7.20. The monthly average for the hour from noon to 1 pm was 384.19 W/m². The hourly beam maximum was 892 W/m², so for this month the average beam radiation was 43% of the maximum. Here are some direct/diffuse numbers close to noon when direct solar radiation is low: 6/197, 7/143, 2/239, 2/274, 4/154, 2/128, 7/208.
Beam radiation in Boston for January days in the years 1984–1990 on days of high solar radiation are shown in Figure 7.23.

Diffuse radiation at Boston on the same days as recorded in Figure 7.23 is shown in Figure 7.24. A general characteristic is apparent and is modeled in by three characteristics shown with heavy, dashed red lines given by $xM+15 \text{ W.m}^2$, where the lower response has $x = 35$, the middle response $x = 59$, and the upper response $x = 83$. 
For days of low insolation, the diffuse radiation is shown in Figure 7.25. The bottom model is given by $155M^{1.2}+15$, the middle model is $90M+10$, and the upper model is $280M+10$; these are shown by the heavy, dashed, red lines.

The sums of beam and diffuse at Boston on days of high solar radiation close to noon for the years of 1984–1990 are shown in Figure 7.26. The simple model

$$IM=1000M^{0.4}-70(7.13)$$
in W/m² is a good approximation, where $M$ is the air mass as defined in Section III of Chapter 6; this approximation is shown as the red, heavy dashed line in Figure 7.26.

Figure 7.27: Clear Sky Beam Plus Diffuse for Boston

Figure 7.28: Correction for Elliptic Orbit
For the years 1996–1990, the beam radiation received on days of high solar radiation in Boston is shown in Figure 7.27. Most of the annual variation of $91 \times 0.644 = 60.4 \text{ W/m}^2$ is due to the elliptical orbit of the Earth, so correcting for this produces the pattern shown in Figure 7.28. Notice the bias towards the first half of the year.

![Figure 7.29: Diffuse V Direct Radiation](image)

For the same years a chart of the diffuse to the beam radiation during the months of January and February is shown in Figure 7.29, a total of 42 days. The number of days of high solar radiation during December for the same years yielded only four days. The sample size is small and so may not be significant, but it does suggest that the second half of the year is less sunny than the first half.

### 7.6 Clear Sky Radiation Data for U.S. Cities


The items in italics below are quoted directly from the *National Solar Radiation Database 1991–2010 Update: User’s Manual*, by Stephen Wilcox:

The 1991–2010 National Solar Radiation Database was produced by the National Renewable Energy Laboratory under the U.S. Department of Energy’s Office of Energy Efficiency and Renewable Energy in collaboration with these partners:

- Atmospheric Sciences Research Center, State University of New York at Albany
- Climate Systems Branch, National Aeronautics and Space Administration
• Clean Power Research
• National Climatic Data Center, U.S. Department of Commerce
• Northeast Regional Climate Center, Cornell University
• Solar Consulting Services, Colebrook, New Hampshire
• Solar Radiation Monitoring Laboratory, University of Oregon.

The measured solar radiation data came from:

• Atmospheric Radiation Measurement Program, Department of Energy
• Florida Solar Energy Center, State of Florida
• Integrated Surface Irradiance Study and Surface Radiation Budget Measurement Networks, National Oceanic and Atmospheric Administration Air Resources Laboratory and Earth System Research Laboratory Global Monitoring Division
• Measurement and Instrumentation Data Center, National Renewable Energy Laboratory
• University of Oregon Solar Radiation Monitoring Laboratory Network
• University of Texas Solar Energy Laboratory.

The sites for the 1991–2010 NSRDB are subdivided into three classes of stations.

• Class I Stations have a complete period of record (all hours 1991–2010) for solar and key meteorological fields and have the highest-quality solar modeled data (242 sites).
• Class II Stations have a complete period of record but significant periods of interpolated, filled, or otherwise lower-quality input data for the solar models (618 sites).
• Class III Stations have some gaps in the period of record but have at least 3 years of data that might be useful for some applications (594 sites).

The data is for every hour of every day. There are in all 49 fields, and 49 columns in the spreadsheet, of which a limited number are used here:

1 date
2 time
16 measured global horizontal (called global here)
18 measured beam (called beam here)
20 measured diffuse horizontal (called diffuse here)

Another field that could be useful is field 26, the dry-bulb temperature, but when one looks at the numbers in this column it is evident that this column is incorrectly assigned.
The 1991-2010 update manual has this important disclaimer:

“Nearly all of the solar data in the original and updated versions of the NSRDB are modeled. The intent of the modeled data is to present hourly solar radiation values that, in the aggregate, possess statistical properties (e.g., means, standard deviations, and cumulative frequency distributions) that are as close as possible to the statistical properties of measured solar data over the period of a month or year. These data do not represent each specific hourly value of solar radiation to the same or equivalent accuracy as the long-term statistics. One must read sections 2.1.5, 2.2, and 2.3 to understand the content of the database and its applicability.”

This disclaimer is important, since NOAA (the National Oceanic and Atmospheric Administration) states, in the on-line communication https://data.noaa.gov/dataset/national-solar-radiation-database-nsrdb (accessed on March 12, 2015).

“The National Solar Radiation Database (NSRDB) was produced by the National Renewable Energy Laboratory under the U.S. Department of Energy’s Office of Energy Efficiency and Renewable Energy. The NSRDB update is a collection of hourly values of the three most common measurements of solar radiation (i.e., global horizontal, beam, and diffuse horizontal) over a period of time adequate to establish means and extremes and at a sufficient number or locations to represent regional solar radiation climates. Nearly all of the solar data in the NSRDB are modeled, and only 40 sites have measured solar data – none of them with a complete period of record. Because of the data-filling methods used to accomplish the goal of serial completeness, NSRDB meteorological data are not suitable for climatological work.”

NREL distinguishes between modeled and measured data. For example, fields 7, 10, and 13 are modeled numbers for global, beam, and diffuse. Fields 16, 18, and 20 are measured numbers for global, beam, and diffuse. To have two sets of numbers, modeled and measured, surely indicates that they are separate and distinct. Further, NREL uses the number –9900 to indicate missing data.

Even more disturbing, within a week of my first accessing the NOAA site https://data.noaa.gov/dataset/national-solar-radiation-database-nsrdb, it was taken down. Does this indicate dissent between NREL and NOAA, both governmental agencies?

The attempt here is to build a clear sky model for useable solar radiation using NREL solar data. The NREL data found most useful was for beam radiation B, and diffuse D radiation. It was found that global solar radiation \( G \) could be accurately modeled using the simple formula \( G = B \cos \phi + D \), where \( \phi \) is the zenith angle of the Sun determined by \( \cos \phi = \sin L \sin \delta + \cos L \cos \delta \cos 15r \); the correlation coefficient using this formula was commonly above 0.98.

Because of the elliptic orbit of the Earth, extraterrestrial radiation has been modeled as 1367 \( (1 + 0.0334 \cos (\mu L - 2.96)) \). One can assume that the variation in solar energy at the Earth’s surface also varies by \( \pm 3.34\% \), or for Boston, with average clear sky beam radiation of 880
W/m², this amounts to ±29 W/m², with a maximum near the beginning of the year and a minimum near the middle of the year.

We have seen that of the extraterrestrial radiation, 1366 W/m² but only about 880 reaches the Earth’s surface. What happened to the rest? If the Rayleigh scattering models are to be believed, then starting with 1366 and an air mass of 1, about 180 is scattered, and for air mass of 2 about 279 is scattered.

A select number of Class I cities in the United States were considered. For every day, the peak radiation for a particular day was chosen provided its value was in excess of 700 W/m². At this same hour the global and diffuse radiation was recorded.

The amount of variability required aggregation of the data, so it was decided that the data gathered would be placed in bins of 30, such that bin 1 covered January 1 through January 30. Bin 12 covered November 27 (or 28 for a leap year) through December 31. Even with this amount of aggregation, the randomness was large, so the annual bin numbers were averaged over the ten-year period 2001 through 2010.

The first city considered was Boston, Massachusetts, and its data is shown in Figure 7.30. What stands out is the lack of seasonal changes in the beam radiation, but a fairly strong variation in diffuse radiation occurs such that under clear sky conditions the diffuse radiation in winter is about half the value in summer. This is contrary to what one would expect. In particular, with Rayleigh scattering of 180 W/m² with an air mass of 1 it is to be expected that at least half of this, about 90 W/m², contributes to diffuse radiation, and this should occur in summer. Instead, it seems to occur in winter when the air mass is over 2, so diffuse should be almost double that in summer, again under clear sky conditions.
The next city considered was Denver, Colorado, the so called “mile high city.” Perhaps there is a small seasonal variation in the beam radiation, and a definitive increase in diffuse radiation in the summer months. The shape of the global radiation is choppy, in contrast to the smoother shape for Boston’s global radiation.

Orlando, Florida, shows a significant reduction in beam radiation and a significant increase in diffuse radiation in summer. It will be seen that for all the cities considered, the diffuse radiation is significantly higher in summer than in winter, counter to what is expected.
Figure 7.33: Kennedy

Figure 7.34: Lincoln
Figure 7.35: Seattle

Figure 7.36: Salt Lake City
Figure 7.37: Detroit

Figure 7.38: Billings
Figure 7.39: Bangor

Figure 7.40: Baltimore
Figure 7.41: Chicago

Figure 7.42: Concord, NH
Figure 7.43: Lexington

Figure 7.44: Albuquerque
Figure 7.45: Raleigh, NC

Figure 7.46: Colombus
Figure 7.47: Reno

Figure 7.48: Hartford
Figure 7.49: Tulsa

Figure 7.50: Phoenix
Figure 7.51: Sacramento

Figure 7.52: Austin
Notice that the average diffuse radiation for the cities studied for days with high beam radiation in winter is about 90 W/m², the same as predicted for Rayleigh scattered energy reaching the earth’s surface when $M = 1$. However, diffuse radiation is almost double that in summer — not what one would predict.

There appears to be a minor seasonal variation in beam energy in most of the cities considered.
All data contained in the cities just studied were modelled, measured data for Denver is shown in Figure 7.54 for the years 1990 to 1998 and the year 2000; the year 1999 had no measured data. There are no distinctive differences between modelled data and measured data for Denver.

Also shown in Figure 7.54 are quadratic polynomial approximations to the three responses. These are:

\[
B = 0.492x^2 - 7.28x + 941.2 \\
G = -14.626x^2 + 182.47x + 325.2 \\
D = -1.397x^2 + 16.13x + 61.44
\]

7.7 Solar Radiation Maps of the World

The solar radiation numbers collected should be appropriate to the way the solar energy is used. Passive solar collection for winter heating is almost always, in the northern hemisphere, through vertical, south-facing glass. Unfortunately, data collected on such a surface is rarely available. For year round collection with photo-voltaics, south-facing arrays, latitude inclined, is considered close to ideal, and this data is available.

![Solar Radiation Map](https://example.com/solar_radiation_map.png)

Figure 7.55
The most important source of solar data in the United States is NREL (National Renewable Energy Laboratory). It has performed sterling work over the years. Its budget should be increased, but in some instances has been cut before its impact on a particular technology can be realized. Case in point, the research work it has performed on the biofuel source algae, far more important to our future than other biofuels such as corn-based ethanol or switch grass.

Figure 7.55 gives the average daily energy falling on a latitude-inclined surface over a complete year in KWH/m², and is based on NREL-created maps. As expected, the lowest numbers are on the Canadian border, and the highest are in the Southwest.
With apologies to our Canadian colleagues, there are no NREL maps for the United States and Canada combined, and the maps of Canada alone are not fully satisfying.

For South America, the average daily energy falling on a latitude-inclined collector over a complete year in KWH/m² is shown in Figure 7.56; this is based on NREL data, with “model estimates of monthly average daily total radiation using inputs derived from satellite and surface observations and cloud cover, aerosol optical depth, precipitable water vapor, albedo, atmospheric pressure and ozone sampled at a 40 Km resolution.”


The solar radiation data for southeast Asia shown in Figure 7.58 is derived from SolarGIS @ 2013 Geomodel Solar s.r.o.

As for the prior figure, the solar radiation data for Australia as shown in Figure 7.59 is derived from SolarGIS @ 2013 Geomodel Solar s.r.o.
Figure 7.59: Australia

Figure 7.60: Africa and the Middle East
Chapter 8

Transmission of Solar Energy through Glazing

The most important diathermanous material is glass, that transparent substance made from the unlikely mixture of ground-up sand, baking soda, and ground limestone — common materials with little worth. It has interesting properties. For example, it is not strictly a solid: it is a super cooled liquid and flows under the force of gravity. This flow rate is small but measurable. Old window panes can be seen to be thicker at the bottom than the top.

Glass was an expensive luxury before the industrial revolution. Now it is a commodity, a product with lots of manufacturers, often competing with each other on low markup products.

The alternate diathermanous materials to glass are clear plastics, particularly acrylics and polycarbonates. These are evaluated for transmissivity in this chapter, and shown to be, in some cases, superior to glass. However, they have other properties that are not satisfactory. For example, polycarbonates age with time, reducing their transmissivity to solar energy. Acrylics have a low forming temperature, possibly making them unsatisfactory over flat plate solar collectors.

The transmissivity, absorptivity, and reflectivity of glass are analyzed using the ray tracing technique. Reflectivity at an air-to-glass interface, or a glass-to-air interface, is determined from Fresnel’s equations for $\rho_\perp$ and $\rho$, which are the reflectivity for polarized light in two orthogonal planes; more about this later.

8.1 A Brief History of Glass

“Before people learned to make glass, they had found two forms of natural glass. When lightning strikes sand, the heat sometimes fuses the sand into long, slender glass tubes called fulgurites, which are commonly called petrified lightning. The terrific heat of a volcanic eruption also sometimes fuses rocks and sand into a glass called obsidian. In early times, people shaped obsidian into knives, arrowheads, jewelry, and money. We do not know exactly
when, where, or how people first learned to make glass. It is generally believed that the first manufactured glass was in the form of a glaze on ceramic vessels, about 3000 B.C. The first glass vessels were produced about 1500 B.C. in Egypt and Mesopotamia. The glass industry was extremely successful for the next 300 years, and then declined. It was revived in Mesopotamia in the 700’s B.C. and in Egypt in the 500’s B.C. For the next 500 years, Egypt, Syria, and the other countries along the eastern shore of the Mediterranean Sea were glassmaking centers” [Steve W. Martin, Ph.D., Professor of Materials Science and Engineering, Iowa State University, on-line at www.texasglass.com/glass_facts/history_of_Glass, as accessed on December 24, 2012]

Early windows were small openings with no glass. Often wooden shutters were used to cut the heat loss on cold nights. Interiors were dark and illuminated by the fire over which food was cooked. Sometimes the openings were covered with oiled animal hides, or parchment, either of which was not particularly transparent, but they did permit some light to enter while reducing the excessive heat loss through an opening.

Glass was a great improvement, but early glass-melting furnaces could barely melt glass. Syrian craftsmen invented the blow pipe in the first century BC and overcame this problem. In this method the glass is softened with heat and a ball of it attached to the end of a blow pipe through which the craftsman blows air to inflate a cylindrical shaped balloon, all the while rotating the expanding balloon while partially supporting it on some flat surface, probably wood. The glass was permitted to harden on that surface, and the cylinder sliced lengthwise, creating curved sheets of glass. This produced a superior window to glass pebbles or paper or animal hides, which were at best translucent.

The Romans adopted the blowpipe method, and its use spread, particularly to the north and west. By about 1000 AD Alexandria, Egypt, became the most important center of glass manufacture. Christian Europe built an astonishing number of churches and cathedrals across the continent, and stained glass windows were used extensively; these used small panes set in lead frames, and portrayed biblical scenes and pious adventures.

By the eighth century Venice had become the glassmaking center of the western world. They borrowed from the technology developed by the Byzantine empire. By the thirteenth century its Glassmakers Guild set out rules and regulations in order to protect its trade secrets. As always, guilds can only protect proprietary information for a limited time, and glassmaking spread across the western world.

By the late sixteenth century, English glassmakers were making glass in Venetian style. In 1674, an English glassmaker, George Ravenscroft, received a patent for lead glass. At that time only the very wealthy could afford glazed windows, usually with small leaded panes.

In the early nineteenth century the demand for window glass was satisfied with crown glass. “The technique of crown glass remained standard from the earliest times: a bubble of glass, blown into a pear shape and flattened, was transferred to the glass-maker’s pontil (a solid iron
rod), reheated and rotated at speed, until centrifugal force formed a large circular plate of up to 60 inches in diameter. The finished ‘table’ of glass was thin, lustrous, highly polished (by ‘fire-polish’), and had concentric ripple lines, the result of spinning; crown glass was slightly convex, and in the centre of the crown was the bull’s eye, a thickened part where the pontil was attached. This was often cut out as a defect, but later it came to be prized as evidence of antiquity. Nevertheless, and despite the availability of cheaper cylinder glass (cast and rolled glass had been invented in the 17th century), crown glass was particularly popular for its superior quality and clarity. From about the mid-17th century the crown glass process was gradually replaced by easier methods of manufacturing larger glass sheets.

In 1688 a process for making plate glass by pouring molten glass over a special table and rolling it out was invented in France; an extensive amount of grinding and polishing was needed. About the end of the First World War, Belgian engineer Emil Bicheroux developed the technique of pouring molten glass through rollers to produce glass sheets of uniform thickness, thereby reducing the amount of grinding needed, and becoming the standard production method.

In the seventeenth century the wood framed casement window became common. It became popular when the sash was counterbalanced with corded weights and pulleys so it could remain in an intermediate position without fastening.

Two great glass buildings were the Palace of Versailles, 10 miles southwest of Paris, built in the 1680s, and the Crystal Palace in London, built for the Great Exhibition of 1851. The original interest of the royalty of France in Versailles was for hunting. Louis XIII built a small hunting lodge there, but it was his successor, Louis XIV (1638–1715), who built the Palace of Versailles and made it the seat of government. He was known as the Sun King, and his embrace of the light makes the Palace of Versailles, whose centerpiece is the Hall of Mirrors, one of the most attractive and imposing places to visit in the world. It is full of great works of art, paintings and sculptures, and innumerable fountains. It is gratifying that a structure of such opulence as the Palace of Versailles, whose original purpose was for the pleasure of a few, has been preserved and now delights multitudes of visitors daily.

Contrast this to the history of another structure of light and glass — the Crystal Palace in London. Prince Albert, consort to Queen Victoria, conceived of a enormous building of glass and iron. It was 1,848’ long and 408’ wide with water fountains containing 11,000 jets. The whole idea was to impress visitors to the Great Exhibition of 1851. After the exhibition it was moved to Sydenham in south east London to a place now called Crystal Palace Park. However, a structure that large was difficult to maintain, and it was declared bankrupt in 1911. A massive fire in 1936 destroyed most of the structure. Now there are plans to rebuild it on the same size and scale as the original.

The refinement of the float method of glass production by Sir Alastair Pilkington in the 1950s is the most successful advance in glassmaking since Bicheroux; 90% of flat glass is
manufactured by this method today. “The contribution to business success in the shape of technology comes in many forms. For Pilkington it came with the help of a washing up bowl. Alastair Pilkington, then technical director of the ‘family’ business, conceived the idea of float glass when he was doing the washing up. He was fascinated by the sight of a plate floating on water and wondered whether the principle could be applied to glass making. Seven years later in 1959 it was. Instead of a floating plate he produced a ribbon of glass by floating the raw materials at higher temperature over a bath of molten tin. The end product was more economical, high class glass for shop windows, cars and mirrors without the distortions. Development costs of £7m (£80m in today’s money) almost broke Pilkington” [Roland Gribben, February 27, 2006].

8.2 Transmissivity of Diathermanous Materials to Light

Glass is not an ideal transmitter of the Sun’s energy. Some energy bounces off the incident surface. Some is absorbed inside the glass due to impurities. Some bounces off the inside surface. The angle of incidence of the incoming radiation is an important determinant for reflectivity.

![Figure 8.1: Refraction of Light](image)

When sunlight is transmitted through glass or acrylic, it is refracted by the surface. Consider a beam of light in air (refractive index \( \eta_1 = 1 \)) with incidence angle of
Figure 8.2: Angles of Refraction

Table 8.1: Incident and Reflected Angles

<table>
<thead>
<tr>
<th>θ1</th>
<th>θ2</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>6.6</td>
</tr>
<tr>
<td>20</td>
<td>13.2</td>
</tr>
<tr>
<td>30</td>
<td>19.5</td>
</tr>
<tr>
<td>40</td>
<td>25.4</td>
</tr>
<tr>
<td>50</td>
<td>30.7</td>
</tr>
<tr>
<td>60</td>
<td>35.3</td>
</tr>
<tr>
<td>70</td>
<td>38.8</td>
</tr>
<tr>
<td>80</td>
<td>41.1</td>
</tr>
</tbody>
</table>

θ₁ striking a surface (such as glass or acrylic) with approximate refractive index η₂ = 1.5, as shown in Figure 8.1. The angle θ₂ of the beam inside the surface is given by Snell’s Law [E. Hecht, Optics, Fourth Edition, Addison Wesley, Boston, 2001]

\[ \eta_1 \sin \theta_1 = \eta_2 \sin \theta_2 \tag{8.1} \]

The relationship between these angles is given in Table 8.1, and the data presented graphically in Figure 8.2.

The refractive index of glass is not 1.5, but is often considered closer to 1.52. Yet even this figure is misleading since the refractive index is a function of the wavelength considered. The refractive index of soda lime glass has been characterized [Matt Considine, http://astro.umsystem.edu/atm/ARCHIVES/MAY00/msg00671.html] as 1.53 at the shortwave end of the visible spectrum and 1.515 at the long wave end of the visible spectrum. Since the center of solar energy is closer to the long wave end of the visible spectrum, we could consider \( \eta_2 = 1.515 \) to be a good overall figure to use for glass. This is verified by the data for BK7 optical glass [//refractiveindex.info/?group=GLASSES] shown graphically in Figure 8.3, where the abscissa is the wavelength in microns. Also shown in Figure 8.3 is a line for refractive index 1.515, the chosen average number for this chapter.
The refractive indices of some common transparent materials are given in Table 8.2. Lucite and Plexiglas are proprietary names for acrylic.

The refractive index of water is temperature dependent. At 0°C it is 1.33346, at 20°C 1.33283, and at 100°C 1.31766.

The wavelength dependence of transmissivity of glass and 1/8” acrylic is shown in Figure 8.4; the abscissa axis is in microns. The glass is 1/4” thick, considerably thicker than most glass sheets, so the transmissivity is on the low side. Notice that the transmissivity window is essentially in the waveband $0.3 \leq \lambda \leq 2.7 \mu$.

As discussed in Chapter 6, approximately 4% of the Sun’s energy is in the waveband $0 \leq \lambda \leq 0.3 \mu$, and about 2.5% is in wavelengths greater than $\lambda = 2.7 \mu$, so approximately 6.5% of the Sun’s energy is automatically rejected by glass.
Also shown in **Figure 8.4** is the transmissivity of 1/8” thick acrylic. Its characteristic has deep nulls and its wavelength cutoff point is about 2.2 μ, but it passes more of the Sun’s energy than glass. To prove this point the characteristics of glass and acrylic were sampled at intervals of 0.05μ, Planck’s law was normalized such that its integral (Stefan-Boltzman law) was unity, and at every sample point the transmission value was multiplied by the Planck value. The result was that glass passes 77% of the Planck law energy, and acrylic 84%.

![Figure 8.4: Wavelength Dependency on Transmissivity](image)

Acrylic, poly(methyl methacrylate), is not a single material; it varies by manufacturer, and some manufacturers have more than one formulation.

Acrylic has a low temperature softening point, and melts at a little above the boiling point of water. It ignites at 460°C. Like glass, its transparency remains high, but it is relatively soft and can easily be scratched, which reduces its optical clarity while having little effect on its solar transmissivity.

Polycarbonates, with trade names such as Lexan and Macrolon, have a high impact resistance but low scratch resistance, which has little effect on solar transmissivity. However, they craze with age, and this can be a major deterrent to their use in solar applications. They melt at about 155°C ≡ 311°F.

Plastic sheets in general have characteristics similar to acrylic, and cause problems for plant growths in plastic-sheeted greenhouses, since plants need energy in certain wavelengths that are filtered out by the plastic; one common phenomenon is for the plant to put a lot of its energy into height in a fruitless attempt to get closer to the light.

Window glass and glass used for other glazing purposes is predominantly silica (silicon dioxide $SiO_2$). A common glass, soda-lime glass, has a composition of 71–75% silica, 12–
16% soda $\text{Na}_2\text{O}$, and 10–15% lime $\text{Ca}$, plus small amounts of other compounds. Of these, the most significant compound is $\text{Fe}_2\text{O}_3$ in an amount about 0.04%, which causes a measurable absorption of solar radiation energy in the glass.

Table 8.3: Absorptivity of Normal and Low Iron Glasses

<table>
<thead>
<tr>
<th>thickness</th>
<th>absorptivity normal iron</th>
<th>absorptivity low iron</th>
</tr>
</thead>
<tbody>
<tr>
<td>3mm or 1/8”</td>
<td>5%</td>
<td>3%</td>
</tr>
<tr>
<td>6mm or 1/4”</td>
<td>12%</td>
<td>6%</td>
</tr>
<tr>
<td>12mm or 1/2”</td>
<td>17%</td>
<td>12%</td>
</tr>
</tbody>
</table>

The iron bond is a major factor to be considered: $\text{Fe}^{2+}$ ions absorb solar energy more strongly than $\text{Fe}^{3+}$ ions. Soda-lime glass includes float glass and low-iron glass.

Iron in solar glass causes a green tint, observed by looking along the exposed edge of a sheet of glass. Low-iron glass was introduced in the 1970s and became widely available by 1990 when it was manufactured by PPG Industries. The absorptivity $a$ of normal iron and low iron, as a function of glass thickness, is given in Table 8.3; these figures are approximate and vary by manufacturer.

For the rest of this chapter we will use absorptivity of 5%, $a_0 = 0.05$.

### 8.3 The Polarization of Light

Light can be considered an oscillating wave propagating along a straight line while in a homogeneous material of refractive index $\eta_1$, until it strikes a different material with refractive index $\eta_2$. The two materials could be air, with a refractive index $\eta_1 = 1$, and glass, with a refractive index $\eta_2 = 1.515$. The photon theory of light assumes that the oscillation of an individual photon while in air is coherent: That is, with respect to the line of travel, the sinusoidal oscillations about that line are in a single plane. If all the photons exhibited oscillations that occur in single plane, then the light is said to be fully polarized.

Sunlight in space is not coherent, so the assembly of planes of oscillation can be considered to be a circle about the direction of travel. When the light enters the Earth’s atmosphere there is some selective scattering and filtering, as discussed in the previous chapter. If the Sun is low in the sky, some photons whose oscillations are perpendicular to the ground are filtered out, and the light is said to be partially polarized.
Consider a beam of light shown as vector $\mathbf{v}_i$ in Figure 8.5 that strikes an air/glass interface. The normal to that interface is shown as vector $\mathbf{v}_n$. The plane containing vectors $\mathbf{v}_i$ and $\mathbf{v}_n$ is $\mathbf{p} = \mathbf{v}_i \times \mathbf{v}_n$. Looking down vector $\mathbf{v}_i$, the envelope of the unpolarized photon oscillations is a circle as shown in the right-hand diagram — any other shape would indicate polarization.

The heavy lines in the right-hand diagram of Figure 8.5 are the directions of parallel $\parallel$ and perpendicular $\perp$ polarization. These heavy lines indicate the plane of oscillation of the light. A Polaroid filter is like a picket fence to these oscillations, and can block the wave or allow the wave through, depending on orientation. This is shown in Figure 8.6, where in the left-hand diagram the picket fence allows the $\parallel$ wave to pass, but blocks the $\perp$ wave in the right-hand diagram.

![Figure 8.5: Polarization of Light](image)

![Figure 8.6: Fence Model for Polarization](image)

Reflection can polarize light. Sunlight at a low angle on snow or water can produce glare to an observer, and this indicates that the light is polarized in a direction parallel to the surface. A polished metallic surface does not cause polarization. Similarly, when light is refracted by passing through an interface into a transparent medium with a different refractive index, some polarization occurs, usually in the plane perpendicular to the interface.

**8.4 Fresnel’s Equations**
The relation between the reflection $\rho$ at the interface between two mediums with refractive indices $\eta_1$ and $\eta_2$, as a function of the angles of incidence $\theta_1$ and refraction $\theta_2$, as given by Snell’s Law, is given by Fresnel’s equations:

$$\rho = \tan^2(\theta_2 - \theta_1) \tan^2(\theta_2 + \theta_1), \rho_\perp = \sin(\theta_2 - \theta_1) \sin(\theta_2 + \theta_1)$$ (8.2)

Notice that the values of $\theta_1$ and $\theta_2$ can be swapped in Equations 8.2 without affecting the reflectivity. Thus, air-to-glass reflectivity is the same as glass-to-air reflectivity.

Sunlight is weakly polarized at best, and can usually be considered unpolarized, so how can we use Fresnel’s equations? Although it is not exactly satisfying, it is common to take the average as

$$\rho = \rho + \rho_\perp = 12 \sin(\theta_2 - \theta_1) \sin(\theta_2 + \theta_1) \{1 + \cos(\theta_2 + \theta_1) \cos(\theta_2 - \theta_1)\}.$$ (8.3)

$\theta_2$, $\rho$, $\rho_\perp$ and $\rho$ are shown in Figure 8.7 as a function of the angle of incidence.

For radiation close to normal, so $\theta_1$ is small, the cosine terms in Equation 13.26 become unity, and so

$$\rho_{\text{normal}} = \sin(\theta_2 - \theta_1) \sin(\theta_2 + \theta_1)$$ (8.4)

Eliminating the compound angle in Equation 8.4 with a well-known trigonometric identity gives
\[ \rho_{\text{normal}} = (\sin \theta_2 \cos \theta_1 - \cos \theta_2 \sin \theta_1 \sin \theta_2 \cos \theta_1 + \cos \theta_2 \sin \theta_1) \]
\[ = (\sin \theta_2 - \sin \theta_1 \sin \theta_2 + \sin \theta_1)^2. \quad (8.5) \]

Next, applying Snell’s law produces
\[ \rho_{\text{normal}} = (\eta_1 \eta_2 \sin \theta_1 - \sin \theta_1 \eta_1 \eta_2 \sin \theta_1 + \sin \theta_1) = (\eta_1 - \eta_2 \eta_1 + \eta_2) \]
and the transmitted energy is
\[ T_{\text{normal}} = 1 - \rho_{\text{normal}} = 1 - (\eta_1 - \eta_2 \eta_1 + \eta_2)^2 = 4 \eta_1 \eta_2 (\eta_1 + \eta_2)^2. \quad (8.6) \]

If \( \eta_1 = 1 \) and \( \eta_2 = 1.515 \), so the two mediums are air and glass, then \( \rho_{\text{normal}} = (0.515 \cdot 2.515)^2 = 0.0419 \) and \( T_{\text{normal}} = 1 - 0.0419 = 0.9581 \).

As can be seen from Figure 8.7, the reflected energy \( \rho \) can go to zero. The incidence angle that produces this is called Brewter’s angle \( \theta_b \), and occurs when \( \tan(\theta_2 + \theta_1) \to \infty \), or \( \theta_2 + \theta_1 = 90 \). Applying Snell’s law \( \eta_1 \sin \theta_1 = \eta_2 \sin \theta_2 = \eta_2 \sin (90 - \theta_1) = \eta_2 \cos \theta_1 \), so
\[ \theta_b = \tan^{-1} \eta_2 \eta_1. \quad (8.7) \]

With \( n_1 = 1 \) and \( n_2 = 1.515 \), \( \theta_b = 56.5726^\circ \).

Also shown in Figure 8.7 is the terminal value for \( \theta_2 \) when \( \theta_1 \) approaches \( 90^\circ \). This value for the given refractive indices is \( 41.3049^\circ \), and is known as the critical angle. Light energy inside the glass at the critical angle will remain trapped inside the glass to bounce between the two surfaces and be lost to absorptivity of the glass.

Shown in Figure 8.8 are the critical angles for materials of interest in solar energy; diamond was included since its refractive index of 2.418 is very high. Flint glass is leaded glass, and the span shown in Figure 8.8 for flint glass runs from 29% to 71% lead. It is interesting to note that refractive index \( \eta \) indicates the reduction \( \nu \) of the speed of light in the material given by
\[ \nu = c \eta. \quad (8.8) \]

where \( c \) is the speed of light in a vacuum.

### 8.5 Extinction Coefficient and Absorptivity of Glass

The extinction coefficient of a medium is the measure of the rate of diminution of transmitted light by scattering and absorption in that medium [V.E. Derr, “Estimation of the Extinction Coefficient of Clouds from Multiwavelength Lidar Backscatter Measurements,” Applied...
Assuming that the intensity of light is proportional to the distance $x$ traveled in the medium, and proportional to the intensity $I$ of light, then

$$dl = -kI \, dx \quad (8.10)$$

where $k$ is known as the extinction coefficient. Integrating over distance $\ell$ produces

$$I\ell = \int_0^\ell -kI \, dx = I_0 e^{-k\ell} \quad (8.11)$$

where $I_0$ is the intensity of light entering the medium. Equation 8.11 is known as Bouguer’s law [J.A. Duffie and W.A. Beckman, Solar Engineering of Thermal Processes, John Wiley, Hoboken, New Jersey, Fourth Edition, 2013], or the Lambert-Beer law, or simply Beer’s law. For glass of thickness $L$ and light at angle $\theta_2$ to the normal, this law becomes

$$IL = I_0 e^{-kL/cos \theta_2} \quad (8.12)$$

Further, only absorption occurs in glass, at least glass suitable for solar collection purposes, and scattering can be ignored, so the energy absorbed is $\alpha = I_0 - IL = I_0(1 - e^{-kL/cos \theta_2})$. Normalizing, so $I_0 = 1$,

$$\alpha(k, \theta_2) = 1 - e^{-kL/cos \theta_2} \quad (8.13)$$

Physicists discuss absorptivity through a partially transparent medium in terms of the extinction coefficient and complex (as in complex numbers) refractive indices. The mediums they consider include window glass. However, glass manufacturers consider simpler models, and typically will quote a single number for absorptivity.
Assuming the refractive index of air is \( \eta_1 = 1 \), for glass it is \( \eta_2 = 1.515 \), and that the angle of incidence on the glass is \( \theta_1 \), then the angle \( \theta_2 \) in the glass as given by Snell’s law [E. Hecht, *Optics*, Fourth Edition, Addison Wesley, Boston, Mass, 2001] is

\[
\theta_2 = \sin^{-1} \left\{ \eta_1 \eta_2 \sin \theta_1 \right\}. \quad (8.14)
\]

The reflectivity at the air/glass interface as given by Fresnel’s equation is

\[
\rho = \frac{12 \sin^2(\theta_2 - \theta_1) \sin^2(\theta_2 + \theta_1) \{1 + \cos^2(\theta_2 + \theta_1) \cos^2(\theta_2 - \theta_1)\}}{1 + \cos^2(\theta_2 - \theta_1) \cos^2(\theta_2 + \theta_1)}, \quad (8.15)
\]


\[
\rho_{\text{normal}} = \lim_{\theta_1 \to 0} \sin^2(\theta_2 - \theta_1) \sin^2(\theta_2 + \theta_1) = 0.0419. \quad (8.16)
\]

The extinction coefficient \( k \) for window glass is in the range from 4 meters\(^{-1}\) to 32 meters\(^{-1}\). For flat glass of thickness 3 mm and the three incident angles 0°, 45° and 90°, the absorbed energy defined by Equation 8.13 is plotted as shown in Figure 8.9: at the incidence value of \( \theta_1 = 90^\circ \), \( \theta_2 = 41.3049^\circ \), the critical angle. However, no light can enter the glass when the incidence angle is 90°, and for angles close to 90° the reflectance is almost 100%, as will be considered after discussion of the simple absorption model.
Also shown in Figure 8.9 with the symbol “o” are the values of $0.0382\cos\theta_2$, for the same three angles $0^\circ$, $45^\circ$, and $90^\circ$, plotted at $k = 13$, and $0.0915\cos\theta_2$ for the same three angles plotted at $k = 32$.

$$\alpha(13, 45) = 0.0431, \alpha(13, 90) = 0.0509, \alpha(32, 45) = 0.1029, \text{ and } \alpha(32, 90) = 0.1200,$$

Compare these numbers to

$$0.0382\cos 27.823 = 0.0432, 0.0382\cos 41.305 = 0.0509, 0.0915\cos 27.823 = 0.1035 \text{ and } 0.0915\cos 90^\circ = 0.1200.$$

The deviations of the simple model from the extinction curve of the points for $\theta_1 = 45^\circ$ and $\theta_1 = 90^\circ$ are minimal. Further, the responses of $\alpha(k, 45)$ and $\alpha(k, 90)$ are virtually straight lines. Therefore, for clear window glass with realistic absorptivity, it is safe to assume that absorptivity can be modeled by the simple expression $a = a_0\cos\theta_2$, where $a_0$ is the absorptivity-to-normal incident radiation.

If we had applied Taylor series to equation 8.13 we would have

$$1-e^{-kL/\cos\theta_2} = kL\cos\theta_2 - 12!(kL\cos\theta_2)^2 + 13!(kL\cos\theta_2)^3 - \ldots - kL\cos\theta_2$$

but $a_0 \neq kL$. With $kL = 13 \times 0.003 = 0.039$ we have $a_0 = 0.0382$, and with $kL = 32 \times 0.003 = 0.096$ we have $a_0 = 0.0915$.

The small deviation of the simple model that occurs for high incidence angles and high absorptivity coefficients is even less significant when one looks at the total reflectivity

$$R = \rho + \rho(1-\rho)21(1-a)21 - \rho2$$

and the total transmissivity

$$T = (1-\rho)21(1-a)1 - \rho2$$

With glass having $a_0 = 0.05$ and $\theta_1 = 89^\circ$, $R = 0.9296$ and $T = 0.0296$. Thus, the simple model is fully satisfactory when considering the absorptivity of glass covers for solar collectors. However, when considering attenuation of solar radiation through the atmosphere, the simple model is unsatisfactory, and the extinction coefficient model reigns [V. E. Derr, “Estimation of the Extinction Coefficient of Clouds from Multiwavelength Lidar Backscatter Measurements,” Applied Optics, vol. 19 #14, pp. 2310–2314, 1980].

8.6 Transmission, Reflection, and Absorption
Consider the transmission of light through a single pane of glass as shown in Figure 8.10. The angle of incidence to the normal is $\theta_1$, and the reflectivity on the outside surface is $\rho$, where $\rho$ is a function of the refractive index of air, the refractive index of glass, and $\theta_1$; $\rho$ increases with $\theta_1$ and with an increase in the differential of the two refractive indices.

For the situation shown in Figure 8.10 with incidence solar energy normalized to unity, $1 - \eta_1$ passes through the outside glass surface. The energy reaching the inner surface of the glass is $(1 - \rho)(1 - a)$, of which $(1 - \rho)(1 - a)(1 - \rho)$ passes through the second surface and $(1 - \rho)(1 - a)^2\rho$ is reflected back inside the glass.

The total energy passing through the glass is

$$T = (1 - \rho)(1 - a)(1 - \rho) + (1 - \rho)(1 - a)\rho(1 - a)(1 - \rho) + (1 - \rho)(1 - a)\rho(1 - a)\rho(1 - a)\rho(1 - a)(1 - \rho)(1 - a)^2\rho + \ldots$$

The total energy reflected out and lost is

$$R = \rho + (1 - \rho)(1 - a)(1 - \rho) + (1 - \rho)(1 - a)\rho(1 - a)(1 - \rho) + (1 - \rho)(1 - a)\rho(1 - a)\rho(1 - a)(1 - \rho) + \ldots = \rho + (1 - \rho)(1 - a)(1 - \rho)(1 - a)^2\rho + \ldots$$
The total absorptivity that heats the glass is

\[ A = (1 - \rho)a + (1 - \rho)(1 - a)\rho + (1 - \rho)(1 - a)\rho(1 - a) + (1 - \rho)(1 - a)(1 - a)\rho + ... = (1 - \rho)a(1 + \rho(1 - a) + \rho^2(1 - a)^2 + \rho^3(1 - a)^3 + ...) \]

From the binomial theorem \(1 + x + x^2 + x^3 + ... = 11 - x\), we simplify the expressions above into

\[ T = (1 - \rho)2(1 - a)1 - \rho2(1 - a)2R = \rho + (1 - \rho)2(1 - a)21 - \rho2(1 - a)2A = (1 - \rho)a1 - \rho(1 - a)(8.20) \]

\[ T + R = \rho + (1 - \rho)2(1 - a)(1 + \rho - \rho)a1 - \rho2(1 - a)2 + (1 - \rho)2(1 - a)(1 + \rho - \rho)(1 - \rho(1 - a))1 + \rho(1 - a) = \rho + (1 - \rho)2(1 - a)1 - \rho(1 - a) \]

\[ T + R + A = \rho + (1 - \rho)2(1 - a)1 - \rho(1 - a) + (1 - \rho)a1 - \rho(1 - a) = \rho + (1 - \rho)((1 - \rho)(1 - a) + a1 - \rho(1 - a) = \rho + (1 - \rho)(1 - \rho + \rho)a1 - \rho(1 - a) = 1 \]

as expected. The energy system for this single sheet of glass is as shown in Figure 8.11.

![Figure 8.11: Single PAne Energy System](image)

Assuming \(\theta_1 = \theta_2 = 0\) and \(a = 0\), then \(T = (1 - \rho)21 - \rho2 = 1 - \rho1 + \rho = 0.95811.0419 = 0.9196\).

The transmissivity \(T\) is a somewhat misleading figure to assess solar gain. A more meaningful measure is effective transmissivity \(T_{\text{eff}} = T \cos \theta\); for example, the solar radiation on the glass when \(\theta = 60^\circ\) is half what it would be if \(\theta = 0\) since \(\cos 60 = 0.5\). Similarly, the effective absorptivity of the glass is \(a = a_0 / \cos \theta\) as discussed previously.

The effective solar radiation and the effective absorptivity in the glass as a function of the angle of incidence are shown in Figure 8.12, assuming the absorptivity of glass to normal incident radiation is 5\%. The actual numbers for \(T_{\text{eff}}, A\), and \(R\) are given in Table 8.4.

Consider next the energy situation with two parallel sheets of glass, as shown in Figure 8.13, where \(R_1 + A_1 + T_1 = 1\) and \(R_2 + A_2 + T_2 = 1\). \(R_1\) is part of the reflected energy that is lost to the environment. \(T_1T_2\) contains the energy passing through the second sheet of glass as a result of \(T_1\) passing through the first sheet of glass, but this is not all the energy transmitted. The reflected energy \(T_1R_2\) bounces
at each interface and will add to energy lost to the environment, absorbed in each sheet of glass, and transmitted to the interior. The energy situation of Figure 8.13 is summarized in the equations

\[ T_d = T_1T_2 + T_1R_2R_1T_2 + T_1R_2R_1R_2R_1T_2 + \ldots \]

\[ = T_1T_2 - R_1R_2.R_d = R_1 + T_1R_2T_1 + T_1R_2R_1R_2T_1 + T_1R_2R_1R_2R_1R_2T_1 + \ldots = R_1 + T_1R_2T_1 + R_1R_2 + \ldots = A_1 + T_1R_2A_1 + R_1R_2.A_d = T_1A_2 + T_1R_2R_1A_2 + T_1R_2R_1R_2R_1A_2 + \ldots = T_1A_2 + R_1R_2.A_d = A_1d + A_2d = A_1 + T_1T_2 - R_1R_2(A_1R_2 + A_2). (8.21) \]
$R_d$ is the solar energy reflected out and lost. The only direct solar gain is $T_d$. The energy absorbed in each layer of glass, $A_{1d}$ and $A_{2d}$, provide some solar benefit by heating the glass, reducing the temperature differential between the interior of the structure and the glass, and so reducing heat losses through the glass. Notice that the energy transmitted $T_d$ is unchanged if the glazing is reversed. However, the reflected energy will be different if the glazing sheets are different, as will the energy absorbed.

As derived in Equation 8.21, the individual sheets of glass can have different properties. However, if the sheets are identical, as will occur in most cases, then Figure 8.13 becomes Figure 8.14, so Equation 8.21 becomes

$$T_d = T_2 R_d = R + T_2 R A_{1d} = A + T R A_{2d} = T A R_{d} = A \{1 + T \mathcal{R}(1 + R^2)\} \quad (8.22)$$

where $\mathcal{R} = 1 - R^2$. These equations could describe the single glazing situation shown in Figure 8.15.

The effective solar radiation and the absorptivity in both layers of glass as a function of the angle of incidence are shown in Figure 8.16.

The actual numbers for $T_{eff}$, $R_d$, $A_{1d}$, and $A_{2d}$ are given in Table 8.5.
Figure 8.14: Energy Distribution through Two Identical Sheets

Figure 8.15: Single Sheet Equivalent for Double Glazing
Table 8.5 as well as ?? were generated using MATLAB, and, as part of the verification that the equations used were accurate, the sum $T_t + R_{1t} + A_{1t} + A_{2t}$ was verified to be unity for all incidence angles.

This in turn leads to analyzing triple glazing as shown in Figure 8.17, comparing it to Figure 8.13 and modifying Equation 8.13 to produce

$$T_t = T_d T_{31} - R_d R_3 R_{t31}$$

where $A_{12t}$ is the composite loss in the two outer layers; the problem with this analysis is that we do not know the heat absorbed in the individual outer sheets. This can be resolved by...
grouping the two inner sheets whose reflectivity is 
\[ R' = R_2 + T_2 R_3 - R_2 R_3, \]
and putting on a single outer sheet to give the triple glazed equations for absorptivity of the outer sheet as

\[ A_{1t} = A_1 + T_1 R'd'A_11 - R_1 R'd', \]

so the absorptivity in sheet #2 is

\[ A_{2t} = A_{1d} - A_{1r}. \]

![Figure 8.17: A Triple Glazing Model](image)

Next suppose all three sheets of glass are identical, then the energy equations become

\[
T_t = T_d T_1 - R d = T_3 R - T_2 R_2 R_t = R_d + T_d R_1 - R d R = R \{1 + T_2 R + T_4 R_2 - T_2 R_2\} A_{1t} = A + T R d
\]

\[ 'A_1 - R R d' = A \{1 + T R (R + T_2) R_2 - T_2 R_2\} A_{12t} = A d \{1 + T d R_1 - R d R\} = A(1 + R R (R + 1)) \]

\[ \{1 + T_2 R_9 R_2 - T_2 R_2\} A_{3t} = T d A_1 - R d R = T_2 A R R_2 - T_2 R_2 (8.24) \]

The actual numbers for \( T_{eff}, A_1, A_2, A_3, \) and \( A_{tot} = A_{1t} + A_{2t} + A_{3t} \) are given in Table 8.6, and shown graphically in Figure 8.18; this data was generated using MATLAB, and, as part of the verification that the equations used were accurate, the sum \( T_t + R_{1t} + A_1 + A_2 + A_3 \) was verified to be unity for all incidence angles.

The maximum value of \( A_{tot} \) is 17.43% and occurs at incidence angle of 68°.

The analysis can continue by adding an outer fourth sheet of glass, sheet #1, to the triple glazed system, which will determine the total transmissivity and reflectivity, as well as the energy absorbed in sheet #1. Grouping sheets #1 and #2 as a double glazed pair, and sheets #3 and #4 as a second double glazed pair, the absorptivity in the outer sheets #1 and #2 is then determined, so the energy absorbed in sheet #2 is established.
8.7 Transmission and Absorption for Normal Radiation

In this section we consider the transmission and absorption of solar energy when its direction is normal to the glass, so \( \rho = 0.0419 \). We also assume the refractive index of the glass is 1.515. First, if the absorption of glass was zero in Equation 8.20, then a single sheet of glass would transmit \( T = (1-\rho)21-\rho^2 = 0.9195 \). Similarly, using Equations 8.22 and 8.24 we find the transmissions for double glazing and triple glazing to be 0.8510 and 0.7920, respectively.
The transmission and total absorption for single, double, and triple glazing as a function of absorptivity of the glass are shown in Figures 8.19 and 8.20 respectively. The transmission figure shows the large reduction in transmissivity for high absorptivity glass and when using triple glazing. As discussed in Chapter 3, single glazing is unacceptable in a temperate climate such as New England, and now one can see the penalty for triple glazing.

Figure 8.19: Transmissivity as a Function of Absorptivity
8.8 Daily Solar Gain through Glazing

For south-facing vertical collectors at latitude 42.3 north, that of Boston, Massachusetts, we consider three times of the year, winter solstice with declination \(-23.45^\circ\), equinox with declination zero, and summer solstice with declination \(23.45^\circ\). The times of day considered range from solar noon, when the solar elevation is maximum, to when the solar elevation is zero or the angle of the Sun on the collector is \(90^\circ\).
The effective transmissivity for single, double, and triple glazing in Boston at winter solstice is shown in Figure 8.21. The absorptivity of each sheet of glass is assumed to be 3%. The length of the day is from 4.33 hours before solar noon to 4.33 hours after solar noon.

Table 8.7: Daily Solar Gain

<table>
<thead>
<tr>
<th>declination</th>
<th>single</th>
<th>double</th>
<th>triple</th>
</tr>
</thead>
<tbody>
<tr>
<td>−23.45</td>
<td>1587</td>
<td>1381</td>
<td>1162</td>
</tr>
<tr>
<td>0</td>
<td>977</td>
<td>722</td>
<td>570</td>
</tr>
<tr>
<td>23.45</td>
<td>189</td>
<td>113</td>
<td>75</td>
</tr>
</tbody>
</table>

Table 8.8: Solar Gains at the Solstices and Equinox

<table>
<thead>
<tr>
<th>declination</th>
<th>solar noon</th>
<th>end collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>−23.45</td>
<td>24.25°/69.82%</td>
<td>56.27°/29.79%</td>
</tr>
<tr>
<td>0</td>
<td>47.7°/42.4%</td>
<td>90°/0%</td>
</tr>
<tr>
<td>23.45</td>
<td>71.15°/9.8%</td>
<td>89.17°/0%</td>
</tr>
</tbody>
</table>

Repeating the same calculations at equinox produces Figure 8.22. Finally, doing the same at summer solstice produces Figure 8.23.
Integrating the solar energy passing through the glazing, single, double, and triple over a sunny day, assuming 280 BTU/ft²-hour strikes a surface normal to the incident radiation, then the energy passing the inner layer of glass is given in Figure 8.7.

Single glazing is unsatisfactory for use in the Boston area. Triple glazing results in substantial loss of incoming radiation. Possibly double glazing is the best compromise, so we concentrate on that. For the three times of the year we calculate the angle of the Sun on the glazing and the efficiency of collection at solar noon and at the end of the collection day. The numbers are shown in Figure 8.8. For example, at winter solstice the end of the collection day occurs when the solar elevation is zero, the angle on the collector is 56.27°, and the efficiency of collection is 29.79% just as the Sun sinks below the horizon.

8.9 Transmissivity through LoE Glass

In Figure 8.4 we saw the transmissivity of 1/4” glass and 1/8” acrylic. There is not much that can be done to improve the transmissivity of either material. Reducing the iron content is the only improvement that is possible, at least over the waveband $0.3 \leq \lambda \leq 2.7$. We know that for wavelengths greater than $2.7 \mu$, glass is opaque. The peak of the energy radiated, as given by Wien’s displacement law, from a human body at 310°K, occurs at $2897310 = 9.35 \mu$, and from a room at 290°K is at $2897290 = 9.9910 \mu$: refer to Figure 6.3.

There are three mechanisms at work in the capture of solar energy through glass:
1. The transmission through the glass of the incoming solar radiation over the solar energy spectrum, essentially from 0.3μ to 2.7μ, to heat the room behind the glass. Glass does this well.
2. The absorption by internal surfaces in the room receiving the solar energy and the isotropic re-radiation of some of this energy.
3. The absorption by the glass of some of this re-radiated energy and in turn its isotropic re-radiation.

This third mechanism is where the problem occurs. Glass has an emissivity of about 0.9, so it is a poor medium to conserve the outgoing energy. What is needed is a selective surface on the glass that has a low emissivity to all re-radiated energy, from 2.7μ to about 30μ, while not attenuating the incoming solar radiation.

### 8.9.1 Standard LoE Glass

LoE coatings had their origins in the 1940s, but no commercial product resulted until the OPEC oil embargoes of 1970s, when the price of a barrel of oil increased by almost 800%. In response, the federal government funded research into energy efficiency of windows.

The ERDA (Energy Research and Development Administration), now DOE (Department of Energy), funded studies into the mechanisms of heat transfer in windows at LBNL (Lawrence Berkeley National Lab). One technology that resulted was energy efficient glass with LoE coatings. “From 1976 to 1983, LBNL received $2 million ($5.5 million in current dollars) in funding from DOE to support industry’s low-E R&D efforts with thin film testing, field testing of LoE prototypes.” [J. Rissman and H. Kennan, “Low-Emissivity Windows,” American Energy Innovation Center, energyinnovation.org/wp-content/uploads/2013/03/Low-E-Windows-Case-Study.pdf, as accessed on December 20, 2013].

A group of graduate students at MIT experimented with a transparent film applied to glass to reduce heat losses, and formed a company to market it. They were not successful, but their work attracted the attention of the DOE. DOE gave them a grant of $700,000 provided they worked with LBNL to develop a transparent film they called Heat Mirror. This became the first commercial LoE glass product in 1981, and was when applied to an interior surface of multiple-pane windows.
Growth in market share is strongly a function of market category. Computer software is probably the most volatile, where a product life cycle may be months. For consumer electronics it is years, but in residential construction it is decades. However, during the 1980s the rise in market share for LoE glass in the window market, as can be seen in Figure 8.24, was spectacular.

LoE glass has substantial advantages when installed on any surface but the side facing the Sun — it reduces the unwanted solar gain and reduces the heat flow to the cold outside. Now, most residential windows in the United States are LoE, and increasingly, building codes are requiring LoE glass in all applications, no exemptions. Further, they are requiring all glass to have a low Solar Heat Gain Coefficient. This is the kiss of death for passive solar heating. It is a mindless mandate imposed by people in charge who do not know what they are doing.

“The Northern zone is the perfect candidate for a LoE glass with a high SHGC and the North/Central and even the South/Central can greatly benefit from high SHGC LoE glass on southern exposures. Unfortunately, even though the average American family spends far more on heating than air conditioning, both Energy Star, LEED, and the International Energy Conservation Code (IECC) seem to be color blind when it comes to space heating and the benefits of LoE glass with a high SHGC [‘Low-e Glass, A Nation Divided,’ https://sunhomedesign.wordpress.com, May 3, 2007].”

“Energy Star requirements for SHGC seem to be all about cooling as the following requirements for SHGC for each climate zone indicates. The IECC is no better, only requiring a SHGC of ≤ 0.40 for any residence with less than 3,500 Heating Degree Days (HDD). LEED only takes the cooling bias further by requiring even lower SHGC’s in the South/Central and Southern climate zones [op cit].”

The SHGC (Solar Heat Gain Coefficient) defines the window’s ability to transmit solar radiation. Theoretically, SHGC lies between 0 and 1. However, the upper limit for single
Glazing is about 0.92; about 4% is lost at each air/glass interface. Additional losses are caused by impurities in the glass such as iron, as discussed earlier in this chapter.

For passive solar heating through glass, a high SHGC is needed. Unfortunately, LoE glass does not provide this. Instead it is used to reduce the unwanted transmission of solar energy through the glass, and/or to reduce the heat loss to the environment through the glass. As such, its SHGC is low and it is not suitable when solar gain is needed.

Glass companies designate the surfaces of glass as surfaces 1 (outside exposed to the weather, 2, 3, and 4 (inner exposed to the interior). LoE glass is manufactured by depositing a thin metal or metallic oxide layer on surface 3; such a coating on an exterior surface would be susceptible to mechanical abrasion or moisture deterioration.

The coatings are of two varieties. Soft coat uses a spluttering technique using the MSVD (Magnetron Spluttering Vacuum Deposition) process; it is not durable and has limited life. Hard (pyrolytic) coat is manufactured under the CVD (chemical vapor deposition) process, which covalently bonds to the glass, and so is very durable.

The trick in making passive solar collection effective is to have more solar heat gain than losses through the diathermanous skin. Certain techniques can be used to reduce the heat losses, such as using LoE (low emissivity) glass, or using blinds or shades when the Sun is gone. These are considered parts of a passive solar system rather than pushing them into the active solar category.

The transmissivity of clear glass and three kinds of LoE glass are shown in Figure 8.25 [Lawrence Berkeley National Laboratory]. There is considerable reduction in transmissivity in the near infrared waveband for the high solar gain LoE glass. This can be seen in the extended wavelength responses of Figure 8.26, which also shows the reflectivity characteristics of LoE glass that results in lower heat loss during cold nights.

As discussed earlier, the total solar energy in the near infrared waveband is 45.845%. The solar energy in specific wavebands is shown in Table 8.9.

Also shown in Figure 8.26 are the average transmissivities over the indicated wave-bands. For example, in the waveband 1.5→2.0 μ, clear glass has an average transmissivity of 0.74, whereas the LoE glass has an average transmissivity of 0.42.
Figure 8.25: Transmissivity of Various Glasses

Figure 8.26: Comparison of LoE to Clear Glass
Using Table 6.1, which gives the percentage of solar energy up to a specified wavelength, we now reproduce this information to give the percentage of solar energy in a specified waveband in Table 8.10.

### Table 8.9: Solar Energy by Waveband

<table>
<thead>
<tr>
<th>waveband</th>
<th>% solar energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7→1.0</td>
<td>22.174</td>
</tr>
<tr>
<td>1.0→1.5</td>
<td>15.214</td>
</tr>
<tr>
<td>1.5→2</td>
<td>5.503</td>
</tr>
<tr>
<td>2→2.7</td>
<td>2.954</td>
</tr>
</tbody>
</table>

### Table 8.10: Energy by Waveband

<table>
<thead>
<tr>
<th>waveband</th>
<th>% energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3→0.5</td>
<td>23.40</td>
</tr>
<tr>
<td>0.5→0.7</td>
<td>24.28</td>
</tr>
<tr>
<td>0.7→1.0</td>
<td>22.17</td>
</tr>
<tr>
<td>1.0→1.5</td>
<td>15.21</td>
</tr>
<tr>
<td>1.5→2.0</td>
<td>5.50</td>
</tr>
<tr>
<td>2.0→2.7</td>
<td>2.95</td>
</tr>
</tbody>
</table>

The average transmissivities in the waveband 0.3→0.5 μ is about 0.65 for clear glass and 0.55 for LoE. Multiplying the energies in each waveband by each transmissivity gives:

- clear → $0.65 \times 23.4 + 0.85 \times 24.28 + 0.75 \times 22.17 + 0.68 \times 15.21 + 0.74 \times 5.5 + 0.76 \times 2.95 = 74.68$
- LoE → $0.55 \times 23.4 + 0.8 \times 24.28 + 0.7 \times 22.17 + 0.62 \times 15.21 + 0.42 \times 5.50 + 0.12 \times 2.95 = 66.31$

These numbers are the percentage of normal incident solar radiation passing through each type of glass. At 880 W/m² this amounts to 657 W/m² for clear glass and 583.53 W/m² for LoE glass. So the clear-day reduction is 73.65 W/m². This is unacceptable for glass facing the equator and intended for passive solar heating.

### 8.9.2 New Generation LoE Glass

A number of companies have developed products that could substantially change the passive solar heating industry, provided misguided energy codes permit them.
Saint Gobain UL Ltd. has developed a product called PLANITHERM®TOTAL+. They term it a “new generation advanced performance LoE glass . . . . A unique combination of multiple metal oxide layers are applied to high quality PLANILUX clear float glass using a magnetically enhanced cathodic spluttering process under vacuum conditions. The resultant microscopically thin and transparent metallic coating very effectively reflects long-wave heat radiation back into a room, thereby retaining heat within a building, whilst maximizing natural light transmission.”

Table 8.11

<table>
<thead>
<tr>
<th>glazing system</th>
<th>T = –30°C</th>
<th>T = –10°C</th>
</tr>
</thead>
<tbody>
<tr>
<td>clear single pane</td>
<td>-19</td>
<td>-3</td>
</tr>
<tr>
<td>clear double pane</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>LoĒ180, air filled</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>LoĒ180, argon filled</td>
<td>10</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 8.12

<table>
<thead>
<tr>
<th>glazing system</th>
<th>visible</th>
<th>SHGC</th>
<th>U air/argon</th>
<th>UV</th>
<th>fading</th>
</tr>
</thead>
<tbody>
<tr>
<td>clear single pane</td>
<td>90%</td>
<td>0.86</td>
<td>1.04/-</td>
<td>0.71</td>
<td>0.84</td>
</tr>
<tr>
<td>clear double pane</td>
<td>82</td>
<td>0.78</td>
<td>0.48/-</td>
<td>0.58</td>
<td>0.75</td>
</tr>
<tr>
<td>LoĒ180, air filled</td>
<td>76</td>
<td>0.72</td>
<td>0.34/0.30</td>
<td>0.50</td>
<td>0.68</td>
</tr>
<tr>
<td>LoĒ180, argon filled</td>
<td>80</td>
<td>0.69</td>
<td>0.31/0.26</td>
<td>0.29</td>
<td>0.63</td>
</tr>
</tbody>
</table>

LoĒ180 is a recent addition to the Cardinal Glass range of LoE windows. They provided the data given in Table 8.11. The temperature of the inner sheet of a window system for the ambient temperature $T$ is as given in column 2; recall from Chapter 3, the temperature gradient across a sheet of glass is negligible. Also, if the temperature of the inner sheet of glass is at room temperature, the heat loss through that window can be considered zero.

Cardinal also provided the data given in Table 8.12, where the second column gives the percentage of visible light transmitted, a number that appears to concern some architects but is of little importance to an engineer concerned with winter heat loss. What is more important is the SHGC and the winter U-factor.

Cardinal Glass defines the fading transmission as the fraction of energy transmitted in a spectral range from 0.3μ to 0.6μ. They point out that the shortwave energy around 0.3μ is responsible for the fading of upholstered furnishings and carpets.

U-Factor — This represents the heat flow rate through a window expressed in BTU/h·ft²·°F, using winter night weather conditions of 0°F outside and 70°F inside. The smaller the number, the better the window system is at reducing heat loss.
Chapter 9

Climate and the Siting of a Solar Structure

Europe, North America, and China are heavily focused on reducing the burning of fossil fuels. It does not appear that nuclear fission energy will occupy a dominant role in this effort, especially since Germany, the biggest economy with the biggest population in Europe, is phasing out its nuclear industry. As discussed in Chapter 1, hydroelectric power, valuable though it is, is a fairly mature industry and so incapable of massive expansion. There are hard limits to the conversion of vegetation to energy. Tidal power is limited to a few locations, and wave power barely registers as a viable energy source.

It is imperative that we find a replacement for fossil fuels as soon as possible. Nuclear fusion could be the answer, but we have been trying to make it work for the last 70 years. We are left with clean, reliable, solar energy, and its offspring, wind energy.

In Chapter 7 we attempted to determine the amount of solar energy reaching the Earth’s surface, but were only partially successful. Here we will try and determine the climatic conditions in various locations in the world so as to maximize the effectiveness of solar energy for residential use. In doing so we will devote more attention in the affluent temperate areas of the world. The reason is simple: Affluent areas use substantially more energy per capita than third world countries, as was discussed in Section V of Chapter 1, and so extensive use of solar energy in these countries is most effective in reducing CO₂ in the atmosphere.

The most populous countries and the most affluent countries in the world are listed in Table 9.1, although not all are both. The numbers are for 2014 [www.worldometers.info]. The population numbers are in millions. Notice that, by some measures, countries that are increasing in population the fastest, as given by the last column in the table, are the most overpopulated already — India and Pakistan, for example.
 Countries whose population is decreasing, such as Russia, Japan, and Germany, may have some economic difficulties as populations age and fewer workers remain to support an elderly population with increasing strains on medical needs.

China is investing heavily in solar energy. It is unlikely that India will be a major user of solar energy in the near future. It does not have the industrial base or the financial resources. The majority of its population is engaged in subsistence farming, and global warming is not a primary concern.

The energy needs in third world countries are vastly different from those in the advanced world. In Europe and North America, a reliable national grid is taken for granted, so power can be drawn from it at any time, and returned to it at any time with grid-connected PA panels. In Haiti there is no reliable electrical grid. The electronic device most valuable to citizens in Haiti is the cell phone, and the problem there is how to charge it. The answer is a small, portable, and inexpensive (by western standards) PV based charger. To own one in Haiti is to have freedom.

This chapter explores the requirements on a stand-alone residential building lot that would make it suitable for solar heating. Possibly the most important factor in determining if a solar house is appropriate is the climate.

We start by discussing climate classifications, notably the Köppen climate classification and the zoning maps popular in Europe and the United States. We will see that these are gross measures, of limited use in determining the energy needs of a residence. Better metrics are the HDD (heating degree days) and the CDD (cooling degree days). Even better are HMD (heating monthly degree days) and CMD (cooling monthly degree days).
Many lots are not suitable to accept solar systems. The area may not receive enough solar radiation (too near the poles, or too much cloud cover). An urban lot shaded by tall buildings is another obvious example.

Even if the lot itself could be suitable, most existing houses present challenges to accepting solar systems. The roof angle and pitch may be wrong. The windows in the house are facing away from the Sun. There is little thermal mass to smooth out the temperature peaks and valleys and so make passive solar heat acceptable. In essence, the ideal situation is to design a house from scratch to accept the solar elements, not as a late-stage add-on, but fully integrated into the fabric of the house. Hence, the main thesis of this work as given by its title, “Building Integrated Solar Systems.”

The Passivhaus design is a total commitment to “building integrated.” More such commitments must take place. It can be grass-roots based, or can be spurred by governmental action. However, a danger exists when building codes, enforceable by law, run contrary to the needs of a solar structure. There are a number of such cases, of which the most glaring example is the requirement that all window glass be LoE: this was discussed in Section IX of Chapter 8.

9.1 The Köppen Climate Classification

The German climatologist Wladimir Köppen published his classification of the world’s climates in 1884. Based on the premise that native vegetation is the best measure of climate, it remains one of the most popular systems in use today. Precipitation and temperature are key factors [T.L. McKnight and D. Hess, Climate Zones and Types: The Köppen System, Prentice Hall, New Jersey, 2000]. The classification starts with an upper case letter (A through E) followed by one or two lower case letters or the letters W or S. Temperatures all relate to low elevations.

Some terms used in this section need to be defined:
biome: a major ecological community type, such as desert, rain forest, or grassland.
chaparral: an ecological community comprised of shrubby plants adapted to dry summers and moist winters.
rain forest: a forest, from tropical to temperate, accepting at least 70” of annual rainfall.
esteppe: a treeless tract of arid land.
tundra: an treeless plain with heavy, black soil supporting flowering dwarf herbs over permafrost in arctic or subarctic regions.

For the convenience of the reader, the temperature equivalents in Celcius to Fahrenheit are given by the conversion formula $y_F=95x_C+32$.

The initial upper case letter in the Köppen classification determines the gross climate class:
A: moist tropical climate with constant high temperatures, with monthly averages in excess of 18°C.
B: arid and semiarid with large temperature range. This class contains two subgroups, S for semiarid/steppe, and W for arid/desert.
C: mid latitudes with warm, dry summers and cool, wet winters.
D: continental climates with moderate to low precipitation and large seasonal variation in temperature.
E: cold climates with most of the year at below-freezing temperatures.

A second, lower case letter designates seasonal temperature and precipitation:
a: hot summers with hottest monthly averages over 22°C and in C and D climates.
b: warm summers with hottest monthly averages below 22°C and in C and D climates.
c: cool short summers with less than four monthly averages over 10°C, and in C and D climates.
d: very cold winters with the coldest monthly average below −38°C.
f: adequate precipitation in all months, with this letter usually following A, C, and D.
h: dry and hot climate with mean annual temperature over 18°C, and in B only.
k: dry and cold with mean annual temperature below 18°C, and in B only.
m: rain forest climate with monsoon cycle, with this letter always following A.
s: dry season in the summer.
w: dry season in the winter.

The climatic regions of the world of interest to us here are the ones in which solar heating can be effective and where solar heating is needed. This excludes the tropics (from 23.45°S to 23.45°N) and near the poles: The arctic/antarctic circles are ±66.55° north and south, but latitudes 60 degrees or more from the equator are unsuitable for solar heating.

The Mediterranean climate is in the class Cs, with a dry warm-to-hot summer and a wet moderate winter. The vegetation, predominantly chaparral biome, has adapted to the dry summers. The temperature range is about 7°C and precipitation about 17”. The latitude range is 30° to 40°. Cs regions include central and southern California, coastal zones of the Mediterranean Sea, coastal southern and western Australia, the Cape Town region of South Africa, and coastal Chile.
The BS class is dry, mid-latitude steppe predominantly of grasslands; a drier climate would make BS into BW, whereas a wetter climate would make BS into tall grass prairie. BS regions include the west central United States and from the steppes of eastern Europe to the Gobi desert and North China. The BS climate is characterized by large seasonal temperature variations.

The eastern and central United States above latitudes 35 are classified as Dfa, a continental climate with adequate rainfall and large temperature changes. Much of western Europe, outside the Mediterranean regions, is also Dfa. We conclude that the regions of the world suitable for solar space heating are the ones shown in Figure 9.1; in the B classification, both s and w are to be included.

There is no ideal system for climate classification. For example, classification Cf, where the C means warm dry summers and cool, wet winters, and the f means adequate precipitation year round, is applied to the British Isles, which is rainy much of the time. The British Isles, particularly the western coastal areas, experiences considerable rainfall and moderate temperatures year round due to the prevailing westerly winds from the Atlantic.

The problem with any classification system is that it tries to cram the diverse climates of a region into one of a limited number of categories.

The Köppen system does not take into account a marine environment. Marine climates can occupy a small area. The Olympic Mountains between Seattle and the Pacific Ocean cause the warm, moist winds from the west to rise and so cool down and produce rain. Seattle gets lots of it, but it is almost desert conditions over the mountains into central Washington State: In this area trees do not grow unless they are irrigated. However, central Washington is the premier area in the United States for growing apples. Orchards here have vertical irrigation pipes with spray heads above the tops of the trees, and the farmer can water the trees as needed.

The Köppen system applied to Europe is shown in Figure 9.2. Western Europe and the British Isles are classified as Cfb. The regions surrounding the Mediterranean Sea are Csa.

Western Europe is fortunate to have an enormous natural resource pumping massive amounts of energy to moderate the winter climate — the Gulf Stream, which starts in the warm waters of the Caribbean and makes straight for Europe. Without the Gulf Stream there would be icebergs in the English Channel.

The benefits of the Gulf Stream are largely dissipated by the time the prevailing westerlies reach Central Europe. Much of Eastern Europe is classified Dsb, a continental climate with large seasonal temperature swings, with warm dry summers. The Mediterranean Sea produces its own climactic zone, called a Mediterranean climate.

The Köppen map for Asia and Australasia is shown in Figure 9.3. The classifications in Asia run from A to D, far more diverse than Europe, which is confined to classifications C and D.
It should be pointed out that the Köppen classification is approximate, and not intended for specific locales. Further, one can see little practical application of the Köppen classification during the design process for a solar structure. Instead, actual measurements are needed.

The Köppen map for North America is shown in Figure 9.4. Little of the land area of Canada lies in the B, C, or D classification, so its areas most suitable for solar space heating are adjacent to the contiguous United States.

Less precise than the Köppen classification is the system shown in Figure 9.5 for Europe. From warmest to coldest (sort of) the categories are Mediterranean, subtropical, humid, humid continental, marine, steppe, and tundra.

A climate map of China using a different set of categories is shown in Figure 9.6. It is no more precise than Figure 9.5.
9.2 Climates Zones by the Number

The Köppen climate classification is of limited use in determining the energy needs of a residence. A little better is the numerically based international climate classification specifically designed to aid energy calculations in buildings: ANSI/ASHRAE/ IESNA Standard 90.1. In this standard there are eight numerical zones based on average temperature and precipitation, where the warmest zone is 1, and the coldest is 8.

There is also a marine classification, which satisfies these four criteria:
1. Mean temperature of coldest month −3°C/18°C, or 27°F/65°F.
2. Mean temperature of warmest month < 22°C, or < 72°F.
3. At least four months with mean temperatures over 10°C, or 50°F.
4. Heaviest precipitation occurs in winter, which is at least three times the amount in the driest of the warmer months.
In the United States, zones 1–2 might be considered the southern United States, with predominantly cooling requirements. Zone 3 is mixed, with probably about the same energy for heating and for cooling on an annual basis. Zone 4 might be considered middle United States, with significant heating requirements. Zones 5–8 have a heating climate, with a small percentage of the total energy load devoted to cooling. Neither of these methods, Köppen or zone, is a satisfactory determinant of the heating or cooling loads for a residence.

There are climate zone maps for all parts of the world. The eight zones are based on HDD (heating degree days) and CDD (cooling degree days) plus humidity and precipitation. These are of great value to gardeners, who use them to determine what plants are suitable for their specific location. For example, will a particular plant survive the New England winter, or will another thrive in the arid American Southwest? However, as far as energy load on a structure is concerned, we are more interested in HDD and CDD than on rainfall, and the marine classification has little value.

HDD is defined with respect to a given base, and the standard base is $18^\circ C = 64.4^\circ F$ or $65^\circ F = 18.33^\circ C$; the small difference is not of concern since the variations in HDD numbers are much greater. Why choose $18^\circ C$ as the base temperature for assessing HDD, when the comfort zone temperature is closer to $21^\circ C$? The reason is that human activity in the house, heat loss from human bodies, plus electrical use add sensible heat to the residence, so at an ambient of...
18°C the interior will be at about 21°C, the heating load is zero, so the heating degree number for that day is zero. However, the winter setting for the thermostat is about 21°C.

Why did Standard 90.1 set the CDD base at 10°C? This is close to the temperature of a refrigerator at 5°C, far too cold to be close to the comfort zone for humans.

It is considered standard operating procedure to set the thermostat in summer above the winter setting. Clothing in summer is lighter than that in winter. A summer thermostat setting will be about 24°C, or perhaps a little higher. The human activity differential of about 3°C will add to the cooling load in summer rather than subtract; this would argue for a cooling base of 21°C, which results in a cooling load of zero at this ambient temperature.

![Figure 9.5: Climate Classifications for Europe](image)

To illustrate the problem with the Standard 90.1 cooling base, we consider the monthly averages of cooling degree days for three southern European cities at 10°C and 21°C, and also include the monthly averages for heating degree days at the standard base of 18°C. These
yearly averages using www.degreeday.net over the time period June 2012 to May 2015 are shown in Table 9.2. The bottom row of this table gives the overall yearly average for all four cities. It appears that these cities have a predominantly cooling climate when cooling base 10°C is used, but become predominantly heating climates when 21°C is used.

The climate zones as defined by Standard 90.1 are given in Table 9.3. The HDD base is 18°C, which is fine, but the CDD base of 10°C is a problem.

![Climate of China](Figure 9.6: The Climate of China)
9.3 Monthly Averages for Heating and Cooling

www.degreeday.net is an excellent site that provides HMD (heating monthly degree days) and CMD (cooling monthly degree days) numbers over a 36-month period ending at the month prior to the current date. When this section was being prepared, the 36 month period was June 2012 through May 2015. This site relies on temperature data from, which gets data from thousands of weather stations worldwide. They recommend using airport data when available since it is the most reliable; aircraft safety is of prime importance, and rapid temperature changes could indicate bad weather conditions.

Table 9.2: Influence of Base Temperature

<table>
<thead>
<tr>
<th>city</th>
<th>cool base10</th>
<th>cool base21</th>
<th>heat base18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rome</td>
<td>2553</td>
<td>457</td>
<td>1479</td>
</tr>
<tr>
<td>Lisbon</td>
<td>2426</td>
<td>281</td>
<td>1159</td>
</tr>
<tr>
<td>Barcelona</td>
<td>1949</td>
<td>405</td>
<td>1168</td>
</tr>
<tr>
<td>average</td>
<td>2703</td>
<td>466</td>
<td>1239</td>
</tr>
</tbody>
</table>
Table 9.3: Climate Zone

<table>
<thead>
<tr>
<th>Zone</th>
<th>CDD and/or HDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CDD &gt; 5000</td>
</tr>
<tr>
<td>2</td>
<td>3500 &lt; CDD ≤ 5000</td>
</tr>
<tr>
<td>3</td>
<td>2500 &lt; CDD ≤ 3500</td>
</tr>
<tr>
<td>4</td>
<td>CDD &lt; 2500 and HDD ≤ 3000</td>
</tr>
<tr>
<td>5</td>
<td>3000 &lt; HDD ≤ 4000</td>
</tr>
<tr>
<td>6</td>
<td>4000 &lt; HDD ≤ 5000</td>
</tr>
<tr>
<td>7</td>
<td>5000 &lt; HDD ≤ 7000</td>
</tr>
<tr>
<td>8</td>
<td>HDD &gt; 7000</td>
</tr>
</tbody>
</table>

Another reason for relying on HMD and CMD numbers is that they depend on temperature only, and temperature is easy to measure, and to measure automatically. Every major city, at least in Europe and the United States, has a number of sites that keep good temperature records. Further, the aggregation over a month smooths out daily extremes without losing seasonal variations.

HMD and CMD are based on daily average temperature deviations from a defined base temperature $T_b$ summed over a complete year. In the United States $T_b = 65°F, 18.3°C; NOAA (National Oceanic and Atmospheric Administration) uses the average of the minimum and maximum daily temperature to determine $T_b$. In Europe, $T_b$ varies from 15.5°C to 19°C, and the daily temperature is usually assumed the same way as NOAA.

Why choose a Fahrenheit scale when Celsius is available? Partly because of the extensive climate databases in the United States, all relying on the same heating base temperature of 65°F, and partly because the European base temperatures vary. Further, the translation from the Fahrenheit number $HDD_f$ to Celsius $HDD_c$ is simply $1 HDD_f = 59HDD_c$.

If the average daily temperature (average of the temperature high and the temperature low) is 50, then $65 - 50 = 15$ is added to HMD. Thus, with a fixed temperature base, the daily addition to the annual HMD is a linear function of the average daily temperature. It is also a linear function of the base temperature: With an average daily temperature of 50 and a base temperature of $65 + x$, then the daily addition to the HMD is $65 + x - 50 = 15 + x$.

It was decided to test the accuracy of the data by using three base values for HMD and CMD. These were 65°F, 70°F, and 75°F. Fitting a quadratic polynomial through the three data points should produce a very small coefficient for the quadratic term.

Given three points $(0, f_1), (1, f_2), (2, f_3)$, the quadratic polynomial passing through these three points, as determined by the Lagrange interpolating polynomial, is
For selected cities in Europe, the monthly HMD averages for all months over 200 HMD were calculated for the 36 months from June 2012 through May 2015, as extracted from www.degreeday.net, with data taken from the three bases 65, 70, and 75°F. These are shown in Figure 9.8, where at 65°F the abscissa value is 0, and at 75°F the abscissa value is 2. The corresponding coefficients for the second order polynomial fit are given below for the indicated cities: for Lisbon the polynomial is \( 1.063x^2 + 147.42x + 345.39 \), and for Moscow the polynomial is \( 1.259x^2 + 145.42x + 883.61 \).

Also shown in Figure 9.8 as the dashed, heavy blue line is the average for the cities chosen. The city Athens was also studied, but its data indicated a predominantly cooling climate, so it was excluded from the study for HMD. The annual average HMD and CMD at their respective base temperatures are 2074 and 704. In fact, few cities in Europe qualify as having a predominantly cooling climate.

The span of the HMD is 2, and measures the temperature from 65 to 75. The span of CMD is 2, and measures the temperature from 75 to 65. Using the collection system of www.degreeday.net, the slope of the HMD from the span of 1 to 2 is the average number of monthly days times the span divided by the actual temperature range, that is, \( \frac{365 \times 12 \times 10}{2} = 152.08 \).
We will test this against the actual data. For the set of European cities considered, the average quadratic is

\[ a_2x^2 + a_1x + a_0 = 1.3329x^2 + 146.1725x + 604.1217. \] (9.2)

The average slope of this polynomial is \( a_2 + a_1 = 1.3329 + 146.1725 = 148.40 \), close to the expected slope of 152.22. The monthly trend lines are essentially straight since \( a_2 \) is small, and the coefficients \( a_1 \) are about the same. The big variation is with the \( a_0 \) coefficients. The three cities with the smallest \( a_0 \) are Lisbon, Rome, and London. Lisbon and Rome have a predominantly cooling climate and are the southernmost of the cities chosen. The three cities with the highest \( a_0 \) are Moscow, Helsinki, and Stockholm, the northernmost of the cities chosen.

Notice that \( a_2 \) in this study varied from 0.023 to 0.091. Even at this high end the contribution of this term to the HMD numbers would be at most \( 0.091 \times 2^2 = 0.364 \), well within the rounding error by defining HMD numbers as integrals.

Next, most major cities in the United States were studied. For the set of U.S. cities considered, the average quadratic is
The average slope of this polynomial is $a_2 + a_1 = 2.29753 + 139.84 = 143.7$, compared to the expected slope of 152.22.

One suspects that the software used in www.degreeday.net rounds the average daily temperatures in one direction. If the rounding was to the next highest integer, 2.1 and 2.9 would round to 3 each, but rounding to the closest integer would produce 2 and 3, and there would be no bias in the calculations.

For the U.S. cities, the coefficient of the quadratic term in this study varies from 1.00 to 14.42. At the low end the contribution to the monthly HMD number amounts to 4, not a significant number but a little disturbing, compared to the tiny number for the European cities. How did this discrepancy occur? Even worse, at the high end the contribution to monthly HMD is 57.68, a large contribution to the total number. It can be argued that the large number is for Orlando, Florida, with its notorious, almost daily, electric storms, and that it is a predominantly cooling climate and so should be excluded from considerations involving HMD. However, the fact remains, where did this discrepancy come from? Is it possible that Weather Underground relies on an unreliable source for U.S. numbers, and the European numbers are from more effectively monitored sites?

Until this century, the predominant application of solar energy in residential use would be thermal. Photovoltaics were out-of-sight expensive, so cooling applications of solar energy were not an option. In recent years the precipitous decline in the cost of a PV generated watt make solar cooling an attractive option in some locations, particularly those with a predominant cooling climate. CDD numbers will be studied next.
The monthly cooling degrees for selected cities worldwide, when the monthly average of degree days is 40 or greater, are shown in Figure 9.9. The coefficients for the quadratic polynomials are shown in Table 9.4. Notice that the coefficients $a_2$ are large, indicating that either there is a problem with the data used by www.degreeday.net, or the correct data was incorrectly processed by that website. The city with the highest $a_2$ coefficient was Athens. When the coefficient $a_0$ was higher than 170, coefficient $a_2$ was within the expected rounding error.

Table 9.4: Quadratic Coefficients for Cooling Cities

<table>
<thead>
<tr>
<th>city</th>
<th>$a_2$</th>
<th>$a_1$</th>
<th>$a_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algiers, Algeria</td>
<td>13.47</td>
<td>-80.81</td>
<td>123.41</td>
</tr>
<tr>
<td>Athens, Greece</td>
<td>47.58</td>
<td>-68.33</td>
<td>156.73</td>
</tr>
<tr>
<td>Bombay, India</td>
<td>3.00</td>
<td>-141.98</td>
<td>240.84</td>
</tr>
<tr>
<td>Cairo, Egypt</td>
<td>8.95</td>
<td>-108.18</td>
<td>195.98</td>
</tr>
<tr>
<td>Calcutta, India</td>
<td>4.43</td>
<td>-129.91</td>
<td>257.00</td>
</tr>
<tr>
<td>Caracas, Venezuela</td>
<td>1.03</td>
<td>-149.10</td>
<td>194.56</td>
</tr>
<tr>
<td>Lisbon, Portugal</td>
<td>1.03</td>
<td>-149.10</td>
<td>194.56</td>
</tr>
<tr>
<td>Mecca, Saudi Arabia</td>
<td>3.35</td>
<td>-139.62</td>
<td>445.21</td>
</tr>
<tr>
<td>Nagasaki, Japan</td>
<td>7.10</td>
<td>-127.28</td>
<td>163.54</td>
</tr>
<tr>
<td>Rangoon, Myanmar</td>
<td>4.19</td>
<td>-135.45</td>
<td>206.71</td>
</tr>
<tr>
<td>Rio de Janeiro, Brazil</td>
<td>13.76</td>
<td>-104.90</td>
<td>115.57</td>
</tr>
</tbody>
</table>
Using the data from www.degreeday.net, the monthly averages $M(i)$, $i = 1, 2, ..., 36$ for Boston from June 2012 to May 2015 are as shown in Figure 9.10. In order to calculate the monthly averages over the three-year period, the first two data points must be moved. This is done with

$$S(k)=M(k+2)+M(k+14)+M(k+26), k=1,2,..., 10$$

$$S(k+10)=M(j)+M(k+12)+M(j+24)$$

so $S$ contains every month of the year for three years. The result is given by the blue response as shown in Figure 9.11. Notice that the three-year average exhibits less randomness than the individual years in Figure 9.10. In this diagram, month 1 is August, and month 6 is January. One would expect the response over a longer period to further attenuate the noise.
The next step is to model the data using the modified sine function

\[ g(i) = n^2 \{1 + \sin(\pi(i-d)45)p\} \] (9.4)

where factors \(d\), \(p\), and \(n\) are determined empirically by trial and error so as to match the points \(S(k)\). That is, the objective is to minimize

\[ \varepsilon = \sum_{k=1}^{n} (S(k) - g(x_i))^2 \] (9.5)

at least as close as the eye can determine. The sample average is

\[ M = \frac{1}{n} \sum_{k=1}^{12} M(k) \] (9.6)
and a zero order approximation will be \( \Sigma k=112 (M(k)-\bar{M})^2 \), so the goodness of fit \( G \) is given by

\[
G=\Sigma k=112 (M(k)-\bar{M})^2 - \epsilon \Sigma k=112 (M(k)-\bar{M})^2, (9.7)
\]

a number between 1, a perfect fit, 0, a useless fit since it is no better than the zero order approximation, and negative numbers indicating that the approximation is worse than useless. This is the principle of a least squares approximation, as was discussed in Section V of Chapter 1.

For the particular data of Figure 9.11, \( G = 0.992 \), a nearly perfect fit, but this is for one city over a three-year period. Some validation of Equation 9.4 based on NOAA data is provided in Section IV.

The model numbers for Phoenix seem strange, but they do fit the actual data.

It is unfortunate that the model for HDD requires three parameters. Attempts to link these, and so reduce the parameters to two, or ideally one, have proven fruitless.
Further research may produce common characteristics between cities. Attempts to do this by latitude failed.

### 9.4 Other Data Sources for HDD and CDD

Another potentially useful source for heating degree days for U.S. cities is the “Comparative Climatic Data” from the National Climatic Data Center of NOAA [http://ggweather/ccd/nrmhdd.htm]. It claims to use 30 years of data from 1981 to 2011. Data for selected cities is given below, with a 65°F base:

<table>
<thead>
<tr>
<th>European city</th>
<th>deg</th>
<th>pow</th>
<th>num</th>
<th>G</th>
<th>lat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lisbon, Portugal</td>
<td>25.0</td>
<td>1.4</td>
<td>320</td>
<td>0.961</td>
<td>38.73</td>
</tr>
<tr>
<td>Rome, Italy</td>
<td>25.0</td>
<td>1.5</td>
<td>400</td>
<td>0.971</td>
<td>41.90</td>
</tr>
<tr>
<td>Bucharest, Romania</td>
<td>22.0</td>
<td>1.7</td>
<td>640</td>
<td>0.987</td>
<td>44.42</td>
</tr>
<tr>
<td>Belgrade, Serbia</td>
<td>23.0</td>
<td>1.4</td>
<td>630</td>
<td>0.970</td>
<td>44.82</td>
</tr>
<tr>
<td>Zurich, Switzerland</td>
<td>23.0</td>
<td>1.1</td>
<td>900</td>
<td>0.972</td>
<td>47.37</td>
</tr>
<tr>
<td>Budapest, Hungary</td>
<td>23.0</td>
<td>1.3</td>
<td>780</td>
<td>0.978</td>
<td>47.47</td>
</tr>
<tr>
<td>Munich, Germany</td>
<td>23.5</td>
<td>1.0</td>
<td>1000</td>
<td>0.973</td>
<td>48.13</td>
</tr>
<tr>
<td>Vienna, Austria</td>
<td>23.0</td>
<td>1.2</td>
<td>800</td>
<td>0.989</td>
<td>48.22</td>
</tr>
<tr>
<td>Paris, France</td>
<td>24.0</td>
<td>1.0</td>
<td>750</td>
<td>0.975</td>
<td>48.85</td>
</tr>
<tr>
<td>Prague, Czech Republic</td>
<td>23.0</td>
<td>0.9</td>
<td>1100</td>
<td>0.980</td>
<td>50.08</td>
</tr>
<tr>
<td>Frankfurt, Germany</td>
<td>23.5</td>
<td>1.0</td>
<td>850</td>
<td>0.987</td>
<td>50.12</td>
</tr>
<tr>
<td>Brussels, Belgium</td>
<td>25.0</td>
<td>0.8</td>
<td>900</td>
<td>0.972</td>
<td>50.87</td>
</tr>
<tr>
<td>London, England</td>
<td>25.0</td>
<td>0.8</td>
<td>800</td>
<td>0.973</td>
<td>51.50</td>
</tr>
<tr>
<td>Warsaw, Poland</td>
<td>25.0</td>
<td>0.7</td>
<td>960</td>
<td>0.978</td>
<td>52.23</td>
</tr>
<tr>
<td>Amsterdam, Netherlands</td>
<td>25.0</td>
<td>0.8</td>
<td>900</td>
<td>0.981</td>
<td>52.37</td>
</tr>
<tr>
<td>Berlin, Germany</td>
<td>23.5</td>
<td>1.0</td>
<td>900</td>
<td>0.983</td>
<td>52.52</td>
</tr>
<tr>
<td>Dublin, Ireland</td>
<td>25.0</td>
<td>0.5</td>
<td>1050</td>
<td>0.947</td>
<td>53.21</td>
</tr>
<tr>
<td>Hamburg, Germany</td>
<td>25.0</td>
<td>0.8</td>
<td>1000</td>
<td>0.974</td>
<td>53.55</td>
</tr>
<tr>
<td>Moscow, Russia</td>
<td>22.0</td>
<td>1.1</td>
<td>1300</td>
<td>0.966</td>
<td>55.75</td>
</tr>
<tr>
<td>Edinburgh, Scotland</td>
<td>25.0</td>
<td>0.5</td>
<td>1100</td>
<td>0.937</td>
<td>55.95</td>
</tr>
<tr>
<td>Stockholm, Sweden</td>
<td>24.5</td>
<td>0.8</td>
<td>1250</td>
<td>0.968</td>
<td>59.33</td>
</tr>
<tr>
<td>Helsinki, Finland</td>
<td>23.5</td>
<td>0.9</td>
<td>1350</td>
<td>0.957</td>
<td>60.17</td>
</tr>
</tbody>
</table>
The data is plotted for Boston and shown as the blue response in Figure 9.12. Also shown is the dashed green line which uses Equation 9.4 with $d = 32$, $p = 1.3$, and $n = 850$, with $G =$
0.9941. Although the fit is excellent, there is an apparent problem with the NOAA data: It is noisy, something not to be expected with a 30-year average, and therefore suspect.

Another site that provides monthly data is www.climate-zone.com/climate. For a comprehensive list of cities across the world, it provides temperature, average, maximum and minimum, as well as HMD, CMD, and a number of other measures such as precipitation and average sunshine. They do not say what the bases are, but they can be calculated from the formulae

\[ Bh = Ta + HDDn, Bc = Ta - CDDn(9.8) \]

where \( B_h \) is the HMD base, \( B_c \) is the HMD base, \( T_a \) is the monthly average temperature, and \( n \) is the number of days in the month. For Boston, Massachusetts one gets the following numbers:

<table>
<thead>
<tr>
<th></th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct-Nov-Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_a )</td>
<td>28.6</td>
<td>30.3</td>
<td>38.6</td>
<td>48.1</td>
<td>58.2</td>
<td>67.7</td>
<td>73.5</td>
<td>71.9</td>
<td>64.8</td>
<td>54.8453336</td>
</tr>
<tr>
<td>HMD</td>
<td>1128</td>
<td>972</td>
<td>818</td>
<td>507</td>
<td>221</td>
<td>32</td>
<td>6</td>
<td>72</td>
<td>321591973</td>
<td></td>
</tr>
<tr>
<td>CMD base</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>113</td>
<td>264</td>
<td>220</td>
<td>66</td>
<td>500</td>
<td></td>
</tr>
</tbody>
</table>

The bottom line uses the formula of Equation 9.8 to determine the base for HMD. The deviation from a base of 65 occurs in the summer months when the numbers for HMD fall below 350.

Notice the coldest month is January, and the hottest month is July. One will find that in most locations in Europe and the United States the coldest/hottest time of the year is close to a month after winter/summer solstice. This is true for predominantly heating climates, but may not be true for predominantly cooling climates.

The numbers for CMD in Boston are not high enough to determine a cooling base. However, for Miami, Florida, here are the numbers from www.climate-zone.com/ climate:

<table>
<thead>
<tr>
<th></th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct-Nov-Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_a )</td>
<td>67.2</td>
<td>68.5</td>
<td>71.7</td>
<td>75.2</td>
<td>78.7</td>
<td>81.4</td>
<td>82.6</td>
<td>82.8</td>
<td>81.9</td>
<td>78.3736691</td>
</tr>
<tr>
<td>HMD</td>
<td>88</td>
<td>51</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0641</td>
</tr>
<tr>
<td>CMD base</td>
<td>156</td>
<td>149</td>
<td>221</td>
<td>306</td>
<td>425</td>
<td>492</td>
<td>546</td>
<td>552</td>
<td>507</td>
<td>412264168</td>
</tr>
</tbody>
</table>

Here we see that Miami has a predominantly cooling climate and the cooling base is 65°F. The formula of Equation 9.8 gives the cooling base as 65.0 provided the CMD number is above 250. As pointed out previously, a cooling base of 18°C, 65°F, is appropriate; I would prefer it be 21°C. However, the cooling base of 10°C as determined by ANSI/ASHRAE/IESNA Standard 90.1 shows very poor judgement.
The www.climate-zone.com/climate website is useful, but has an unfortunate error. It permits you to choose temperatures in Fahrenheit Centigrade, but in the Centigrade tables they use HMD and CMD numbers are from the Fahrenheit table.

9.5 Southern Exposure

In northern (southern) latitudes, a southern (northern) exposure is essential for a residence to be suitable for solar heating. South-facing slopes are therefore more suitable than north-facing slopes. How does one determine the direction of due south? Call in a surveyor, but that is expensive and that surveyor may have little experience in determining north from south. Find someone who can shoot a star (good luck!).

An approximate way to determine the direction of due south is to determine the shadow cast by a vertical stick at noon: the use of a spirit level to ensure that the stake is vertical is essential. But what is meant by noon? There was a full discussion of clock time noon, longitude noon, and the “equation of time” (EoT) in Chapter 5.

As a function of latitude, the azimuth at 15 minutes from solar noon is shown in Figure 9.13 for winter solstice W, eqinox E, and summer solstice S. The azimuth values at S, especially for $L = 30^\circ$, is startling: at $L = 30$ and S, the azimuth is 26.89°.

The message is not to use the shadow cast by a vertical stick to determine a north/south line in summer. What else can be done to minimize directional error? The ideal is to know the value of EoT and the difference between clock time noon and longitude noon.
The EoT at certain Julian days helps establish guidelines:
EoT = 0 J106 (April 16) J165 (June 14) J245 (September 2)
EoT has a maximum deviation from clock noon on February 12 of –14:20, on May 14 of +3:44, on July 26 of –6:25, and on November 3 of 16:23.

![Figure 9.14: Soar Elevation from Shadow](image)

The shadow of length $x$ cast by a vertical stake of height $h$ above level ground can be used to determine the elevation angle of the Sun, as shown in Figure 9.14. The hypotenuse of the triangle has length $h^2 + x^2$, so $\cos \alpha = \frac{1}{1} + \frac{x^2}{h^2}$.

### 9.6 Magnetic Declination

An alternative to using the shadow cast by the Sun on a stake is to use a compass, but for accuracy this requires a good compass mounted on a steady platform with good sight lines that can be staked, plus knowledge of the magnetic declination in the area, remembering that declination changes with time.

Magnetic declination is the difference between magnetic north and true north. If the magnetic declination is east of true north it is considered positive; when west of true north it is considered negative. Solar declination is the same regardless of location on the Earth’s surface, but magnetic declination varies considerably depending on location.

Iron is the dominant magnetic material, and it is iron in the Earth’s core that produces the Earth’s magnetic field. A compass needle points approximately north, plus or minus 30° or thereabouts. If you know your location, you can go online to determine the magnetic declination.

The Chinese were the first to use magnetic declination, around 700 AD. It took a long time for the rest of the world to catch up. In 1510 Georg Hartman determined the declination of Rome. In 1700 Edmund Halley produced contour lines of equal declination on a map; since then such maps have been in constant use.
Magnetic declination changes with time at a much faster rate than the precession of the equinoxes. For example, the changes in declination for two Canadian cities over time is shown in Figure 9.15 [http://gsc.nrcan.gc.ca/geomag/field/magdec_e.php]. The rate of change in declination for Churchill (latitude/longitude approximately 59/93) is $+1^\circ$ every 3.6 years for the years 1720 to 1820, but after 1840 the declination declines. Contrast this with the angle of obliquity, which changes by 50.3” per year, or $1^\circ$ over 51.57 years. On the other hand, the declination change at Ottawa (latitude/longitude approximately 45/75) has hovered around $-10^\circ$ for the last 300 years.

Magnetic declination will have increased variance the nearer one gets to the poles. In latitudes between 50° south and 50° north (the U.S.–Canadian border) the variance is typically around 2° every 100 years.

The variability of the magnetic declination for the United States is shown in Figure 9.16, and for Europe in Figure 9.17.
9.7 Visibility Distances

The time of sunrise, as given in meteorological reports and commonly reported in the local newspaper, is the time when the Sun rises above a flat horizon. So, at sunrise the Sun has elevation angle of zero. As far as a house is concerned, it is impossible to have a situation where the Sun lights the house with an elevation of less than zero. Even for a house perched at the top of a south-facing cliff, which will be illuminated by the Sun before a house in the valley below, the lowest elevation angle of the Sun is, at best, marginally below zero.

The Earth’s radius at the equator is 6378 km, while the radius at the poles is 6357 km. The average radius is 6371 km. If \( r \) in Figure 9.18 is the radius of the Earth, and \( x \) is the height of the observer, then the furthest distance at which the observer can see a point on the Earth’s surface is along a sight line tangential to the surface passing through this point, so this tangent line is orthogonal to the radius at that point. Therefore, \( \cos \theta = \frac{r}{r+x} \) and \( \sin \theta = \frac{y}{r+x} \). Since \( x \neq r \), we solve for \( y \) by squaring and adding as
\[ \cos^2 r \sin^2 r = (r+x)^2 + (yr+x)^2, \text{ so } 1 = 2(r+x)^2 + 2y(r+x)^2, \text{ giving } y = 2rx + x^2 \]

\[ y^2 \times 6371100x = 127.42x(9.9) \]

where \( x \) is in meters and \( y \) is in kilometers. Thus, an observer at height 10 m can see an object 35.7 km away, assuming a perfectly flat terrain. The square root in Equation 9.9 means that to double the visibility distance requires the observer to be four times as tall. Alternatively, in Imperial units,

\[ y \times 79135280x = 1.4987x(9.10) \]
where $x$ is in feet and $y$ is in miles. If $x = 6$, then $y = 1.4987 \times 6 = 3.00$ miles, and to see to 6 miles requires the observer to be at 24 feet.

The rate of rise or fall in elevation of the Sun is greatest at dawn and dusk. At latitudes $50^\circ$ and $40^\circ$ the Sun is $3.42^\circ$ and $5.5^\circ$ above the horizon, respectively, at half an hour after sunrise or before sunset at winter solstice. Also, since cloudiness is higher at dawn and dusk, compared to cloudiness nearer the middle of the day, and since the azimuth angles are greatest near dawn and dusk such that the reflectance from south-facing glazing is enhanced, it can be assumed that the Sun’s energy at elevations less than $5^\circ$ is negligible. Thus, landscaping should seek to maximize the collection of solar energy for elevations above $5^\circ$ and neglect any collection for lower elevation angles.

The first requirement for any residence to capture solar energy is exposure to the Sun, meaning southern exposure in northern latitudes and northern exposure in southern latitudes. Again, with apologies to those in southern latitudes, we avoid a tedious repetition by writing for northern latitudes only.

South-facing slopes are desirable locations since solar exposure is considerably more likely to occur there than on a north-facing slope. Even slopes to the east and west could be problematic.

One of the biggest mistakes in solarium design is to ignore the summer heat gain by glazed areas. Glass on a south-facing roof is particularly difficult. Vertical glass is much less of a problem when it is facing due south. The best for minimal summer gain is vertical south-facing glass with an overhang; the geometry needed for this overhang is discussed in Section III of Chapter 11.

## 9.8 Design With Nature

Rachel Carson’s *Silent Spring*, published in 1962, awoke the nation to an impending crisis. There were a few other voices calling in the wilderness, including Barry Commoner, Ralph Nader, and Paul Ehrlich. We were unaware of global warming caused by the profligate burning of fossil fuels. At that time, only the Audubon Society, whose focus then was on birds, and the Sierra Club, then concerned with the scenic west, joined the Conservation Foundation in having any focus on ecology. In 1967 Ian McHarg, 1920–2001, was commissioned by Russell Train, President of the Conservation Foundation, to write a book linking ecology and planning. The book, *Design With Nature*, was a sensation; the book is still in print.

The great philosopher of science and technology, Lewis Mumford (1895–1990), in the introduction to the twenty-fifth anniversary edition, wrote: “Design With Nature is a notable addition to a handful of important texts that begin, at least in Western tradition, with Hippocrates’ famous medical work on Airs, Waters and Places: the first public recognition that
man’s life, in sickness and in health, is bound up with the forces of nature, and that nature, so far from being opposed and conquered, must rather be treated as an ally and friend, whose ways must be understood, and whose counsel must be respected.”

McHarg describes the value and endurance of the simple dune grass. It is the best defense against ocean erosion on a sandy shore, and far better than concrete barriers for at least two reasons — aesthetics and utility. Grasses survive and stabilize where concrete collapses. The arguments he presents, augmented with diagrams and photographs, is compelling.

Seawalls and groins block the incoming wave and throw its energy back against the next wave, with the unintended consequence of augmenting the energy in the following wave. McHarg explained how beaches are moving entities, sometimes augmenting one section while attenuating its upwind neighbor.

The overall thesis of McHarg’s book considers how we interact with our environment, so it makes sense to question what we are doing and how to make our footprint beneficial rather than detrimental. It is in this context that siting and landscaping considerations should be paramount before one begins construction of a residence.

Dubbed the Ash Wednesday Storm due to the date of its arrival, the 72-hour-long storm of 1962 battered Long Beach Island, New Jersey, known locally as LBI. It occurred at the height of spring tide, which augmented its destructive force. At the time it was the worst storm to hit LBI in history, but this is no longer true. One beneficial side effect was the change in public attitude to building on a sand dune overlooking a beach — before OK, after bad.

McHarg described: “Houses are built upon dunes, grasses destroyed, dunes breached for beach access and housing . . . . From the fifth to the eighth of March, 1962, there came retribution. A violent storm lashed the entire northeast coast from Georgia to Long Island. For three days sixty-mile-an-hour winds whipped the high spring tides across a thousand miles of ocean . . . . Three days of storm had produced eighty million dollars\(^1\) worth of damage, twenty-four hundred houses destroyed or damaged or damaged beyond repair, eighty-three hundred houses partially damaged.”

Superstorm Sandy hit the New Jersey and Long Island, New York, shores on October 12, 2012, causing an estimated $68 billion in damage. Much of that damage was the result of human interference with the natural environment, much of it in the early and middle years of the twentieth century. Now, by and large, we know better.

“In the nineteen-sixties and seventies, coastal-engineering philosophy shifted, as reflected in ‘Design with Nature’ . . . . The new thinking advocated protecting coastlines with less invasive ‘soft structures,’ such as beaches and dunes. Even the word that engineers used for beach building, ‘nourishment,’ suggested a holistic, almost maternal concern for a beach’s well-being. Unlike a seawall, a beach absorbs wave energy as the water sinks into the sand, and a dune spreads and dissipates the wave along its hips and shoulders rather than flinging it back
against the next wave. Of course, the hard structures built by previous generations of coastal engineers remained in place, which meant that the artificial beaches were likely to erode more rapidly, and would have to be continually replaced. But with so much of New Jersey’s economy based on beach tourism, political leaders had no choice but to persist in this Sisyphean task” [J. Seabrook, “The Beach Builders; Can the Jersey Shore be Saved,” The New Yorker, July 22, 1013].

Europe is more progressive with respect to ecology than the United States; 2013 they need to be since they have a far denser population density and have experienced environmental degradation due to that and industrial despoiling. The United Kingdom, with the Town and Country Planning Act of 1947, has sought to balance economic development and environmental quality: Particular concerns were urban sprawl and pollution. It helped preserve many of the traditional villages by limiting housing developments around them while preserving the core functionalities, such as post office, library, and school. A person, particularly an elderly person, can walk to all amenities, and a stroll of a few hundred yards bring them to the country. Yes, there are ugly “estates” with cookie cutter houses and little consideration for environment or aesthetics, but the villages are largely preserved.

Another feature of the English and Welsh countryside is the 225,000 km of public footpaths. These footpaths are part of the Queen’s highway, and are afforded the same protection in law as other highways. Hikers obtain an Ordinance Survey map showing features in extraordinary detail, including public footpaths and royal mail boxes, as well as individual houses. Covering an area of about 130×130 km², the map makes it possible to pinpoint a feature within 100 to 200 meters. These footpaths often cross a farmer’s field, and if a farmer tries to close access over cropland he or she will be brought up before the “local bill,” the magistrate presiding over that district. However, public foot paths are for walkers only, and riding a bicycle on one is committing trespass and the same “local bill” would impose sanctions if a civil case was brought by the landowner. However, it is not that the Brits are down in cyclists as there is a network of bridle paths that permit horses and bicycles.

The Town and Country Planning Act permits its individual counties to develop their own planning system while preserving their core principles. Greenbelt legislation was added in 1955. In the U.K. a property owner cannot unilaterally develop property. The act effectively nationalized the right to develop land. Local authorities are required to create maps of the land under their aegis, indicating those areas that could be developed and those areas that cannot be developed, and imposing restrictions on those in between. It is a system that is largely accepted, with a little grumbling, since it is seen to provide benefits that outweigh its strictures.

The post-World War II suburban and rural development in the United States has followed a vastly different pattern. Local planning boards require houses on large lots of land with no regard for many of the core requirements of a society. As a result, an automobile is essential and it is usually impossible to walk across open land as can be done in England and Wales.
Some communities have introduced cluster zoning, and a few towns, such as Lincoln, Massachusetts, have done it well. In other towns with planning boards possessing less foresight, developers have been permitted to build grid subdivisions on smaller lots than are otherwise permitted with no concept of the benefits of a properly designed cluster with adjacent open land.

Table 9.7: Noon-Time Solar Elevation

<table>
<thead>
<tr>
<th>latitude</th>
<th>elevation</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>36.55</td>
</tr>
<tr>
<td>35</td>
<td>31.55</td>
</tr>
<tr>
<td>40</td>
<td>26.55</td>
</tr>
<tr>
<td>45</td>
<td>21.55</td>
</tr>
<tr>
<td>50</td>
<td>16.55</td>
</tr>
</tbody>
</table>

9.9 Clearances on Flat Sites

The solar elevation at solar noon at winter solstice as derived in Section VIII of Chapter 5 is $90 - L - 23.45$. The elevations as a function of latitude are given in Table 9.7.

At latitude 42.33°, the latitude of Boston, the winter solstace elevation of the Sun at noon is 23.8°, while the equinox elevation is 47.66°. We assume that the solar collectors must be free from obstructions, as seen at ground level at elevations up to 23.8° due south and from 5° above the horizon at ±60° from south.

The azimuth when the solar elevation is 5° for significant declinations on the winter side of the year are shown in Figure 9.19 for latitudes 30 through 50. Notice that at equinox the azimuths are close to 90°, when the direct solar gain in south-facing solar collectors is zero. Thus, on the summer side of equinox, solar collection ceases at solar elevations above 5°.

The length of the solar collection day is given in Figure 9.20 for declinations from equinox to winter solstice and latitudes 30 to 50.

When a site is prepared for a house, attention should be given to the sizes to which trees and shrubs will grow. The sapling may become a beautiful tree in 15 years’ time, except it is in the wrong place. Luckily, when solar heating is needed the most, deciduous trees have no leaves and subtend a small portion of the sunlight. However, a dense grouping of deciduous trees must not block the winter Sun from the collectors. No evergreen trees must block the Sun. The mature sizes of some common trees are given in Table 9.8.
9.9.1 Individual Residences

Referring to Figure 5.27, one sees that for Boston at winter solstice the azimuth angle is about 50° when the Sun’s elevation is at 5° (near dawn and dusk), so the 60° from south condition is not invoked. For an object 10m tall, the minimum distance from the structure so as not to violate the 5° rule is 10/ tan 5 = 114.3 m. The same object due south of the structure must be 10/tan 23.8 = 22.6 m away from the structure.
The necessary clearance area from objects 3m tall is shown in Figure 9.21. The arcs are 35m in radius centered at the southeast and southwest corners of the house. The points of the clearance area are rounded to produce a clearance rectangle on the south side of the 8×12m structure of 65×19m; this results in minimal loss of solar energy while reducing the clearance area moderately.

Table 9.8: Height of Trees

<table>
<thead>
<tr>
<th>tree</th>
<th>mature size</th>
</tr>
</thead>
<tbody>
<tr>
<td>maple</td>
<td>15–23m</td>
</tr>
<tr>
<td>oak</td>
<td>18–39m</td>
</tr>
<tr>
<td>beech</td>
<td>15–23m</td>
</tr>
<tr>
<td>birch</td>
<td>15–23m</td>
</tr>
<tr>
<td>poplar</td>
<td>23–30m</td>
</tr>
<tr>
<td>elm</td>
<td>23–30m</td>
</tr>
<tr>
<td>pine</td>
<td>18–30m</td>
</tr>
<tr>
<td>spruce</td>
<td>15–30m</td>
</tr>
</tbody>
</table>
Figure 9.21: Clearance from 3m Objects
Assuming that the house must be no closer than 3m from a lot line, the minimal lot size and shape is given in Figure 9.22. This lot approximates a 65×19 rectangle with two triangles 25×30m lopped off, leaving an area of 1275m².

For a flat site, the acceptable vegetation for a solar house which collects solar energy down to the ground level is shown in 9.23. The house footprint is assumed to be about 12×8m.

The designations are as follows:

A – unlimited mixed wood
B – shrubs to 3m
C – shrubs to 6m
D – small trees to 12m
9.9.2 Communities of Solar Homes

Clustering solar sites with lot size and shape given in Figure 9.22 would result in a grid pattern as shown in Figure 9.24. As can be seen, the sites without the triangular corners do not result in a closer clustering than rectangular sites 65×19m, encompassing 1930m².

Further, since a house is directly in front of by 19m a solar house, this means that the house in front cannot be taller than 1922.510=8m.
However, other considerations, including the need for appurtenant structures such as garages, could make this lot size too small: More realistic is a rectangular lot of 65×30m. Further, a subdivision of solar houses requires land for streets, and each house must be connected, by paths or by driveway, to a street. A subdivision of solar houses with driveways, paths, garages (some attached, some detached), is shown in Figure 9.25. The lot size is now reduced to 59×30 = 1770m². Further reductions can occur by locating driveways towards the front (south side) of the structure, but caution must be exercised.

![Figure 9.24: A Cluster of Solar Houses](image)

Figure 9.24: A Cluster of Solar Houses
The problem with the close clustering of single family houses and setbacks from lot lines is that solar exposure requires limitations on adjoining lots. This can be achieved in a subdivision by deed restrictions on the placement and heights of buildings. It is more difficult to have the same restrictions on trees.

Assuming a maximum solar angle at winter solstice of 24°, the minimum distance of a structure in front of a solar house on a flat site is shown in Figure 9.26. For a single-storey structure, the minimum distance is 13m. For a two-storey structure, the distance is 19m.
9.10 Clearances on Sloping Sites

In this section we derive the angles on south-facing sites that are planar and clear of all obstructions. Suppose the maximum slope is $\lambda^\circ$ at an angle of $\sigma^\circ$ to due south as shown in Figure 9.27. South is in the x direction, and east is in the y direction, so $v_1 = [\cos \sigma \sin \sigma 0]$, which is a unit vector. To find $v_2$ in Figure 9.27 it is necessary to determine plane $q$ in which vectors $v_1$ and [001] lie. This is

$q = v_1 \times [001] = [\cos \sigma \cos \sigma 0] \times [001] = [\sin \sigma - \cos \sigma 0]$. Rotating $v_1$ by angle $\lambda$ in plane $q$ produces
Suppose an object ℓ high exists at the end of unit vector \( \mathbf{v}_2 \), and the angle at the top of that object subtends with respect to the horizontal is \( \varsigma \). Then

\[
\ell = \sin \lambda + \tan \varsigma \cos \lambda. \quad (9.12)
\]

The situation is shown in Figure 9.27. The object, be it a tree or a structure, of height \( \ell \) at distance unity from a solar collector at ground level, presents most shading problems near dawn or dusk. To avoid shading, the solar elevation \( \alpha \geq \varsigma \).

In northern latitudes, north-facing slopes present some difficulties for solar collection. Clearance distances are increased, so typically lot sizes need to be larger. South-facing slopes are much easier. As espoused by Ian McHarg [Design With Nature, 1967], flat land should be devoted to agriculture whenever possible, and the hillsides permit a tighter clustering of houses while still maintaining privacy. Now, in this era of solar energy, south-facing slopes are most desirable.

In some states there are laws to encourage particular uses. For example, Chapter 40B in Massachusetts encourages low-income housing by permitting developers to avoid compliance with certain local ordinances. Sometimes these laws are clumsy and cause undesirable results, but often they result in the necessary incentive for public good.

Some would argue that leaving landowners alone to decide for themselves and allow market forces to prevail is the only way to go, but experience shows otherwise. A developer can take a farm and split it up into as tight a grid of houses as zoning will permit, and this is done regardless of the terrain.

South-facing slopes present another advantage. With a solar house on a south-facing slope, the south-facing windows have a height advantage over structures to the south, which not only
affords a better view than that from a flat site, it also increases the feeling of size and associated privacy. Two solar houses on a south-facing slope are shown in Figure 9.28.

A house exposed to high winds will lose significantly more heat in winter than a well-protected house. Winds impacting a house can be considerably reduced using wind breaks, which can be natural or designed.

A windbreak can be a stand of trees or shrubs. Consider a stand of three or more rows of trees or shrubs at uniform height $H$. Within the envelope on the downwind side, which reaches out distance $4H$ as shown in Figure 9.30, the wind velocity is reduced by at least 60% The stand does not need to stop all air flow though it. The objective is substantial reduction, not total subjugation. The windbreak should extend at least 30m on either side of the site to account for

9.11 Wind Protection
air turbulence at the end of the windbreak [as shown in csfs.colostate.edu/windbreakdesign, as accessed June 8, 2011].

Improperly designed windbreaks can funnel and concentrate the wind. For example, a new building on the MIT campus in Cambridge, Massachusetts, funneled the wind to such an extent that even athletic males were swept off their feet; the solution was a windbreak within that funnel designed as a piece of sculpture. Paths through windbreaks should not align with prevailing winds, and any path should not provide a straight line for the wind.

Does an effective windbreak necessarily create solar shading on the solar elements of the structure? The answer is no. First, the windbreak is most effective on the windward side of the house, and less so on the leeward side, but can still be effective. Windbreaks on the southeast and/or southwest sides of the house will be problematic, but less so to the south. The house I built for my daughter’s family, described in Section III of Chapter 10, has tall trees as windbreaks on the east, west, and north of the house. The house itself is on the crest of a strong south-facing slope with a clear view to the distant horizon. It has no windbreak solar shading.

\footnote{It would be many times this in today’s dollars.}
Chapter 10

Solar Structures from Early to Modern

There are, according to some estimates, 100,000 solar structures in the world. Regardless of the actual number, it is clear that the number will increase rapidly. In this chapter we attempt to document some of the earliest solar structures and identify what is effective in the design of an energy efficient structure.

The previous chapter considers the mechanisms and databases used to determine heating and cooling loads. These provide a basis to guide us in the design of a solar-sensitive structure, and possibly go further in designing a house that uses the Sun to reduce or even eliminate its use of fossil fuels.

Some of the early solar structures had modest expectations for solar heating. More often the Sun was for daylighting, while very little consideration was given to heat. A burst of enthusiasm occurred in the early 1900s, with installations containing flat plate solar collectors as well as passive solar. The energy crises of the 1970s and U.S. federal tax credits for solar installations provided a second burst of activity in solar heated houses, both passive solar and/or active solar, mainly using flat plate collectors. Unfortunately, a change in administration led to the ending of tax credits and diminished the appetite for domestic solar heating systems.

Some consider tax credits for renewable energy with disfavor. “If it cannot make it on its own it is not worth doing” goes the thinking. There are three strong arguments against this mind set:
* The fossil fuel industry enjoys substantial tax subsidies in spite of enormous profits, so why should they be given an unfair advantage over renewables?
* The real cost of fossil fuels (which includes global warming, pollution, and adverse health effects) is substantially greater than the cost of fuel at the pump.
* The proven reduction in greenhouse gas emissions by 17% from 2005 to 2014 in the United States is due in large part to the growth of renewables, in particular PV panels and wind turbines, helped in part by the $90-billion stimulus in the first few months of the Obama administration.

The importance of the principles behind the passivhaus design require a separate section, which completes this chapter. The large number of passivhauses built in Europe is evidence of
their value. The United States is lagging in the adoption of these principles and the embrace of the simple, effective, and economic use of passive solar heating.

It can be argued that the most important component for the majority of solar collection systems is the availability of affordable sheet glass. The development of durable double glazing has made passive solar houses and efficient flat plate solar collectors possible.

10.1 Double Glazing

The ability for a passive solar heated structure to succeed depends to a major extent in the multiple glazing system. Single glazing is unacceptable. It is double or triple glazing that makes a passive solar structure possible.

The history of glass was discussed in Section I of Chapter 8. For persons of means, and with the use of fossil-fueled furnaces, heat losses from a structure became less important than an architectural statement, and glass was often an essential part of that statement. In particular, the picture window of fixed glass with an area of 20 ft\(^2\) or greater, usually surrounded by operable windows, became popular. One problem with the picture window was that it was a single sheet, and this led to three major problems:

* as discussed in Chapter 3, such windows lose a lot of energy to the outside in cold weather,
* single panes result in condensation on the inside of the glass in cold weather, and
* a picture window inhibits the use of a storm sash.

Storm sashes are hung over the window at the beginning of the heating season, and so act as a form of double glazing, with some serious disadvantages. Rarely does a storm sash fit tightly, with the result that there is considerable air flow between the sash and the window. Also, dirt and dust trapped between the sash and window typically must be tolerated all winter until the sash is taken down, washed, and stored until the next winter.

Double glazing has made solar heating possible in the temperate zones, predominantly between the latitudes 30 to 50. It was invented by Charles D. Haven in the 1930s. He started work on it in 1930 and termed it Thermopane. The failure of the bond between the spacer and the glass was a problem that he was unable to solve without adequate financing. In 1934 Libbey-Owens-Ford bought him out, and retained him to perfect the seal. In 1937 he succeeded using a alloy of aluminum, titanium, and copper as the seal. This Thermopane withstood extreme environmental testing, and the product came on the market in November 1937. Haven was awarded patent 2,138,164 in 1938, which was assigned to the Thermopane Company.
The construction of a double glazed panel is shown in Figure 10.1. The metal spacer is a hollow, rectangular structure of an aluminum alloy in which holes are drilled on one side, which will face the interior air gap. The interior of the spacer is filled with a desiccant such as silica gel or zeolite; a desiccant absorbs water vapor that could mist the interior surfaces of the glass panels. The sealant between the spacer and the glass is polysulphide or silicon.

10.2 Historic Use of Solar Heating

Early human uses of solar energy appear to involve concentrating lenses and plane mirrors. In 1515 Leonardo da Vinci made sketches of a solar concentrator using mirrors — he failed to reduce it to practice. In 1747 Comte de Buffon demonstrated in Paris a solar concentrator using 168 mobile mirrors, each 20cm x 20cm. The chemist Antoine Laurent de Lavoisier (1743–1794) used convex lenses to concentrate solar energy to melt metal. The present day interest in multiple mirrors to concentrate solar energy on a relatively small target is in solar towers: these are tall towers with thousands of heliostats focusing the Sun’s energy to the top of the tower.

Dennis R. Holloway [www.dennisr海湾architect.com/SimpleDesignMethodology] documented one of the oldest houses in the world, built between 24,000 BC and 12,000 BC in Mezhirich, Ukraine of mammoth bones and tusks. “The massive bones acting as thermal mass, were covered with insulating layers of hides. The south-facing opening, covered in stretched skins allowed low level solar energy to temper the interior.”

As reported by John Perlin of the University of California, Santa Barbara [http://cecsb.org/index.php/blog/item/solar-architecture-in-ancient-china], “Six thousand years ago, Neolithic Chinese villagers had the sole opening of their homes face south. They did this to catch the rays of the low winter sun to help warm the indoors. The overhanging thatched roof kept the high summer sun off the houses throughout the day so those inside would stay cool.”

Barry Stevens [barryonenergy.wordpress.com/2011/06/19/solar-power] described a third century BC use of solar energy by Egyptians: “According to Ecomall.com, ‘the ancient Egyptian Pharaohs solar heated their palaces by capturing solar energy in black pools of water by day and draining the hot water into pipes in the floor of the palaces at night. This kind of architecture heated homes at night while keeping the temperature low during the day.’”
The early Romans were the preeminent engineers of their epoch. The Greeks gave us the best philosophers, but the Romans gave us the arch, aqueducts, and their famous bathhouses. Many of these bathhouses had openings to the south to let in both light and heat.

“Baths were especially popular among the Romans but demanded a great amount of heat. From the times of the early empire onward, most faced the afternoon sun in wintertime when they had maximum use. They also had their large windows covered with either transparent stone like mica or clear glass, a Roman invention of the 1st century AD, one of the great breakthroughs in building and solar technology. Transparent glass, the Romans discovered, acts as a solar heat trap, admitting sunlight into the desired space and holding in the heat so it accumulates inside. Recent experiments show that on a clear winter day in Rome the sun entering a properly oriented, glass-covered bath would keep the temperatures inside above 100 degrees Fahrenheit throughout the day, the desired temperature in the caldarium, the hot bath [www.motherearthnews.com/renewable-energy/solar-energy-in-ancient-rome].”

“An often cited early example of solar design awareness is the ‘Megaron House’ described by Socrates in the year 400 BC. Numerous other examples can be found, i.e., the New England ‘salt box’ of the seventeenth century or Swiss farm houses of the eighteenth century. In the twentieth century, the term ‘solar house’ became popular and following the first oil shock of 1973, the term ‘passive solar buildings’ was coined. In all these examples, the basic principles are the same; maximize the south exposure of a building to capture as much solar heat as possible and insulate the enclosure to keep the heat in [www.springerreference.com/docs/html/chapterdbid, as accessed December 7, 2013].” Megaron means “big room,” with the big room facing south.

By passive solar we mean capturing solar energy through south-facing windows. We all know how rapidly the interior of a parked automobile will heat up in direct sunlight. So does a house with windows to the south on a sunny day in winter. Thermal mass is needed to store that heat for the night-time hours — more on this in the next chapter.

Socrates is reputed to have said: “Doesn’t the sun shine into houses facing south in winter, whereas in summer the Sun wanders over us and the roof so that we have shade? Because this is comfortable, then south-oriented rooms should be built higher in order not to shut out the Sun, whereas the north rooms should be lower because of the cold north wind [op cit].” These principles are applied to the Megaron House, as shown in Figure 10.2.
The Justinian code is “the collections of laws and legal interpretations developed under the sponsorship of the Byzantine emperor Justinian I, AD 529 to 565. Strictly speaking, the works did not constitute a new legal code. Rather, Justinian’s committees provided basically two reference works containing collections of past laws and extracts of the opinions of the great Roman jurists. Also included were an elementary outline of the law and a collection of Justinian’s own new laws [www.britannica.com].” This code provided the legal foundation for modern Europe. Within this code there are “sun rights” to the owners of sunrooms on houses and public buildings to ensure access to the Sun.

When Rome fell we entered the Dark Ages for almost a thousand years. The lessons taught us by the Greeks and the Romans about solar sensitivity were forgotten. European cities were rarely planned — they just happened. Contrast this to China in the same period, where south meant warmth and health, so buildings were designed to welcome the Sun.

The Anasazi people in the American west lived in cliff dwellings, and many of these were south-facing to take advantage of the winter Sun. Cliff dwelling had another advantage — security from marauders. The word anasazi may have derived from “ancient outsiders” or “ancient ones.” They live in the four corners area, named since four states meet at a single point, of the American southwest. They built substantial dwellings in the cliffs, and seemed to
prefer substantial overhangs to shield them from the summer Sun. They thrived from about 700 to about 1300 AD, and then abruptly disappeared.

They had no need to climb down the cliffs to get water. They channeled and captured water as it trickled down near their dwellings. There was a drought around 1300 that could have forced them to move, or it could have been caused by the decline of some other resource.

Indigenous peoples were often ingenious peoples. They had to be to survive. We are fortunate that such a need rarely occurs today. However, we have too often ignored our environment, with catastrophic results: Ian McHarg gave the example of building on and so de-stabilizing the protective sand dunes along the New Jersey shore, as discussed in the prior chapter.

Even in cases with no direct adverse effect on our environment, all too often we have had no solar sensitivity. Why face a house away from the Sun?

Urban blight in Europe intensified with industrialization that began in northern England. Great wealth aggregated to the factory owners, and poverty to the masses. This was so eloquently described in the works of Charles Dickens. Overcrowded conditions and lack of sanitation led to outbreaks of tuberculosis, cholera, typhoid fever, and smallpox. Efforts to relieve this misery began in the nineteenth century, including adequate sanitation and adequate space and light in the living quarters, which naturally led to some measure of solar sensitivity.

The wealthy Europeans used conservatories more as trophies than as places to grow plants. Plants were grown in greenhouses and tended to by the common man. The notion of using glass houses for heat was not considered. In essence, using solar energy for warmth began in Europe and in the New World in the twentieth century.

10.3 Solar Heating for Residences in the United States

In the United States, central heating became the standard in the early years of the twentieth century, permitting architects the opportunity to use large panes of glass with little consideration of heat loss; after all, the fossil fuel-fired furnace has virtually unlimited capacity. Glass was used for natural lighting and aesthetics.

Between the Great Wars, there was an acute housing shortage in Germany [K. Butti and J. Perlin, A Golden Thread: 2500 Years of Solar Architecture and Technology, Van Nostrand Reinhold, New York, 1980]. The Zeilenbau (row-house) plan was conceived. The main axis of the row houses, called terraced houses in Britain, was oriented north-south, so each unit had windows that faced east or windows that faced west. This was meant to maximize sunlight. The problem is in the latitudes of Germany in winter the Sun rises far south of east and sets far south of west. Light enters, but little solar energy enters each dwelling at any time of day.
The Hugo Haring house was built in the early 1930s. Its single storey design, elongated along the east-west axis, became an emblematic style. The windows on the south side were floor to ceiling. The windows on the north side were small and set high in the wall.

Following his experience with an all-glass structure he designed for the 1932 Chicago World Fair, the architect George Keck designed a house for his friend, the real estate developer Howard Sloan. The long south side of the house was all glass, with roof overhang to shade the glass from the summer Sun. Sloan was proud of his house, and set out to promote it. He is credited with naming it a “solar” house [K. Butti and J. Perlin, *A Golden Thread: 2500 Years of Solar Architecture and Technology*, Van Nostrand Reinhold, New York, 1980]. The name stuck.

Sloan built a housing development in the Glenview district of Chicago. Half the houses were solar, and the other half were conventional. The solar houses far outsold the conventional
houses, so his next development was all solar.

Illinois architect George Keck built a number of solar structures in the late 1930s into the 1940s using large south-facing glass, first with single glazing and later with the new thermopane glass. Shown in Figure 10.4 is a Keck-designed solar house, built near Chicago, Illinois in 1939, and now known as the Sloan house.

In about 1941, George Keck built another solar house for Dr. and Mrs. Hugh Duncan in Flossmore, Illinois in 1947; it was considered ground breaking. A team from the Illinois Institute of Technology tested it over the winter of 1941–1942.

![Figure 10.5: The Duncan House](image)

MIT (Massachusetts Institute of Technology) built six solar heated structures. The first was underwritten to the tune of $650,000 by wealthy Bostonian Godfrey Lowell Cabot in 1938. The house was built in 1940 and has 41 m² of drain-back flat plate collectors feeding a 66 m³ water storage tank; it was probably the first solar heated house in the U.S. It was more of a laboratory than a residence. The basement housed a 17,400-gallon water tank. The tank was surrounded by insulation, while 14 flat plate collectors were mounted on the roof, which was inclined 30° to the horizontal. Three sheets of glass covered each collector. The front of the house was supposed to face due south, but the magnetic compass reading was interpreted incorrectly with the result that the house was 7° off.

Arthur T. Brown (1900–1993), was an early solar pioneer, made famous for his powers of observation and deduction. He designed a house whose client insisted that the outside be painted black. When he walked past the south side of the house on a sunny day he noticed how much heat radiated from the black wall. When he designed the Rosenberg house he used wall-to-ceiling glass on the south side so that in winter the Sun would penetrate deep into the house and heat a concrete block wall, which, of course, was painted black. The floor was concrete to
increase the thermal mass, and was insulated from below. In a sense the house was the first application of the Trombe wall concept, except that the dead space between the glass and the wall is inaccessible in a Trombe wall, but was the principle living space in Brown’s realization.

Figure 10.6: MIT Solar House 1

Figure 10.7: The Rosenberg House
Notice the large difference between the east-west axis and the north-south depth of the house. This is inefficient, as will be explored later in this chapter.

In the 1930s Professor F.W. Hutchinson of Purdue University in Lafayette, Indiana, directed the building of a solar heated house, possibly the first in the United States designed for solar heat. The heating system was composed of direct solar radiation through single glazing. When the Sun was shining, the house was warm. Otherwise, in winter, when night fell or cloud cover was present, the house was cold. It did prove that in northern locations single glazing is not satisfactory since the heat loss over a typical winter day far exceeds the solar heat gain. Of course, the windows can be draped, screened, or shuttered to reduce the night time heat loss, and with this strategy and double or triple glazing, passive solar heat can work.

Why is Professor Hutchinson important in the world of solar heating? It is partly because in 1945–1947 his team at Purdue University was commissioned by Libbey-Owens-Ford to study residential solar heating. Of course, the company was motivated to promote its Thermopane system. The study found that one can expect a large range in temperatures — hot on a sunny day, cold in the night. This indicated that a solar house should have high thermal storage.

The book *Your Solar House*, (Simon and Schuster, New York, 1947), was a compilation prepared by Maron J. Simon of designs, including plans and isometric renditions, from 49 leading architects of solar houses. Each house is designed for a geographic region, and every region of the continental United States is represented. The houses were meant to appeal to the moderate-priced market, and cost no more than $150,000 in pre-war dollars.

It appears that the dominant architectural influence over these 49 architects was Frank Lloyd Wright. One of the airy, sprawling concepts of this most favorite of American architects is shown in Figure 10.8: notice the cantilever roofs and the low ratios of height to length and depth to length.

The floor plan of a Frank Lloyd Wright house is shown in Figure 10.9. Notice the single storey, and the large length-to-depth ratio. Also he avoided an interior corridor by connecting the east and west ends of the house with a glass wall gallery, bringing the outside in.

In *Your Solar House* Delaware architects Victorine and Samuel Homsey presented a two storey solar house, shown in Figure 10.10, that is closer in design to a modern house, with a lot
of glass on the south side and limited glass elsewhere. Notice the three wing-like exterior shades above the south side glass to reduce unwanted solar penetration during the summer side of the equinoxes.

The Connecticut architect Douglas Orr proposed in *Your Solar House* the single storey solar house that is sketched in Figure 10.11. If the garage is deleted the house is 77’ long and 25’6” deep, for an aspect ratio of 3:1. Notice the long corridor from kitchen to master bedroom. Corridors are features that divide, rather than connect, and should be minimized when at all possible. Further, single storey structures need corridors more than two storey structures.

![Figure 10.9: Floor Pan of a Frank Lloyd Wright Solar House](image-url)
The American colonial era house that lends itself to solar heating is the saltbox, as shown in Figure 10.12. The house faces the winter Sun with two levels of windows, while the other side of the house is one storey high and has limited windows. This structure provides two levels of passive solar heating, and in addition can have active solar panels, either flat plate solar collectors or photovoltaic panels, on the south-facing roof.
With modern double glazing and effective blinds, the saltbox should make a comeback. It is far better in energy capture and usage than the central entrance garrison colonial.

Notice the central chimney in the saltbox house typically had several flues, one for each of the fireplaces in the rooms surrounding the large central masonry structure. These fires were kept running all winter, fueled by wood or possibly coal, not exactly the model to reduce carbon footprint.

George O.G. Löf, Director of the Solar Application Laboratory of the State University in Fort Collins, Colorado, built a house for the Saint Gobain Corporation in Denver, Colorado in 1958 using partial active solar heat. The hot air from a 600 ft$^2$ collector on the flat roof of a 2,050 ft$^2$ living area heated 11 tons of gravel in a tall column. The original efficiency was low, mainly due to the fact that the solar panels were a late addition to a conventionally heated house. Subsequent improvements raised its efficiency considerably, but nowhere near self sufficiency. George Löf subsequently received a $250,000 grant from the National Science Foundation to plan and conduct a solar heated and cooled house. The main problem with the plans was complexity — no contractor would undertake the work.

Saint Gobain Corporation built a 1,000 ft$^2$ house in Boulder, Colorado in 1947 using a 463 ft$^2$ solar air collector on the 27° roof. This collector provided 20–25% of the heating needs of the house.

Harold Hay of Phoenix, Arizona used an intriguing system for the solar heating and cooling of his 1967-built house. Shallow pools of water covered the flat roof of the one storey structure. Insulation and moveable panels on the roof exposed or covered the water for heating or cooling, partially by evaporation. The temperature was maintained at a minimum of 65°F in winter.
An MIT house was built in Lexington, Massachusetts in 1958. The hot water solar collector was 16’ × 40’ on the 60° to horizontal south-facing roof. The house was rated at needing 30,000 BTU/hour for a 70°F temperature differential (interior to exterior temperature), and the solar panels provided about half the necessary heat; a back-up furnace supplied the rest. The design was overly complicated, and as such it was impractical. Another criticism was that it used a traditional house design modified for solar use.

Boston heiress Amelia Peabody was the underwriter for the last MIT house. It was built in Dover, Massachusetts. Dr. Maria Telkes designed the solar heating system with the eutectic salt sodium sulfate decahydrate; eutectic salts are discussed further in Section 4.5 of Chapter 11. Phase-change problems with these salts caused the solar system to be abandoned after three years. None of the MIT houses are presently using solar power.

Harry E. Thomason of Washington, D.C., built his first solar heated house in the 1950s. His original house used a 840 ft² trickle hot water collector system that ran into a hot water tank surrounded by rock to increase the thermal mass. The 1,500 ft² living area was heated by passing air over the heated rock. A back-up oil furnace used less than 50 gallons per heating season. Later designs used heat pumps and had improved efficiencies. [H.E. Thomason, “Experience with Solar Houses,” Solar Energy, vol. 10 #1, pp. 17–22, Jan–March 1966].

A self-sufficient, 672 ft² living area house was constructed in Amado, Arizona by Donovan and Bliss in 1954. A screen type, woven, black cloth air solar collector of 315 ft² fed an underground bin containing 65 tons of 4” rocks. Temperatures in the bin ranged from 90°F to 140°F, and the storage capacity was 1.35×10⁶ BTU. The daily heating load was 136,000 BTU, so this house was the best of the early solar houses in providing the highest percentage of the heating need by solar means.

The Steve and Holly Baer house in Albuquerque, New Mexico, used passive solar heat with improved storage. It was constructed in the late 1960s. The southern glass wall of the house was covered with 55-gallon water-filled drums, painted black on the outside and white on the inside. Flap-down panels lie on the ground during the day with reflective surfaces to increase insolation on the drums. During the night time hours the flaps are closed to reduce the heat loss from the glass. Steve started in construction by using car roofs as panels for dome-like structures. He read Daniels’ book [F. Daniels, Direct Use of the Sun’s Energy, Yale University Press, New Haven, 1964] and it led to “zomes,” airy homes employing solar elements.

The Lawrence Gardenshire house at State College, New Mexico, built in 1953, had a 1,100 ft² living area and a 457 ft² solar air collector at a 45° angle. Two tons of glauber salt, a phase change material that has high heat capacity, was used for storage. The heat load was 50,000 BTU/hour for a 70° temperature differential; the solar system could provide 50% of this load. The glauber salt had problems with phase separation.
Glauber salt was the favorite storage medium for Dr. Marie Telkes. She used 470 ft\(^3\) with a heat capacity of \(4.7 \times 10^6\) BTU in her solar heated house in Dover, Massachusetts. The house, built in 1947, had a living area of 1,390 ft\(^2\) and a vertical air solar collector of 740 ft\(^2\). As with the Gardenshire house, phase separation occurred in the glauber salt reducing the effectiveness of the system, so auxiliary heat was required. The building is no longer solar heated.

Solar I is a solar powered house built by the University of Delaware under the direction of Dr. Karl W. Boer. The 24 air-heated panels, three of which contain cadmium sulfide electric solar cells, over an area of about 700 ft\(^2\), send heat to the \(10^6\) capacity tank containing glauber salt. The small two-bedroom house is not architecturally pleasing, looking like a lean-to shed with nothing to lean-to. A 81% heating efficiency was claimed.

Since the energy crisis of 1973, many solar houses have been constructed, both privately and by institutions. Most government funds have gone to institutions. Most of the structures are ugly and expensive. It is evident that the high heat load of these early houses, reflecting standard building practice of the times, was high in air leakage and low R values in the walls and through the roof. For example, the Gardenshire house of 1,390 ft\(^2\) leaked 50,000 BTU/hour. G.O.G. Löf and R.A. Tybout argued that 100% solar heat was not possible [Natural Resources Journal, vol. 10 #2, p. 268, 1970, and Solar Energy, vol. 14, p. 253, 1973], but their analyses were based on the leaky/lossy construction techniques that were standard fare at that time.

Modern designs would leak far less than these early attempts at 100% solar heating. Conclusion: While it is true that solar heating is easier today than when the brave pioneers experimented with it, it is also true that airtight windows, better building materials, and tighter designs have significantly reduced the heat load on modern residences. This leads to the debate: Is it better to use as much solar energy as necessary, or to reduce heat loss as much as possible? The obvious answer is both, and use a prudent combination of both as long as they are cost effective.

In 1978, Gene Leger constructed a super-insulated house in Pepperell, Massachusetts with double stud walls, and a really tight design. He claimed an annual heating bill of $38.50 without the use of solar heating.

Martin Holladay, the Green Building Advisor, in his blog “Forgotten Pioneers of Energy Efficiency” [www.greenbuildingadvisor.com/blogs/dept/musings, April 17, 2009], identified two structures as leading the pack towards energy efficiency via superinsulation. The first was the Saskatchewan Conservation House, built in 1997, with R40 walls, R60 roof, heat recovery ventilation, and triple glazed windows. The other, built in the same year, was the Gene Leger house. Both houses focused on reducing air infiltration in unexpected places. Neither of them used solar elements to any extent.

Massachusetts physicist William A. Shurcliff issued a press release in June 1979 extolling the virtues of these houses and pointing out that they were built without a large thermal mass, large
south-facing windows, or weird architecture. It seemed that there was a conceptual battle between the “maximize solar” forces and the superinsulation advocates: The solar forces won, at least for a while.

A decade later, the father of the Passivhaus, Wolfgang Feist, acknowledged the pioneering achievements shown in these two houses and the work of Shurcliff and others. The net-energy-zero houses of today are well insulated, so the heating load is small, and can be provided by simple passive, solar elements installed with simple and pleasing architecture.

A house uses a substantial amount of electrical energy, the majority of which converts to sensible heat to keep the house warm. Incandescent lighting alone in a super-tight house can keep it warm. Refrigeration was an energy hog in the 1970s, less so today with Energy Star appliances. Referring to Section II of Chapter 2, we see that a couple will dissipate between them about 200 watts of sensible heat, enough to heat a living room in a super insulated house.

The total number of partially or totally solar heated buildings, past and present, in North America as of December 1974 was less than one hundred. Many of these structures have since been demolished or converted to take conventional heating plants.

There are many more solar houses than indicated here, and many of them are net-energy-zero or near-net-zero. A net-energy-zero house provides all its energy needs from renewable sources that include solar thermal, photovoltaics, wind energy, and geothermal.

Shurcliff [W.A. Shurcliff, Super-Solar Houses: Saunders Low-Cost, 100% Solar Designs Brick House Publishing Co., Andover, Massachusetts, 1983] refers to Norman B. Saunders in his title, and freely acknowledges the extensive help he received from him. The book is about three houses, the Shrewsbury House constructed in 1981, the Cliff House constructed in 1983, and the All-Solar-Too House designed for the 1982 Gardenway Passive Solar Design Competition. The solar heating systems in all three houses were designed by Saunders.

As Shurcliff wrote in the introduction to his 1986 paper [W.A. Shurcliff, “Superinsulated Houses,” Annual Review of Energy, vol. 11, pp. 1–24, 1986], “Until the oil embargo of 1973 there was little interest in saving heat. Architects, builders, money lenders, and home buyers gave the subject little attention. Most of the existing stock of houses had little or no insulation and even in the newest houses the insulation consisted, typically, of only 3 1/2” of fiberglass. Also, houses were loosely constructed: cold air could leak in easily through cracks around windows, doors, and sills. Warm indoor air could escape equally easily. On windy days in winter, infiltration and exfiltration could account for as much as half of the house’s entire heat loss.”
The two storey passive house shown in Figure 10.13 was constructed in 2012 by the husband and wife team, Andrea and Ted Lemon. Andrea was the designer who worked with energy consultant Marc Rosenbaum. The footprint of the house is 46’×26’. The house is super insulated and uses heat recovery ventilation. The side-wall structure uses 2×6 studs, with 3/16” ZIP System sheathing, over which are two layers of 2” rigid foam. The floor of the lower level is a 4” reinforced concrete slab with two layers of 3” foam insulation over compacted fill. Similarly, the concrete foundation is insulated inside with two layers of 3” foam. The circular staircase saves space, but could be a problem in getting large pieces of furniture or a box spring upstairs, not to mention a guest with mobility problems. The crosshatched area on the upper level is an atrium, open to the lower level, a nice design feature I favor since it enables good air circulation.

The Appalachian Energy Center of the Appalachian State University has developed the Affordable Passive Solar Planbook for North Carolina [http://energy.appstate.edu/programs/70]. The small structure shown in Figure 10.14 is called the Catawba Valley plan. The single storey, passive solar design is 50’ long and 26’ deep, containing 1132 square feet.
The floor plans of the Macneil house, designed and constructed by me in Massachusetts in 2010, is shown in Figure 10.15. The foundation of the house is 50’×26’. There are five half levels in the house, and four are shown here. A lower walk-out level has a game room and the mechanical room. The four flights of stairs join the half levels in the atrium area, which is open from the dining area at elevation 0’0” to the master bedroom floor at elevation 13’4”. The atrium area is crosshatched in the plan. Operable windows connect the master bedroom and the front bedroom to the atrium, so the passive solar heat generated in these two rooms flows towards the ceiling of the atrium. The atrium also collects the passive solar heat generated on the south side of the dining and living areas.

A thermostatically controlled fan duct system in the atrium ceiling pumps heat to a rock storage system under the dining and kitchen area. A perforated manifold distributes the air along the south side of the dining area. A similar manifold on the north side of the house, in the kitchen area, sucks air out of the rock store. It is connected to the same thermostat as described previously. The rock store is approximately 24’×23’×3’, and is insulated from the foundation by 2” closed cell foamboard. The foundation is also insulated on the outside by 2” foamboard. Thus, the kitchen dining area has radiant floor heating, courtesy of the Sun.

The passive solar heating system provides part of the heat load of the house. The rest is provided by a geothermal system in the front yard, under the septic system. Four 500’ long 1” HDPE pipes are set about 7’ beneath the surface, each with its own trench. The trenches are about 100’ long, and the pipe is folded back on itself so that each trench has four parallel runs about 1’ apart. A 3-ton (36,000 BTU) heat pump in the mechanical room extracts heat from the water in the pipes and distributes it through an air handler to the house.
There was some concern about this geothermal system, since it was innovative and untested. The experts who were consulted said it would not work. The pipes should be set a minimum of 10’ apart, and while my trenches are 10’ apart the four pipes in each trench would cause chilling that would freeze the ground and disable the system. To monitor its performance, a thermometer was set on the return side of the heat pump. Since it extracts heat from the field, one would expect the return temperature to be lower than the ambient temperature at 7’ below the surface, 49°F. Since its installation, the return temperature has never dropped below 47°F, indicating that the field has more than enough capacity for the size of the heat pump.

The reason for placing it under the septic system was twofold. First, the preparation for the septic system requires the topsoil and the subsoil be removed before the infiltrator field is laid, so at least half the work involved in trenching for the geothermal system was needed to
install the septic system. The other reason was to use the water that flows out of the septic tank into the infiltrator field to moisten the material around the HDPE pipe, and so improve the thermal conductivity between the ground and the fluid in the pipe. To further enhance this affect, the septic tank was insulated on its sides and top with 2” closed cell foamboard.

In 2011 the 3.8 kW PV system was installed on the roof over the master bedroom area. It generates an average of 420 kWh/month.

There is backup heat. It is a heat exchanger in the domestic hot water tank, and that tank is heated by natural gas. It has only been used in the month of February, 2015, when the PV panels were covered with snow and ice, so the power generated that month was only 6 kWh, and the amount of passive solar heat was measly. The gas bill for the month was $140, while the usual bill, for hot water and cooking, amounts to about $30/month. It should be pointed out that the geothermal system provides two additional benefits beyond space heating, and these are preheating the domestic hot water and providing air conditioning in summer — the system reverses, so that in summer the heat pump puts heat back into the ground.

The Macneil house was not constructed to Passivhaus standards. The walls are 2 x 6 with fiberglass insulation. There is effectively 14” of fiberglass in the attic floor and the roof. The house was stick built, but extra care was taken to reduce infiltration losses. The natural infiltration is such that HRV is not required; the house is about 2800 ft², far larger than the typical Passivhaus. Innumerable cans of Great Stuff were used, and polyethelene sheeting was used on the side walls. The result is an energy efficient house, but not a net-energy-zero house.

The 1952 President’s Materials Policy Commission on Resources for Freedom, in a report entitled “Promise of Technology,” predicted the sale of solar heating units to reach 13 million by 1975, costing between $2,000–$3,000 per installation. Obviously, that prediction did not pan out.

It is more difficult to efficiently heat a single family residence than a large commercial building as a function of the number of occupants. The heat loss from a building is mainly through its skin, and the skin area per occupant is far greater in a detached residence than with a large apartment building.

If a society employs a building technique or design over many generations without significant change, then that architectural feature is suitable for that climate. For example, houses in arid regions can be made out of mud whose high thermal mass reduces daily temperature variations better than concrete and at a fraction of the cost. A Cape Cod style house is suitable for New England, but not Florida. A flat roof works in Florida but not Seattle.

Function and form are connected, and can be complimentary. Some would establish the form and leave the function for a later consideration. This is wrong and leads to inadequate performance. This often happens in the design and construction of large office buildings, when there is little possibility of temperature or ventilation adjustment for individual cubicles.
Buildings should be synthesized — designed from the parameters controlling the total internal environment.

Most early solar heated houses were not particularly successful. The reason was not the failure of the methods and techniques to capture the Sun’s energy. Rather, the problem was the excessive heat losses from the structure due to poor building components and standards. The successful solar structures were tightly constructed and well insulated.

10.4 Solar Heating for Residences in the Rest of the World

Sales of solar collectors in the European and Middle Eastern geographic regions were dominated until the early 1990s by countries such as Israel, Greece, and Cyprus. Installed collectors were about 1m$^2$ per inhabitant in Cyprus, but only 0.02m$^2$ per inhabitant in the European Union [www.ebhe.gr/library/ESIF_Info_Booklet_eng.pdf, as accessed on May 18, 2011]. Since then the northern European market has blossomed, with cloudy Germany being the greatest market. The market has been dominated by water heating. Since individuals, spurred by environmental concerns, were the primary clients, the systems tended to be small and at a relatively low cost.

“As a sector, heating and cooling remains the largest contributor to the final energy demand in 2050. The renewable heating and cooling market, comprising residential and industrial biomass as well as solar thermal and geothermal applications, is predicted to take off fast. Together, they represent approximately 21% and 45% of the total final energy consumption in 2030 and 2050 respectively [‘RE-thinking 2050,” European Renewable Energy Council, 2011].”

Heating and cooling residential and industrial buildings present great potential for energy efficiencies. “In 2020 over 25% of heat consumed in the European Union could be generated with renewable energy technologies . . . . By 2030 renewable heating and cooling technologies could supply over half the heat used in Europe [Common Vision for the Renewable Heating & Cooling Sector in Europe, Renewable Heating and Cooling European Technology Platform (RHC–Platform), 2011].”

The Green Building Store’s construction division built the first cavity wall passivhaus in the U.K. in 2010 for a cost of £141K. It was also one of the first certified passivhauses constructed in the U.K. The plan of the two storey, 118 m$^2$, three bedroom house is shown in Figure 10.16 [www.greenbuildingstore.co.uk/page–denby-dale-passivhaus].

“Solar thermal energy, together with biomass and geothermal energy, is a major source of heating and cooling in Europe. It is an extremely convenient source of heating, based on a simple concept enhanced by cutting edge technology. Thanks to technological progress solar
thermal has become not only a better option for more traditional applications, such as domestic hot water production, but also an attractive alternative for new and more advanced applications such as industrial process heat . . . The ‘Solar Active House’ concept is very promising, since it meets the ‘Nearly-Zero Energy Building’ requirements that will become compulsory in the EU in 2020. The R&D priorities aim to reduce costs, increase the solar fraction per building and to improve the reliability of solar thermal systems” [“Strategic Research Priorities for Solar Thermal Technology,” http://www.rhc-platform.org/fileadmin/Publications/Solar_Thermal_SRP_single_page.pdf].

It is not clear how the term “solar fraction” is defined in the prior paragraph. We define solar fraction as the amount of solar energy striking a unit surface compared to a unit surface normal to the incident radiation. Solar fraction will be used extensively in the next chapter.

The 1950s in Europe and United States was the era of low energy costs, and the promise, not fulfilled, of incredibly low-cost nuclear powered electricity. In Asia, energy prices did not plummet as in North America and the solar heater industry prospered. The simple glass-covered box with water tanks of the turn of the century was highly successful as were also simple plastic bag collectors resembling an air mattress but water filled.

By 1969 over 4,000,000 solar collectors were installed in Japan. With introduction of super tankers and falling oil prices, the solar industry in Japan suffered a setback, just as it did in California earlier.

Because of the Middle East oil embargo of 1972, the solar industry enjoyed a revival. In 1979 the second oil shock again helped solar. In Japan more than 10 million solar powered water heaters have been installed.
After a slow start, Britain has become a powerhouse in renewables. The BedZED (Beddington Zero Energy Development) with a number of houses is one example. After construction it was found that 60% of the renewable features could have been installed at 20% of the cost; this should be considered a valuable learning experience, and a clear justification for the “building integrated” principle. The salient features of BedZed include:

* passive solar heating
* PV systems
* communal biomass CHP (Combined Heat and Power) plant
* low infiltration rates with HRV

In his excellent doctoral thesis, Alex Glass [A. Glass, “Analysis of Building-Integrated Renewable Energy Systems in Modern UK Homes,” Doctor of Engineering Thesis, University of Manchester, 2010] identified some net-zero or near-net-zero houses and gave their U-factors and energy usage. The “Alderney” detached house with ground floor of 56.4m$^2$, second floor of 47.8m$^2$, and volume of 25.4m$^2$ has the annual energy requirements of 5278 kWh, and hot water 3500 kWh. It has 2.5m$^2$ of flat plate solar thermal and 400W of a micro wind turbine. It is designated “A” in Table 10.1. The “Buckingham” detached house is “B.” The “Edinburgh” detached house is “E.” The “Palmerston” semi-detached house is “P.” The “Washington” semi-detached house is “W.” All have thermostats set at 21°C.

In Table 10.1 the numbers are conductivity in W/m$^2$·K: This is very low by U.S. standards. The wall conductivity of 0.28 W/m$^2$·K or resistivity of 3.57 m$^2$·K/W (20.3 ft$^2$·h·°F/BTU) is considerably better than required in the United States.

constructed near the city of Celle, 30km north of Hanover, the 3Litre houses in Germany have the low annual energy demand of 34 kWh/m$^2$ for the sum of the space heating and the ventilation requirements. They get their name because 34 kWh is the heat content of 3 liters of fuel oil. The houses are prefabricated with lightweight timber construction. The sidewall has studs with 20 cm of fiberglass plus 8.5 cm of polystyrene board on the outside. The inside walls are plaster on the ground floor and limber on the second. The wall has conductivity $U = 0.15$ W/m$^2$·K, or resistivity $6.66$ m$^2$·K/W (38.86 ft$^2$·h·°F/BTU). The roof has 24 cm fiberglass and $U = 0.18$. The concrete slab is laid on top of 12-cm rigid foam insulation with $U = 0.31$. Polyethelene sheet covers the interior walls before the plaster board of limber is affixed. The

<table>
<thead>
<tr>
<th>U-factors</th>
<th>A</th>
<th>B</th>
<th>E</th>
<th>P</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>wall</td>
<td>0.28</td>
<td>0.28</td>
<td>0.28</td>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>roof</td>
<td>0.16</td>
<td>0.18</td>
<td>0.16</td>
<td>0.22</td>
<td>0.16</td>
</tr>
<tr>
<td>ground</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>window</td>
<td>2.57</td>
<td>1.55</td>
<td>1.59</td>
<td>1.58</td>
<td>1.59</td>
</tr>
<tr>
<td>doors</td>
<td>1.59</td>
<td>1.38</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
</tr>
</tbody>
</table>
joints in the roofing membrane sheets are hermatically sealed. The energy demands are heating 26.9 kWh/m², ventilation with heat recovery ventilation 6.6 kWh/m² (these add up to 33.5, qualifying them for the 3Litre house designation), and hot water 32.8 kWh/m².

10.5 Efficient Building Envelopes

We have criticized some of the designs for solar houses as presented in Your Solar House. Too long. One storey. Space dividing corridors. But we liked the saltbox design. In this section some attempt is made to define an efficient design and use this as a basis for a solar heated structure.

The only shape that prevents a manhole cover from falling through is circular. The circular manhole cover is both efficient and effective. The rectangular door is also efficient and effective. The rectangular joist or rafter is also efficient and effective — it is easy to cut from a log, and easy to plane to a standard dimension. Manufactured wooden structural members can be made in the shape of a capital I to save on material and so save on cost. The rolled steel I beam is even more efficient in material.

What is the most efficient envelope for a dwelling? The scallop seems to have solved that one, carrying its house on its back. So has the turtle. Until recent history, the purpose of clothing was for warmth and protection: now clothing is for modesty and for adornment as well as for comfort.

The heat lost from a dwelling is a function of the surface area exposed to the environment as well as the thermal properties of the exterior walls, windows, and roof. Reducing the surface
area of the structure and increasing the thermal resistance of the walls are the keys. Thermal resistance was discussed in detail in Chapter 3, and need not be revisited here.

The Anastazi people had the potential for the most efficient dwelling envelope, the cave, but they lacked the one material that could make it happen — glass. A cave dwelling with a glass face is shown in Figure 10.18. Notice the overhang that can exclude the summer Sun but collect energy from the winter Sun over the complete glass face. Also, although the exposed rock can vary in temperature from −18°C to 40°C plus, the rock at a depth of 2m or more is at a constant year-round temperature of about 10°C. This means that the temperature differential for comfort is only 11°C. Further, human activities inside the cave raise the temperature by about 2°C, so the solar heat needed is minimal.

![Figure 10.18: The Cave Dwelling](image)

Not many people would want to live in a cave. The Megaron House discussed in Section II of this chapter is an improvement over the cave, and a partially underground house can have an open, airy feel. A post and beam structure that is bermed on three sides is shown in Figure 10.19. The arrows show how the Sun’s rays penetrate through the predominantly glass wall on the south side. Three of the four walls of the lower level are poured concrete, and these are underground. Further, the concrete should be insulated on the outside with closed cell foamboard: the exposed concrete on the inside provides thermal mass. To reduce the expense of bringing in material for the berm, it is best to construct this house on a south-facing slope. This design is energy efficient.

Terraced structures, called townhouses in more affluent districts, provide an efficient building envelope, made even more so if the string of dwelling units lies on an east-west axis so the street fronts face south. The main reason is that, with the exception of the end units, the immediate neighbors on the east and west sides insulate these sides from heat loss in winter. Provided the walls between units are built so as to prevent sound transmission, terraced houses provide more privacy than can be found in many housing estates of single family or semi-detached houses.
10.6 The Passivehaus Design

Developed in Germany, and now popular in Austria and Scandinavia as well, the Passivehaus designs have proved to be effective, efficient, and cost effective. The first house to be built to this standard was constructed in 1991 in Darmstadt–Kranichstein, Germany. Now there are more than 20,000 houses built to this standard, and there is a movement to make the Passivehaus Standard required for all new housing units. Initially somewhat expensive, the additional cost over a conventionally built house is now between 5 and 7%.

As much as possible of the frame of a passivehaus is factory constructed, including mechanical elements such as plumbing and wiring. The design criteria include:

* Less than 15 kWh/m$^2\cdot$yr for space heating
* Less than 15 kWh/m$^2\cdot$yr for space cooling
* Less than 42 kWh/m$^2\cdot$yr for space heating and cooling, hot water and electricity
* Less than 120 kWh/m$^2$/yr as primary energy.

Insulation standards include R40 for the basement and walls and R60 for the roof. Heat recovery ventilators (see Section V of Chapter 2) with efficiencies of at least 80% are required since the air changes per hour must be 0.6 or less at 50 Pa [W. Post, “Reducing Energy Use of Houses,” which can be found at http://uppervalleysierraclub.org/articles/213-articles/243-reducingenergyusehomes] or ACH50. Recall from Section VIII of Chapter 4 that at ACH50 the air leakage is about 20 times that of natural air leakage. An ultra-tight house is a necessity at ACH50 ≤ 5, as is heat recovery ventilation.

The passivehaus requirements are considerably more stringent than existing U.S. standards. It is interesting to note the different response in Europe to government-imposed regulations. Europeans are willing to accept such regulations, whereas the reaction in some locations in the United States is that the government is interfering with the free market.

Figure 10.19: A Bermed Post and Beam House
The essentials of a passive solar house in a predominantly heating climate, in order of importance, are:
* exposure to the Sun, meaning a clear view towards the equator, and with the main axis of the house within 10° of true east/west,
* vertically mounted glass, double or triple glazing, on the winter Sun side, with HSGC, and some method to reduce heat loss through the glass during winter nights such as automated roller blinds, and an architectural overhang to shield the glass from summer Sun,
* high thermal mass inside the structure,
* tight construction with low air infiltration,
* residents who understand how to avoid wasting energy.

It also provides a number of benefits [op cit]:
* “Improved comfort and reduced energy costs to consumers
* Quality of environment and climate protection
* Security of energy supply and independence from energy imports
* Cheaper than investing in increased energy capacity
* Lower greenhouse gas emissions.”

Some of the strategies that either do not work, or are of limited value, include:
* Trombe walls,
* LoE glass on the passive solar face,
* too much glass on the walls not facing the winter, noontime Sun,
* glass on conservatories or “Sun rooms” that are set at a roof angle, from a low pitch to 24 pitch,
* high-tech systems that require monitoring and adjustments by the residents.
Chapter 11

Passive Solar Collection

The amount of solar energy reaching the Earth’s surface was discussed in Chapter 5. The energy is there, and now the problem to be addressed is how to capture it. The simplest collection method is passive — permitting the Sun’s energy to pass through a diathermanous skin into the structure, where it is converted to heat.

In Chapter 10 we looked at early to modern solar structures from an architectural standpoint. In Chapter 7 we calculated how much solar energy reaches a vertical surface and a latitude inclined surface at zero yaw angles. Here we consider surfaces with non-zero yaw angles.

This chapter addresses solar heating a structure in a predominantly heating climate. This covers most of North America and most of Europe, but will probably not be of interest for structures in the tropics, and some locations close to the tropics with a predominantly cooling environment.

Following the presentation in Chapter 7, we use the solar fraction and half day solar gain as the most important measures for determining the performance of a passively heated solar structure. Knowing the daytime solar gain is not enough, and we need to also calculate the losses during the nighttime hours through the glass. This shows the importance of using blinds or drapes to reduce unacceptable heat losses over a cold winter day.

An essential feature of a passive solar house is a high thermal mass, usually achieved by masonry materials such as concrete or rock. Still considered here as passive solar heat is the use of a fan system to either circulate the solar heated air, or to duct the heated air to another internal area such as a rock storage.

Solar radiation can be used for a variety of human needs. One such need is the continual need for potable water. Another is a way to heat water for bathing and cooking. Often these needs are more urgent in or near the tropics. The need for space heating is outside of the tropics.
11.1 Solar Gain on Vertical Collectors with Non-Zero Yaw Angles

It was shown in Chapter 5 that the angle of the Sun ($\theta$), whose elevation is $\alpha$ and azimuth is $\beta$, on a surface with angle to the vertical of $\phi$ and with a yaw angle $\gamma$, is

$$\cos \theta=\cos \alpha \cos \phi \cos(\beta-\gamma)+\sin \alpha \sin \phi.(11.1)$$

However, there is no good reason, and only bad reasons, to slope the south-facing glass at any angle $\phi$ other than zero. Passive solar collection requires the collectors to be vertical and facing the equator, or close to it. If the collectors are inclined forward, with the top of the glass further south than the bottom of the glass, solar gain is reduced. If the collectors are inclined backward, with the top of the glass further north than the bottom of the glass, solar gain is increased in winter, but the solar gain in summer will probably be intolerable.

Thus, with $\phi = 0$, Equation 11.1 becomes

$$\cos \theta=\cos \alpha \cos(\beta-\gamma).(11.2)$$

In Section I of Chapter 7 we investigated the solar fraction and half day solar gains for collectors at zero yaw angle. Here we investigate the solar energy on vertical surfaces at latitudes from 30 to 60 and with a moderate yaw angle. As we did in Chapter 7, there is no accounting for the reduction in solar energy due to air mass.

The three essential times of the year are winter solstice, equinox, and summer solstice, marked W, E, and S in the diagrams. The abscissa in the diagrams that follow is the time from solar noon and the responses are until the solar collection day ends. The daily gain is the sum of the half day gains for the positive and negative yaw angle $\gamma$ chosen.

For all yaw angles and with latitudes outside of the tropics, the collection day ends

* at winter solstice when $\alpha = 0$.
* at summer solstice when $\beta = 90 + \gamma$.

For all positive yaw angles the collection day at equinox ends at $\tau = 6$, $\alpha = 0$, and $\beta = 90$. For all negative yaw angles the collection day at equinox ends at $\beta = 90 + \gamma$.

The time $\tau$ and azimuth $\beta$ are given at the end of the collection day at winter solstice. $\tau$ and elevation $\alpha$ are given at the end of the collection day at summer solstice. $\tau$ and $\alpha$ are given at the end of the collection day at equinox.
Figure 11.1: $L = 30 \gamma = 10$

Figure 11.2: $L = 40 \gamma = 10$

Table 11.1: Gains at Latitude 30

<table>
<thead>
<tr>
<th>latitude 30</th>
<th>W</th>
<th>E</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>yaw 0</td>
<td>3.43+3.43</td>
<td>1.91+1.91</td>
<td>0.21+0.21</td>
</tr>
<tr>
<td>yaw ±10</td>
<td>3.84+2.92</td>
<td>2.55+1.33</td>
<td>0.45+0.11</td>
</tr>
<tr>
<td>yaw ±20</td>
<td>4.12+2.33</td>
<td>3.10+0.91</td>
<td>0.90+0.06</td>
</tr>
<tr>
<td>yaw ±30</td>
<td>4.29+1.67</td>
<td>3.57+0.62</td>
<td>1.58+0.04</td>
</tr>
</tbody>
</table>
Figure 11.3: L = 50 $\gamma = 10$

Figure 11.4: L = 60 $\gamma = 10$
Figure 11.5: L = 30, γ = 10

Figure 11.6: L = 40, γ = 10
Figure 11.7: \( L = 50 \) \( \gamma = 10 \)

Figure 11.8: \( L = 60 \) \( \gamma = 10 \)
Figure 11.9: $L = 30 \gamma = 20$

Figure 11.10: $L = 40 \gamma = 20$
Figure 11.11: \( L = 50 \) \( \gamma = 20 \)

Figure 11.12: \( L = 60 \) \( \gamma = 20 \)
Figure 11.13: $L = 30 - \gamma = 20$

Figure 11.14: $L = 40 - \gamma = 20$
Table 11.2: Gains at Latitude 40

<table>
<thead>
<tr>
<th>Latitude 40</th>
<th>W</th>
<th>E</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>yaw 0</td>
<td>3.49+3.49</td>
<td>2.46+2.46</td>
<td>0.73+0.73</td>
</tr>
<tr>
<td>yaw ±10</td>
<td>3.83+3.03</td>
<td>3.01+1.84</td>
<td>1.09+0.49</td>
</tr>
<tr>
<td>yaw ±20</td>
<td>4.12+2.52</td>
<td>3.616+1.34</td>
<td>0.90+0.33</td>
</tr>
<tr>
<td>yaw ±30</td>
<td>4.14+1.91</td>
<td>4.04+0.95</td>
<td>2.25+0.2</td>
</tr>
</tbody>
</table>
Table 11.3: Gains at Latitude 50

<table>
<thead>
<tr>
<th>Latitude 50</th>
<th>W</th>
<th>E</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>yaw 0</td>
<td>3.30+3.30</td>
<td>2.93+2.93</td>
<td>1.33+1.33</td>
</tr>
<tr>
<td>yaw ±10</td>
<td>3.55+2.96</td>
<td>3.55+2.29</td>
<td>1.77+0.98</td>
</tr>
<tr>
<td>yaw ±20</td>
<td>3.68+2.52</td>
<td>4.06+1.74</td>
<td>2.29+0.71</td>
</tr>
<tr>
<td>yaw ±30</td>
<td>3.71+2.01</td>
<td>4.44+1.26</td>
<td>2.91+0.50</td>
</tr>
</tbody>
</table>

The daily gain at latitude 40 and for yaws of 0°, 10°, 20°, and 30° at the three significant times of the year are given in Table 11.1. Similar tables for latitudes 40, 50, and 60 are shown in the next three tables.

In a predominantly heating climate, the solar gain should be maximized, but this is reduced by a yaw other than zero. The penalty is small for small yaw angles and for small latitudes. For example, at $L = 30$ and $\gamma = 10$, the daily solar gain at winter solstice is reduced from 6.46 to 6.36. The numbers for $L = 50$ are 5.78 and 5.69, also a small loss. Even at yaw angles of 30, the reduction is from 6.46 to 5.59, and at latitude 50 from 5.78 to 5.01.

Since the penalty in solar gain for having a non-zero yaw angle, at least up to 20° of yaw, is so small, this suggests a strategy for situations where some occlusion of the Sun occurs in early morning or late afternoon for a significant part of the year. It is possible to choose a yaw angle to face the passive solar elements away from the obstruction to reduce its adverse effects. However, the penalty for moderate to high yaw angle is unwanted summer gain.

In the analysis of this section so far it is assumed that the clear sky solar intensity is the same near dawn and dusk as at solar noon. A more realistic model for solar intensity, based on hard data that uses air mass $M$, was proposed in Section V of

Table 11.4: Gains at Latitude 60

<table>
<thead>
<tr>
<th>Latitude 60</th>
<th>W</th>
<th>E</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>yaw 0</td>
<td>2.55+2.55</td>
<td>3.31+3.31</td>
<td>1.93+1.93</td>
</tr>
<tr>
<td>yaw ±10</td>
<td>2.66+2.36</td>
<td>3.92+2.66</td>
<td>2.43+1.51</td>
</tr>
<tr>
<td>yaw ±20</td>
<td>2.70+2.06</td>
<td>4.41+2.06</td>
<td>2.97+1.14</td>
</tr>
<tr>
<td>yaw ±30</td>
<td>2.65+1.77</td>
<td>4.78+1.53</td>
<td>3.56+0.83</td>
</tr>
</tbody>
</table>

Chapter 7 as

$$IM=1000M^{0.4}-70,(11.3)$$

Here we use
for the air mass as derived in Section III of Chapter 6, and multiply the solar fraction by $M_0.4M_0.4$ where $M_0$ is the air mass at solar noon given by

$$M = (r^2 h^2 \cos^2 \Psi + 2rh+1)^{1/2} - rh \cos \Psi.$$ (11.4)

For latitudes 30 to 60 the next four figures show the solar fraction and the half day solar gains for the three significant times of the year with and without air mass reduction; the responses for air mass reduction are shown as dashed lines in the figures.

As expected, the reduction in solar fraction due to air mass is greatest at winter solstice, and least at summer solstice.

### 11.2 Solar Gain Through Glass

We have investigated solar gain on fixed, vertical, south-facing surfaces as a function of latitude. We also showed the effect of a yaw angle — when the orientation was not facing due south (for northern latitudes). In this section we consider what happens when the fixed, vertical surface is glass.
There is enough evidence that single glazing is unsatisfactory. In the United States, it is virtually impossible to purchase single-glazed windows. Further, the penalty, in the high purchase price and the reduction in solar transmission make triple glazing a poor choice for passive solar collection purposes. Thus, only double glazing will be considered here.

Before, we defined the solar fraction with respect to solar energy striking a surface at a specific orientation. Here we redefine solar fraction to be the ratio of the beam solar energy...
transmitted through the double glazing to the beam radiation. The absorptivity of each layer of glass to normal radiation is 5%.

For latitude 30, the solar fractions at winter solstice, equinox, and summer solstice are shown in Figure 11.21 as a function of the time away from solar noon. Also shown next to each response is the solar gain over the half day, the area under that response curve. Assuming constant insolation over the half day period, we can multiply this number by the insolation number to obtain the half day collection in W/m².

The solar fractions at latitude 40 are shown in Figure 11.22. Notice the increase in gains at all seasons. Finally, the responses for latitude 50 are shown in Figure 11.23: Here the largely unwanted summer gain may be a problem, requiring some form of summer shading. The summer heat gain at latitude 60, although an increase over the situation at latitude 50, may not
be as much of a problem since the summer ambient temperature will often be low enough to generate heating degree days.

11.3 Shading
It is possible to reduce the unwanted solar gain in summer without compromising the wanted solar gain in winter. The process in which an overhang is constructed over the south-facing glass is given the name “shading.” Consider the geometry shown in Figure 11.24, where the overhang is distance $y$ over the top of the glass and sticks out from the structure distance $x$. The relationship between $x$ and $y$ depends on the maximum elevation $\theta_1$ of the Sun at winter solstice, where $y = x \tan \theta_1$. Also, at summer solstice $y + h = x \tan \theta_1 + h = x \tan \theta_2$, where $\theta_2$ is the elevation of the

![Figure 11.23: Gains through Glass at Latitude 50](image-url)
Sun at solar noon, at summer solstice. Thus $x(tan \theta_2 - tan \theta_1) = h$, yielding

$$y = h1 - tan \theta_1 / tan \theta_2. \quad (11.6)$$

<table>
<thead>
<tr>
<th>latitude</th>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0.239</td>
<td>0.177</td>
</tr>
<tr>
<td>32</td>
<td>0.319</td>
<td>0.219</td>
</tr>
<tr>
<td>34</td>
<td>0.402</td>
<td>0.256</td>
</tr>
<tr>
<td>36</td>
<td>0.487</td>
<td>0.287</td>
</tr>
<tr>
<td>38</td>
<td>0.574</td>
<td>0.312</td>
</tr>
<tr>
<td>40</td>
<td>0.663</td>
<td>0.331</td>
</tr>
<tr>
<td>42</td>
<td>0.753</td>
<td>0.344</td>
</tr>
<tr>
<td>44</td>
<td>0.844</td>
<td>0.350</td>
</tr>
<tr>
<td>46</td>
<td>0.934</td>
<td>0.350</td>
</tr>
<tr>
<td>48</td>
<td>1.025</td>
<td>0.344</td>
</tr>
<tr>
<td>50</td>
<td>1.115</td>
<td>0.331</td>
</tr>
</tbody>
</table>

A plot of $x$ and $y$ for latitudes 30° through 50° is shown in Figure 11.25, when $h = 1.9$ m. Dimension $y$ reaches a maximum of 0.351 when the latitude is 45°.
The clear height $h$ of the glass is set at 1.9 m since this is the height of a sliding glass door, and tempered, insulated glass panels to fit such a door are available at modest cost, at least in the United States. Any size that differs from a stock size costs a premium.

![Figure 11.25: Shading Dimensions](image)

The stock size in the United States for the glass in an 8’ sliding glass door is 46”×76”. The dimension 74.5” is the clear height of the glass used by this author in a number of solar houses. The vinyl framing parts cover the edges of the glass by 3/4”, leaving 74.5” exposed to the Sun.

Figure 11.5 gives the actual sizes for $x$ and $y$. The $y$ values are not a problem, and can easily be incorporated in a typical residence. However, the $x$ values for latitudes greater than 44° present an architectural aesthetic problem. An overhang of 0.6 m is no problem, but an overhang greater than 0.8 m is not pleasing to look at. An overhang of over a meter needed for proper shading at latitudes 48° is downright unacceptable. Thus, a compromise is required. Luckily, for larger latitudes the summer solar gain is often needed.

### 11.4 Absorption of Glass — Not All Is Lost

The absorptivities of double glazing as a function of solar fraction at the three significant times of the year — winter solstice, equinox, and summer solstice — are shown in Figure 11.26: the latitude is 30. A1 refers to the outer sheet and A2 the inner.
Extracting the numbers from Figure 11.26 and putting them in Table 11.6, we see that the absorptivity of solar radiation in the glass is significant at winter solstice, is almost the same as solar transmissivity at equinox, and is 23 times greater than solar transmissivity at summer solstice, albeit 23 times a very small number.

The heating effect on the glass in winter reduces the heat loss through the glass from the interior heated space. In fact, if the temperature of the inner glass panel is the same as room temperature, then the conductive and convective heat losses from the room will be zero.

Table 11.6: Latitude 30

<table>
<thead>
<tr>
<th>trans</th>
<th>abs1</th>
<th>abs2</th>
<th>tot</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.099</td>
<td>0.328</td>
<td>0.225</td>
<td>0.553</td>
<td>26.35%</td>
</tr>
<tr>
<td>0.674</td>
<td>0.447</td>
<td>0.207</td>
<td>0.654</td>
<td>97.03</td>
</tr>
<tr>
<td>0.01</td>
<td>0.181</td>
<td>0.050</td>
<td>0.231</td>
<td>2310</td>
</tr>
</tbody>
</table>

The heated glass also will significantly reduce the radiative heat losses from the room, but it is complicated to determine by how much. The emissivity of the glass is high, so across the air gap the radiative heat loss dominates the other two mechanisms of heat transfer. But radiation from the inner glass panel directs 50% of the energy in, and 50% out. Similarly, the room-radiated energy is absorbed in the glass, helping to heat the glass.
For latitude 40, the absorptivities of double glazing as a function of solar fraction are shown in Figure 11.27.

Extracting the numbers from Figure 11.27 and putting them in Table 11.7, we see that the absorptivity of solar radiation in the glass remains significant in winter, and has increased in summer.

For latitude 50, the absorptivities of double glazing as a function of solar fraction are shown in Figure 11.28. Extracting the numbers from Figure 11.28 and putting them in Table 11.8, we see that the absorptivity of solar radiation in the glass remains significant in winter, and has increased in summer.

<table>
<thead>
<tr>
<th>Trans</th>
<th>Abs1</th>
<th>Abs2</th>
<th>Tot</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>2.316</td>
<td>0.284</td>
<td>0.206</td>
<td>0.490</td>
</tr>
<tr>
<td>Equinox</td>
<td>1.119</td>
<td>0.431</td>
<td>0.223</td>
<td>0.654</td>
</tr>
<tr>
<td>Summer</td>
<td>0.134</td>
<td>0.298</td>
<td>0.111</td>
<td>0.409</td>
</tr>
</tbody>
</table>
Shown in Figure 11.29 is the net solar gains at latitudes 30, 40, and 50 with incoming insolation at 880 W/m² and 440 W/m² through double and triple glazing. Each glass sheet absorbs 5% of the normal incident solar radiation. The net solar gain is the total radiation passing through the glass to the interior, plus the energy absorbed in the glass minus the heat lost through the glass due to the temperature differential from interior to exterior of 40°F = 22.2°C; with 3/16” glass and 1/2” air gaps the heat losses are 73.3 W/m² for double glazing and 49.3 W/m² for triple glazing.

Table 11.8: Latitude 50

<table>
<thead>
<tr>
<th></th>
<th>trans</th>
<th>abs1</th>
<th>abs2</th>
<th>tot</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>winter</td>
<td>2.326</td>
<td>0.232</td>
<td>0.176</td>
<td>0.408</td>
<td>17.54%</td>
</tr>
<tr>
<td>equinox</td>
<td>1.556</td>
<td>0.416</td>
<td>0.232</td>
<td>0.648</td>
<td>41.65</td>
</tr>
<tr>
<td>summer</td>
<td>0.418</td>
<td>0.345</td>
<td>0.154</td>
<td>0.499</td>
<td>119.38</td>
</tr>
</tbody>
</table>

Also given is the half day solar gain in kilowatts. As can be seen, there is little difference between the gains for double and triple glazing, reinforcing the conclusion that triple glazing is not worth the cost.
11.5 Solar Blinds

Building code requirements in the United States and elsewhere require a minimum percentage of wall areas for fenestration. For example, Massachusetts code requires a minimum of 10% of above-grade wall area to be windows or glass doors: they also set a maximum of 25%.

Windows and glass doors provide light, ventilation, and means of egress in case of emergency. Unfortunately, they are also the primary source, per unit area, of heat loss in winter and heat gain in summer.

The approximate R values in ft²·hr·°F/BTU for materials associated with windows are:

- wood ~1/inch
- fiberglass ~4/inch
- still air ~5
- heavy drapes ~?
- insulating blinds ~1-5
- shutters ~?
- glass - single ~1, double
Glass itself has an R value of less than 0.2, so it has effectively no insulating value. It does have surface resistances to reduce heat losses; see Chapter 3. Fiberglass has resistance four times greater than wood, not because of the insulating properties of the glass itself, but because of its ability to trap air and hold it still around the fibers. Drapes, blinds, and shutters use the same strategy — hold the air still.

Drapes can provide color and texture to a room or meeting hall, as well as effectively eliminate exterior light and noise. They can also be expensive and obtrusive even when drawn. Shutters require physical operation from the exterior twice a day, once to open and once to close; for this reason most people consider this unacceptable. There are many choices of blinds available. The standard office blind has slats, either vertical or horizontal, that can be rotated to let in more or less light: for these blinds, the insulation factor is usually not a concern. However, there are some products that have insulating slats, usually elliptical or airfoil shaped that can be rotated to let in light when needed or cozy up to each other to prevent air flow through them — such slats should always be horizontal.

![Figure 11.30: Honeycomb Blinds](image)

On a solar collection glass walls, insulated blinds that trap air are the right choice. They are economical and effective. Some, such as the Duette Architello line of blinds by Hunter Douglas, have R factors up to 7 plus. When drawn, blinds can be unobtrusive. There are many styles. The most common is honeycomb. Single, double, and triple layer honeycomb blinds are shown in Figure 11.30. The insulation value lies with the still air trapped in the honeycombs. The material of the blind should be such that sagging is minimal, a problem in particular for the double and triple layer blinds; it is important that the cells do not collapse. The blind must have a bottom rail to keep the blind straight and to provide enough gravitational force to allow the blind to deploy downwards so as to reach the bottom of the window frame.

The insulating blinds can reduce the heat losses in winter through the material of the blind to an acceptable level. The problem, however, is heat flow around the sides, top, and bottom. The top heat loss is easy to control by the method of affixing the blind to the top of the window frame. The bottom air control means that when down the blind must touch the bottom of the window frame, or be enclosed with blocks to reduce air flow. The main concern is leakages around the sides of the blinds. A gap of 1/8” permits excessive heat losses. Some manufacturers have plastic tracks that can be affixed to the sides of the window frame and in
which the blind slides. This author prefers a simpler, cheaper, more aesthetically pleasing and just as effective solution to winter heat loss — this is described next.

The construction detail for the south-facing glass and the insulating blinds is shown in Figure 11.31. The blinds are attached to the top frame so no air flow occurs at the attachment. The bottom of the blind has a bar, possibly made of wood, that keeps the bottom edge straight and acts as a weight to ensure that the blind is fully extended when needed. The glass is mounted outboard in a manufactured h-shaped frame that is pre-drilled for nailing into the supporting wooden frame. Wooden blocks are nailed to the sides of the vertical supports as shown. When closed in winter, cold air is trapped between the blind and the window, and cold air is denser than warm air, so the blinds are pushed against the side blocks, effectively trapping the cold air and increasing the insulating value of the blinds.

![Figure 11.31: Construction Details for Blinds](image)

The situation to be avoided at all costs is shown in Figure 11.32. There is an air gap above the top roller, and an air gap below the bottom rail, and these permit a closed cycle of warm room air entering the gap between the roller blind and the window.
As this air cools it falls to the bottom of the window and exits back in to the room. This situation can result in greater heat losses than would occur with no blinds at all.

11.6 Thermal Storage

In order for passive solar space heating to work, it is essential that the internal thermal mass be high. Otherwise, the interior winter temperature will be hot when the Sun is shining and cold when it is not. Masonry walls, insulated from the outside, can provide this storage.

Heat storage in rock with 30% voids requires \( 62.437 \times 0.7 = 2.41 \) times the volume of the equivalent storage in water.

1 gallon of water weighs 8.345 lb, 1 ft\(^3\) of water weighs 62.4 lb, 1 ft\(^3\) of water \(\equiv\) 7.49 gallons.

11.6.1 Direct Storage

The simplest solar thermal system is shown in Figure 11.33. The incoming radiation passes through the glass and reaches the thermal mass in the masonry floor and the masonry wall. It is important that the thermal mass be sufficient to absorb enough energy to prevent overheating on a sunny day.
It is important that there is no thermal bridging from the masonry to the environment. A concrete floor can be insulated from below with foamboard. If the masonry wall is the foundation for the structure, it must be insulated on the outside with foamboard.

Table 11.9

<table>
<thead>
<tr>
<th></th>
<th>specific heat BTU/lb-°F</th>
<th>density lb/ft³</th>
<th>unit heat capacity (no voids) BTU/ft³-°F</th>
</tr>
</thead>
<tbody>
<tr>
<td>water</td>
<td>1.00</td>
<td>62.4</td>
<td>62.4</td>
</tr>
<tr>
<td>scrap iron</td>
<td>0.112</td>
<td>489</td>
<td>55</td>
</tr>
<tr>
<td>magnetite</td>
<td>0.165</td>
<td>320</td>
<td>53</td>
</tr>
<tr>
<td>aluminum</td>
<td>0.215</td>
<td>168</td>
<td>36</td>
</tr>
<tr>
<td>concrete</td>
<td>0.27</td>
<td>140</td>
<td>38</td>
</tr>
<tr>
<td>rock</td>
<td>0.205</td>
<td>180</td>
<td>37</td>
</tr>
<tr>
<td>brick</td>
<td>0.20</td>
<td>140</td>
<td>28</td>
</tr>
<tr>
<td>sand</td>
<td>0.20</td>
<td>95</td>
<td>19</td>
</tr>
<tr>
<td>window glass</td>
<td>0.16</td>
<td>161</td>
<td>25.8</td>
</tr>
</tbody>
</table>

The orientation of the glass is due south. Shown in Figure 11.33 is an overhang above the glass to help reduce the solar gain in summer; it also protects the side of the house from the weather. As calculated in Section VII, this overhang may not be necessary to reduce the summer gain since it is so small. For the Boston latitude, the total daily solar gain on a sunny day is 4.36 Kwh/m² at winter solstice and 0.36 Kwh/m² at summer solstice for double glazing and 5% low iron glass.
11.6.2 Rock Storage

A variant to direct storage is to duct the heated air into a rock store. This requires a fan to take the heated air at the top of the wall, pass it through the rock store, and return the cooler air to the bottom of the wall as shown in Figure 11.34. To ensure adequate air flow through the rocks, a second fan on the return side is recommended.

![Figure 11.34: Rock Storage of Solar Energy](image)

11.6.3 Trombe Walls

Invented in 1881 by Edward Morse, but named after Frenchman Felix Trombe, a Trombe wall system is an elegant form of passive solar heating. Inside the south-facing glass, an insulated wall isolates the incoming solar radiation from the interior of the structure, so the wall heats up. Louvers at the top and bottom of the wall permit the heated air between glass and wall to flow as shown in Figure 11.35. The louvers only open one way, so that when the air temperature between glass and wall falls to the interior temperature, the louvers close and air flow ceases.

As shown, the louvers are passive and open when the heated air pressure causes the individual blades to swing inward in the top set, and swing outward in the bottom set. Often the louvers are powered, opening when the temperature between glass and wall is higher than the interior temperature. Another possibility is a louvered fan.
Regardless of which is chosen, the passive louvre, the powered louvre, or the fan with louvre, the success of a Trombe wall system depends to a large extent on the effective R value of the louvre system when at rest — an R-factor of at least 2 \( \text{ft}^2 \cdot \text{hr}/\text{BTU} \) is needed. Unfortunately, getting definitive numbers from the manufacturer appears difficult. Further, the powered unit should be quiet when in operation so as not to be objectionable.

Figure 11.35: The Trombe Wall
Some of the problems with Trombe wall systems were pointed out by Shurcliff [New Inventions in Low Cost Solar Heating, Brick House Publishing, Andover, Massachusetts, 1079]. One major problem is the inability to have access to the inside surface of the glass for the purpose of cleaning. Further, the wall blocks natural light, and could make the interior gloomy. The time constant is long, meaning it takes a long time for the morning Sun to warm the room behind the wall.

Shurcliff suggests some enhancements to make the Trombe wall perform better. Adding vertical fins painted black on the east side and reflective on the west side permits heat from the morning Sun to enter the room rapidly, while the afternoon Sun heats the wall. A plan view of the fin system is shown in Figure 11.36. However, the bottom line on Trombe walls is “forget it.” Far superior is the duct system with rock storage.

### 11.6.4 Water Storage

Water is cheap and has the highest thermal capacity of all common materials. However, to use water storage as a way to store passive solar energy commonly requires:

* an insulated tank to hold the water
  * an air-to-water heat exchanger
  * a fan system to push the heated air through the heat exchanger

All these are expensive, and there is a further problem — the potential for algae growth in the water since it will not reach the sterilization temperature of 140°F, 60°C.

An intriguing water-based passive collector was devised by Shawn Buckley of Massachusetts Institute of Technology in 1974 [op. cit.]. A water tank is divided by what Buckley called a
septum, as shown in Figure 11.37 with the side exposed to the Sun being far narrower than width of the tank on the other side of the tank. The water level in the tank is just below the top of the septum. There is a small gap between the septum and the bottom of the tank.

A mineral oil layer, whose specific gravity is about 0.9, is on top of the water. When the solar energy heats the water in the narrow tank, expands and flows over the top of the septum, being replaced by cold water flowing under the septum. At nighttime there is heat lost from the narrow tank to the environment, and the water in the narrow tank shrinks, pulling oil into this tank until equilibrium when there is no water movement. The septum acts as an insulator for the wide tank. The inside wall of the wide tank acts as a radiator for the room behind it.

11.6.5 Eutectic Salts

Water has two phase change states. One occurs as the temperature drops below 0°C and the water becomes ice, the other when the temperature rises above 100°C and water becomes steam. Water has a specific heat capacity of 1 cal/g-°C. However, to freeze water requires 80 cal/g, and to boil water requires 540 cal/g.
Adding salt to water depresses the temperature at which pure water freezes. Above a particular temperature for a given salt percentage, the salt solution is all liquid. As that temperature drops, at some point solids are formed, either ice or solid salt. If the temperature is dropped even further, the mixture becomes all solid. Surprisingly, a solution of 23.3% salt produces the lowest temperature, −21.1°C, at which the mixture is all liquid, called the eutectic temperature. The situation is shown in Figure 11.38.

Given a eutectic mixture of 23.3% salt, a drop in temperature from just above the eutectic temperature to just below it changes the solution from all liquid to all solid.

Sodium Sulfate $\text{Na}_2\text{SO}_4$ is a white crystalline solid. Adding water to it produces sal mirabilis (miraculous salt) $\text{Na}_2\text{SO}_4 \cdot 10\text{H}_2\text{O}$, now known as Glauber’s salt after Johann Rudolf Glauber, the seventeenth century Dutch/German chemist. The phase change occurs at 32.4°C, 90.3°F, and requires 83 cal/g. Thus, it appears to be a good medium for storage of heat, particularly in a structure heated by solar collectors.

![Figure 11.38: Phase States in Eutectic Salt](image)

There are some problems with Glauber salt. The salt is corrosive, posing difficulties with storage. The thermal conductivity of the material is low. Finally, after a number of melting/freezing cycles, there is a tendency for phase separation, inhibiting the phase change, as was discussed in the prior chapter.
11.7 Solar Stills

Clean drinking water is a problem in hotter and dryer parts of the world, the horse latitudes — see Section 8.1 of Chapter 1. There may be water, but it is not potable. On the seashore there is unlimited water, but it cannot be drunk. The quality of the water in many lakes and rivers is often poor. Unfortunately, the areas without clean water are too often the areas without electricity to boil or distil the water. Thus, the need for a self-sufficient solar still is large.

“Cooper, in his efforts to document the development and use of solar stills, reports that Arabian alchemists were the earliest known people to use solar distillation to produce potable water in the sixteenth century.” [P.I. Cooper, “Solar Distillation State of the Art and Prospects, - Solar energy in the Arab World, pp. 311–330, 1983]. “But the first documented reference for the device was a made in 1742 by Nicolo Ghezzi of Italy, although it is not known whether he went beyond the conceptual stage and actually built it.” [J. Gordes & H, McCracken, “Understanding Solar Stills,” Volunteers in Technical Assistance, Arlington, Virginia]

Distillation of water using solar radiation can be very important. Water in some regions is in very short supply, and some of the water available may be contaminated. The United States Department of the Interior has recognized the problem of distilling fresh water from sea water [J.W. O’Meara, “Objectives and Status of the Federal Saline Water Conversion Program,” in Advances in Chemistry, American Chemical Society, 1960]. A full discussion of water distillation plants using solar energy is given by F. Daniels [Direct Use of the Sun's Energy, Yale University Press, 1964].

The largest solar still was built in Chile in 1870 to provide drinking water for men and animals working the nitrate mines. The still covered 5,400m² and used heavily contaminated water from wells. About 19,000 liters of drinking water were distilled a day; this is a little over 3.6 l/m² of still area. After a railroad was built to supply the mines, the urgent need for a solar still was eliminated, but it remained in use until 1912.
Solar stills that are simple to construct can be very effective. A clear plastic cover over a pool of sea water or contaminated water traps some of the Sun’s energy, which vaporizes the water that condenses on the cover and trickles down into collection troughs, as show in Figure 11.39. Daniels shows that about 0.35 cm of water the area of the still can be extracted daily [op. cit.]. For a still with an area of 10 m$^2$, the volume of distilled water collected in a day is about $10^{0.35} \times 1000$ m$^3$, and water weighs 1000 kg/m$^3$, so the amount of distilled water is 3.5 kg/day.

An even simpler solar still, shown in Figure 11.40, works effectively on damp ground, or in the damp sand by the seashore. It can even work in the Sahara desert as long as moisture exists near the surface. Air leaks should be reduced to the minimum.

An even simpler one is shown in Figure 11.41. A circular depression is dug in the damp ground and in the bottom center into which a collection bucket is placed, partially covered so it will not tip over as it collects water. A sheet of clear plastic is placed over the depression and held in place by rocks or whatever is available. In the center of the sheet, over the bucket, a weight is placed so that the lowest point of the plastic sheet is the bottom of the weight. When the Sun shines on the sheet, evaporation occurs, and when the water vapor comes in contact with the underside of the cool plastic sheet it condenses. The pure, distilled water runs down the sheet to the center point and drips into the bucket.
The three solar stills discussed work effectively in the tropics where the Sun is always high in the sky. None of them would be effective at latitudes 40 or greater.

The eliodomestico solar water distiller shown in Figure 11.42 was voted one of the 25 best inventions for 2012 by *Time* magazine. It was developed by freelance designer Gabriele Diamanti, and is available for about $50 using simple materials such as earthenware and recycled plastics — this cost will plummet when it is mass produced, and to the credit of its inventor, it is open sourced. It is intended to distil sea water, and will only work in areas with high ambient temperatures. The device is about 12–14” in diameter and about 15” high. A black tank containing sea water at the top of the eliodomestico is exposed to the Sun. As the tank heats up the vapor pressure inside rises, and pure water vapor travels through the distillation tube to an unpressurized tank below. It is claimed that it produces about 6 liters of distilled water a day.
11.8 Solar Cookers and Ovens

Early attempts to capture solar energy were hampered by the absence of glass. We saw in Chapter 8 that glass became a thriving industry in Europe by the second millennium. By the eighteenth century, glass became an architectural statement for the wealthy. They constructed conservatories — showcases for plants. Greenhouses were for the utilitarian purpose of growing things.

We saw in Chapter 11 that the French–Swiss scientist Horace de Saussure invented the hot box. In 1767 he constructed five glass boxes of different sizes and without bottoms, so they could be stacked on top of one another. Placed under full Sun on a black wooden table, the interior reached a high enough temperature to cook food. He is considered the father of the solar box cooker.

Box solar ovens such as that shown in Figure 11.43 can be constructed out of common but durable materials, such as wood. Cooking temperatures can be as low as 80°C, and rarely does the temperature need to reach 170°C. Wood slowly chars at a rate of 0.3–0.5 cm/h at temperatures above 120°C [www.tcforensic.com.au]. The auto-ignition temperature is higher than this, but a wooden oven cannot be made to withstand temperatures slightly above the boiling point of water. Thus, wooden ovens are slow cookers.

An early cooker in the shape of a parabolic, polished aluminum dish about 1.2m in diameter focused solar energy on the bottom of a pot. It was developed at the Physical Laboratory of India, and was not a success [M.L. Ghai, “Application of Solar Energy for Heating, Air Conditioning and Cooking,” in Solar Energy Research, edited by F. Daniels and J.A. Duffie, University of Wisconsin Press, 1955]. It was capable of producing 500 W, but was not a success in spite of its modest cost of about $15. Experimentation with metalized mylar dishes, using a tubular steel frame, at the University of Wisconsin, with the same overall size and

The problem with all these designs was probably the difficulty of orienting in two axes, particularly to track solar azimuth. It is much easier to orient trough-type collection with reflectors that can be flat, of circular section, or best of all, parabolic.

Dr. Marie Telkes designed a box oven with aluminum flaps to direct solar energy through a glass window [“Solar Stoves,” in Transmission Conference on Use of Solar Energy: The Scientific Basis, Rome, 1961]. The collector area was about 0.5m² and the oven was capable of reaching about 200°C in full sunlight. When the oven was up to temperature, or when cloud cover occurred, an insulating pad was placed over the glass.

The box of a solar oven should be big enough to hold the pots that are typical in the area in which it is used. A hinged side door enables the pots to be placed inside the oven. The hinged top cover can be adjusted to reflect the sunlight at the optimum angle, and the hinges are tight so that the cover stays where placed. The inside of the top is reflective.
A somewhat more ambitious solar oven using adjustable and flexible wings to concentrate the Sun’s rays is shown in Figure 11.44. These wings could be polished aluminum, thin enough to be flexible. The orientation occurs at both ends of the box-shaped oven, which could be about 1m long, and 0.4m in width and height. A hinged door in the back provides access.

The orientation mechanism is shown in Figure 11.45. The orientation bar pivots about a point on the end plate coincident with the focal line and can be held in a number of discrete places by a locking pin. The flexible reflecting wings are affixed at one end to a spacer bar, which itself is held to the orientation bar. Glass the length of the box accepts the solar energy, either directly or by reflection from the reflective wings. The limits of movement of the orientation bar would be from about 10° to 30° to the vertical, so it is designed for use in the tropics only.

The bottom of the oven is on a horizontal plane, and the door of the oven faces north for northern latitudes and south for southern ones. The sides of the box could be constructed from two aluminum sheets with insulation inside and held together with fasteners and spacers. The glass would be a single pane and tempered.

A large number of solar cookers are in use in India and in China: By some estimates over 100,000. Most developing countries in the tropics use them. A popular solar cooker in Kenya is the CoolKit panel cooker [www.psp.org/e21/media/solar_cooking_v105_tn]

In 1987 SCI (Solar Cookers International), dedicated to the importance of solar powered ovens for third world countries, was formed by a group in southern California.
Their purpose was humanitarian, not profit, and they embraced the endorsement of UNESCO (United Nations Educational, Scientific and Cultural Organization) in sponsoring international solar cooking conferences.

A durable solar cooker developed at the Norwegian University of Science and Technology (NTNU) in Trondheim was recently announced [G. Zorpette, “Solar Cookers Get Hot,” IEEE Spectrum, Sept. 12, 2014]. Its claim to fame is the placement at the focal point of a parabolic reflective dish a container of a salt that undergoes phase change at 220°C, a suitable temperature for cooking. This enables the cooker to store enough energy to cook on a cloudy day. A thermal efficiency of 80% is claimed.

“The project began in 2011 with Asafaw Haileselassie Tesfay, a mechanical-engineering researcher at NTNU. Tesfay came to Norway from his native Ethiopia, which has a particular need for a solar cooker. Eighty five percent of Ethiopians lack access to electricity. And the country has been severely deforested: 40 percent of Ethiopia’s land was once covered by forest; the figure is now 4.6 percent. Ethiopian cuisine, too, has a very specific requirement: a cooker that can reach a temperature between 200°C and 250°C. Such a temperature is needed to bake injera, a spongy pancake that is eaten with most meals in Ethiopia [op. cit.].”

The container holding the phase-change salt is mounted on a frame, and the cooking pot rests on top of this. The frame itself does not move, but the parabola can be oriented with respect to
this frame. The parabola covers an arc of about 300°. A gap is on the high side to enable the
cook to have easy access to the pot without shading the pot from the Sun.

Solar cooking, while valuable in third world countries, demands little attention in developed
countries. However, there is a world-wide need for hot water.

11.9 Solar Water Heaters

Possibly the first one to invent an apparatus for heating domestic water by the Sun (called
SHW, a solar hot water system), such that it was piped from the place it was heated, was
Clarence Kemp. In 1881 he obtained U.S. patent #451,384 for “Apparatus for Utilizing the
Sun’s Rays for Heating Water,” and soon thereafter founded the C.M. Kemp Manufacturing
Company in Baltimore, Maryland, which marketed the heater under the name Climax. Before
his invention, the common way to heat domestic hot water was with pots over the kitchen
stove. The pots were carried to the bathroom and poured into the bathtub. Typically, a bath was
taken in tepid water a few inches deep. Not very satisfactory.

An improvement over the kitchen stove was a “wet-back” wood stove with a coiled pipe
inside. The disadvantage of this was that while hot water was needed year round, space heat
was only needed in the winter, so in the warmer seasons the unwanted heat from the wood
stove was uncomfortable.

The Climax system was mounted on the roof of the house. Inside a glazed frame, four 30-litre
galvanized steel tanks painted black and containing water were heated by the Sun. At the end of
the collection day, water was drawn by gravity from those tanks to be used for bathing, clothes
washing, dish washing, etc. Afterwards the tanks were refilled from the domestic cold water
supply. More than 1,600 Climax units were sold at a cost of about $25 by the turn of the new
century [Mc Donald & Bills, 2007]. The tanks had some insulation, but cold Maryland winters
could freeze the water in the tanks. For that reason, Kemp sold out to a southern California
company, and he died in 1911 a rich man. Kemp’s basic design is still in use today.

In Pasadena, California, a real estate developer named Frank Walker devised a major
improvement over on-the-roof SHW. He built the unit inside the roof with glazing above, so
heat losses were considerably reduced: Building integrated is the way to go, both in cost and
efficiency, as is advocated here. These units could be placed in colder environments without
freezing. Also, the system could use a wood stove as backup.

However, most consumers are highly price-sensitive, and when West Coast natural gas was
discovered in the 1920s, the market for SHW plummeted. As SHW died in the West, it boomed
in the Southeast: about half the houses in Miami used SHW by the 1940s [op. cit.].
Roof-mounted water heaters are common in countries not subject to winter freezing, such as Israel. During the fuel and electricity shortage soon after the formation of the State of Israel in the 1950s, engineer Levi Yissar seized the opportunity. The government forbade the heating of water between 10 pm and 6 am, making SHW that delivers maximum output at the end of the day desirable. Yissar formed the NerYah Company in 1953 to manufacture water heaters of his design. Within 15 years, about 5% of households in Israel had purchased 50,000 SHW units.

Following the energy crisis in the 1970s, in 1980 the Israeli Knesset passed a law requiring the installation of SHW in all new homes, with a few exceptions. Now 85% of households heat their water from the Sun, making Israel the world leader, and reducing energy from fossil fuels by 3%.

Here is what is happening elsewhere, worldwide, in SHW:

**Spain** required all new buildings starting in 2005 to install PV, and SHW systems in the following year.

**Australia**, with moderate winters, is moving in the same direction as Spain.

**Mediterranean countries** are doing the same, with **Cyprus** having over 30% of homes with SHW.

**Columbia**, thanks to the engineering work at Las Gaviotas, has installed over 40,000 SHW systems.

**China** has over 30 million SHW installed, with efficient evacuated tubes a significant fraction.

For colder winter environments, the Kemp type of SWH is not appropriate. China can have cold winters, and their use of evacuated tubes, particularly with selective surfaces to reduce radiation losses, is the wave of the future in SHW for northern Europe and the northern tier of the United States.
Chapter 12

Non-Concentrating, Active Solar Collectors

If a solar collector has a fixed orientation, it cannot be a concentrating collector with a concentrating factor of three or higher. This chapter considers fixed-orientation active collectors only.

The only solar systems considered in this text are solar ovens, described in Section IX of Chapter 11, and these require orientation capability in both azimuth and elevation. The most common application of a solar oven is in rural areas of third world countries, usually without access to an electrical grid or any other services common in advanced countries.

As discussed in Chapter 11, an essential feature of a passively solar heated structure is high thermal mass. Similarly, a high thermal mass to store the solar heat is necessary for a structure with active solar heat, such as flat plate collectors, evacuated tube collectors, heat pipes, or any other type of thermal solar collectors. In fact, any type of solar thermal collector designed to be the primary heat source of the structure needs to be able to store several days’ heat in some type of thermal mass.

Passive solar heat is typically stored in masonry materials, such as concrete or rock where air is the vehicle, doing its job by natural convection or with a fan. The transport vehicle for active solar heat can be air, or a liquid such as water or oil, and the heat is contained in a tank.

As we have seen, the heat capacity of water is superior to all other common materials, but it has drawbacks, such as bacterial growth or corrosion — it is not called the universal solvent for nothing.

12.1 Solar Gain on Sloping Surfaces

It is desirable to maximize the solar gain of a solar collector, but what exactly does this mean? In a predominantly heating environment, it could mean maximizing the winter solar gain. In a predominantly cooling environment, a solar thermal system could heat water, but with no need for heating the structure.
For a grid-connected photovoltaic system in which the utility company returns the same amount per kilowatt hour generated throughout the year, it means maximizing the annual number of kilowatt hours generated. The conventional wisdom says that the ideal angle for year-round solar gain is latitude inclined, and this will be assumed here.

If the yaw angle of the collector is zero, then the angle $\theta$ of the Sun to the normal to the collector, as given by Equation 5.26, is given by

$$
\cos \theta = \cos \alpha \cos \beta \cos \phi + \sin \alpha \sin \phi \tag{12.1}
$$

where $\phi$ is the tilt angle of the collectors to the vertical.

The solar collection day ends when either

* the solar elevation becomes negative, or
* the angle between the Sun and the collector surface becomes negative.

In the latter case, we test for this by determining that this angle increases, which is impossible unless the Sun is behind the collector.

![Figure 12.1: Solar Gains](image)

The half day solar gains at various latitudes as a function of the tilt angle are shown in Figure 12.1.

Shown in the next four figures are the solar fractions for latitudes 30 through 60 when the sum of the latitude $L$ and the tilt angle to the vertical $\phi$ are 90; that is, the collectors are latitude inclined. The solid lines are when there is no accounting for air mass, so the air mass is unity, and the dashed lines are when the solar fraction is attenuated by $M_0.4M_0.4$ with
\[ M = (r^2 h^2 \cos^2 \psi + 2rh + 1)^{1/2} - rh \cos \psi \] (12.2)

for the air mass, as derived in Section III of Chapter 6, and the air mass at solar noon of

\[ M_0 = (r^2 h^2 \cos^2 (L - \delta) + 2rh + 1)^{1/2} - rh \cos (L - \delta) \] (12.3)

where the zenith angle is then \( \psi = L - \delta \), the azimuth \( \beta = 0 \), and the solar elevation is \( \alpha = 90 - L + \delta = \phi + \delta \), so the angle of the Sun on the collector is \( \cos \theta = \cos \alpha \cos \phi + \sin \alpha \sin \phi = \cos(\alpha - \phi) = \cos \delta \). Since the declination at winter solstice is \(-23.45\) and at summer solstice is \(\delta = 23.45\), then at solar noon the solar fractions are the same for these times of the year, as it will be for declinations at a specific \( \pm \delta \).

![Solar fraction graph](image)

Figure 12.2: Latitude 30

Also shown in Figures 12.2 through 12.5 is the sum of the half day solar gains at the three times of the year, where \( tot_1 \) is for solar gain with no reduction due to air mass, and \( tot_M \) is for solar gain attenuated by \( M_0 \).04M_0.4 \).
For all latitudes and no consideration of air mass, the response at winter solstice is the same as at summer solstice until the winter response terminates when $\alpha = 0$. For example, at $L = 30$ the winter responses end at $\tau = 5.03$. The effect of air mass is a modest reduction in solar fraction at summer solstice and equinox, but a large reduction at winter solstice, particularly at larger latitudes.
We will now investigate the latitude inclined phenomenon where the response at winter and summer is the same until the winter response ends when the solar elevation reaches zero. The solar gains for the full range of latitudes are shown as the blue responses in Figure 12.6, and it
can be seen that for all latitudes and all declinations on the summer side of equinox, the gains are the same. However, on the winter side of equinox the solar gains are a function of latitude. At equinox the solar gain is 3.82. The collection day ends when $\alpha \geq 0$ on the winter side of equinox, but ends when $\theta \geq 90$ on the summer side.

Using the air mass model, the solar gain at equinox is 3.42 for all latitudes. On the summer side of equinox the solar gain is greater at higher latitudes, but the opposite is true for the winter side of equinox.

### 12.2 Air Collectors

Facade collectors use the skin of the building to collect solar energy. They can be passive, like the Trombe wall system discussed as a passive solar heating system in Chapter 11, or they involve fans, as will be discussed in this section. Also discussed in Chapter 11 is a semi-passive heating system in which solar energy entering from the south-facing windows heats the interior space, rises to the ceiling(s), and is ducted into a rock store with the aid of one or more fans. Since it employs fans, it could be strictly considered an active system, but since it directly heats the interior of the structure we classify it as passive. Here the facade collectors pass the heated air directly to a thermal storage area without initially heating the interior spaces.

A simple facade collector is shown in Figure 12.7. The air between the south facing glass and the plate with high absorptivity to solar radiation is heated when the Sun is shining. A sensor (not shown) is attached to the back of the absorbing plate so that when it is warm the fan turns on, sucking the heated air down into the rock storage. This system can also work on an inclined surface, such as a roof.
A variant on the air-to-rock heat exchange is to employ an air-to-water heat exchanger, thus storing the solar energy in a water tank; this tank can be a preheater for the domestic water supply. As a further refinement, when the structure calls for space heat, the heat exchanger can be bypassed so the hot air flows directly into the rooms.

12.3 Liquid Carrying Flat Plate Solar Collectors

12.3.1 Trickle Type Collectors
Trickle type collectors such as that shown in Figure 12.8 can take several forms. Harry E. Thomason of Washington, DC was an early advocate for such collectors, with water trickling down corrugated and blackened metal panels. He built several houses in the Washington DC area using this type of active collector. The blackened surface does not need to be corrugated provided the complete surface can be wetted. Suitable wetting agents may be required in the water; sunlight itself is a good wetting agent. A problem with this type of trickle collector is that condensation can form on the inside of the glass, blocking some of the incoming solar radiation, and reducing the efficiency of the collector. Other objections are the engineering problems of controlling flow to fully wet the collector surface.

Fully wetted trickle collectors can eliminate the previous objection. One such collector uses two metal sheets that are in close proximity, with water flowing in between. The upper metal surface has high absorptivity to the solar energy. The Ron Shore house in Snowmass, Colorado uses this system; in this case the metal sheets are corrugated. A variant on this is to make the upper surface of the lower metal sheet have high absorptivity to the solar energy and change the top cover to transparent plastic sheeting. Either way, it is difficult to ensure full wetting of the bottom sheet, and service durability is suspect.

The refractive index of water as a function of wavelength varies: from close to 1.4 at the wavelength of $\lambda = 0.2 \, \mu$, dropping to a low of about 1.15 at $\lambda = 2.8 \, \mu$, rising precipitously to 1.48 at about $\lambda = 3.5 \, \mu$, drops again to a low of about 1.14 at about $\lambda = 15 \, \mu$, then rises continually for longer wavelengths [http://refractiveindex.info/?shelf=main&book=H2O&page=Hale]. However, such detail is not needed here, and we can use the commonly accepted refractive index of water as 1.33.

In Chapter 8 we saw that reflectivity of solar energy to normal incident radiation at an air-to-glass interface is about 4%. Taking the two refractive indices of glass and water as 1.515 and 1.33, respectively, the reflection to normal incident radiation is

$$\rho = \frac{(\eta_1 - \eta_2\eta_1 + \eta_2)^2}{(1.515 - 1.331.515 + 1.33)^2} = 0.1852.845 = 0.004228, (12.4)$$

or approximately 0.4%, a tenth of the reflectivity between glass and air. Thus, if there is water but no air in the space between the blackened sheet and the glass, then the reflectivity between glass and metal sheet can be ignored. However, the glass itself will be at the same temperature as the water, resulting in unacceptable heat losses on cold but sunny days. The solution is to employ double glazing on the collector.

### 12.3.2 Closed Loop and Fully Wetted Collectors

Revere Brass and Copper, Bridgeport, Connecticut, make a copper plate in which copper pipes are brazed with a single inlet and outlet for easy hookup. This plate is expensive and has the additional disadvantage that the temperature between the plate and the tube can be significant, reducing the operating efficiency.
The Roll-Bond division of Olin Brass produced an integral tube-in-plate made of copper or aluminum. The plate is in reality two plates metallically heat and pressure bonded over their entire surface except for the tube pattern laid down by a silkscreen process. Air pressure at a single point expands and forms the tubes. Any tube pattern can be produced. Inlet and outlet pipes are provided. The bonding procedure is trouble free, and the fluid in the tubes can be at high pressure without fear of leaks. The temperature differential between a point on the external surface and the inside surface of the integral tube is small, even with high fluid flow rates at low temperatures.

The roll-bond process was invented by Leland H. Grenell and described in U.S. patent #2690002, issued in September 28, 1954, and assigned to Olin Industries. Some years later, Olin sold off, or dropped, its Roll-Bond division. Since then, the only players in this game appear to be a number of Chinese companies. The primary application for these panels is as evaporators in refrigeration, and the material used is aluminum. The interest and need for solar energy solutions in recent years has encouraged these companies to expand their roll-bond panels to include panels designed as the heart of flat plate solar collectors.

The roll-bond technology permits the rapid production of panels of almost any size and tube pattern. A good width for residential use would be 22” — studs and rafters are typically 24” on center, so they produce a spacing of 22 1/2”. Not coincidentally, the Olin panels circa 1970 were available in 22” widths. The length of the panel can be chosen to suit the length of the rafters or studs, allowing enough space to connect to the feed and return tubes at the ends of each panel.

A hybrid system with both active and passive solar elements is shown in Figure 12.10. The active system with roll bond, flat plate solar collectors delivers heat via a heat exchanger into a water tank. The fluid in the roof-mounted collectors should be oil to reduce the dangers of corrosion and failure with a water based system, and to protect against freezing problems.
The roll-bond panels can also be used as room radiators. That is, these panels can be affixed to walls. When a room thermostat demands heat, a circulating pump draws heat from the water tank as shown in Figure 12.10. These tall radiators are far more effective than baseboard radiation since they can deliver heat at considerably lower temperatures of the collector — 100°F rather than 180°F. There is one disadvantage, though, and that is the tall radiators on one wall will create an air flow cycle that can produce uncomfortable conditions even when the room temperature is acceptable; for this reason, radiator panels on opposite walls may be desirable.

The use of passive solar collection necessitates thermal mass. This mass should be such that there is adequate air flow around it. Further, if it is on an exterior wall it should be insulated on the outside.

12.4 Stratified Water Tanks

As discussed in Chapter 2, the average person in North America requires about 12kWh/day to heat water, which amounts to about 18% of the total energy needs of a residence. Space heating takes about 42% of the energy needs, so water and space heating are 60% of the total energy use according to the U.S. Energy Information Administration[www.eia.gov/consumption/residential/data/2009]. This leaves 40% to service the appliances; most of this is for refrigeration, with only 6% for air conditioning.
Stratified water tanks are an effective way to cut down the energy used to heat water. The first part of the strategy is to reduce the flow rate of cold water entering the bottom of the tank to reduce the mixing of this water to the warmer water above the entry point. When used with a solar collection system, the next part of the strategy is to heat the water in the tank externally with a mantle, as shown in Figure 12.11.

Also shown is another feature — a radiator system getting its heat from the mantle. As shown the radiator is tall, as was discussed with Figure 12.10.

12.5 Selective Absorptivity/Emissivity Surfaces

12.5.1 A Statement of the Problem

The absorption of solar energy in the usual flat plate collector is on a black painted or plated or anodized surface. Such a surface is usually non-selective, i.e., the emissivity at one wavelength is the same as the emissivity at another wavelength.
The coating on the flat surface exposed to the Sun should absorb as much of the Sun’s energy as possible, meaning that its absorptivity over the waveband 0.3\( \mu \) to 2.5\( \mu \) should be close to 1. In order to improve the overall efficiency of the collector, it is desirable that the same coating have low emissivity to the re-radiated infrared. Luckily, there are such selective surfaces. For example, a deposition technique for roll-bond panels comprised of 99.5% Al and 0.5% AlMnZr with solar absorptivity of 94% and infrared emissivity of 14% was developed by M. Moeller under NASA grant NNX09AB39G, and was awarded to the Smithsonian Astrophysical Observatory.

There are selective absorptivity/emissivity surfaces. White paint reflects most solar energy. However, at temperatures of about 330°K, the emissivity is about 0.9, so it is typical in Mediterranean climates to whiten the outside walls of masonry houses to reflect most of the sun’s energy yet to re-radiate the infrared. The high thermal mass moderates the daytime/nighttime temperature differential.

### 12.5.2 Some Early Results

Enhanced efficiency of collection will result from a back plate with a high absorptivity (greater than 0.9) but a low emissivity (less than 0.1) to the reradiated infrared energy. Such a
surface was developed by NASA at its Marshall Space Flight Center, Alabama [report # M-TU-74-3, May 10, 1974]. The collector is plated with a bright metal such as nickel and polished. This surface has a low emissivity and absorptivity to both solar and infrared energy. However, nickel is good conductor of heat. The bright surface of the collector is then plated with a very thin (0.1μ to 1μ) black, porous layer. This black layer is thick compared to the wavelengths of visible light, so most solar energy is absorbed. However, the infrared energy is at longer wavelengths for which the thin black surface appears transparent, and so the low emissivity to the infrared is about 0.06. No mention has been made of this invention in recent years, so it can be surmised that it either does not work or that other systems exist that are simpler, or more efficient, or both.

A complicated experimental system was suggested by NASA at their Marshall Space Flight Center [NASA Technical Brief B73-10485]. No glazing is needed. Instead, two coolants are used, one at a high temperature and one at a low temperature. The low temperature coolant flows through louvers that are perpendicular to the back plate. Solar energy bounces off the sides of these louvers so that most of the energy reaches the saw-tooth-shaped rear absorber containing the high temperature coolant. Efficiencies of 90%, even with high temperature differentials between the back plate and the outside environment, are claimed.

Another complicated solar collection system was developed by the same group [NASA Technical Brief B73-10484]. This system uses two temperatures of coolant as did the previous system. Three sheets of glass are used. The low temperature coolant passes through the two inner sheets of glass to capture the solar energy absorbed by these two layers. The solar energy passing through these inner sheets reaches the V-shaped grooved rear absorber though which the high temperature coolant flows. The claimed efficiency of 95% is unrealistic.

What is the purpose of NASA Technical Briefs? As stated in the heading of each brief, “NASA Tech Briefs announce new technology derived from the U.S. space program. They are issued to encourage commercial application.” It appears that nobody was interested in developing the two NASA Briefs discussed.

Water-based systems have advantages and disadvantages. The specific heat of water is defined as unity — higher than other common materials such as concrete (0.27 BTU/lb °F), rock (0.205 BTU/lb °F), or brick (0.20 BTU/lb °F). A significant disadvantage of water is its corrosive potential.

### 12.5.3 Cermat Surfaces

Cermat is a mixture of ceramic and metal at the atomic level, meaning that an atom of a ceramic material is bonded to an atom of a metal. When sputtered on a substrate it can absorb over 90% of solar energy but have an emissivity to the re-radiated infrared of less than 5%. The cermat coating is typically about 0.25μ thick, slightly less than the short wavelengths of
visible light. It is metal rich at the substrate and ceramic rich at the surface, with a continuous gradation throughout the thickness of the coating: such a gradation is illustrated in Figure 12.12, where only metal atoms are on the substrate and only ceramic atoms are exposed to the Sun.

![Figure 12.12: The Cermat Surface](image)

Sputtering is a highly technical procedure. Molecules of an ionized gas under low pressure displace target atoms, which then bond to a substrate and create a thin, uniform, durable film.

Cermat surfaces are durable, with a service life of up to 30 years. They are also stable even at temperatures up to 500°C. However, they cannot be applied in situ, meaning that this is a disadvantage for building-integrated systems. Also, it is an expensive procedure.

### 12.5.4 Selective Coating Films

Films that are applied to collector surfaces are commercially available, and could well be the surface of choice until spray paints with good absorptivity/emissivity characteristics and service durability are developed.
Using technology developed for the space industry, the Israeli company Acktar has a selective film with application in flat plate solar collectors and concentrating solar collectors. They claim a solar absorptivity of 98%, but this does not match the reflectance characteristics provided by them and duplicated here in Figure 12.13. They do say that the re-radiated emissivity is about 11%, and this is consistent with Figure 12.13.

Solmax produces foils for flat plate solar collectors with solar absorptivity of 95–99%, although this figure may be inflated since the adhesive to the substrate has thermal resistance. Further, the adhesive cannot withstand temperatures over 180°C. Its emissivity to re-radiation is 4–10%.

12.5.5 Optically Selective Paints and Sprays

Thermofin has a selective coating they call Crystal Clear that has absorptivity 94–96%, and emissivity to re-radiation of 7–10% [www.thermaf.in.com/coat_intro.shtml]. It is a quartz encapsulated bimetallic alloy.

Solkote has absorptivity 88–94%, and a poor emissivity to re-radiation of 20–49%. It is a spray that can easily be applied in situ for a building-integrated application. It has good high temperature durability [www.solec.org/solkote/solkote-technical-specifications].

The characteristics of Thermalox are similar to those of Solkote, with somewhat better absorptivity at 96% but worse emissivity to re-radiation of 52%, and is easy to apply on site [www.dampney.com/ProductLine/tabid/123/AT/View/PID/2/Thermalox-250.aspx].
12.6 The Bread Box Solar Collector

The bread box system, as shown in Figure 12.14, sometimes called the batch solar collector or an integral collector, is an integral, all-in-one collector that heats the domestic pressurized water for space heating applications of pre-heating domestic hot water. It is an insulated glazed box (bread box) facing the mid-day Sun, containing a storage tank that is directly heated by the Sun. This system cannot be employed where freezing conditions can occur.

12.7 Evacuated Tube Collectors

In order to reduce convective heat losses, one strategy is to use a vacuum between concentric tubes as shown in Figure 12.15. The inner tube is has a coating designed to absorb the maximum amount of solar energy. This coating could also have selective properties, so its absorptivity to the incoming solar radiation is high, but the heat loss to the re-radiated infrared is low. Inside the inner tube is a water-based liquid or oil.
Figure 12.14: The Bread Box Collector

- Cold water feed
- Hot water output
- Hot water tank

Figure 12.15: An Evacuated Tube Collector

- Fluid
- Glass tube
- Energy absorbing tube
- Vacuum
Typically, a rack of such tubes is constructed with a manifold at each end, and the assembly mounted on a south-facing roof. The problem with this is that it is susceptible to mechanical damage, such as being struck with a branch of a tree. If the inner tube is punctured, the complete system could pump its fluid out.

The efficiency of the evacuated tube collector is reduced by the ratio of diameters of the tubes. The device intercepts solar energy at the width of the outer diameter, but the solar collection occurs over the diameter of the inner diameter. For this reason, care should be taken before accepting a manufacturer’s efficiency numbers.

![Figure 12.16: Reflective Enhancement for Tube Collector](image)

Efficiency can be increased by installing a semicircular reflector, as shown in Figure 12.16. There is a twofold benefit: Additional solar energy is reflected onto the collector, and 50% of the energy radiated from the collector is reflected back to it.

For the reflector to have best effect, the axes of the cylinders are assumed to be east–west to eliminate consideration of solar azimuth. This reflector is inclined to 40° to the horizontal, and this is suitable for a latitude of about 40°. The angles for latitude 40 at winter and summer solstices at solar noon are close to 26° and 76°, respectively. These angles are shown on the right-hand side of the reflector; as can be seen, there is no shading by the reflector on the collector. On the other side of the reflector the problem of shading would occur at low solar angles, so here at an elevation of 5° there is no shading; this works here for the ratio of diameters shown, but other ratios may require something other that a semicircular reflector.
12.8 Heat Pipes

The heat pipe is similar to the evacuated tube described in the prior section, but here the inner tube is copper or aluminum, and holds enough of a vacuum so the liquid inside boils at a reduced temperature. The liquid is alcohol, mercury, or water with additives.

The inside walls of the tube are coated or constructed to act as a wick. The wick can be a sintered metal powder, or fine grooves to substantially increase the interior surface area.

When solar energy heats this liquid it vaporizes, soaks the wick, and rises to the top of the tube, where a manifold removes heat so the vapor turns to liquid, which trickles back down the tube to the bottom. The minimum tilt angle to the horizontal should be 25° or more in order for the internal fluid to flow back to the bottom. The cycle continues as long as the Sun is shining on it. Much higher temperatures are produced than with flat plate collectors, but because of the vacuum the temperature differentials along the length of the tube are small [www.aavid.com/product-group/heatpipe/operate].

There are no moving parts, so there are no maintenance problems [www.ocmodshop.com/heat-pipes-explained]. However, heat pipes may have a problem with self starting.
Chapter 13

Photovoltaic Panels

The underlying principles behind semiconductors in general, and photovoltaic cells in particular, are presented in this chapter. The physics of these devices is far more complicated than the simplified descriptions given here, and those who need an in-depth study should consult the textbooks and papers in scientific journals. A single short chapter, such as this is, cannot do justice to such a complex field. However, it can present the fundamentals needed to appreciate the state-of-the-art in photo-voltaics, and define the promise of high efficiency with third generation PV cells.

We start with a brief history of semiconductors, a term attributed to Alessandro Volta in 1782. Rectification with metal sulfides was discovered by Karl Ferdinand Braun in 1874, but it took until the 1930s for Russell Ohl at Bell Telephone Laboratories at Holmdel, New Jersey to observe a strange phenomenon while experimenting with silicon rectifiers for use in radar detectors. A crack, in the crystalline structure of the silicon held impurities. When a bright light was shone on the crack, the current flowing through the silicon increased with the impurity phosphorus, but with the impurity boron the current decreased. Phosphorus produced a surplus of negatively charged electrons, afterwards called an n-type region, while boron produced a dearth of electrons, later called holes, and called a p-type region. The p-n junction was the ultimate result.

Experimentation with p-n junctions developed led to the invention of the point contact transistor. The next big evolution was the invention of the field effect transistor that made integrated circuits possible. The silicon diode that was capable of generating electricity when exposed to sunlight, the photovoltaic (PV) cell, was next. Thus, we have come in a circle, with Russell Ohl observing the PV effect, which led to the transistor and the integrated circuit, and now to PV cells with increasing efficiency.

In the 1950s the first application for PVs was for Earth-orbiting satellites; they were expensive, but money was no object. Manufacturing costs had dropped sufficiently by the 1970s such that PVs were charging navigation and telecommunication equipment. By the 1980s, PVs were powering hand-held calculators and radios.
One of the early PV-powered residences was the University of Delaware’s grid connected “Solar One.” When the Sun produced electrical power from the PV panels, these panels also produced heat, which was passed by fans into bins of phase change material. Solar One is a fine example of a hybrid building-integrated solar system.

Semiconductor technology evolves rapidly. Until this century, the cost of PV arrays made its use for residential purposes unlikely. We have now moved from first generation crystalline to second generation thin film and a dramatic drop in cost. Extensive research is being conducted on the development of third generation cells with substantially higher efficiency. Also, a discard from the past that seemed to have little future has emerged as a low cost contender whose improvement in performance since 2009 has been astounding — perovskites.

Without a doubt, the future will see continued expansion of PV generated electricity. This subject, including the technical details regarding grid-connections with inverters, is considered in the next chapter.

### 13.1 Semiconductor Fundamentals

Although the term semiconductor originally meant the construction with materials whose electrical conductivity was somewhere between an insulator and a conductor, its application now dominates its meaning → the crystal of silicon doped with some impurity in layers to produce an electrical effect that can be harnessed to amplify a signal or act as an electronic switch. Other materials, such as germanium or gallium arsenide, can be used for the base crystal, but only recently perovskites appear to have been given a chance to challenge the dominance of silicon.

The study of semiconductors, or solid state devices, starts with the simple model of the atom shown in Figure 13.1. The electrically neutral atom contains a positively charged nucleus surrounded by orbiting negatively charged electrons. The electrons lie within specific bands, usually three, with the outer band called the valence band. The forces holding the electrons in the bands is weakest in the valence band. When an electron in the valence band breaks free from the atom, it can move around between surrounding atoms, and may be captured by another atom that has previously lost a valence electron. Electric current is caused by the movement of electrons.

The energy level of the bands of electrons is measured in electron volts and is shown in Figure 13.2. The energy of an electron exists in discrete states, so an electron cannot exist in the spaces between bands. An insulator, such as pure silicon, has a large energy gap between the valence band and the conduction band. When a valence electron in this insulator becomes more energetic, due to heat or electric charge, it can jump into the conduction band. Few electrons make this jump, so little electric current can flow. However, in a metal the conduction band
partially overlaps the valence band, so electrons are drifting freely at all times, and electric current is easily induced.

Silicon is a crystal with a regular pattern of atoms. It has four valence electrons that are tied to electrons in surrounding electrons with what are called covalent bonds. These covalent bonds are difficult to break, and indicate the large energy gap between the valence and conduction
bands. Suppose a small number of atoms of some impurity, such as arsenic with five valence electrons, is introduced into the silicon matrix.

![Figure 13.3: A Loose Electron](image1)

![Figure 13.4: A Loose Hole](image2)

Each arsenic atom is surrounded by four silicon atoms, and the covalent bonds in the silicon remain intact as shown in Figure 13.3. However, the fifth valence electron of the arsenic has a weak bond to its parent, and heat or electric charge can easily break this bond. The energy state of this donor electron when held to the atom is just beneath the conduction band. Silicon doped with arsenic, or any other element with five valence electrons, is called an n-type semiconductor.

Silicon can also be doped with atoms having three valence electrons, and this causes a covalent bond to a neighboring silicon atom to be missing, as shown in Figure 13.4; this missing electron is called a hole. The way to think of it is as a positively charged particle that...
can move around like an electron. The hole is ready to accept an electron, and we show this acceptor energy state as just above the valence band energy. Thus, it takes a small amount of energy for a valence electron of a silicon atom to be boosted to the acceptor level, satisfying the hole in the impurity but creating a hole elsewhere.

![Energy Gaps Diagram](image)

Figure 13.5: Energy Gaps

The energy gaps for insulators, semiconductors, and conductors is shown in a simplified fashion in Figure 13.5.

### 13.2 p-n Junctions

When p-type and n-type materials are bonded together into what is known as a p-n junction, electrons drift from the n-type material to the p-type, and holes drift from the p-type material to the n-type, creating the electric charge as shown in Figure 13.6.

Suppose metallic contacts are bonded to each end of the p-n junction and connected to a voltage source such as a battery. Current is induced that depends on the voltage and the polarity of the voltage. The p-n junction and the resulting voltage/current characteristics are shown in Figure 13.7. If the voltage reduces the inherent charge across the junction, due to the drift of electrons and holes, we say the junction is forward biased. However, when the polarity of the voltage is reversed, the inherent charge is increased and the junction is said to be reversed biased. $I_s$ is the maximum current that can flow when the junction is reverse biased. The device is known as a diode and is used extensively in electronic circuits, such as power supplies, which convert alternating current into direct current that can be used in computers, television sets, and radios.
13.3 The First Transistors

If p-type material is sandwiched between two n-type semiconductors, we produce an npn transistor as shown in Figure 13.8. In a typical operation, one pn junction is forward biased and the other is reverse biased; in this mode the voltages across the semiconductor are as shown. The forward biased n-type material is called the emitter, the central p-type material is called the base, and the reverse biased n-type material is called the collector. Under the influence of emitter-to-base voltage $-V_e$, electrons drift into the base to become what is known
as minority carriers. In the base there are the inherent holes (called majority carriers). The number of carriers, majority and minority, in the base is small, so the chances of their combination is small. The majority carriers can drift into the emitter across the forward biased junction, but the reverse bias on the base-collector junction prevents their drift into the collector. However, what is a barrier to the majority carriers in the base is a steep downhill slope to the minority carriers, which are swept across the base-collector junction into the collector. The small emitter voltage causes current to flow in the high voltage collector circuit, and this effect can be used to amplify small signals into large signals. Some minority carriers in base drift to the metallic contact of the base to produce base current, but the base current is much less than the collector current.

The transistor was invented in 1948 at Bell Telephone Laboratories by William Shockley, Walter Brattain, and John Bardeen. It rapidly replaced the vacuum tube: the problem with the vacuum tube was its large size (about 1/3 of the size of a light bulb) and its heavy consumption of electricity due to its internal heater. Soon after its introduction, many engineers believed that the transistor was a low power, low frequency device whose application was limited. Now we know that circuits using transistors can be designed for high power, and that transistors are capable of handling higher frequencies than the vacuum tube.
The usual way of drawing a transistor is shown in Figure 13.9. The npn transistor has an arrow pointed out of the emitter. The pnp transistor, constructed with n-type material sandwiched between p-type materials, has its emitter arrow pointed inward.

A simple amplifying circuit using an NPN transistor is shown in Figure 13.10. It does not take much energy for current to flow into the base when the base emitter junction is forward biased. The small input signal produces a large current variation in the emitter and so a large current change in the collector. This current produces a large voltage change across the load resistor. Thus, a small input signal becomes a large output signal. Power is drawn from the power supply $V_s$ to amplify the signal.
13.4 The Field Effect Transistor

A simpler and faster type of transistor is the FET — field effect transistor. It works by restricting the conducting channel in semiconductor material by the bias voltage on the gate, which produces a non-conducting region, shown in Figure 13.11 as the shaded region below the gate, through which no current can flow from the source to the drain. It is much like the effect of a faucet on the flow of water in a pipe. The transistor invented at Bell Labs was a bipolar device. The FET is a unipolar device and it and its variants are the workhorses of the modern computer chip.
The standard for spacing of components on a modern computer chip is the half-pitch, which is half the distance between identical features in an array. It was first used to define spacings on dynamic memories in 2008, and by 2012 was adopted to describe processor chips.

In semiconductor work, smaller is faster. The half-space in 2012 was 22nm (nanometers, where a nanometer is $10^{-9}$m, $10^{-3}$μ). The Broadwell chips from Intel have 14nm half-space. There is a lower limit to the half-space: Make it too small and electrons flowing from the source to the drain can flow to the gate. Individual features can be considerably less than the half-pitch. On an Intel 14nm chip the insulation thickness protecting the gate may be as small as 0.5nm. This is approaching the space occupied by two silicon atoms \[www.economist.com/news/science-and-technology/2165205, \text{May 26, 2015}\].

We are packing more transistors onto a single chip: there are indications that the number is over five billion. Perhaps we are approaching the limit when Moore’s law no longer works. For example, the cost per transistor has been rising recently, blamed in part by the expensive equipment needed to manufacture these tiny transistors, and also because of the decreasing yields [op. cit.].

Some believe that gallium arsenide will replace silicon as the crystalline basis for the semiconductor of the future, since it is faster and takes less power; power is becoming a dominant consideration with modern chips since they have so many transistors in a tiny piece of real estate that must be cooled. Some are excited with the promise of perovskites. New semiconductor devices are being invented at a rate that makes one wonder if they can ever catch up to the fast-paced technology. However, in this era of specialization, manufacturing engineers will let the specialists determine new frontiers of science. The job of the manufacturing engineer is to apply proven technology to make products.

### 13.5 Photovoltaic Cells

Most semiconductor devices are grown from a single crystal of silicon from the top down. Thus, a P-N diode is made by doping the pure silicon wafer to make it N-type, then doping the top to make it P-type.
A photodiode is made in this way, as shown in Figure 13.12. Metallic conductors are fused to the silicon, with the anode on the P-type material and the cathode on the N-type. The thickness $x$ of the P-type material is small compared to normal diodes, to permit light penetration. Distance $x$ also controls the response characteristics. When a photon (bundle of light energy) penetrates the P-layer, it is absorbed by forming an electron-hole pair. The electron migrates to the N region, and the hole to the P region, producing current when conductors are attached to the anode and cathode to complete a circuit. This is the photovoltaic effect.

![Figure 13.12: The Photodiode](image)

Photodiodes have been used extensively for decades in low power applications such as solar powered watches and calculators, and most modern cameras use photocells to estimate light levels. Remote regions use banks of photocells to control traffic lights and even provide electrical power to buildings. However, the efficiency of the early photovoltaic cell is low, less than 10%, and this inhibited its use as an alternative energy source for the electrical power utilities. Now, with increased efficiency, and a dramatic reduction in cost, the calculus has changed.

### 13.6 Energy in Photons

Light energy is described as bundles of energy called photons. Photons are wavelength specific, and the energy $E$ in a photon is given by $E = \frac{hc}{\lambda}$, where $\lambda$ is the wavelength of the photon, $c = 2.998 \times 10^8$ m/sec is the speed of light, and $h = 6.626 \times 10^{-34}$ J·sec is Planck’s constant; $hc = 1.99 \times 10^{-25}$ J·m; see J.L. Gray, “The Physics of the Solar Cell,” in *Handbook of Photovoltaic Science and Engineering*, edited by A. Luque and S. Hegedus, John Wiley & Sons, 2003. The energy in a photon as a function of wavelength is shown in Figure 13.13.

In semiconductor parlance, one uses eV, electron volts, to describe energy. 1 eV is the energy required to raise an electron through 1 V, so 1 eV $\equiv 1.602 \times 10^{-19}$ J and $hc=1.99\times10^{-25}1.602\times10^{-19}=1.24\times10^{-6}$ eV·m=1.24 eV·μ, resulting in $E=1.24\lambda$ where $E$ is in electron volts and $\lambda$ is in microns.
Silicon has a bandgap of 1.11 eV, so in order for a photon to create a hole/electron pair, it must have a wavelength no greater than

\[ \lambda = \frac{hc}{E} = 6.626 \times 10^{-34} \times 2.998 \times 10^8 \times 1.12 \times 1.6 \times 10^{-19} = 1.11 \times 10^{-6} \] (13.1)

in meters. Integrating Planck’s law from zero to 1.11 μ shows that the amount of solar energy at or below 1.11 μ is 78.7% of the total solar energy. Silicon was and is the favorite choice for transistor substrates, and now we can see that it is the favorite choice, at the present time, for PV substrates.

![Graph showing the relationship between electron voltage and wavelength](image1)

**Figure 13.13:** Electron Voltage of Light as a Function of Wavelength

![Graph showing the relationship between band gap and maximum wavelength](image2)

**Figure 13.14:** Maximum Wavelength

Shown in **Figure 13.14** is the relationship between the maximum wavelength for a photon to boost an electron into the conduction band. Also shown in this figure are the band gaps for...
some PV materials. Germanium (Ge) has 0.67 eV, Silicon (Si) at 1.11 eV, gallium arsenide (GaAs) has 1.43 eV, and amorphous silicon (aSi) has 1.7 eV.

A photovoltaic diode, whose standard symbol is shown in Figure 13.15, can convert a photon to electricity only if the energy in that photon is greater than the bandgap of the photocell material. For example, the band gap of silicon is 1.1 eV (electron volt). If the photon energy exactly matches the bandgap, then all the energy in that photon is converted to electricity. If the photon energy is less, then all of its energy is converted to heat. The photon energy above the band gap is converted to heat.

When a photon hits a PV cell, it can be reflected and its energy lost. It can also be absorbed in the semiconductor, and if its energy in electron volts is sufficient to knock an electron or a hole into the conduction band, then electricity is generated. Further, if the photon energy is greater than the bandgap, the surplus energy is released as heat. It is also possible for the photon to have energy at multiples of the bandgap, with the result that multiple releases of electrons or holes results — this is an unlikely occurrence.

A first analysis could lead one to recommend a material with a low band gap, but this is a mistake. The output power is the product of the bandgap voltage and the current generated, and studies have shown that cells with energy gaps of 1.4 to 1.5 eV have the highest efficiencies.

13.7 Responses of First through Third Generation PV Cells

A typical photovoltaic cell generates about 0.5 volts, so in order to produce a useable output, 34 → 36 cells are strung in series, producing an open circuit voltage of about 18 volts. However, the voltage that produces maximum power is 14 → 16 volts, suitable to charge a standard 12-volt battery. The characteristics of the array of cells is shown in Figure 13.16. The current as a function of voltage is essentially flat from zero volts to about 14 volts, and then the current sharply drops off. However, maximum power occurs on the knee of the drop off. The power curve shows this clearly, with the peak power of 60 watts occurring at \( \nu_{mp} = 16 \) volts. Notice the open circuit voltage \( \nu_{oc} = 19 \).

The fill factor of a PV cell is defined as maximum obtainable power \( \nu_{oc} \times \nu_{isc} \). A typical cell has a fill factor >0.7. A high fill factor means low series resistance and high shunt resistance, with the result that there is low internal heat losses.
The upper limit of the efficiency of a single junction photocell under non-concentrating conditions is 33.7% for a cell with bandgap 1.1 eV, as for silicon; this is the Schockley-Queisser limit [W. Schockley and H.J. Queisser, “Detailed Balance Limit of Efficiency of p-n Junction Solar Cells,” *J. Applied Physics*, vol. 32, pp. 510–519, March 1961]. This limit is not reached by any single junction photocell; in practice, typical PV panels have efficiencies of about 15%, but can reach 22%.

The normalized responses for some of these materials is given in Figure 13.18 [op. cit.]. Notice that when the bandgap is large, the wavelength of the peak response is short.
Silicon permits some light to penetrate based on the wavelength of that light. The depth of penetration in microns on a logarithmic scale as a function of wavelength in microns is shown in Figure 13.19, with minimum penetration at the short wavelengths and maximum penetration at the long wavelengths.

Comparing Figures 13.13 and 13.19 leads to the suggestion that cell material that absorbs high eV light be placed on top of cell material that absorbs low eV light. In this way the more incoming solar radiation can be captured. Ultimately, a multilayer stack of photocells with
successively lower band gaps could capture most of the Sun’s energy. This is the strategy being investigated for the third generation of photovoltaic panels.

A common way of describing the effectiveness of a photocell to generate electricity is with QE, quantum efficiency, which is the percentage of incoming photons that are successful in knocking an electron or a hole into the conduction band and so generate electricity. QE is measured in electrons per photon. Another way of describing efficiency is SE, spectral response, measured as the ratio of current $A$ in amperes generated in the photocell to the incident solar power on the cell in watts $W$; the unit of SE is $A/W$.

![Figure 13.20: Ideal and Typical Response for a Photocell](image)

The ideal response of spectral response in $A/W$ with respect to wavelength $\mu$ is the triangular response shown in Figure 13.20. The upper limit occurs at $\lambda = h\nu/E_g$ where $E_g$ is the bandgap electron voltage of the semiconductor. The peak of the triangle occurs at the transmissivity of the medium above the photocell — typically glass. For a single layer of glass, this transmissivity is typically 0.86. Also shown in Figure 13.20 is the response of a typical silicon photocell with bandgap of 1.12 eV. Since the cutoff is at 1.11 $\mu$, then all solar energy with wavelengths above 1.11 $\mu$ is rejected, and this amounts to about 22%. The relation between QE and SE is $QE = SE\lambda \times h\nu E = 1240\lambda SE \mu W/A$.

The spectral responses of typical PV cells are shown in Figure 13.21 [H. Field, “Solar Cell Spectral Response Measurement Errors Related to Spectral Band Width and Chopped Light Waveform,” NREL Report, 26th IEEE Photovoltaic Specialists Conference, Anaheim, California, Sept. 29–Oct. 3, 1997]. The code for the cell types is:

A GainP; B a-Si; C CdRe; D GaAs; E InP; F multi-Si; G mono-Si; H ZnO/CIGS
The ideal characteristics of some PV cells are given in Figure 13.22. GaAs is the clear winner as far as efficiency is concerned.

Multilayer PV cells are called third generation. The first generation cells are the single bandgap p-n junctions; their cost is high, and their efficiency hits the Schockley-Queisser limit. The second generation cells are also single bandgap devices, so they hit the same efficiency limit as the first generation cells, but substantially reduce cost by using thin film technology. Third generation cells break through the single bandgap limit, but since they are a laboratory novelty rather than an established manufacturing process, they are extremely expensive at present.
However, there is no doubt that Moore’s law applies to PV cells, just as it does to integrated circuits. Gordon Moore, one of the co-founders of Intel Corporation, the largest processor chip manufacturer in the world, predicted that computing power would double every 18 months, or the cost of a computing bit would be halved every 18 months (these two are equivalent). Therefore, expect affordable multilayer PV arrays by or before 2022.
Consider a three layer PV system with the top layer having bandgap $eV_1$, and subsequent layers $eV_2$ and $eV_3$, where $eV_1 > eV_2 > eV_3$. The PV is receiving three photons, $A$, $B$, and $C$, with energies $eVA$, $eVB$, and $eVC$, respectively, where $eVA \geq eV_1$, $eVB \geq eV_2$, and $eVC \geq eV_3$. The absorptance of the photons by the individual layers is shown in 13.23.

The University of New South Wales showed a cost/efficiency projection for all three generations of PV cells [GCEP Symposium, Stanford University, October 2008], and this is shown in Figure 13.24. The efficiency projections are obtainable in the short term, but the cost projection is a little speculative at this time.

### 13.8 Perovskite Solar Cells

Will we ever have a hydrogen economy? Certainly not with the present means of generating hydrogen, which is to mix methane and steam at high temperature to produce $H_2$ and $CO_2$. A more promising approach is to use PV to split water into hydrogen and oxygen, a process called photocatalytic decomposition. What is needed is a PV system that has a band gap in excess of 1.23 V, something that single-junction silicon PV cannot do. Multiple cells can be stacked, but the process is very expensive and there is no way to scale it up to produce meaningful quantities of $H_2$. This is where there is a growing sense in the scientific community that perovskite solar cells are the answer. They have a bandgap voltage of 1.5 V, enough to split water.

Perovskite solar cells are a potential game changer in the fight against $CO_2$ emissions into the atmosphere and its central role in global warming. And the reason is not its cost, efficiency, or longevity, even though cost is modest, efficiency is fast rising, and longevity is rapidly improving: It is its potential to produce hydrogen from water by electrolysis.
Perovskite solar cells have been around since about 2009. They initially had a disappointing efficiency of less than 4%, and an even more disappointing service life of a few hours. Now efficiencies of close to 20% are reported, and new manufacturing techniques have shown stability after 2000 hours of service.

The main problem with PV generated electrical power as it stands today is storage — there are few options for bulk storage of electrical energy, so the reality is to use it or lose it. Another problem is that some parts of the world where solar energy is abundant are distant from population centers, and passing electricity over long distances via power lines is not efficient. However, hydrogen can be stored in tanks and transported efficiently by water or rail.

“Michael Grätzel at the Swiss Federal Institute of Technology in Lausanne (EPFL), Switzerland, one of the pioneers of perovskite cells, has now shown, along with his colleagues, that using just two cells in parallel is enough to start the hydrogen fizzing out of water. Their research was published this week in Science. His team connected the cells to cheap, efficient electrocatalysts of their own design that use the solar power to split water. One electrode delivers electrons to water, splitting the molecules into hydrogen gas and hydroxide anions. The other electrode uses those hydroxide anions to produce oxygen gas and release more electrons back to the circuit [“Cheap Solar Cells Offer Hydrogen Hope,” Mark Peplow, IEEE Spectrum, September 25, 2014].”

The recent interest and publicity surrounding perovskites is justified because of three factors: 1. The astounding increase over a few years in energy conversion efficiency from photon to electrical generation.
2. The low cost of the native material, and its possibility of becoming a spray-on-glass PV panel.
3. The high band-gap voltage that could enable perovskites to be the only carbon neutral method for generating hydrogen from water.

13.9 Building Integrated Photovoltaics

We have seen that solar thermal, be it passive solar or active solar with fluid-carrying flat plate collectors, has been successfully integrated into the fabric of a residence with significant cost savings. Now we are approaching the point where BIPV (Building Integrated PhotoVoltaic) is poised to become the accepted system, rather than the standard BAPV (Building Applied PhotoVoltaic). PIPV is part of the structure. On a roof it acts to keep the weather out.

BIPV has two advantages over PAPV. The first is economic, the most important driver in construction decisions. The second is aesthetics. Until recently the BIPV advantages have been in commercial installations, in particular for roofing, curtain walls (meaning vertical glass), and skylights. Semi-transparent wall facades can work in a commercial building, but are not likely to be accepted in a residence.

The standard residential PV installation is rack mounted. First, an array of columns are affixed through the roof shingles into the rafters, onto which a metal grid is set. The columns are short, so that from a distance it makes the complete array look unobtrusive. The rack is designed to clip standard sized PV panels on top. The columns permit air flow beneath the PV panels that help to cool them and make them operate more efficiently; see the next section on temperature sensitivity. BIPV systems restrict air flow below them, and will tend to run hotter than a rack mounted system.

BIPV must satisfy two qualification and design standards. As an appliance it must satisfy bodies such as the Underwriters Laboratory, and as a structural element it must satisfy the International Code Council, whose building codes have been adopted by all 50 U.S. states “as a roofing material that dictates its performance on stability, wind resistance, durability, and fire safety [T. Lowder, “The Challenges of Building-Integrated Photovoltaics,” www.renewableenergyworld.com/articles/2012/05/the-challenges-building-integrated-photovoltaics, May 8, 2012].”

The big game changer in BIPV was the announcement in 2015 by Dow Chemical Company that in early 2016 they would launch a PV shingle system called POWERHOUSE™ Solar System 2.0, known as PH 2.0 for short, to select markets in the United States. With size and shape like a regular asphalt roof shingle, it is nailed to the plywood sheathed roof the same way.
The vital statistics of the PH 2.0 are:

- peak power 39 Wp
- reveal height 13.3”
- reveal width 41.6”
- thickness 0.7”
- weight 18.5 lbs

PH 2.0 is an evolution of an earlier BIPV shingle, and a substantial improvement. The system is certified by Underwriters Laboratories to withstand rain, hail, and wind, and comes with a 20-year warranty, backed by Dow Chemical [www.dowpowerhouse.com/news/2015/20150812a].

A rule of thumb for conventional PV panels is that they generate 10Wp/ft². Dow shingles produce 12Wp.ft² of PV generating area, which is a little less than the revealed area of the shingle. Thus, we can assume that the generating capacity of Dow POWERHOUSE™ Solar System 2.0 shingles is about the same as conventional PV panels.

PH 2.0 uses state-of-the-art copper indium gallium selenide solar cells. One of its best attributes is that individual shingles snap together to automatically make their own electrical connections. “A typical residential cluster of 350 solar shingles on a roof could slash one’s household electric bill by 40–60 percent. Such an installation can cost a homeowner over $20,000 [www.scientificamerican.com/article/im-getting-my-roof-redone-and-heard-about-solar-shingles].”

Other BIPV roof mounted systems are manufactured. For example, Certainteed, a St. Gobain company, offers a module system called Apollo II, with 14 monocrystalline silicon cells/module. Luma Resources has a 54”×16”×2” module that is mounted on an existing roof so that air flows beneath it.

The aesthetic appeal of BIPV can possibly overcome the prejudice in some jurisdictions regarding PV arrays. For example, Florida v PV - check e-mails to Michelle Putko

### 13.10 Temperature Sensitivity

The efficiency of a present-day photovoltaic cell is temperature sensitive. Above 25°C, electrical output drops by up to 0.5% for every degree rise. A panel at 60°C will generate about 17% less than a panel at 25°C.

One company, Massachusetts-based Sundrum, working closely with photovoltaic system installers, has a water-carrying flat-panel system that sits behind the photovoltaics and acts to cool them. Further, the warmed water is used to preheat the domestic hot water. The effectiveness of cooling the photovoltaic panels can be seen from Sundrum data taken on July 25, 2009 as shown in Figure 13.25. The increase in electrical power generated is measurable,
but the big winner is the amount of thermal energy captured to preheat the domestic hot water supply.

Figure 13.25: Increase in Electric Production by Cooling

13.11 Concentrating Photovoltaic Arrays

Since 1998, a dozen or more startup companies using methods of concentrating sunlight on photovoltaic cells have been formed. A considerable amount of venture capital has been poured into this sector. The concentrating methods include mirrors, lenses, and parabolic reflectors, and they all need orientation mechanisms, either one axis for the parabolic trough reflectors (see Chapter 13) or two axes. Every one axis concentrator has a horizontal east-west axis and tracks the Sun’s elevation.

As was shown in Chapter 5, the maximum solar elevation at winter solstice for Boston is about 23°, and as discussed in Section VIII of Chapter 9, the elevation should be commonly above 5°, so the angular elevation change is 18°.

Two-axis orientation is required for high concentrations, one for elevation, and the other for azimuth. Over a collection day the azimuth angle changes far more than elevation angle, typically at least ±60°.
The problem with using a convex lens to concentrate the Sun’s rays on a photovoltaic cell is weight. Glass is heavy. The way around this is to use a Fresnel lens. In Figure 13.26 is a standard convex lens and its Fresnel equivalent. Figure 13.27 shows a Fresnel lens focusing the incoming light onto a photocell. The intensity of light on the cell would result in excessive heating, which must be alleviated with a heat sink.

![Figure 13.26: The Fresel Lens](image)

![Figure 13.27: Fresnel Lens Concentrator](image)

The efficiency of a CSC is considerably higher than the same cell under direct sunlight. How can this be? If photons require a high enough eV to knock an electron from its valence position, and photons work in isolation, then CSC should not produce an enhanced efficiency. However,
if under high concentrations photons can “gang up” on an electron by combining their eV, then concentration has distinct advantages as far as efficiency is concerned.
Chapter 14

Smart Grids, FiTs, and Net Metering

It probably happens in all parts of the inhabitable world. We endure the mistakes of prior generations. Many mistakes were epic, such as two world wars in the twentieth century, genocidal campaigns either ethnic or religious-based, and state sponsored barbarism and terrorism. Edward Gibbon, in his classic book, *Decline and Fall of the Roman Empire*, probably put it best:

“History is indeed little more than the register of the crimes, follies, and misfortunes of mankind.”

We are the cause of climate change that threatens our planet, but steps to counter this are not sufficient.

We often do not realize the mistakes we have made until their effect is difficult to correct. For example, containing the Mississippi River with levies carries the silt and nutrient laden water far out to sea, thus starving Louisiana of the materials necessary to prevent encroachment of salt water into the bayous. With sea level rising due to global warming, much of lower Louisiana will end up under salt water.

We build roads that are obsolete soon after their completion. We built bridges to carry vehicles of the horse and buggy age. A local town decided to rebuild a bridge over a major highway, one year before that highway was scheduled to be expanded; that one year old bridge was then destroyed and a new bridge constructed over the highway.

The power distribution grid began in the horse and buggy era, and has not adapted to the existing reality. Even worse, there is too much resistance on the part of the power distribution industry to accepting more and more electrical power from renewable sources. This is the subject of this chapter.

There is an alphabet soup of acronyms and abbreviations associated with electrical power generation. Some of them are:

ANSI — American National Standards Institute
14.1 History of Electrical Utilities

Electrical power distribution occurred randomly and erratically. Some entrepreneurs won, but most lost, and their efforts are lost to history. This is true of all business enterprises: Of ten businesses started in the United States, within two years seven are gone, two are the walking dead, and one succeeds. Is it worth it for society? Most definitely. That set of the one-in-ten winners becomes a major engine for growth. As regards electrification, when the winning strategy became clear the advanced world made it a necessity for all its citizens.

A German immigrant named Alfred Dolge invested and ran some factories in Brock-ett’s Bridge, New York in the mid nineteenth century. Like most factories of its time, they were located to take advantage of water power. In 1980 he installed a water powered dynamo to provide electricity to his factories [www.edisontechcenter.org/Dolgeville].

Charles Brush (1849–1929), began experimenting and inventing at a young age. By the time he graduated from high school he built his first arc light. He received his degree from the University of Michigan in engineering. In the 1870s he improved on a German dynamo design and was awarded U.S. patent 189,997. The Brush Electric Company got into the business of providing arc lamps, often powered by Brush dynamos.

One of Brush’s customers was the California Electric Company (now PG&E) which in 1879 used two DC generators to power arc lamps via transmission lines. This is probably the first instance of a utility selling electric power from a central plant. Brush then set up utility companies in New York, Boston, Montreal, and a number of other cities. By 1983 the streets of New York City were lit with 1500 arc lamps, supplied and powered by Brush dynamos [https://en.wikipedia.org/wiki/History_of_electric_power_transmission].
In 1882, led by Thomas Edison, the Edison Illuminating Company built the Pearl Street power station, which supplied 110 V DC power for incandescent lamps to 82 customers. Its range was within a circle of radius about 2 km. The power it supplied was cheaper than power from Brush’s dynamos. By 1884 the Pearl Street plant had 508 customers. It was also the world’s first co-generation plant, since it used waste steam to heat nearby buildings [https://en.wikipedia.org/wiki/Pearl_Street_Station].

The limitations of low voltage DC power soon became apparent. Power could only be efficiently distributed over short distances. Conversion of DC power to high voltage, transmitted to a point near the customer, then converted back to low voltage to feed buildings, is inefficient and expensive. Thomas Edison made it a hard fight with the advocates of AC power, but ultimately lost out to the transformer. The transformer enabled AC power to be efficiently transformed and transmitted long distances at high voltage, and so low current, and then efficiently transformed to low voltage close to the customer at the far end.

William Stanley (1858–1916), developed the first practical transformer. The center of Great Barrington, Massachusetts, was electrified by Stanley in 1886. He used a Siemens steam engine to drive an AC generator, and, of course, he used his own transformers. The system he built is very similar to what we use today [http://ethw.org/Milestones:Alternating_Current_Electrification, 1886].

Stanley was also instrumental in the success of Westinghouse Electric, headed by George Westinghouse, in the building of a hydroelectric plant at Niagara Falls that sent its power to Buffalo 20 miles away [www.edisontechcenter.org/WilliamStanley].

What was happening in the nineteenth century in Europe with regard to electrical utilities? Croatia, while part of the Austro-Hungarian Empire, constructed a DC power plant in 1880 or thereabouts. It constructed its first AC plant in 1895 using two 550 kW generators running at 42 Hz. The primary and secondary voltages were 3000 V and 110 V, respectively, and the power supplied 340 street lights and some houses [http://ethw.org/2007_IEEE_Conference_on_the_History_of_Electric_Power].

The center of the town of Godalming in the U.K. was lit with arc lamps in 1881, demonstrating England’s first public electricity supply. The lamps were activated by a Siemens DC generator, which was powered by a water wheel on the River Wey. At about the same time the town of Brighton in England installed AC power for the incandescent light bulbs of shop owners. The DC Godalming scheme failed, but the AC Brighton venture was a success to be duplicated all over the country [www.zum.de/whkmla/sp/0809/kyungmook/km2.html].

All successful business ideas attract competitors, and in the United States numerous utility companies sprang up following the Westinghouse model. The plants were hydroelectric powered, if available, or coal fired if not. It soon became apparent that size mattered, and the bigger the power capability the better [https://power2switch.com/blog/how-electricity-grew-up-a-brief-history-of-the-electrical-grid].
The early utility companies in the United States were private and independent. There was little or no cooperation between companies in the same city, so each company would string wires at considerable expense and unsightliness. Also, there was no incentive for these companies to service the rural customer.

This changed during the Great Depression. Franklin D. Roosevelt took office in 1933. In an attempt to grow the stalled economy, he instituted the New Deal, and a central part of this initiative was improving and expanding the national infrastructure. Hydroelectric projects, like the Hoover Dam, greatly increased the electrification of America, as did the formation of the publicly owned Tennessee Valley Authority. The Roosevelt Administration made rural electrification a national objective.

Further, the presence in a city of a number of independent operators was neither efficient nor effective, so by 1935 each electric company would be granted the exclusive right to generate power in a region provided everyone wanting power in that region was serviced. They became controlled utilities. One control put on them was the need to justify price increases in return for a specified profit margin. Those companies were essentially the same as the companies are today [https://power2switch.com/blog/how-electricity-grew-up-a-brief-history-of-the-electrical-grid].

There is no national electricity grid in the United States. Instead, the country is divided into three regions, the Eastern Interconnected System, the Western Interconnected System, and, living up to its name as the Lone Star State, the Texas Interconnected System. These grids also interact with our neighbors to the north and south, Canada and Mexico [http://burnanenergyjournal.com/power-grid-technology].

Europe and most other countries in the world have a secondary voltage between 220 and 240 volts, but the Americas and Japan have secondary voltages between 100 and 127 volts. Europe uses large, three phase transformers capable of handling 300 to 1000 kVA in going from primary to secondary voltage, so the secondary feeds a whole neighborhood. North American transformers are single phase units capable of handling only 25 to 50 KVA, so the secondary handles about 10 customers. Also with twice the voltage, the secondary can feed comparable loads over four times the distance, and the three phase permits double the length of the secondary, so the European secondary can run eight times the distance of the American secondary [http://electrical-engineering-portal.com/north-american-versus-european-distribution-systems].

Remote rural areas in Australia, New Zealand, South Africa, and parts of Canada use earth return, so they can operate with a single wire feed. This cannot be done when the energy density is high. As will be discussed in Section II, current to ground commonly means a fault has occurred. In fact, to prevent electrocution in kitchens, bathrooms, swimming pools and other areas that can be wet, electrical codes in Europe and the United States require all circuits in these areas to have ground fault interrupters that automatically shut off power whenever a current to ground is detected. The European RCD (Residual Current Device) will trip when the
fault to ground is 10-300 mA (milliamps), smaller than the typical currents in properly operating circuits, and operate within 25-40 milliseconds. The U.S. GFI (Ground Fault Interrupter) will trip with a ground fault of 4–6 Ma in 25 milliseconds.

14.2 The Power Factor Problem

AC (Alternating Current) has a sinusoidal waveform that repeats 50 times a second (50 Hz, as it does in Europe) or 60 times a second (60 Hz, as it does in America). Power is the product of voltage and current, but how can this be determined when both voltage and current are changing? The answer, in part, is to determine the RMS (Root Mean Square) value of voltage and current and multiply these together to get power. There is a caveat to this called power factor.

Given a sinusoidal voltage \( V = A \sin \omega t \), the RMS value over period \( T \) seconds is given by

\[
VRMS = \frac{1}{T} \int_{0}^{T} A \sin \omega t \, dt.
\]

Using the trigonometric identity \( \sin 2\omega t = 2 \sin \omega t \cos \omega t \), then

\[
VRMS = A \frac{1}{2} \left[ t - \frac{\sin 2\omega t}{2\omega} \right]_{0}^{T} = A \left[ \frac{1}{2} - \sin 2\omega T / 4\omega T \right]
\]

and for large \( T \) the second term is zero, so \( VRMS = A/2 \).

The power factor is the cosine of the angle between voltage and current. When in-phase, the power factor is 1. When out of phase by 90° the power factor is 0. Shown in Figure 14.1 are three sine waves. The blue wave has a peak value of 2 and an RMS (Root Mean Square) value of 1. The green wave has a peak value of unity and is in phase with the blue wave. The red wave lags the blue and green waves by \( \pi/4 \) radians, or 45°, and the power factor is
\[
\cos 45 = \frac{1}{2} = 0.7071; \text{ the power factor is the cosine of the phase angle between voltage and current.}
\]

Some old electric power meters would integrate the current, assume a constant voltage, and multiply the two together. Newer meters measure current and voltage, multiply the two together, and then integrate over time. If the power factor of the load is 1.0, so the voltage and current are in phase, the product of voltage and current is watts, duly recorded by the meter. If the power factor of the load is 0, then the angle between the voltage and the current is 90°, the product of the voltage and current is zero, and the meter records zero watts.

As an example, suppose the load current is 20 amps and is being delivered at 115 volts to a residence. If the power factor of the load was 1.0, the meter would see 20 x 115 = 2300 W or 2.3 kW, and would integrate this over time, so over a day the meter would have recorded 2.3 x 24 = 55.2 kWh. If the power factor was 0, the meter would record zero watts and 0 kWh. However, the current needed at the power station to be sent down the line produce the 0 kWh is the same as if that current was in phase with the voltage, so the pair of them produce useable power to the residence.

With capacitive loads the current leads the voltage. Few, if any, normal loads serviced by the utility company are capacitive. Most domestic loads are mildly inductive, but some industrial loads can be substantially inductive. Electrical motors are the dominant culprit. Utility companies set a lower limit on the power factor of the load drawn by a customer, usually at 0.95. The reason is that the power lines carry more current than they should for the billable kWh of power provided.

\[
\text{Figure 14.2: Real and Reactive Power}
\]

The effect of reactive power is illustrated in Figure 14.2. The power lines carry current for the apparent power while only delivering the real power.

An industrial customer whose load has a power factor below the 0.95 threshold will be charged more per kWh, or in some cases this customer could be required to install active capacitive loads to improve the overall power factor of the loads that they draw. Although it is common for residential loads to have power factors lower than 0.95, power companies ignore the problem, following the rubric that is too difficult, or contentious, to fix.

An analysis of the losses due to inductive loads that make the power factor fall below the acceptable level of 0.95 illustrates a problem for the power company [J. Weikert, “The Why of
Voltage Optimization, “Power System Engineering, www.nreca.coop/wpcontent/uploads/2013/08/TS_Volt_VAR_January2013.pdf]. Weikert considered a factory receiving 1.76 MW of real power at power factor 0.8. The substation 10 miles away is connected to the factory with 4/0 cable. The line voltages at the substation and the factory are 12.47 kV and 10.82 kV, respectively. The substation produces 2.5 MVA and 2.0 MW of real power and supplies 200 amps to the factory. Power loss along the line is a function of the resistance $R$ of the line and the current $I$ that it carries, the so-called $I^2R$ losses. By adding a capacitor bank to bring the power factor up to 0.95, the losses on the line drop considerably, and the substation can reduce its production voltage while still giving an acceptable voltage level at the factory. Thus, the MVA needed to be produced to feed the factory drops from 2.5 MVA to 2.05 MVA.

Like most utility companies, National Grid, in the U.S. Northeast, bills for the electricity that passes through its lines under two categories, supply services and delivery services. It is possible to contract with a vendor other than National Grid for supply services. Delivery services are the exclusive domain of National Grid, and in a May/June, 2015 bill they break these down into six categories:

- customer charge $10.00
- distribution charge at 4.557 c/kWh for the first 2,000 kWh, 6.329 for the rest
- transition charge at −0.154 c/kWh
- transmission charge at 2.276 c/kWh
- energy efficiency charge 0.987 c/kWh
- RE charge 0.05 c/kWh

Thus, the total delivery service charges amount to $10 plus 7.716 cents per kWh for the first 2,000 kWh, and 9.488 c/kWh for the rest.

If the customer uses 1 kWh for that month, they would be charged $10.77 for delivery services. If the customer uses 10,000 kWh for that month they would be charged $10.00 plus $154.32 for the first 2,000 kWh and $759.04 for the remaining 8,000 kWh, a grand total of $923.36 for delivery services. Wear and tear on the lines is not a function of the flow of electrons through those lines, so delivery costs should have a larger flat fee and a much smaller fee per kWh. Putting it bluntly, National Grid is gouging its customers.

14.3 Standard Connections from Grid to Residence

The responsibility of the U.S. customer regarding maintenance of overhead electrical service to a residence has changed in recent years, although most people were unaware of it happening. The customer would make a call to the utility company about a downed service feed, and the company would send a crew to fix the problem at no expense to the customer.
Now the customer is responsible for the weatherhead, the service entry cable from the weatherhead to the meter box, and the meter box itself. Does the company expect the homeowner to make the repairs themselves, or to find a contractor and pay that contractor? One would expect the company to do what they traditionally did before the era of exorbitant and inexplicable delivery service fees.

A typical service entry in National Grid territory is shown in Figure 14.3. The power is a split 115 V single phase. That is, it has two anti-phase live lines, so using these two wires gives 230 V to power high-energy-use devices such as electric stoves and water well pumps. For 115 V appliances, one of the live wires is used with the neutral. The ground is provided by a ground rod tied to the meter by a ground wire.

![Figure 14.3: A Typical Electrical Service Entry](image)

A braided cable attached to a solid anchor point on the sidewall of the house supports the three wire service from the nearest pole on the street. Under normal operation, no current flows to ground. The three wires pass under the weatherhead, which is at the top of a service mast that feeds these wires into the meter box. A necessary part of the connection of wires to the weatherhead are the drip loops.

### 14.4 Inverters and Their Electrical Connections
The power from the photovoltaic panel, either a single 60-watt panel, or more commonly larger arrays that generate several kilowatts, can be used off-line, as shown in Figure 14.4. Here, the panel(s) feed the battery, or set of batteries, through a forward biased diode. The purpose of the diode is to prevent power leaking back to the photovoltaic panel(s) at nighttime or during periods of low insolation. Since most appliances and lighting use alternating current, an inverter is used to convert the DC power to AC.

![Figure 14.4: Off-Grid Storage and Usage](image1)

Devices that converted AC to DC were historically called “converters” and therefore devices that did the reverse, converting DC to AC, were called “inverters.”

Power storage with batteries is problematic. The energy density of a battery is limited: see Chapter 16. Also, some batteries off-gas hydrogen and oxygen, creating a potentially explosive situation. The best strategy, one employed by almost all residential photovoltaic systems, is to connect the arrays to the power grid as shown in Figure 14.5.

![Figure 14.5: Grid Connected PV System](image2)
Since the inverter provides a barrier to back-feeding the panels, and there is no battery that could be drained, the diode is not needed here. There is a central distribution board to which all the residential loads are connected. A meter from the inverter records the power generated by the Sun. Another meter connects the utility-produced power to the distribution board. When more Sun power is available than is needed for the house, the distribution board sends power back down the power lines, essentially running the utility meter backwards. Some utilities give you back the same per kilowatt hour as they charge, so for an installation with utility power there is every incentive to go on-line; this is called net metering.

The load on the electric grid can be classified as motor load and non-motor load. A characteristic of a motor load is that it decreases by about 2% for every 1% drop in frequency. Non-motor loads are largely unaffected by frequency, so a useful rule of thumb is that the overall load drops by the same percentage as the frequency [A. Gardiner, “Frequency Control,” www.texasre.org/CPDL/Frequency%20Control_operations%20seminar%20(2)].

![Figure 14.6: Energy Load, Ontario, Canada](image)

Provided there is a balance between generation and load, the frequency on the grid will not change. However, load changes throughout the day. An example is the energy loads over a day for Ontario, Canada, for the year 2008 as shown in Figure 14.6. Four dates were chosen, July 10, January 10, April 10, and October 10. The lowest energy use occurs between 2 and 4 am on all four dates. The highest energy use occurs on July 10 a little after noon. The peak of energy use on January 10 occurs about 5 pm, and for October 10 a little after 7 pm. The flattest daily load occurs on April 10. “The demand for electricity varies throughout the day and across the seasons; smart grids can reduce these peaks and optimize system operation [www.iea.org/publications/freepublications/publication/smartgrids_roadmap].”

Utilities should offer off-peak pricing, as is done extensively in Britain. Hot water heaters can draw energy during night hours at greatly reduced rates. Clothes and dish washers should be able to be programmed to operate during off-peak hours. EVs should automatically charge at
night. Even better, the utility company should be able to control when power is fed to a particular EV, with the owner having the ability to override if needed. In other words, with the consent and support of the customer, the utility company can send its surplus power to domestic loads, and shut off power to certain domestic loads when overall power loads peak.

14.5 Smart Grids

The system developed by William Stanley in the last decades of the nineteenth century survives, as discussed in Section I, with little change today. We are now in a rapidly changing world, one in which RE is increasingly becoming a critical factor. The old system cannot cope. We will identify the major problems before we consider the possible solutions.

Electric motors of 115 Volts operate efficiently provided the voltage is within ±5% of this voltage. A voltage substantially outside this range, lower as well as higher, can result in overheating and reduced service life.

Typical voltage ranges within +5% and –8.33% of nominal as measured at the point of common coupling between the customer and utility company. Allowable frequency variations for 60 Hz service are within 59 to 61.5 Hz. Deviations outside these ranges will likely cause the utility company to shut off power to the affected branch.

There are what are called Balancing Authorities to manage the changing demand. In North America there are over 100 such authorities, which themselves are over-seen by Reliability Coordinators. “The relationship between Reliability Coordinators and Balancing Authorities is similar to that between air traffic controllers and pilots”

The British National Grid has an obligation to control the system frequency of 50.00 Hz within ±1%. They do this dynamically and non-dynamically. “Dynamic Frequency Response is a continuously provided service used to manage the normal second by second changes on the system, while Non Dynamic Frequency Response is usually a discrete service triggered at a defined frequency deviation [www.2.nationalgrid.com/uk/services/balancing-services/frequency-response].”

Britain has an interesting method called FCDM (frequency control demand management) for controlling the grid frequency. Customers under this system agree to accept interrupted power “for a 30 minute duration, where statistically interruptions are likely to occur between approximately ten to thirty times per annum. FCDM is required to manage large deviations in frequency which can be caused by, for example, the loss of significantly large generation [op. cit.].”
The sometimes unpredictable changes in load on an electric grid require a semi-instantaneous response that may be difficult to achieve. Following a report by Oak Ridge National Laboratory we illustrate the problem in Figure 14.7. The top graphic shows a changing load, the middle graphic the response in power generation, and the bottom graphic the resulting frequency response. There are three blips in the load and three responses in power generation. The generation response to the first blip is answered late, and the effect is a change in frequency, which is not desirable. The generation responses to the second and third blips in load are timely, with the result that there is no corresponding change in frequency [B. J. Kirby, J. Dyer, C. Martinez, R.A. Shoureshi, R. Guttromson, J. Dagle, and “Frequency Control Concerns In The North American Electric Power System,” web.ornl.gov/~webworks/cppr/y2001/rpt/ 116362].

With interconnected power generation from utility companies, when one company’s power goes down it can trigger a chain reaction. Often, a single event does not cause a problem. It is when two or more incidents occur and things get out of control.

The massive power outage in the American Northeast and Canada was caused by a system failure and operator error, the same deadly combination that led to the Three Mile Island nuclear incident. The date was August 1, 2003 and the time mid-afternoon on a hot day. Loads on the grid were high, and high-voltage transmission lines in Northern Ohio were so hot that they sagged down into the branches of trees below. The alarm should have automatically sounded, but a software bug prevented this. The operators at FirstEnergy Corporation failed to take appropriate action, and by 4:05 p.m. most of eight states and Southeast Canada were without power. 50 million people were without power, many for two or more days [http://mentalfloss.com/article/57769/12-biggest-electrical-blackouts-history].

A number of measures were taken after this incident, including the passage of regulations that carry substantial fines for noncompliance. There has been no similar power failure since 2003, but there are some warnings regarding the grid of the future with distributed power, particularly from solar and wind generation.
Some customers of an electric utility company many choose to have a back-up generator. By code, this can only be activated by a single master switch after all connections to the grid are severed. There is then no possibility that power from the generator goes to the grid and endangers line workers attempting to fix the fault. The generator is considered a power island, as is a grid connected PV system, and there must be anti-islanding measures in place so when grid power is interrupted, the islands are disconnected from the grid.

Some key phrases regarding inverters that connect DERs to the electrical grid are “ramp rate control,” “fault ride through,” “anti-islanding protection,” and “reactive power control,” also known as VAR (volt amperes reactive) control. We will discuss these next.

Utility companies have indicated a concern with the variability of RE connected to the grid. For example, a single PV array grid connected through an inverter to the grid is delivering 10 kW until a cloud passes over, and then its output drops to zero. If this is repeated synchronously over many such systems it may be difficult for the utility to respond. For this reason, what is being proposed is ramp rate control on such independent producers.

However, PV power producers are spread over a broad geographic area, so a cloud occluding one small area has little effect on the totality of PV power entering the grid.

The only way to provide ramp-rate control for an RE system is to have adequate battery storage. This could become unnecessarily expensive if the utility requires 100% compliance to a strict standard. “However it is significant that a small relaxation in the compliance requirement from 100% to 98% results in a significantly smaller battery system requirement, particularly for the large system. This corresponds to sizing the battery for all but the most difficult days [www.solon.com/us/Application_Sheets/IEEE_Energy_Storage_Ramp_Rate].”

A similar problem for the utility is how to handle a transient change in voltage or current. If it is allowed to persist without any action taken, it could mean ignoring a major problem until it is too late to recover. It could mean taking power off a transmission line for a minor problem. There are proposals to require distributed generators to provide an FRT (fault ride-through) capability. Like ramp-rate control, this means having a significant battery capability tied to the inverter.

Rather than responding to a minor transient on a power line by disconnecting from that line, a DER should have FRT (fault ride-through) capability. Such a capability is illustrated in Figure 14.8. Time is the abscissa and voltage the ordinate. The voltage output of the DER as a function of time is shown by the heavy red line. At time $t_1$, the line voltage drops, possibly to zero, while the DER continues to produce power to the line, typically between 15 and 25% of the standard line voltage until time $t_2$, where $t_2 - t_1$ 0.2 seconds. If the line voltage begins to recover as shown, the DER must stay connected and remain connected provided the line voltage reaches 75 to 90% of the line voltage. If this does not occur, so the line voltage is below this, and between 0.7 and 1.2 seconds have elapsed, then the DER may disconnect.
Islanding in utility parlance is the grid connection of independent power producers on the grid after grid power is down. This is said to cause a dangerous situation for utility workers trying to fix a fault who expect lines they are working on to be dead. For this reason, inverters are required to have anti-islanding capability. However, line workers must always assume that lines are live, and some studies have shown the danger to such workers is minimal. Further, with a tiny amount of power being generated by the independents, when the utility power is off, the DERs will have no capability to service the normal load and will automatically assume a fault and shut down. The real danger for line workers is when working on high tension lines, and these are never directly energized by the small RE producers.

What is a smart grid? In most cases this is spoken as a catch phrase with little meaning, a shopping list of generalizations with no specificity. For example, here is the list of attributes of a smart grid [http://energy.gov/oe/services/technology-development/smart-grid]:

* Self-healing from power disturbance events
* Enabling active participation by consumers in demand response
* Operating resiliently against physical and cyber attack
* Providing power quality for twenty-first-century needs
* Accommodating all generation and storage options
* Enabling new products, services, and markets
* Optimizing assets and operating efficiently

Much better, and from the same citation, is that “Smart grid” generally refers to a class of technology people are using to bring utility electricity delivery systems into the 21st century, using computer-based remote control and automation. These systems are made possible by two-way communication technology and computer processing that has been used for decades in other industries. They are beginning to be used on electricity networks, from the power plants and wind farms all the way to the consumers of electricity in homes and businesses. They offer many benefits to utilities and consumers — mostly seen in big improvements in energy efficiency on the electricity grid and in the energy users’ homes and offices [op. cit.].”
Modern communication technology should inform an automated system owned by the utility company about a potential problem before it becomes a real problem. It can also advise the customer on how to use power more effectively, such as water heating during off-peak hours at reduced rates.

An important study that came out of MIT in 2011 considers the evolution of the U.S. electric grid over the next two decades [“The Future of the Electric Grid: an Interdisciplinary MIT Study,” https://mitei.mit.edu/system/files/Electric_Grid_Full_Report.pdf, 2011]. It made a series of observations and recommendations. Among the observations, presented here without change, are:

* The U.S. does not have a comprehensive national electricity policy, and regulatory regimes differ substantially among states.
* The U.S. electric utility industry has historically devoted a very small fraction of its revenues to R&D, instead relying primarily on its suppliers for innovation.
* University power engineering programs have languished over the past several decades due to the increasing popularity of other electrical engineering subdisciplines and a lack of research funding to support graduate students.

Among the findings, also presented here without change, are:

* As a result of the layering of historical policy decisions and the lack of a comprehensive, shared vision of system structure or function, the U.S. electric power system today operates under a fragmented and often inconsistent policy regime.
* Data are not available to quantitatively and accurately assess the reliability of the U.S. electric grid, particularly its changes over time. However, what data are available indicate the reliability of the U.S. grid is in line with that of other developed countries.
* Devising and deploying mechanisms to provide incentives for investment in flexible generation and for operating flexibly within the system will become increasingly important as the penetrations of wind and solar generation increase.
* Efficiently increasing the penetration of grid-scale renewable generation while maintaining reliability will require modifications to power system design and operation.
* High penetration of distributed generation complicates the design and operation of distribution systems. Net metering provides a subsidy to distributed generation, and utilities have inadequate incentives to make investments necessary to accommodate it.
* Because of its aging workforce and the nature of emerging challenges, the electric utility industry faces a near-term shortage of skilled workers, particularly power engineers. While this problem has been widely recognized, it remains to be seen whether efforts to deal with it will prove adequate.
* New technologies have the potential to improve the reliability and efficiency of bulk power systems by enhancing operators’ ability to observe and control these systems. Technologies similarly can enhance distributions systems and make demand more responsive to real-time costs, but effective use of these technologies will require changes in regulatory policy.

The term “smart grid” has been used to refer to a wide variety of electric grid modernization
efforts and ideas over the past several years. While uses of the term vary throughout industry, government, and the public, it is perhaps best described as the expanded use of new communications, sensing, and control systems throughout all levels of the electric grid. Many industry websites have been created to try to make sense of the flood of “smart grid” ideas, concepts, and products originating from industry, organizations, and individuals. These websites include Smart-Grid.gov (www.smartgrid.gov), the Smart Grid Information Clearinghouse (www.sgiclearninghouse.org/), and IEEE’s Smart Grid site (http://smartgrid.ieee.org/).

The U.S. Title XIII of the Energy Independence and Security Act of 2007 established the development of the smart grid as national policy and identified it as a broad collection of ambitious goals:
1. Increased use of digital information and controls technology to improve reliability, security, and efficiency of the electric grid.
2. Dynamic optimization of grid operations and resources, with full cybersecurity.
3. Deployment and integration of distributed resources and generation, including renewable resources.
4. Development and incorporation of demand response, demand-side resources, and energy efficiency resources.
5. Deployment of “smart” technologies (real-time, automated, interactive technologies that optimize the physical operation of appliances and consumer devices) for metering, communications concerning grid operations and status, and distribution automation.
6. Integration of “smart” appliances and consumer devices.
7. Deployment and integration of advanced electricity storage and peak-shaving technologies, including plug-in electric and hybrid electric vehicles, and thermal-storage air conditioning.
8. Provision to consumers of timely information and control options.
9. Development of standards for communication and interoperability of appliances and equipment connected to the electric grid, including the infrastructure serving the grid.
10. Identification and lowering of unreasonable or unnecessary barriers to adoption of smart grid technologies, practices, and services.

As discussed earlier, most electrical loads are inductive, so as far as the utility company is concerned, it is desirable to encourage capacitive loads that can improve the overall power factor. Alternatively, it would be desirable for a power generator at the load end of the transmission line to produce power where the current leads the voltage. This is possible with the inverter, as discussed next.

“Georgia Power requires small solar generators to use advanced inverters to provision reactive power in exchange for compensation. Similarly, a group of Western utilities is working to make advanced inverters mandatory for all new solar facilities within their service territories. . . . A report by the Oak Ridge National Lab found that distributed voltage control significantly outperforms centralized voltage control. Reactive power suffers 8 to 20 times greater line losses than real power, and those losses increase as a line is more heavily loaded
Most inverters used by DERs have reactive generation capability. Using these would save utility companies significant expense. However, the reactive power produced results in loss of normally compensated power. There needs to be a national policy to require utility companies to provide adequate compensation for the reactive power from DARs.

“Enhanced inverter functionality is also desired to achieve FRT capability, so that PV and distributed renewable generators could contribute to grid stability during system disturbances where the grid voltage or frequency may go outside the normal operating ranges. Existing voltage trip settings prescribed in IEEE 1547 are conservative, forcing generators to trip off line quickly to avoid islanding. Conversely Germany has developed standards that require inverters to provide FRT capability and dynamic reactive support. In fact, Germany has determined that it is necessary to retrofit 315,000 existing inverters to achieve enhanced inverter functions including low voltage ride-through (LVRT) capability in order to provide grid support during faults and voltage recovery during post-fault conditions to avoid a system blackout. The cost of the retrofit is approximately $300,000,000. [“Inverter Technical Standards Proposal,” www.sdge.com/sites/default/files/documents/2060692059/Inverter%20Technical%20Standards%20White%20Paper%20August%207%202013.pdf, August 2013].”

“Ride-through may be defined as the ability of an electronic device to respond appropriately to a temporary fault in the distribution line to which the device is connected. A fault typically derives from conditions or events that are extreme or unanticipated and cause an unintended short-circuit [E. Malashenko, S. Appert, and W. alMukdad, “Advanced Inverter Technologies Report,” www.cpuc.ca.gov/NR/rdonlyres/6B8A077D-ABA8-449B-8DD4-CA5E3428D459/0/CPUC_AdvancedInverterReport_2013FINAL.pdf].”

14.6 The Future of Solar Electrical Generation

What is the future of electrical generation by photovoltaic means? Well, the future is here.

Until the present century, photovoltaic-generated electric power was a curiosity. It was used to power hand-held calculators and other small devices, but was too expensive to be considered as a substantial source of real power. Crystalline PV cells are first generation technology. We are now in second generation technology with thin-film PV: no more efficient than crystalline, but a lot less expensive.

The cost of the generation of electricity from photovoltaic panels is dropping rapidly. In 1974 the cost of the PV watt was about $75. In 2014 it had dropped to $0.75, and is destined to drop further. The so-called soft costs (the frame, support system, and glazing) are a significant
fraction of the overall cost, but these are dropping rapidly due to improved methodologies and, of course, competition.

With tax incentives and declining prices, and pay-back time frames as low as five years, homeowners are installing PV arrays on their roofs at an ever increasing rate. Most installations in the advanced world are grid connected, so the electric meter can run backwards when the Sun is shining.

The number of PV installations in Asia have increased exponentially. China, Japan, and Thailand have doubled the number of installations in 2014 alone. Contrast this to Europe and the United States where growth rates have stalled; we have yet to recover from the 2008 recession.

In places too remote to have electric power from a utility company, some homeowners have elected to install a standalone PV array. The problem they face is how to store the energy; the energy storage problem is the subject of Chapter 16.

Reliable electric power does not exist in many third world countries. In some cases there is no electrical power at all. The most desirable modern device is the cell phone, and the problem is how to get electrical power to charge it. Cell phones do not require much power, and small PV units are available at modest cost for the task.

Without a doubt, the future will see more and more PV installations, from those on residential roofs to massive fields of arrays that can generate megawatts. The next boost could come when third generation PV cells, with their high efficiency, become available and affordable. It could be with perovskites, either by electrical generation or by hydrogen generation.

The total output in GW in recent years of installed PV solar systems is shown in Figure 14.9. This growth will continue for a number of years until generation exceeds the ability to absorb it.

The annual PV solar installed in three recent years is given in Table 14.1. The leader in 2011 and 2012 was Germany, largely due to an overly generous governmental financial incentive, the FiT; see the next section. The reduction in 2013 occurred since this incentive was reduced because set goals were achieved. The leader is now China, and this is expected to continue into the foreseeable future. Japan and the United States are #2 and #3 respectively [http://solarcellcentral.com/markets_page.html].

The astounding decline in the per watt cost of a PV cell in recent years is shown in Figure 14.10. Much of this is due to overcapacity. PV manufacturers saw the major increase in demand of 172% in 2010, following the recession year of 2009, and added production capacity so as not to be left behind. There were plenty of warnings regarding overcapacity, which were ignored. As a result, by 2011 the price of crystalline silicon solar cells fell by about 60% as manufacturers tried to unload their inventories. Some manufacturers filed for
bankruptcy [http://solarcellcentral.com/markets_page.html]. There is little chance that prices will rebound.

![Figure 14.9: Installed PV Output](image1)

![Figure 14.10: Cost per Watt of a PV Cell](image2)

“Although the multiple applications of photovoltaics make it difficult to compare system costs, data from California’s Sacramento Municipal Utility District’s (SMUD’s) Photovoltaic Program offers one example of cost reduction. The SMUD PV Program promotes a range of photovoltaic applications including grid-connected roof-mounted systems (SMUD’s PV Pioneer Program) and grid-connected ground-mounted installations at substations and parking lots. Each year, SMUD issues a request for proposals to provide the installed roof- and ground-mounted systems on a turnkey basis [E.S. Piscitello and V.S. Bogach, “Financial Incentives for Renewable Energy Development,” Proceedings of an International Workshop, February 17–21, 1997, Amsterdam, Netherlands, World Bank Discussion Paper No. 391].”
Table 14.1: Installed PB in Recent Years

<table>
<thead>
<tr>
<th>country</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>2.8</td>
<td>3.5</td>
<td>12.9</td>
</tr>
<tr>
<td>Japan</td>
<td>1.3</td>
<td>2.0</td>
<td>6.9</td>
</tr>
<tr>
<td>United States</td>
<td>1.9</td>
<td>3.4</td>
<td>4.8</td>
</tr>
<tr>
<td>Germany</td>
<td>7.5</td>
<td>7.6</td>
<td>3.3</td>
</tr>
<tr>
<td>Italy</td>
<td>6.7</td>
<td>3.8</td>
<td>1.4</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0.8</td>
<td>0.9</td>
<td>1.5</td>
</tr>
<tr>
<td>India</td>
<td>0.6</td>
<td>1.0</td>
<td>1.1</td>
</tr>
<tr>
<td>Australia</td>
<td>0.8</td>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td>France</td>
<td>1.8</td>
<td>1.1</td>
<td>0.6</td>
</tr>
<tr>
<td>ROTW</td>
<td>3.5</td>
<td>6.0</td>
<td>6.4</td>
</tr>
<tr>
<td>TOTAL</td>
<td>27.7</td>
<td>30.1</td>
<td>39.6</td>
</tr>
</tbody>
</table>

Table 14.2: Costs-Installation and per kWh

<table>
<thead>
<tr>
<th>year</th>
<th>cost</th>
<th>kWh cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>$8.78</td>
<td>25–26</td>
</tr>
<tr>
<td>1994</td>
<td>$7.13</td>
<td>20–21</td>
</tr>
<tr>
<td>1995</td>
<td>$6.87</td>
<td>19–20</td>
</tr>
<tr>
<td>1996</td>
<td>$6.21</td>
<td>17–18</td>
</tr>
<tr>
<td>1997</td>
<td>$5.34</td>
<td>15–16</td>
</tr>
<tr>
<td>1998</td>
<td>$5.07</td>
<td>14–15</td>
</tr>
<tr>
<td>1999</td>
<td>$4.50</td>
<td>13–14</td>
</tr>
<tr>
<td>2000</td>
<td>$4.00</td>
<td>11–12</td>
</tr>
<tr>
<td>2001</td>
<td>$3.42</td>
<td>10–11</td>
</tr>
<tr>
<td>2002</td>
<td>$3.18</td>
<td>9–10</td>
</tr>
<tr>
<td>2003</td>
<td>$2.98</td>
<td>8–9</td>
</tr>
</tbody>
</table>

“SMUD did not place much emphasis on the ‘high value’, but small and labor-intensive, ‘niche’ markets, preferring to reduce market costs for the majority of its customers [Donald W. Aitken, “Smud’s PV Program: Past, Present, and Future: Report to the Sacramento Municipal Utility District,” Final Report, Donald Aitken Associates, December 2000].” Extracted from that report is the data in Table 14.2. The cost is the total installed cost, which includes about 13% of SMUD-related costs. The final column is the 30-year projected energy cost per kWh, showing the dramatic decline in costs in recent years. Care should be exercised in relying on the data for recent years, since from 2000 on the data is projected.

SMUD is a Muni, a municipal utility, which is invested in PV. It will be interesting to see how it copes with the increasing percentage of grid-connected renewables. It is likely that most if not all the Munis like SMUD will adapt to the new reality and become power management companies.
The service life of a PV panel is 25-plus years. The output of these panels slowly degrades with time, so that after 25 years they produce only 80% of their initial capability. However, given the dramatic drop in the cost of a PV module in recent years, and with the promise of higher efficiency panels to come, sometime during that 25 year service life, the owner will probably think about replacing these panels.

The future will not be with single junction crystalline silicon. It could be in multi-junction silicon. It could be in organic PV. It could be in perovskite PV cells. The societal challenges in the future will be in the handling of the multitude of micro-producers of electricity, and the development of energy storage systems with capacity to match one or more days’ output of all the RE (Renewable Energy) producers combined.

The problem with electricity generated is that it must be used immediately. That is, at the present time, storage of a day’s worth or more of generating capacity is not an option. There are schemes for energy storage, but these have limited capacity. Flywheels have been tried, with limited success. When the terrain permits it, pumped hydroelectric storage has proven successful, as in Wales, but it can hold only a small fraction of the daily electrical needs of the nation. Considerable attention is being devoted to the improvement in the energy storage capabilities of batteries; there are great improvements in energy density of batteries, but there is a long, long way to go before battery storage becomes really useful.

Energy storage is the subject of Chapter 16.

### 14.7 FiTs and Net metering

A number of countries encourage the installation of PV on roofs of residences, of concentrated solar power, of wind energy—or some other type of renewable—with FiTs (feed in tariffs). The electricity generated by the RE system needs to be grid connected. The FiT program is run by the utility company supplying electrical power. They offer three benefits to the consumer:

* a generation tariff,
* an export tariff, and
* savings on the electric bill.

For a fixed period of time, usually 20 years, the utility company guarantees a fixed rate per kWh generated by the consumer’s RE system, and guarantees to pay the consumer a fixed rate for every kWh exported to the utility.
The average national electricity prices around the world in U.S. cents is shown in Figure 14.11 [http://shrinkthatfootprint.com/average-electricity-prices-kwh]. Since the FiT benefit is dependent on the electricity cost in the country where the RE system is being installed, it is evident that locking in the tariff in Denmark is much better than in France.

The German tariffs are distinct for each type of RE. For small installations these are:

- hydropower 12.52
- landfill gas, sewage gas, and mine gas 8.42
- biomass 13.66
- geothermal 25.20
- onshore wind 8.90
- offshore wind 15.40
- PV 12.88

The numbers after each category is the tariff paid out in 2014 in euro cents [www.germanenergyblog.de/?page_id=16379]. One wonders why there is such a variation in rates. Why is geothermal worth over double the average of the other categories? Why is offshore wind worth 73% more than onshore wind? The exchange rate in August 2015 was $1 = 88c euro.

The French government decided in 2011 to reduce the FiTs due to the unexpectedly high rate of PV connected to the grid. Grid-connected PV increased from 81 MW at the end of 2008 to 261 MW at the end of 2009 and reached 1025 MW on December 31, 2010. They had budgeted for 1100 MW by the end of 2012 and expected 5400 MW of photovoltaic energy by 2020.
The reduced FiTs in euro cents for residential buildings were 0 to 9 kW @ 46c, 9 to 26 kW @ 40.25c for BIPV (building integrated PV), with 0-36 kW @ 30.35c and 36–100 kW @ 28.83c for simplified BIPV.

In 2014 a further reduction occurred in February, with the BIPV rate reduced for 0–9 kW to 28.51c [www.pv-tech.org/news/french_energy_regulatory_commission_to_announce_new_tariff_rates].

The energy performance of an English and Welsh property is rated from A to G, where A is the best and G the worst. The rating is based on the energy used per square meter of floor area. The FiT in England and Wales is based on the property’s energy rating, and the best tariff goes to a property rated D or better [www.wmepec.co.uk/PDF/EPC%20Explained%20Final%202018%20July%202011].

The tariff also depends on the size of PV installation, with systems 4 kW or less getting the best rate. For the latter half of 2015, a structure rated in band A through D with a PV array generating 0 to 4 kW got 12.47p per kWh, and if it generated 4 to 10 kW it got 11.30p per kWh. A system generating 0 to 10 kW but in band E or below got 5.94p per kWh. A system generating more kWh than it consumes generates an export tariff of 4.85p per kWh. The duration of the tariff is 20 years, and the rate is guaranteed [www.which.co.uk/energy/creating-an-energy-saving-home/guides/feed-in-tariffs-explained/feed-in-tariff-savings-and-earnings].

Consider a 4 kW installation costing £6,500 on a south-facing roof at angle 30° to horizontal in Birmingham on a house rated in band D and registered in August 2015. It would get an annual generation tariff income of £441, an annual export tariff income of £83 (based on a deemed export rate of 50%), and an annual fuel-bill savings of £120. This amounts to a total of £643 a year [op. cit.].

If the property is rated below band D it would probably be beneficial to carry out energy efficiency improvements before applying for the 20-year FIT [http://tools.energysavingtrust.org.uk/Generating-energy/Getting-money-back/Feed-In-Tariffs-scheme-FITs/Energy-Performance-Certificates-and-the-Feed-in-Tariff].

Japan introduced a FiT in 2012 that guaranteed -?- 42/kWh, leading to 1.2 million applications, mostly for solar PV. Japan’s power utilities balked. Kyushu Electric Power in the sunny south had 72,000 applicants in 2014 who were trying to beat the deadline for a cut in the tariff to -?- 32/kWh. In response, the company cited reliability issues. The other utility companies followed suit. METI (the Ministry of Economy, Trade and Industry) is backing the utilities. Critics say METI should have learned from Germany’s mistakes and set the guaranteed tariff much lower. What is now in doubt is the goal of having renewables produce 20% of electricity by 2030 [D. McNeill for EURObiZ, http://www.japantoday.com/category/opinions/view/japans-feed-in-tariff-program-becomes-a-solar-shambles, February 14, 2015].
China announced in August 2011 that it would put in place a FiT of one RMB/kWh, $0.156/kWh in U.S. dollars, not particularly generous by European standards. At also announced a somewhat higher rate, but made the cut-off date prior to its August announcement [J. Gifford, www.pv-magazine.com/archive/articles/beitrag/china-introduces-fit-100004087/329/#axzz3jYPJJAPO, September 2011].

“Recent media reports quoting Chinese officials abreast with information about the solar power policy in the 13th Five Year Plan (2016–2020) state that financial support in the form of high tariffs and even incentives to equipment manufacturers would go by the end this decade [M. Chadha, http://cleantechnica.com/2014/10/17/china-phase-financial-support-solar-power-sector-2020, October 2014].” The objective is to reduce carbon intensity by 40–45% from 2005 levels by 2020.

Many countries use FiTs, including the advanced countries and the most populous. The problem with this type of incentive is that it encourages oversized systems that produce more energy than needed. The fixed purchase price in the FiT over 20 years is a long time for the utility company to pay out [http://solarcell_central.com/markets_page.html]. Germany is now phasing them out since they have reached the desired goal. The U.S. states with FiTs are California, Oregon, Washington, Vermont, and Maine.

In April 2015, since 98 MW of grid-connected PV electricity was added in the United Kingdom in January through March 2015 the DECC (Department of Energy and Climate Change) in the U.K. cut the FiT tariff by 28% for July–September 2015 to 0.062E/kWh, $0.067 USD [P. Tisheva, “UK standalone solar FiT to be cut by 28%,” http://renewables.seenews.com/news/uk-standalone-solar-fit-to-be-cut-by-28-473792, April 2015]. This will have a chilling effect on potential new customers for roof-top PV.

RPS (Renewable Portfolio Standards) are the requirement that a U.S. state gets a specified minimum percentage of its energy from renewables. For example, the RPS for California is 20% by the end of 2013, 25% by the end of 2016, and 33% by the end of 2020. Most of the contiguous U.S. states have RPSs. Those not participating are the southeastern states plus Idaho, Wyoming, and Nebraska [“Feed–in tariff: A policy tool encouraging deployment of renewable electricity technologies,” www.eia.gov/todayinenergy/detail.cfm?id=11471, May 30, 2013].

Most of the contiguous U.S. states require net metering: Net metering means that a grid-connected RE system gets back from the utility company the same per kWh as the utility company charges for its power. The exceptions are Mississippi, Alabama, Tennessee, and South Dakota, with no policy on net metering. Texas, South Carolina, and Idaho have voluntary standards.

The European countries projected to have the highest percentage of electricity produced by PV are given in Tables 14.3 and 14.4. The second column is the percentage of energy from renewable sources in final consumption of energy for 2005. The third column is the target for
the percentage of energy from renewable sources in final consumption of energy in 2020. The source of this data is EREC (the European Renewable Energy Council) [“RE Technology Roadmap: 20% by 2020,” www.erec.org/fileadmin/erec_docs/Documents/Publications/Renewable_Energy_Technology_ Roadmap.pdf]. It appears that Denmark, by generating 43% of all their electricity from renewables in 2012, has already surpassed the 2020 target.

Table 14.3: Percentage from PV

<table>
<thead>
<tr>
<th>country</th>
<th>2005</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>39.8</td>
<td>49</td>
</tr>
<tr>
<td>Latvia</td>
<td>34.9</td>
<td>42</td>
</tr>
<tr>
<td>Finland</td>
<td>28.5</td>
<td>38</td>
</tr>
<tr>
<td>Austria</td>
<td>23.3</td>
<td>34</td>
</tr>
<tr>
<td>Portugal</td>
<td>20.5</td>
<td>31</td>
</tr>
<tr>
<td>Denmark</td>
<td>17</td>
<td>30</td>
</tr>
<tr>
<td>Estonia</td>
<td>18</td>
<td>25</td>
</tr>
<tr>
<td>Slovenia</td>
<td>16</td>
<td>25</td>
</tr>
<tr>
<td>Romania</td>
<td>17.8</td>
<td>24</td>
</tr>
<tr>
<td>Lithuania</td>
<td>15</td>
<td>23</td>
</tr>
<tr>
<td>France</td>
<td>10.3</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 14.4: Percentage from PV Continued

<table>
<thead>
<tr>
<th>country</th>
<th>2005</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spain</td>
<td>8.7</td>
<td>20</td>
</tr>
<tr>
<td>Germany</td>
<td>5.8</td>
<td>18</td>
</tr>
<tr>
<td>Greece</td>
<td>6.9</td>
<td>18</td>
</tr>
<tr>
<td>Italy</td>
<td>5.2</td>
<td>17</td>
</tr>
<tr>
<td>Ireland</td>
<td>3.1</td>
<td>16</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>9.4</td>
<td>16</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1.3</td>
<td>15</td>
</tr>
<tr>
<td>Poland</td>
<td>7.2</td>
<td>15</td>
</tr>
<tr>
<td>Slovenia</td>
<td>6.7</td>
<td>14</td>
</tr>
<tr>
<td>Netherlands</td>
<td>2.4</td>
<td>14</td>
</tr>
</tbody>
</table>

“Created in 2000, EREC is the umbrella organization of the European renewable energy industry, trade and research associations active in the sectors of bioenergy, geothermal, ocean, small hydropower, solar electricity, solar thermal, and wind energy. EREC represents the entire renewable energy industry with an annual turnover of more than 40 billion Euros and more than 400.000 employees [op. cit.].”
EREC studied the field of renewables, determined their market size in 2006, and then projected out to 2020 as shown in Table 14.5. The last column gives the projected annual growth rate. In the column for 2006 the designation GWp means peak output in GW.

By 2012, 22% of electrical power in Germany was produced from renewables, predominantly from wind and rooftop PV. The average cost of power in Germany is about $0.387/kWh, about twice what it should be, caused in large part by the costs of FiTs. The decision to close nuclear facilities after the Fukushima Daiichi disaster in Japan is also part of the problem. The increased the burning of the soft coal lignite to cover the shortfall caused a spike in carbon emissions in 2012. Profits in the electric utility industry have plummeted [R. Elberg, “In Germany, Net metering Brings Unintended Consequences,” www.navigantresearch.com/blog/in-germany-net-metering-brings-unintended-consequences, October 18, 2013].

Table 14.5: Energy from Renewables

<table>
<thead>
<tr>
<th>category</th>
<th>2006</th>
<th>2020</th>
<th>annual growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>wind</td>
<td>47.7 GW</td>
<td>180</td>
<td>8.5%</td>
</tr>
<tr>
<td>hydroelectric</td>
<td>106.1 GW</td>
<td>120</td>
<td>0.8%</td>
</tr>
<tr>
<td>photovoltaic</td>
<td>3.2 GWp</td>
<td>150</td>
<td>23.6%</td>
</tr>
<tr>
<td>biomass</td>
<td>22.3 GW</td>
<td>50</td>
<td>5.2%</td>
</tr>
<tr>
<td>geothermal</td>
<td>0.7 GW</td>
<td>4</td>
<td>14.9%</td>
</tr>
<tr>
<td>solar thermal electric</td>
<td>-</td>
<td>15</td>
<td>31.1%</td>
</tr>
<tr>
<td>ocean</td>
<td>-</td>
<td>2.5</td>
<td>17.5%</td>
</tr>
</tbody>
</table>

It can be argued that the birth of FiTs began in the United States in 1978 when Congress passed PURPA (the Public Utilities Regulatory Policy Act), which allowed independent RE generators to connect to the grid. A tariff was set on what these independents would be paid, essentially creating the FiT [P. Gipe, “Time to Break Free of Net-Metering; We Need a ‘FiT’ Policy for Renewable Energy to Soar,” http://energyblog.nationalgeographic.com/2013/12/26/break-free-net-metering, December 2013].

The FiT is a big incentive to install PV on your roof, and in Europe in particular has been successful in bringing a significant percentage of renewables into the electrical generation equation. The United States has taken the route of net metering, and to date the amount of renewable energy being fed back to the grid is limited; more about this later in this section.

In 2006 the United States started a 30% RE investment tax credit. This tax credit ends in 2016. Some states in the Union have tax credits or other types of encouragement for RE installations. In 2011 the Commanwealth of Massachusetts provided a 17% tax credit for the installation of a 3.8 kW PV system on the house I built. The total tax credit, federal and state, amounted to 47%. In addition, Massachusetts provides SRECs (state renewable energy credits); when my PV
system generates 1000 kWh it earns 1 SREC, and that can be sold at auction, and typically brings in about $450.

The monthly output from my roof-mounted 3.8 kW PV is shown in Figure 14.12. One sees a strong seasonal component to this output. The mean monthly output is shown by the green dashed line, and is 413 kWh. This mean would be somewhat higher if it was not for the month of February 2015. The total production for this month was 6 kWh, and was due to the panels being covered with snow and ice.

![Figure 14.12: Monthly PV Output](image)

The winter of 2014–15 was possibly the most brutal ever in recorded history in the American Northeast.

Most consumer-owned electrical generation in the United States is grid connected, and the consumer enjoys net metering. Net metering means that a grid connected RE system gets back from the utility company the same per kWh as the utility company charges for its power. The price is not guaranteed, but the consumer effectively has the second and third benefit of the FiT.

The tax credits, the SRECs, and the net metering in the United States should be a guarantee of success in producing a significant percentage of renewables in electrical generation. The fact that this has not happened is due, in part, to program caps.

There are program caps on the amount of distributed RE the IOUs (investor owned utilities) are willing to accept tied to the grid. The most common program cap is based on a percentage
of the IOU peak demand, capacity, or load, based on a reference year, which is usually the previous year; 20 states have this cap, set typically at 5%. California limits the renewables to 5% percent of the non-coincident peak demand, which is the sum of individual customer peak demands. Some states have capped the renewables to a maximum number of MW: Maryland caps net metering renewables at 1500 MW, and New Hampshire at 50 MW.

The program caps are far too low. Under the FiT schemes, Europeans have far higher percentages of electricity from renewables. Raise the cap put on net metering and the problem is solved.

Already there are strong signals from electric utility companies that they will not continue with the expansion of net metering. A case in point can be found in Hawaii. Electricity is generated by the utility companies in Hawaii using fuel oil, all of which must be imported. The spike in fuel oil prices in 2012 meant that the average consumer in a moderate-sized house paid $250/month or more for electricity. This made roof-mounted PV very attractive, so now Hawaii leads the nation in grid-connected PV. In some locations of Hawaii the renewables generate up to 20% of the peak energy demand.

In September 2013, HECO (Hawaiian Electric Company) refused to issue permits to grid-connected PV on any circuit with 120% or more of the minimum daytime load generated by renewables. This meant that about 50% of residents trying to connect were not permitted. HECO argued that the excess power would cause voltage spikes that would damage equipment. However, a study by NREL (the National Renewable Energy Laboratory) showed otherwise, and the state intervened and required HECO to issue permits [A.C. Mulkern, “Hawaii utility commits to highest-in-nation level for rooftop solar on circuits, but can grid handle it?” E&E Publishing, www.eenews.net/stories/1060016274, April 2015].

It is evident to a number of analysts that the business model of the IOU is out of date, and is unsustainable. They are trying to figure out how to survive and stay profitable. It is not the duty of government to assist them in this. Instead, government should see a future in which renewables have an ever-increasing share of energy production in all sectors of the economy in order to reduce our carbon footprint to sustainable levels. If the IOUs cannot do this, they should be pushed aside.

Already, some IOUs, such as National Grid in the American Northeast, are splitting their business into two parts — power generation and delivery service. Looking at costs in recent months, I see the costs of these two parts are fairly evenly split. Delivery service charges do not seem to have any basis in reality. The electrons passing through the copper wires do not wear out, and neither do they wear out the wires. If a customer uses 1000 kWh/month instead of 1 kWh/month, does the costs of delivery increase by a factor of 1000? It is this type of abuse by the IOU that leads to customer dissatisfaction and ultimately to the demise of that IOU.

Most electricity in the United States is provided to individual residents by IOUs. Only 14.7% is provided by Munis [http://publicpower.com/floridas-electric utilities-2/], a percentage
likely to increase significantly in the future. Why? Because a municipal electric utility company has a “business model featuring local ownership and control, non-profit operations, superior customer service and generally lower rates [Massachusetts Municipal Wholesale Electric Company 2011 Annual Report].” That report paraphrases from a 1933 campaign speech of Franklin Delano Roosevelt as follows:

“When a community is not satisfied with the service rendered or the rates charged by a private utility, it has the undeniable right, as one of its functions of government, to establish its own electric utility, by vote of the electorate.”

The dissatisfaction with IOUs is increasing. “Florida municipal utilities’ average outage is less than one hour. For the two largest private utilities in Florida, Florida Power and Light and Florida Power Corporation, the average customer outage is more than two hours [op. cit.]”

A story in the Los Angeles Times under the title “California electric bill shock: Private firms charge way more than public utilities,” www.latimes.com/local/lanow/la-me-ln-electric-bills-differ-private-companies-charge-more-than-public-utilities-20150614-story] compared California electric bills for October 2014 at different usage levels:

<table>
<thead>
<tr>
<th>Usage</th>
<th>Public Average</th>
<th>Private Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>$621.38</td>
<td>$351.08</td>
</tr>
<tr>
<td>1000</td>
<td>$276.88</td>
<td>$163.59</td>
</tr>
<tr>
<td>500</td>
<td>$102.29</td>
<td>$75.54</td>
</tr>
<tr>
<td>200KWh</td>
<td>32,32</td>
<td>$30.29</td>
</tr>
</tbody>
</table>

“Salary and benefits paid to executives at investor-owned utilities — generally higher than those paid by public agencies — also affect rates, consumer advocates say. Public salaries criticized at the Los Angeles utilities department were $220,000, compared with $11.6 million in cash and equity in 2014 for the CEO of PG&E, an IOU [op. cit.]”

“Publicly owned utilities exist to serve customers. Period. There are no stockholders, and thus no profit motive. Our electric prices do not include a profit. That means Nebraska’s utilities can focus exclusively on keeping electric rates low and customer service high. Our customers, not big investors in New York and Chicago, own Nebraska’s utilities. You can see the difference every month, in your electric bill. Our costs to generate and deliver electricity do not include a profit markup. As a result, electricity costs in Nebraska are well below the national average [Nebraska Power Association, www.nepower.org/who-we-are/public-power].”

Among the U.S. cities considering forming municipal electric power are Boulder, Colorado, Minneapolis, Minnesota, and Santa Fe, New Mexico. The reasons are more environmental than economic. They perceive that IOUs are dragging their feet to expand energy production from renewables.
Sometimes the conversion from private to public encounters problems. A case in point is the muni Progress Energy that took over in 2005 in Winter Park, Florida. It initially lost money and had to raise rates, but investments in infrastructure have paid off and it is now making money and using it to bury 10 of its 80 miles of cable [D. Cardwell, “Cities Weigh Taking Over From Private Utilities,” New York Times, March 13, 2013].

One argument that IOUs make to retain control is that they would be left with stranded costs — they need more time to recoup the investments that they have made. Rarely is this a valid argument. Their infrastructures are aging and their capital improvements scant.

Another argument used by some IOU companies is that to compare Munis with IOU is unfair, since the Munis are non-profit and exempt from taxes, whereas the IOUs are for-profit and so subject to the 35% corporate federal tax, and to state taxes. But who pays 35% of their profits to the federal government? Very few companies, and certainly not the typical IOU. Far more significant is the enormous compensation packages given to IOU executives.

<table>
<thead>
<tr>
<th>State</th>
<th>cum 2012</th>
<th>new 2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massachusetts</td>
<td>207.3</td>
<td>123.2</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>164.3</td>
<td>31.3</td>
</tr>
<tr>
<td>Florida</td>
<td>116.9</td>
<td>21.9</td>
</tr>
<tr>
<td>Ohio</td>
<td>79.9</td>
<td>48.3</td>
</tr>
<tr>
<td>Oregon</td>
<td>56.4</td>
<td>20.6</td>
</tr>
<tr>
<td>Tennessee</td>
<td>45.0</td>
<td>23.0</td>
</tr>
<tr>
<td>Illinois</td>
<td>42.9</td>
<td>26.7</td>
</tr>
<tr>
<td>Connecticut</td>
<td>39.6</td>
<td>7.5</td>
</tr>
<tr>
<td>Vermont</td>
<td>28.0</td>
<td>16.3</td>
</tr>
<tr>
<td>Georgia</td>
<td>21.4</td>
<td>8.2</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>21.1</td>
<td>8.2</td>
</tr>
<tr>
<td>Michigan</td>
<td>19.9</td>
<td>11.1</td>
</tr>
<tr>
<td>Washington</td>
<td>19.5</td>
<td>7.2</td>
</tr>
<tr>
<td>Louisiana</td>
<td>18.2</td>
<td>11.9</td>
</tr>
<tr>
<td>Minnesota</td>
<td>11.3</td>
<td>6.5</td>
</tr>
<tr>
<td>Virginia</td>
<td>10.5</td>
<td>5.2</td>
</tr>
<tr>
<td>South Carolina</td>
<td>4.6</td>
<td>0.5</td>
</tr>
<tr>
<td>Indiana</td>
<td>4.4</td>
<td>0.9</td>
</tr>
</tbody>
</table>

The Huffington Post in February, 2014, identified 26 major companies that were given a negative federal tax for the years 2008 though 2012. These companies included PG&E, which paid at a tax rate of –16.7%, and Pepco Holdings, a company with equity interest in PG&E.
That is a negative sign that you see [www.huffingtonpost.com/2014/02/25_corporation-tax-rate_n_4855763], adding to PG&E’s profits of about $7 billion.

The state of Minnesota, which is a leader in advocacy for RE, published a list showing the total installed PV in GW per state in 2012, and the amount installed in 2012 alone


Because 70% of the U.S. grid is older than 25 years, positioning energy storage on the load side of the grid will make the grid more robust [op. cit.]. Why is the infrastructure of the grid not being updated as a matter of course? The answer is in the business model of the IOU, which is 100 plus years out of date. Bill McKibben in The New Yorker, June 29, 2015, had a piece called “Power to the People; Why the Rise of Green Energy Makes Utility Companies Nervous.” He cites the expansion of Solar City, whose PV installations have doubled every year for the last seven years, and points out that by 2016 “solar will be the fastest-growing new source of energy in the country, approaching half of new capacity,” and will ultimately be the dominant energy generator. “As customers began to generate more of their own electricity from the solar panels on their roofs, utility revenues would begin to decline, and the remaining customers would have to pay more for the poles and wires that keep the grid alive,” creating “a death spiral.”

McKibben cites a reporter for Vox, David Roberts who covers energy matters: “If you’re in a business where the customer is the public-utility commission, and after that your profits are locked in by law, it’s the sleeppiest business sector there is, if you could even call it a business sector. They build power plants, sit back, and the money comes in.”

These are a sampling of the overwhelmingly pro Muni and anti IOU comments that can be found online. The defenders of IOUs provide a counterset of arguments that, at least from the vantage point of this observer, are not convincing.

### 14.8 A Regulated Future

Some of the biggest and most successful companies in United States did not exist a generation ago. Google, Amazon, Facebook, Netflix, and now Uber and Airbnb are just a few examples. Some of those that existed and thrived in the latter part of the twentieth century are gone.

Two prominent examples on the gone side are Digital Equipment Corporation and Wang Laboratories, the #2 and #3 computer companies in the world in 1980, both Massachusetts
based-companies. Digital Equipment was co-founded in December 1957 by MIT engineers Ken Olsen and Harlan Anderson. Wang Labs was founded in 1955 by Harvard graduate An Wang with a Ph.D. in physics.

Massachusetts in 1980 was the heart of the minicomputer industry. Digital Equipment in 1989 was the largest private employer in the state with 34,000 employees. As a high-tech center, Massachusetts was bigger than Silicon Valley in California. What went wrong? What happened was launching of the IBM PC in August, 1981 with an operating system developed by MicroSoft Corporation (the hyphen was soon dropped). Within a decade the PC became the dominant computer platform across the world.

Both Digital Equipment and Wang Labs failed to adapt. “Digital developed a reputation as an insular company that rejected innovative ideas from beyond its borders . . . . So while the personal computer was welcomed in places like Seattle and Silicon Valley, the minicomputer giants of Massachusetts would not embrace the new technology until it was far too late [Hiawatha Bray, The Boston Globe, February 15, 2011].” Wang Labs filed for bankruptcy in 1992.

In the 1960s the largest company in America was General Motors. The saying at that time was “what’s good for General Motors is good for the U.S.A.” It’s “As California goes, so goes the nation,” with respect to presidential election. In June, 2009, General Motors declared bankruptcy. The recession gripping the world was blamed for its demise, and indeed many companies went under at that time. GM was saved by a government bailout, subsequently repaid, leading some wags to say “GM now stands for Government Motors.” However, GM was hemorrhaging market share for decades by failing to match the quality and value of the Japanese imports. It share of the U.S. market was 51% in 1962, 31% in 1997, and 18% in 2013.

A book that came out in 1970 was Future Shock by Alvin Toffler: its main thesis was that change was happening at an increasing rate. This is happening to business, particularly since we are in an interconnected, global economy where many of the old rules no longer seem to be appropriate.

In Haiti and similar places, they have adapted to life without electrical power. In the advanced countries of the world, we will never accept that condition. So what will happen if the IOUs refuse to accept the reality that they must become power managers, not power producers? I have told myself a million times not to exaggerate. The only answer is that they must be pushed aside so that Munis or publicly owned and operated companies take over.

“One of the largest ‘socialist’ enterprises in the nation is the Tennessee Valley Authority, a publicly owned company with $11 billion in sales revenue, nine million customers and 11,260 employees that produces electricity and helps manage the Tennessee River system. In 2013 President Obama proposed privatizing the T.V.A., but local Republican politicians, concerned with the prospect of higher prices for consumers and less money for their states, successfully
opposed the idea . . . . More than 2,000 publicly owned electric utilities supply more than 25 percent of the country’s electricity, now operating throughout the United States . . . . In one of the most conservative states, Nebraska, every single resident and business receives electricity from publicly owned utilities, cooperatives or public power districts. Partly as a result, Nebraskans pay one of the lowest rates for electricity in the nation [G. Alperovitz and T.M. Hanna, “Socialism, American-Style,” New York Times, July 23, 2015].”
Chapter 15

Architectural Considerations for Capturing Solar Energy

Water power, biofuels, and geothermal systems exist independently of the architecture of the house they help power. Wind power could be roof mounted, but this is unusual; the rotating blades produce noise and vibration, a condition that will not be tolerated by occupants in the house below.

On the other hand, a significant component of success in utilizing solar thermal energy in residential construction is the architecture. This is the subject of this chapter. Also important is the siting and exposure to the Sun that was discussed in Chapter 9. Present day photovoltaic panels are installed over a completed roof, and works best when the roof faces the Sun at the correct pitch; refer to Section I of Chapter 11.

Solar thermal energy can be used to heat a residence in winter, and to heat water year round. These may impose different architectural requirements, as will be seen. Solar energy can also be used to generate photovoltaic-based electricity, and this requires somewhat different construction requirements.

Architecture styles vary according to the climate of the region. What is appropriate for Santa Fe, New Mexico will not work in New England, and vice versa.

15.1 Basic Requirements for a Passively Solar Heated Residence

The first requirement for any solar heated residence is exposure to the Sun, meaning southern exposure in northern latitudes and northern exposure in southern latitudes.
Again, with apologies to those in southern latitudes, we avoid a tedious repetition by writing for northern latitudes only.

As discussed in Chapter 9, south-facing slopes are desirable locations, since solar exposure is considerably easier to obtain than on a north-facing slope. Even slopes to the east and west could be problematic.

One of the biggest mistakes in solarium design is to ignore the summer heat gain by glazed areas. Glass on a south-facing roof is particularly difficult. Vertical glass is much less of a problem when it is facing due south. The best for minimal summer gain is vertical south-facing glass with an overhang.

An essential element with passive solar heating is a high thermal mass. The choices are water, masonry, or eutectic salts. As discussed in Chapter 11, water has a high thermal mass — much higher per unit volume than rock or concrete. Water is cheap to the point of being free, so it is tempting to make this the first choice. There are, however, significant drawbacks. First, the water must be contained in some type of tank that is not degraded or corroded by the water. Second, an air-to-water heat exchanger is required, with electrical needs. Third, the temperature to be anticipated in the tank will be below 60°C, 140°F, too low to kill off any bacterial action: pasteurization temperature is 72°C, 162°F.

15.2 Architecture for Natural Air Flow
A passively solar heated house should have a fairly open interior layout. This precludes the colonial designs so popular in New England. Instead, a contemporary layout provides a more effective architecture. Such a design is shown in Figure 15.2. The structure is basically a split level structure, split vertically near the north/south center line, and one side displaced by half a level, close to 5’. The two sides are then opened out to create a central atrium with half flights of stairs joining the next level up or down. In the plan view on the top of the four flights of stairs are stacked, so only two flights of stairs are visible in the plan view. This is my design for the house I built for my daughter’s family.

The south glass face accepts the solar heat that rises to the top of the atrium area where it is ducted to a rock storage bin shown in Figure 15.3. The duct feeds a manifold almost the full length of the storage bin; such a manifold can be 6” sewer pipe with large holes along its length facing into the rock bin. A thermostatically controlled fan pumps the solar heated air into
the rock. A matching manifold on the opposite side of the bin has a duct leading to a wall register; a fan in this duct shares input from the same thermostat, so both fans operate simultaneously.

![Figure 15.3: Rock Storage with Air Distribution Manifolds](image)

It is suggested that the rocks be of a nominal 3” to 6” size, clean and free of fines, dressed on top with 3/4” to 1 1/2” stone, and in turn dressed again with peastone. The concrete slab poured on top will not pass through the peastone.

Care must be taken to avoid mold in the rock store. This means that the water table must be well below the bottom of the rock. It is suggested that beneath the rock there be a sandwich of sand, plastic sheet, and sand. The plastic sheet should be larger than the floor area so it can wrap up the sides.

The most difficult part of a structure with two or more levels is the traffic flow to and from the stairs. To be specific, with an open interior layout, it is challenging, but worth the effort. With the colonial style house, whose individual rooms function independent of each other and with corridors feeding these individual rooms, it is simple. Having designed a number of solar
structures, I can happily say I have never designed one with a corridor. Sometimes it means having a shared space at the bottom or top of a set of stairs, but it can be done.

There are fairly strict rules that must be followed in the design of stairs. One golden rule is that twice the riser height plus the tread depth must be close to 25”, and the ideal is for the tread to be 10” so the riser is 7.5”. This design is shown in Figure 15.4 and represents a half flight between the levels in Figure 15.2. The top step is set against the joist of the upper floor and set 7.5” below the plywood surface above that joist, so the total rise is 60”. There are a total of eight risers on each half flight.

![Figure 15.4: Standard Stair Layout](image)

For outside steps the same 25” formula applies. For example, a path too steep to have an inclined surface, about a slope of one in six, or 9.5°, requires steps. If the riser is x” the tread is 25 − 2x”; a riser of 2” pairs with a tread of 21”, and a riser of 5” pairs with a tread of 15”. Stairs/steps with risers greater than 9” do not work.

Interior stairs are usually custom cut to fit the situation. For example, suppose the total rise is 57”, then for eight risers there will be 7 1/8”. The treads are typically not a problem since the bottom of the stringer does not have to be at the edge of the lower floor - it can move in or out.
15.3 Glazing for Passive Solar

A relatively inexpensive item in the United States is 46” x 76” tempered insulated glass, since this is the size and type used in 8’ sliding glass doors. Economically prudent practice is to use standard sizes and types whenever possible.

The only acceptable orientation for passive solar heating is vertical and facing due south in northern latitudes. The vertical orientation is a compromise that works well when combined with an overhang feature in the construction as shown in Figure 15.5; see Section III of Chapter 11 on shading. Here the 24° angle on the solar aperture is appropriate for latitudes 42°.
The frame of the glass eats 2” off each side, so the effective height of the glass is 72”. The solar aperture determines the solar gain, and in this case is 68”, so about 8% of the Sun’s energy is lost. However, the summer Sun at solstice has a minimum angle of a little over 70°, and as can be seen in Figure 15.5 the overhang is shading most of the glass, so only 45.6” is irradiated. This makes the effective aperture about 10”. Also, the angle of the Sun on the glass reduces the transmissivity through the glass by 70%. The bottom line is that the summer solar gain through the south-facing vertical glass is negligible, far less than that through unshaded...
glazing on the east and west sides; to latitudes greater than 35°N the solar gain on north-facing glazing at any time of the year is negligible in summer and zero at other times of the year.

Orientable glazing is not a realistic option for passive solar. This orientation can be in the rotation about a horizontal axis, but as can be seen from the prior analysis, the gain is not worth the effort. If the glazing is rotated about a vertical axis, the glass can be oriented to face the Sun as it swings over a large azimuth angle. One problem is in shading — one glass panel partially shades the next panel from the Sun, as can be seen in Figure 15.6. However, the most serious problem is how to permit the glass to move along a circular arc while sealing against air and water infiltration.

Reflective surfaces could increase passive solar gain. An early solar design, possibly not realized, proposed optically reflective hinged panels that folded down during the day, and were closed up at night to reduce the nighttime heat loss. The problems with this notion are:

* The reflective flaps probably could not be walked on, limiting access from that side of the structure,
* The internal glare would be intolerable.

The glare problem can be totally alleviated by incorporating the flaps with a Trombe wall system as shown in Figure 15.7. The limited access would not become a problem if the
reflective flap and Trombe wall was constructed over ground that cannot be accessed. For example, if the structure was built on a cliff with solar orientation out over the void.

![Diagram of reflective flap and Trombe wall](image)

**Figure 15.7: A Reflective Flap to Increase Solar Collection**

### 15.4 Roof Mounted Solar Elements

#### 15.4.1 Requirements for Photovoltaic Panels

Photovoltaic panels are typically mounted above an existing roof. For a standard pitch roof with asphalt shingles, 4” to 6” aluminum pillars spaced 32” to 48” apart are attached through the shingles by stainless steel 1/4” × 3 1/2” lag bolts into the rafters as supports for a horizontal rail system; a vertical rail system would be problematic since the rafters run up the roof not across it. Flashing around the pillars ensures against leakage. The pillars must be carefully placed within 1/4” of nominal; otherwise, the bank of collectors attached to the rails
will not properly align and the appearance will be unacceptable. Each photovoltaic panel has clips that are attached to the rails. Any height inconsistency caused by minor variations due to the rafters can be adjusted with stainless steel washers. The rails provide electrical connection as well as physical connection — the rails are electrically grounded.

Natural air flow under and over the collectors provides cooling. As discussed in Chapter 13, the efficiency of a typical photovoltaic panel is about 15%, so the remaining 85% heats the panel. Further, these panels can be cooled using the Sundrum system described in Chapter 13. For metal pitch roofs the rails are mounted on pillars that are themselves clipped to the metal roof, so there is no roof penetration. Similarly, for flat roofs there is no roof penetration; the aluminum rack system can be ballasted using concrete blocks.

15.4.2 T-Bar Glazing Systems

A popular method for constructing a glazed roof is to use a T-bar system to hold the glazing. The T-bars run down the roof on top of the rafters or similar structural elements; they are not used as purlins that run horizontally between successive sheets that run up the roof. The advantage of the T-bar system is that minor leaks flow into a trough pointed downwards on the roof, and so drain out without penetration into the structure of the building.

Columbia Aluminum Products uses the T-bar system to produce a broad range of sloped glazing systems, be they small scale conservatories or large commercial installations, such as shopping malls. A simplified version of their hardware is shown in Figure 15.8.
A number of other companies use T-bar methodology to offer glass or clear plastic roofs, usually for shopping malls, train stations, or other large commercial projects. To find them use Google with the key words “roof glazing systems.” A few results are shown below:

“Sapa 5050 roof glazing system consists of 50 mm wide insulated aluminium profiles. The system has a longitudinal insulating strip which provides good insulation and support for the glass holders. The horizontal, vertical and dividing profiles come in different depths. Rebate for single glass or for double and triple glazing units . . . . The 5050 system makes it possible to give the roof virtually any form. Glazed roofs can have large spans without compromising the static stability [www.sapagroup.com/lt/company-sites].”

“CrystaLite is a Northwest Washington manufacturer of high quality skylights, roof glazing, sunrooms, and railing systems. Our suppliers are also local businesses and employ local workers . . . . Our primary supplier of aluminum is SAPA Inc, based in Portland. They receive their raw billet from regional suppliers in Ferndale and The Dalles; and 80% of their raw material is mined by local workers within 500 miles. Cardinal IG is our primary source of glass, based in Hood River. Likewise, they receive their supplies from Cardinal FG in Winlock, who manufactures float glass created from raw sands mined from local quarries [www.crystaliteinc.com].”

“From modular rooflights to large area roof glazing systems, our products help you to make the most of daylight. Roofglaze designs, supplies and installs rooflights for all types of projects —
from retail parks and leisure centres, to offices, schools, hospitals and private homes.

“We work as partners with world-leading rooflight manufacturers — as well as offering our own high quality products manufactured at our Cambridgeshire factory. Our solutions range from modular Flatglass and polycarbonate rooflights, to bespoke skylights for major projects like London’s V&A Museum and the refurbishment of Debenhams, Oxford Street [www.roofglaze.co.uk].”

It is interesting that the interest in glass roofs seems to be considerably greater in rainy Britain than in other countries. Their web sites show beautiful sunny spaces under glass, usually framed with aluminum bars of various designs. On example is the Glazatherm system from the Lonsdale Metal Company.

Some companies specialize in greenhouses or conservatories, usually as add-ons to existing structures: Shed roofs are popular in this case since the difficulties of weatherproofing the top and the bottom are somewhat alleviated. Flashing systems at the top are fairly easy to implement on a case-by-case basis. At the bottom the system can be left open provided the method to hold the glazing is strong enough to prevent slippage of the glazing panels — one would not want to have the panels slip out from the wall, or even worse, slide out completely.

To prevent slippage of the panels, a shoe system can be employed, as shown in Figure 15.9. The glazing extends beyond the T-bar support system and is held in the shoe.

Weep holes in the shoe ensure adequate drainage. Thermal conductivity loss through the shoe may be a problem, but this can be minimized by supporting the shoe on a wooden sill so it does not suck heat directly from the structure below.
There may be substantial heat loss by conduction through the T-bar shown in Figure 15.8. To prevent this, a thermal break can be constructed in the T as shown in Figure 15.11.

### 15.4.3 Multi-Layer Plastic Roofs

Polycal uses multiwall polycarbonate sheets with a fairly simple aluminum hardware system, as shown in simplified form in Figure 15.12. Each sheet of glazing is sandwiched between a base and a cap, with gaskets to bear against the glazing to provide the weatherproofing. The central affixing screw could be a machine bolt or a self-tapping sheet metal type of screw. The head of the screw should be truss so the width of the head can hold a soft sealing washer for weatherproofing.
The Polycal system can rest on an aluminum frame to produce a solarium or a commercial atrium. It could, for our purposes, rest on wooden rafters that comprise the roof of the house. Deglas® also uses a multi-sheet, from two to four layers, but in acrylic, and their hardware system is similar to that of Polycal. An alternative to the seal washer is the clip-on cap shown in Figure 15.13.

Starlight Glazing Systems of Hurlock, Maryland, has an aluminum frame system with multiwall polycarbonate panels 166 mm, 5/8’’ thick, in sizes up to 4’ wide and 16’ long. The extruded aluminum glazing bars are integral to the glazing panels. The frames interlock on the sides to be weatherproof. It is not clear how the system can be made weatherproof at the top.
15.4.4 Requirements for Solar Thermal Collectors

When active solar collectors are installed, they should in all probability be placed on or in the roof of the house. Complete packages come with glazing, solar collection elements (flat plate or otherwise), frame, insulation, and hardware to affix to the roof. They are expensive, but have the advantage that the homeowner can contract for a turnkey operation. A significant part of the expense is the duplication of structural members. The roof of the house requires rafters, plywood, insulation, flashing, and shingles. Instead, wouldn’t it be desirable to use the house structure for the frame, glazing, and insulation of the active solar elements? The savings from doing this could result in a savings of 70% of the turnkey cost.

15.5 Plumbing for Water Heating

The need for hot water is universal, from the hottest to the coldest climates. Some schemes for using the Sun to heat water were discussed in Section X of Chapter 11. Now we tie the water heating to the house architecture.

Plumbing codes require adequate pipe sizes to avoid excessive drop in pressure at an outlet far from the water source within the structure. How realistic are these codes, and do they waste both water and energy by being too stringent? We consult the Copper Tube Handbook,
produced by the Copper Development Association, Inc., which can be found at www.copper.org, to answer these questions. According to them, the pipe sizes chosen are dependent on the available main pressure, the pressure required at individual fixtures, water demand in each fixture, pressure losses due to friction, and velocity limitations to reduce noise and erosion.

Table 15.1

<table>
<thead>
<tr>
<th>size</th>
<th>OD</th>
<th>ID</th>
<th>wall</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>3/8”</td>
<td>0.500</td>
<td>0.402</td>
<td>0.049</td>
<td>0.269</td>
</tr>
<tr>
<td>1/2”</td>
<td>0.625</td>
<td>0.527</td>
<td>0.049</td>
<td>0.344</td>
</tr>
<tr>
<td>3/4”</td>
<td>0.875</td>
<td>0.745</td>
<td>0.065</td>
<td>0.641</td>
</tr>
<tr>
<td>1”</td>
<td>1.125</td>
<td>0.995</td>
<td>0.065</td>
<td>0.839</td>
</tr>
</tbody>
</table>

Table 15.2

<table>
<thead>
<tr>
<th>size</th>
<th>OD</th>
<th>ID</th>
<th>wall</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>3/8”</td>
<td>0.500</td>
<td>0.430</td>
<td>0.035</td>
<td>0.198</td>
</tr>
<tr>
<td>1/2”</td>
<td>0.625</td>
<td>0.545</td>
<td>0.040</td>
<td>0.285</td>
</tr>
<tr>
<td>3/4”</td>
<td>0.875</td>
<td>0.785</td>
<td>0.045</td>
<td>0.455</td>
</tr>
<tr>
<td>1”</td>
<td>1.125</td>
<td>1.025</td>
<td>0.050</td>
<td>0.655</td>
</tr>
</tbody>
</table>

They also cite static pressure problems, but this is only significant in tall buildings. A pressure of 40 lb/in² will lift water a vertical distance of 92.4’.

Copper is traditionally the most common pipe used in domestic plumbing, but it now has a competitor in the plastic pipe called PEX. Here we concentrate on copper. The outside diameter (OD) of a copper pipe is 1/8” larger than its nominal size, so a 1/2” nominal pipe has a 5/8” OD, and a 3/4” nominal pipe has a 7/8” OD. The common wall thicknesses are rated as type K or type L. K has the thicker wall, and should be used for all domestic supplies at mains pressure, typically about 40 lb/in²; pressure should not exceed 80 lb/in² to avoid damage to fixtures. Type L can be used for hydronic heating systems such as feeding baseboard radiators.

The dimensions and weight of type K copper pipe are given in Table 15.1. The dimensions and weight of type L copper pipe are given in Table 15.2.

The specific heat for copper is 0.09 BTU/lb·°F, or 0.39 kJ/kg·°C, or 0.092 kcal/kg·°C.

Water demand per fixture is rated [www.copper.org] in gallon/minute as

- drinking fountain 0.75
- lavatory faucet 2.0
- lavatory faucet, self closing 2.5
A general rule of thumb is that a 1/2” main can feed up to three 3/8” branches, a 3/4” main can feed up to three 1/2” branches, and a 1” main can feed up to three 3/4” branches. However, plumbing codes often require the minimum feeder pipes for a bathroom to be 3/4”. The ID for a 3/4” type K pipe is 0.745”, so the length of such pipe containing 1 U.S. gallon of water is $122C=44.16'$, the length of a typical house.

The pressure loss in psi/foot due to friction for type K pipe as a function of flow rate in gpm provided the flow rate does not exceed 8’/second is given in Figure 15.14 [see The Engineering Toolbox]; notice the log scale, for which the responses are linear. These pressure drops can be characterized by

\[
D_{1/2} = e^{1.8431-4.605} \\
D_{3/4} = e^{1.853-6.303} \\
D_1 = e^{1.870-7.689}
\]

Figure 15.14: Friction Losses in Copper Pipe
The friction loss for elbows and tees can be defined in terms of the friction loss in feet of regular pipe, and this is given in Table 15.3.
Chapter 16

Methods of Energy Storage

The burning of fossil fuels is the prime culprit in global warming. Unfortunately, the fossil fuels have two substantial advantages over renewable energy sources — portability and energy density. It is easy to tote a 10-liter can of gasoline that contains about 120kWh of energy. A large, lead 60-amp-hour acid car battery with volume about ten liters has a maximum of 0.72kWh of energy. To make things even worse, to drain a lead acid battery is to seriously damage it; that is, it will no longer be able to store 60 amp-hours of energy.

Not only do fossil fuels have a high energy density, they can also be stored for long periods of time before use with little or no degradation. Coal is dug out of the ground after its formation millions of years before, and can be stored for an unlimited time before being combusted. That is, the characteristics of most fossil fuels are high energy density and long-term stability. This is not the case with renewable energy, whose characteristic is use it now or lose it.

Thus, the challenge that could determine our future is finding an efficient and economical way of storing energy produced from carbon neutral, renewable sources.

16.1 Electrical Generation Capacity

Renewable energy resources that can be tapped when needed include hydroelectric, biofuels, and geothermal. Those that cannot wait for a time of convenience are solar PV and wind energy, and these can be considered intermittent power producers: If the energy generated cannot be used or stored, it is wasted.

Table 16.1: World Electric Production

<table>
<thead>
<tr>
<th>source</th>
<th>TWh</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>coal</td>
<td>8,390</td>
<td>40</td>
</tr>
<tr>
<td>gas</td>
<td>4,744</td>
<td>23</td>
</tr>
<tr>
<td>hydro</td>
<td>3,619</td>
<td>17</td>
</tr>
<tr>
<td>nuclear</td>
<td>2,344</td>
<td>11</td>
</tr>
<tr>
<td>oil</td>
<td>879</td>
<td>4</td>
</tr>
<tr>
<td>other</td>
<td>1,042</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 16.2: Electrical Capacity

<table>
<thead>
<tr>
<th>Country</th>
<th>Capacity (GW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>1,505</td>
</tr>
<tr>
<td>U.S.</td>
<td>1,053</td>
</tr>
<tr>
<td>Japan</td>
<td>287</td>
</tr>
<tr>
<td>Russia</td>
<td>239</td>
</tr>
<tr>
<td>India</td>
<td>223</td>
</tr>
<tr>
<td>Germany</td>
<td>178</td>
</tr>
<tr>
<td>Canada</td>
<td>139</td>
</tr>
<tr>
<td>France</td>
<td>130</td>
</tr>
<tr>
<td>Italy</td>
<td>124</td>
</tr>
<tr>
<td>Brazil</td>
<td>119</td>
</tr>
<tr>
<td>Spain</td>
<td>102</td>
</tr>
<tr>
<td>South Korea</td>
<td>87</td>
</tr>
<tr>
<td>U.K.</td>
<td>76</td>
</tr>
<tr>
<td>Iran</td>
<td>65</td>
</tr>
</tbody>
</table>

The 2014 U.S. breakdown is fairly similar, with coal at 39%, gas at 27%, nuclear at 19%, hydro at 6%, other renewables at 7%, with 1% left for petroleum and 1% for other gases [“What is U.S. electricity generation by energy source?” www.eia.gov/tools/faqs/faq.cfm?id=427&t=3, U.S. Energy Information Agency]. The total U.S. electricity generated in 2014 was about 4.093 Twh.

The total worldwide generating capacity in 2014 was 5,492 GW [www.tsp-data-portal.org/TOP-20-Capacity#tspQvChart]. The top producers of electrical generating capacity in GW [www.cia.gov/library/publications/the-world-factbook/rank-order/2236rank] are given in Table 16.2.

### 16.2 The Energy Density Problem
For energy storage purposes in which the charge/discharge cycle requires taking the storage device into deep discharge, the lead acid battery is unsatisfactory. Luckily, other types of battery can be discharged without damage, but their energy density remains a problem.

Table 16.3: Installed PHS

<table>
<thead>
<tr>
<th>Country</th>
<th>Installed PHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japan</td>
<td>27.438</td>
</tr>
<tr>
<td>China</td>
<td>21.545</td>
</tr>
<tr>
<td>U.S.</td>
<td>20.858</td>
</tr>
<tr>
<td>Italy</td>
<td>7.071</td>
</tr>
<tr>
<td>Spain</td>
<td>6.889</td>
</tr>
<tr>
<td>Germany</td>
<td>6.388</td>
</tr>
<tr>
<td>France</td>
<td>5.894</td>
</tr>
<tr>
<td>India</td>
<td>5.072</td>
</tr>
<tr>
<td>Austria</td>
<td>4.808</td>
</tr>
<tr>
<td>South Korea</td>
<td>4.700</td>
</tr>
<tr>
<td>U.K.</td>
<td>2.828</td>
</tr>
<tr>
<td>Switzerland</td>
<td>2.687</td>
</tr>
<tr>
<td>Taiwan</td>
<td>2.608</td>
</tr>
<tr>
<td>Australia</td>
<td>2.542</td>
</tr>
</tbody>
</table>

Consider a typical, energy efficient residence in New England over a winter day. The heat load for that house will be about 10kW, or 240kWh/day, equivalent to the energy in about 22 L of fuel oil; the energy density of gasoline is about 9.6kWh/L, and in fuel oil or diesel oil about 10.8kWh/L. If that energy was provided by car batteries that could be replaced every day, the number of batteries needed would be \( \frac{240 \times 24}{0.72} = 333 \). This would require the services of a rigger, a massive amount of space with adequate ventilation to prevent the concentration of potentially explosive gasses, and a large truck. In other words, storing energy in batteries for space heating purposes is presently impossible.

It is the problems involving the storage of electrical energy that limit the amount of wind or photovoltaic energy that can be generated. It is relatively easy for a traditional power station to increase or decrease its output. This is not possible for electricity generated by wind or sunlight.

Power stations have a particular problem. They are required to deliver power to their consumers yet have no control on demand. They do have estimates of demand based on experience, but even a 10% deviation from that estimate leads to power outages or power wastage — once power is generated, unless it is used by its customers, it is lost. What is needed are energy storage systems that can contain a significant fraction of the daily needs of their customers. Apart from hilltop reservoirs, no such systems exist.

The energy densities in Wh/kg for existing systems are:
### 16.3 Mechanical Energy Storage Techniques

Energy stored by mechanical means can be potential or kinetic. Hydroelectric projects use water held behind dams (potential energy) to flow through pipes to pick up velocity, converting the potential to kinetic energy, to power water turbines. As discussed in Chapter 1, hydroelectric power is a mature industry, with most rivers capable of generating significant amounts of power already doing that generation. This type of energy capture and production has a major drawback — when the reservoir is depleted, energy production ceases.

There is a significant promise in tidal energy in a limited number of locations in the world, as was discussed in Chapter 1. However, the problem is that the power they generate is intermittent, and therefore not a good partner to solar and wind energy. The Bay of Fundy has the greatest potential, but a very limited effort to harness this energy has been made, possibly because it would totally change the ecology of the bay.

#### 16.3.1 Pumped Hydroelectric Storage

Hilltop reservoirs are closed cycle systems with two substantial benefits. They have a high cycle efficiency, and they are capable of rapid power production.

Almost all of the approximately 2.5% of electrical power that passes through the electrical grid is generated by hilltop reservoirs [R. Walawalkar and J. Apt, “Market Analysis of Emerging Electric Energy Storage Systems,” DOE/NELT-2008/1330 final report, Carnegie Mellon University, July 2008]. The term that is now used for these facilities is PHS (pumped hydroelectric storage). This 2.5% amounts to 23.4 GW, and is 95% of all grid connected energy storage.

The top producers of installed PHS Capacity in GW [C-J. Yang, “Pumped Hydroelectric Storage,” http://people.duke.edu/~cy42/phs,] are given in Table 16.3.

“The earliest PHS in the world appeared in the Alpine regions of Switzerland, Austria, and Italy in the 1890s. These early designs used separate pump impellers and turbine generators. Since the 1950s, a single reversible pump-turbine has become the dominant design for PHS. The development of PHS remained relatively slow until the 1960s, when utilities in many
countries began to envision a dominant role for nuclear power. Many PHS facilities were intended to complement nuclear power in providing peaking power [op. cit.].”

![Simplified Schematic of a PHS Facility](https://example.com/schematic.png)

The need for an energy producer that can work with intermittent power generators, such as solar PV and wind energy, to support a reliable electrical grid is increasing, and PHS is such a producer. What is needed is a hill, and Japan has many, so it leads the world with 27 GW of storage and generating capacity in PHS. China is not far behind with 21 GW, and many more hills to harness. Low-lying countries such as Holland have little opportunity to develop PHS.

A simplified schematic of a PHS facility is shown in Figure 16.1. Two reservoirs are connected by a large pipe. The power plant moves water along the pipe. It pumps water up to the upper reservoir when there is surplus power, and the same mechanism becomes an electrical generator when electrical power is needed. Also needed is a surge chamber shunt to eliminate damaging pressure surges, and a breaker to prevent unwanted water flow.

“The main problem with gravitational storage is that it is incredibly weak compared to chemical, compressed air, or flywheel techniques. For example, to get the amount of energy stored in a single AA battery, we would have to lift 100 kg 10 m to match it. To match the energy contained in a gallon of gasoline, we would have to lift 13 tons of water one kilometer high [T. Murphy, “Using Physics and Estimation to Assess Energy, Growth, Options,” http://physics.ucsd.edu/do-the-math/2011/11/pump-up-the-storage].” However, what PHS has is volume.

Water in the upper reservoir has potential energy given by $mgh$, where $m$ is its mass, $g$ is the acceleration due to gravity, and $h$ is the head. A cubic meter of water weighs 1000 kg, so if it is 100 m above a turbine its potential energy is $1000 \times 9.81 \times 100 = 0.981 \times 10^6$ kg·m/s²·m; the dimension kg·m²s⁻² is the same as the joule. Suppose this 1m³ of water passes through the
turbine in 1 second; then, assuming 100% efficiency, the turbine would generate $0.981 \times 10^6$ W. In other words, with a head of 100m and a flow rate of a little above $1m^3/s$, the turbine will be generating 1MW.

“Ironically, what originally motivated pumped storage installations was the inflexibility of nuclear power. Nuclear plants’ large steam turbines run best at full power. Pumped storage can defer surplus nuclear power generated overnight (when consumption is low) to help meet the next day’s demand peak.

“Japan’s utilities chose to install variable-speed pumps so that the plants could also help stabilize their grids. Whereas single-speed pumps rotate synchronously with the grid, variable-speed pumps’ asynchronous motor-generators can adjust a plant’s charging and discharging to simultaneously balance power supply and demand, thus regulating the grid’s AC frequency [P. Fairley, “A Pumped Hydro Energy-Storage Renaissance: Storage Plants Built for Nuclear Power are Being Revamped for Wind and Solar,” http://spectrum.ieee.org/energy/policy/a-pumped-hydro-energystorage-renaissance, March 2015].

“40 PHS plants in U.S. producing more than 20GW, nearly 2% of the overall generating capacity. Compare this to Europe’s 5% and Japan’s 10%. World wide, PHS produced over 100GW in 2009 [http://energystorage.org/energy-storage/technologies/pumped-hydroelectric-storage].”

A plant in Goldisthal, Germany operates eight PHS. The largest generates 1.05GW, and all eight combined can generate 2.81GW with full cycle efficiencies of up to 80% [http://advantage-environment.com/framtid/pump-power-stores-energy].

In use since 1929, the Ffestiniog PHS in North Wales can generate 360MW semi-instantaneously (within a minute).

The Dinorwig plant, constructed between 1976 and 1982, scooped a cavern out of the Elidir Mountain in Wales to produce a head of 600m with a capability of generating up to 1,800MW. This plant was designed to be off-line most of the time, but to satisfy power surges within 16 seconds such as occur at the half-time of an important football (soccer) game when thousands of electric kettles are used to make tea.

An even larger project in Tianhuanping, China produces 1,800MW from a maximum head of 600m. There are two hilltop reservoirs in this complex, with a 7m-diameter conduit from each, each of which branches into three pipes 3.2m in diameter to feed the six 306MW turbines. The system was completed in 2000.

The world’s largest PHS facility is in Bath Country, Virginia, with the capability of generating over 3GW. The height difference between the reservoirs is 385m and water can flow at the rate of 852m$^3$ [http://advantage-environment.com/framtid/pump-power-stores-energy].
The Danish architecture firm Gottlieb Paludan proposed building artificial islands surrounding a deep tank that can be pumped out by wind turbines set on top of the islands. Sea water is allowed into the tank to generate electricity when more power is needed than the turbines can provide [http://www.economist.com/node/21548495?frsc=dg%7Ca].

![Diagram of Grid Connected Energy Storage Systems](image.png)

Figure 16.2: Grid Connected Energy Storage Systems

Following the mode of presentation in “Grid Energy Storage” by the U.S. Department of Energy, December 2012, Figure 16.2 shows the importance of PHS. The 5% of grid-connected storage that is not PHS is broken down, with thermal storage generating 36% of the 5%, CAES (compressed air energy storage) 35%, batteries 26%, and flywheels only 3%.

The numbers quoted in this section come from a number of different sources, and will vary. Part of the differences is due to data being collected at different times, and part may be due to differing data collection techniques, but overall the numbers are in agreement.

PHS is a well established and reliable method for storing energy. Early plants had efficiencies of about 60%, but in modern plants the efficiencies are above 80%. With 127GW of installed capacity worldwide, PHS is the dominant method for storing utility-size energy, dwarfing all the other energy storage schemes by orders of magnitude.

To put this amount of installed PHS in context, the electrical generating capacity of the world is 5,492GW. Over 99% of bulk energy storage worldwide is with PHS, according to The Economist [www.economist.com/node/21548495?frsc=dg%7Ca]. The reason the United States has 95% PHS generated, a smaller percentage than the rest of the world, is due partly to topography — are there suitable hills near population centers — and partly to technology. PHS uses simple and well-established technology, and the other methods for energy storage are more complicated and technically challenging for emerging countries.
PHS in China can store a full day of electrical production, the United States a little over half a day, but Japan less than three hours. Japan produces 10% of its electricity from PHS, Europe 5%, and the United States just 2% [source: Energy Storage Association online with Storage Capability for Pumped Hydroelectric].

The value of PHS is in its ability to change from idle to full generating capacity in seconds, and the ability to store surplus energy during nighttime hours and generate electricity during peak times. This daily surplus and peak load cycle is relatively consistent.

We have previously noted that hydroelectric facilities are a mature industry, so few new opportunities exist in the world for new plants. The same is true of PHS, so the future of grid-connected energy storage will not be with PHS.

It could be argued that traditional hydroelectric plants are a form of energy storage. A dam across a river contains water in a reservoir that is created above the dam. However, this is not the way most perceive them. The water that powers a hydroelectric plant is used once, but hilltop reservoirs operate with recycled water — the power-generating water was pumped to the top of the hill from a ground-level reservoir using surplus electrical power.

Lack of water could cripple some hydroelectric plants. The 14-year drought in the American West, in particular the Colorado River basin, has resulted in Lake Mead being at 40% of capacity. 43m down is the lowest level since the Hoover Dam was constructed in the 1930s to create Lake Mead, the nation’s largest reservoir. It is not rainfall that fills Lake Mead, it is the snowpack in the Rocky Mountains, and the winter of 2014–2015 provided less than half the normal amount [K. Siegler, “As Lake Mead Levels Drop, The West Braces For Bigger Drought Impact,” www.npr.org/2015/04/17/400377057/as-lake-mead-levels-drop-the-west-braces-for-bigger-drought-impact].

About 70% of the water from the Colorado River and its reservoirs goes to growing crops to feed the nation, with the rest to major cities such as Las Vegas, Phoenix, San Diego, and Denver. Nevada is the nation’s driest state, and to build a city like Las Vegas in the middle of a desert is looking increasingly unwise.

### 16.3.2 Compressed Air Energy Storage

Although dwarfed by PHS, CAES (compressed-air energy storage) is the second biggest bulk energy storage scheme. CAES is an effective way of handling short term peak electrical loads. What is needed are two things:

* A sealed underground chamber or similar repository;
* Proximity to main power lines.

A schematic of a CAES system is given in Figure 16.3. During off peak hours, energy is taken from the grid to power a motor and an air compressor. The compressed air is injected into the
sealed chamber. During short periods of peak demand, compressed air from the chamber is piped up to operate a turbine, which feeds an electrical generator that is attached to the grid. The amount of energy that can be stored is fairly small, typically much less than a PHS can provide, so it should only be operated in circumstances when a power demand overload is expected, and expected to be very short term.

A plant in Huntorf, Germany with 290MW capability was first CAES plant in the world when constructed in 1978. It operates at a pressure of 100 bar with storage volume of 300,000 m$^3$.

A plant in McIntosh, Alabama can generate 110MW from a natural salt cavern with volume of over 3 million m$^3$, using pressures from 45 to 74 bar. A smaller plant in Italy produces 25MW. Other plants in Israel and another in America are in the works.

“According to RWE, a German utility, the Huntorf plant is only 42% efficient, and the one in Alabama is only slightly better. The problem is that air heats up when pressurized and cools down when expanded. In existing CAES systems energy is lost as heat during compression, and the air must then be reheated before expansion. The energy to do this usually comes from natural gas, reducing efficiency and increasing greenhouse-gas emissions [www.economist.com/node/21548495?frsc=dg%7Ca].”
“As with hydro storage, efforts are under way to adapt the basic concept of CAES to make it more efficient and easier to install. RWE is working with GE, an industrial conglomerate, and others to commercialize a compressed-air system that captures the heat produced during compression, stores it, and then reapplies it during the expansion process, eliminating the need for additional sources of heat. Having proven the theoretical feasibility of this concept, the partners must now overcome the technical hurdles, which include developing pumps to compress air to 70 times atmospheric pressure, and ceramic materials to store heat at up to 600°C. The aim is to start building a 90MW demonstration plant in Strasfurt, Germany, in 2013, says Peter Moser, the head of RWE’s research arm [www.economist.com/node/21548495?frsc=dg%7Ca].

“Several smaller outfits are also developing more efficient forms of CAES. SustainX, a company spun out of Dartmouth University’s engineering school and supported by America’s Department of Energy (DOE) and GE, among others, has developed what it calls ‘isothermal CAES’, which removes heat from the compressed air by injecting water vapor. The water absorbs the heat and is then stored and reapplied to the air during the expansion process. And rather than relying on salt caverns, SustainX uses standard steel pipes to store the compressed air, allowing its systems to be installed wherever they are needed. The firm has built a 40 kilowatt demonstration plant and is partnering with AES, a utility, to build a 1–2MW system. General Compression, a Massachusetts-based company also backed by the DOE, has developed an isothermal CAES system focused on providing support to wind farms. With the backing of ConocoPhillips, an energy giant, it is building a 2MW demonstration plant in Texas. [www.economist.com/node/21548495?frsc=dg%7Ca].”

### 16.3.3 Kinetic Storage

An effective method for storing kinetic energy is the flywheel. Flywheels have been used to smooth out the irregularities of power output of certain types of machinery. The James Watt steam engine used a flywheel to smooth the output power. I have an old Gravely model L walk-behind tractor that has one cylinder, and to avoid output pulses a large internal flywheel is employed.

The kinetic energy stored in a flywheel is given by \( E = \frac{1}{2} I \omega^2 \) where \( I \) is its moment of inertia and \( \omega \) is the angular velocity in radians/sec. The energy density of a flywheel system is of the order of 120Wh/kg, about the same as lithium ion batteries. Further, unlike batteries, they can charge and discharge at a high rate. Since most power interruptions last for a few seconds, flywheels can effectively bridge this problem.

Until recently, steel flywheels with traditional mechanical bearings were the only game in town; these are low speed, measured in thousands rpm. Now, carbon-fiber rotors on magnetic bearings using magnetic levitation are the method of choice: Because of their tiny frictional losses in their bearings they can spin at 50,000 rpm. Air friction could be the dominant
component of loss, but this is eliminated by encasing the flywheel and its electrical components inside a vacuum chamber.

We are all familiar with a permanent magnet, typically made of iron, a ferromagnetic material. The material is not naturally magnetic, but becomes so under a strong magnetic field. It can be demagnetized under an alternating magnetic field of diminishing strength. Soft iron is easily magnetized, but rapidly loses its magnetism when the enabling magnetic field is removed.

The ability of a material to be magnetized, and the type of magnetism, is a function of the tendency of electrons in the valance band to form magnetic dipoles, and this is dependent on the type of motion of those electrons. Before iron is magnetized, the dipoles are randomly oriented. After magnetism the dipoles line up, creating what we call a north pole and a south pole. Unlike poles attract, while like poles repel. A magnetic field flows between the poles, and can be observed with the iron filing experiment. Place a bar magnet under a horizontal sheet of paper, sprinkle iron filings on the paper, and tap the paper so the iron filings jump. The filings realign into lines that show the pattern of the magnetic field.

The term dipole is unfortunate since it implies that there exist monopoles. A bar magnet has two poles, north and south. Cut it in two pieces and you get two bar magnets, each with a north pole and a south pole.

Levitation should be simple, or so we think! Since like magnetic poles repel, then all we need to do is to set up a system of like poles and we have a frictionless system. Unfortunately, the physics does not cooperate, and the problem was exposed by the theorem presented in 1842 by Samuel Earnshaw [“On the Nature of the Molecular Forces that Regulate the Constitution of the Luminiferous Ether,” Transactions of the Cambridge Philosophical Society, vol. 7, pp. 97–112, 1842]. Actually, it is the corollary of the theorem that is the spoiler: stable equilibrium means that to move some matter a little from its static position, then there must be a force that returns it to its original state. If that force does not exist, there is no stable state. The system of like magnetic poles has no such force and so no stable state.

Magnetic levitation is possible because of something called diamagnetism. There are three types of magnetic materials whose electrons react differently to a magnetic field. In ferromagnetic materials some of the electrons spin in parallel and so retain some of their magnetism when the magnetic field is removed; the most common ferromagnetic materials are iron and nickel. In paramagnetic materials the material becomes magnetic in the presence of the magnetic field, but loses all of this magnetism when the field is removed; the most common paramagnetic materials are aluminum, platinum and uranium.

We know that a bar magnet attracts a ball of soft iron. That is, put the ball close to the north pole of a magnet and it is attracted. Put the ball close to the south pole of the magnet and it is attracted. However, put some diamagnetic material near the bar magnet, either north or south pole, and it is repelled, and this force of repulsion can produce the phenomenon known as magnetic levitation. Carbon, copper, and lead are weakly diamagnetic, but pyrolytic graphite is
strongly diamagnetic, so placing a plate of this material over a strong rare-earth magnet causes levitation.

The bar magnet is made of ferromagnetic material, with dipoles that are aligned to create a north pole and a south pole. Apply a magnetic field to paramagnetic material and a weak force of attraction occurs, and no magnetic dipole moment is produced. Similarly, the atoms in diamagnetic material have no permanent magnetic dipole moment, and when a magnetic field is applied, a weak magnetic force is produced in the opposite direction to the applied field, so a repulsive force results. It is interesting that most materials have diamagnetic properties, but the force is too weak to be observed, and certainly too small to overcome the force of gravity. However, place the diamagnetic material in a strong magnetic field and levitation can result.

A popular experiment in levitation is to place a small frog, which is weakly diamagnetic, on top of an electromagnet, and it is levitated.

The lifetime of flywheels with magnetic levitation is effectively unlimited. They have high energy density of 100–300Wh/Kg. Efficiencies, the ratio of power out to power in, can be as high as 90% Modern flywheels store from 3kWh to 130kWh.

A number of companies offer flywheel energy storage systems. Beacon Power of Tyngsborough, Massachusetts marketed a 25kW flywheel that spins between 8,000 and 16,000 rpm. At the higher speed the rim of the flywheel travels at the equivalent of about 1,500 mph. The unit is designed for a 20-year life with hundreds of thousands of charge/discharge cycles. Vycon’s flywheels are capable of rotating at up to 36,750 rpm when fully charged, and down to 10,000 rpm at full discharge: This flywheel can be cycled from full charge to discharge and full charge again in 60 seconds. Innovative technologies employed by these or other companies reduce iron losses in the electrical motor/generator.

Unfortunately, Beacon Power, the biggest player in the game, did not have a sustainable business plan despite $43 million in government assistance and a contract for $29 million to build a 20 MW plant in Pennsylvania. They filed for bankruptcy in November 2011. Rocklans Capital, a private equity firm, agreed to purchase the assets of Beacon Power in 2013 for $30.5 million, and so return to the government about 70% of the government loan.

Flywheel energy storage is not suitable for long or bulk storage by power plants such as occurs with hilltop reservoirs. Instead, their use is to smooth out peaks and troughs in energy demand. Flywheels could become the method of choice for residences that have no grid access, and it is possible that the flywheel discussed in the next paragraphs will be the vehicle used.

A new flywheel design by Silicon Valley guru Bill Grey called Velkess (for Very Large Kinetic Energy Storage System) shows serious promise due to its projected cycle efficiency of 80% and cost competitive with hilltop reservoirs. “The Velkess improves on traditional flywheels by better managing the natural ‘wobble’ of a spinning mass. Traditional flywheels have been very expensive because engineers align the natural axis of the wheel’s rotation with the desired
rotation of the generator. Thus, they are always struggling to minimize the natural wobble of the wheel using very expensive magnets and bearings, high-precision engineering and materials like high-grade carbon fiber or rigid steel. Beacon’s flywheel for grid storage cost a whopping $3 million per MWh [C. Nedler, “Turn Up the Juice: New Flywheel Raises Hopes for Energy Storage Breakthrough,” Scientific American, April 2013].

“Instead of trying to fight the wobble, Gray redirected it by suspending the wheel within a gimbal — the same concept that makes a gyroscope work. The gimbal in the Velkess is asymmetrical, so the two axes of rotation — the flywheel axis as well as that of the rotor, which drives the brushless, inducting DC motor — are not on the same plane, and have different periods of frequency. This dampens the resonance effects that make traditional flywheels hard to control (a resonant disturbance in one of the planes can intensify until the device shatters). With the gimbal, resonance in one plane is translated into the other, which is non-resonant at the same frequency. Accordingly, only very loose engineering tolerances — about one sixteenth of an inch — are required to build the device [op cit].

Based on the research work of John Vance, a retired professor at Texas A&M University, Velkess uses cheap fiberglass, so the unit cost is projected to be about 1/10 the cost per Kwh stored of the Beacon flywheels. The rotor is designed to flex with the result that less energy is lost to friction. The discharge rate is 3kW until it is fully discharged at 9,000 rpm. The 15kWh stored will power a small ultra-tight residence.

16.4 Thermal Storage Systems

Thermal energy storage can occur using three different techniques. The simplest is sensible heat storage where heat is added when end user heat demand is low, and heat is extracted to meet user demand. Phase change material takes advantage of the large energy transfer needed to change from one state to another; these are sometimes called latent heat storage systems.

In these two methods the thermal storage is local. That is, it is not suitable for a large commercial facility at a remote location. It is used extensively in residential passive solar energy systems as discussed in Chapter 11. It could also be employed in large commercial buildings in regions in which diurnal temperature swings are high, so storing excess solar heat during the day reduces the direct heating effect of the Sun, while the low nighttime temperatures can be bridged by retrieving the daytime heat stored.

The third method is geothermal, and here the thermal energy is commonly taken out, not pumped in. However, the temperate geothermal systems employ a heat pump that can extract heat from the ground for space heat of the structure in winter, and can pump heat back into the ground for air conditioning (to mitigate the unwanted solar heat) in summer.
16.5 Electrical and Electro-Mechanical Storage Systems

16.5.1 A Brief History of Electricity

Magnetism was discovered before electricity arrived on the scene. The Chinese invented the magnetic compass during the Qin dynasty between the 200 years before and after the Christian era. The word electric was coined by William Gilbert, physician to the royal court of England, in his work *De Magnete* published about 1600: In this work he separated static electric forces from magnetic forces.

An early discovery in the field of stored electricity was the Leyden jar, a glass jar filled with water and surrounded with metal foil outside and inside and having a rod in the insulating stopper to touch the water. The metal foils do not touch. The jar is charged electrostatically. Rubbing a comb through dry hair produces such a charge. Walking over a carpet in a house with low humidity can result in electrostatic shock when touching an electrically grounded conductor, such as a metal faucet. The Leyden jar was the first electrical capacitor.

Much earlier in history, and basically lost as a building block on which to build knowledge, was Thales of Miletus (640–546 BC) who observed that when amber was rubbed it could attract light objects. Interestingly, the word electricity derives from “elektron,” Greek for amber.

Englishman Henry Cavendish (1731–1810) while experimenting with capacitors, discovered the principle of electrical potential (voltage), something that he measured with a gold leaf electrometer, but he had no means of measuring electrical current. It was left to successors, such as Oersted, Coulomb, Ohm, Faraday, and Volta to flesh out understanding of the relationship between charge (coulombs), current (amperes), and voltage.

Johann Karl Friedrich Gauss (1777–1855) born in Braunschweig in Germany, had his name given to the unit of magnetic flux. Gauss’s Law, formulated in 1835 but not published until 1867, can be stated as: The electric flux passing through a closed surface is proportional to the enclosed charge. Wilhelm Eduard Weber, born in Wittenberg, Saxony in 1804, was a follower of Gauss, and it is the unit of magnetic flux that carries his name.

Hans Christian Oersted of Denmark was born in 1777, and in the winter of 1819–1820 he found the link between electric current flow in a wire and magnetism by discovering, during a series of lectures, that a current-carrying wire could deflect a magnetic needle.

Frenchman Andre-Marie Ampere (1775–1836) used Oersted’s discovery to show how a force is created between two current-carrying wires. He then developed the relationship between electricity and magnetism now known as Ampere’s Law: The electric flux through any closed
surface is proportional to the total charge enclosed within the surface. A magnetic equivalent, also known as Ampere’s Law, is: The net magnetic flux out of any closed surface is zero. The unit of current, the ampere or amp for short, is named in his honor.

Georg Simon Ohm, born in Bavaria in 1789, while experimenting with the newly discovered field of electromagnetism by determining the extent of current passing through a conductor (by measuring the magnetic flux it created), found that the type of conductor affected the flux. This led to the all important Ohm’s Law: The voltage $V$ across a current-carrying conductor is the product of the current $I$ passing through the conductor and its resistance $R$, or $V = I R$.

Alessandro Volta (1745–1829) in 1800 soaked paper in salt water and placed plates of zinc and copper on opposite sides of the paper, noting that the reaction produced electricity. The unit of electrical potential, the volt, is named after Volta. In 1807 Humphrey Davy showed that the reaction was chemical since pure water would not produce electricity, and so he argued that chemical reactions are electrical in origin. Humphrey Davy was born in Penzance, Cornwall, in 1778 and was considered to be Britain’s greatest scientist, but is possibly honored most for his inspiration to others.

Most notable amongst those inspired was Englishman Michael Faraday (1791–1867), who discovered electromagnetic induction, the foundation of electromagnetic field theory, whose governing equations were codified by Maxwell. One of these laws, Faraday’s Law, can be stated: The EMF (electromotive force, voltage for short) generated in a current-carrying wire is proportional to the rate of change of magnetic flux.

In 1821 Faraday demonstrated magnetic flux lines by sprinkling iron filings on a piece of paper resting on a bar magnet, then tapping the paper to make the filings jump and fall into lines of flux — this is the iron filing experiment so common in high school physics classes. He also made the first electric motor in the same year, the first electric generator a decade later. He also invented the transformer and the solenoid.

The first to coin the word “battery,” as well as a number of terms used to describe electrical quantities and devices that have become standard today, was Benjamin Franklin, in around 1752.

At the beginning of the nineteenth century a number of experimentalists, such as Nicholas Gautherot of France in 1801, and William Cruikshank of England in 1802, produced crude lead-acid batteries.

Charles-Augustin Coulomb (1736–1806) was born in France and educated as a military engineer. He invented the torsion balance that was far more sensitive than existing devices. However, his most important discovery was in the field of magnetism. His formula, known as Coulomb’s Law, defines the force $F$ between two point charges $q_1$ and $q_2$ separated by distance $r$ as $F \propto q_1 q_2 r^2$. Cavendish independently discovered the same law before Coulomb, but omitted to publish it, so history awards Coulomb the honor.
The unit of electric charge is the Coulomb (abbreviated C). Ordinary matter is made up of atoms that have positively charged nuclei and negatively charged electrons surrounding them. The charge on a single electron is $-1.602 \times 10^{-19}$.

English physicist James Prescott Joule (1818–1889) was the first to establish that energy could exist in a number of forms, and that they were interchangeable; this showed the path to the formulation of the first law of thermodynamics. He established that the heat produced by current $I$ in a wire of resistance $R$ is $I^2R$. With William Thompson, later Lord Kelvin, he found that when a gas expands its temperature drops, a discovery that led to refrigeration. The unit of energy carries his name, the joule.

It was Lord Kelvin (1824–1907) who established the first and second laws of thermodynamics. His name is given to the Kelvin temperature scale and the determination of the lowest possible temperature, absolute zero. His work helped Maxwell formulate his famous laws.

Samuel Morse invented the electric telegraph in 1844, and for the first time in history semi-instantaneous long-distance coded communication became possible. Alexander Graham Bell (1847–1922) invented the telephone in 1876 with patent number 174,465. He constructed the first telephone exchange in 1878. By 1884 Boston and New York City were connected by telephone. Interestingly, Bell offered his invention to the Western Union telegraph company, but they saw no future in the telephone and turned him down. The Bell Telephone company grew into the giant AT&T, and Western Union languished.

Perhaps the two most influential persons to mathematically describe electrical phenomena are Kirchhoff and Maxwell. Gustav Robert Kirchhoff (1824–1887) produced two fundamental principles in 1846, now known as Kirchhoff’s laws, that define how electrical circuits work:

* The algebraic sum of electrical currents entering a node is zero
* The algebraic sum of voltages in a loop is zero

where a node is the joint of connectivity between two or more circuit elements, such as resistors, capacitors, and inductors. Nodes also could be transistors or other non-symmetric devices, and a loop is a closed path of circuit elements in which no circuit element is traversed more than once to arrive back to the starting point.

In 1864, Scottish scientist James Clerk Maxwell (1837–1879) consolidated and clarified prior work by Gauss, Ampere, Faraday, and others to produce what are now known as Maxwell’s equations, the foundation of electromagnetic field theory — by providing a unifying theory for magnetism, electricity — and light.

Nicola Tesla (1856–1943) born to Serbian parents, invented the “Tesla coil,” the transformer that could convert low voltage power to high voltage power that substantially reduced resistive losses and enabled power to be transmitted over long distances. He also invented the
AC electrical generator, which George Westinghouse (1846–1914) used to produce AC power in the world’s first major hydroelectric power station at Niagara Falls in 1895.

There are a number of units to describe magnetic flux. All are based on Faraday’s law, which can be presented mathematically as

$$\varepsilon = d\phi/dt \quad (16.1)$$

where $\phi$ is the magnetic flux, so $d\phi/dt$ is the rate of change of that flux. This means that the voltage induced in the open circuit ends of conductive wire in a circle in a magnetic field decreasing at the rate of 1 weber/second is 1 volt. If the conductive wire completes $n$ loops, then the voltage produced across the ends is $n\varepsilon$.

The SI unit for magnetic flux density is the tesla, defined as 1 tesla = 1 weber/m$^2$. The outdated cgs (centimeter-gram-second) unit is the gauss, defined as 1 gauss = 1 maxwell/cm$^2$, and 1 gauss = 10$^{-4}$ tesla.

We have reason to be grateful to Thomas Alva Edison (1847–1931) for the invention of the incandescent light bulb, as we have reason to condemn these energy hogs. Prolific inventor though he was, Edison occasionally got it wrong: He backed the wrong horse when he bet that the future lay in direct current power stations rather than alternating current.

Other giants in electricity include:

* Heinrich Hertz (1857–1894) who was the first to transmit the radio waves predicted, but never demonstrated, by Maxwell. This proved that light was a form of electromagnetic energy. The frequency of an alternating current is measured in Hertz.
* Guglielmo Marconi (1874–1937) who shared the Nobel prize in physics with Ferdinand Braun in 1909. In 1901 he transmitted a radio signal from Poldhu, Cornwall, to St. John’s, Newfoundland.
* Albert Einstein (1879–1955) developed the theory of relativity that tied energy to mass in the simple equation $E = mc^2$, which led to the atomic age, and, to his great regret, the atomic bomb.

The greatest invention of John Ambrose Fleming (1849–1945) was the thermionic valve, also known as the vacuum tube. He termed it a valve since it allowed current flow in one direction only.

The chief claim to fame of Lee de Forest (1873–1961) was his invention in 1906 of the Audion, which is now called the triode tube, the first electronic device that could take an input signal and amplify it. This made radio possible, taking a small input signal from an antenna and amplifying it until it was strong enough to power loudspeakers.

It was the triode that derailed the nascent semiconductor development. In particular, the semiconductor P-N junction was known at the beginning of the twentieth century, yet it took
half a century before three scientists at Bell Telephone Laboratories put 2 and 2 together by joining two P-N junctions, to create the point junction transistor.

In 1945, a conference at Bell Telephone Laboratories discussed the uses and future of semiconductors, and a solid state group led by William Schockley and Stanley

Figure 16.4: The Vacuum Tube
Morgan was formed. Schockley postulated that it could be possible to construct a semiconductor triode. His early experiments failed, but it was John Bardeen who suggested that electron flow was trapped under the semiconductor surface. A third member, Walter Brittain, joined the group, and it was his idea that demonstrated a small but measurable gain. The point contact device with one forward biased P-N junction joined to another reverse biased P-N junction was demonstrated to senior management in 1947, and the transistor was born. Bardeen, Schockley, and Brittain shared the Nobel prize in physics in 1956.

In 1948, a memorandum invited selected staff members at Bell Telephone Laboratories, including the inventors of the device, to produce a name for the semiconductor triode. The winning word, transistor, is a compaction of the three words transductance, transfer, and varistor.

Great though the transistor was, in its discrete state it needed to be replaced by what is now called the integrated circuit — a monolithic crystal containing all the circuit elements and interconnections. Two electrical engineers, working independently, obtained patents on the integrated circuit. Robert Noyce, as co-founder of Fairchild Semiconductor Corporation, using silicon, was awarded patent #2,981,877, and Jack Kirby of Texas Instruments, using germanium, was awarded patent #3,138,742: They filed for these patents in 1959. Their two companies, after battling for supremacy for several years, decided to collaborate by cross-licensing their technologies. The global market in integrated circuits is estimated at over $1 trillion annually,
Some of the great names in the history of electricity are given in the timeline shown in Figure 16.6.

16.5.2 Supercapacitors

A capacitor has two metallic plates separated by an insulating dielectric. It is an electrostatic device that stores energy. \( E = \frac{1}{2} CV^2 \) when there is a voltage difference \( V \) between the two metallic plates: \( C \) is the capacitance of the device measured in Farads, and this is a function of the area of the metallic plates, the dielectric constant, and the thickness of the dielectric film. The metallic plates are very thin so that the sandwich of plate-dielectric-plate-dielectric can be rolled into a tight cylinder.

Supercapacitors use two layers of porous electrodes, rather than smooth metallic plates, to increase the surface area. These electrodes are suspended in an electrolyte solution and separated by a very thin dielectric. The result is a much larger capacitance for a given volume. Another advantage of supercapacitors is that, unlike batteries, they have an almost unlimited cycling.

Now a major improvement in supercapacitor construction has been announced. An edited version of the report by Sebastian Anthony of March 19, 2012
A team of international researchers have created graphene supercapacitors using a LightScribe DVD burner. These capacitors are both highly flexible and have energy and power densities far beyond existing electrochemical capacitors, possibly within reach of conventional lithium–ion and nickel metal hydride batteries.

The team, which was led by Richard Kaner of UCLA, started by smearing graphite oxide — a cheap and very easily produced material — films on blank DVDs. These discs are then placed in a LightScribe drive (a consumer-oriented piece of gear that costs less than $50), where a 780nm infrared laser reduces the graphite oxide to pure graphene. The laser-scribed graphene (LSG) is peeled off and placed on a flexible substrate, and then cut into slices to become the electrodes. Two electrodes are sandwiched together with a layer of electrolyte in the middle — and voila, a high-density electrochemical capacitor, or supercapacitor as they’re more popularly known.

Now, beyond the novel manufacturing process — the scientists are confident it can be scaled for commercial applications, incidentally — the main thing about LSG capacitors is that they have very desirable energy and power characteristics. Power-wise, LSG supercapacitors are capable of discharging at 20 watts per cm$^3$, some 20 times higher than standard activated carbon capacitors, and three orders of magnitude higher than lithium–ion batteries. Energy-wise, we’re talking about 1.36 milliwatt-hours per cm$^3$, about twice the density of activated carbon, and comparable to a high-power lithium–ion battery.

These characteristics stem from the fact that graphene is the most conductive material known to man — the LSG produced by the scientists showed a conductivity of 1738 siemens per meter (yes, that’s a real unit), compared to just 100 siemens for activated carbon. The performance of capacitors is almost entirely reliant on the surface area of the electrodes, so it’s massively helpful that one gram of LSG has a surface area of 1520 square meters (a third of an acre). As previously mentioned, LSG capacitors are highly flexible, too, with no effect on its performance.

These graphene supercapacitors could really change the technology landscape. While computing power roughly doubles every 18 months, battery technology is almost at a standstill. Supercapacitors, which suffer virtually zero degradation over 10,000 cycles or more, have been cited as a possible replacement for low-energy devices, such as smartphones. With their huge power density, supercapacitors could also revolutionize electric vehicles, where huge lithium–ion batteries really struggle to strike a balance between mileage, acceleration, and longevity. It’s also worth noting, however, that lithium–ion batteries themselves have had their capacity increased by 10 times thanks to the addition of graphene. Either way, then, graphene seems like it will play a major role in the future of electronics.”
16.5.3 Superconducting Magnetic Energy Storage

Energy can be stored in the magnetic field of a coil of superconducting material cooled cryogenically; the temperature must be below its superconducting critical temperature. These are SMES (superconducting magnetic energy storage) systems.

The salient characteristics of an SMES system is low energy density, about 10 kJ/kg, but with a very high power density. This makes it a good candidate to smooth out transient pulses on an electrical grid. They have a fast charge, and an even faster discharge — capable of megawatts semi-instantaneously. Their round-trip efficiency is greater than 95% [P. Tixador, “Superconducting Magnetic Energy Storage: Status and Perspective,” IEEE/CSC & ESAS European Superconductivity News Forum, No. 3, January 2008].

The high capital cost is the primary inhibitor to the widespread use of SMES systems. This quote from a recent MIT online report on superconducting magnets for grid-scale storage summarizes the problem. “SMES has long been pursued as a large-scale technology because it offers instantaneous energy discharge and a theoretically infinite number of recharge cycles. Until recently, however, the material costs for SMES devices have been prohibitively high for all but very small applications. Now a project funded by the U.S. Department of Energy (DOE) could pave the way for SMES technology that offers megawatt hours of energy storage. Such capacity is becoming increasingly necessary for electricity grids that need to balance the intermittency of renewable energy sources [www.technologyreview.com/s/423227/superconducting-magnets-for-grid-scale-storage].”

16.6 Batteries and Battery Technology

16.6.1 The Focus

Sandia National Laboratory in Albuquerque was tasked with developing “Batteries for Specific Solar Applications” in 1978. Additional responsibilities involving battery technologies have been added over the years, including the Exploratory Battery Development & Testing Program (EDT) in the early 1980s. EDT became the Utility Battery Storage Program (UBS) in 1991, which became the Energy Storage Systems Research Program (ESS) in 1996. ESS works closely with universities and industries, as well as other governments.

Without doubt, a considerable effort involving massive sums of money and substantial brain power has been devoted to improving existing battery technology and developing new technologies. The main goal is to increase energy density. A secondary goal is to improve the cycle life limitation: When a battery is partially or totally discharged its life is shortened. For
example, the lead acid battery commonly used in automobiles is seriously damaged if totally discharged.

16.6.2 Common Batteries

A battery is an electrochemical cell in which chemical energy is converted to electrical energy. The chemical used to fuel a battery is in a sealed container and is not replenished, unlike the fuel cells in which chemical fuel is continually added. The configuration of a lead-acid cell, the type commonly used in automobiles, is shown. The electrolytic fluid is a solution of sulfuric acid. The positive terminal is connected to the lead-peroxide anode submersed in the fluid. The negative terminal is connected to the lead cathode also immersed in the fluid.

Consider the decomposition of two molecules of sulfuric acid into four hydrogen cations and two sulphate anions. When an external circuit is completed between the positive and negative terminals, one molecule of lead-sulphate is deposited on the anode by the chemical action of four cations of hydrogen and one sulphate anion. Two molecules of water are also produced, so the fluid becomes less acid. This chemical action requires two electrons. At the cathode the remaining sulphate anion combines with lead to form lead sulphate and release two electrons. The chemical and electrical balance is shown.

\[
\begin{align*}
2H_2SO_4 & \rightarrow 4H^+ + 2SO_4^- \\
\text{anode} & \rightarrow PbO_2 + 4H^+ SO_4^+ + 2\varepsilon = PbSO_4 + 2H_2O \\
\text{cathode} & \rightarrow Pb + SO_4^- = PbSO_4 + 2\varepsilon
\end{align*}
\]

When the cell is charged, the process is reversed, sulfuric acid is formed, lead peroxide deposited on the anode, and lead deposited on the cathode.
The fully charged voltage of the lead-acid cell described is between 2.09 and 2.15 volts. The automotive lead-acid battery is made up of six cells in series to produce a little over 12 volts at its terminals. The battery is designed for a long, reliable life, but a deep discharge severely limits its life, particularly during freezing weather, since the freezing point of the electrolytic
fluid is highest when the battery charge is lowest. The lead-acid battery has an energy density of about 10.5 watt-hours per pound weight.

The acid contained in the lead-acid battery makes it unsuitable for uses in which the orientation of the battery is variable, such as in a flashlight. A dry cell such as the Leclanché dry cell shown is used in this application. The core of the Leclanché dry cell is a mixture of ammonium chloride (the electrolyte), carbon, and manganese dioxide. The output of this cell is 1.54 volts with an energy density of 20 Wh/kg, provided currents are kept low and intermittent.

16.6.3 Battery Types

There is a broad field of batteries, but a limited field to satisfy a particular application. Here we look at batteries that have the best energy density accompanied with economy per watt-hour stored. This eliminates batteries for watches, calculators, or hearing aids. Those remaining, together with their typical applications, include the following:

The lead acid battery is over 150 years old. Its primary application is in automobiles and trucks since it can deliver a high current to power starter motors, and the weight of the battery is not of primary concern. It is reliable and inexpensive. It must be kept upright to prevent the acid in the battery from leaking out, which precludes its use in portable tools or flashlights.

It also off-gasses hydrogen and oxygen, particularly during a charging cycle, which is not a major problem in a truck but of major concern if a bank of them are used as a primary means to store intermittent power from renewables. The gel cell lead acid battery reduces or eliminates this off-gassing, but the charging rate must be reduced to avoid damage.

There are many variations of lead acid batteries, of which a few are listed here. Additives such as antimony, calcium, and selenium can improve battery performance, but it will still retain most of the characteristics of the standard automotive battery. Deep-cycle lead acid batteries have applications in forklift trucks and golf buggies; they require stronger plates since charging these produces more heat than the conventional battery.

Leclanché cells, also known as dry cells or zinc-carbon cells, produce 1.5 volts. They are the familiar, low-cost D, C, AA, and AAA batteries used in flashlights, clocks, toys, and remote controls. Their biggest problem is their tendency to leak, particularly when left in an unused device over a long period when discharged; the fluid leaked is corrosive and can damage the device. Also, they do not take recharging well.

Leclanché cells do not perform well at low temperatures. I need to change the AAA batteries every few months in winter in an electronic exterior thermometer that sends its signal to an interior display screen, which also reads the interior temperature; the AAA batteries in the interior device last for years.
Alkaline batteries, also known as alkaline manganese dioxide primary cells, deliver 1.5 volts, and are the alternative to Leclanché cells. They have twice the capacity, a much longer life, and some can be recharged. Until relatively recently they suffered from the “memory effect,” meaning that the battery could not be fully recharged. This problem has now been overcome. They also perform well at low temperatures, so I should convert to alkaline for my remote thermometer.

NiCad, nickel cadmium batteries, were once popular in portable tools due to their small size and high power output when needed. They have now lost out to nickel metal hydride and lithium batteries.

The premier application for NiMH, nickel metal hydride batteries, is in digital cameras, MP3 players, and GPS units. They are rechargeable and can withstand a large number of charge/recharge cycles with little memory problems, but have an overall lifetime, regardless of use, of about five years. Thus, they are the battery of choice when the device they are in is in continual use.

NiZn, or nickel zinc, batteries have an output of 1.65 volts, and have good cycle life. They can also be rapidly charged, and can be fully discharged without damage, so they are popular in lawnmowers and electric scooters where their heavy weight is not of paramount importance.

NiFe, nickel iron or nickel alkaline batteries, have a long life of about 20-plus years if not physically abused. They can be fully discharged for long periods and can endure a large number of charging/recharging cycles without damage. However, they have a low voltage, will self discharge at a rate higher than most other batteries, and are bulky. Their principal application is in forklift trucks that are typically recharged daily.

There is a large family of Li-ion batteries, including LiCoO2, LiFePO4, LiMn2O4, LiNiCoAlO2, LiNiMnCoO2, and Li4Ti5O12. Lithium is the lightest of metals, and is the most chemically reactive, so it has substantial advantages as well as presenting new challenges. The worldwide market is over $11 billion and is expected to be about $60 billion by 2020.

Lithium ion batteries have high energy density, long shelf life, and the ability to work at low temperatures. They have a low self-discharge rate, and have no recycling memory problems, but do have a poor recycling record. They also have the propensity to overheat and possibly explode when they are driven hard. In 2013 the Boeing Dreamliner fleet was temporarily grounded because of fires in their lithium ion batteries. Also, they have the propensity to explode when exposed to moisture.

Japan is experimenting with flow batteries that pump chemicals from storage tanks to the active part of the battery to generate electricity.
16.6.4 Energy Density in Batteries

The CPUC (California Public Utilities Commission) was required by law in October 2010 to establish energy storage targets. They set the target for 2021 at 1.325 GW. Other states will follow the lead of California since it is recognized that the uncertainty and intermittent nature of renewables, particularly PV and wind, require a much larger percentage of energy storage capacity than we presently have [“Grid Energy Storage,” U.S. Department of Energy, December 2013].

Worldwide there are considerable efforts to improve the energy density of batteries. Shown in Table 16.4 is the number of initiatives and the funding in millions of dollars for fiscal years 2009 through 2012 in the United States [op. cit.].

In April 2015, Tesla Motors CEO Elon Musk, announced that his company had produced Powerwall, a wall-mounted lithium ion home battery approximately 1.2 x 0.9 meters and 0.2 m thick. Its cost was $3000 for the 7 kW model, and $3500 for the 10 kW model, not including inverters and power connections [www.teslamotors.com/powerwall] It is intended to be charged predominantly from solar panels and has a backup electricity supply, and is grid connected. It automatically switches to its own power during blackouts. We discussed fault ride-through capability for DARs in Section ?? of Chapter 14. If the utility companies require such DARs to have this capability, it will require backup capability, and the Powerwall could be a good candidate for this.

There is considerable interest in the California start-up company Seeo, particularly since Seeo has attracted a number of well-heeled investors, including the Korean electronics giant Samsung. The objective of this start-up is achieving an energy density of 400 Wh/kg, about double the energy densities for existing batteries in EV vehicles [R.L. Hales, “Will Seeo’s 400 Wh/kg Battery Fulfil Expectations?” http://cleantechnica.com/2014/12/28/will-seeos-400-whkg-battery-fulfil-expectations, December 2014]
The tech company called Envia Systems has developed a lithium-ion battery with an energy density of 400 Wh/kg, considerably higher than competing batteries at 245 Wh/kg; for example, the batteries in the Nissan Leaf are about 120 Wh/kg. Compare this to gasoline, with about 12,000 Wh/kg [www.openthefuture.com/2012/02/record_battery_energy_density].

A research team at the University of Tokyo School of Engineering has developed a new lithium ion battery with energy density of 2,570 Wh/kg. This is about seven times the energy density of current lithium ion batteries. What the team did was add cobalt to the lithium oxide crystal structure, which enhances oxide and peroxide production, leading to superior performance. The research team also claimed faster recharge times. [C. DeMorro, “New Battery Boasts 7 Times More Energy Density,” http://cleantechnica.com/2014/07/30/new-battery-boasts-7-times-energy-density, July 2014].

A battery technology with a lot of promise and a lot of problems is the lithium/air battery, with a theoretical energy density of 5000 Wh/kg. Another is the lithium/seawater battery based on protected lithium electrodes. The PolyPlus Battery Company has developed protected lithium electrodes that are remarkably stable and with the energy density potential of 11,000 Wh/kg [www.polyplus.com/technology.html].

16.7 Fuel Cells

A fuel cell can be the reverse of the electrolysis process discussed in the prior section. That is, hydrogen and oxygen are combined in the fuel cell to produce electricity as shown in Figure 16.9. The porous anode and cathode permit hydrogen on one side and oxygen on the other to join in the central electrolyte to produce water. The energy released is mainly electrical. The energy balance is 285.83 kJ/mol of hydrogen producing 237.13 Kj/mol electricity and 48.7 Kj/mol waste heat. The energy efficiency is 237.13/285.83=0.8296, nearly 83%, far higher than burning hydrogen to power an electrical generator. The only problems are

* how to reliably make a large scale fuel cell,
* how to efficiently produce and store hydrogen.

Fuel cells have been predicted to have a significant role in producing green energy. However, technical problems are more difficult to overcome than its advocates fore-saw. The first commercial successes are to be found in the combined heat and power (CHP) systems developed in Japan and Germany.

The efficiency of the residential CHPs developed by Panasonic Corporation in association with Asaka Gas in producing electricity is only 35%, less than the 40% efficiency produced by the power grid, but the CHP system captures most of the 65% waste heat produced to heat the residence. The German government spearheaded the Callux residential CHP system in 2008.
16.8 Hydrogen Generation and Storage

50 million tonnes of hydrogen are produced annually, and with a growth rate of 10% per year. About half of this is used to make ammonia that is used to produce fertilizers. The other half is used in hydrocracking — converting heavy petroleum into lighter fractions to be used as gasoline or other fuels.

The State of the Union speech of President George W. Bush in 2003 touted hydrogen as a fuel of the future, and anticipated mass-produced hydrogen-powered cars by 2020. In response, the U.S. Department of Energy allocated $318 million for fuel cells and hydrogen production. However, the National Academy of Sciences considers the 2020 goal “unrealistically aggressive.”

The sources to produce hydrogen are fossil fuels, with almost half from natural gas. Only 4% is produced using electrolysis, the process of passing a DC current from one electrode, the positive terminal called the anode, through an electrolyte solution to another electrode, the negative terminal called the cathode. Hydrogen is released at the cathode. An electrolytic cell using ammonia $NH_3$ as the electrolyte is shown in Figure 16.10. As far as energy balance is concerned, unless the source of electricity is inexpensive (as was touted in the early days of nuclear power) or transient (use it or lose it), electrolysis is not destined to play an important
part in the energy needs of the nation. Liquid ammonia is 17.8% hydrogen by mass. The chemical reaction can be described by

\[
2NH_3 \rightarrow 3H_2 + N_2
\]

Water can also be the electrolyte in place of ammonia. Regardless, efficiency of hydrogen production by electrolysis is low. A possible alternative is to use the Sun to split water. What is required is a solar cell with a large bandgap, somewhat higher than that of a silicon PV cell. Perhaps perovskites, discussed in Section VIII of Chapter 13, could one day do the job with respectable efficiency.

It takes a lot of energy to produce hydrogen as a fuel source. At one time it was assumed that nuclear fission would produce electricity so cheap it was not worth metering it. Now we know better. Fission has another major problem, and that is the radioactive byproducts with an almost infinite half life. Now the promise for the future could be nuclear fusion (the power of the Sun) with the promise of more benign radioactive byproducts with short half lives. However, success in producing self-sustaining nuclear fusion is always predicted to be 20 years away.

On May 6, 1937, the Hindenberg Zeppelin was attempting to land at Lakehurst Naval Air Station in New Jersey when electrostatic charge on its exterior ignited its paint, which then ignited its hydrogen. The dirigible exploded, killing 36 people. This was the end of passenger-carrying airships.
Hydrogen has a high energy density by weight, but low energy density by volume, as Table 16.5 shows [L. Becker, “Hydrogen Storage,” available online at www.csa.com/discoveryguides/hydrogen].

Storage is a problem with hydrogen since it is a tiny molecule. If hydrogen is to be used to fuel vehicles it needs to be stored at ambient temperatures in a tank capable of handling high pressure (up to 700 bar) without leakage over at least 1000 cycles. This may be unrealistic. Instead, reversible chemicals such as metal hydrides (alanates) $NaAlH_4$ or $AlH_4$ including, which can be stored at atmospheric pressure and ambient temperature, could be the answer.

<table>
<thead>
<tr>
<th>material</th>
<th>H atoms/cm³ x10^22</th>
<th>% H by weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_2$ gas at 200 bar</td>
<td>0.99</td>
<td>100</td>
</tr>
<tr>
<td>$H_2$ liquid at 20° K</td>
<td>4.2</td>
<td>100</td>
</tr>
<tr>
<td>$H_2$ solid at 4.2° K</td>
<td>5.3</td>
<td>100</td>
</tr>
<tr>
<td>$MgH_2$</td>
<td>6.5</td>
<td>7.6</td>
</tr>
<tr>
<td>$Mg_2NiH_4$</td>
<td>5.9</td>
<td>3.6</td>
</tr>
<tr>
<td>$FeTiH_2$</td>
<td>6.0</td>
<td>1.89</td>
</tr>
<tr>
<td>$LaNi_5H_6$</td>
<td>5.5</td>
<td>1.37</td>
</tr>
</tbody>
</table>
Index

A

AAC block wall, see Autoclaved aerated concrete block wall
Active solar collectors, see Non-concentrating, active solar collectors
Air gap, heat flow across, 93–98
Air infiltration, 148–155
   ACH50, 151
   air leakage from conditioned spaces, air leakage from, 150–152
   air leakage around electrical outlets, switches, and lights, 153
   air leakage around windows, 152–153
   duct leakage, 153–155
   furnace, energy losses from, 150
   Venturi effect, 150
Albedometer, 235
Alkaline batteries, 529
American Plywood Association (APA), 123
American Society for Testing Materials (ASTM), 147
Ampere’s Law, 518
Anode, 532
Apollo II module system, 454
Appliances (residential), 79–84
Architectural considerations, 489–503
   glazing for passive solar, 493–496
   natural air flow, architecture for, 490–493
   passively solar heated residence, basic requirements for, 489–490
   plumbing for water heating, 501–503
   roof mounted solar elements, 496–501
      multi-layer plastic roofs, 499–500
      photovoltaic panels, requirements for, 496–497
      solar thermal collectors, requirements for, 501
      T-bar glazing systems, 497–499
ASHRAE Fundamentals
air changes per hour, 61
glazing losses, 91
heat loss through walls and roofs, 108
heat loss through windows, 103
thermal resistance of air spaces, 96
vision statement, 156
zone of floor temperatures, 162
ASTM, see American Society for Testing Materials
Atmosphere, transmission of light through, 209–239
air mass and dimming, 219–222
coherent scattering, 229
diathermanous material, 209
incoming solar radiation, 210–214
black body, 210
Planck’s law, 210, 211, 212
Stefan-Boltzmann law, 213
Wien’s displacement law, 212
instruments for measuring solar radiation, 231–235
albedometer, 235
cold end, 231
photosynthetic photon flux density, 235
pyranometers, 231–234
pyrheliometer, 234
pyrogeometers, 234–235
radiometer, 235
shadow ring pyranometer, 234
thermopile, 231
zenith reading, 235
isotropic and circumsolar radiation, 235–239
cloudiness, 239
concentrating photovoltaics, 238
concentrating solar power, 238
direct normal irradiance measurements, 238
Mie scattering, 229
Rayleigh scattering, 223–229
solar constant, 215
solar radiation in the atmosphere, 215–219
solar radiation measures, 230–231
beam radiation, 230
circumsolar radiation, 230
diffuse radiation, 230
global horizontal radiation, 230
global latitude inclined, 230
global-vertical south radiation, 230
normal radiation, 230

Autoclaved aerated concrete (AAC) block wall, 134–135

B

Balancing Authorities, 467
BAPV, see Building Applied Photo-Voltaic
Basements (residence)
  heat losses from, 113–116
  sealing and insulating, 145
Bathtub curve, 81
Batteries and battery technology, 526–531
  alkaline batteries, 529
  battery types, 528–529
  common batteries, 526–528
  D cell battery, 527
  dry cells, 528
  energy density in batteries, 530–531
  flow batteries, 529
  focus, 526
  lead acid battery, 527, 528
  Leclanché cells, 528
  lithium-ion batteries, 529
  nickel cadmium batteries, 529
  nickel iron batteries, 529
  nickel metal hydride batteries, 529
  zinc-carbon cells, 528
Beam radiation, 230
Beer’s law, 294
Bituminous coal, 2
Bouguer’s law, 294
Bread box solar collector, 430
Brewer’s angle, 293
Brick wall construction, 129–130
British Thermal Unit (BTU), 89
“Brown coal,” 2
Building Applied PhotoVoltaic (BAPV), 453

C

CAA, see Clean Air Act
CAES, see Compressed air energy storage
CAFE regulations, see Corporate Average Fuel Economy regulations
Canadian Lumber Standard (CLS), 89
Cane break, 124
Catawba Valley house, 369
Cathode, 532
CD, see Cooling degree days
Cement production, energy use in, 41
Chemical industry, energy use in, 38–39
Chernobyl, 36
CHP systems, see Combined heat and power systems
Chunnel, 32
Circumsolar radiation, 230
Cirrus clouds, 255
Clean Air Act (CAA), 42
Clean Water Act (CWA), 42
Climate, siting of solar structure and, 315–351
  clearances on flat sites, 342–348
    communities of solar homes, 346–348
    individual residences, 343–345
  clearances on sloping sites, 348–350
  climates zones by the number, 320–323
  design with nature, 339–342
  HDD and CDD, other data sources for, 332–334
  Köppen climate classification, 317–320
  magnetic declination, 336
  monthly averages for heating and cooling, 324–332
  southern exposure, 334–336
  visibility distances, 337–339
  wind protection, 351
CLS, see Canadian Lumber Standard
C.M. Kemp Manufacturing Company, 415
Coherent scattering, 229
Cold joints, 128
Combined heat and power (CHP) systems, 531
Commercial sector, energy use in, 45
Compressed air energy storage (CAES), 512–514
Concentrating photovoltaics (CPV), 238
Concentrating solar power (CSP), 238
Concrete block walls, 131
Construction techniques (residential), see Residential construction techniques
Cookers (solar), 411–415
Cooling degree days (CDD), 316
Corporate Average Fuel Economy (CAFE) regulations, 43
Coulomb’s Law, 519
Covalent bonds, 437
CPV, see Concentrating photovoltaics
Critical angle, 293
Crystal Palace Park, 286
CSP, see Concentrating solar power
CWA, see Clean Water Act

D

D cell battery, 527
Decipol, 72
Devonian Shale deposit, 11, 12
Diamagnetism, 515
Diathermanous material, 209
Diffuse radiation, 230
DIN 1946-6 standard (Germany), 62
Diode, 440
Direct normal irradiance (DNI) measurements, 238
Doldrums, 22
Domestic water usage, 75–79
Douglas Fir Plywood Association, 123

E

Early to modern solar structures, 353–378
double glazing, 354–355
efficient building envelopes, 375–376
historic use of solar heating, 355–358
Passivehaus design, 377–378
solar heating for residences in the rest of the world, 372–375
solar heating for residences in the United States, 358–372
Catawba Valley house, 369
competition, 368
Macneil House, 370
row houses, 358
Sloan house, 360
tax credits for renewable energy, 353
Thermopane, 354
townhouses, 376
Earthen materials, 133
Eastern Standard Time, 199
EIFS insulation system, see Exterior Insulation and Finish Systems insulation system
Energy Independence and Security Act of 2007, 472
Energy sources, energy uses, and global warming, 1–53
barrages, 29
bituminous coal, 2
“brown coal,” 2
composition of fossil fuels, 5–6
deuterium, 36
direct solar energy, 21
energy usage and income, 17–20
energy use in the United States, 37–46
cement production, 41
chemical industry, 38–39
commercial and residential sectors, 45
food production, 45–46
government regulation and the reaction from industry
industrial sector, 37–43
paper and pulp industry, 40–41
petroleum production, 37–38
steel and aluminum industries, recycling and energy savings in, 39–40
transportation in the modern world, 43–44
environmental costs of fossil fuels, 13–16
fossil fuels, uses and reserves, 7–13
Fukushima Daiichi disaster, 36
geothermal energy, 32–34
global warming, 46–53
effects of greenhouse gases, 49–53
greenhouse gases, 46–49
history of fossil fuels, 25
indirect solar energy, 21–28
biomass and biofuel potential, 28
doldrums, 22
horse latitudes, 22
hydroelectric power, 25–28
prevailing westerlies, 22
roaring forties, 22
tropics, 22
wave power, 23–24
winds of the world, 22–23
Marcellus and Devonian Shale areas, 11, 12
Mount St. Helens, volcanic eruption at, 32
nuclear fission, problems with, 36
nuclear power, 34–37
Old Faithful, 34
‘passive stability,’ 35
peat, 2
planet Earth, 20–21
retorting, 13
“ring of fire,” 32
thermonuclear bomb, structure of, 36
tidal power, 29–32
ENERGY STAR® appliances, 80–82
Energy storage, methods of, 505–534
batteries and battery technology, 526–531
alkaline batteries, 529
battery types, 528–529
common batteries, 526–528
D cell battery, 527
dry cells, 528
energy density in batteries, 530–531
flow batteries, 529
focus, 526
lead acid battery, 527, 528
Leclanché cells, 528
lithium-ion batteries, 529
nickel cadmium batteries, 529
nickel iron batteries, 529
nickel metal hydride batteries, 529
zinc-carbon cells, 528
electrical and electromechanical storage systems, 518–525
  brief history of electricity, 518–523
  supercapacitors, 524–525
  superconducting magnetic energy storage, 525
  triode tube, 521
 electrical generation capacity, 505–506
 energy density problem, 506–508
 fuel cells, 531–532
 hydrogen generation and storage, 532–534
  anode, 532
  cathode, 532
  integrated circuit, 523
 latent heat storage systems, 517
 mechanical energy storage techniques, 508–517
  compressed air energy storage, 512–514
  diamagnetism, 515
  flywheels, 516
  kinetic storage, 514–517
  magnetic levitation, 515
  permanent magnet, 516
  pumped hydroelectric storage, 508–512
  Velkess flywheel design, 516
 thermal storage systems, 517
 Enthalpy transfer, 70
 Envia Systems, 530
 EPA National Pollutant Discharge Elimination System, 42
 Equation of Time (EoT), 165, 207–208
 Eutectic salts, 407–408
 Evacuated tube collectors, 430–432
 Exterior Insulation and Finish Systems (EIFS) insulation system, 141
 Exterior walls, construction techniques for, 121–129

F

Faraday’s law, 521
Favelas, 121
FCDM, see Frequency control demand management
Federal Water Pollution Control Act, 42
Feed in tariffs (FiTs), 477
Field effect transistor (FET), 442–443
Flint glass, 293
Flow batteries, 529
Flywheels, 516
Food production, energy use in, 45–46
Fossil fuels
  composition of, 5–6
  environmental costs of, 13–16
  history of, 25
  uses and reserves, 7–13
Foundations (residence), heat losses from, 113–116
Frequency control demand management (FCDM), 467
Fresnel’s equations, 291–293
Fuel cells, 531–532
Fukushima Daiichi disaster, 36
Fulgurites, 284
Furnace, energy losses from, 150

G

GBPN, see Global Buildings Performance Network
Geothermal energy, 32–34
Glauber’s salt, 407
Glazing, transmission of solar energy through, 283–313
  brief history of glass, 284–286
  fulgurites, 284
  obsidian, 284
  petrified lightning, 284
daily solar gain through glazing, 306–308
extinction coefficient and absorptivity of glass, 293–296
  Beer’s law, 294
  Bouguer’s law, 294
Fresnel’s equations, 291–293
  critical angle, 293
  Flint glass, 293
  Snell’s law, 293
polarization of light, 290–291
transmission and absorption for normal radiation, 304–305
transmission, reflection, and absorption, 296–304
transmissivity of diathermanous materials to light, 286–290
transmissivity through LoE glass, 308–313
Heat Mirror, 309
  new generation LoE glass, 312–313
  standard LoE glass, 308–312
Glazing losses, fundamentals of, 91–93
Global Buildings Performance Network (GBPN), 158
Global horizontal radiation, 230
Global latitude inclined, 230
Global-vertical south radiation, 230
Global warming, 46–53, see also Energy sources, energy uses, and global warming
Gnomon, 200
Great Stuff spray, 127, 128
Greenhouse gases, 46–49
Greenwich Mean Time (GMT), 199

H

Heat flow (residence), see Residence, heat flow from
Heating degree days (HDD), 316
Heating monthly degree days (HMD), 316
Heat Mirror, 309
High-rise housing project, 121
Hindenberg Zeppelin, 533
Horse latitudes, 22
Hydrogen bomb, structure of, 36
Hydrogen generation and storage, 532–534

I

Indirect solar energy, 21–28
Industrial sector, energy use in (United States), 37–43
  cement production, 41
  chemical industry, 38–39
  government regulation and the reaction from industry
  paper and pulp industry, 40–41
  petroleum production, 37–38
  steel and aluminum industries, recycling and energy savings in, 39–40
Instruments for measuring solar radiation, 231–235
  albedometer, 235
  cold end, 231
  photosynthetic photon flux density, 235
pyranometers, 231–234
pyrneliometer, 234
pyrogeometers, 234–235
radiometer, 235
shadow ring pyranometer, 234
thermopile, 231
zenith reading, 235

Insulated concrete formwork, 134
Insulation types and methods (residential), 135–141
  cellulose insulation, 137
  EIFS insulation system, 141
  fiberglass batts, 136–137
  foam and foam board, 138
  multi-layer radiant barriers, 139–141
  radiant barriers, 139
  urea formaldehyde foam, 138

Integrated circuit, 523
Internal environment (residence), see Residence, internal environment of
International Space Union Conference, 165
Investor owned utilities (IOUs), 483

J

Julian calendar, 170–172

K

Kepler’s laws, 172
Kinetic storage, 514–517
Kipp-Zonen 121B shadow ring pyranometer, 234
Kirchhoff’s laws, 520
Kitchen faucets, flow rate of, 77
Köppen climate classification, 317–320

L

Lambert-Beer law, 257, 294
Latent heat storage systems, 517
Lawrence Berkeley Laboratory (LBL), 151
Lead acid battery, 527
Linke turbidity coefficient, 257
Lithium-ion batteries, 529
LoE (low emissivity) glass, 87, 308
Longitude noon, 165
Low voltage ride-through (LVRT) capability, 473
Luma Resources, 454

M
Magnetic levitation, 515
Manufacturing Energy Consumption Survey (MECS), 41
Marcellus Shale deposit, 11, 12
Masonry walls (residential), 129–135
    AAC block wall, 134–135
    brick wall construction, 129–130
    concrete block walls, 131
    early cements, 130
    earthen materials, 133
    insulated concrete formwork, 134
    polystyrene block, 134
    poured concrete foundations and walls, 130–131
    rammed earth, 133
    rock construction, 132–133
Maxwell’s equations, 520
Mie scattering, 229
Mill certified plywood, 123
Minority carriers, 440
Mount St. Helens, volcanic eruption at, 32

N
National Ambient Air Quality Standards (NAAQS), 42
National Highway Traffic Safety Administration (NHTSA), 43
National Pollutant Discharge Elimination System (NPDES) permit program, 42
National Renewable Energy Laboratory (NREL), 209, 263
National Solar Radiation Data Base (NSRDB), 263
Net metering, see Smart grids, FiTs, and net metering
Nickel cadmium batteries, 529
Nickel metal hydride batteries, 529
Noise control (residence), 84–85
Non-concentrating, active solar collectors, 417–433
   air collectors, 422–423
   bread box solar collector, 430
   evacuated tube collectors, 430–432
   heat pipes, 433
   liquid carrying flat plate solar collectors, 423–425
       closed loop and fully wetted collectors, 424–425
       trickle type collectors, 423–424
   selective absorptivity/emissivity surfaces, 426–430
       ceramit surfaces, 428–429
       early results, 427–428
       optically selective paints and sprays, 430
       selective coating films, 429–430
       statement of the problem, 426–427
   solar gain on sloping surfaces, 418–422
       stratified water tanks, 425–426
       universal solvent, 417
Normal radiation, 230
NPDES permit program, see National Pollutant Discharge Elimination System permit program
NREL, see National Renewable Energy Laboratory
NSRDB, see National Solar Radiation Data Base
Nuclear fission, problems with, 36
Nuclear power, 34–37

O

Obliquity vector, 165
Obsidian, 284
Oil shale, 12
Old Faithful, 34
“Old Ironsides,” 5
Oriented strand board (OSB), 124
Ovens (solar), 411–415

P

Paper and pulp industry, energy use in, 40–41
Passivehaus design, 377–378
Passive solar collection, 379–416
   absorption of glass, 396–399
shading, 393–396
solar blinds, 399–402
solar cookers and ovens, 411–415
solar gain through glass, 390–393
solar gain on vertical collectors with non-zero yaw angles, 380–390
solar stills, 408–411
solar water heaters, 415–416
thermal storage, 402–408
direct storage, 402
eutectic salts, 407–408
Glauber’s salt, 407
rock storage, 404
Trombe walls, 404–406
water storage, 406

Peat, 2
Permanent magnet, 516
Perovskite solar cells, 451–453
Petrified lightning, 284
Petroleum production, energy use in, 37–38
Photocatalytic decomposition, 451
Photosynthetic photon flux density, 235
Photovoltaic (PV) panels, 435–456
  building integrated photo-voltaics, 453–454
    Apollo II module system, 454
    POWERHOUSE Solar System 2.0, 453
concentrating photovoltaic arrays, 455–456
diode, 440
energy in photons, 444–446
field effect transistor, 442–443
first transistors, 440–442
minority carriers, 440
perovskite solar cells, 451–453
photocatalytic decomposition, 451
photovoltaic cells, 443–444
p-n junctions, 439–440
responses of first through third generation PV cells, 446–451
semiconductor fundamentals, 436–439
  covalent bonds, 437
  energy bands, 437
  n-type semiconductor, 438
valence band, 436
temperature sensitivity, 454–455

PHS, see Pumped hydroelectric storage

Planck’s law, 211, 212

Plywood, development of, 123–124

p-n junctions, 439–440

Polystyrene block, 134

Poured concrete foundations and walls, 130–131

POWERHOUSE™ Solar System 2.0, 453

Prevailing westerlies, 22

Prime meridian, 165, 170

Progression of the equinoxes, 185

Pumped hydroelectric storage (PHS), 508–512

PV panels, see Photovoltaic panels

Pyranometers, 231–234

R

Rammed earth, 133

Rayleigh scattering, 223–229

Red Leaf Resources, 11

Residence, heat flow from, 87–116
  empirical nature of heat loss calculations, 90–91
  glazing losses, fundamentals of, 91–93
  heat flow across an air gap, 93–98
  heat losses from a residence, 107–116
    basements and foundations, 113–116
    structural integrated panels, 112–113
    walls and roofs, 107–112
  thermal resistance and conductance, 99–102
  thermal terms, 87–90
  windows, 102–106
    heat loss through windows, 103–105
    R-factors, 105
    spacers and their influence on heat loss, 105–106
    warm edge spacers, 105

Residence, internal environment of, 55–86
  decipol, 72
  domestic water usage, 75–79
    kitchen faucets, flow rate of, 77
septic systems, 77
electrical fixtures and appliances, 79–84
  appliance failure rate, 81
  bathtub curve, 81
  burn-in period, 81
  electrically generated lighting, 82–84
  ENERGY STAR® appliances, 80–82
electrical use and its contribution to sensible heat, 56–57
enthalpy transfer, 70
heat recovery ventilation, 68–70
human needs and demands, 58
humidity sources in a residence, 70–71
illumination and daylighting, 73–75
noise control, 84–85
return on investment and service life, 85–86
  major domestic appliances, 86
  State Renewable energy Credits, 86
  unavoidable costs, 85
sick building syndrome and VOC mitigation, 71–72
temperature levels, 59–60
Tubular Daylighting Device, 74
ventilation guidelines, 60–67
  air changes per hour, 61–62
  air flow through exterior doors, 65–66
  air infiltration, 64–65
  air velocity, 62
  DIN 1946-6 standard (Germany), 62
  relative humidity, 62–64
  respiration and carbon dioxide, 67
  revolving storm door, 65
Residential construction techniques, 117–163
  air infiltration, 148–155
    ACH50, 151
    air leakage from conditioned spaces, 150–152
    air leakage around electrical outlets, switches, and lights, 153
    air leakage around windows, 152–153
    duct leakage, 153–155
    furnace, energy losses from, 150
    Venturi effect, 150
  energy codes and their effect on construction methods, 155–159
errors in energy codes, 159–161
exterior walls, construction techniques for, 121–129
  balloon framing, 127
  bamboo and other grasses, 124–125
  box frame, 126, 127
  cane break, 124
  cold joints, 128
  mill certified plywood, 123
  plywood, development of, 123–124
  post and beam framing, 127
  structural integrated panels, 128–129
  stud wall construction, 125–128
  wooden walls, 121–123
fire, civilization, and simple dwellings, 119–121
  favelas, 121
  high-rise housing project, 121
  population increase, 120
  shanty towns, 121
  yurts, 121
“high tide” mark, 145
insulation guidelines and their
  influence on construction, 162–163
insulation types and methods, 135–141
  cellulose insulation, 137
  EIFS insulation system, 141
  fiberglass batts, 136–137
  foam and foam board, 138
  multi-layer radiant barriers, 139–141
  radiant barriers, 139
  urea formaldehyde foam, 138
masonry walls, 129–135
  AAC block wall, 134–135
  brick wall construction, 129–130
  concrete block walls, 131
  early cements, 130
  earthen materials, 133
  insulated concrete formwork, 134
  polystyrene block, 134
  poured concrete foundations and walls, 130–131
  rammed earth, 133
rock construction, 132–133
roof construction and problems, 141–143
roof construction techniques, 141–142
roof problems, 144–145
unwanted heat and humidity flow through roofs, 142–143
wildfire that reaches a house, 144
sealing and insulating basements, 145
water damage and vapor migration, 145–148
Residential sector, energy use in, 45
Retorting, 13
Rovolving storm door, 65
“Ring of fire,” 32
Roaring forties, 22
Rock construction, 132–133
problems, 141–143
techniques (residential), 141–142
Roof(s)
construction and problems, 141–143
heat losses through, 107–112
mounted solar elements, 496–501
Row houses, 358
Royal Observatory (Greenwich), 165

S

SBS, see Sick Building Syndrome
Seasons and solar angles, 165–208
analemma and how to capture it, 202–207
angle of declination, 166
angle of the Sun on the Earth’s surface, 188–191
angle of the Sun on solar collectors, 192–193
change in solar angle with time, 197–198
ecliptic plane, rotation of the Earth around the Sun in, 165
equation of time approximation, 207–208
equation of time and sundials, 199–202
   Eastern Standard Time, 199
gnomon, 200
   Greenwich Mean Time, 199
   sundial, 200
geometry of the ellipse and Kepler, 172–175
importance of time, 167–169
Julian calendar, 170–172
longitude, 169–170
longitude noon, 165
obliquity and the seasons, 166–168
obliquity vector, 165, 183–187
points, vectors, and planes, 175–180
prime meridian, 165, 170
progression of the equinoxes, 185
rotation of vectors, 180–183
seasonal angles, 199
solar collection day, 193–197
Semiconductor fundamentals, 436–439
covalent bonds, 437
energy bands, 437
n-type semiconductor, 438
valence band, 436
Septic systems, 77
SERI, see Solar Energy Research Institute
Shanty towns, 121
SHW, see Solar hot water system
Sick Building Syndrome (SBS), 71–72
Sloan house, 360
Smart grids, FiTs, and net metering, 457–488
FiTs and net metering, 477–487
future of solar electrical generation, 473–477
history of electrical utilities, 458–461
inverters and their electrical connections, 465–467
net metering, 466
power factor problem, 461–463
regulated future, 487–488
smart grids, 467–473
Balancing Authorities, 467
frequency control demand management, 467
low voltage ride-through capability, 473
ride-through, 473
Smart Grid Information Clearinghouse, 472
VAR control, 469
soft costs, 474
standard connections from grid to residence, 463–464
Snell’s law, 287, 293
Solar angles, see Seasons and solar angles
Solar Energy Research Institute (SERI), 209
Solar gain and solar attenuation, 241–282
   air temperature, pressure, and clouds, 252–256
      cirrus clouds, 255
      troposphere, 253
albedo and specularity, 247–252
clear sky radiation data for U.S. cities, 263–279
   Albuquerque, 273
   Austin, 277
   Baltimore, 271
   Bangor, 271
   Billings, 270
   Boston, 266
   Chicago, 272
   Columbus, 274
   Concord, NH, 272
   Denver, 267
   Detroit, 270
   Hartford, 275
   Kennedy, 268
   Lexington, 273
   Lincoln, 268
   Orlando, 267
   Phoenix, 276
   Raleigh, NC, 274
   Reno, 275
   Sacramento, 277
   Salt Lake City, 269
   Seattle, 269
   Tulsa, 276
   Wichita, 278
models for radiation dispersion in the atmosphere, 256–258
   Lambert-Beer law, 257
   Linke turbidity coefficient, 257
solar fraction and solar gain, 242–247
solar radiation data for Boston, 258–263
solar radiation maps of the world, 279–282
Solar hot water system (SHW), 415
Solar radiation measures, 230–231
   beam radiation, 230
   circumsolar radiation, 230
diffuse radiation, 230
   global horizontal radiation, 230
global latitude inclined, 230
global-vertical south radiation, 230
   normal radiation, 230
Solatube Daylighting Systems, 74
State Renewable energy Credits (SRECs), 86
Steel and aluminum industries, recycling and energy savings in, 39–40
Stefan-Boltzmann law, 213
Structural integrated panels, 128–129
Stud wall construction (residential), 125–128
Sundial, 200
Sun King, 285
Supercapacitors, 524–525

T

Tax credits for renewable energy, 353
T-bar glazing systems, 497–499
TDD, see Tubular Daylighting Device
Terraced houses, 358
Tesla Motors, 530
Thermionic valve, 521
Thermonuclear bomb, structure of, 36
Thermopane, 354
Tidal power, 29–32
townhouses, 376
Transportation in the modern world, 43–44
Triode tube, 521
Trombe walls, 404–406
Troposphere, 253
Tubular Daylighting Device (TDD), 74

U

Universal solvent, 417
U.S. cities, clear sky radiation data for, 263–279
Albuquerque, 273
Austin, 277
Baltimore, 271
Bangor, 271
Billings, 270
Boston, 266
Chicago, 272
Columbus, 274
Concord, NH, 272
Denver, 267
Detroit, 270
Hartford, 275
Kennedy, 268
Lexington, 273
Lincoln, 268
Orlando, 267
Phoenix, 276
Raleigh, NC, 274
Reno, 275
Sacramento, 277
Salt Lake City, 269
Seattle, 269
Tulsa, 276
Wichita, 278

U-Seal, 145
U.S. Geological Survey (USGS), 39
U.S. residences, solar heating for, 358–372

V

Vacuum tube, 521
Valence band, 436
Vapor barrier, ASTM definition of, 147
Velkess flywheel design, 516
Ventilation guidelines (residence), 60–67
Venturi effect, 150
Volt amperes reactive (VAR) control, 469

W
Walls (residence), heat losses through, 107–112
Warm edge spacers (WES), 105
Water heaters (solar), 415–416
Water heating, plumbing for, 501–503
Wien’s displacement law, 212
Wildfires, 144
Windows (residence), heat flow from, 102–106
World Meteorological Organization (WMO), 235
World residences, solar heating for, 372–375

Y

Yurts, 121

Z

Zehnder ComfoAir, 70
Zenith reading, 235
Zinc-carbon cells, 528
Zone of floor temperatures (ASHRAE), 162