Basic Electronics and Electric Power
Denard Lynch, P. Eng.
© 2014
# Table of Contents

1 Review of Direct Current and Voltage ................................................................. 1  
  1.1 Useful Circuit Comparisons: ................................................................. 1  
  1.2 Review of Resistor Behaviour in a Direct Current (DC) Circuit ................. 1  
  1.3 Review of Capacitor Behaviour in a Direct Current (DC) Circuit .......... 2  

2 Inductors ................................................................................................................. 4  
  2.1 Inductor Behaviour in a Direct Current (DC) Circuit ......................... 7  
  2.2 Calculating the inductance, L, of a coil from physical properties .......... 8  
  2.3 Inductors in Series and in Parallel ......................................................... 10  
  2.4 Voltage Across an Inductor .................................................................. 11  
  2.5 Inductive Transient Behaviour ............................................................. 12  
  2.6 Energy Stored in an Inductor or Capacitor ....................................... 17  
  2.7 Generalization of exponential transient behaviour ......................... 17  

3 Capacitors ................................................................................................................. 20  
  3.1 Capacitors in a Steady-state DC Circuit .............................................. 20  
  3.2 Adding Capacitors in Series or Parallel ............................................ 21  
  3.3 Capacitive Transient Behaviour ............................................................ 22  
  3.4 Energy stored in a Capacitor: ................................................................. 24  
  3.5 Inductive and Capacitive Transients – SUMMARY .......................... 26  
  3.6 Determining Initial and Final Values in R-C and R-L Circuits: ........ 28  

4 Alternating Current (AC) and Power .................................................................... 29  
  4.1 Average and Effective (RMS) Values ...................................................... 36  
    4.1.1 Average Value: ............................................................................. 36  
    4.1.2 Effective Value: ........................................................................... 36  
  4.2 Adding and Subtracting AC Voltages and Currents ......................... 39  
  4.3 Opposition to Flow in AC Circuits ......................................................... 41  
    4.3.1 Impedance, Z................................................................................. 43  
  4.4 Power in AC Circuits ............................................................................. 49  
    4.4.1 Resistor ....................................................................................... 51  
    4.4.2 Inductor ....................................................................................... 52  
    4.4.3 Capacitor ..................................................................................... 53  
    4.4.4 AC Maximum Power Transfer ..................................................... 56  
  4.5 Power Factor Correction ........................................................................ 56  
  4.6 Transformers ............................................................................................... 60  
    4.6.1 Impedance Matching ..................................................................... 62  
    4.6.2 Schematic Representation: ............................................................ 64  
    4.6.3 Two additional considerations ..................................................... 65  
  4.7 Three Phase Power (3Φ Power) ............................................................... 66  

5 Electronic Devices and Integrated Circuits ......................................................... 68  
  5.1 Single, Electronic Devices: Diodes and Transistors ............................. 68  
    5.1.1 Diodes: ....................................................................................... 68  
    5.1.2 Transistors: ................................................................................ 76  
    5.1.3 Field Effect Transistors: ............................................................. 84  
  5.2 Operational Amplifiers ............................................................................. 90  
    5.2.1 Configurations: ........................................................................... 92  
    5.2.2 Frequency Response: ................................................................. 94  

# Table of Examples

| Example 2-1: | Inductance of an Air-core Coil | 8 |
| Example 2-2: | Inductance of an Iron-core Coil | 8 |
| Example 2-3: | R-L Transient using R_L | 18 |
| Example 3-1: | R-C Transient Response | 25 |
| Example 3-2: | R-C Example 1 | 25 |
| Example 5-1: | Simple Op Amp Circuits | 96 |
| Example 5-2: | Tandem Inverting Op Amps | 98 |
| Example 6-1: | An Unregulated Full-wave Supply | 108 |
| Example 6-2: | A Linear Regulated Supply | 114 |
| Example 6-3: | A SMPS Supply | 121 |
| Example 8-1: | A Motor/Generator Bar & Rail Example (2006 midterm) | 134 |
| Example 8-2: | Motor Example using the Simple Model | 137 |
| Example 8-3: | An Example using the Simple Generator Model | 138 |
| Example 9-1: | Battery Sizing Example 1 | 148 |
| Example 10-1: | PV Example 1: based on an example by (Solar Energy International) | 152 |
| Example 10-2: | PV Example 2 | 153 |
| Example 11-1: | Wind Power Example 1 | 159 |
# Table of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1: Ohm's Law Analogies</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Figure 1.2: Resistors in Parallel</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Figure 1.3: Resistors in Series</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Figure 1.4: R-C in Parallel</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Figure 1.5: R-C in Series</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Figure 2.1: Induced Voltage in a Coil</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Figure 2.2: Example of an Air-core Coil</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Figure 2.3: R-L in Parallel</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Figure 2.4: R-L in Series</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Figure 2.5: Induced Voltage in Inductor due to Changing Current</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Figure 2.6: R-L Charging Transient</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Figure 2.7: R-L Voltage and Current during Charging Transient</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Figure 2.8: R-L Discharge Transient</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Figure 3.1: R-C Charging Transient</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Figure 3.2: R-C Voltage Transients</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Figure 3.3: R-C Current Transients</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Figure 4.1: An AC Sinusoidal Waveform</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Figure 4.2: Example of Lead - Lag Phase Difference</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Figure 4.3: A simple AC Circuit Representation</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Figure 4.4: Current Direction in an AC circuit</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Figure 4.5: AC Example 1</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Figure 4.6: V, I Waveforms</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Figure 4.7: AC Example 2</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Figure 4.8: Phasor Diagram</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Figure 4.9: Impedance Diagram</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>Figure 4.10: Instantaneous Power in Resistive Load</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Figure 4.11: Instantaneous Power in an Inductive Load</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Figure 4.12: Instantaneous Power in a Capacitive Load</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>Figure 4.13: Complex Power Components</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>Figure 4.14: A Simple Transformer Fig. 22.10 Invalid source specified</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>Figure 5.1: Diode Schematic Symbol</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>Figure 5.2: 1N4007 DiodeInvalid source specified</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Figure 5.3: Ideal Diode I-V Curve</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Figure 5.4: Diode Model - Exponential</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Figure 5.5: Diode Model - Linear-Piecewise</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Figure 5.6: Diode Model - Constant Voltage</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Figure 5.7: NPN and PNP BJT Structure Invalid source specified</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>Figure 5.8: BJT Schematic Symbols Invalid source specified</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>Figure 5.9: Typical Transistor Packages Invalid source specified</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>Figure 5.10: MOSFET Construction Invalid source specified</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>Figure 5.11: Transistor Symbols: a) bipolar, b) FET, c) MOSFET, d) dual-gate MOSFET, e) Inductive Channel MOSFET Invalid source specified</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>Figure 5.12: FET Example 1</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>Figure 5.13: FET Example 2</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>Figure 5.14: Ideal Op Amp Model</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Figure 5.15: Inverting Op Amp</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>Figure 5.16: Non-inverting Op Amp</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>Figure 5.17: Voltage Follower Op Amp</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>Figure 5.18: Gain versus Bandwidth for TL082 Op Amp (Texas Instruments)</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>Figure 5.19: Op Amp Simple Circuits Eg. a)</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>Figure 5.20: Op Amp Simple Circuits Eg. b)</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>Figure 5.21: Op Amp Simple Circuits Eg. c)</td>
<td>97</td>
<td></td>
</tr>
</tbody>
</table>
1 Review of Direct Current and Voltage
Consider the following analogies comparing fluid, electric and magnetic circuits.

1.1 Useful Circuit Comparisons:

![Diagram of fluid, electricity, and magnetism analogies]

The parallel between magnetic circuits and (particularly) electric circuits is very useful for analyzing and designing magnetic circuits!!
In all these cases, there is a simple relationship between three parameters in the circuit as summarized in the following table:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fluids</th>
<th>Electricity</th>
<th>Magnetics</th>
</tr>
</thead>
<tbody>
<tr>
<td>“force” that causes something to flow</td>
<td>pressure, $P$</td>
<td>electromotive force, $E^*$, measured in Volts, $V$</td>
<td>magnetomotive force (mmf), $\Phi$, measured in Ampere-turns, $A$</td>
</tr>
<tr>
<td>that which “flows”</td>
<td>liquid</td>
<td>current, $I$, measured in amperes, $A$</td>
<td>Magnetic flux, $\Phi$, measured in Webers, $W$</td>
</tr>
<tr>
<td>that which impedes flow</td>
<td>restriction like a valve, kink or hydraulic motor</td>
<td>Resistance, $R$, measured in Ohms, $\Omega$</td>
<td>Reluctance, $\Re$, measured in amp-turns/Weber</td>
</tr>
</tbody>
</table>

*In an electric (schematic) diagram, sources are labelled “$E$”, whereas voltage drops across sinks are labelled “$V$” when required.*

1.2 Review of Resistor Behaviour in a Direct Current (DC) Circuit
Consider some simple examples of the behaviour of these parameters:
1) Given a common voltage (pressure) across different values of resistance, $I(\text{flow}) \propto \frac{1}{R}$.
   In the circuit shown in Figure 1.2, $\frac{3}{4}$ of the current (6A) would flow through $R_1$, and $\frac{1}{4}$
(2A) would flow through $R_2$. Note that the 6V source would supply the total 8A into that node, verifying Kirchhoff’s Current Law (KCL).

![Figure 1.2: Resistors in Parallel](image)

Of course, we would measure exactly the same voltage (6V) across each resistor.

2) If the resistors were in series, as in Figure 1.4, the current would be common to both, and the voltage would be divided across each in proportion to their value ($V_r \propto R$). In this case, we would measure $\frac{1}{4}$ of the source voltage (1.5V) across $R_1$, and $\frac{3}{4}$ (4.5V) across $R_2$, which corresponds to Kirchhoff’s Voltage Law (KVL).

![Figure 1.3: Resistors in Series](image)

This behaviour in a Direct Current (DC) circuit is described by Ohm’s Law: $E = IR$.

A note about voltage polarity and current direction: in a passive component like a resistor (a sink for the energy), the “pressure” (voltage) is higher on the input end, and lower” on the output. In both cases just considered, the polarity of the voltage drop across the resistors would be positive on top in the figures. For a source of energy like a battery however, the current flows out of the positive terminal and back into the negative terminal once it has been routed through the circuit.

### 1.3 Review of Capacitor Behaviour in a Direct Current (DC) Circuit

Capacitors are one of the three basic electrical elements: resistors, capacitors, and inductors (discussed later). The voltage measured across a capacitor is proportional to the charge on its “plates”, which are separated by a dielectric material. The energy stored in a capacitor is in the form of an electric field.

We can make the following observations about capacitors in DC circuits:
- The Capacitance, C is a function of the physical characteristics of the capacitor:
  \[ C = \varepsilon_r \varepsilon_0 \frac{A}{d} \]
  where \( A \) = the plate area (m\(^2\)), \( d \) = the distance between the plates (m), \( \varepsilon_0 \) = permittivity of free space \((8.854 \times 10^{-12} \text{ C}^2/\text{N} \cdot \text{m}^2)\), and \( \varepsilon_r \) is the relative permittivity of the dielectric between the two capacitor plates (a unit-less number giving the ratio of the permittivity of the dielectric to that of a vacuum, with a vacuum considered essentially the same as free space).

- The voltage across a capacitor is proportional to the charge, \( q \) \( (V_c \propto q) \)
- \( q \) = (current, \( I \))(time, \( t \)) (e.g. 1A flowing for 1 second = 1 Coulomb of charge)
- 1 Coulomb of charge on a 1 Farad Capacitor will result in 1 Volt across the capacitor.
- At “steady state” in a DC circuit (i.e. once the charge on the plates is built up and no more is being added), a capacitor “looks like” an open circuit. (N.B.: while it is “charging up”, this is not the case. Transient behaviour will be discussed later.)

Consider the following examples:

1) A capacitor in parallel with a DC source and a resistor as shown in Figure 1.4. In this case (once a “steady-state” is reached), there will be 6A flowing through the resistor \((6\text{V}/1\Omega)\) but no current flowing into or through the capacitor. We will, of course, measure the same 6V across either element.

2) The same capacitor in series with the resistor as shown in Figure 1.5. In this case, there will be no current flowing in the circuit, as the capacitor (once charge up, acts like an open circuit. Since no current is flowing through the resistor, there is no voltage drop across it (\(\Omega\)’s Law), therefore we must see a 6V drop across the capacitor to satisfy Kirchhoff’s Voltage Law (positive polarity on top).
2 Inductors

The 3rd basic electrical element is the inductor. It operates on electro-magnetic principles. Recall that we can use electricity to create a magnetic field (e.g. an electromagnet), and we can also use a magnetic field to produce electricity (e.g. a “generator”).

Where capacitors build up and store energy in an electric field, an inductor builds up and stores energy in a magnetic field. They are physically constructed of a coil of wire wound around a core. The core is often made of a ferromagnetic material, but can be just air or some other dielectric material.

The scientific principle behind inductor operation is Faraday’s Law of Magnetic induction, which means that a current flowing in a wire produces a magnetic field around the wire, and a changing (or moving) magnetic field will exert a force on a charge making it want to move. If this “induced current” (moving charge) encounters any restriction, it builds up a “pressure” (voltage) to push the current through, and in this way the coil acts like a source.

In inductors, and other electromagnetic devices, both the creation of a magnetic field from a flowing current and the induction of a voltage from a changing magnetic field occur at the same time. Consider this simple sequence of causes and effects:

- A current, I, flowing in a coil produces a magnetic flux, \( \Phi \)
- A changing current leads to a changing flux: \( \Delta I \Rightarrow \Delta \Phi \)
- A changing flux will induce a voltage: \( \Delta \Phi \Rightarrow E \)
- A voltage will cause a current to flow: \( E/R \Rightarrow I \)

Therefore, as you cause a current to flow in a coil (inductor), the changing magnetic flux caused by the increasing current induces a magnetic field (\( \Delta \Phi \)) that induces a voltage that will put pressure on the charges in the wire. A companion law, Lenz’s Law, tells us that the polarity (or direction) of the induced voltage is such that it opposes the current flow that is trying to create it.

Considering the voltage induced in a coil, Faraday’s Law tells us: 

\[
E_{Mag} = -N \frac{d\Phi}{dt}
\]

for the voltage induce in a loop of wire with \( N \) turns. (Note: the negative sign may be interpreted as indicating that the induced voltage opposes the flow of current which is creating it, as predicted by Lenz’s Law.)

Substituting the relationships between current and magnetic flux, and the physical properties of the coil and its core, we have an expression of the voltage observed across an inductor, \( E_{Mag} \), in terms of the physical properties of the coil and the current passing through it:

\[
E_{Mag} = \frac{\mu_0 \mu_s N^2 A}{l} \frac{dI}{dt}
\]

Equation 2.1

where:
- \( \mu_0 \) is the permeability of free space (4\( \pi \) x 10\(^{-7} \) Wb/At-m)
- \( \mu_r \) is the relative permeability of the core material (if not air)
- \( N \) is the number of turns of wire in the coil
- \( A \) is the area of the coil cross-section in \( m^2 \)
- \( l \) is the length of the coil in m
- \( \frac{dI}{dt} \) is the rate of change of current with time

Define the constant properties of the coil as its *inductance*, \( L = \frac{\mu_r \mu_0 N^2 A}{l} \), we can re-write Equation 2.1 as:

\[
E_{Mag} = L \frac{dI}{dt} E \quad \text{Equation 2.2}
\]

Where the letter ‘L’ is used to represent “self inductance”, usually referred to as just “inductance”, and the unit of measurement is the **Henry** (H). (E.g. A rate of change of current of 1 ampere per second in a 1 Henry inductor will produce a voltage of 1 Volt across the inductor.)

Figure 2.1: Induced Voltage in a Coil

Figure 2.1 depicts how a coil moving through a magnetic field (grey box) would gradually envelop more of the magnetic flux until it is entirely filled, and then reduce as the coil passes out of the field to the right. As the flux *changes* (bottom graph), a voltage is induced. Note that when the coil is “full” and the flux is not *changing*, there is no
voltage induced. Also note that the induced voltage polarity changes depending on whether the flux is *increasing* or *decreasing*.

![Diagram of an Air-core Coil](image)

**Figure 2.2: Example of an Air-core Coil**

For an air-core coil, the reluctance “outside” the coil is considered small compared to the reluctance “inside” the coil (R_i >> R_o). Thus, for an air-core inductor where the length is >> than the diameter (i.e. l/d > 10), the inductance can be estimated using:

\[
L = \frac{\mu_0 N^2 A}{l}
\]

Equation 2.3

This formula assumes an ideal, infinitely long coil. For coils with a diameter-to-length ration of 10 or more, this is accurate within approximately 4%. (Note: for l/d < 10, you can apply Nagaoka’s correction factor to improve accuracy.)

In summary:
- a steady-state (non-changing) current will produce a magnetic field, but since it is *not changing*, it will *not* induce any voltage. This, in turn, means...
- in a static, DC circuit, an ideal inductor (V_L=0) looks like a *short circuit* (a practical inductor has *some* resistance due to the wire in its coil)
2.1 Inductor Behaviour in a Direct Current (DC) Circuit

Consider the following examples of inductor in DC circuits:

1) Inductor in parallel with a source and resistor

![Figure 2.3: R-L in Parallel](image)

In this case, the resistor would still handle the 6A of current (6V/1Ω) as expected. An ideal inductor would be problematic, as it would represent a “short circuit” directly across the source leading to, ideally, an infinite current! Even a practical inductor can have a very low resistance, and should never be connected directly across a source to avoid damaging equipment and creating an unsafe situation with fire or excess heat.

2) Inductor in series with a resistor in a DC circuit

![Figure 2.4: R-L in Series](image)

At steady-state, \( \frac{dI}{dt} \) (by definition) is zero (0V), therefore there must be 6V across the resistor to satisfy KVL. The resulting 6A flows through both the resistor and inductor.
2.2 Calculating the inductance, $L$, of a coil from physical properties.

The inductance, an indication of how strongly a coil opposes changes in its current, can be calculated from the physical characteristics of the inductor:

$$ L = \frac{\mu N^2 A}{l} \quad \text{Equation 2.4} $$

where $N$ is the number of turns of wire in the coil, $A$ is the area of the coil in meters$^2$, and $\mu$ is the permeability of the material in the core of the coil. If the core is air or some other dielectric, $\mu = \mu_0$ is the permeability of free space: $4\pi \times 10^{-7}$ Wb/At-m. If the core is filled with a ferromagnetic material, the permeability, $\mu = \mu_0 \mu_r$, $\mu_r$ is the relative permeability of the material (ratio of its permeability to $\mu_0$), and can be in the thousands in some cases.

Example 2-1: Inductance of an Air-core Coil

An air-core coil… find the inductance, $L$.

Check: $l/d = 12.5$ – OK.

$$ L = \frac{\mu_0 AN^2}{l} $$

$$ L = \frac{(\mu_r)(.006^2)(\pi)(120^2)}{(.15)} $$

$$ L = 13.64 \mu H $$

Example 2-2: Inductance of an Iron-core Coil

A 100 turn coil wound on a 1cm X 1cm laminated sheet steel core (S.F. = .93) with dimensions as shown. Calculate the inductance, $L$. The absolute permeability ($\mu_0 \mu_r$) is .0025 Wb/At-m.

$$ L = \frac{\mu N^2 A}{l} \quad \text{Equation 2.4} $$

$$ L = \frac{(0.0025)(100)^2 (.01m)^2 (.93)}{.1m} = 23.25mH $$
Notes:
1) With a ferromagnetic core, the leakage that would occur with an air core, lowering the inductance from that predicted by the “ideal” formula is not really a problem, but if this was an air core coil, we should check the ratio of the diameter to the length (as in Example 1)

2) The permeability of most ferromagnetic materials (cores) is not linear. The material saturates once a certain level of magnetic flux is created, and any incremental current over the amount require to “fill” the core will produce progressively less incremental flux. When completely saturated, the permeability of the ferromagnetic material approaches that of air (non-ferromagnetic). Most practical inductors are designed to work in an approximately linear region of flux creation to ensure the inductance is predictable.

3) Area, A: this core is laminated, which means it’s constructed of strip of metal separated by a thin film of insulating material. The effective area is less than the nominal dimensions would suggest. The Stacking Factor (S.F.) gives the effective area as a fraction of the nominal area. In this case, we are given 0.93 or 93% of 1cm X 1cm.

4) Use base units to calculate the inductance, L, in Henries
2.3 Inductors in Series and in Parallel

Adding inductors in series or parallel in a circuit is the same as for resistors:

For inductors in series: \( L_T = L_1 + L_2 + L_3 + \ldots \)

![Series Inductors Diagram]

For inductors in Parallel: \( \frac{1}{L_T} = \frac{1}{L_1} + \frac{1}{L_2} + \frac{1}{L_3} + \ldots \)

![Parallel Inductors Diagram]

Example:
Find the total equivalent inductance of the circuit shown below:

![Circuit Diagram]

(ANS: 3.83H)

1) add 4H and 8H in parallel: \( L' = \left( \frac{1}{4H} + \frac{1}{8H} \right)^{-1} = 2.667H \)

2) add this in series with the 3 and 1H inductors: \( L^* = 3 + 2.667 + 1H = 6.667H \)

3) add this again in parallel to the 9H: \( L_T = \left( \frac{1}{9H} + \frac{1}{6.667H} \right)^{-1} = 3.83H \)
2.4 Voltage Across an Inductor

Let’s go back again to the expression for the voltage induced across an inductor with a changing magnetic field caused by a changing current flowing through it:

\[ E_{Mag} = L \frac{dI}{dt} \quad \text{Equation 2.2} \]

(For an inductor in a circuit, \( E_M \), the magnetically-induced voltage observed across the inductor, is often represented as \( V_L \) (or \( v_L \)). The ‘E’ designation (a source voltage) is actually correct in this instance as the “induced voltage” is acting like a source, but convention usually labels voltage across a passive component in a circuit with ‘V’.)

If we have a changing current flowing through the inductor (refer to Figure 2.5), we will observe a “self-induced” voltage across it (its attempt to stop or at least resist the change in current).

Figure 2.5: Induced Voltage in Inductor due to Changing Current

from 0s – 1s:
\[ L \frac{di}{dt} = (1H) \frac{1A}{1s} = 1V \]

from 1s – 2s
\[ L \frac{di}{dt} = (1H) \frac{-1A}{1s} = -1V \]

(Note: We could work backwards from the voltage to find out what the current was by integrating!)
2.5 **Inductive Transient Behaviour**

So far we’ve considered *static* conditions. Now let’s look at *transient* conditions (i.e. what happens from the time a change is made in the circuit [parameters] until steady-state is again reached.)

Let’s start by examining our inductor in a simple R – L circuit as sown in Figure 2.6: R-L Charging Transient:

Before we close the switch, it is obvious there is no current in the circuit and no voltages to observe across the components.

Let’s assume we close the switch at an arbitrary time, \( t = 0 \), to apply a voltage, \( E \), to our R – L circuit. Using static (DC) analysis, we can determine that since the inductor will oppose any change in current (i.e. want to keep it at zero) \( I_L \) will still be zero at \( t = 0^+ \), just after we close the switch. Once steady-state is reached, the current through the inductor, and resistor, will be \( E/R \) (recall the inductor looks like a “short”). (In practice, the wire used in a “practical inductor” has some resistance, so \( V_L \) won’t quite be 0. It can be modelled as:

but we’ll assume an ideal inductor for this analysis (\( R_{\text{internal}} = 0 \Omega \))

Once the switch is closed and we are in our “time period of interest”, we can use KVL to write an equation using the voltages across the components as a function of time:

\[
E - V_R - V_L = 0
\]

substituting current-based expressions for voltage:

\[
E - IR - L \frac{di}{dt} = 0
\]

Re-arranging and solving for \( I \) as a function of \( t \):

\[
i_L(t) = \frac{E}{R} \left(1 - e^{-Rt/L}\right)
\]
Which represents the current through the inductor (or the resistor, since they are in series). Note that I’ve used a lower-case ‘i’ for current to indicate that it is changing with time.

The ratio of \( \frac{L}{R} \) is defined at the **time constant**, \( \tau \), of the circuit, and is measured in **seconds**. Making this substitution, the expression becomes:

\[
i_L(t) = \frac{E}{R} \left(1 - e^{-t/\tau}\right) \hspace{1cm} \text{Equation 2.5}
\]

The time, \( t \), in this case starts from just after the switch closes until steady-state is again reached.

Recall that because of the electromagnetic nature of an inductor, the current cannot change instantly, so if it is 0A just prior to the switch closing (\( t = 0^+ \)) it is 0A just after (\( t = 0^- \)). Our expression can be verified by testing the values at \( t = 0^- \) and \( t = \infty \).

\[
i_L(0) = \frac{E}{R} \left(1 - e^{-R(0)/L}\right) = 0, \quad \text{and} \quad i_L(\infty) = \frac{E}{R} \left(1 - e^{-R(\infty)/L}\right) = \frac{E}{R}
\]

We know from our DC analysis that \( \frac{E}{R} \) is just the final steady-state current, \( I_f \), so we could write the expression using that term, which does not restrict our expression to the simplified circuit assumptions we originally used, but does assume no pre-existing conditions:

\[
i_L(t) = I_f \left(1 - e^{-t/\tau}\right) \hspace{1cm} \text{Equation 2.6}
\]

To derive an expression for the **voltage** across the inductor as a function of time during this transient period, we simply use our original expression, \( v_L = L \frac{di}{dt} \), substitute the expression (2.5) we just derived for \( i_L(t) \):

\[
v_L(t) = L \frac{d\left(\frac{E}{R} \left(1 - e^{-t/\tau}\right)\right)}{dt} = \frac{E}{R} \frac{d\left(1 - e^{-t/\tau}\right)}{dt} = \frac{E}{R} \left[ -\frac{1}{\tau} \right] e^{-t/\tau}
\]

and solving for \( v_L \) yields:

\[
v_L(t) = E e^{-t/\tau} \hspace{1cm} (3)
\]

where \( \tau \) is again the time constant, \( \frac{L}{R} \).
We observe again that at the very beginning of this period of interest when \( t = 0^+ \), the initial voltage, \( V_i = E \), and as we would expect at \( t = \infty \), \( V_f = 0 \). We can again generalize this expression by using \( V_i \) instead of \( E \):

\[
v_L(t) = V_i e^{-t/\tau} \quad \text{Equation 2.7}
\]

Note, however, that in the general case, the initial voltage during discharge must be determined from the circuit components and Kirchhoff’s Voltage Law once the source, \( E \), has been removed, and generally has no correlation to the original source voltage.

So, what does this behaviour look like if we plot it over time (“charging transient”):

![Graph of charging transient](image)

**Figure 2.7: R-L Voltage and Current during Charging Transient**

Note: exponential transient are generally considered to have reach “steady-state” after 5 time constants (99.3% of their final value).
We have looked at the behaviour when the circuit is “charging” (energy is being built up in the magnetic field), so now let’s look at what happens when that energy is dissipated (discharged). This will occur when we step the voltage that’s impressed across the circuit from E to 0V (which is shown in Figure 2.8 by disconnecting the source and connecting the resistor to ground):

Figure 2.8: R-L Discharge Transient

Using some simple circuit analysis and Kirchhoff’s Voltage Law (KVL) we can write an expression for the voltage from the time just after we change the circuit (drop the voltage to 0), \( t = 0^+ \). (Note: we’ve reset the clock to a new arbitrary \( t = 0 \) to simplify our derivation.)

\[
-V_R + V_L = 0 \\
-(I_R) + L \frac{di}{dt} = 0; \text{ recall: } I_i = i(t=0^+) = i(y=0^-)
\]

Again, solving for \( i_L \):

\[i_L(t) = I_i e^{-t/\tau'} \quad \text{Equation 2.8}\]

where \( \tau' \) is the time constant for the decaying circuit (which may not be the same as in the charging circuit! It is always safe to use \( R_{Th} \) from the inductors perspective for both the charging and discharging cases. For a refresher on Thévenin, see ).

Substituting (5) for \( i \) into \( v_L = L \frac{di}{dt} \) and solving we get:

\[v_L(t) = V_i e^{-t/\tau'} \quad \text{Equation 2.9}\]

where \( \tau' = \frac{L}{R'} \), \( R' \) is the Thévenin resistance in this circuit; \( V_i = v_L(t=0^+) \). Note: use KVL in the above circuit to solve for \( V_i = -I_i R' \).
We can again use these expressions to verify our static analysis for initial and final values for voltage and current:

\[
\begin{align*}
    i_L(t = 0^+) &= I_i e^0 = I_i \\
    i_L(t = \infty) &= I_i e^\infty = 0 \\
    v_L(t = 0^+) &= V_i e^0 = V_i \\
    v_L(t = \infty) &= V_i e^\infty = 0
\end{align*}
\]

What does this look like over time?

Again remember: \( \tau \) and \( R_{th} \) may not be the same in the decaying circuit as in the charging circuit!
2.6 Energy Stored in an Inductor or Capacitor

Finally, let’s look at the energy that’s stored in these elements. For an inductor, the energy is stored in the form of a magnetic field due to the current in its coil. For a capacitor, the energy is stored in an electric field due to the voltage across its plates. Recall that \( P = V \times I \). Also recall that Work (Energy) = Power X Time, or: \( W = P \text{ } dt \).

Combing, we have:

\[
W = \int_{t=0}^{\infty} v(t)i(t) \text{ } dt \quad \text{Equation 2.10}
\]

Substituting the expressions developed for voltage and current through an inductor and a capacitor during the charging phase (for the simple Thévenin equivalent circuits considered above), and evaluating the integral from the beginning of the charging phase \( (t=0) \) until steady-state is reached, we arrive at the following expressions for the energy stored in an inductor or a capacitor, respectively, (assuming zero initial conditions):

\[
W_L = \frac{1}{2} LI^2 \quad \text{Equation 2.11}
\]

\[
W_C = \frac{1}{2} CV^2 \quad \text{Equation 2.12}
\]

2.7 Generalization of exponential transient behaviour

It turns out that many things in nature respond exponentially to step changes in input (those where the “pressure”, for example, in/decreases with time as the “flow” continues). In these cases, the behaviour can be described using only initial and final values during the period of interest (which is typically from when the circuit is “changed, \( t=0 \)” until steady-state is reached).

\[
x(t) = X_f - (X_f - X_i) e^{-t/\tau} \quad \text{Equation 2.13}
\]

Use KVL, KCL and \( \Omega \)’s Law to determine the initial and final values, \( X_i \) and \( X_f \) respectively, of the parameter of interest, and this expression will always describe the parameter’s behaviour over time. Note that the “final values” required in this expression are those which would be realized at steady –state, even though we may only be interested in a portion of that interval.
Example 2-3: R-L Transient using $R_{Th}$
Consider a simple example with this circuit:

i) Write the expressions for $i_L(t)$ and $v_L(t)$ for the charging and discharging phases (after the switch is closed and after it is re-opened respectively). Assume there is no current or voltage in the circuit elements prior to switch closure.

ii) Find the current in the inductor 20\(\mu\)s after the switch is initially closed

iii) Sketch the current and voltage through the inductor, as a function of time, for both phases

### Charging phase
- $R_{Th} = 470\Omega$
- $\tau = \frac{L}{R_{Th}} = \frac{0.1H}{470\Omega} = 21.3\mu s$
- $I_f = \frac{5V}{470\Omega} = 10.6mA$
- $I_i = 0$
- $V_i = E_{Th} = 5V$
- Use $I_i$ and KVL
- $V_i = 0$ (I is steady)
- Use $x(t) = X_f - (X_f - X_i) e^{-t/\tau}$
- $i_L(t) = 10.6 - (10.6 - 0) e^{-t/21.3\mu s} mA$
- $i_L(t) = 10.6 e^{-t/21.3\mu s} mA$

### Discharging (decaying) phase
- $R_{Th} = (470+220) = 690\Omega$
- $\tau' = \frac{0.1H}{690\Omega} = 14.5\mu s$
- $I_i = I_f (charging) = 10.6mA$
- $I_i = 0$
- $V_i = -I_i R_{Th} = -(10.6mA)(690\Omega) = -7.3V$
- because it is opposite in polarity from the charging phase
- $V_i = 0$ (I is steady @ 0)
- $i_L(t) = 0 - (0 - 10.6) e^{-t/14.5\mu s} mA$
- $i_L(t) = 10.6 e^{-t/14.5\mu s} mA$
\[ v_L(t) = 0 - (-05)e^{-t/\tau} \text{V} \quad \quad \quad v_L(t) = 0 - (0 - (-7.3))e^{-t/\tau} \text{V} \]
\[ v_L(t) = 0 - (-05)e^{-t/21.3\mu s} \text{V} \quad \quad \quad v_L(t) = -7.3e^{-t/14.5\mu s} \text{V} \]

ii
Use \( i_L(t) = 10.6 \left( 1 - e^{-t/21.3\mu s} \right) \text{mA} \) from “Charging” above:
\[ i_L(20\mu s) = 10.6 \left( 1 - e^{-20\mu s / 21.3\mu s} \right) \text{mA} = 6.46\text{mA} \]

iii

\[ I_L(\text{mA}) \]

\[ V_L(\text{V}) \]

Sw closed
Sw opened

Use

-充电：
\[ i_L(t) = 10.6 \left( 1 - e^{-t/21.3\mu s} \right) \text{mA} \]

-放电：
\[ i_L(t) = -7.3e^{-t/14.5\mu s} \text{mA} \]
3 Capacitors

Now let’s look at the transient behaviour of a capacitor when subjected to a step change in circuit conditions.

First, let’s again review capacitors in a static DC environment:

Recall that the capacitance of an ideal parallel-plate capacitor, \( C = \varepsilon_r \varepsilon_0 \frac{A}{d} \), where \( A \) = the plate area (\( \text{m}^2 \)), \( d \) = the distance between the plates (\( \text{m} \)), \( \varepsilon_0 \) = permittivity of free space (\( 8.854 \times 10^{-12} \text{ C}^2/\text{N} \cdot \text{m}^2 \)), and \( \varepsilon_r \) is the relative permittivity of the dielectric between the two capacitor plates (a unit-less number giving the ratio of the permittivity of the dielectric to that of a vacuum, with a vacuum considered essentially the same as space).

Also recall that the voltage across an ideal capacitor is proportional to the charge held on the plates and inversely proportional to the amount of the capacitance: \( V_c = \frac{Q}{C} \) (\( Q \) in coulombs and \( C \) in farads will give \( V \) in volts). Since the amount of charge is a function of the current and how long it has been flowing, we can also write: \( V_c = \frac{i \cdot t}{C} \). An incremental increase in voltage can be expressed: \( dV_c = \frac{i \cdot dt}{C} \), and re-arranging, we get an expression for the current: \( i = C \frac{dV_c}{dt} \) (note the “duality” with the expression for voltage across an inductor?)

3.1 Capacitors in a Steady-state DC Circuit

In a “steady-state” DC circuit, a capacitor appears like an open circuit, with no current flowing through it and a voltage across it determined by the other circuit components and their configuration.

![Diagram of capacitor circuit]
Redraw the schematic $V_{C1} = (72V) \frac{2\Omega}{(2 + 7\Omega)} = 16V$

$V_{C2} = (72V) \frac{7\Omega}{(2 + 7\Omega)} = 56V$

Alternatively, we could have figured out the current flowing through R1 and R2 (8A) and determined the voltages across the capacitors from that.

### 3.2 Adding Capacitors in Series or Parallel

Capacitors in parallel add simply (like resistors or inductors in series):

$$C_{Total} = C_1 + C_2 + C_3 + ...$$

$$C_{Total} = 8\mu F + 6\mu F + 12\mu F = 26\mu F$$

Capacitors in series add like resistors or inductors in parallel:

$$C_{Total} = \left( \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + ... \right)^{-1}$$

$$C_{Total} = \left( \frac{1}{8\mu F} + \frac{1}{6\mu F} + \frac{1}{12\mu F} \right)^{-1} = 2.667\mu F$$
3.3 Capacitive Transient Behaviour

It is mathematically simpler (and a better parallel to the inductor case), to consider a simple R–C circuit with a current source instead of a voltage source in order to derive expressions for the transient response to a step change in voltage. Converting a voltage course with a potential \( E \) and internal resistance, \( R_I \) results in the circuit shown in Figure 3.1:

\[ \sum I = 0, \quad I_{\text{Source}} - I_{R1} - I_{C1} = 0 \]

Recalling that conversion of a voltage source with potential \( E \) and internal resistance \( R_I \) will result in a current, \( I = E/R_I \), we can write this equation in terms of the source, \( E \), and the voltage across the capacitor, \( V \), to derive an expression very similar in form to that derived for the current through an inductor:

\[
\frac{E_{\text{Source}}}{R_I} - \frac{V}{R_I} - C \frac{dv}{dt} = 0
\]

\[
v_C(t) = E \left(1 - e^{-t/\tau} \right)
\]

\[
i_C(t) = \frac{E}{R} e^{-t/\tau}
\]

\[
n_C(t) = i_t e^{-t/\tau}
\]

\[
i_C(t) = -\frac{V_i}{R_{th}} e^{-t/\tau}
\]

\[
v_C(t) = V_t e^{-t/\tau}
\]

Solving this equality for \( V \), the voltage across the capacitor yields:

\[
v_C(t) = E \left(1 - e^{-t/\tau} \right)
\]

where \( \tau \) is the time constant for this circuit: \( R_I C \)
Now, recalling that \( i_C = C \frac{dv}{dt} \), and using the expression for voltage across the capacitor derived above, we can also develop an expression for \( i_C \) for this same circuit as a function of time:

\[
i_c = C \frac{dV}{dt} = C \frac{d}{dt} \left( E \left( 1 - e^{-\frac{t}{\tau}} \right) \right)
\]

solving for \( i_C \), we are left with:

\[
i_C(t) = \frac{E}{R} e^{-t/\tau} \quad \text{Equation 3.1}
\]

where we observe that \( E/R \) is the initial, or maximum current that will flow in the circuit at time, \( t = 0^+ \), just after the switch is closed. Note that \( \tau \), the time constant, is \( RC \) in this case, where again, \( R \) is the Thevenin equivalent resistance seen by the capacitor.

In keeping with previous forms, this can be written in terms of the final value:

\[
i_C(t) = I_i e^{-t/\tau} \quad \text{Equation 3.2}
\]

For the decaying transient phase, it can also be shown that the current, \( i_C \) is:

\[
i_C(t) = -\frac{V_i}{R_{Th}} e^{-t/\tau'} \quad \text{Equation 3.3}
\]

Where \( V_i \) and \( I_i \) are the conditions at \( t = 0^+ \). The voltage, \( v_C \) is given by:

\[
v_C(t) = V_i e^{-t/\tau'} \quad \text{Equation 3.4}
\]

Again, the time constant, \( \tau' = RC \), is that of the decaying circuit, which may not be the same as the charging circuit (always check \( R_{Th} \) if in doubt).

We can now plot these phases over time ("charging and discharging transients"):
Again, we have ignored the possibility of any pre-existing conditions (like a charge on the capacitor) in order to simplify the derivation. The implied assumption is that the voltage was 0 at time $t=0$ (i.e. when the circuit conditions for the capacitor changed). Of course, in many situations the initial voltage is not 0, and this pre-existing condition must be taken into consideration. The general expression will again always give the correct expression as long as the starting and ending values are determined correctly:

$$x(t) = X_f - \left(X_f - X_i\right)e^{-t/\tau} \quad \text{Equation 2.13}$$

### 3.4 Energy stored in a Capacitor:

In a derivation very similar to that used to determine the energy stored in an inductor (using the expressions previously derived for the voltage across and current through a capacitor), we can determine that the energy stored in a capacitor will be:

$$W_c = \text{Power} \times \text{time} = \int_0^t i_c v_c \, dt = \frac{CV^2}{2}$$
Example 3-1: R-C Transient Response

Let’s look at an example similar to the inductor’s case:

![Circuit Diagram]

Example 3-2: R-C Example 1

i) Write the expressions for $i_C(t)$ and $v_C(t)$ for the charging and discharging phases (after the switch is closed and after it is re-opened respectively). Assume there is no current or voltage in the circuit elements prior to switch closure.

ii) Find the voltage across the capacitor 40 $\mu$s after the switch is initially closed.

iii) Sketch the current through and voltage across the capacitor, as a function of time, for both charging and discharging phases.

**Charging phase**

$R_{Th} = 470\Omega$

$\tau = R_{Th}C = (470\Omega)(.1\mu F) = 47\mu s$

$V_f = E_{Th} = 5V$

$V_i = 0$

$I_i = \frac{E_{Th}}{R_{Th}} = \frac{5V}{470\Omega} = 10.6mA$

$I_f = 0$  (C is fully charged)

Use

$x(t) = X_f - (X_f - X_i) e^{-\frac{t}{\tau}}$

$i_C(t) = 0 - (0 - 10.6)e^{-\frac{t}{\tau}}mA$

$i_C(t) = 10.6e^{-\frac{t}{47\mu s}}mA$

$v_C(t) = 5 - (5 - 0)e^{-\frac{t}{\tau}}V$

$v_C(t) = 5\left(1 - e^{-\frac{t}{47\mu s}}\right)V$

**Discharging (decaying) phase**

$R_{Th} = (470+220) = 690\Omega$

$\tau' = (R_{Th})(C) = (690\Omega)(.1\mu F) = 69\mu s$

$I_i = \frac{V_f}{R_{Th}} = \frac{-5V}{690\Omega} = -7.25mA$

$I_f = 0$

$V_i = V_f = 0$  (C is fully discharged)

$i_C(t) = 0 - (0 - (-7.25))e^{-\frac{t}{\tau}}mA$

$i_C(t) = -7.25e^{-\frac{t}{69\mu s}}mA$

$v_C(t) = 0 - (0 - (5))e^{-\frac{t}{\tau}}V$

$v_C(t) = 5e^{-\frac{t}{69\mu s}}V$
ii

Use \( v_C(t) = 5 \left( 1 - e^{-\frac{t}{47\mu s}} \right) \) V from “Charging” above:

\[
v_C(40\mu s) = 5 \left( 1 - e^{-\frac{40\mu s}{47\mu s}} \right) V = 2.87V
\]

iii

Again remember: \( \tau \) and \( R \) may not be the same in the decaying circuit as in the charging circuit! (And in this case again, they’re not!)

### 3.5 Inductive and Capacitive Transients – SUMMARY

(Jackson: Ch 13(C), 17(L). Boylestad Ch 10(C), 11(L))

Changing current, \( i \), and voltage, \( v \), in an inductor or capacitor:
Transient Response timeframes:

Previous conditions $t = -t$ to $0-

Time period of interest $t = 0$ to $t (\infty)$

$t = 0$ (time when the circuit changes from L’s or C’s perspective)

Use $x(t) = X_f - (X_f - X_i) e^{-\frac{t}{\tau}}$, where $X_f$ and $X_i$ are the final and initial values during the period of interest, and $\tau$ is the time constant ($\frac{1}{R_{th}}, \frac{1}{R_{th}C}$) and $R_{th}$ is the resistance “seen” by the inductor or capacitor again, during the period of interest.
3.6 Determining Initial and Final Values in R-C and R-L Circuits:

Initial

For L

\[ i_L(t=0^+) = i_L(t=0^-) \]
\[ v_L(t=0^+) \text{: use old current}(t=0^-) \text{ and new circuit}(t=0^+) \text{ and KVL} \]

For C

\[ v_C(t=0^+) = v_C(t=0^-) \]
\[ i_C(t=0^+) \text{: use old voltage}(t=0^-) \text{ and new circuit}(t=0^+) \text{ and KVL} \]

Final

At steady-state, the voltage across and inductor or the current through a capacitor is always 0 (\( v_L(t=\infty)=0V \), \( i_C(t=\infty)=0A \)).

For the final current in an inductor or the final voltage across a capacitor, use basic circuit analysis, remembering that at steady-state (SS, \( t=\infty \)), an inductor looks like a short and a capacitor looks like an open.

\[ L_{SS}: \quad C_{SS}: \]

ss = steady-state
4 Alternating Current (AC) and Power

Up until now, we’ve considered circuits that were powered by direct current (DC) sources (i.e. the voltage of the source did not vary with time). When we considered transient behaviour, we started to consider electrical quantities that did vary with time, at least during a transition period between one static state and the next. We are now going to consider sources where the voltage varies continuously with time.

While there are a number of time varying voltage situations, we are going to focus mainly on the case where the source varies periodically, continuously and the polarity alternates. We will look specifically at the sinusoidal voltage and current that provides electrical energy to home and industries around the world.

Recall that in the DC case, our source voltage, or current, were constant with time, unless we changed the circuit somehow. Then there would be a “transient” period of adjustment until the voltage and current returned to a steady value. In this case, the magnitude and polarity were the only parameters we needed to know to describe the voltage or current. (E.g.: V = +5V. In fact, the ‘+’ is usually assumed, even if it is not explicitly stated.)

In the Alternating Current (AC) case, there are a number of additional parameters needed to adequately describe voltage or current. As implied above, AC is generated or produced such that the signal is sinusoidal. A general parametric equation that describes and AC waveform is:

\[ e_m = E_{\text{Max}} \sin(\omega t) \]

This signal (voltage) is considered sinusoidal, because its value varies as the sine of the argument, \(\omega t\), and it is periodic because there is a specific pattern that repeats every \(T\) seconds (\(T\) is called the period of this waveform, measured in seconds).

Refer to the plot of the waveform: \(e(t) = Asin(\omega t)V\) as a function of time (Figure 4.1). Consider the following terms commonly used to describe periodic signals:

- \(A\), the amplitude = \(E_{\text{Max}}\), also known as the ‘zero-to-peak’ signal level (\(V_{0-p}\))
- \(\omega\) the angular frequency, measured in Radians per second (R/s),
- \(f\), frequency (= the number of ‘cycles’ or repeated patterns per second), measured in cycles per second or Hertz (Hz). Note, therefore, \(\dot{\omega} = 2\pi f\)
• $T\left(= \frac{1}{f}\right)$, the period, or length of time in seconds required to complete one ‘cycle’.

Figure 4.1: An AC Sinusoidal Waveform

The argument for the trigonometric function in these parametric expressions can be given in several, equivalent ways:

$$\omega t = (2\pi f t) = \left(\frac{2\pi}{T} t\right)$$

While AC signals can have any frequency, power systems select and use one frequency, and control it very strictly. In North America the frequency is 60 Hz.

In the DC case, we only need to know the magnitude and polarity of a signal to be able to combine them or determine their interaction. In the AC case, the polarity is constantly (and periodically) alternating, but different signals can be displaced in time, even if they have the same frequency. This results in a phase difference, which must be considered when determining the interaction between multiple sinusoidal waveforms.

This phase difference is represented by adding a phase offset or phase angle, $\phi$, to the argument in the general parametric equation:

$$v(t) = A\sin(\omega t + \phi)V$$

where $\phi$ may be a positive (leading) or negative (lagging) angle, and is, by convention, <180°. **
Consider the examples: \( Asin(\alpha t + \phi_1) \) and \( Asin(\alpha t - \phi_2) \). As we can see in Figure 4.2, the phase-offset signals are identical except that they are shifted slightly on the horizontal axis. Remember, the resulting argument for the sine function must always reduce to a consistent angular argument, even though it is not uncommon to have \( \omega \) expressed in R/s and the phase offset, \( \phi \) given in degrees. In addition, although we are plotting the function with respect to time, it is often convenient to show the equivalent angular argument, in radians or degrees, on the horizontal axis. The number shown on the horizontal (time) axis in this case is the net argument of the trigonometric function, \((2\pi f t + \phi)\) or \((\alpha t + \phi)\).

![Image](425x115.png)

**Figure 4.2: Example of Lead - Lag Phase Difference**

**Representation of Sinusoidal Signals**

In the *Direct Current* (DC) environment, the current direction and voltage polarity do not change with time, but in the *Alternating Current* (AC) environment, both the direction of current and the voltage polarity change twice each cycle. DC, non-time-varying values were typically represented using upper-case letters (e.g. \( V \), \( I \)). AC, time-varying values are typically represented using lower-case letters (e.g. \( v \), \( i \)).

As a convenient convention, voltage sources in AC circuits are shown with a ‘+’ indicating the source of positive current during the positive half cycle. (Figure 19) During one half of the cycle, the ‘+’ terminal is positive and the current flows out of it as it would from a DC source. During the other half cycle, the polarity is reversed and the ‘+’ terminal is actually negative in polarity and the current is flowing into the source.

![Image](459x115.png)

**Figure 19: A simple AC Circuit Representation**
Another interpretation is that when the value of the expression $e(t) = A\sin(\omega t)$ is positive, current flows out of the ‘+’ terminal. When the value of the expression is negative, it flows in. Consider our earlier expression: $e(t) = A\sin(\omega t)V$ plotted in Figure 20. When the expression yields a positive value, the current flow is also positive according to the polarity indication on the source. During the other half-cycle, it is the opposite.

![Figure 20: Current Direction in an AC circuit](image)

Figure 20: Current Direction in an AC circuit
AC Signals – Example 1:

Given \( e(t) = 10\sin(1256.6t - \pi/3) \text{V} \), find:
- \( f \) and \( T \) (frequency and period)
- \( t \) when \( e \) is first 0V
- \( t \) when \( e \) is at its first maximum
- \( e(\@ t=0) \)
- expression for \( i(t) \)
- plot \( e(t) \) and \( i(t) \).

The general expression for a sinusoidal signal is:

\[ e(t) = E_{\text{Max}} \sin(\omega t + \phi) \text{V} \]

By inspection, we see that:
- \( E_{\text{M}} = 10 \text{V} \)
- \( \omega = 1256.6 \text{R/s} \)
- \( \phi = -\pi/3 \text{R} \ (-60^\circ) \)
- \( \omega = 2\pi f \)
- \( T = 1/f = 5 \text{ms} \)

\( E(t) = 0 \) when \( \sin(1256.6t - \pi/3) = 0 \), or when the argument, \( (\omega t + \phi) = 0^\circ, \pi(180^\circ), 2\pi (360^\circ) \), etc.

\[ 1256.6t - \pi/3 = 0 \]
\[ t = \frac{\pi/3}{1256.6 \text{R/s}} = 0.83 \text{ms} \]

\( e(t) \) is a maximum when \( \sin(1256.6t - \pi/3) \) is maximum, or when \( (\omega t + \phi) = \pi/2 (90^\circ), 3\pi/2 (270^\circ) \), etc.

\[ 1256.6t - \pi/3 = \pi/2 \]
\[ t = \frac{5\pi/6}{1256.6 \text{R/s}} = 2.08 \text{ms} \]

you consider maximum magnitude, then it occurs every 2.5ms thereafter.

\( e(0) = 10\sin(1256.6(0) - \pi/3) \text{V} = 10\sin(-\pi/3) = -8.66 \text{V} \)

In the AC environment, Ohm’s Law still applies:

\[ i(t) = \frac{v(t)}{R} = \frac{10\sin(1256.6t - \pi/3)}{4 \Omega} = 2.5\sin(1256.6t - \pi/3) \text{A} \]

Plotting both waveforms (Figure 4.6):
An additional parameter: \( I_{\text{peak-peak}} = 5 \text{A} \)

Another question: When is \( e \) first 5V?

\[
10\sin\left(1256.6t - \frac{\pi}{3}\right)V = 5V
\]

\[
\sin\left(1256.6t - \frac{\pi}{3}\right)V = 0.5V
\]

\[
\sin^{-1}(0.5) = 1256.6t - \frac{\pi}{3} : \; t = 1.25ms
\]
AC Signals – Example 2:

Given the general parametric expression for an AC sinusoidal signal:

\[ e(t) = A\sin(\omega t + \phi)V \]

and the information given on the accompanying graph, determine the actual expression for this AC waveform.

---

Given the general parametric equation for a sinusoid: \( e(t) = A\sin(2\pi f t + \phi) \); and

\[ e(0) = 147V; \quad e(5.555ms) = 0V; \quad \phi = \pi/3 = 60^0 \]

Find the expression for the waveform given in the plot.

---

i) \[
147 = A\sin\left(2\pi f \left(0 + \frac{\pi}{3}\right)\right) = A\sin\left(\frac{\pi}{3}\right)
\]

\[
A = \frac{147}{\sin(60^0)} = 169.7V
\]

\[ e(5.555ms) = 0V \]

\[
169.7\sin\left(2\pi f \left(0.005555 + \frac{\pi}{3}\right)\right) = 0V , \text{ therefore} \sin\left(2\pi f \left(0.005555 + \frac{\pi}{3}\right)\right) = 0
\]

\[
sin\theta = 0 \text{ at } 0, \pi, 2\pi \text{ etc.}; \text{ take } \pi \text{ by observation:} \\
2\pi f \left(0.005555 + \frac{\pi}{3}\right) = \pi
\]

\[
f = \frac{\pi - \frac{\pi}{3}}{2\pi \left(0.005555\right)} = 60Hz
\]

\[
\omega = 60Hz \left(2\pi\right) = 377R/s
\]

\[
T = \frac{1}{f} = \frac{1}{60Hz} = 16.67ms
\]

\[ e(t) = 169.7\sin(377T + \pi/3)V \]
4.1 Average and Effective (RMS) Values

When working with continuous, time-varying signals (e.g. AC voltages and currents), it is often inconvenient to use their instantaneous values to determine the effect they have on circuit components and their effectiveness in transferring energy. To simplify these tasks, we use two statistical properties of the waveforms: the average (or mean) value, and the effective (or RMS) value. The average value, as the name would suggest, is simply the mean value of the signal over time. This statistical value is often useful for determining parameters like torque, force, or magnetic field strength resulting from an AC current. When considering energy, the electrical power is related to the square of either the voltage or the current, and we must consider a measure that reflects this “effect”. This is the effective value of a signal, and is determined using what is statistically the variance of the signal. The method usually used to derive this value involves taking the square Root of the Mean of the Square of the signal, the acronym of which leads to its common name: the RMS value.

4.1.1 Average Value:
For a time varying, periodic signal, the average value is given by:

\[
\text{Average} = \frac{\int v(t) \text{ (or, area under the curve)}}{\text{time base}}
\]

Statistical Note: Continuous, periodic signals typically exhibit “stationarity”, a characteristic that means that for the first two statistical moments, namely the mean and the variance, the value determined over one period, T, is the same as the value over all time.

This means that for a continuous, periodic signal, the average value can be defined as:

\[
\bar{v}(t) = \frac{1}{T} \int_{0}^{T} v(t) \, dt.
\]

An similar definition is used for current, i(t), or any other continuous, periodic variable.

In the DC world, the average DC values are often a reasonable indicator or energy or force. In the AC world, however, the average value of the sinusoidal waveforms we encounter is zero, while we know intuitively that the power is definitely not. We must consider another statistical characteristic of an AC waveform to determine the real effect it has.

4.1.2 Effective Value:
The definition of effective value is an equivalent value that has the same heating effect as a DC voltage or current of the same value. This means it delivers the same power as a DC signal source of the same value. This is due to the “squared” quantities that determine the power in electric circuits. Recall that in the DC world,
\[ P = \frac{V^2}{R} = I^2 R \]

and similarly, in the AC world,

\[ p(t) = \frac{v(t)^2}{R} = i(t)^2 R \]

Consider the following plot of a voltage waveform and the resultant instantaneous power, \( p(t) \) into a 1\( \Omega \) resistor. We note that the power is i) always positive, and ii) varies at twice the frequency as the voltage (or current). The trigonometric identity:

\[ (\sin \theta)^2 = \frac{1}{2} - \frac{1}{2} \cos 2\theta \]

also illustrates this result.

Because the heating effect is proportional to the square of the voltage or current, it does not increase linearly, and we must consider the average of the squared quantity to determine an equivalent value. The effective value is defined as:

\[
\text{Effective Value (RMS)} = \frac{1}{T} \int_{0}^{T} v(t)^2 \, dt
\]

This value is commonly referred to by the acronym RMS, because we take the square Root of the Mean of the voltage or current Squared to arrive at the appropriate equivalent value. Using this RMS value to determine the power in a circuit will actually give us the average power; identical to what we would get if we used the same DC value.

If we take the \( v(t) = V_m \sin \left( \frac{2\pi}{T} t \right) V \), the parametric equation for a sine wave, find the RMS (Effective) value is equal to:

\[ V_{RMS} = \frac{V_M}{\sqrt{2}} \]
This factor, \( \frac{1}{\sqrt{2}} \) (or .707) is commonly used to simply determine the RMS (effective) value of a sinusoidal AC signal.

For example, the RMS value for the voltage waveform described by:

\[
v(t) = 170 \sin(277t) V
\]

is 120.2V, the typical household branch circuit voltage! Note again, that by convention this would be written as: \( V = 120V \) or \( V = 120V\text{AC} \), using an upper-case V to signify that this is the effective value, or has the equivalent heating power of a 120V DC source.
4.2 Adding and Subtracting AC Voltages and Currents

As mentioned previously, when combining DC voltages or currents, we only need to know the polarity (voltage) and direction (current). In the case of alternating voltages and currents, the polarity and direction are periodically changing. There are mathematical and graphical ways to combine these, but these can be quite cumbersome and time consuming, especially when there are a number of cases to determine.

Fortunately, George Proteus Steinmetz developed a method of combining AC signals which greatly simplifies this task; we know it as “Phasor Notation”. It is based on the more complicated general, mathematical trigonometric method of combination, but is greatly simplified when applied to sinusoidal voltages or currents of exactly the same frequency. Remembering the general parametric equation for a sinusoidal signal, we notice that if the frequency of all the signals concerned is exactly the same, they will differ only in magnitude and phase angle. It is these two characteristics that are used to do a transformation of the signal into a two-dimensional, complex valued number which can be mathematically combined quite simply by doing complex addition or subtraction to combine the signals. Once combined, they can be converted back in to a time function representation should that be required to determine time-related parameters.

General transformation of a sinusoidal signal from the time domain to the phasor domain:

<table>
<thead>
<tr>
<th>Time Domain to Phasor Domain Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time</strong>: ( v(t) = A\sin(\omega t + \phi) V )</td>
</tr>
<tr>
<td><strong>Phasor</strong>: ( V = \frac{A}{\sqrt{2}} \angle \phi )</td>
</tr>
<tr>
<td><strong>Phasor</strong> (polar form), ( V = \frac{A}{\sqrt{2}} \angle \phi )</td>
</tr>
<tr>
<td><strong>Phasor</strong> (rectangular form), ( A \cos(\phi) + j\frac{A}{\sqrt{2}} \sin(\phi) )</td>
</tr>
</tbody>
</table>

*Note: phasors are conventionally shown as RMS values, but the concept works equally well for maximum (i.e. peak) values.*
Reverse transformation from the *phasor* domain back into the *time* domain:

**Table 4.2: Phasor to Time Conversion**  
**Phasor Domain to Time Domain Conversion**

**Phasor:** \( V = B \angle \theta V \) (polar form) or \( V = (X + jY)\) (rectangular form)

**Time:**  
\[ v(t) = \sqrt{2} B \sin(\omega t + \theta) V, \text{ or } v(t) = \sqrt{2} \sqrt{X^2 + Y^2} \sin(\omega t + \tan^{-1}\left(\frac{Y}{X}\right)) V \]

Of course in the reverse case, there is know way of determining \( \omega \) without additional information!

**The Phasor Diagram:**  
Phasors are just vectors in a 2-D complex plane. A Phasor Diagram can be used to illustrate their relationship and addition. An example:

**Final note:** Phasors are transformations of *time-varying signals only*. Other AC circuit elements, as we will discuss next, will also be represented as complex-valued numbers, but are **not** considered phasors.
4.3 Opposition to Flow in AC Circuits

In a steady-state DC circuit, resistors, capacitors and inductors behave relatively simply: resistors obey Ohm’s Law, capacitor look like open circuits, and inductors look (almost) like short circuits. If the AC world, these elements can react a little differently.

Resistors still obey Ohm’s Law, according to the instantaneous voltage or current at any given time. Thus, a resistor connected to an AC voltage source will allow a current to flow according to Ω’s Law at any instant, resulting in a sinusoidal current of exactly the same frequency and phase as the voltage, and a magnitude predicted by Ω’s Law at any instant.

For inductors that are connected to an AC source:

\[
V_L = L \frac{dI}{dt} = L \frac{d}{dt} \left( I_M \sin \omega t \right) = \omega L M \cos \omega t = V_M \cos \omega t = V_M \sin \left( \omega t + 90^\circ \right) \text{ where } V_M = \omega L M
\]

Note that the voltage leads the current by 90° (or, current lags voltage: eLi).

Define the term \( X_L = \omega L \), as the Inductive Reactance, which is the Alternating Current world's “opposition to flow” for an inductor, and is measured in ohms, Ω. (L is the inductance in Henries, and \( \omega \) is in Radians/second). Note also that the reactance is directly proportional to both the frequency and the Inductance.

\( X_L = \frac{V_M}{I_M} \) which illustrates its equivalence to resistance, or “opposition to flow”.

Similarly for Capacitors:

\[
i_c = C \frac{dV}{dt} = C \frac{d}{dt} \left( V_M \sin \omega t \right) = \omega C V_M \cos \omega t = I_M \cos \omega t = I_M \sin \left( \omega t + 90^\circ \right) \text{ where } I_M = \omega C V_M
\]

Note that here the current leads the voltage (voltage lags current, \( iCe \)) by 90°. The quantity, \( \frac{V_M}{I_M} = \frac{1}{\omega C} \), where C is in Farads (F) is called the Capacitive Reactance, \( X_C \) and is the Alternating Current world's equivalent of Resistance, R, for a capacitor and is measured in ohms, Ω. Note that in this case, the reactance is inversely proportional to frequency and capacitance.

Note that R, \( X_L \) and \( X_C \) are all scalar quantities representing opposition to flow, and will all obey Ω’s Law at any instant in time.
If we have converted our voltages and currents to phasors, we need to do a conversion of R and X so that the “opposition to flow” will obey Ω’s Law perfectly and simply in the Phasor Domain.
### 4.3.1 Impedance, Z

Define one more new term: **Impedance, Z**, which represents the “opposition to flow” in the phasor domain. It is also measured in Ohms (Ω), and is defined as:

\[
Z = \frac{\text{phasor\_voltage}}{\text{phasor\_current}}, \text{ or } Z = \frac{V}{I}
\]

(Note: since Z is just a ratio, it can be calculated from either RMS or Max values.) Since the phasor voltage and phasor current in this definition are complex valued, the impedance, Z, is also complex valued. It is also measured in Ohms (Ω) and can represent the opposition to flow of any element or combination of elements.

Using our previous determinations of the reactance of the basic elements (R, L, C) and the phase relationship of the voltage and current when they are connected to an AC source, we can develop an expression for the impedance of each type.

For **resistors**, since voltage and current are in phase, this is straightforward:

\[
Z_R = \frac{V\angle 0^\circ}{I\angle 0^\circ} = Z\angle 0^\circ \Omega = X_R\angle 0^\circ = R\angle 0^\circ \Omega
\]

For **inductors**, since current is 90° behind the voltage (eLi)

\[
Z_L = \frac{V\angle 0^\circ}{I\angle -90^\circ} = Z\angle 90^\circ \Omega = X_L\angle 90^\circ \Omega = \omega L\angle 90^\circ \Omega = j\omega L
\]

For **capacitors**, since current is 90° ahead of the voltage (iCe)

\[
Z_C = \frac{V\angle 0^\circ}{I\angle +90^\circ} = Z\angle -90^\circ \Omega = X_C\angle -90^\circ \Omega = \frac{1}{\omega C}\angle -90^\circ \Omega = -j\frac{1}{\omega C}
\]

For any **series** combination of R-L-C in AC circuits, the impedance, Z = R +j(X_L –X_C), a linear combination of the individual elements’ impedance.
Impedances in series or in parallel ad exactly like resistors in series or in parallel, except that we must use the complex-valued numbers.

Using this complex impedance, \( Z \), along with phasor voltage and current representations, **Ohm's Law** can be applied in a straightforward fashion to A.C. circuits made up of combinations of resistive, inductive and capacitive components. **Note**: that although \( Z \) is complex, it is *not a phasor* as it is not a sinusoidally varying quantity.

Note: that reactance(\( X \)) and resistance(\( R \)) are scalars and *not* complex valued. While impedance is a complex vector which can be represented on a 2-D plane, \( Z \) does *not* represent a time-varying quantity (like a phasor does).

Impedance, \( Z \), can also be represented and added on a 2-D plane called an Impedance Diagram (similar to, but *not* a Phasor Diagram!):
Finally our favourite laws. These are virtually identical to those for the DC world, except we use (complex-valued) phasors:

**Ohms Law for AC Circuits:**

\[ V = IZ, \quad \text{and} \quad Z = \frac{V}{I} \]

**Kirchhoff’s Laws for AC circuits:**

**KVL:** The \( \Sigma \) phasor voltages around a loop = 0

**KCL:** The \( \Sigma \) phasor currents into a node = 0

Norton and Théven equivalent and current/voltage source conversion also work exactly the same as the DC counterparts, except that phasor and complex impedance values are used.

With phasor voltages and currents and [complex] impedances, we can apply almost all of the same rules and Laws we used in DC circuits in the AC world.

### Table 4.3: Complex-number Math Summary

<table>
<thead>
<tr>
<th>Rectangular Coordinates</th>
<th>Polar Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A + jB) + (C + jD)</td>
<td>( + )</td>
</tr>
<tr>
<td>(A + jB) - (C + jD)</td>
<td>( - )</td>
</tr>
<tr>
<td>(A + jB) ( \times ) (C + jD)</td>
<td>( \times )</td>
</tr>
<tr>
<td>(A + jB) ( \div ) (C + jD)</td>
<td>( \div )</td>
</tr>
</tbody>
</table>

Add like vectors; usually easiest to convert to rectangular coordinates then add. If \( \theta \) is the same for both, can add magnitudes

As above, subtract like vectors, but easiest to convert to rectangular coordinates first. If \( \theta \) is the same for both, can subtract magnitudes

\[
\begin{align*}
&\quad (A\angle \theta_1)(B\angle \theta_2) = (A\times B)\angle(\theta_1+\theta_2) \\
&\quad (A\angle \theta_1)/(B\angle \theta_2) = (A/B)\angle(\theta_1-\theta_2)
\end{align*}
\]

**NOTE:** quantities in the phasor domain can be, and usually are, given in **effective** or **RMS** terms. Generally assume that any Phasor quantities are **RMS** unless otherwise stated.
### Table 4.4: Phasor Representation Summary

<table>
<thead>
<tr>
<th>Time Domain</th>
<th>Phasor Domain (Polar)</th>
<th>Phasor Domain (Rectangular)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Asin(}\omega t+\theta)$</td>
<td>$A \angle +\theta$</td>
<td>$A\cos\theta + jA\sin\theta$</td>
</tr>
<tr>
<td>$A$</td>
<td>$A$</td>
<td></td>
</tr>
</tbody>
</table>

- For $A\angle +\theta$, the vector extends $A$ units from the origin at an angle of $\theta$ radians from the positive real axis.

<table>
<thead>
<tr>
<th>Time Domain</th>
<th>Phasor Domain (Polar)</th>
<th>Phasor Domain (Rectangular)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Asin(}\omega t-\theta)$</td>
<td>$A \angle -\theta$</td>
<td>$A\cos\theta - jA\sin\theta$</td>
</tr>
<tr>
<td>$A$</td>
<td>$A$</td>
<td></td>
</tr>
</tbody>
</table>

- For $A\angle -\theta$, the vector extends $A$ units from the origin at an angle of $\theta$ radians from the negative real axis.

### Equations

- $A\cos(\omega t+\theta) = A\sin(\omega t+\theta+90^\circ)$
- $A\sin(\omega t+\theta) = A\cos(\omega t+\theta-90^\circ)$
- $-A\sin(\omega t+\theta) = -A\cos(\omega t+\theta+180^\circ)$

### Phasor Operations

- If $A\angle +\theta = (A\cos\theta + jA\sin\theta)$, then $A\angle -\theta = (A\cos\theta - jA\sin\theta)$
Eg. 1: An R-L-C series AC circuit:

Given the current magnitude of 2.48A, and the reactance values:

\[ V_R = 116.6V \]
\[ V_L = 93.5V \]
\[ V_C = 65.7V \]

**Note:** these are scalar magnitudes!

- a) Determine the impedance of each element, and their total impedance as seen by the source.
- b) Determine the phasor voltage across each element and the phasor voltage of the source

**Another Note:** the default assumption is that any values for V or I are RMS, and component values are for reactance an in \( \Omega \)s unless specified otherwise.

We can now write the impedance of each element (note that the reactances, \( X \), are already given):

\[ Z_R = 47 \angle 0^\circ \Omega, \quad Z_L = 37.7 \angle 90^\circ \Omega, \quad \text{and} \quad Z_C = 26.5 \angle -90^\circ \Omega \]

And now calculating the phasor voltages across each element: *(Note: we’ll assume the values are now in RMS for convenience.)*

\[ V_R = IZ_R = (2.48 \angle 0^\circ A)(47 \angle 0^\circ \Omega) = 116.6 \angle 0^\circ V \]
\[ V_L = (2.48 \angle 0^\circ A)(37.7 \angle 90^\circ \Omega) = 93.5 \angle 90^\circ V \]
\[ V_C = (2.48 \angle 0^\circ A)(26.5 \angle -90^\circ \Omega) = 65.7 \angle -90^\circ V \]

And the phasor total is:

\[ E = V_R + V_L + V_C = (116.6 +j0) + (0 +j93.5) + (0 -j65.7) = 119.9 \angle 13.4^\circ V \]

The total impedance is:

\[ Z_T = Z_R + Z_L + Z_C = (47 +j0) + (0 +j37.7) + (0 -j26.5) \]
\[ = (47 +j11.2) \Omega = 48.3 \angle 13.4^\circ \Omega \]

We can do a check:

\[ Z_T = V/I = 119.9 \angle 13.4^\circ V/2.48 \angle 0^\circ A = 48.3 \angle 13.4^\circ \Omega \checkmark \]
Eg. 2: A series - parallel AC circuit

Given:
\[ E_S = 35 \angle 23^0 V = (32.2 + j13.68)V, \]
\[ f = 30kHz, \]
\[ R = 10\Omega, \quad C = .22\mu F, \quad L = 100\mu H \]


First, calculate the impedances:
\[ Z_C = -\frac{j}{\omega C} = -\frac{j}{2\pi (30kHz)(.22 \times 10^{-6} F)} = 24.11 \angle -90^\circ \Omega = (0 - j24.11)\Omega \]
\[ Z_{RL} = Z_R + Z_L = 10 + j\omega L = 10 + j(2\pi \cdot 30kHz \cdot 100 \times 10^{-6} H) = 21.34 \angle 62.05^\circ \Omega = (10 + j18.85)\Omega \]

Calculate the branch currents and voltages across the elements:
\[ I_C = \frac{E_S}{Z_C} = \frac{35 \angle 23^0 V}{24.11 \angle -90^\circ \Omega} = 1.45 \angle 113^\circ A \]
\[ I_{RL} = \frac{E_S}{Z_{RL}} = \frac{35 \angle 23^0 V}{21.34 \angle 62.05^\circ \Omega} = 1.64 \angle -39.05^\circ A \]
\[ V_C = I_C Z_C = (1.45 \angle 113^\circ A)(24.11 \angle -90^\circ \Omega) = 35 \angle 23^\circ V = E_S \]
\[ V_R = I_{RL} Z_R = (1.64 \angle -39.05^\circ A)(10 \angle 0^\circ \Omega) = 16.4 \angle -39.05^\circ V \]
\[ V_L = I_{RL} Z_L = (1.64 \angle -39.05^\circ A)(18.85 \angle 90^\circ \Omega) = 30.91 \angle 50.95^\circ V \]

The total source current is just the sum:
\[ I_S = I_C + I_{RL} = 1.45 \angle 113^\circ A + 1.64 \angle -39.05^\circ A = .7687 \angle 23.25^\circ A \]

We can also check to make sure the voltages across \( R \) and \( L \) add to \( E_S \):
\[ E_S = V_R + V_L = 16.4 \angle -39.05^\circ V + 30.0 \angle 50.95^\circ V = 35.0 \angle 23.0^\circ V \]

Finally, we can calculate the total impedance of the circuit:
\[ Z_T = \frac{E_S}{I_S} = \frac{35.0 \angle 23.0^\circ V}{.7687 \angle 23.25^\circ A} = 45.5 \angle -0.25^\circ \Omega \]

Note: These phasors and impedances can be plotted in their respective 2-D complex planes. However, not all values can be graphically added. \( I_C \) and \( I_{RL} \) will add correctly to \( I_S \), and \( V_R \) and \( V_L \) will add to \( E_S \), and \( Z_R \) and \( Z_L \) will total \( Z_{RL} \), other combinations may require a separate approach. To graphically add \( Z_C \) and \( Z_{RL} \) for instance, we would have to convert them to admittances and add them on an Admittance Diagram.
**Eg. 3: Some more practice with AC circuits**
Find the unknown voltages for the following circuits and express you answer in polar notation. (Note: there are two possible answers for (b); provide both.)

![AC Circuit Diagram](image)

a) Common current, therefore $V_C$ must be $90^\circ$ behind (lagging) $V_R$. Their vector sum must also equal $E_S$. (Refer to the Phasor diagram)

Magnitude of $|V_R| = \sqrt{(150^2 - 80^2)} = 126.9$ V, and the angle for $V_C$,

$$\theta = -\cos^{-1}\left(\frac{80}{150}\right) = -57.8^\circ$$

so:

$$V_C = 80\angle -57.8^\circ V$$

$$V_R = 126.9\angle (-57.8^\circ + 90^\circ) = 126.9\angle 32.1^\circ V$$

b) $V_C$ and $V_L$ are $180^\circ$ apart (common current), and the only way their sum can be at $0^\circ$ is if they are also both on the horizontal axis, with an angle of $0^\circ$ or $180^\circ$. Since $V_L + V_C = E_S$, if $V_C$ is $200\angle 0^\circ V$, $V_L$ must be $50\angle -180^\circ V$. If $V_C = 200\angle -180^\circ V$, then $V_L$ must be $350\angle 0^\circ V$

### 4.4 Power in AC Circuits
As in the DC world, power in the AC world at any instant is \((\text{voltage}) \times (\text{current})\). In the DC case, where the voltage and current were not changing with time, this was a straightforward calculation. In the AC case, where the voltage and current are not only changing with time and alternating polarity, but they may also be out of phase with each other. As a result, the instantaneous power is a function of time, and is not necessarily “in sync” with either the voltage or the current!

To examine this from a mathematical point of view, consider a basic expression for the instantaneous power as a function of time:

\[
p(t) = i(t)v(t) = I_M \sin(\omega t)V_M \sin(\omega t + \theta)W
\]

Applying some trigonometry and rearranging, we have:

\[
p(t) = \frac{I_M V_M}{2} \left( \cos(\theta) - \cos(2\omega t + \theta) \right)W
\]

where we note that \(\frac{I_M V_M}{2} = \frac{I_M}{\sqrt{2}} \frac{V_M}{\sqrt{2}} = I_{\text{RMS}}V_{\text{RMS}}\), and \(\cos(\theta)\) is a constant term and \(\cos(2\omega t + \theta)\) is a time-varying sinusoid, twice the frequency of the voltage or current, and with a phase offset.

If we look at the average power, we can use our expression that defines the average value, and we note that the average value of the time-varying sinusoidal portion of this expression is 0, and the average power is simply:

\[
\bar{p}(t) = I_{\text{RMS}}V_{\text{MS}} \cos(\theta)W
\]

where \(\theta\) is the angle (phase difference) between the voltage and the current and the angle of the complex impedance, \(Z\), that represents the load. We are usually interested in the average power delivered, so this expression is very useful. However, comparing this to the expression for power in the DC world \((P=IV)\), we note the modifying term \(\cos(\theta)\). This term is known in the AC world of “complex power” as the Power Factor (PF), and indicates the portion of the total Apparent Power, \(S\), (e.g. the magnitude of which \(=I_{\text{RMS}}V_{\text{RMS}}\)) that is actually doing some real work. As we will see shortly, the reactive elements in an AC circuit (inductors and capacitors) absorb some of this apparent power so it is not available for real work.
Let's look at the instantaneous power in the time domain with our three basic elements (R, L, C) and observe the effect of the Power Factor.

### 4.4.1 Resistor.
Voltage and current are in phase, so $\theta=0$, and:

$$p(t) = \frac{I_M V_M}{2} \left( \cos(0) - \cos(2\omega t + 0) \right)$$

and the average power is:

$$\bar{p}(t) = I_{RMS} V_{RMS} \cos(0) = I_{RMS} V_{RMS}$$

which is the same as we would expect from the DC case. Looking at this graphically:

As our intuition would suggest, the power delivered in this resistive case is all positive, and is known as real or active power, $P$, and is measured in Watts ($W$).
4.4.2 Inductor
The voltage leads the current by $90^\circ$ (eLi), so:

$$p(t) = \frac{I_M V_M}{2} \left( \cos(90) - \cos(2\omega t + 90) \right)$$

and our average power is:

$$\overline{p(t)} = I_{RMS} V_{RMS} \cos(90) = 0W$$

We note that we still have a voltage and current present, but we have no real power. A look at this situation graphically may help us understand:

One way to interpret this is to consider that during the positive half-cycle of the instantaneous power, energy is being stored in the magnetic field of the inductive element, and during the negative half-cycle, the energy is being returned to the circuit as the field collapses.

Although no real work is being done (assuming an ideal inductor), that current may have to flow through resistance in other parts of the circuit, which will dissipate real power.
and require real energy input that will be wasted. Minimizing this type of loss is a subject we’ll be looking shortly.

### 4.4.3 Capacitor.

The capacitive case is very similar to the inductive case, except that the phase shift between the voltage and current is opposite that of the inductor, and the resulting instantaneous power is inverted (180°) from the inductive case. The instantaneous and average power are:

\[
p(t) = \frac{I_M V_M}{2} (\cos(-90°) - \cos(2\omega t - 90°))
\]

\[
\overline{p(t)} = I_{RMS} V_{RMS} \cos(-90°) = 0W
\]

Again, taking a graphical look at this:

![Figure 4.12: Instantaneous Power in a Capacitive Load](image)

In both the purely inductive and purely capacitive case, we can see that there is energy transfer occurring, but no “real” power being used. This component of the overall power “consumed” by the reactive elements (C, L) is known as reactive power, \(Q\), and is measured in volt-amperes-reactive or \(\text{VARs}\).

To summarize the “types” of power in an AC circuit, we have:
• **Apparent Power, S=VI**: (Complex Power, \( S = \text{phasor voltage, } V \times \text{the complex conjugate of the phasor current, } I^* \)) what would *apparently* be the power if we simply multiplied the magnitude of the voltage times the magnitude of the current, This is measured in **volt-amperes** (VAs). \( S \) is a complex-valued number representing a vector in a two-dimensional complex plane. Together with phasor voltages and currents and complex impedance, \( Z \), it completes the set of complex-valued “transforms” that simplify analysis and calculations in the AC environment.

• **Real or active power, \( P \)**: (\( \theta = 0^0; \) P.F. = 1 or “unity”), measured in **watts** (\( W \)). This is the real, scalar component of the complex power, \( S \).

• **Reactive power, \( Q \)**: (\( \theta = \pm 90^0; \) P.F. = 0), measured in **volts-amperes-reactive** (VARs). This is the imaginary, scalar component of the complex power, \( S \). Reactive power is, by convention, considered *positive* for inductive loads, and *negative* for capacitive loads. These are *subtractive* in an AC circuit, as inferred by the opposite polarity indicated on the instantaneous power (time) graph.

In keeping with the theme used in phasors and complex impedance, we can represent this complex power we find in AC circuits in a two-dimensional complex plane. The triangle formed by drawing the active power along the real axis and the reactive power along the imaginary axis is completed with the apparent power as the hypotenuse. The angle, of the Apparent Power in this Power Triangle is the *same as phase angle* between the voltage and current, and the *same angle* as the complex impedance! Figure 4.13 shows a Power Triangle, and a graphical representation of the various components of Complex AC Power. Table 4.5 gives a summary of the relationships of the various Complex Power components.
Recall that the **Power Factor** is the cosine of $\theta$, and when applied to the **Apparent Power**, gives the real (active) power in the circuit. Since the cosine of $\theta$ in these two quadrants is ambiguous (same for $\pm \theta$) the Power Factor must be qualified by indicating whether it's **lagging** ($1^{\text{st}}$ quadrant – inductive loads) or **leading** ($4^{\text{th}}$ quadrant – capacitive loads).

(E.g.: if the load impedance, $Z$, is $100 \angle 45^\circ \Omega$, the P.F. = 0.707 lagging)

P and Q are scalar quantities that follow the DC rules for calculation, and can be represented on orthogonal axes in the Complex Power plane. The Apparent Power, S, is a vector quantity in the complex plane, and can be manipulated as such.

**Error! Reference source not found.** gives a summary of the rules and relationships governing the behaviour of $P$, $Q$, and $S$.

### Table 4.5: Complex Power Relationships

- $|P| = \frac{V^2}{R} = I^2 R$
- $|Q| = \frac{V^2}{X} = I^2 X$
- $|S| = \sqrt{P^2 + Q^2}$
- $P_{\text{Total}} = \sum P_i$
- $Q_{\text{Total}} = \sum Q_i$
- $\bar{S}_{\text{Total}} = \sum \bar{S}_i$
- $\bar{S} = P + jQ_{\text{Ind}}, \text{ or } P - jQ_{\text{Cap}}$
- $\bar{S} = \sqrt{P^2 + Q^2} \angle \theta, \text{ where } \theta = \tan^{-1} \left( \frac{Q}{P} \right)$
- $\bar{S} = VI^*, \text{ where } I^* = \text{Complex Conjugate of } I \left( e.g., (1 \angle \theta)^* = 1 \angle -\theta \right)$
\[ Power Factor, F_p = \frac{|P|}{|S|} = \cos(\theta) \]

4.4.4 AC Maximum Power Transfer

As in the DC case, the ratio of the impedance (resistance) of the load compared to the impedance of the source will affect the magnitude and efficiency of the energy transfer. In the AC world, **Maximum Power Transfer** will take place when the impedance of the load is the complex conjugate of the source: \( Z_{\text{Load}} = Z^{*}_{\text{Source}} \).

A relative maximum occurs when \( R_{\text{Load}} = \sqrt{R_{\text{Source}}^2 + (X_{\text{Source}} - X_{\text{Load}})^2} \).

4.5 Power Factor Correction

As we have seen when we look at phasor diagrams of currents and voltages, the phase difference between two currents that are summing, for instance, can lead to a resultant that is less than the magnitudes of either individual current. For this reason, a load impedance, \( Z_L \), in an AC circuit that has a significant phase angle (i.e. between the voltage and the current) can have a higher current than what is needed to provide the real power required. This extra current must often be “pushed” through long transmission lines where losses occur. The power company has an interest in making all the current it “pushes” provide real work for the end customer. For this reason, we usually try to “correct” the Power Factor by actually adding additional current draw at a phase angle that “cancels” other reactive current and reduces the overall current requirement and thus, line losses.

**Procedure:**

*Note: the key to this process is remembering that the real power, \( P \), must remain the same.*

1. Determine the “power triangle (i.e. \( \Sigma P, \Sigma Q, \) and \( S \)) for the device or combined load. (Drawing the power triangle usually helps too!)
2. Determine the new apparent power, \( S_{\text{New}} \): using the improved power factor, \( F_{P\text{New}} \):
   \[ |S_{\text{New}}| = \frac{P}{F_{P\text{New}}} \]
3. Determine the new reactive power, \( Q_{\text{New}} \), the net \( Q \) that will be left after correction:
   \[ Q_{\text{New}} = \sqrt{|S_{\text{New}}|^2 - P^2} \]
4. Calculate the reactive power, \( \Delta Q \), that must be “added”:
   \[ \Delta Q = Q_{\text{New}} - Q_{\text{Old}} \]

A negative result indicated that “negative VARS” must be added (i.e. a capacitor); a positive result indicates “positive VARs” must be added (i.e. an inductor).
For a load fed from a voltage source, $V_S$, the impedance of the reactor is:

$$X_R = \frac{|V_S|^2}{\Delta Q} = \frac{1}{\omega C} = \omega L$$

The reactive element must be placed across the source in parallel with the load (so the voltage at the load will not be affected by the addition of the element). Since, for a purely reactive element, the apparent power, $S$, is equal in magnitude to the reactive power, $Q$, the rating of the element is $|\Delta Q|\text{VA} @ V_S\ \text{Volts}$ (e.g. 1200VA, 600VAC)

A few final points to keep in mind:
- Power ($P$ and $Q$) are additive, no matter where in a load they occur, and regardless of the voltage or physical arrangement,
- The real (active) power must always stay the same,
- The corrective element should always be attached where it will not affect the voltage applied to the load (this usually mean in parallel with the load, across the “mains”)
- The balancing of the inductive and capacitive parts of the load actually creates something close to a “resonant” circuit; a load where the inductive and capacitive parts of the impedance cancel each other (at the supply’s frequency) on the Z-Diagram, making the combined load “look like a purely resistive load.

An example will illustrate the process and the result.
Example #1
The loading of a factory on a 1000-V, 60-Hz system includes:
1. 20-kW heating and incandescent lighting (unity power factor)
2. 10-kW of induction motors (0.7 lagging power factor, \( \eta = 0.9 \))
3. 5-kW fluorescent lighting (0.85 lagging power factor)

a) Establish the power triangle for the total loading on the supply.
b) Determine the power-factor capacitor required to raise the power factor to unity.
c) Determine the change in supply current and power savings from the uncompensated to the compensated system.

\[
\begin{array}{|c|c|c|c|}
\hline
\text{Load} & \text{P(W)} & \text{Q(VAR)} & \text{S(VA)} \\
\hline
\text{Heating} & 20,000 & \text{0} & 10\sqrt{2}\text{kVA} = 14286 \\
\text{Motors*} & 10,000 & 10,202 & 10\sqrt{2}/.7 = 14286 \\
\text{Flourescent} & 5,000 & 3098.7 & 5\sqrt{2}/.85 = 5882 \\
\hline
\text{total} & 35,000 & 13,301 & 37,441 \\
\hline
\end{array}
\]

* could have also interpreted this as output power and then divided by \( \eta \) to get 11,111W – OK.

b) Need to add -13,301VARs (a capacitor) to bring net VARs to zero and the Power Factor to 1.

\[
X_C = \frac{V^2}{Q_C} = \frac{1000^2}{13,301} = 75.188\Omega; \omega = 377, C = 1/(\omega X_C) = 34.96\mu F
\]

c) \( S = VI^*; I_{\text{orig}} = (S/V)^* = \frac{(37441\angle20.81VA)}{(1000\angle0V)} = 37.441\angle20.81A \\
I_{\text{new}} = (S_{\text{new}}/V)^* = \frac{(35000\angle0VA)}{(1000\angle0V)} = 35.000\angle0A \\
\Delta I = (37.441\angle20.81A) - (35.000\angle0A) = 13.29\angle90A, \text{ which we can observe is simply the current through the added power correction capacitor (1000V/75.188\angle-90\Omega). However, we would normally be interested in how the reduction in current affects our real power losses in the system. Since the real power losses are mainly } I^2R \text{ losses, we would be interested in the magnitude of the current, i.e. } P = |I|^2R. \text{ In this case we would see a reduction in current of:} \\
(37.441 - 35.000) / 37.441 = 6.5\%, \text{ and a power savings of:} \\
(37.441^2 - 35^2) / 37.441^2 = 12.6\%
Example #2
(2004 Final)
A single-phase 6hp motor is connected to a 240VAC, 60Hz supply. On average, the motor is loaded to 75% of its rated power. Its efficiency is 81%, and the name plate rating states 6.9kVA at full load.

a) What is the Power Factor, \( F_P \), for the motor without any correction?

b) What element (C or L, in and value in F or H) would be required to improve the Power Factor to 0.9 lagging?

c) What is the minimum VA rating for this correction element?

---

a) The power factor for the motor can be determined from the data given:
\[ |S|_{\text{max}} = 6900 \text{VA}; \quad P_{\text{max}} = \frac{(6\text{hp} \times 746\text{W/hp})}{.81} = 5525.9\text{W} \]
\[ F_P = \frac{P}{S} = 0.80 \text{ lagging} \] (motors are generally inductive)

b) Since it is generally inductive, we would need to add capacitance to increase the power factor:
Input power, \( P \), at 75% load = .75(5525.9W) = 4144.4W
\[ Q_{\text{.8}} = \tan(\cos^{-1} .8)P = 3107.7\text{VAR}_{\text{ind}} \]
\[ S_j = 4144.4/.9 = 4605.4\text{VA}, \quad Q_j = \tan(\cos^{-1} .9)P = 2007.5\text{VAR}_{\text{ind}}. \]
\[ \Delta Q = Q_{\text{new}} - Q_{\text{old}} = 2007.5\text{VAR} - 3107.7\text{VAR} = -1100.2\text{VAR} \]
\[ X_C = \frac{V^2}{Q_C} = \frac{240^2}{1100\text{VAR}} = 52.35\Omega; \quad C = \frac{1}{\omega X_C} = 50.66\mu\text{F} \]

c) Minimum rating should be 240V, 1100VA.
4.6 Transformers

Recall from our discussion on inductors, that there was a two-way relationship between electricity and magnetism. More specifically, we deduced from Faraday’s law that a changing current in a coil of wire (with or without a ferromagnetic core) would “induce” a voltage in that same coil. The phenomenon was called induction, and the property of the coil was called inductance, more specifically self-inductance, because the voltage was induced in the same coil as the flux that produced it, and tended to counteract the causal agent (current). Since the induced voltage is a direct result of the changing magnetic flux, a voltage can be induced in any coil that is “linked” by that changing flux. This property is called mutual-inductance. While this phenomenon can be an unwanted byproduct in some circuit situations, this property is exploited in devices called transformers to a very useful end: it is used to transform voltage either up or down to suit different purposes. (They are also useful for isolation and impedance matching, which will be discussed later.)

A transformer is a device in which two coils are deliberately arranged so that the flux from one coil is maximally coupled to a second coil, usually on the same core. Assume for the moment that there is a flux established in one of two closely coupled coils, \( \Phi_p \), then the voltage induced in this coil, \( e_p \), is given by:

\[
e_p = N_p \frac{d\Phi_p}{dt} = L_p \frac{di_p}{dt}
\]

where the subscript ‘p’ means the primary of two coils. Similarly, the voltage induced in the second coil would be:

\[
e_s = N_s \frac{d\Phi_s}{dt}
\]

where ‘s’ indicated the secondary coil.

If we define a coefficient of coupling, \( k \), as \( k = \frac{\Phi_s}{\Phi_p} \), where \( \Phi_p \) is the originally induced flux and \( \Phi_s \) is the portion of the flux which couples the second coil. Clearly \( \Phi_s \) can never be greater than \( \Phi_p \), so \( k_{\text{max}} \leq 1 \). Modern transformers are “tightly coupled”, with \( k \approx 1 \). Most of the useful formulas describing transformer behaviour are developed here assuming that \( k = 1 \), but they could easily be modified to reflect the performance of a “loosely coupled” coil (\( k < 1 \)).
The first useful operational formula:

\[
e_p = N_p \frac{d\Phi}{dt}, \quad \text{and} \quad e_s = N_s \frac{d\Phi}{dt}
\]

and \( \frac{e_p}{e_s} = \frac{N_p}{N_s} \), where \( \frac{N_s}{N_p} \) is called the turns ratio, \( n \)

therefore, \( e_s = ne_p \) \hspace{1cm} (Equation 4.1)

Note that a turns ratio, \( n > 1 \) is called a step-up transformer, a turns ratio \( < 1 \) is called a step-down transformer, and an \( n = 1 \) is called an isolation transformer.

While the voltage depends on changing flux, the flux depends only on the current, \( I \) (Ampere’s Law for magnetic circuits).

The magnetomotive force, \( \mathcal{Z} \): \( \mathcal{Z} = NI = \Phi \mathcal{R} \) or \( \Phi = \frac{NI}{\mathcal{R}} \), where \( \mathcal{R} \) is the reluctance (opposition to flow of flux).

The primary flux: \( \Phi_p = \frac{N_p I_p}{\mathcal{R}_p} \), and the secondary flux must be: \( \Phi_s = \frac{N_s I_s}{\mathcal{R}_s} \).

Since we have assumed \( k = 1 \), \( \Phi_p = \Phi_s = \Phi \), and since they are wound on the same core, \( \mathcal{R}_p = \mathcal{R}_s \).

Eliminating the equal elements, \( \Phi, \mathcal{R} \), and equating the right sides:

\[
N_s I_s = N_p I_p, \quad \text{and} \quad I_s = I_p \frac{N_p}{N_s}. \hspace{1cm} \text{Recalling that} \quad \frac{N_s}{N_p} \quad \text{is the turns ratio}, \ n, \ \text{we have:}
\]

\[
I_s = \frac{I_p}{n} \hspace{1cm} \text{(Equation 4.2)}
\]

Equations (4.1) and (4.2) are the general transformation equations for an ideal transformer.
Ex: Consider a physical transformer with \( N_p = 100 \) and \( N_s = 200 \):

![Figure 4.14: A Simple Transformer](image)

If \( e_p = 10V \), \( i_p = 2A \), find \( e_s \) and \( i_s \).

**Ans:** the turns ratio, \( n \), is \( \frac{N_s}{N_p} = \frac{200}{100} = 2 \)

\[
e_s = e_p n = (10V)(2) = 20V
\]
\[
i_s = \frac{i_p}{2} = \frac{2A}{2} = 1A
\]

4.6.1 Impedance Matching

There are occasions when we would like a load impedance to “look” either higher or lower than is actually is. One application is to match a load, like a loud speaker, to a transmission line for optimal power transfer.

Again consider the transformer shown above, but now consider the load, \( Z_L(R_L) \). For the ime being, let us assume a purely resistive load (i.e. voltage and current in phase). If \( i_s \) and \( e_s \) are as determined above: 1A and 20V respectively, therefore \( v_L \) and \( i_L \) are equal to these values, the load resistance, \( R_L = (20V)/(1A) = 20\Omega \), by definition. If we want to determine what \( R_L \) “looks like” from the primary side, we can consider the input voltage and current: \( R_p = (10V)/(2A) = 5\Omega \).

Generally, \( R_p = \frac{e_p}{i_p} \), and since \( e_p = \frac{e_s}{n} \) and \( i_p = \frac{i_s}{n} \), we can substitute giving us:

\[
R_p = \frac{e_s}{i_s^2} \text{, and since } \frac{e_s}{i_s} = R_s \text{, we have:}
\]

\[
R_p = \frac{R_s}{n^2} \text{ or } R_s = n^2 R_p \tag{Equation 4.3}
\]

---

1 Fig 22.10, Boylestad’s Circuit Analysis, Second Canadian Edition
where again \( n \) is the turns ratio, \( \frac{N_s}{N_p} \). This relationship is also true for complex impedance, \( Z \), as well as resistance, \( R \).

**Example:**
Consider again the general configuration shown at right:

If the primary winding is 50t and the secondary is 200t (assume \( k =1 \)), the input voltage, \( e_p = 10V \) and the load resistor is 100\( \Omega \), find \( e_s, i_s \) and \( i_p \).

**One way:** the turns ratio is 200/50 = 4, and \( R_p = \frac{R_s}{n^2} = \frac{100\Omega}{4^2} = 6.25\Omega \)

\[
\begin{align*}
    i_p &= \frac{e_p}{R_p} = \frac{10V}{6.25\Omega} = 1.6A \\
    i_s &= \frac{i_p}{n} = \frac{1.6A}{4} = 0.4A \\
    e_s &= ne_p = 4(10V) = 40V
\end{align*}
\]

**Another way:**

\[
\begin{align*}
    e_s &= ne_p = 4(10V) = 40V \\
    i_s &= \frac{e_s}{R_L} = \frac{40V}{100\Omega} = 0.4V \\
    i_p &= ni_s = 4(0.4A) = 1.6A
\end{align*}
\]
4.6.2 Schematic Representation:

In a schematic diagram, a transformer is usually indicated by the symbols below. The vertical lines indicate an iron or ferromagnetic-core.

![Transformer Symbols]

The “dot” generally means when *that* terminal on the primary coil is positive, the “dot” terminal on the *secondary* is also positive.
4.6.3 Two additional considerations...

4.6.3.1 Excitation current:

The voltage induced in the transformer primary is a reaction to oppose some changing flux. For any flux to exist implies some current must be flowing(!) even if there is no circuit attached to the secondary to conduct any corresponding current. The current that must flow to produce this “excitation” flux is called the excitation current. Any additional current that flows produces additional flux. This is usually in response to external circuits (as shown in the above examples) and serves to transfer power from a source, through the transformer to a load attached to the secondary.

4.6.3.2 Frequency Relation:

It can be shown that: 
\[ E_{p_{\text{max}}} = 2\pi f N_p \Phi_{\text{max}} \]
where \( f \) is the frequency, \( N \) is the number of turns and \( \Phi \) is the flux. Remember, to get an induced voltage the flux must be changing, and note that this voltage will be higher for a given flux if the frequency of the changing flux does up. Also note from our previous example that the input power (10V*1.6A = 16W) is transferred to the load (40V*.4A = 16W) without any loss (At least in an ideal transformer). If we want to increase the power transferred to the load by increasing the voltage, \( E \), we need to increase either \( N \), \( \Phi \) or \( f \). There are physical limits to the number of turns that can be wound on a core; the flux can only be increased so much before the core saturates and the amount of current required to increase the flux further is excessive. Increasing the frequency, when possible, is an effective way to transfer more power through a transformer without increasing its physical size while still retaining the performance advantages (voltage step-up, step-down or impedance matching). Alternatively, if we increase the frequency we could transfer the same power with a much smaller physical core.

Another way to consider this phenomenon, is that the rising flux in the primary stores a certain amount of energy in the magnetic field which is then transferred to the secondary coil and to the load. For a certain physical volume of core, there is a maximum flux that can be achieved before saturation, so there is only so much energy that can be transferred per “flux cycle”. Therefore, we can transfer more energy through the transformer by increasing the number of flux cycles per second (i.e. frequency).

Final Note:

While we have only considered a simple 2-winding iron-core transformer in this discussion, air-core and multi-winding cores are also in widespread use in many circuit applications.
4.7 Three Phase Power (3Φ Power)
\[ y = 100\sin^2(377x) + 100\sin^2(377x + 2.0944) + 100\sin^2(377x - 2.0944) \]
5  Electronic Devices and Integrated Circuits

This installment of EE204 Class Notes covers a few common electronic devices often used in electronic circuits, as well as two popular and commonly used integrated circuits. These are just a few examples of the many devices employed in electronic circuits used in control and instrumentation application, communications, entertainment and power generation and distribution. Also popular are “embedded microcontrollers”, small, dedicated computers used for a specific task or application. An introduction to embedded microcontrollers is a future topic in this course.

5.1  Single, Electronic Devices: Diodes and Transistors

We have already studied the basic electrical elements: resistors, capacitors and inductors. These are known as “passive” elements, because they only react to circuit voltages and currents that are imposed on them, and operate on basic electric and magnetic principles. There is a family of devices that operate based on the characteristics of special materials: semiconductors. These devices can either “react” to circuit conditions, somewhat like the basic elements, or “create” circuit conditions by their behavior. Semiconductors, as the name would suggest, are materials that are partially conductive. Their conductivity is often a function of how they are placed in a circuit, the specific conditions to which they are subjected, or controlling voltages or currents. Semiconductors are typically man-made by adding specific impurities (called “doping”) to a silicon-based crystal, or other special material, that alter its conductivity.

5.1.1  Diodes:
The first device we will consider is the semiconductor diode, commonly referred to simply as a diode. This device acts like a switch, and is conductive when its terminals are one polarity (forward biased), and not conductive when the polarity is reversed (reverse biased). Diodes are very commonly used to convert alternating current to direct current in power applications, and also in communications circuits. They are also used to control direction of current flow in other circuits for a variety of special purposes. The schematic symbol used to represent a diode is shown in Figure 5.1. This is a “two-terminal” device, with an anode at one end, and a cathode at the other.

![Figure 5.1: Diode Schematic Symbol](image)

When the anode is positive with respect to the cathode, the diode will conduct (switch on). When the polarity is reversed, it will not conduct (switch off). The physical shape and size of a diode can vary considerably, but anode and cathode terminals are usually indicated by either a contrasting band at the cathode end (as shown by the silver band on the 1N4007 diode shown in Figure 5.2), or with an imprinted diode schematic symbol on the case.
The Current - Voltage characteristics (I-V curve) of an ideal diode is shown in Figure 5.3.

Note that when the diode is forward biased (i.e. anode positive; cathode negative), current will flow (unrestricted in the “ideal” case, like a short circuit) to satisfy the surrounding circuit conditions. When the polarity is reversed (anode negative, cathode positive), no current will flow, regardless of the circuit conditions. Of course this is for an “ideal” diode. A practical diode, because of the material characteristics of the device, has an I-V characteristic that can be modelled in different ways, depending on the accuracy required. The most accurate, and most difficult to use, is an exponential model. Figure 5.4 illustrates an I-V curve using the exponential model. It very closely resembles a real diode I-V curve.
Most times, a simpler model can be used to represent a diode and calculate the necessary ratings and parameters. One very popular model is called the Piecewise-Linear Model. A graphical comparison of the I-V curve using this model, and the schematic equivalent are shown in Figure 5.5.

Figure 5.4: Diode Model - Exponential

\( V_F \) is the forward bias voltage required before the diode starts to conduct, and \( R_D \) is the internal resistance of the conducting material. \( R_D \) defines the slope of the conducting part of the Piecewise-Linear \((1/R_D)\).

One of the simplest models for a diode is the Constant Voltage model. It is exactly the same as the Piecewise Linear model, except that the internal resistance of the diode, \( R_D \), is assumed to be 0Ω. The I-V curve resulting from this last model is shown in Figure 5.6.

Figure 5.5: Diode Model - Linear-Piecewise
Figure 5.6: Diode Model - Constant Voltage

For this course, we will usually use the Piecewise-Linear or Constant Voltage model for a diode.
The key characteristics of a diode are summarized in Table 5.1.

Table 5.1: Key Diode Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward Voltage, $V_F$</td>
<td>The forward bias voltage required before the diode start to conduct significantly, or the typical voltage across the diode at a stated current level. Silicon power diodes are typically assumed to have a forward voltage of 0.7V, but other materials and types of diodes can have significantly different forward voltage levels.</td>
</tr>
<tr>
<td>Reverse Breakdown Voltage (reverse blocking voltage)</td>
<td>The maximum reverse bias voltage before the diode conducts in the reverse direction (will usually lead to diode destruction)</td>
</tr>
<tr>
<td>Average Forward Current</td>
<td>The average forward conduction current the diode can safely handle continuously without overheating and breaking down</td>
</tr>
<tr>
<td>Peak Forward Surge Current</td>
<td>The maximum forward current the diode can handle for a very short time without damage</td>
</tr>
<tr>
<td>Reverse Current</td>
<td>The maximum “leakage” current we can expect in the reverse direction when the diode is reverse biased at its maximum reverse blocking voltage</td>
</tr>
<tr>
<td>Diode Resistance, $R_D$, or Forward Characteristic</td>
<td>The I-V characteristic of the diode. In the Linear – Piecewise representations, this can be estimated by a resistance over the typical operating range.</td>
</tr>
<tr>
<td>Reverse Recovery Time</td>
<td>The time it takes for the diode to stop conducting once the bias is changed from forward to reverse.</td>
</tr>
</tbody>
</table>

Note that for some applications, other characteristics may be of concern. Also note that diodes are usually quite sensitive to temperature, and should be de-rated when operating, or expected to operate, at temperatures higher than normal.

**Some Types of Diodes**

One of the most common diode applications has been to rectify current (change from AC to DC) in power supplies (we will consider this application in more detail later in this course). Diodes for this application are usually made form a silicon-based semiconductor and referred to as power diodes.

Because of the bulk of material required to handle the heavy currents, power diodes cannot turn off fast enough for some high frequency detection applications. Other materials and manufacturing processes are used to satisfy the requirement.

The reverse breakdown voltage can be used to advantage, and some diodes, called Zener diodes, are deliberately manufactured with a specified reverse conduction voltage and built to operate in this region. Zener diodes can be used to provide a voltage reference or a simple regulated voltage.
Light Emitting Diodes (LEDs) have been used for years as indicators, but are now becoming increasingly use for illumination. The different materials and processes required to obtain different wavelengths and light output from diodes often means the forward voltage can be several volts instead of several tenths of volts.
Diode Example 1:
Determine i) the average current ii) peak current rating, iii) the reverse breakdown voltage, and iv) the power dissipated in the diode in the following application.

i) 120VAC is the RMS value. This means the peak value of this sinusoid is \( \sqrt{2} \times 120 = 169.7 \) V. The average of a half-wave rectified sine wave is

\[
\frac{A}{\pi} = 54.0 \text{V} \quad \text{and} \quad \overline{I} = \frac{54.0}{50} \Omega = 1.08 \text{A} 
\]

ii) the peak current will occur when the sinusoid reaches its peak voltage: 169.7 V, so \( I_{\text{peak}} = \frac{169.7}{50} \Omega = 3.39 \text{A} \)

iii) the maximum reverse voltage the diode will be subjected to is the peak voltage of the sinusoid when the diode is reverse biased, so \( V_R = 170 \text{V} \)

iv) To determine power, we should look at a model: piece-wise linear. This is a little trickier because the signal is only passed through the diode during half the cycle. Remember, the RMS voltage is used to determine power, so we can go back to the original 120VAC. The RMS current is just 120/50 = 2.4A. If we look at some single, power diodes, we find the Vishay ES07D has an average current capability of 1.2A, a peak current capability of 30A and a reverse voltage capability of 200V. While \( R_D \) is not given directly, we can estimate it from other spec’s. One spec’ shows a maximum forward voltage of 980mV @ 1A. If we make the reasonable assumption, using our model, that is approximately .7V at 0A, we can calculate

\[
R_D \cong \frac{\Delta V}{\Delta I} = \frac{0.98 - 0.7}{1} = 0.28 \Omega 
\]

\[
P = I^2 R = (2.4)^2 (0.28) = 1.61 \text{W} 
\]

again, if the diode was conducting full time. Since it’s conduction half the time, and we consider average power in AC circuits, the average power that the diode has to dissipate is 0.8W.

In practice, the power dissipation capability in a diode is usually implied in its current ratings.
Diode Example 2:
Determine the peak current rating and reverse voltage rating for the diode in the following circuit.

\[ V_R = 12\text{V}, \quad I_{\text{Peak}} = 0.6\text{A} \]

Diode Example 3:
A LED requires 10mA of current for proper illumination. If it is to be driven by the 5VDC output from a microcontroller, determine the value of R that will result in the proper current when the output pin goes “high” (i.e. 5V). The forward voltage for the LED is 1.9V

\[ 310\Omega \]
5.1.2 Transistors:

Transistors are also semi-conductor devices, built using similar materials and techniques as diodes. The chief difference is that while diodes are two terminal devices, transistors are generally three terminal devices. This third terminal is used to control the flow of current between the other two terminals. While this control can be used to achieve a proportional control of the current for amplified and comparator applications, it is also very commonly used to either turn the transistor fully on or fully off, thus acting as a “solid-state switch”.

Their use as solid-state switches is very popular in electro-mechanical control applications, and will be the sole application focus in this course. Some of the key advantages of solid-state switches are:

- No mechanical parts to wear (long life)
- Small size
- Speed
- Quiet operation
- Tolerance to wide range of operating conditions
- More efficient in many cases
- Easy interface to electronic control systems (e.g. microcontrollers)

Some of the disadvantages are:

- Susceptibility to electromagnetic and electric field noise
- Temperature sensitivity and heat dissipation
- Requires some electrical knowledge

As in most things, weighing the advantages against the disadvantages for a particular application will dictate the choice, but solid-state switching in becoming increasingly popular in an ever-increasing number of applications.

Of the many specialized types of transistors, we will focus on two: bipolar junction transistors (BJT), and field effect transistors (FET). While there are many characteristics of importance for specific applications, for switching we will focus on the two most important: those that ensure it will survive in a particular environment (e.g. voltage, current and power ratings), and those that describe its performance (e.g. what is required to “turn the switch on and off”, speed of operation, saturation voltage drop etc.).

**Bipolar junction transistors:**

“The invention of the BJT in 1948 at the Bell Laboratories ushered in the era of solid-state circuits, which led to electronics changing the way we work, play, and indeed, live.”

Invalid source specified.

BJTs have three terminals called the base, emitter and collector. The schematic symbols for BJTs are shown in . Whereas diodes (two terminal devices) where constructed of two layers of semiconductor: a “P-doped” layer and an “N-doped” layer, BJ transistors have three alternating layers: either N-P-N or P-N-P. This means there are two major types of BJTs called “NPN” and “PNP”, named after the order of their “layers”. Their operation differs in only one important aspect. To understand this, let’s fist look at how they operate.
In a diode, a “forward bias” across the P-N junction allowed a current to flow as long as the bias voltage was greater than the forward voltage required for that particular diode.

In a bipolar junction transistor, if we forward bias the base-emitter junction, a current will also flow, like in a diode, from the P to the N material. When this base-emitter current flows, it allows a much larger current to flow between the collector and the emitter, provided there is a source of current and the necessary voltage. These are usually provided by the surrounding circuit.

Now look at the two “layer” depictions for an NPN and PNP transistor shown in Figure 5.7.

![Figure 5.7: NPN and PNP BJT Structure](Invalid source specified.)

In the NPN case, a positive bias across the base-emitter junction will allow a positive current to flow from the collector to the emitter. In the PNP case, a negative bias across the base-emitter junction will cause a negative current to flow from the collector to the emitter. Thus, an PNP circuit will operate in a circuit exactly the same as an NPN transistor if the supply polarity is reversed! The schematic symbols for the two types are shown in Figure 5.8

![Figure 5.8: BJT Schematic Symbols](Invalid source specified.)

The “arrow” in the symbol can be thought of as a diode which, when forward biased, will allow current to flow. The advantage of a BJT transistor is that that a very small current flowing through the base-emitter “diode” will allow a very large current through the collector-emitter path. Thus a BJT is a current gain device. Like a diode, when the transistor is conducting from collector to emitter, it is like a closed switch. When it isn’t conducting, it’s like an open switch. The difference is, we can now control when the switch is open or closed by forward biasing the base-emitter junction and “injecting” a little current.

How much current can flow through this “switch” is a function of how much current the device can handle, and how much is injected into the base. The ratio of the collector current, \( i_C \), to the base current, \( i_B \), is known as the common emitter current gain, \( \beta \). For
typical BJTs, $\beta$ ranges from the 10s to the 100s and even up to 1000s in some cases, 50 – 200 being a typical range for medium power transistors. $\beta$ is usually referred to as the hybrid forward common-emitter current gain, $h_{fe}$.

BJTs come in a variety of industry standard cases, depending on their application, current and heat dissipation requirements. Transistors which have a large current carrying capacity are usually in physically larger packages, usually with a metal tab or back that can attached to an external heat sink to absorb and dissipate the heat produced by the current heating in the transistor junctions. Several examples are shown in Figure 5.9.

![Figure 5.9: Typical Transistor Packages](Invalid source specified.)

Some of the key characteristics of bipolar junction transistors are show in Table 5.2 along with a brief explanation of their meaning.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_{fe}$ or $\beta$</td>
<td>The current gain of the device $= \frac{i_C}{i_B}$</td>
</tr>
<tr>
<td>Collector-base voltage, $V_{CB}$</td>
<td>Maximum voltage (reverse bias) between these terminals before the device will breakdown</td>
</tr>
<tr>
<td>Collector-emitter voltage, $V_{CE}$</td>
<td>Maximum voltage allowed from collector to emitter</td>
</tr>
<tr>
<td>Emitter-base voltage, $V_{EB}$</td>
<td>Maximum reverse bias</td>
</tr>
<tr>
<td>Collector current, $I_C$</td>
<td>Maximum collector current before device damage</td>
</tr>
<tr>
<td>Base current, $I_B$</td>
<td>Maximum bas current before device damage</td>
</tr>
<tr>
<td>Dissipation</td>
<td>Maximum power dissipation. This is often given at a set ambient temperature (e.g. 25$^\circ$C) with no heat sinking, and also with the case kept at a set temperature (i.e. with an external heat sink)</td>
</tr>
<tr>
<td>Switching times</td>
<td>The time it takes the transistor to turn on and turn off after the forward bias, $V_{BE}$, changes</td>
</tr>
</tbody>
</table>
Given that the BJT is a current device, we must determine the required current that must be “switched” in a circuit to determine the transistor’s specifications. An example will illustrate this.
**BJT Example 1.1:**

In the following application, an NPN transistor switch is being used to control a high-current relay. Determine the voltage and current specifications for the transistor, and the value and ratings for the base and collector resistors, if required. The 12VDC AZ8 relay winding has a DC resistance of $80\,\Omega$ and has a “pull-in current” of $100\,mA$. The controller has a 5VDC high signal on the TS pin to control the switch.

First, find the current requirement to close the relay:

\[ I_{\text{Relay}} = \frac{12\,V}{80\,\Omega} = 150\,mA \]

With a 12V battery, this should be the maximum collector current.

If we start by assuming an $h_{fe}$ of 50, this means we need a minimum base current of:

\[ I_B = \frac{I_C}{h_{fe}} = \frac{150\,mA}{50} = 3\,mA \]

Based on these preliminary current calculations, we can look for a suitable transistor. The 2N3904 (as shown) is a reasonable start, as it has a $I_{C\text{-max}}$ of 200mA, and a minimum $h_{fe}$ of 30 at $I_C=100\,mA$.

Recalculating our $I_B$ with an $h_{fe}$ of 30 gives an $I_B$ of 5mA. $I_{B\text{-max}}$ for the 2N3904 is spec’d at 50mA, so providing our controller can output this much, this could work.

The 2 numbers given for the collector-emitter saturation voltage, $V_{CE\text{(Sat)}}$, are .2V @ 10mA and .3V @ 50mA collector current. This implies a bulk resistance of approximately:

\[ \Delta V/\Delta I = \frac{.1\,V}{.04\,A} = 2.5\,\Omega \]

Extrapolating the $V_{CE\text{(Sat)}}$ for 150mA gives us:

\[ V_{CE\text{(Sat)}} = .3 + R\Delta I = .3 + (2.5\,\Omega)(.1\,A) = .55\,V \]

We can now calculate the collector resistor to give us 150mA through the relay coil:

\[ R_C + R_{\text{Relay}} = \frac{12\,V - .55\,V}{150\,mA} = 76.3\,\Omega \]

Since the relay coil resistance is $80\,\Omega$, we don’t need a collector resistor at all. The coil current with the $80\,\Omega$ is still ~143mA (>100mA pull in), so the relay will still operate fine.
Since the output of our controller is 5V, we can calculate $R_B$ based on a $I_B$ of 5mA. The $V_{BE}$ forward voltage given in the spec sheet is 0.95V at a base current of 5mA and collector current of 50mA. We’ll assume 1V to be a little conservative:

$$R_B = \frac{5V - 1V}{5mA} = 800\Omega \ .$$

We would take the next lower standard value of 750\Omega, giving a slightly increased base current and slightly higher emitter current (but not a significant effect on the $V_{CE(sat)}$ or power).

Finally, we compare voltages and power in this circuit to the maximum spec’d for the 3904.

- $V_{EB}$ should never be reverse biased in this configuration, and the spec is 6V
- $V_{CB}$ and $V_{CE}$ should never exceed the battery voltage, 12VDC + $V_D$, the forward voltage of the clamping diode (assume ~1V max), so 13V. The spec’s for these are 60V and 40V respectively.
- The collector power dissipation is somewhere between $(150mA)^2(2.5\Omega) = 56mW$ and $(150mA)(.55V) = 83mW$, both well below the 625mW spec’d for room temperature operation.
BJT Example 1.2:
In the following application, a PNP transistor switch is being used to control a high-current relay. Determine the voltage and current specifications for the transistor, and the value and ratings for the base and collector resistors, if required. The 12VDC AZ8 relay winding has a DC resistance of 80Ω. The controller has a tri-state output, TS, that can be made “active low”, and a floating high impedance otherwise.

In this variation of the control circuit, we show a 2N3906, which is a PNP switching transistor with virtually the same spec’s as the NPN 2N3904 except with reverse polarity. All calculations will be the same (no RC, voltages, currents and power spec’s all OK) except for the base drive circuit, which has some special considerations.

In this case, we need to bring the base negative with respect to the emitter to turn the transistor on. The way to do this is to have the TS pin go low (i.e. 0V) when you want to turn the fan on. This is easy, the problem is turning it off when you want it to stop!. If we allow the TS pin to go high (5V w.r.t. ground), the base-emitter junction is still very forward biased (12v – 5V = 7VDC!). We could use another smaller NPN transistor as a trigger, but in this case (assumed) our controller has the capability of making that output pin “active low” (i.e. it will be grounded, 0V, when it is “low”), and have it go into a high impedance “tri-state” (~∞Ω) when it is not. Then all we need to do is make sure we can turn the transistor off when we don’t want it on. We do this by placing a bleed resistor, R3, between the base and collector so that when it is not pulled to ground through RB, it will pull the voltage down to the same as the collector (+12V) and the transistor will stay off.

We need only calculate the values for RB and R3 so that we get the requisite 5mA base current when the TS pin is grounded and VCE is greater that the required VBE (~1V). If we select a 1KΩ value for R3, it will draw \[ \frac{IV}{1k\Omega} = 1mA \] when the transistor is on. This, combined with 5mA current, means: \[ R_b = \frac{12V - 1V}{6mA} = 1833\Omega \]. The next lower standard value (5%) is 1.8KΩ; a very close fit.

Modeling
The main element of a typical BJT model is a simple current sink (resistor) as an input, and a dependent current source for an output. Modeling of this sort is rarely required in switching applications and will not be addressed any further at this time.
5.1.3 Field Effect Transistors:

Another very useful type of three-terminal semiconductor device which can be used in switching applications is the filed effect transistor (FET). Its basic operation is very much like the bipolar junction transistor in that is has one terminal that can be used to control current flow between the other two terminals. The main difference is that while the BJT is a current controlled device, the FET is a voltage controlled device. For a FET, the conducting channel is controlled by the field created by applying a voltage to one of the terminals, thus the “field effect” name. The techniques used to fabricate FETs results in a device with an extremely high input impedance (~10^{15} \Omega), and an very low “on” resistance (~100m\Omega) and thus a very low “saturation” voltage. This has made the FET a popular choice for switching applications as well as established this technique as the mainstay for modern integrated circuit technology.

Figure 5.10 shows a typical physical structure for these devices. For “enhanced-mode” MOSFETs (metal-oxide semiconductor FET), the “channel” labeled in the figure is induced in the P-type substrate by a field created by the voltage between the gate terminal and the substrate electrode. (Another type of FET is the “depletion mode”, where the channel exists with no bias applied, and is “pinched off” by a reverse (negative) bias between the gate and source.) The substrate electrode is usually tied to the source terminal internally. When biased appropriately, current will flow from the drain to the source, provided the proper circuit conditions are in place.

![MOSFET Construction](image)

**Figure 5.10: MOSFET Construction** Invalid source specified.

This figure is for an “N-Channel” FET. They also come in the “P-Channel” type where, as in the BJT case, all polarities are just reversed.

Notice in Figure 5.11 that a variety of schematic symbols can be used, depending on the particular type of FET. In all cases, however, the arrow is pointing in for the N type and out for the P type; the reverse of the BJT. The BJT symbol is shown for comparison. Also note the similarity between: base and gate, source and emitter, and collector and drain. These terminals perform similar functions on each device. There are other variations of the symbol used, but all show similar elements of the base, source and drain. In addition, some show other elements added in the package, like a zener diode to provide device protection under certain circuit conditions.
The key parameters of FET devices are, again, similar to those of BJTs: current capacity and voltage breakdown limits. Table 5.3 summarizes the important characteristics for FETs used in a switching application.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_D$, Continuous drain current</td>
<td>As stated, continuous current capability of source-drain path</td>
</tr>
<tr>
<td>$I_{D\text{Max}}$, Peak drain current</td>
<td>Pulsed or short term current capability. Important because many loads (e.g. motors) draw much more current to start than when running.</td>
</tr>
<tr>
<td>$V_{GS}$, gate source max voltage</td>
<td>Determined from circuit characteristics. Rating should be high enough so it won’t be over-voltage in operation.</td>
</tr>
<tr>
<td>$V_{DS}$, max drain-source voltage</td>
<td>As above. Should also consider transient voltages that might occur, especially when switching inductive loads.</td>
</tr>
<tr>
<td>$R_{DS}$, static, drain-to-source on-resistance</td>
<td>Use to determine “on” saturation voltage. (Also related to power dissipation.)</td>
</tr>
<tr>
<td>$V_{GS @ I_D}$, gate – source voltage for desired drain current</td>
<td>Usually not given as a single number, but often graphed. Ensure that drive voltage is sufficient for switched current required for load.</td>
</tr>
<tr>
<td>$P_D @ T$, case power dissipation at a given temperature</td>
<td>As with BJTs, capability is usually de-rated with higher operating temperatures. Switching on/off times will affect power dissipation. Driver source impedance will</td>
</tr>
</tbody>
</table>
affect on/off times due to gate capacitance.
May be critical for high-speed applications.
Also has implications for power dissipation.

Consider some switching application examples.
FET Example 1:

An electronic controller is used to control a 12VDC motor which is driving a fan. The motor draws 3A when running at its rated 2200rpm, and draws approximately 10 times as much (30A) when stalled or during start-up. Determine a suitable switching element to control the motor given the controller will output a 5V high signal with a rise time of 0.1 \( \mu \)s.

Figure 5.12: FET Example 1

Some basic analysis about voltages and currents:
1. When Q1 is off, the maximum voltage across the FET (\( V_{DS} \)) is 12VDC.
2. When the switch is on, the only voltage across Q1 (D-S) is the saturation voltage (we’ll check this later, once we have a candidate FET).
3. Since there is a clamping diode across the motor, the maximum voltage when the switch first turns off will be \(~12.7V\)
4. If the motor is on continuously, the draw, as given, is 3A.
5. Start-up current is never more than the stall current (0 rpm; \( E_s/R_{motor} \)): 30A
6. Since we only have 5V to turn the FET on enough, we’d like to find a FET with a \( I_D \sim 30A \) when \( V_{GS} \) is 5V.

A FET like the IRL2703 has a continuous current of 24A at 25\(^0\)C, and a peak current of 96A, more than adequate. Although the peak is tested with a 300\(\mu\)s pulse, it also notes (as is usually the case) that the peak rating is limited by temperature, so 30A during startup should be OK – we can check heating later.

A look at the \( I_D \) vs \( V_{GS} \) curve shows an \( I_D \) of \(~30A\) at 25\(^0\)C and about 24A @ 175\(^0\)C. Again sufficient.

\( V_{DS} \) is given as 30V, so as long as we control the flyback voltage from the motor, we’re below.

\( V_{GS} \) is given as \( \pm 16V \). We are driving with only 5V, and there is no other circuit reason why it should get any higher.

Other notes:
R_G is chosen to limit the draw from the source should something go wrong with the FET and yet keep the impedance low so the FET will turn on faster (charge up the gate C faster).

R_2 is added as good design practice to make sure the FET will stay off should the source be interrupted by drawing the gate voltage to ground.

We’ve added a 78L05 linear voltage regulator to our circuit so we can run the controller from the same source as the motor. It provides a regulated 5VDC output with an input from anything >~6V. The Cs are added to provide a little source filtering, and R_3 is added to dissipate some of the (wasted) power at the controllers maximum current draw of 20mA ((12V-6V)/0.02A = 3000Ω).
FET Example 2:

Variation on Example 1, but using a P-channel FET.

As in the BJT PNP case, we have the problem of a negative grounded controller and a “positive grounded” switch. In this case, we chose to resolve the problem by using an NPN “Darlington Pair” transistor (basically 2 NPNs connected to each other in series to provide a higher current gain). It is part of an IC package that has 7 Dairingtons with integrated base resistors and common emitters (as well as integrated shunt diodes on each transistor). We need only one of the 7 for this switch.

The current and voltage requirements are identical to the first example, with all the required spec’s being the same but occasionally reverse polarity. (Actually, the basic structure of FETs is such that they can almost be used in either polarity, but once you commit to connect the substrate to one of the terminals, it becomes dedicated for use with only one polarity.)

Also, the rationale for the gate resistor and the associated pull-down resistor are the same.
5.2 Operational Amplifiers

There is often a requirement in instrumentation, audio, communications and related application to increase the strength or magnitude of a signal. This is usually done to increase the power in the signal so that it can be used to drive other circuitry or devices. Active gain devices like vacuum tubes, bipolar and FET transistors have been used in the design of amplifiers for many years. The era of integrated circuits provided new opportunities to create an amplifier in a small package with a number of desirable amplifier characteristics. While the family of amplifiers known as “operation amplifiers” can be, and were, built using discrete components, the integrated circuit “Op Amp” provides a very versatile and convenient “gain block” that can be used in a variety of applications.

They operate by following one general rule: **the output will do whatever it can to keep the difference between the input terminals at 0V.** If the + terminal is at a higher voltage than the – terminal, the output will increase in a positive direction until the situation is corrected or the output reaches the positive maximum. If the + terminal is lower than the – terminal, the output will move in a negative direction until the difference is reduced to 0V or it reaches the negative voltage limit. The reason for the “inverting” and “non-inverting” names will become evident as we consider typical circuit configurations later in this section.

Consider some of the key characteristics of an “ideal” amplifier:

- high gain (can multiply the power of a very weak signal to many times its input value)
- wide bandwidth (can amplify very low frequency and very high frequency signals equally well)
- high input impedance (puts very little load on the driving circuit, thereby not affecting its operation)
- low output impedance (is not affected by a heavy load, and can maintain its output level regardless)
- low distortion (provides consistent, linear amplification at all levels of input and output)

In addition, there are other related advantages like high efficiency, a high voltage output range, small size, low cost etc. A simple model is shown in Figure 5.14.

While not “ideal” by strict standards, modern Op Amps come suitably close such that they are very widely used, and in many circuit applications are close enough to ideal that they can be assumed to be so for design purposes. The TL082 op amp you will deal with in the lab, for instance, has an open loop (maximum) gain >200,000, can amplify frequencies from DC to ~3MHz, and has a distortion in the audible range of < .006%! Note that its ability to amplify DC is very useful in control applications where it can be used to compare and amplify (multiply)
static, DC levels. It has FET transistors as input stages, and as a result, an input resistance of $\sim 10^{12}\Omega$. It can handle a short circuit output indefinitely (this is really self-protection, as clearly there will be no useful voltage output with a 0\(\Omega\) load!). Finally, the TL082 is very cost effective, quantity priced at <$.16, for a case that contains 2 separate op amps!

The modern Op Amp is also very versatile, and can be configure to provide gain or comparison in a wide variety of applications. As with most solid-state devices, op amp ICs are designed and built for different voltage and current ratings for different situations. While there are quite a number of different configurations and applications, we will focus on only three: the inverting amplifier, the non-inverting amplifier, and the voltage follower (buffer) amplifier. Table 5.4 gives a summary of the key characteristics.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum supply voltage</td>
<td>Op Amps are usually supplied by two sources of opposite polarity. A ±18V rating for instance means a maximum $V_{CC+}$ to $V_{CC-}$ rating of 36V.</td>
</tr>
<tr>
<td>Maximum output voltage swing</td>
<td>The maximum output voltage limits. This is a capability restriction, not an exposure limit. It would ideally be the same as the supply, but in practice it is less because of the internal “circuit overhead”.</td>
</tr>
<tr>
<td>Maximum differential input voltage</td>
<td>The maximum voltage difference the device can withstand between its two input terminals.</td>
</tr>
<tr>
<td>Unity-gain bandwidth [product]</td>
<td>The product of the voltage gain, $A_v$, and the frequency at which the gain starts to drop from its maximum design value. E.g. an op amp with a 3Mhz unity-gain bandwidth configured for a voltage gain of 3 would start falling below 3 after the frequency reached 1MHz.</td>
</tr>
<tr>
<td>Input resistance</td>
<td>What the input of the amplifier would look like as a resistor (with no other circuitry attached to that input). It reflects how much it would load down the source driving it.</td>
</tr>
<tr>
<td>Slew rate</td>
<td>The maximum rate at which the voltage at the output will change. This is related to the frequency characteristics, but is really a limit on the frequency/voltage swing combination.</td>
</tr>
</tbody>
</table>
5.2.1 Configurations:

Three common configurations will be discussed in more detail below. In addition to power supply connections, op amps typically have a minimum of 3 other terminals: a non-inverting input (commonly labelled ‘+’), an inverting input (commonly labelled ‘-‘) and an output terminal. Op amps are a very versatile “gain block”, that can be moulded to perform as required by arranging the appropriate circuitry around them (connecting other components to their terminals).

5.2.1.1 Inverting Amplifier Configuration:

The circuit shown in Figure 5.15 is know as “inverting” because an increase in voltage at the input terminal (Input+) will cause an inverse change (i.e. decrease) at the output. The non-inverting input is connected to ground in this case (the same point at Input-). This ties the + terminal to 0V and, if the supplies are symmetrical (i.e. same magnitude), the output can vary an equal amount in either direction to try to keep the – terminal at 0V as well. If R2 was not there, any small difference impressed at the – terminal would be amplified by the full “open loop” gain of the amp (>200,000), which may not be what is desired. If we “close” the loop by adding a feedback resistor (R2) between the output and the input, the output terminal can influence the voltage at the – input terminal (e.g. through the voltage divider formed by R2 and R1). Applying the voltage divider rule to obtain the voltage at the – terminal, V2, we have:

\[ V_2 = V_{in} + \left( V_{out} - V_{in} \right) \frac{R_1}{R_1 + R_2} \]

Recalling that V2 must equal 0V, the same as V1 and multiplying through by (R1+R2) we have:

\[ V_{in} R_1 + V_{in} R_2 + V_{out} R_1 - V_{in} R_1 = 0 \]

Finally, rearranging, we have an expression for the voltage gain for this configuration:

\[ A_v = \frac{V_{out}}{V_{in}} = -\left( \frac{R_2}{R_1} \right) \quad \text{Equation 5.1} \]

where the negative sign indicates the 180\(^0\) phase shift and inverting nature of the amplifier circuit.

![Figure 5.15: Inverting Op Amp](image)
Since the inverting input is maintained at 0V (or the same voltage as the + input) by the feedback from the output, it is considered a “virtual ground”. This means it “appears” that the input resistor, $R_1$, is “grounded, which in turn means the input resistance (impedance) is simply $R_1$.

The output resistance (impedance) is $R_2$ in parallel with the output resistance of the amp itself. The output is modelled as a voltage source with an output voltage $= A_V(\Delta V_{in})$. An ideal source has an internal resistance of 0Ω (i.e. Thévenin equivalent). While not ideal, modern op amps are very close, and an assumption of a 0Ω output impedance is a good approximation.

### 5.2.1.2 Non-inverting Configuration:

The second popular configuration is wired as shown in Figure 5.16. You will note that the feedback must still be connected to the inverting (-) input. This is so “negative” feedback is provided, which leads to a stable, balanced result. (“positive” feedback would cause “run away”, never reaching a stable voltage level.)

![Figure 5.16: Non-inverting Op Amp](image)

In this case, instead of connecting the non-inverting input to a fixed voltage (e.g. 0V in the inverting circuit in Figure 5.15), it becomes the input terminal. When the input voltage varies, the output changes to move the other input so the difference is still 0V (although the voltage on both inputs may be any value within the allowable input voltage limits). In this case, $V_2 (-)$ must be equal to $V_{in}(+)$, and $V_2$ is the result of the voltage divider between the output voltage and ground:

$$V_2 = V_{in} = \left(\frac{V_{out}}{R_2}\right) \frac{R_1}{R_1 + R_2}, \text{ re-arranging:}$$

$$A_V = \frac{V_{out}}{V_{in}} = \left(\frac{R_1 + R_2}{R_1}\right) = \left(1 + \frac{R_2}{R_1}\right) \quad \text{Equation 5.2}$$

The input impedance is simply the input resistance of the op amp, usually very high. If we require something less, we can simple put a resistor across the input (input+ to ground). The output resistance, as in the inverting configuration, is ~0Ω.
5.2.1.3 Voltage Follower Configuration:

The final configuration we will consider is shown in Figure 5.17, and is known as the voltage follower, or buffer amplifier. It is used to provide impedance isolation, or a “buffer”, between a source with a high output impedance and a load with a low input impedance. It is a simple extension of the non-inverting configuration with $R_1 = \infty$, and $R_2 = 0 \Omega$. Using the voltage gain expression for the non-inverting case and substituting $\infty$ and $0 \Omega$ for $R_2$ and $R_1$ respectively, we find, as expected, the voltage gain for the voltage follower, $A_V = 1$.

![Figure 5.17: Voltage Follower Op Amp](image)

Simple examination of the circuit also shows that the input impedance is still that of the op amp, and the output impedance is still $\sim 0 \Omega$ (the very high input resistance of input 2 (-) in parallel with $R_3$ and the very low Thévenin impedance of the output).

5.2.2 Frequency Response:

While a truly ideal amplifier would amplify any frequency at any level, practical op amps are not quite that capable. They have a very high gain over a narrow band of frequencies, and a lower gain over a broader band of frequencies. The relationship between the gain and the “bandwidth” (the range of frequencies over which the gain remains constant) is monotonically inverse; that is as the bandwidth goes up, the gain goes down. This is usually expressed as the “gain-bandwidth product”. The figure of merit that is usually given in a specification sheet is often called the “Unity-gain Bandwidth” or “Unity Gain Product”. While not a perfect correlation of the whole frequency band, it is a linear, inverse relationship over a good part of it. The “unity-gain bandwidth” is usually supported by a graphical description as shown in Figure 5.18. Notice that the gain is flat up to about 20 Hz (also note the log-log scale), and then “rolls off” after that until it reaches a gain of 1 (unity) at 3MHz (thus the 3 MHz Unity-gain specification). This specification would predict a gain of 10 at a bandwidth of 300 kHz (log(300) = 2.48, which corresponds very well to the graph at a gain of 10).

You will also note the phase shift shown on this graph. This is the result of some reactive phase effects as well as some latency (delay) through the op amp. Since the amplifier basically responds to voltage, any phase shift essentially appears as a delay. If the delay is enough to cause a $180^\circ$ phase shift, the feedback will no longer be negative, but will be in phase with the input thus providing positive feedback, which could lead to instability and oscillation. It is important for stability that an op amp does not feed back in phase if the gain is $> 1$. Most op amps, including the TL082, have a latency or phase
shift that is small enough so they will not feed back in phase until the gain is \(< 1\) which avoids oscillation or other instability.

5.2.3 Slew rate limitations:

The “slew rate” is the maximum rate at which the output voltage can change, and is usually given in units of Volts per \(\mu\)sec (V/\(\mu\)s). This parameter limits the maximum output voltage at a given frequency. The slew rate is basically the slope of the output waveform. For a general sine wave: \(A \sin \omega t\), the slope is \(d(A \sin \omega t) = \omega A \cos \omega t\). The maximum slope is when \(\cos \omega t = 1\), and is \(\omega A\). Therefore the product of the frequency and the amplitude cannot exceed the slew rate of the amp or severe distortion will occur. For example, for the TL082 the minimum slew rate is specified as 8 V/\(\mu\)s. If the top frequency that was to be amplified was 800 kHz, the peak voltage allowed would be:

\[
A_{\text{peak}} = \frac{8 \text{V}}{2\pi \left(800 \text{kHz}\right)} = 1.59 \text{V},
\]

which may limit the gain that can be expected without distortion.

5.2.4 Power dissipation:

For small signal applications, power dissipation is rarely a problem. The TL082, for instance, has a dissipation capability of 490 mW @ 85\(^\circ\)C. The maximum supply current is 2.8 mA, and even at the maximum supply voltage of \(\pm 18\) V, this is only \(~ 100\) mW if all the power is dissipated in the case. In many high power applications, the op amp is used to drive external components that will handle most of the power and can be sized accordingly. Power op amps do exist, and power handling capability should be considered in those cases. A reasonable estimate is to determine the maximum power consumed by the op amp less the power delivered to the load. This difference must be
dissipated by the op amp itself. Power op amps usually have packages that will facilitate connection to thermal sinks (heat sinks) to help dissipate any excess power.

### 5.2.5 Examples:

**Example 5-1: Simple Op Amp Circuits**

For a) to c) below, find the gain, and the input and output impedance of each of the op amp circuits below:

![Diagram of Op Amp Circuit](image)

**Figure 5.19: Op Amp Simple Circuits Eg. a)**

a) In this case, we have an inverting configuration (the input is going to the inverting input). The gain is simply \( \frac{-R_2}{R_1} = -7.5 \). Since the (-) input acts as a virtual ground, the input impedance is \( R_1 = 10\,\text{k}\Omega \). \( R_3 \) is in series with the very low output impedance of the op amp, so \( R_{\text{out}} = 5\,\text{k}\Omega \)

![Diagram of Op Amp Circuit](image)

**Figure 5.20: Op Amp Simple Circuits Eg. b)**

b) Since the input is tied to the (+) input, this is a non-inverting configuration, and the gain is given by \( A_v = \left(1 + \frac{R_1}{R_2}\right) = \left(1 + \frac{75\,\text{k}\Omega}{10\,\text{k}\Omega}\right) = 8.5 \). The input impedance is 100k\Omega resistor in parallel with the input impedance of the op amp (very high), so it is still 100k\Omega. Looking back into the output, \( R_2 \) is virtually grounded at the (-) input and in parallel with the internal output impedance of the amp, but since \( R_{\text{int}} \) is very low, the output impedance is still \( \approx 0\Omega \)
c) This also looks like a non-inverting configuration, but since the feedback resistor is 0Ω, it is the special case called a voltage follower, with a gain of 1. R3 and R1 in the input and output respectively are somewhat misleading. Their placement is such that they actually have no effect, and the input impedance is still that of the op amp, \( \approx \) infinite (>10\(^{12}\)Ω), and the output impedance is also that of the op amp: \( \approx 0\)Ω.

d) In a) above, at what frequency does the gain start to “roll off” (decrease) if the TL082 op amp has a Unity Gain Bandwidth of 3MHz?

\[
\left( A_V \right) \left( BW \right) = 3 \text{MHz}; \quad BW = \frac{3 \text{MHz}}{A_V} = \frac{3 \text{MHz}}{75} = 400 \text{kHz}
\]

e) The “slew rate” for the TL082 op amp is given as typically 13V/µs. If it is desired to obtain a sinusoidal output signal with a \( V_{\text{0-peak}} \) of 5V, what is the maximum allowable frequency?

\[
f_{\text{max}} = \frac{slew \_ rate}{2\pi V_{\text{peak}}} = \frac{13V}{2\pi (5V)} = 413.8 \text{kHz}
\]
Example 5-2: Tandem Inverting Op Amps

Objective: Design an op amp circuit using the TL082 integrated circuit that will provide a voltage gain of -10, have an input resistance of 1MΩ, and output impedance of 10kΩ and be capable of amplifying a maximum frequency of ≥ 700kHz at the specified gain of -10.

First, we note that the gain required is a positive 10, so we must either use a non-inverting configuration, or a back-to-back inverting configuration. Then note, from the device specifications, that the TL082 IC has 2 op amps in the one package (this will turn out handy). Also note the Unity-Gain Bandwidth is 3MHz. This implies a maximum flat gain frequency limit of 300kHz at a voltage gain of 10. This means we will need to use both op amps in the package to achieve a gain of 10 with a minimum frequency capability of 700kHz. To keep inverting (-10), we will need to design one as inverting and the other non-inverting.

1) First, determine the maximum flat gain with a frequency limit of 700kHz:

\[
(A_v)(BW) = 3MHz; \quad A_v = \frac{3MHz}{700kHz} = 4.286
\]

this means we must design each stage with a voltage gain of no more than |4.38| if we want to maintain the 700kHz bandwidth. If both stages were designed to this limit, the overall gain of the two stages in tandem would be 4.3^2 ≈ 18.5, so this combination is capable of meeting our objective of |10|.

2) Select the gain for the 2 stages that will result in an overall gain of -10 while keeping each one below ≈ |4.3|.

Select +2.5 and -4 (any combination that gives a product of -10 will do, but this will give us some margin on each stage and simplify the arithmetic)

Refer to Figure 5.15 and Figure 5.16 and the associated gain expressions:

Inverting: \[ A_v = \frac{V_{out}}{V_{in}} = -\left(\frac{R_2}{R_1}\right) \] \text{Equation 5.1}\n
Non-inverting: \[ A_v = \frac{V_{out}}{V_{in}} = \left(\frac{R_1 + R_2}{R_1}\right) = \left(1 + \frac{R_2}{R_1}\right) \] \text{Equation 5.2}\n
For an inverting gain of |4|, we can select the feedback resistor, R_2 as 40kΩ, and determine the value for R_1 as 10kΩ

(Note: in theory, any resistor values of that ratio would provide the require feedback characteristic and gain. The actual choice is a compromise between a resistance so high that the very small input current affects the voltage divider ratio, and so low that the drain on the output is enough to affect the output voltage even with a very small output resistance of the op amp. For the TL082, selections in the 10s of kΩ or even the 100s of kΩ are a good compromise.)
For a non-inverting gain of $|2.5|$, we select $R_2$ and $R_1$ to give a ratio of $\frac{R_2}{R_1} = (A_v - 1) = 1.5$.

Select $R_2$ as $15\,\text{k}\Omega$ and $R_1$ as $10\,\text{k}\Omega$. The circuit configuration with these values is shown in Figure 5.22.

![Figure 5.22: Tandem Inverting Op Amp Example](image)

3) Finally, we need to address the input and output impedance requirements. As previously discussed, the input impedance of a non-inverting configuration is the input impedance of the op amp; $\approx 10^{12}$ ohms in this case. We can set a specific input impedance by putting a resistor in parallel with the input of the amp. For the required $1\,\text{M}\Omega$ input impedance, we can put a $1\,\text{M}\Omega$ resistor in parallel with the input ($10^{12}\Omega$ in parallel with $1\,\text{M}\Omega$ is still very close to $1\,\text{M}\Omega$).

Recalling that the output impedance of a good op amp is very low ($\approx 10\,\text{s}$ of $\Omega$s), we can simply put a $10\,\text{k}\Omega$ resistor in series with the output so that it will “look like” $10\,\text{k}\Omega$ ($10$ of $20\,\text{ks}$ in series with $10\,\text{k}\Omega$s is still very close to $10\,\text{k}\Omega$).

These input and output impedance setting resistors are also shown in Figure 5.22 as $R_5$ and $R_6$ respectively.
6 Energy Sources and Power Conversion

This Section will discuss some of the ways energy is provided to electronic and electromechanical devices. In most cases, the voltages required for various purposes are varied and usually different than the available source. AC sources must be converted to DC, voltages must be changed to supply electronic controllers (5VDC, 3.3VDC), solenoid valves, motors, relays (12VDC, 24VDC etc.). In other cases, batteries and other forms of storage devices are used to supply stored energy, or devices like generators and photovoltaic cells are used to convert mechanical, wind or solar energy to electrical energy. In all these cases, there is usually a need to control the output voltage of these generators or converters within certain tolerances in order to ensure the proper functioning of the electrical devices.

Topics that will be covered, at least to some extent, include:

1. Power Conversion
   a) AC (mains) Power Supplies
      i. Half wave rectifiers
      ii. Full wave rectifiers
      iii. Switching supplies
   b) Filtering of Rectified or Switched Power Supplies
   c) Voltage Regulation
      i. Linear Regulators
      ii. Regulated Switch Mode Power Supplies (SMPS)

2. Power Generation and Storage
   a) Mechanical Generators
      i. Wind energy
   b) Batteries
      i. Lead-acid
      ii. Carbon-Zinc
      iii. NiCad
      iv. NiMH
      v. LiPo
   c) Solar Cells
6.1 AC (mains) Power Supplies

Alternating Current (AC), and its extensive distribution system, is the most common way to distribute electrical energy in most parts of the world. In Canada, for instance, the annual electrical energy consumption is estimated at approximately 500 Billion kWh. (U.S Energy Information Administration) This compares to an annual estimated energy consumption from all sources of approximately 3800 Billion kWh (kilowatt-hour equivalent). Assuming an average conversion efficiency of 35% for all uses except heating, and making some adjustment for the petroleum used for electricity production, this means energy consumption from petroleum and electricity are roughly comparable. It is no wonder that this extensive electrical energy distribution system is commonly converted to drive electrical gear and devices in homes, businesses, institutions and industry.

AC to AC and AC to DC conversion

A very common method of providing DC power to devices is by using rectifiers to “polarize” the alternating input. “Rectifier” is the generic term for devices that allow current flow in one direction, but not the other (i.e. one way valves). The most common devices in use today are the diodes that were considered in a previous section of this course. The DC output from a diode arrangement is usually close to the peak AC input voltage. Because of this, the AC voltage is often converted to an appropriate level using a transformer (also previously studied) prior to rectification. Finally, the “pulsed” nature of a simply rectified AC waveform is often unsuitable for most DC equipment, and some “filtering” must be applied.

6.1.1 Unfiltered Half-wave Rectifier Configuration

Although half-wave and full-wave configurations were considered during the investigation of diode properties, they are repeated here for reference, and with the additional components and considerations required of a “power supply”. Figure 6.1 shows a typical half-wave configuration. A 120 VAC 60 Hz input and a step-down transformer with a 10:1 turns ratio, will result in a peak secondary output voltage of 17V minus the forward voltage drop of the diode used.

![Figure 6.1: Half-wave Rectifier Configuration](image)

An approximation of the resulting output waveform is shown in Figure 6.2. A 1V forward voltage drop for the diode was assumed, leaving a peak voltage of 16V. The average output is approximately 5.1V, although it is clearly a “pulsed” output that ranges
from 0 to 16 V. While this might be an adequate source for some tolerant applications, a much “steadier” output is usually required, and can be achieved either through a full-wave rectifier or at least some filtering. (Note: a typical load is simply represented in the schematic by a resistor.)

![Graph](image)

**Figure 6.2: Half-wave Rectified 60 Hz AC**

### 6.1.2 Unfiltered Full-wave Bridge Rectifier Configuration

A full-wave bridge rectifier configuration utilizes the full AC cycle by arranging four diodes as shown in Figure 6.3. This schematic again shows a step-down transformer with a 10:1 turns ratio, resulting in a peak output of 15V from a 120 VAC source if we consider a 1V forward voltage drop across each of the two diodes in current path for each half of the cycle (17V peak – 2V).

![Schematic](image)

**Figure 6.3: Full-wave Bridge Rectifier Configuration**

The output waveform for a full-wave configuration is shown in Figure 6.4, where the peak is now only 15V (17V – 2V_D@1V). The output, while still “pulsed”, has an average output voltage of \((2/\pi \times 15V = 9.55V\). This type of output is often adequate for powering DC motors, solenoids or relays, but would not be “smooth” enough for most electronic equipment.
6.1.3 Unfiltered Full-wave Rectifier Configuration
An alternate way of providing a full-wave output with only two diodes instead of four is shown in Figure 6.5. This configuration also has the slight advantage that there is only one diode voltage drop in series with the load for each half cycle, but the disadvantage is that a centre-tapped secondary is required on the transformer. Note also that the transformer in Figure 6.5 has an overall 5:1 turns ratio to achieve approximately the same output voltage (It is effectively a 10:1 ratio for each “half” of the secondary so that the 17V peak is impressed across one of the diodes during each half cycle). Compared to the bridge configuration, the peak output voltage is 16V.

![Figure 6.5: Full-wave Centre Tap Configuration](image)

In the next Section, we investigate how we can use a capacitor to provide some filtering, thus achieving an output much closer to a true Direct Current.

6.1.4 Filtering

The characteristics of a capacitor can be used to “maintain” the voltage between the “replenishing” peaks of the rectified waveform. Figure 6.6 shows the placement of a filter capacitor across the output terminals, in parallel with the load. Note that the type of
capacitors commonly used for filtering applications are called “electrolytic” because of the type of construction. The materials used result in a smaller physical size, but also in a device that is polarized, which means that they have a positive and negative terminal, and must connected properly in the circuit or they will not perform properly and will be severely damaged.

![Diagram of Filtered Full-wave Bridge Configuration]

**Figure 6.6: Filtered Full-wave Bridge Configuration**

Analyzing the filtered circuit in Figure 6.6, we observe that when the diodes are reverse biased, the source side of the circuit is isolated from the output side (i.e. the diode switches are “off”), and the capacitor and load can be considered alone. During this period, the capacitor will discharge through the load, its voltage decreasing exponentially as it does according to the familiar expression:

\[
v_L(t) = v_C(t) = V_o e^{-t/\tau} \quad (6.1)
\]

where \( V_o \) is the nominal output voltage, \( v_L = v_C \) is the output voltage across the load as it decays over the time period involved, and \( \tau \) is the time constant \( R_L C \), and \( t = 0 \) at the start of the discharge period, when the sinusoid is at its peak. The output “ripple voltage”, \( V_R \), is defined as the variation in the output voltage (i.e. the difference between the peak and the lowest voltage before the capacitor is charged again). While we can use equation (1) to calculate this variation, the reasonable assumption that the voltage decays linearly over ~ a full half period if \( V_R \) is \( \ll V_o \) results in the following expression for determining the capacitor size for a desired ripple voltage:

\[
C = \frac{I_{\text{Max-Out}}}{fV_R} \quad (6.2)
\]

where \( I_{\text{Max-Out}} \) is the maximum expected output current, \( V_R \) is the maximum desired, or tolerable, output ripple, and \( f \) is the “replenishing” frequency (e.g. 60 Hz for a half-wave configuration and 120 Hz for a full-wave configuration with a 60 Hz AC source). The predicted output waveform for the power supply shown in Figure 6.6 is shown below in Figure 6.7.
Figure 6.7: Filtered 60 Hz AC (Ripple Voltage)

This waveform is based on a 120 VAC 60 Hz input, a 17 VDC peak output, and a 10,000 \( \mu F \) filter capacitor (note polarization). This results in a measured \( V_{Ripple} \) of ~ 0.75V at the output with an assumed maximum draw of 1 A (i.e. a 17\( \Omega \) load). Using these parameters with equation (6.2) specifies a filter capacitor:

\[
C = \frac{I_{Max-Out}}{fV_R} = \frac{1A}{(120V)(.75V)} = 11,111\mu F \quad (6.3).
\]

Note that this is both reasonably close and errs on the “safe” side of the actual 10,000\( \mu F \) capacitor assumed in the plot calculation. This small difference would not significantly affect the result, and component variation (a 10% tolerance in value is not uncommon) could have a similar effect. This calculation is at maximum load of 1A; anything less and the ripple voltage will be less than the .75V limit specified.

The “charge” and “discharge” cycle occurs over a half cycle of the input frequency, or 8.333 ms in this case. (Note, this would be over a full cycle for a half-wave configuration.) Figure 6.8 is an expanded view of two “peaks” from a filtered output similar to that depicted in Figure 6.7. Observe that the waveform shows a portion of the whole cycle where the load is receiving current from the capacitor only (discharge) and a much shorter portion where the load and the capacitor are receiving current from the supply through the diode(s) (charge). This allows separating the period in to the charge and discharge portions, as labelled in Figure 6.8.

What is needed is to find the time at which the decaying capacitor voltage intersects the rectified sinusoid that will “re-charge” it. One way to get a reasonable estimate is to simplify the discharge curve by assuming it is linear (a reasonable assumption for the
early part of the exponential delay) and will reach the maximum discharge by the time the sinusoid intersects (refer again to Figure 6.8). The discharge time is from the peak of the last sinusoidal “half-cycle-pulse” until the sinusoid reaches \( V_{\text{Peak}} - V_{\text{Ripple}} \). If we “restart the clock” at the beginning of the second rising sinusoid, we can use the expression

\[
\nu(t) = V_p \sin(\omega t) = V_{\text{Peak}} - V_{\text{Ripple}}
\]

to describe the intersect point in that portion of the cycle. We must add to that \( \frac{1}{4} \) of the period of the source waveform (60 Hz in this case, not the 120 Hz we used in the capacitor calculation.

![Figure 6.8: Filtered 60 Hz AC (10μF Capacitor)](image)

Also note, for a half-wave configuration, we need to add \( \frac{3}{4} \) of a period). The total discharge time can be determined:

\[
V_p \sin(\omega t) = (V_p - V_R), \text{ solving for } t
\]

\[
t = \frac{\sin^{-1}\left(\frac{1-V_R}{V_p}\right)}{2\pi/T}
\]

\[
t_{\text{dis}} = \frac{T}{4} + \frac{\sin^{-1}\left(\frac{1-V_R}{V_p}\right)}{2\pi/T} \quad (6.4)
\]

where the inverse sign is in radians. If the inverse sine function is calculated in degrees:

\[
t_{\text{dis}} = T \left\{ \frac{1}{4} + \frac{\sin^{-1}\left(\frac{1-V_R}{V_p}\right)}{360^\circ} \right\} \quad (6.5),
\]
where $T$ is the period of the input AC waveform (16.667ms for a 60Hz source). Also remember that the $\frac{1}{4}$ term becomes $\frac{3}{4}$ for a half-wave rectified waveform.

Recall that the “replenish” period is $T/2$ of the source waveform for a full-wave configuration, and $T$ for a half-wave configuration. The charge time is then simply:

$$t_{\text{charge}} = \frac{T}{2} - t_{\text{dis}} \quad \text{for the full-wave case, and}$$

$$t_{\text{charge}} = T - t_{\text{dis}} \quad \text{for the half-wave case.}$$

If, for example, we use the previous example parameters (60 Hz, full wave rectified source, peak output voltage of 15V, a 1 Ampere draw at peak voltage (15Ω load), and a 10,000μF capacitor) the associated decay (discharge) time is ~ 7.5 ms. This leaves 0.833 ms to replenish the charge lost from the capacitor and bring $v_C$ back to the peak output voltage. This information can be used to determine the average and peak current capability of the diodes.

Since the amount of charge lost from the capacitor during the discharge portion of this cycle must be replenished during the charge portion, the “time-current” integral ($Q = I \times t$) must be the same in each part of the cycle ($\text{charge} Q = \text{discharge} Q$). Again, the reasonable assumption that $V_R$ is very small compared to $V_O$ simplifies analysis. In the example used here, 1 A is being drawn for 7.5 ms. so $(1A)(\frac{7.5\text{ms}}{0.833\text{ms}}) = 9A$ plus the 1A also drawn by the load during this period = 10 A that must pass through the diode during the “charging” portion of this cycle. The average is $\frac{8.33}{10A} = 1A$. Observe that the smaller the proportion of time available to “re-charge” the capacitor, the higher the ratio between the peak current and the average current.

This relationship can be manipulated into the more general expression:

$$I_{\text{Peak}} = I_{\text{Out-Max}} \left(1 + \frac{t_{\text{dis}}}{t_{\text{charge}}}\right) \quad (6.6)$$

where $I_{\text{Peak}}$ is the peak diode current required, $I_{\text{Out-Max}}$ is the maximum output current capability required of the supply, and $t_{\text{dis}}$ and $t_{\text{charge}}$ must be determined graphically or by using Equation (6.5) and (6.6).

The key parameters of a mains-fed power supply of the type described above are summarized in Table 6.1.
Table 6.1: Key Power Supply Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_O$</td>
<td>The desired output voltage for the supply</td>
</tr>
<tr>
<td>$I_O$</td>
<td>The maximum output current required</td>
</tr>
<tr>
<td>$V_R$</td>
<td>The maximum allowed output voltage variation at the maximum output current</td>
</tr>
<tr>
<td>$C_f$</td>
<td>The size of filter capacitor required to achieve the desired $V_R$. It must also have a voltage rating of at least $V_O$</td>
</tr>
<tr>
<td>$I_D$</td>
<td>The average current rating for the rectifier diodes</td>
</tr>
<tr>
<td>$I_{Peak}$</td>
<td>The peak forward current rating for the rectifier diodes</td>
</tr>
<tr>
<td>$PIV$</td>
<td>The Peak Inverse Voltage the rectifier diodes must withstand</td>
</tr>
<tr>
<td>Load Regulation</td>
<td>$\Delta V_O$ as $I_L$ changes</td>
</tr>
<tr>
<td>Line Regulation</td>
<td>$\Delta V_O$ as $V_{in}$ changes</td>
</tr>
</tbody>
</table>

Regulation will be considered in more detail in the next section. Completing the example design used above will also demonstrate the process of ensuring a reliable and capable power supply.

Example 6-1: An Unregulated Full-wave Supply

Determine the parameters for an off-line (120VAC source) 12VDC, 3A power supply with a maximum ripple voltage of 2%. Use a step-down transformer to design an unregulated “raw” supply with these specifications.

Expanding a bit on the parameters:
2% ripple is .24V or ±.12 volts, so the output can vary (ripple) from 11.88 to 12.12V.

We have 3 common choices for configuration: ½ wave, full-wave and full-wave bridge. Since we are short of diodes and would like to keep the filter capacitor as small as possible, let’s choose a full wave with a centre-tapped secondary transformer (only 2 diodes required for full-wave).

Typical silicon power diodes have a conducting threshold of ~0.7 volts, and a $V_F$ of ~1.1V at 3A (1N4007 datasheet). We should design for worst case, so we need another 1.1 volts peak from the transformer so we have 12.12 at the output. The RMS voltage we need out of the transformer is

$$\frac{12.12V + 1.1V}{\sqrt{2}} = 9.35V_{RMS}$$

![Figure 61: 1N4007 $V_F$ Characteristic](image)
Transformers are usually spec’d with an output voltage in RMS at a specified secondary current. Looking at the table in Figure 6.10, we find model 14A - 56 – 20 with a CT secondary of 20\text{V}_{\text{RMS}} (10\text{V} each side of centre) at a current of 2.8A. This isn’t enough and we should continue the search, but for now we’ll assume we can find one similar, spec’d at 3A secondary. (Actually, to get a 3A \textit{average} secondary current, we need and RMS current of 3.33A (3A average means: \[\frac{3A}{\sqrt{2}} = 4.71A_{\text{Peak}} = \frac{4.17A_{\text{Peak}}}{\sqrt{2}} = 3.33A_{\text{RMS}}\])

This model will give us an extra .65\text{V}_{\text{RMS}} which may be helpful to overcome unforeseen losses.

The next step is to calculate the size of the filter capacitor. With the familiar configuration shown in Figure 6.11, we can use the commonly accepted expression:

\[C_F = \frac{I_{\text{Out}}}{fV_{\text{Ripple}}} \quad (6.7)\]

\(C_F\) is the size of the filter capacitor in Farads, \(I_{\text{Out}}\) is the maximum average output current, \(f\) is the frequency of the “re-charging” (e.g. for a full wave rectified 60 Hz source, the capacitor is “re-charged” 120/s, so \(F = 120\)), and \(V_{\text{Ripple}}\) is the maximum desired (peak-to-peak) variation in the output voltage.

For your design we have:

\[C_F = \frac{I_{\text{Out}}}{fV_{\text{Ripple}}} = \frac{3A}{(120)(.24V)} = 104,170\mu F\]

Of course we will probably have to use the next higher commercially available size.

Find the \textbf{peak current rating for the diodes}:

Our average current rating for the rectifier diodes is 3A, as implied by the output parameter. To determine the \textit{peak} current, we must somehow estimate the ratio of the “discharge” time to the “charge”. First calculate the discharge and charge times:
We can now calculate our required peak current capability for the diodes:

\[ I_{\text{peak}} = I_{\text{Max-Out}} \left( 1 + \frac{t_{\text{dis}}}{t_{\text{charge}}} \right) = 3A \left( 1 + \frac{7.80}{.53} \right) = 47.2A \]

In summary (Table 6.2), we would need components with the following spec’s:

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transformer</td>
<td>120VAC primary with 18.7VAC CT secondary (9.35V per half) @3.33A average</td>
</tr>
<tr>
<td>Diodes (1)</td>
<td>I_{\text{Avg}}=3A; I_{\text{Peak}}=47.2A</td>
</tr>
<tr>
<td>Capacitor</td>
<td>105,000(\mu)F; 12.2WVDC (working volts DC), higher if we account for the probably higher no-load voltage coming from the transformer secondary. Eg. if the internal resistance was .4(\Omega), we would have another 1.2V peak at no load compared to 3A. If spec’s are not available for this parameter, it may have to be measured or estimated. In this size range, electrolytics are the only reasonable economic choice today. These are polarized, and should never be subjected to a reverse voltage.</td>
</tr>
</tbody>
</table>
6.2 Power Supply Regulation:

In many applications it is desired, or even necessary, to keep the output voltage very close to its nominal value. This provides stability of operation for many appliances (e.g. motor speed, light output) and also provides for safe operation of equipment by limiting the over-voltages to which it may be exposed.

Most often, regulation is added to the same “raw” DC source explored in the last section. An additional advantage of regulation is that it will usually regulate for both varying output loads (current) and varying input voltages (e.g. both varying line voltages and varying “raw” DC voltages as may be experiences because of losses in the associated components, like the transformer in the last example). Indeed, two main figures of merit for a regulated power supply are i) how well it maintains the output voltage over a varying load range, and ii) how well it maintains the output voltage over a varying input voltage range.

6.2.1 Linear Regulation

Recall one of the properties of a bipolar transistor was that there was a characteristic voltage drop (diode drop) from the base to the emitter. This junction behaves very much like a diode. While in diode applications we considered that $V_F$ would increase slightly if $I_B$ increased, the reverse is also true, if we increase the forward bias, $V_{BE}$, the current will increase, and this base current is multiplied by the $h_{fe}$ of the transistor. Referring to Error! Reference source not found., this means that the output voltage (connected to the emitter) will stay almost constant regardless of the current drawn through it. If the voltage tries to drop, this will increase the forward bias across the BE junction causing more base current to flow which will cause more emitter (and collector) current to flow. Essentially, the output voltage will stay reasonably close to $V_B - V_{BE}$, regardless of how much current is “drawn” though the collector/emitter, and also regardless of what the input voltage, $V+$ is! While this “regulation” is not based on any overt feedback, we might consider it to be the result of intrinsic feedback based on the transistors junction characteristics. We may be able to improve its performance if we could add some additional feedback, say, by monitoring the collector current or actual output voltage compared to some standard and then increasing the base voltage just enough to compensate for the slight drop due to the increase of $V_{BE}$ with increasing $I_C$.

These principles have been exploited in a family of popular linear regulator ICs that come in a variety of set ± voltages or in an adjustable configuration. All feedback and compensation, including some internal protection for over current and over temperature conditions, is neatly packaged in a three-terminal device: one terminal for input, one for output and one common (ground, like the base voltage in Error! Reference source not found.)
found.). They are housed in the same styles of cases used for transistors, with various sizes and materials for dissipating any heat that may generated.

Figure 6.13 shows a typical fixed (i.e. one voltage), positive (opposed to negative) regulator application. The manufacturer usually recommends placing two capacitors, as shown, as physically close as possible to the IC case to eliminate any high speed “noise” on the input or output leads that may result in unstable operation of the internal circuitry which may lead to uncontrolled, high frequency oscillation of the output. Other than these, there are no external components required.

The optional resistor shown in the schematic is to reduce the dissipation demands on the regulator itself. A simple power analysis will determine the dissipation requirements for the device itself. If, for example, the regulator is capable of delivering 1A at 5VDC output, and the input voltage is 20VDC on average, there is (20V)(1A)=20W going into the device and (5VDC)(1A) coming out, the difference, 15W, must be dissipated in the device itself! This can lead to high temperatures in the device and surrounding enclosure which could damage the device if not removed effectively. In situations where there is a significant voltage difference between the input and output voltage, we can remove some of the power dissipation from the device by lowering the voltage by using a ballast (dropping) resistor. This technique will be demonstrated in a subsequent example.
The key characteristics and parameters of such linear regulators are summarized in Table 6.3.

**Table 6.3: Summary of Linear Regulator Characteristics**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_o$</td>
<td>Same as any power supply: the specified nominal output voltage</td>
</tr>
<tr>
<td>$I_o$</td>
<td>The maximum specified output current capability</td>
</tr>
<tr>
<td>Drop Out voltage</td>
<td>The minimum difference between $V_{in}$ and $V_{out}$. As with Op Amps, there is some “circuit overhead” that requires some voltage in excess of the output voltage.</td>
</tr>
<tr>
<td>Maximum Input Voltage</td>
<td>The maximum input voltage allowed for safe operation</td>
</tr>
<tr>
<td>Line regulation</td>
<td>How much the output voltage varies with a change in input voltage</td>
</tr>
<tr>
<td>Load Regulation</td>
<td>How much the output voltage varies with a change in output current (usually specified over most of the maximum current range)</td>
</tr>
<tr>
<td>Line Transient Response</td>
<td>How the output voltage varies in response to abrupt changes in the input voltage (see Figure 6.14)</td>
</tr>
<tr>
<td>Load Transient Response</td>
<td>How the output voltage varies in response to abrupt changes in load current (see Figure 6.14)</td>
</tr>
</tbody>
</table>

**Figure 6.14: Transient Response Curves for LM2940**
Example 6-2: A Linear Regulated Supply

Provide a regulated, 5VDC, 1A supply using a linear regulator IC. As an input, use a full wave rectifier similar to Figure 6.11.

Modify the supply shown in Figure 6.11 by using a smaller filter capacitor, which will result in more ripple, but save room and expense. If we reduce the size to 10,000μF from 100,000μF, we will increase our ripple ($V_r$) from .24V to ~2.4V (re: Equation (5)). Assume all voltages remain the same as determined, even though we will be drawing a maximum of only 1A instead of 3A. This means our raw input voltage to the regulator will vary from 12.12 to 9.72V (12.12V – 2.4V)

We must design the rest of our circuitry to achieve the desired output with a minimum 9.73V input.

We will use the standard circuit configuration given in Figure 6.13 (and repeated at right for convenience).

Select the LM2940-5, with a 5V/1A output. Provided we are willing to accept the specified performance specifications of this model (summarized below), all that remains is to follow the manufacturers recommendations for de-coupling capacitors. On the input side, a regular, non-polarized .47μF capacitor is recommended. On the output side however ($C_2$), a minimum 22μF “low ESR” capacitor is recommended. “ESR” stands for Equivalent Series Resistance, and is a measure of how quickly the capacitor can charge and discharge, which affects its ability to respond to very high frequency variations in the voltage impressed upon it. This regulator relies on a very stable output voltage with very little high frequency “noise” in order to maintain its stability internally.

Power Check: We assumed above that we would have a maximum $V_{in}$ of 12.12V and a minimum of 9.72V. Assume no dropping resistor to start with. At the high end, the regulator must dissipate:

\[ (12.12V - 5V)(1A) = 7.12W \]

and at the lower end:

\[ (9.73V - 5V)(1A) = 4.73W \]

We can deflect some of this from the device into an external resistor ($R_1$), but we must make sure we don’t lower the input voltage at the maximum current to below the drop out voltage or the output voltage will not be maintained.

Referring to the datasheet, the maximum specified dropout voltage is 0.8V at 1A output current. This means we must not allow $V_{in}$ to drop below 5.8V to ensure the 5V output is maintained. This means a maximum voltage drop of 9.72V-5.8V = 3.92V. At 1A, a 3.92Ω resistor would provide this drop. A check of standard resistor values and supplier stores indicates a 3.6Ω, 5W 5% tolerance is available (the 5% is important, as a 3.6Ω 10% that is at the upper limit of its tolerance range would be 3.96Ω - too much). If we use a 3.6Ω, 5W resistor, we can expect the following:

At 12.12V, the device must dissipate

\[ (12.12V - (3.6\Omega \cdot 1A) - 5V) (1A) = 3.52W \]
At 9.73V, it must dissipate \( (9.73V - (3.6\Omega \cdot 1A - 5V))(1A) = 1.13W \)

Since the device dissipation varies constantly between the two extremes, we can take the average, and the device must handle 2.33W power dissipation.

The resistor must always dissipate 3.6W \((1A^2R)\), at least at maximum current draw. Thus we’ve reduce the heat dissipation requirement in the device from an average of 5.9W to 2.3W by the addition of an additional external component. It may also be easier to physically place the resistor away from the other circuitry for cooling.

The specifications are now dependent on the IC’s spec’s:
- Maximum input voltage: 26V
- Output voltage: 4.85 – 5.15V (device variation)
- Line regulation: 50mV \((V_{in} +2\) to 26V
- Load Regulation: 50mV \((I_o 50mA \) to 1A)
6.3 Regulated Switch-Mode Power Supplies

In a previous section, we noted one of the disadvantages of linear power supply regulation was the inefficiency (wasted energy). Switch-Mode Power Supplies (SMPS) exploit some of the characteristics of inductors and transformers to realize a more efficient, flexible design. Improvements in solid-state devices (like transistors and diodes) have made SMPS the common design choice for almost all conversion applications. This section discusses the basic principle of operation and the key characteristics of SMPSs.

6.3.1 Operating Principle

Recall that capacitors and inductors temporarily store energy in another form (and electric field in a capacitor, and a magnetic field in an inductor). While either could in theory be used, the inductor proves to be advantageous because of its ability to “self generate” any required voltage to “discharge” the energy stored in its magnetic field. Typically use a small, light inductor is used, exploiting the relationship between frequency, flux and voltage represented by:

$$E_{P\text{-max}} = 2\pi f N_p \Phi_{\text{max}} \quad (6.8)$$

Note that as $f$ increases, $N$ and $\Phi$ (related to inductance) will decrease for a given energy transfer rate (power).

Regulating the output is typically achieved by varying the amount of energy “loaded” into the magnetic field in order to accommodate variations in the input supply or the output load. By designing suitable circuitry, we are able to extract energy from a source with one voltage, temporarily store it in a magnetic field, and then “discharge” it at another voltage into a load.

Unlike the linear power supplies considered previously, SMPSs also have the advantage that the only losses in the voltage “conversion” process are the relatively minor component losses (e.g. core losses in the inductor), which can be kept very low leading to very high efficiencies.

SMPSs can have a variety of configurations, but they may be categorized as one of two basic types: the flyback or buck-boost converter, and the forward or buck converter. A third common designation is the push-pull converter, which is a variation of the forward converter. These all exploit the same principles, but vary in how the components are arranged depending on whether the desired output voltage is higher or lower than the source voltage. Exploring the operation of one configuration, the flyback converter, will adequately demonstrate the basic principles of operation which apply to other configurations as well.

6.3.2 Flyback Converters:

A flyback converter is a fairly simple architecture that is easy to construct and manipulate. To understand the basic design, considering the inductor storage element is a good starting point. You may recall that the energy stored in an inductor is:
\[ W = \frac{LI^2}{2}, \quad (6.9) \]
and that the voltage across an inductor is:
\[ v_L = L \frac{di}{dt}, \quad (6.10) \]

Using these basic relationships from inductor theory and simple circuit analysis techniques such as Kirchhoff’s Voltage Law, you can understand how a flyback converter operates and calculate the requirements for the required components. The basic flyback configuration is shown in Figure 6.15. When the switch is closed (i.e. \( V_{sw} = 0V \)), \( V_{in} \) is present across the inductor (i.e. \( v_L = v_{in} \)), \( L \), and the diode, \( D1 \), is reverse biased (i.e. no current flows to the output from the source). The current builds up in the inductor according to the expression given in Equation (2) (re-arranged):
\[ di = \frac{V_{in}}{L} dt. \]
If \( V_{in} \) and \( L \) are at least reasonably constant, and the current starts at 0A, it will rise linearly as long as the switch is closed (\( \Delta t \)) and the inductor does not saturate, up to some peak current, \( I_{peak} \), letting us express this change as:
\[ I_{peak} = \frac{V_{in}}{L} \Delta t. \quad (6.11) \]
Rearranging, we have an expression for the time required to “charge” the inductor to a certain level of energy:
\[ \Delta t_{charge} = \frac{L}{V_{in}} I_{peak} \quad (6.12) \]

![Figure 6.15: Simple Flyback circuit](image)

When the switch is opened, the magnetic field that was built up in the inductor core material will collapse, inducing a voltage across the inductor of a polarity opposite to \( V_{in} \) which will forward bias \( D1 \) and allow current to flow, delivering energy to the load, \( R_{Load} \), and charging the output filter capacitor, \( C \). The current generally decreases linearly again according to Equation \( I_{peak} = \frac{V_{in}}{L} \Delta t. \quad (6.11) \), and using KVL we find that \( V_L \) must be equal to \( (V_o + V_{D1}) \), where \( V_{D1} \) is the forward voltage drop across \( D1 \) while it is conducting and \( V_o \) is the desired output voltage across the load. We can rearrange this equation to find the time required for the current, and thus the magnetic field, to fall from \( I_{peak} \) to zero:
\[ \Delta t_{\text{Discharge}} = \frac{L}{V_o + V_{D1}} I_{\text{peak}} \quad (6.13) \]

It is also important that we eliminate any residual build-up of magnetic flux in the inductor core from one charge cycle to the next so the inductor does not saturate. To do this, we must either allow sufficient time for the current to fall to zero each cycle or employ other techniques to dissipate any remaining magnetic energy in the core.

If we make the simplifying assumption that \( V_{D1} \) is negligible compared to \( V_O \), we can rearrange equations (10) and (11) to get an expression for the duty cycle (the ratio of the charge time to the total time):

\[ \gamma = \frac{t_{\text{Charge}}}{t_{\text{Charge}} + t_{\text{Discharge}}} = \frac{1}{1 + \frac{V_{in}}{V_{out}}} \quad (6.14) \]

Recall from equation (10) that the energy that can be stored in the inductor, (and thus transferred to the load during one “cycle”) is a function of the time the inductor is “charged”, and from equation (12) we note that the charge time is a function of the ratio of \( V_{in} \) to \( V_{out} \). This is an unwanted limitation of this simple configuration.

The advantage of the flyback configuration is that the full input voltage, \( V_{in} \), can be used to “charge” the inductor, regardless of the output voltage, and the inductor will generate what ever voltage is require to discharge this energy when the field collapses. Thus, the output voltage, \( V_o \), may be either lower (buck, or step-down) or higher (boost, or step-up) than \( V_{in} \). The amount of charge put into \( C \) will depend in the energy that was stored in the inductor (given in Equation (1)), which is proportional to \( I_{\text{peak}} \) which in turn is dependent on the length of time the switch was closed, \( \Delta t \). By adjusting the length of time or how often the switch is closed, we can provide just enough energy to the output capacitor to make up for the current that was drawn by whatever load, \( R_L \), we might attach to the output. Of course, between charges the voltage across the capacitor will decrease, giving \( V_o \) a “ripple”, similar to the conventional full-wave rectifiers studied previously. The amount of ripple will depend on the size of \( C \), the load current, and how often the energy is “replenished”. A distinct advantage of a SMPS is that the frequency can by made very high, since it can be controlled by the circuitry, and not limited to the power line frequency. This means the size of capacitor required for a given \( V_{\text{Ripple}} \) is very small compared to a 60 Hz line-fed supply, and the inductor can also be very small as only a very small amount of energy needs to be transferred each cycle as long as the “cycle rate” is very high.

There are two other considerations with the simple configuration shown in Figure 6.15. First, there is no electrical isolation between the input and output. When dealing with off-line power supplies (i.e. directly connected to a commercial AC power source), \( V_{in} \) could easily be several hundred volts, which presents a safety hazard. The second issue is that \( V_o \) is inverted with respect to \( V_{in} \), which will make it much more complicated to monitor \( V_o \) and adjust the parameters of the “charge” cycle and regulate the output voltage accordingly. Figure 6.16, also shows a simple version of the flyback configuration that uses a transformer as the inductance, which addresses both of these
problems. The operation is identical except that the energy built up by current in the primary when the switch is turned on is recovered through current flowing in the secondary when the switch is turned off. The currents and voltages, as you will recall, are related by the turns ratio of the transformer. If we modify $V_{in}$ and $V_{out}$ in equation (13) with the transformer turns ratio, this also provides the advantage of allowing approximately half the total cycle for each of the charge time and discharge time thus optimizing the energy transfer for a given inductance.

The other “improvement” made in Figure 6.16 is that a FET is used as a switch and put in the “ground” side of the circuit. This facilitates turning the “switch” on and off very quickly with an electric signal. You can also see that, because the secondary is electrically isolated, we are no longer faced with the same safety hazard, and the negative side can potentially be connected to the same signal ground as the switch control circuitry making it easier to use feedback from the output to control the switch. (In practice, an opto-isolator is often used to provide additional isolation between the output and input, especially when the control circuitry is powered from $V_{in}$.) A simple feedback loop is also shown to illustrate the basic mechanism typically used to control the output voltage.

![Figure 6.16: Transformer-based Flyback Configuration](image)

Figure 6.16 shows a summary of the voltages and current across the main components during one “cycle”. The voltage across the primary (i.e. inductance) can be obtained by applying KVL while the switch is closed, and then open. A similar analysis gives the voltage across the switch, which is a key consideration when specifying the FET parameters. The peak primary current is also needed to determine the FET switch requirements, and the peak output current is needed to specify the required diode characteristics. Not shown is the peak inverse voltage (PIV) across the diode, which could also be obtained applying KVL during the “charge” part of the cycle when it is reverse biased and not conducting.

![Figure 69: SMPS Voltages and Currents](image)
6.3.3 Key Characteristics

Most of the characteristics that apply to other power supplies also apply to SMPSs, namely input and output voltage, power handling capability (i.e. output current), and both static and dynamic line and load regulation. In addition, because of the high frequencies involved, the electromagnetic fields generated have the possibility of causing electromagnetic interference (EMI). While this is not generally considered a “power supply” specifications, it may be of concern if the supply is to operate in the vicinity of other equipment that may be adversely affected by EMI.
Example 6-3: A SMPS Supply

![Diode Bridge Circuit Diagram]

Figure 6.18: PS Example 3 - SMPS

a) Calculate the voltage and current requirements for the diodes, capacitors and FET switch in the following power supply circuit given the 120VAC input and a required 12 VDC output at 3A maximum current draw. Also calculate the required turns ratio and power handling capability for the transformer. You can assume the supply regulates the output by monitoring the output voltage and varying the duty cycle of the switch input accordingly. You can also assume that the source voltage (120VAC) remains constant.

b) Calculate the frequency required for the pulse width control to achieve the required power output. Assume the transformer primary has a saturation current of 1A and a primary inductance of 200\( \mu \)H (i.e. the core of the transformer will stay in the linear region and not saturate with a current up to 1A).

Start at the source input side: the peak voltage impressed across the diode bridge will be: \((120V)\left(\sqrt{2}\right) = 170V_{\text{peak}}\). Let’s assume a maximum of 1V for \(V_F\) for all diodes in this circuit. Placing 170V DC (instantaneous) across the AC inputs and using KVL around the loops involving the diode bridge tells us that the Peak Inverse Voltage (PIV) they will be subjected to is 169V. Going further with this loop analysis, we find the maximum voltage across the filter capacitor is 168VDC. We will have to go a bit further to find the current requirements.

Start with the output power: 3A @ 12V = 36W. This is a “steady” (continuous) 3A under full load. Of course our supply will only be providing current to the output filter capacitor, \(C_2\), part of the time. We need to figure out how this 3A steady output translates into a linear “ramp” with a peak current as shown for \(I_{\text{Out}}\) in Figure 69.

Recall that the slope (m) of this ramp is \(L/V_f\). Also recall that the secondary voltage is \(nV_{\text{Primary}}\), and the secondary current is \(I_{\text{Primary}}/n\). Also recall from our study of inductors, that the inductance is proportional to \(N^2\). With a transformer we have two different coils on the same core, and so the inductance of the secondary is \((L_{\text{Primary}})(n^2)\). If we make these substitutions into equation (13) above, we see that if we use a turns ratio that exactly matches the input voltage to the output voltage, the “\(\Delta t\)” for both the charging and discharging parts of the cycle are equal, and we will be able to maximize our power throughput by using half the cycle time to “charge” the inductor, and the other half to push this energy into the load side of the circuit. (In
practice, we would probably cap the duty cycle at ~45% to allow a little margin for variation in components and line voltages.)

**Explore further to see the advantage of a transformer:**
Assume all we had was an inductor with the same inductance as the primary: 200μH.

Using equation (11): \[ \Delta t_{in} = \frac{200 \mu H}{170V}(1A) = 1.18 \mu s \] . Using the same inductance for discharge with a lower voltage: \[ \Delta t_{out} = \frac{200 \mu H}{12V}(1A) = 16.67 \mu s \]. With these times required to charge and discharge at the maximum current, we can have a maximum frequency of:

\[ T = 1.18 \mu s + 16.67 \mu s = 17.85 \mu s \quad f = 56.0 kHz \]

Recall the amount of energy stored (transferred) during one cycle is:

\[ W = \frac{LI^2}{2} = \frac{(200 \mu H)(1A)^2}{2} = 100 \mu J \]

The power throughput is just the energy/time (1J/s = 1W):

\[ P = Wf = (100 \mu J)(56.0 kHz) = 5.6W \]

Unfortunately, the power required is: \[ P_{required} = (12V)(3A) = 36W \]

Now let’s assume we can balance the times by using a transformer to make the voltages the same for charging and discharging times:

\[ T' = 1.18 \mu s + 1.18 \mu s = 2.36 \mu s \]

In this case, our maximum frequency is: \[ f = 423.7 kHz \], and the maximum power we could transfer is:

\[ P = Wf = (100 \mu J)(423.7 kHz) = 42.4W \]

Which would satisfy our requirement. 😊

Although we would normally allow some margin for line variations, we were allowed the assumption of a constant source supply voltage. However, we cannot simply assume that the input voltage will always be the 168V peak we would see across the filter capacitor because this voltage will “ripple” between “charges” of that capacitor (at 120Hz with the full-wave rectifier shown). We need to work backwards to find the average current drawn from the source supply and determine the ripple so we know the minimum voltage at the input and use that to determine the turns ratio. This may require some iteration unless we are constrained by, say, the filter capacitor size.

To simplify our calculations, let us assume a duty cycle of exactly 50%, and a frequency that will give us the output power required at this frequency (1/T).

Applying our knowledge of average values, we can determine that the peak is 4 time the average (with our simplifying assumptions). This means a peak secondary current of 12A.

Since we were told the maximum primary current before
saturation was 1A, this means we need a turns ratio of no more than $(1A/12A) = .0833$. To provide for the 12V output and 1V drop across the diode, we need 13V (average) at the secondary. If we design for a maximum 0.2V ripple, the peak output must be 13.1V ($13V+.2/2$). Using the turns ratio we determined to meet the current constraints, we will need a minimum primary voltage of:

$$V_{\text{Primary}} = \frac{13.1V}{.0833} = 157.3V$$

Since we have a peak of 168V on the primary side, we can withstand a ripple of (168-157.3=10.7V). The average current on the primary side (similar to the secondary reasoning) is $\frac{1}{4}$ the peak, or 0.25A. We can use these values to calculate the required filter capacitor on the input side:

$$C_1 = \frac{.25A}{(120Hz)(10.7V)} = 194.7\mu F$$

We can use the minimum input voltage to find our “charge” time:

$$\Delta t_{in} = \frac{200\mu H}{157.3V} (1A) = 1.27\mu s$$, and $f = (2\times1.27\mu s)^{-1} = 393.25kHz$

($\Delta t_{chg}$ is $\frac{1}{2}$ the cycle, assuming ideal conditions, so $T = 2\Delta t$.)

A check of our output power at a maximum input current (1A peak in the primary), minimum input voltage (157V) at the low end of the ripple.

$$P = Wf = (100\mu J)(393.25kHz) = 39.3W$$

Calculation of the output filter capacitor size will finish our calculations:

$$C_2 = \frac{3A}{(393.25kHz)(.2V)} = 38.1\mu F$$

Note the relatively small size of both filter capacitors compared to an off-line” supply alone!

Finally, again using KVL, we can determine the voltage requirements for the FET switch:

$$+V_{in} + (V_{O/n}) - V_{DS} = 0; \quad +168V +13V/.0833 -V_{DS} = 0; \quad V_{DS} = 324V$$

In summary, the component specifications would be:

1. Transformer: $L_{pri}=200\mu H; I_{Sat}=1A; n=.0833$
2. Input filter cap: 195\mu F, 168VDC
3. Bridge diodes: 169VDC PIV, $I_{avg}= .25A$
4. Output filter cap: 40\mu F, 13VDC, Low ESR
5. Output diode: 26V PIV, $I_{avg}=3A, I_{Peak}=12A$, fast switching type
6. FET, $V_{DS}=324VDC$, $I_{Peak}=1A$
7 Embedded Microprocessors and Controllers (Microcontrollers)

This Section will provide a very brief overview of microprocessors and discuss some of their uses in control applications. It will provide the basic understanding required to utilize the capabilities of this technology in the many electro-mechanical and process control applications that exist today. It is not intended to be a substitute for a proper study of processor architecture, fabrication and operation that would be required to design specific, detailed hardware or write control code with anything but a high-level programming language.

The example application around which this section is built is the speed control of a small DC motor. Also relevant to this microcontroller application will be a separate section explaining the basic operation of motors and generators.

7.1.1 Terminology

A number of related terms are often used to refer to “computer-based” equipment and processes. For the purposes of this overview, the following working definitions will be adopted:

- Computer: A device capable of interpreting a set of instructions and inputs, and producing an output dependent on those instructions and inputs.
- Software: The instructions given to a computer, that exist only temporarily and may be modified by the operator to alter the operation of a computer.
- Firmware: Instructions, as described in “software”, that are stored in a non-volatile way such that they are “remembered” even when power is completely removed from the device.
- Microcomputer: A computer based on a microprocessor
- Microprocessor: An integrated circuit that performs arithmetic and logic operations under the control of a set of instructions.
- Microcontroller: A microprocessor-based device that is used specifically in control applications by acquitting input signals and setting appropriate output signals. The terms microcontroller and microprocessor are often used almost interchangeably as most modern microprocessor integrated circuits include enough “input / output” circuitry that they can be used directly in many control applications.
- Embedded microcontroller: A microprocessor-based device that is usually built into another device for the purpose of controlling its operation or behaviour.

7.1.2 Microprocessor Hardware:
The “processor”, or “microprocessor” in most cases today, is the “brains” of a computer. It interprets and follows the instructions that control how it is to acquire information, process information, and provide a desired output in a desired form. Its main component is called an “Arithmetic – logic unit” (ALU) that performs, as the name would suggest, all the required calculations and makes the logical decisions necessary to achieve the desired result.

Figure 7.1 shows the block diagram of a Atmega644P microcontroller from Atmel Corporation.

![AVR Microcontroller Block Diagram](image)

**Figure 7.1: AVR Microcontroller Block Diagram**

This integrated circuit microcontroller is a member of their “AVR” family, which is very popular in embedded applications including many automotive uses. The central block (labelled AVR CPU, or Central Processing Unit) is essentially the ALU (arithmetic–logic unit) and basic control circuitry. It is surrounded by a variety of “assisting” blocks or circuitry that may be broadly categorized as follows:

- **Memory**: Various places where instructions and data are stored either permanently or temporarily. This IC includes SRAM (static random access memory, which can be read or written by the CPU and is used to temporarily store data and results), FLASH memory (permanent storage that can be easily re-written and is usually used to store program instructions), and EEPROM (electrically erasable programmable read only memory, which is another form of permanent storage used to maintain relatively permanent settings). Also included
are a number of “registers” which are sometime special places in RAM and sometimes separate pieces of memory used to store interim results).

- **I/O (input/output):** This includes circuitry designed to communicate digitally with other devices in either serial (e.g. two wire interface – TWI) or parallel form (8-bit parallel port communication), and circuitry designed to communicate serially using other protocols (e.g. the universal serial asynchronous receiver transmitter (USART)). This model also includes analog to digital conversion circuitry which will accept a continuous signal voltage as input, and convert it to an 8-bit or 10-bit digital representation.

- **Timers:** These are, as the name suggests, “clocks” that can be set, started, and stopped under program control to enable a number of time-related functions like delays or even pulse-width modulation of an output.

- **House keeping:** This includes a reference clock that synchronizes and controls the speed of CPU operations, power supervision that enables power saving modes to improve efficiency and lower power requirements, “watch dog” timer(s) that helps prevent permanent the unit from getting caught in an endless “wait” or “loop” situation by resetting the unit, and debugging assistance (JTAG (joint test action group) interface or OCD (on chip debugger))

- **Bus:** There is also a common communications path or “bus” that is used to pass data between the various blocks of the microcontroller. This communication is usually between the CPU and the other blocks, but in some circumstances may be possible directly between the blocks without using the CPU. Also note in this case that the CPU has direct communication with some of the memory which can improve overall operational speed.

While there are a variety of different functions or specific circuitry that could be optionally included in a microcontroller, and a variety of architectures (the layout of blocks and how they communicate), most devices share a common set of functions which are then supplemented by additional circuitry to allow for either a broad range of applications or very specific applications. The chip used as an example here would be considered a ‘general purpose’ microcontroller, while there are others that have specific peripheral circuitry to make them suited for applications like communication or lighting control.

### 7.1.3 Electrical Connections and Power

Most modern microprocessors are fabricated using a variant of complementary metal oxide semiconductor (CMOS) technology. The desire for faster, lower power processing has resulted in very small devices with narrow elements and connections requiring lower supply voltages. While this has achieved the goals of improving speed and lowering power consumption, other improvements were required to maintain drive capability. Even so, the output current of today’s devices is limited, and may have to be buffered (i.e. boosted) to interface with some external devices.

Microprocessors (often abbreviated as μP, or μC for microcontrollers) are “digital” devices whose inputs and outputs are considered either “high” or “low” (i.e. either a logic “1” or a logic “0”). Nominally these “levels” correspond to the supply voltage (V<sub>CC</sub>) or
ground (0V) respectively. However circuit conditions and internal impedances will affect the actual voltages at the I/O interfaces, and a logic high or low are usually specified as being over or under specific thresholds (e.g. with a $V_{CC}$ of 5V, 1 is > 4.5V and 0 is < .5V. These thresholds may be specified as a fraction of $V_{CC}$. Note: these values vary.)

Some key electrical characteristic to consider when using, and designing I/O connections for microcontrollers are summarized in Table 7.1.

For the most part, microcontrollers which are adequately powered for the desired frequency of operation will interface easily and reliably with other logic devices with the same $V_{CC}$. One should always check specifications to make sure that the “high” sent by one device is properly interpreted as a “high” at the receiving end.

### Table 7.1: Key μP Electrical Connection Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{CC}$</td>
<td>The supply voltage for the device. With modern materials, this can vary over a range. For example, the Atmel ATMega644P used as an example here will operate with a supply voltage between 1.8 and 5.5VDC. However, this affects the maximum operating frequency (e.g. at 1.8V $V_{CC}$, the AT644P will only operate at 4Mhz. It will only operate safely at 20Mhz with a supply voltage &gt;4.5V.)</td>
</tr>
<tr>
<td>Max frequency</td>
<td>This usually refers to the clock frequency. Most μPs will operate over a range of frequencies. Also, different μPs will execute an instruction in a different number of clock cycles. The processing power is sometimes given in MIPS (million instructions per second), which accounts for the number of clock cycles required for instructions. Often different types of logic operations require a different number of clock cycles. Also, as noted above, the operating frequency is related to the supply voltage.</td>
</tr>
<tr>
<td>I/O pins current-voltage (IV) characteristic</td>
<td>The amount of current a device I/O pin can sink (“absorb” from an external source) or source (supply to an external load) is often specified in relation to the resultant voltage at that pin. These limits may also vary with the supply voltage ($V_{CC}$). While an I/O pin may be capable of sinking or sourcing a given current, the resultant voltage may not stay within the specified limits to be read as a “high” or “low”. The relationship between the pin voltage and current must be considered when designing logic-level connections.</td>
</tr>
<tr>
<td>I/O pin maximum input voltage</td>
<td>Sometimes given in terms of $V_{CC}$ (e.g. $V_{cc} + .5V$); similar to breakdown voltage limits in other solid-state devices. This is not necessarily the same for all pins on a device.</td>
</tr>
</tbody>
</table>
7.2 Software and Code Development

What we have just discussed are some aspects of the hardware involved in a microcontroller application. The other half of the “package” is the software that controls how the hardware reacts to various inputs and conditions to produce the result desired. The software is what makes a microcontroller flexible and adaptable to different inputs and desired outputs. While the hardware is set early in the design and is very difficult to change after a piece of equipment is manufactured, the software (set of instructions) can even be changed “in the field” to adapt to knowledge gained through operating experience or from unexpected changes in conditions and constraints.

While most will recognize the term “software” (or “computer program” or “code”), not as many recognize how much effort, complexity and sophistication is involved in developing the “instructions” which control the behaviour of modern computing devices.

At the hardware level, instructions are stored as “ones and zeros” in some type for storage device (usually called some kind of memory or register), and executed by using those bits to enable or disable logic gates that pass or block a “bit” of information from moving from one side of a gate to the other. At the very basic level, this process is very simple, but the tremendous number of such operations that have to happen simultaneously and sequentially to effect even a trivial a result make the process very complex! In the first days of electronic computers, “coding” used to be done at this “machine” level (i.e. ones and zeros), but today most coding is written using what is called a “high level” language. Higher level programming languages each have their own vocabulary (allowable words) and syntax (grammar/formatting rules), but they all have the same objective: to allow someone to develop program code using a more natural language with more powerful words (instructions), which means the commands and instructions resemble words or abbreviations much like our everyday language instead of a series of “ones and zeros”. These “high level” instructions are written in a certain order with the appropriate syntax (i.e. grammar rules or format) so that they can be converted to “machine code” to be stored in the computer’s memory and executed when required. This “higher level” code is converted to machine code using, not surprisingly, a computer and a software program! A whole discipline and industry has evolved around the study and development of software structures, languages, architectures and applications. As with other “technologies”, an associated language and terms have accompanied that evolution. Some of these terms will be used and explained briefly in the following discussion.

One of the first higher “levels” was assembly code. This was a small step, basically just replacing the ‘ones and zeros’ with a mnemonic (abbreviation) representing one, or a few, instructions to try to make it easier to remember what the instruction(s) did. For example, “ADC” in AVR assembler means “add two registers with carry”, and “RJUMP” means “relative jump”. You can see that using this type of language still required an intimate knowledge of processor operations, even though it was much easier to use than pure machine code. Mnemonics instructions often had arguments (other pieces of information that needed to accompany the instruction, like addresses or values). Once the program was written following the required syntax and arguments, it would have to be converted into machine code and then stored in the computer’s memory. To allow
programmers to focus more on functionality and less on hardware operation, significantly higher level languages were developed. Names like Fortran, Basic, Pascal, C, C++, C# and other variants are used to develop program code are everyday words in the computer industry. Tools like these are computer programs themselves, often including features that help check syntax, debug (find and fix errors) and even simulate execution (also a way to “debug”). A relatively recent step in code development, especially for embedded applications, has been what is generally know as an Integrated Development Environment (IDE), that usually includes a high-level programming language, a compiler that will convert the high-level code to machine code for a specific microprocessor (or variety of microprocessors), debugging or simulation tools and a facility to download the machine code to the “target” CPU. Almost all higher level languages have some common attributes and categories of instructions, and each has their own peculiarities which can be advantageous in certain circumstances or environments. A basic appreciation of common themes and categories will allow us to take advantage of their power with a minimum of overhead.

A general structure of a “computer program” (“software”) showing some typical categorizations of program elements is shown in Figure 7.2. If consists of the initialization or setup of the hardware, the “main” part of the program that does most of the logical and computing work, and subsets of the code that are grouped to improve the efficiency and operation of the overall program.

![Figure 7.2: Common Program Component Categories](image)

The program, through hardware interaction, will respond to various inputs, manipulate or use the information, and interface with the hardware again to produce an output.
program can methodically execute the code in the prescribed sequence, testing for certain conditions, or it can be organized and written to respond “on demand” to “interrupts” that can be generated by certain external hardware conditions (e.g. a button being pressed) or by internal conditions (e.g. an internal timer reaching a pre-set value). In either case, the code is organized to “jump” to a certain spot where there is code written to specifically address the condition that generated the interrupt. A similar and popular use of segregating a portion of code and “jumping” to it when it is needed is subroutines (or functions) where specific code is grouped together because it performs a particular function or task that is needed repeatedly to support the overall program objectives.

Another more specific way to look at the categories of program elements is to categorize the “commands” or words of the language into the types of operations they perform. Table 7.2 shows some common terms and a categorization of program instructions with a brief description of how they are used. Note that there are a multitude of descriptions and categorizations that would be equally valid; this is just one view.

Learning to program efficiently requires understanding the meaning of the “terms” in the language, the “grammar/formatting” rules (syntax) and nuances of how instructions work together. For embedded systems programming, an appreciation of the hardware interaction is also essential.
<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math (Arithmetic/Logic)</td>
<td>There are two main sub-categories: mathematical function (e.g. add, subtract, divide, multiply etc.) and logical tests that are often done for determine how the program should proceed (e.g. is the number stored in one location larger, smaller or equal to another, logical AND or OR two pieces of data).</td>
</tr>
<tr>
<td>Program Control</td>
<td>These are commands and instructions that help determine choices in the program flow, or control how the program proceeds. These include very commonly used and powerful commands like if(), while() and for() statements that can alter which instructions are executed under certain conditions. Often uses logical tests to determine the next step or how to alter the instruction path.</td>
</tr>
<tr>
<td>Data Handling, I/O</td>
<td>Commands or instructions that facilitate reading data from one place (e.g. an input pin or port, a memory location, an internal register) or write data out to any port, memory location or register, or move data between any of these locations and others.</td>
</tr>
<tr>
<td>Initialization and setup</td>
<td>Sets registers and conditions that prepare the system for the environment and tasks it will be required to do. This could include making an interface (pin or port) either an input or output, setting power options, clock rates etc.</td>
</tr>
<tr>
<td>Libraries / Extensions</td>
<td>A computer language, especially a high-level language, usually has a standard “vocabulary” that can be understood and processed (i.e. debug, generate code…) within the IDE. Libraries or extensions are packages of additional code for the IDE that provides additional vocabulary, often to handle specific operations (e.g. handling LCD or keyboard i/o, network communications via specific protocols etc.)</td>
</tr>
<tr>
<td>Interrupts</td>
<td>Interrupt capability is typically built into the hardware at fabrication, and the instructions related to its use consist mostly of enabling, disabling, resetting or assigning code to execute when they occur.</td>
</tr>
<tr>
<td>Timers</td>
<td>As with interrupts, usually built into the hardware at fabrication and set up at initialization for different purposes. Timers often trigger “internal” interrupts that signal the start or completion of a more complex task (e.g. generating a PWM output. μCs often include a “watch dog timers” (WDT), which is intended to reduce the possibility of the μC “freezing” because it is caught in a loop and cannot exit. The WDT will reset the CPU if it doesn’t get confirmation that the program is running normally at specified intervals.</td>
</tr>
<tr>
<td>Subroutines/Functions</td>
<td>These are segments of code that are used repeatedly and often and are “called” when needed to perform a specific function (e.g. service an interrupt or multiply two numbers etc.)</td>
</tr>
</tbody>
</table>
8 Motors and Generators

Motors and generators are both electro-mechanical devices that facilitate the conversion between mechanical and electrical power. This Section will discuss the basic electro-magnetic principles behind the operation of motors and generators. Although this will be done in the context of DC motors and generators, the same principles apply to AC motors and generators. It builds on some of the same principles previously discussed when considering inductors and transformers.

8.1 Electromagnetic Principles Applied to Motors and Generators

Recall the two sides of the relationship between electricity and magnetism:

i) a moving charge creates a magnetic field, and

ii) a magnetic field exerts a force on a moving charge

These principles from elementary physics are the basis for the design of electric motors and generators.

A brief review of the differences between electric and magnetic field effects on a charged particle

<table>
<thead>
<tr>
<th>Electric Fields</th>
<th>Magnetic Fields</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_e$ is in the direction of the field</td>
<td>$F_m$ is perpendicular to the field</td>
</tr>
<tr>
<td>$E_e$ will act on a particle regardless of velocity, and will accelerate a particle</td>
<td>$F_m$ only affects a charge in motion, but will not change its velocity (i.e. no acceleration)</td>
</tr>
<tr>
<td>$F_e$ does work to displace/move a charge, and the energy of the charge is affected</td>
<td>$F_m$ does no work on the charge, and the particles energy is not affected</td>
</tr>
</tbody>
</table>

Recall: Work = Force acting through a distance. In a magnetic field, $F \cdot dS$ (dot product) $= F \cdot vel \, dt = 0$, i.e. no work is done since there is no force in the direction of the motion.

To explore $F_m$ a little further, from physics we know:

$$F_m \propto q, \, \text{vel}, \, B,$$

where: $B =$ strength of the magnetic field (Teslas = Wb/m$^2$),

$\text{vel}$ is the velocity of the particle(s) (m/s), and

$q =$ the magnitude of the charge on a particle (coulombs).

Also, $F_m$, the force due to the magnetic field, is perpendicular to both $\text{vel}$ and $B$.

$$F_m = q \cdot \text{vel} \times B \quad (8.1)$$

Now consider the charged particles in a wire in the presence of a magnetic field of concentration $B$. If the wire (charges) and the magnetic field are static, there is no force or effect on the charged particles. However, if the charged particles are all moving with a velocity of $v_{drift}$, (i.e. a current is flowing) then the force on each one is given by

$$F_m = q \cdot \text{vel} \times B \quad (8.1)$$

above, where $q$ is the magnitude of the charge on one particle, $\text{vel}$.
is its velocity and $B$ is the density of the field. The total number of charge carrying particles in the wire is equal to $nAl$, where $n =$ number of particles/m$^3$, $A =$ cross-sectional area in m$^2$, and $l$ is the length of the material in m.

Then the total force (due to the magnetic field, $B$) on the wire is:

$$F_{m\text{-total}} = (q \cdot vel \times B)(nAl) \quad (8.2)$$

where $q$ is the magnitude of charge per carrier in coulombs/carrier

$n$ is the density of the charge carriers per m$^3$

$A$ is the cross-sectional area in m$^2$, and

$\nu_{drift}$ is the velocity of the charges in m/s.

and recall that current is the amount of charge moving past a given point in a given time:

$$I = qnAv_{drift} \quad (8.3)$$

(Note: that all units in this expression cancel to leave coulombs/sec, and 1 C/s = 1 ampere.)

Substituting $I = qnAv_{drift}$ (8.3) into $F_{m\text{-total}} = (q \cdot vel \times B)(nAl)$ (8.2) results in the expression for the force on a wire in a field, $B$:

$$F_{mag} = Il \times B \quad (8.4)$$

where $I$ is the current in A,

$l$ is the length of the wire in the filed in meters, m,

and

$B$ is the magnetic field density in Teslas, T.

Now let us consider the effect of Faraday’s Law, given that we have a wire moving perpendicular to a magnetic field. Faraday’s Law predicts that a voltage will be induced in the wire that is proportional to the total “flux linkages” that are being “cut” (i.e. the velocity of the wire is perpendicular to the field) by the moving wire. The rate at which linkages are being “cut” is proportional to i) the length of the wire, ii) the velocity of the wire and iii) the strength of the field (i.e. the amount of flux per area).

We can develop an expression for the magnitude of the voltage if we consider the situation depicted in Figure 73 where a disconnected wire of length $l$ is moving through a magnetic field of strength $B$ with a velocity $vel$.

Recall (equation $F_m = q \cdot vel \times B$ (8.1)) the force on a moving charge in a magnetic field is: $F_m = q \cdot vel \times B$. Also recall, that the force on a charge in an electric field is $F = qE$.  

©Denard Lynch 2012-2014
For a natural balance to occur without a “current” flow, the force due to an electric field must balance the force due to the magnetic field, or:

\[ qE = q \cdot \text{vel} \times B, \]

and since the voltage, \( V \), observed over a length, \( l \), is \( E_l \).

Multiplying both sides of the equation above by \( l \):

\[ qV = q(E_l) = ql \cdot \text{vel} \times B, \]

and removing the common “\( q \)” from each side we are left with the general expression for motional emf:

\[ V = E_{\text{motional}} = l \cdot \text{vel} \times B \quad (8.5) \]

As in the case of the force on a moving charge, this is a cross product expression, indicating that the vectors involved must be orthogonal to each other. If we apply the Right Hand Rule (RHR) for cross products to this situation, the resultant “vector”, \( E_m \), can be thought of as a “battery” or “source” where the +ve direction of the vector points towards the positive terminal (i.e. indicates the +ve polarity).

If the bar in the sketch above was a 1m long conductor moving through a field of density 1T pointing into the page as shown, with a velocity (perpendicular) of 1m/s, the motional emf, \( E_m \) would be 1V, with the positive end on the top as indicated.

Now if an electric circuit was connected to this “battery” or generator, current would flow according to \( \Omega \)’s Law. However, once a current is flowing, we have our other relationship (equation \( F_{\text{mag}} = I \times B \) (8.4)) to contend with.

Equations \( F_{\text{mag}} = I \times B \) (8.4) and \( V = E_{\text{motional}} = l \cdot \text{vel} \times B \) (8.5) are the working relationships that will let us explore the basic principles of operation of motors and generators. We will start by using a “bar & rail” model to exercise the formulas.

### 8.1.1 Application to Motors and Generators

These principles, as applied to motors and generators, can be illustrated by using the common physics approach of a “bar & rail” system.

**Example 8-1: A Motor/Generator Bar & Rail Example (2006 midterm)**

Consider the bar&rail system shown in the diagram below to answer the following questions. Assume that the mass of the bar is negligible and that it is resistance-less. The rails are 1m apart, and the field density is 1T.
a) The switch has been in position 1 (as shown) for long enough for the system to reach steady state. i) Determine the velocity and direction of the bar. ii) Is it acting as a motor or a generator?

Note that an initial current of 20A (20V/1Ω) flowing downward through the bar will cause an $F_{\text{mag}}$ to the right (“Right Hand Rule”). This motion will induce a voltage in the bar in opposition to the battery. The system will reach a steady velocity when the velocity (and therefore $E_M$) is enough to limit the current so that $F_{\text{mag}} = Mg$.

$$Mg = F_{\text{mag}} = I \times B; \text{ rearranging: } I = \frac{Mg}{l \times B} = \frac{kg(9.81 m/s)}{(1m)/(1T)} = 9.81A$$

Using KVL: $E_m + (9.81A)(1Ω) - 20V = 0; \ E_M = 10.19V$

$$E_M = l \cdot \text{vel} \times B; \ vel = \frac{E_M}{l \cdot B} = 10.19 m/s \text{ to the right.}$$

Since the electrical source is delivering power to the bar and the mass is being lifted (mechanical work being done), the system is acting at a motor.

b) The switch is now moved to position 2 and the system has once again reached a steady-state. i) What is the direction and velocity of the bar? ii) Is it acting as a motor or a generator? iii) Do a complete power balance to prove that energy is conserved. (Be sure to specify all the mechanical sources/sinks.)

With no external source of current, the mass will pull the bar to the left until the induced $E_M$ is large enough to drive enough current through the other 1Ω resistor to generate a balancing $F_{\text{mag}}$.

Again, at steady state, $Mg = F_{\text{mag}} = I \times B; \ I = \frac{Mg}{l \times B} = \frac{kg(9.81 m/s)}{(1m)/(1T)} = 9.81A$

Using KVL: $E_m + (9.81A)(1Ω) = 0; \ E_M = -9.81V$

$$E_M = l \cdot \text{vel} \times B; \ vel = \frac{E_M}{l \cdot B} = 9.81 m/s \text{ to the left.}$$
Since there is only a source of mechanical power (the mass dropping), the system is now acting as a [generator].

Power:

**Mechanical:**
- **Source:** Mass falling providing force $Mg = 9.81\text{N}$.
  
  $P = F(\text{vel}) = (9.81\text{N})(9.81\text{m/s}) = 96.2\text{W}$
- **Sink:** the bar being pulled with a force, $F_{\text{mag}} = 9.81\text{N}$.
  
  $P = (9.81\text{N})(9.81\text{m/s}) = 96.2\text{W}$

**Electrical:**
- **Source:** Bar (generator) $(E_m)(I) = (9.81\text{V})(9.81\text{A}) = 96.2\text{W}$
- **Sink:** Resistor heating $I^2R = (9.81)^2(1\Omega) = 96.2\text{W}$

### 8.2 Motor and Generator modelling

While the Bar&Rail model is useful for applying the basic physics principles, we will use an alternate model to explore the relationship between voltage, current and power in DC motors and generators.

#### 8.2.1 A Simple Motor Model

Consider the simple model in Figure 8.2:

![Figure 8.2: DC Motor Model](image)

The source provides electrical power which the motor converts to mechanical power. As in all cases, the power inputs and outputs must balance. If we assume that all the electrical inefficiencies are represented by the $I^2R$ losses in $R_{\text{int}}$, then we make the following observations, and calculations, about the electrical power:

a) The total power in is $P_{\text{in}} = (E_s)(I_s)$

b) The electrical losses are $P_{e-loss} = I^2R$

c) The electrical power remaining to be converted to mechanical power:

$d$) The “motor voltage” $V_M = k\omega$, where $k$ is the motor constant, $\omega$ is the shaft speed of the motor in Radians/second

e) The maximum mechanical power: $P_{\text{max-mech}} = P_{e-out} = (k\omega)(I_s))$

f) The mechanical output power: $P_{m-out} = \tau\omega$
The mechanical losses (friction etc.): \( P_{\text{m-loss}} = P_{\text{max-mech}} (P_{\text{e-out}}) - P_{\text{m-out}} (\tau \omega) \)

Any difference between the maximum mechanical power and the actual output power would be due to mechanical losses (e.g. friction). A simple example will illustrate its utility.

**Example 8-2: Motor Example using the Simple Model**

Consider a small DC electric motor represented by the model in Figure 8.2. It is powered by a 12V battery and draws 200mA and spins at 8000 rpm with no load. When the motor is stalled \((\omega=0)\), it draws 2A of current. The “voltage constant”, \(k\), for this motor is .013 V-s/R. a) What is the output power if the motor is loaded so it draws 1A of current? b) What is the output torque?

Make the assumption that any mechanical losses are constant at any motor speed.

a)

When the motor is stalled \((\omega=0)\), the motor internal voltage is 0V, and the internal resistance is:

\[
R_{\text{int}} = \frac{E_{\text{s}} - 0V}{I_s} = \frac{12V}{2A} = 6\Omega
\]

At 8000 rpm, the loss in the electric side is \(I^2R = (.2A)^2(6\Omega) = .24W\). The overall power supplied by the source is \(P_{\text{in}} = (12V)(.2A) = 2.4W\). Since .24W is lost on the electrical side, that means the other 2.16W must be wasted in mechanical losses since there is no actual load on the motor. Since we made the assumption that this would be constant at all rpm, we can use this as the mechanical loss when the motor is loaded as well.

If the motor is loaded so that is draws 1A, this means the motor voltage is 6V:

\[
\frac{E_{\text{s}} - V_M}{R_{\text{int}}} = I_s ; \quad \frac{12V - V_M}{6\Omega} = 1A ; \quad V_M = 6V
\]

Since \(V_M = k\omega\), \(\omega = 462R/s \ (= 4407\text{rpm})\).

The output power is simply the input power minus the losses:

\[
P_{\text{out}} = (12V)(1A) - (1A)^2(6\Omega) - 2.16W = 3.84W
\]

Note that the maximum mechanical power would be \(V_MI_s = 6W\), then minus the 2.16W mechanical losses, which results in exactly the same output power of course.

b)

\[
\tau = \frac{P_{\text{out}}}{\omega} = \frac{3.84W}{462R/s} = 8.31mN \cdot m
\]

**8.2.2 A Simple Generator Model**

We can consider a similar, model for generators. In fact, both motors and generators follow the relationships given by equations \(F_{\text{mag}} = I \times B\) (8.4) and
Both involve a force due to a current in a magnetic filed, and both involve an induced voltage. Consider the generator model shown in Figure 8.3.

The input is a source of mechanical power which the generator converts to electrical power. Again, the power inputs and outputs must balance. If we again assume that all the electrical inefficiencies are represented by the $I^2R$ losses in $R_{int}$, and the mechanical losses are constant at any generator speed, then we make the following observations, and calculations:

a) The total power in (mechanical) is $P_{m-in} = \tau \omega$

b) The mechanical losses (friction etc.): $P_{m-loss} = P_{mech} (\tau \omega) - P_{e-convert}$

c) The electrical losses are $P_{e-loss} = I^2R$

d) The mechanical power remaining to be converted to electrical power, $P_{m-convert}$ is:

$$P_{e-convert} = \tau \omega - P_{m-loss}$$

e) The “generator voltage” $V_G = k \omega$, where $k$ is the generator constant, $\omega$ is the shaft speed of the generator in Radians/second

f) The total converted electrical power: $P_{m-in} - P_{m-losses}$

g) The electrical output power $P_{e-out} = (k \omega I_L) - (R_{int}I_L^2)$

Consider another illustrative example.

**Example 8-3: An Example using the Simple Generator Model**

A DC generator has a 5cm radius drum fixed to its input shaft with a fine rope wound around it. The generator shaft is turned by unwinding the rope. The generator’s internal resistance has been measured and found to be 1.2Ω. Mechanically, the generator is 85% efficient (i.e. 85% of the input power is available to be converted to electricity). The load connected to the generator is a 12V, 35W headlight. The generator is also tested at 100 rpm and no load and the terminal voltage is 4V.

The rope is connected to the owner’s dog. How fast must the dog run away with the rope to provide the required voltage for the headlight? How much force must the dog provide while running to power the headlight to full brightness (i.e.35W)? (Note: no animals were harmed in the execution of this example!)
From the data given, we can determine the “generator constant, \( k \). With no load connected, there is no voltage drop between \( V_G \) and the output terminals, so:

\[
k = \frac{V_G}{\omega} = \frac{4V}{\left(100 \text{ r/min}\right)\left(2\pi \text{R/min}/r\right)\left(1 \text{ min}/60s\right)} = 0.382\frac{V_s}{R}
\]

Full load current (35W@12V) = 35/12 = 2.92A. At this current, the voltage drop across the internal resistance is (2.92A)(1.2Ω) = 3.50V. This means we need a generated voltage of 12+3.5 = 15.5V so we have 12V at the load.

\[
\omega = \frac{V_G}{k} = \frac{15.5V}{0.382\frac{V_s}{R}} = 40.6 \frac{R}{s}
\]

From this we can determine how fast the dog must run. The rope must “unwind” at a rate that will turn the shaft at 40.6R/s. The velocity of a point on a circle of radius \( r \) rotating at \( \omega R/s = \omega r \) which is the velocity at which the rope must “unwind”, or (40.6R/s)(.05m) = 2.03m/s (=7.31 km/hr. For you runners, this is about a 8minute kilometre pace.). Probably a reasonable pace for our best friend, but how much force must it expert while doing this?

There are two ways to figure this out, but first we need to figure out the input power required:

\[
P_{e-total} = 35W + (2.92A)^2(1.2\Omega) = 45.2W, \text{ and } P_{in} = \frac{45.2W}{.85} = 53.2W
\]

1) First way: \( P = F \cdot vel \), and \( F = \frac{P}{vel} = \frac{53.2W}{2.03m/s} = 26.2N \)

2) Second way: \( P = \tau \omega = \left(F \cdot r\right) \omega \), and \( F = \frac{P}{r \omega} = \frac{53.2W}{(0.05m)(40.6 \frac{R}{s})} = 26.2N \)

26.2 Newtons is the force it would take to hold up a 2.7kg mass; not unreasonable for our furry friend, provided it’s had it’s Purina that day…

A summary of motor and generator function with respect to the power input and output is shown in Figure 8.4
Figure 8.4: Energy Sources and Sinks in Motors and Generators
9 Batteries: Electro-chemical Energy Storage and Conversion

While energy has been stored in mechanical form for a very long time, storing electrical energy wasn’t practical until the storage battery was developed although there is some evidence that this may have occurred over 2000 years ago! (Buchmann). Alessandro Volta, after whom the term “Volt” is named, is generally credited with the invention of the first chemical battery (circa 1800). In a sense, a battery is a “converter”, converting the energy in chemical bonds to electricity when it is supplying energy (discharging), and absorbing and storing electrical energy in chemical bonds when it is being charged. In some cases the chemistry is not practically reversible, so the battery acts more like a generator than a storage device, using a chemical or molecular reaction (fuel) to generate an electrical potential and deliver electrical energy. In this sense disposable batteries (i.e. primary cells) and fuel cells are alike; they can generate electrical energy but can’t practically convert electrical energy back into the same chemical or molecular form. A “storage battery” (secondary cell or rechargeable) is the technical term that is generally used to describe a device that can “convert” electrical energy into another form that can be temporarily stored and then extracted at a later time.

9.1 Battery Terminal Voltage

What is commonly referred to as a “battery” is really a “battery of cells”. A battery can contain only one cell, but are often packaged or used together to provide a desire terminal voltage or capacity by connecting the cells in series / parallel combinations. Each cell has a characteristic voltage and capacity depending on the materials used (i.e. the chemistry: anode, cathode and electrolyte materials) and the physical size. The actual terminal voltage of a cell can vary depending on the “state of charge” or chemical conditions. For instance, a “fully charged” lead acid cell (e.g. used in car batteries) can have a voltage up to ~2.1V. However, as the chemical reaction progresses and electrical energy is extracted, the voltage gradually decreases, ultimately to a very low voltage if the reaction is allowed to proceed almost to its limit. In practice, batteries are usually recharged before their terminal voltage drops to the point where it will no longer supply the energy needed to satisfy the demand placed on it by attached equipment or circuitry. Also in practice, the rate at which the chemical reaction can proceed, and thus the rate at which current can be supplied, also reduces as the chemical state changes. Both of these degradations are specific to the materials used and discharge rates and are relevant characteristics to consider when selecting an appropriate energy source. Isador Buchmann gives the nominal voltage for a number of common battery chemistries, summarized in Table 9.1.
Table 9.1: Key Battery Parameters

<table>
<thead>
<tr>
<th>Type</th>
<th>Nominal Voltage per Cell</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbon-zinc</td>
<td>1.5</td>
<td>Primary cells (not rechargeable). Common “dry cells” used in many “throw away” applications. There are 6 in series in a 9V battery. They are also packaged in other terminal voltages (i.e. # of cells)</td>
</tr>
<tr>
<td>Alkaline</td>
<td>1.5</td>
<td>An improved replacement for carbon-zinc with the same nominal terminal voltage. Some alkalines can be recharged under some conditions. For normal alkalines this is not recommended.</td>
</tr>
<tr>
<td>Silver-oxide</td>
<td>1.6</td>
<td>Another “primary” cell</td>
</tr>
<tr>
<td>Lead-acid</td>
<td>2.1</td>
<td>There are 6 in series to make a 12V car battery. Nominal full charge voltage would be 12.6V. These have a long history, but the disadvantage of using liquid sulphuric acid as an electrolyte. “Gel cells” try to mitigate this problem by using an electrolyte that is in a gel form so the orientation of the battery is more forgiving. Other variations have also been developed.</td>
</tr>
<tr>
<td>NiCad (nickel-cadmium)</td>
<td>1.2</td>
<td>Very popular rechargeable battery for portable applications. Maintains it terminal voltage well through most of its discharge range.</td>
</tr>
<tr>
<td>NiMH (nickel-metal hydride)</td>
<td>1.2</td>
<td>An improvement over NiCd, with a higher energy density and fewer environmental concerns with the materials used.</td>
</tr>
<tr>
<td>Lithium</td>
<td>3.0</td>
<td>Another primary (not rechargeable) cell with a high energy density and attractive discharge and storage characteristics.</td>
</tr>
<tr>
<td>Lithium-ion</td>
<td>3.6</td>
<td>Currently very popular in the portable market, especially as costs come down and performance improves. Very high energy density and other favourable characteristics. Many material variations, but all similar in operation.</td>
</tr>
</tbody>
</table>

9.2 Battery Capacity

“Battery capacity” refers to the total amount of energy can be stored in, or, more critically, can be extracted from the battery. The common unit of measure is the Ampere-hour or Amp-hour (A-h), or for smaller batteries, the milliampere-hour (mAh). This figure in itself does not provide enough information for some applications, and there are other characteristics that indicate how the battery will perform under certain circumstances. “Cold cranking amps” (CCA) is one example commonly cited for automotive batteries. It is an indication of the internal resistance of the battery and its reserve capacity and thus the maximum current it can supply for starting an engine in
cold conditions. For example, a CCA of 550A means a battery should be able to supply a cranking (starter motor) current of 550A for at least 30 seconds at a temperature of \(-18^\circ\text{C}\) without the terminal voltage dropping below 7.2V.

Another figure of merit for batteries in their energy density, usually given in Watt-hours per kilogram (Wh/kg). Figure 9.1 gives an example of energy densities for 3 common types of batteries (lead-acid, nickel-based and lithium-based). In general, a higher energy density is more expensive to produce, but in applications where less weight and smaller size are desirable, the extra cost is an acceptable design trade-off. High energy density battery chemistry like the lithium family have made small devices like mp3 players and cell phones commercially practical and attractive. Finally, getting the stored energy out of the battery depends on a number of additional factors such as the age and condition of the battery, the temperature and the discharge rate. Manufacturers usually specify energy content and discharge capabilities for a new battery under optimum conditions. In practice, these specifications must be de-rated to account for the conditions encountered in the particular application in which the battery is used, and for the normal “aging” of the battery through use.

### 9.3 Battery Charging and Discharging (Cycling)

As mentioned above, a number of factors affect the total amount, and the rate at which energy can be extracted from a battery. There are also factors affecting battery life (how many times it can effectively be discharged and re-charged before its capacity is significantly compromised).
9.3.1 Discharging:

In general, the amount of energy that can be extracted from a fully charged battery decreases as the discharge rate increases. For modern rechargeable batteries however, the total energy available is quite consistent over a fairly broad range of discharge rates. Exceeding the manufacturer’s maximum discharge rate for a battery will reduce both the total energy available significantly and reduce the life of the battery (i.e. the subsequent maximum capacity). Recommended discharge and charge rates are usually given in term of the battery capacity, $C$. For instance, if the battery capacity is given as $1500mAh$, and the charge rate is given as $1C$, the battery can be charged at a rate of $1500mA$ ($1.5A$) and, all conditions being ideal, be recharged in a maximum of 1 hour. If the recommended discharge rate is given as $C/10$, the battery would be expected to perform to capacity specifications if the load current was $150mA$ ($1500mA/10$). While a battery is discharging, its terminal voltage drops. This is typically used as a measure of the State of Charge (SoC; the fraction or percentage of maximum capacity remaining), in order to manage the battery properly when it is being re-charged. Unfortunately, various conditions, including the discharge rate, affect the terminal voltage, complicating the determination. Figure 9.2 illustrates the variation in terminal voltage at different states of charge for a lead-acid battery. Note that the discharge rate ($C/x$) significantly affects the terminal voltage even at the same SoC.

![Graph showing battery voltage vs state of charge](image)

**Figure 9.2: Battery Terminal Voltage (V) vs State of Charge (%) (mathscinotes)**

Each time a battery is discharged, its capacity decreases. Its remaining capacity after some time in service is the best indicator of battery “health”, but not an easy assessment to make. A battery’s “life” is typically defined as the point at which it cannot be recharged to an acceptable capacity (compared to its original capacity). Manufacturers
will also estimate the number of “charge-discharge” cycles that can be expected before the capacity suffers significantly. Various factors affect the capacity and the cycle life. Figure 9.3 shows the effect of “Depth of Discharge” (DoD; how much of the capacity is drawn from the battery before it is re-charged) for lead-acid batteries. Note that the expected cycles is quite low if the battery is totally discharged each time. If only half its capacity is discharged each cycle, the number of expected cycles is approximately 10 times more. Simple math shows that the total energy that can be delivered from the battery if it is re-charged twice as often is at least 5 time greater! Battery “sizing” must consider such relationships in order to provide an optimized, economic solution. (Note: a “Deep-cycle Battery” is one that is constructed to reduce the aging effects of deep discharges. This advantage, of course, comes at a cost, but may be a reasonable trade-off in circumstances where frequent charging is not possible.)

<table>
<thead>
<tr>
<th>Depth of Discharge</th>
<th>Starter Battery</th>
<th>Deep-cycle Battery</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>12–15 cycles</td>
<td>150–200 cycles</td>
</tr>
<tr>
<td>50%</td>
<td>100–120 cycles</td>
<td>400–500 cycles</td>
</tr>
<tr>
<td>30%</td>
<td>130–150 cycles</td>
<td>1,000 and more cycles</td>
</tr>
</tbody>
</table>

Figure 9.3: Depth of Discharge vs Life Cycles for Lead-acid Batteries (Buchmann)

Figure 9.4 shows similar data for a particular Lithium-ion battery. While not as drastic as for the lead-acid type, it is still advantageous to re-charge as often as practical. While this may seem to apply generally to secondary batteries, an exception is some types of nickel-based batteries where deeper discharges will optimize useful battery life. This is because of what is know as the “memory effect”: the reduced ability for the battery to discharge past a previous depth of discharge point (SoC).

<table>
<thead>
<tr>
<th>Depth of discharge</th>
<th>Discharge cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% DoD</td>
<td>300 – 500</td>
</tr>
<tr>
<td>50% DoD</td>
<td>1,200 – 1,500</td>
</tr>
<tr>
<td>25% DoD</td>
<td>2,000 – 2,500</td>
</tr>
<tr>
<td>10% DoD</td>
<td>3,750 – 4,700</td>
</tr>
</tbody>
</table>

Figure 9.4: Depth of Discharge vs Life Cycles for Li-ion Batteries (Buchmann)

Other conditions can also affect battery performance. Temperature, for example can have an effect on capacity, discharge and charge performance. This varies for different battery types. The effect of temperature on battery capacity for a lead-acid battery is shown in Figure 9.4.
9.3.2 Battery Charging

The objective of re-charging a battery is to bring it back to a state of charge (SoC) at which it has stored the specified amount of energy. Re-charging a depleted battery is generally the reverse of discharging it; almost all the same considerations apply. There are, of course, variations in sensitivity to these conditions depending on the battery type. Lead-acid batteries, for example, are generally considered to have an approximately linear “terminal voltage versus SoC” curve (betterytender.com), and are reasonably tolerant to “overcharging”. A NiCad battery, in contrast, has an obviously non-linear V vs SoC relationship (Figure 9.6), and is quite intolerant to overcharging. Most current charge management systems will monitor appropriate parameters and control the charging process to protect the battery and ensure a maximum useful life.

Figure 9.5: Battery Capacity Variation with Temperature (mathscinotes)

Figure 9.6: NiCad Terminal Voltage vs State of Charge (Jaycar Electronics)
There are two general approaches to modern charging management: voltage-based and current-based. In both cases, the state of charge is monitored by observing the terminal voltage. A typical charging scenario uses a combination of the following steps:

1. Apply a specified voltage (voltage-based) or current (current-based) to the battery for a predetermined length of time
2. Stop charging and “rest” the battery for specified time
3. Discharge the battery at a certain rate for a specified time
4. “rest” again for a specified time
5. check the terminal voltage
6. resume charging or quit charging based on results.

Chargers for different types of secondary batteries will look for parameter levels specific to that type and make charging decisions accordingly. Recent “smart” batteries integrate charge management right inside the battery itself. As knowledge and experience grow, estimating SoC and managing charging are becoming more accurate and sophisticated.

9.4 Battery Life

Battery life is affected by all the considerations discussed above. The conditions that accrue that bring about the end of the useful life of a battery can be summarized as follows:

- Decreased Capacity: As indicated above, with use the maximum storable energy decreases. At some point, it is too low to provide reliable service and should be replaced.
- Increasing Internal Resistance: This is usually due to chemical changes in the internal components of the battery and will limit both charging and discharging rates.
- Increasing Self-discharge Rate: Even new batteries will have some “self discharge” or leakage that will gradually reduce the SoC over time even with no load. As batteries age and are repeatedly cycled, this self-discharge rate will increase to the point where it is excessive and the battery cannot “hold a charge” for any useful length of time.
- Premature Voltage Cut-off: This is the result of the effect illustrated in Figure 9.2 where the terminal voltage will decline prematurely if the discharge rate is excessive. In monitored situations, this will lead to the load being disconnected before the stored energy is exploited. This is not so much an indication of the end of the life of a battery as much as the improper matching of load characteristics and battery capability.
Example 9-1: Battery Sizing Example 1:

A solar charger is used to maintain the battery that supplies a fence charger. The device draws 2A@12V for 100ms every 4s for the fence output, and 20mA continuously for the timer circuitry. A 12V lead-acid battery is to be used to provide the power, and a 24W/12V solar panel is available to re-charge the battery. We wish to limit the depth of discharge (DoD) to 10% of the battery capacity, C.

a) What is the minimum capacity, C, in A-h, for the battery?
b) How many hours of adequate sunlight, on average, are required to maintain the battery at the target state of charge (SoC)"
c) If we select a 20A-h battery for this system, and the solar charging system happens to fail, what will the SoC be after one week of not charging?

a) The load consists of:
\[(20\text{mA})(24\text{hrs}) = 0.48\text{A-h}\]
\[(2A)\left(\frac{1s}{4s}\right)(24\text{hrs}) = 1.20\text{A-h}\]

**Total:** \(1.68\text{A-h}\)

10% DoD = \(0.1C\); \(\frac{0.1C}{1.68\text{A-h}} = \frac{C}{x}\); \(C = 16.8\text{A-h}\)

b) The load draws 1.68A-h per day at 12V, and that much must be restored each day. Since we have a 24W/12V solar panel, it can supply 2A at the MPP (and with sufficient insolation). The time required to replenish 1.68A-h is: \(\frac{1.68\text{A-h}}{2A} = 0.84\text{hr} = 50.4\text{min}\)

c) The capacity required to run the system for 7 days is: \((1.68\text{A-h})(7)=11.76\text{Ah}\). The state of charge is the energy remaining in the battery, so

\[SoC = \frac{20 - 11.76}{20A - h} = .412 = 41.2\%\]
10 Photovoltaic Electrical Generation

Exploiting solar energy to replace other forms of energy generation hold some promise. While solar energy can be used directly as heat, this discussion is limited to its use to generate electrical power. This type of electrical generation is called photovoltaic. The mechanism of electrical generation from solar radiation exploits the property of some materials to “release” electrons when “excited” by photons. A ‘p’ and ‘n’ doped silicon wafer is commonly used in solar cells. Like batteries, they have a terminal voltage that is characteristic of the material used (approximately 0.5V for silicon) and the cells may be arranged in series and parallel arrays (or modules) to obtain desired terminal voltages (in series) and current capacities (in parallel). Like batteries, they “convert” energy form one form to another (solar to electrical in this case), but unlike batteries, that do not require re-charging and can only be “fuelled” by photon energy. In other ways, “photovoltaic arrays” have unique characteristics that must be appreciated in order to utilize them properly as sources of electrical power.

10.1 Characteristics and Specifications

Photovoltaic (PV) arrays are characterized primarily by their output characteristics. Table 10.1 summarizes the key parameters, usually given for standard test conditions and assuming a specified insolation (commonly 1000W/m²).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Power</td>
<td>The maximum power output that can be obtained from the array under optimum illumination conditions. The output power depends on the “operating point” on the array’s I-V curve. While other operating points are useable, one specific voltage and current combination will lead to the maximum power output.</td>
</tr>
<tr>
<td>V&lt;sub&gt;MP&lt;/sub&gt;</td>
<td>The terminal voltage at the maximum power point on the I-V curve</td>
</tr>
<tr>
<td>I&lt;sub&gt;MP&lt;/sub&gt;</td>
<td>The available load current at the maximum power point on the I-V curve</td>
</tr>
<tr>
<td>V&lt;sub&gt;OC&lt;/sub&gt;</td>
<td>The maximum terminal voltage with no load attached to the array (I&lt;sub&gt;O&lt;/sub&gt;=0)</td>
</tr>
<tr>
<td>I&lt;sub&gt;SC&lt;/sub&gt;</td>
<td>The maximum current available with the output terminal shorted (V&lt;sub&gt;O&lt;/sub&gt;=0)</td>
</tr>
</tbody>
</table>

Note that in the normal operating range, the terminal voltage does not vary significantly once a certain minimum threshold of illumination is present, but the current output is directly proportional to the amount of available illumination up to a certain maximum.

10.1.1 Factor Affecting Performance

There are a few key factors that will affect the performance of a PV array:
1. **Insolation**: This is simply the amount and intensity of the sunlight striking the array, which is proportional to the “strength” of the sun. In turn, the strength of the impinging sunlight is dependent on:
   a. The incident angle of the sun’s rays. Optimum performance is obtained when the incident illumination is perpendicular to the cells.
   b. Distance from the source (sun). This varies with latitude of the location and the season. For example, geographical points in the northern hemisphere are closer to the sun in the summer, and those further south are closer than those further north.
   c. Atmospheric and local conditions. The sun’s energy can be diffracted and absorbed by water vapour, pollution and dust and shaded by clouds, leaves, buildings etc.

   Insolation data is available in tabular form for various geographic locations. It is usually given in units of kW-h/m²/day.

2. **Load resistance**: The load resistance will determine the operating point on the array’s I-V curve, and thus the power than can be delivered. Appropriate array and load matching (design) will optimize performance.

3. **Cell Temperature**: Array power output is usually specified at a certain cell temperature (note cell not ambient temperature). $V_{oc} \propto \frac{1}{T}$, and $V_{oc}$ will typically decrease approximately 0.5% for every $1^\circ C$ rise in cell temperature. This must be considered to ensure that the terminal voltage does not fall below that required to deliver the required power.

Referring to the typical I-V curve shown in Figure 10.1, note that $V_{MP}$ and $I_{MP}$ bound a rectangle that is the maximum power. Any point on this curve will define the I and V bounds that define the power delivered at that point, and the point on the curve that results in the largest area of rectangle is the maximum power point.

![Figure 10.1: Typical PV I-V Curve (Solar Energy International)](image)

Like all energy alternatives, photovoltaic sources have advantages and disadvantages. Table 10.2 provides a quick summary including points from (Solar Energy International).
### Table 10.2: Photovoltaic Advantages and Disadvantages

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliable</td>
<td>Cost</td>
</tr>
<tr>
<td>Durable</td>
<td>Efficiency</td>
</tr>
<tr>
<td>Low maintenance</td>
<td>Variability of supply (needs illumination)</td>
</tr>
<tr>
<td>No fuel costs</td>
<td>Requires energy storage for low light times</td>
</tr>
<tr>
<td>Quiet operation</td>
<td>Limited knowledge/experience with PVs</td>
</tr>
<tr>
<td>Modular</td>
<td>Environmental issues with materials and manufacture</td>
</tr>
<tr>
<td>Safe (no combustible fuels)</td>
<td></td>
</tr>
<tr>
<td>High altitude performance</td>
<td></td>
</tr>
</tbody>
</table>
Example 10-1: PV Example 1: based on an example by (Solar Energy International)
A PV system has available 8 PV arrays, each capable of 24V, 10A, and 8 storage batteries, each 12V, 160Ah. There is also a charge controller available to manage the charging of the batteries. The system is to supply 48VDC to a constant load. Determine the wiring connections and system capacity.

Solution: The combination of series/parallel connections shown will provide the required 48VDC on the array side as well as a 48VDC output from the battery bank.

Note: V adds in series; I and A-h add in parallel. 2 PV modules in series will provide 48V, 10A. 4 of those in parallel will still be 48V, but 4 X 10A = 40A out.
4 batteries in series will provide 48V, 160 A-h. 2 of those in parallel boost the capacity to 320Ah@48V

Let’s examine how much “constant load” this arrangement can supply.

Consulting some tables for insolation data, we can find some data for Minot ND, where the minimum and maximum values are 3.4 and 9.7 kWh/m$^2$/day respectively, assuming the arrays are always kept perpendicular to the sun. These figures are the total solar energy available (on average) per day. We can interpret this as being equivalent to having a minimum of 100W/m$^2$, the array specification standard, for 3.4 or 9.7 hours per day on the worst and best months.

This means each array can provide $(24V)(10A)(3.4hr) = 816$W-h/day or 6.528kWh/day for all 8 in the worst month. $(18.624$kWh/day in the best month)

Constant load capability = \[
\frac{6528Wh}{24h} = 776W
\]

This of course assume a 100% efficiency throughout the system, and would have to be adjusted for inefficiencies in the batteries, solar tracking electrical distribution or conversions.

**Example 10-2: PV Example 2**

A photovoltaic array has the following specifications:
- Maximum Power: 24W@ 1000W/m$^2$ insolation
- $V_{MP} = 12V$
- $I_{MP} = 2A$
- $V_{OC} = 16V@25^0C$
- $I_{SC} = 2.5A$

a) What is the open circuit voltage if the cells are operated at 80$^0C$?

b) If the array is 40cm X 50cm, what is the solar conversion efficiency?

c) What is the capacity (in A-h) of the array if we assume the equivalent of 7 hours of sufficient insolation per day on average?

d) How much energy per day can be harvested (at the MPP)?

**Solution:**

a) $V_{OC}$ decreases at .5%/C rise. $V_{OC}$ is spec’d at 16V@25$^0C$. At 80$^0C$ it will be:

$V_{OC} = V_{25} \left(1 - \left(\frac{80}{25}\right)\left(0.005\right)\right) = 11.6V$

b) Input power is:

$P_{in} = 1000W/m^2 \times (4m)(.5m) = 200W$ : $\eta = \frac{P_{out}}{P_{in}} = \frac{24W}{200W} = 12\%$

c) It will produce a maximum of 24W/hr for 7 hours per day, so 168A-h@12V

d) $W = Pt = (168A=h)(12V) = 2.016kW-h$
11 Wind Powered Electrical Generation

Like solar energy, the energy in the wind has been utilized by people for thousands of years. Sailing ships and windmills are two obvious examples that have been used long before electricity was harnessed. Conversion of wind energy into electrical energy is a relatively new concept, having gained popularity only over the last 100 years or so. “Wind chargers” were popular on prairie farms early in the last century, and the remnants of the towers used to support them are still seen occasionally in the Saskatchewan landscape.

Wind has some similarities and some differences to solar energy as an alternative to traditional forms of generations like fossil fuels, hydro-electric or thermo-nuclear generation. Like solar, it needs an energy input, but does not require a fuel as such. The source of energy input is also irregular, thus requiring energy storage to provide a continuous supply. Also like solar, our knowledge and experience are still somewhat limited. Unlike solar, noise is a concern, and it is more difficult to distribute the generation because a specific locale and significant physical structures are often required. Further, good geographic locations for wind harvesting are usually away from users so transmission lines can be an issue and added expense. In operation, wind generators do not generate greenhouse gasses, but noise pollution and effects on wildlife are still being studied.

The convenience of electrical energy and improved knowledge and technology have stimulated the development of the wind generation industry as we continue to look for alternate sources of electrical energy to satisfy an ever increasing demand. The detailed electric and mechanical details of generators specifically designed for wind applications is beyond the scope of this course. This section will discuss the basics of wind energy and its conversion to electricity.

11.1 The Energy in Wind

The energy in the wind is basically the kinetic energy of a moving mass, and the objective of wind generation is to extract some of that energy and convert it to electricity. The final conversion to electrical form is usually done using a mechanical generator which is basically the same as other mechanical generators and not specific to wind generation. The basic considerations related to energy extraction from the wind will be discussed here.

The kinetic energy in a moving mass is:

\[ W = \frac{1}{2}mv^2 \]  

Equation 11.1

where \( W \) is the energy in Joules, \( m \) is the mass in kg and \( v \) is the velocity in m/s.

Energy is commonly extracted from a moving air mass as it goes by the plane of the “extraction device”, a rotor (propeller) currently being the device of choice, although other alternatives are constantly under study. A rotor arrangement will extract energy from the mass of air going by a circular-shaped plane, ideally perpendicular to the wind direction. The mass of air going through this circular plane is:
\[ m = \rho (\pi r^2) v \quad (11.2) \]

where \( m \) is the mass in \( kg \), \( \rho \) is the density of the air in \( kg/m^3 \), \( r \) is the radius of the propeller arrangement in \textit{meters} and \( v \) is the velocity of the wind in \( m/s \). Note Equation (11.2) gives a result in \( kg/s \).

Substituting (11.2) into (11.1), we have an expression for the \textit{power} in a wind of velocity \( v \) passing through a circular plane of radius \( r \) (since the mass is given in \( kg/s \), we have energy/time which is power):

\[ P = \frac{1}{2} \left( \rho (\pi r^2) v \right) v^2 = \frac{1}{2} \rho \pi r^2 v^3 \quad (11.3) \]

where we note the energy is proportional to velocity\(^3\).

Of course this is the total energy in the wind within a cylindrical “stream tube” perpendicular to the plane of the rotors. Examining Equation (11.1) we note that the energy is clearly proportional to the velocity. If we remove all the energy the velocity must become 0, which is clearly not possible. If the velocity is 0 after the rotors, we must “stop” the wind, and no mass will go through the rotor plane and thus no energy can be extracted.) At the other extreme, the velocity stays the same, then the exiting energy is the same as that entering and no energy is extracted, which is of no benefit. The key is to extract \textit{some} of the energy.

Extracting \textit{part} of the energy in the wind has some additional considerations. If we assume an input velocity of \( v_i \) and an exit velocity of \( v_o \), then the \textit{change} in energy is:

\[ \Delta W = \frac{1}{2} m (v_i^2 - v_o^2) \quad (11.4) \]

The mass of the wind that passes the rotor plane is now:

\[ m = \rho (\pi r^2) v_{\text{average}} = \rho (\pi r^2) \left( \frac{v_o + v_i}{2} \right) \quad (11.5) \]

Again, combining the two equations, and reversing the \( v^2 \) terms to yield a \textit{positive} power output:

\[ P = \frac{1}{4} \rho (\pi r^2) (v_o + v_i) (v_i^2 - v_o^2) \quad (11.6) \]

A related consideration is the consistency of the stream tube. The mass of the air is the same before and after the rotors, but the velocity is reduced. Looking at Equation (11.5) we see that to keep the mass the same the radius must increase. This results in an effect somewhat like that shown in Figure 11.1.
Practically, there is turbulence created before and after the rotor plane and to the sides of the stream tube.

At both extremes (all energy extracted and no energy extracted) the power harvested is zero. By examining the power that can be harvested with various ratios of $v_o$ to $v_i$ we find that the resulting relationship has a maximum at approximately $1/3$ ($v_o\approx0.3v_i$) which harvests approximately 60% ($16/27$ to be precise) of the available wind energy. This is known as Betz’ Law, after the German physicist who identified and published this relationship in 1926.

The proof of Betz’ Law is based on what is described as the “reasonable assumption” that the effective mass passing through the plane of extraction (i.e. rotor) may be represented by using the average velocity of the wind: $\left(\frac{v_i + v_o}{2}\right)$, as shown in Equation 11.5. If we consider an expression for the ratio of the extracted power to the power in the undisturbed (incoming) wind using Equations 11.6 and 11.3 respectively, we have (after cancelling common terms in the numerator and denominator):

$$\frac{P_{\text{extracted}}}{P_{\text{max-in}}} = \frac{1}{2} \left( 1 + \frac{v_o}{v_i} \right) \left( 1 - \left( \frac{v_o}{v_i} \right)^2 \right)$$

Equation 11.7

where $\frac{v_o}{v_i}$ is simply the ratio of the incoming to exit wind velocities. Solving this expression for the maximum yields $\frac{16}{27}$ (Betz’ Law) at a velocity ratio of $\sim \frac{1}{3}$. One may rationalize that the greater the velocity difference, the more turbulence and the less effective mass is available for energy extraction. This counters the increasing kinetic energy available as the velocity difference increases, thus allowing the opportunity for a maximum.

The design of extraction devices considers this relationship and attempts to optimize energy extractions as well as balancing other design considerations.

### 11.2 Geographic Site Considerations

Designing for wind harvest obviously requires a location where there is an adequate supply of the resource. Meteorological data is a valuable source of relevant information.
along with supplemental information on wind behaviour. Three key factors will be discussed below: wind patterns, terrain effects and wind sheer.

11.2.1 Wind Patterns:
The two parameters of interest in power generation are velocity and direction. These can be summarized in a graphical form using a “wind rose”, which illustrates the percentage of the time the wind in a specific locale comes from a given direction and the energy it contains. Figure 11.2, from the European Wind Atlas, shows a typical wind rose. The rose can be divided into 6, 12 or 16 equal sectors (“petals”). The radii of the outer wedges indicates the relative frequency with which the winds comes from that direction (e.g. if the radius in one direction is half that of another, the wind comes from the latter twice as often). The radii of the second wedges are simply the first multiplied by the average velocity. The inner most (3rd, and solid) wedge is the first multiplied by the cube of the velocity, thus showing the average energy from that direction. These are however relative indications, and absolute figures are necessary to do detailed design planning for a generator installation. Historical data is usually supplemented by on-site measurements before serious design commitments are made.

11.2.2 Terrain Effects:
The wind will be affected by the friction with the surface over which it travels. The amount of “friction” will depend on the characteristics of the surface, which generally indicates how smooth or rough it is. This characteristic is often referred to as the “roughness class” or “roughness length” which are mathematically related. (The roughness length is the height above the surface where the wind speed would theoretically be zero (Danish Wind Industry Association). For the smoothest surface, water, this is considered to be \(~0m (0.2mm)\). For a large city with trees and tall buildings, this is \(~1.6m\).) The higher the class or length, the more friction and thus the more the wind will be slowed as we near the surface. This parameter can be used to calculate the expected wind speed at a certain height about the terrain given the speed at a specified height. (This will be illustrated with an example below.) Roughness length is usually obtained from tabulated data which is based on meteorological data. The “roughness” can also be summarized graphically using a “roughness rose”. This is identical in concept to a “wind rose” and gives the average roughness going out in each direction from a particular location. Of course there are many other considerations related to terrain that should be considered in an effort to predict wind performance (e.g. valleys, hills, protrusions etc.)

11.2.3 Wind Sheer:
Wind sheer is the name of the effect of the surface friction that causes the wind to slow down at lower heights above the terrain (as we get closer to the surface). At an altitude
of approximately 1km above the terrain, the wind is not significantly affected by the roughness of the surface (Danish Wind Industry Association). The effect is quantitatively described by the following relationship:

\[ v_x = v_{ref} \left( \frac{\ln \left( \frac{H_x}{H_0} \right)}{\ln \left( \frac{H_{ref}}{H_0} \right)} \right) \]  

(11.8)

where \( v_x \) is the wind velocity at height \( x \) above the terrain, \( v_{ref} \) is the wind velocity at a known height, \( H_{ref} \), \( H_x \) is the height of the wind velocity, \( v_x \), and \( H_0 \) is the roughness height of the terrain. All variables are given in comparable units (e.g. meters, and meters/second).

For example, if we know the wind velocity at a height of 10 meters is 10 m/s, and the roughness height of the terrain is .03 (prairie like: only gentle hills, no fences or hedges, only a few scattered buildings) and we want the wind velocity at a height of say 30m:

\[ v_{30} = (10) \left( \frac{\ln \left( \frac{30}{.03} \right)}{\ln \left( \frac{10}{.03} \right)} \right) = 11.9 \text{ m/s} \]

This would be a low wind sheer. As the roughness height increases, the wind sheer also increases. For example, using a roughness height of .4 (towns and villages with trees, many building or very rough terrain) in our example above gives us a velocity at 30m of 13.4m/s. In general, the rougher the terrain, the more wind sheer, or the more the wind speed increases as height increases.

### 11.2.4 Wind Parks (Farms)

As mentioned above, there is turbulence induced in the air flow around an “energy extractor” like a rotor-driven generator. This turbulence is induced in all directions, with the worst being directly behind, and the least being directly in front. Ideally, undisturbed airflow would be desirable. For this reason, when a number of wind generators are installed in the same general area (a ‘wind farm’ or ‘wind park’), they are spaced so that the turbulence has time (and distance) to dissipate before another extraction is attempted. According to the Danish Wind industry Association, a common rule of thumb is to allow from 5 – 9 rotor diameters between generators in the wind direction, and 3 – 5 rotor diameters horizontally. The final decision will depend on several factors including the design and size of the rotors (how much turbulence they will generate), the working wind speed, and the available space.

### 11.2.5 Wind Cut-off

A typical wind generator is capable of handling a range of wind speed. At very low speeds, the generator will not function enough to contribute useable power. At very high wind speeds, the mechanical and electrical stresses may damage the installation. For this reason, wind generators typically have a “cut-in” and “cut-out” wind speed specified for operations. Outside of this “wind window”, the rotors are turned so that they stop
extracting energy altogether. We expect that improvements in design will continue to broaden this “window”.

**Example 11-1: Wind Power Example 1**

A wind generator with a 10m diameter rotor arrangement and a maximum output capacity of 30kW is mounted on a 30m rigid column. The prevailing wind is from the north-east 70% of the time and its equivalent average velocity is 5.5m/s @ 5m height for 10 hours per day (i.e. this means exactly 5.5m/s wind for 10 hours and then nothing the rest of the time). The “cut-in” and “cut-out” wind speeds for this generator are 2m/s and 10m/s respectively. The average roughness length of the terrain upstream from the installation is 0.3m. (Assume the density of the air in this local is 1.225kg/m$^3$. Also assume the average velocity at the generator height is constant over the rotor area.)

a) What is the average wind velocity at the generator height?

b) What is the total energy in this wind?

c) What is the maximum power that can be extracted from this wind?

d) If the wind speed is reduced only 2m/s, how much energy has been extracted?

**Solution:**

a) This is a “wind sheer” question, and we have all the parameters we need to calculate the wind speed at 30m (we can assume this will be the average over the whole rotor area):

\[
v_{30m} = v_{ref} \left( \frac{\ln \left( \frac{H_v}{H_o} \right)}{\ln \left( \frac{H_{ref}}{H_o} \right)} \right) = 5.5 \text{ m/s} \left( \frac{\ln \left( \frac{30m}{3m} \right)}{\ln \left( \frac{5m}{3m} \right)} \right) = 9.0 \text{ m/s}
\]

b) From \( P = \frac{1}{2} \rho \pi r^2 v^3 \) and \( \rho = 1.225 \text{ kg/m}^3 \) (11.3) notes:

\[
P = \frac{1}{2} \rho \pi r^2 v^3 = \frac{1}{2} \left( 1.225 \text{ kg/m}^3 \right) \pi \left( 5m \right)^2 \left( 9m/s \right)^3 = 3510 \text{ kV}
\]

c) According to Betz’ Law, a maximum of 16/27 can be harvested, so

\[
3510 \text{ kV} \left( \frac{16}{27} \right) = 2080 \text{ kV}
\]
d) Recall that Betz’ maximum was when the outgoing wind speed was \( \sim \frac{1}{3} \) of the incoming, so we’re well above that limit if we reduce it from 9m/s to 7/m/s. Using Eq 11.6:

\[
P = \frac{1}{4} \rho \pi r^2 (v_i + v_o)(v_i^2 - v_o^2) = \frac{1}{4} \left(1.225 \frac{kg}{m^3}\right) \pi (5m)^2 (9 + 7)(9^2 - 7^2) = 12.315kW
\]
11.3 Appendix A: Thévenin 101

Summary:
The purpose of this note is to provide a brief guideline on when and how to use Thévenin and Norton equivalents to help with circuit analysis. It provides some simple steps for determining the equivalents as well as an example to demonstrate the process. Finally, it provides a couple of cautions about the limitations of using equivalents.

Why use equivalents?
Simply stated, an equivalent circuit can be used to simplify determination of the voltage across, or current through a specific element. In order to do this, we replace a rather complicated network of components with a simple 2-element replacement that acts exactly the same as far as the target component is concerned.

Using the equivalent network to find the voltage or current through the target component is usually not a problem. The challenge is determining the equivalent itself, and sometimes when it will be helpful or necessary to do so.

When to use equivalents:
There are many techniques that can be used for circuit analysis, and they will all give the same result. While there is often some advantages in using one method versus another in a specific circuit situation, which one you use will depend on the questions you are trying to answer and your personal familiarity and expertise at applying them. This note will focus on the analysis you will typically have to do for transient D.C. circuits. In these cases, a Thévenin or Norton equivalent is often very useful. You will typically require an equivalent if the network between a source(s) and the target component (i.e. the component of interest) has a combination of series and parallel components. Another, but less obvious way to tell is if the circuit is such that varying the current or voltage in the target component will affect the voltage or current at the component. For example, if we’re interested in the current through the 470kΩ resistor in the circuit above, we can see that changing the current in it will change the voltage across it and thus affect the contribution for the various sources in the network. This is a good hint that an equivalent circuit will be helpful! At worst, you can’t go wrong by using an equivalent circuit (it will always perform the same as what it replaces), you may just do a bit more work than necessary. Of course this assumes that you determine the equivalent correctly!

Definitions:
A Thévenin or Norton equivalent circuit consists of a source (voltage or current respectively) and one resistor, forming a two-terminal network, that will behave exactly the same from the target component’s perspective as the more complex network it replaces. It is very important to understand that the source and the resistor in the equivalent network are generally not the same as any original component, even though they may have the same values! This means, for example, you cannot infer that the current through an original resistor is the same as the current going through R_Th, even

though it may have the same value. In some cases some relevant information can be deduced from the equivalent, but this should be done cautiously!

**How to Determine Equivalents:**
A simple series of steps will help derive an equivalent circuit:
1. define two terminals isolating the component of interest (target component),
2. “remove” the target component, and determine the resistance “looking back” into the terminals with the sources replaced by their internal resistance (a short for voltage sources and an open for current sources),
3. determine the “open circuit” voltage (Thévenin) or short circuit current (Norton) at the terminals,
4. use the source and resistance to build a series network (voltage source, resistor =, Thévenin) or parallel network (current source, resistor = Norton) and attach it to the two terminals in place of the original network.

Admittedly, steps two and three can cause some challenges with some circuits. In these cases, some practice with circuit analysis will help determine the necessary component values.

Finding the equivalent resistance (2), \( R_{Th} \) or \( R_N \), is usually just a matter of reducing resistor networks. This topic is well covered elsewhere and not the prime target of this discussion.

Finding the open circuit voltage or short circuit current (3) can often be aided by using Thévenin<->Norton source conversion themselves. A brief example will illustrate how.

Consider the circuit example shown above. Let’s say we’re interested in the current through the 4.7k\( \Omega \) resistor. We first define two terminals around the component. (Note: I’ve also “flipped” the 3V source so that it is a +3V facing the other way.)

Next we remove the component and replace the sources by their “equivalent impedances”. Looking back into the terminals, we “see”

\[
(100k + 200k)/1M\Omega, \text{ so } R_{Th} = 230.7k\Omega
\]

Next we need to determine the open circuit voltage at the terminals. Here we should recognize the conversion relation between Thévenin and Norton equivalents:
Note this is the same as between all equivalents: same $V_{OC}$, same $I_{SC}$!

First, notice a “Norton Equivalent” on the right that could be converted to a Thévenin Eq. and combined in series with the other “Thévenin Eq. on the right side:

Also note that the polarities between the equivalent’s terminals must also match!

At this point we have two choices: use the voltage divider rule to determine $V_{OC}$; or convert both the remaining Thévenin Eq. to Norton Eq., combine them into 1 Norton Eq. and convert back to 1 Thévenin Eq.

**Voltage divider first:**

Note the two sources are opposing in polarity which leaves $(9-3.6)$ 5.4V to drop across the two resistors. Using simple voltage divider rules, and observing appropriate polarities, we can determine that 4.15V drops across the 1MΩ resistor, and the rest across the 300kΩ resistor. This leaves us with a voltage between our terminals of interest of -4.85V (assuming the bottom terminal is the reference), and the following equivalent circuit..

**Using Source Conversions:**

Combining:

and converting back:
We can now simply calculate the required value: -20.6μA (assuming the positive direction is defined as down through the resistor)

**Final Summary and Cautions:**
The steps suggested above will help you derive Thévenin and Norton equivalents for D.C. circuits containing resistors and independent sources, and use them to simplify circuit analysis. The most important word of caution, as stated above, is that once the equivalent substitution has been done, the original components are not there. In general, you cannot use the equivalent components to deduce any parameters (current through, voltage across) for the original components that were replaced. This is also true of simple Thévenin <-> Norton source conversions; even though the resistances are the same value, they are **not the same resistors**!!